

**Engineering of Cytochrome P450s CYP109E1 and
CYP109A2 from *Bacillus megaterium* DSM319 for the
production of vitamin D₃ metabolites**

Kumulative Dissertation

zur Erlangen des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät

der Universität des Saarlandes

von

Ammar Abdulmughni

Saarbrücken

2018

Tag des Kolloquiums : 14.08.2018

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

Abstract

Active vitamin D₃ metabolites play an essential role in the maintenance of calcium and phosphorus homeostasis. The conventional chemical synthesis of this metabolite is time-consuming, environmentally unfriendly and often results in low yield. Therefore, the biotechnological production of active vitamin D₃ metabolites is of great importance to the pharmaceutical industry. Hereby, cytochrome P450 enzymes have the potential to achieve this goal.

The present work reports on the optimization of a biotechnological process in *Bacillus megaterium* MS941 for the production of vitamin D₃ metabolites. On that account, two cytochrome P450 enzymes were used as biocatalysts, namely CYP109E1 and CYP109A2 from the Gram-positive bacterium *Bacillus megaterium* DSM319. Both enzymes were subjected for functional and structural characterization in order to optimize their activity and/or regio-selectivity towards vitamin D₃. In terms of hydroxylation activity, it has been shown that the conversion of vitamin D₃ with CYP109E1 results in the formation of several derivatives, while CYP109A2 shows clearly a higher regio-selectivity towards 25-hydroxylation. The elucidation of the crystal structure of both enzymes provides detailed insights into the geometry of these enzymes. By means of molecular docking, site-directed mutagenesis was successfully performed, resulting in the creation of mutants with higher regio-selectivity compared to the wild type, in particular when using CYP109E1. The enhancement of the regio-selectivity of this P450 resulted in a higher yield of the valuable metabolite 25-hydroxyvitamin D₃. A two-fold increase of the 25-hydroxyvitamin D₃ yield was achieved after the substitution of I85 by tryptophan or alanine. The engineering of CYP109A2 resulted also in a slight increase in the 25-hydroxyvitamin D₃ yield compared to the wild type. Summarizing, it has been shown that CYP109E1-I85W, CYP109A2-, - or CYP109-T103A-based whole-cell systems are a competitive methods for the production of 25-hydroxyvitamin D₃ among the previously established biotechnological processes.

Zusammenfassung

Aktive Vitamin-D₃-Metaboliten spielen eine wesentliche Rolle bei der Aufrechterhaltung der Calcium- und Phosphor-Homöostase. Die herkömmliche chemische Synthese dieser Metaboliten ist zeitaufwendig, umweltschädlich und führt oft zu geringer Ausbeute. Daher ist die biotechnologische Herstellung von aktiven Vitamin-D₃-Metaboliten für die pharmazeutische Industrie von großer Bedeutung. Der Einsatz von Cytochromen P450 gilt hierbei als vielversprechend.

Die vorliegende Arbeit beschäftigt sich mit der Optimierung eines biotechnologischen Prozesses in *Bacillus megaterium* MS941 zur Produktion von Vitamin-D₃-Metaboliten. Als Biokatalysatoren wurden zwei Cytochrome-P450-Enzyme verwendet, nämlich CYP109E1 und CYP109A2 aus dem Gram-positiven Bakterium *Bacillus megaterium* DSM319. Beide Enzyme wurden einer funktionellen und strukturellen Charakterisierung unterzogen, um ihre Aktivität und / oder Regio-Selektivität gegenüber Vitamin-D₃ zu optimieren. Bezüglich der Hydroxylierungsaktivität wurde gezeigt, dass die Umwandlung von Vitamin-D₃ mit CYP109E1 zur Bildung mehrerer Derivate führt, während CYP109A2 deutlich eine höhere Regio-Selektivität bezüglich einer 25-Hydroxylierung zeigt. Die Aufklärung der Kristallstruktur beider Enzyme liefert detaillierte Einblicke in die Geometrie dieser Enzyme. Mit Hilfe des molekularen Dockings wurde eine ortsgerichtete Mutagenese erfolgreich durchgeführt, wodurch im Vergleich zum Wildtyp, insbesondere von CYP109E1, Mutanten mit höherer Regio-Selektivität hergestellt wurden. Die Optimierung der Regio-Selektivität dieses P450 führte zu einer höheren Ausbeute des wertvollen Metaboliten 25-Hydroxyvitamin-D₃. Die Produktion von 25-Hydroxyvitamin-D₃ konnte durch die Substitution von I85 durch Tryptophan oder Alanin verdoppelt werden. Die Optimierung von CYP109A2 führte ebenfalls zu einem leichten Anstieg der 25-Hydroxyvitamin-D₃-Ausbeute im Vergleich zum Wildtyp. Zusammenfassend wurde gezeigt, dass CYP109E1-I85W-, CYP109A2-, oder CYP109-T103A-basierte Ganzzellsysteme eine wettbewerbsfähige Methode zur Herstellung von 25-Hydroxyvitamin-D₃ unter den bisher etablierten biotechnologischen Verfahren darstellen.

Danksagung

Ich bedanke mich herzlich bei all den Menschen, die mich immer unterstützt und mir dadurch ermöglicht haben, diese Arbeit zu beenden.

Mein besonderer Dank gilt meiner Doktormutter Frau Prof. Rita Bernhardt. Ich danke ihr für die Möglichkeit, dass ich meine Doktorarbeit in ihrem Arbeitskreis anfertigen konnte. Ich bedanke mich bei ihr auch für die großartige wissenschaftliche Betreuung und die wertvollen Anregungen und Diskussionen sowohl in beruflichen als auch persönlichen Angelegenheiten. Von ihr habe ich sehr viel gelernt.

Prof. Dr. Gert-Wieland Kohring danke ich herzlich für die Übernahme des Korreferates. Ich bedanke mich bei Dr. Frank Hannemann für Unterstützung und die hervorragende wissenschaftliche Beratung.

Ganz besonders bedanke ich mich bei Frau Dr. Ilona Jóźwik und Dr. Andy-Mark Thunnissen von der Reichuniversität Groningen für die ausgezeichnete und unkomplizierte Zusammenarbeit.

Ich bedanke mich bei allen Mitgliedern der Arbeitsgruppe für die angenehme und hervorragende Zusammenarbeit. Mein Dank gilt besonderes Dr. Simone Brixius-Anderko, Lisa König, Benjamin Stenger, Natalia Putkaradze, Dr. Martin Litzenburger und Tanja Sagadin. Ich bedanke mich auch bei Birgit Heider-Lips und Antje Eiden-Plach für ihre stetige Hilfsbereitschaft und ihr Interesse, welche auch über die Arbeit hinausgehen.

Im Speziellen bedanke ich mich bei Dr. Mohammed Milhim und Philip Hartz für die zahlreichen und nützlichen Tipps und sehr interessanten Diskussionen in allen möglichen fachlichen und nicht-fachlichen Gebieten.

Für ihre Hilfe bei allen bürokratischen Angelegenheiten möchte ich mich auch bei Gabi Schon bedanken.

Meinen Freunden Danae, Didi, Ghamdan, Jannis, Katrin, Lisa, Shadi, Sherif und seiner Frau Iva, Sheikh, Tarik und Verena danke ich für die Unterstützung. Ich bin sehr froh, dass es euch gibt. Ich danke besonders meiner Esther dafür, dass sie immer an mich glaubt und mir bei schwierigen Zeiten zur Seite steht.

Mein größter Dank gilt meiner großen Familie, insbesondere meinen Eltern, ohne deren bedingungslose Liebe, Geduld und Unterstützung all das nicht möglich wäre.

An dieser Stelle möchte ich an einen wunderbaren Menschen erinnern, meinen besten Freund Musaab Al-Tuwaijary, dem ich immer mein Herz ausschütten konnte. Ich vermisse dich.

Contents

Abstract.....	I
Zusammenfassung.....	II
Danksagung.....	III
Contents.....	V
Scientific contributions.....	VII
1. Introduction.....	1
1.1 Protein engineering.....	1
1.2 Vitamin D ₃	2
1.3 Metabolism of vitamin D ₃	4
1.4 Action of the hormone 1 α ,25-dihydroxyvitamin D ₃	6
1.5 Cytochrome P450 monooxygenases.....	6
1.6 Electron transfer in cytochrome P450 systems.....	9
1.7 Structure of P450 enzymes.....	9
1.8 Cytochrome P450s for the production of active vitamin D ₃ metabolites.....	13
1.8 Aim of this work.....	14
2. Scientific articles.....	15
2.1 Abdulmughni et al (2017a).....	15
2.2 Abdulmughni et al (2017b).....	26
2.3 Jóźwik et al (2016).....	44
2.4 Putkaradze et al (2017).....	72
3. General discussion	103
3.1 Bioconversion of vitamin D ₃ by CYP109E1 and CYP109A2.....	104
3.2 Structural aspects of CYP109E1 and CYP109A2.....	108
3.3 Engineering of CYP109E1.....	109
3.4 Engineering of CYP109A2.....	112

4. Outlook.....	117
5. Appendix.....	119
6. Abbreviations.....	122
7. References.....	123

Scientific contributions

The work is based on four original research papers reproduced in chapter 2 with permission of Journal of Biotechnology (2.1 Abdulmughni et al 2017a), FEBS Journal (2.2 Abdulmughni et al 2017b and 2.3 Jóźwik et al (2016)) and Appl Microbiol Biotechnol (2.4 Pukaradze et al (2017)).

- 2.1 **Abdulmughni A**, Jóźwik IK, Putkaradze N, Brill E, Zapp J, Thunnissen AW, Hannemann F, Bernhardt R. (2017a). Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D₃ hydroxylase. *J Biotechnol.* 243, 38 –47.

The author performed all *in vitro* and *in vivo* experiments with CYP109E1 and its mutants, designed and carried out the site-directed mutagenesis. Furthermore, the author participated in the product purification and to writing the manuscript.

- 2.2 **Abdulmughni A**, Jóźwik IK, Brill E, Hannemann F, Thunnissen AMWH, Bernhardt R (2017b) Biochemical and structural characterization of CYP109A2, a vitamin D₃ 25-hydroxylase from *Bacillus megaterium*. *FEBS J.* 284, 3881-3894.

The author cloned, expressed and purified CYP109A2. He performed all *in vitro* data including the kinetic parameters, designed the vector for whole-cell conversion and performed all *in vivo* experiments. In addition, the author participated to writing the manuscript.

- 2.3 Jóźwik IK, Kiss FM, Gricman Ł, **Abdulmughni A**, Brill E, Zapp J, Pleiss J, Bernhardt R, Thunnissen AW (2016) Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J.* 283, 4128–4148.

The author designed and carried out the site-directed mutagenesis, expressed and purified the mutants of CYP109E1, performed the *in vitro* experiments with CYP109E1 mutants and participated in the interpretation of the results. In addition, the author participated to writing the manuscript.

- 2.4. Putkaradze N, Litzenburger M, **Abdulmughni A**, Milhim M, Brill E, Hannemann F, Bernhardt R (2017) CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*. *Appl Microbiol Biotechnol.* 101, 8379-8393.

The author performed the docking experiments and participated to writing the manuscript.

1. Introduction

1.1. Protein engineering

‘Enzymes as catalysts are of key importance in biotechnology, similar to the role of nucleic acids as carriers of genetic information’ (Buchholz et al., 2005). They are widely used in the industrial production of bulk chemicals and pharmaceuticals. The conventional chemical synthesis often requires complex procedures under harsh conditions; in contrast, enzymes perform their reactions with high specificity and enantio-, regio- and stereo-selectivity under mild conditions. Despite their attractive features, enzymes often do not meet the requirement of an industrial process and, therefore, different enzyme properties usually have to be improved. Out of this necessity, a new multidisciplinary field named protein engineering was born, which comprises different science areas: molecular biology, biochemistry, crystallography, bioinformatics and biotechnology.

Generally, there are two basic approaches for protein engineering: directed evolution and rational design. The first one involves different molecular biology methods that mimic the natural evolution via random mutagenesis (e.g. chemical mutagenesis, error prone PCR) and/or DNA recombination (e.g. gene shuffling, staggered extension process (StEP)) (Zhao et al., 2002). In iterative cycles, mutations are randomly introduced to create a library of genes followed by screening to identify enzyme variants that exhibit the desired function. The advantage of this strategy is that no structural information of the protein is needed. On the other hand, the screening procedure could be as difficult as looking for a needle in a haystack. However, directed evolution has been successfully applied to optimize biocatalysts in terms of their specificity, stability, activity. Successful applications have been reviewed by (Arnold, 2017; Böttcher and Bornscheuer, 2010; Zhao et al., 2002).

In rational protein design, mutations are introduced to specific protein sites (site-directed mutagenesis) and, therefore, the availability of the enzyme structure and knowledge about structure-function relationship is required. The advanced progress in protein crystallography and computational methods can enhance the potential of protein engineering via site-directed mutagenesis (Illanes et al., 2012; Röthlisberger et al., 2008). Rational design has been successfully used to overcome the enzyme limitations in industrial applications. The thermostability of α -amylase from *Bacillus licheniformis* and formate dehydrogenase from *Pseudomonas* sp. 101 has been improved by removal of asparagine residues and

introduction of additional hydrophobic contacts, respectively (Declerck et al., 2000; Rojkova et al., 1999). Further examples for successful protein engineering via site-directed mutagenesis are reviewed by (Böttcher and Bornscheuer, 2010). Recently, the combination of directed evolution and rational design techniques supported by computer-aided methods opens new possibilities for protein engineers (Li and Cirino, 2014).

In recent years, the cytochrome P450 monooxygenases (CYP or P450) has been recognized as potential biocatalyst for pharmaceutical and biotechnological industries since they can catalyze the oxidation of inactivated hydrocarbons in a stereo-selective manner (Bernhardt, 2006; Bernhardt and Urlacher, 2014; Hazel M Girvan & Andrew W Munro, 2016). However, P450s enzymes have some properties that limit their industrial applications to few examples (Julsing et al., 2008). For this reason, cytochrome P450 enzymes have been engineered by directed evolution and rational design to improve their activities, expression levels, stabilities and substrate specificities as well as regio- and stereo-selectivities. In this context, the most investigated cytochrome P450 enzymes is the bacterial fatty acid hydroxylase CYP102A1 (P450 BM3) from *Bacillus megaterium*. CYP102A1 has been engineered to catalyze the hydroxylation of non-natural substrates as diverse as pharmaceuticals, terpenes and alkanes. CYP102A1-F87A was generated by site-directed mutagenesis, which is able to catalyze the hydroxylation of fatty acids at the position ω -4. In contrast, the wild type enzyme catalyzes the hydroxylation at ω -1, ω -2 and ω -3 (Oliver et al., 1997). An overview of CYP102A2 mutants can be found in (Whitehouse et al., 2012). Numerous protein engineering studies have been performed with cytochrome P450 enzymes, providing different examples where different enzyme properties have been successfully improved and in other cases they have not. Further examples for P450s engineering are reviewed in (Gillam, 2008; Kumar, 2010; McIntosh et al., 2014; Yasuda et al., 2018).

1.2. Vitamin D₃

The fat-soluble vitamin D₃ is a secosterol, which is synthesized from the precursor 7-dehydrocholesterol (7-DHC) during exposure to sunlight (Holick et al., 1980). The photolytic conversion of 7-DHC results in the formation of previtamin D₃ that is further converted to vitamin D₃ by heat-induced isomerization (Okano et al., 1977; Holick et al., 1977) (**Figure. 1.1**). Vitamin D₃ can also be obtained naturally from the diet. Nevertheless, only a few natural sources contain significant amounts of this important compound. The flesh of fatty fish such as salmon and tuna as well as the fish liver oils count to the best

naturally sources for vitamin D₃. Although more and more foods are fortified with vitamin D₃, the sunlight-dependent endogenous synthesis of vitamin D₃ remains the main source for most populations (Calvo et al., 2005). Therefore, environmental factors such as latitude and weather conditions affect vitamin D₃ synthesis (Spiro and Buttriss, 2014). Moreover, individual characteristic, such as skin pigmentation, age and outdoor activity can also influence vitamin D₃ status. Due to this, vitamin D₃ insufficiency exists worldwide including Europe (Duso et al., 2004; Spiro and Buttriss, 2014).

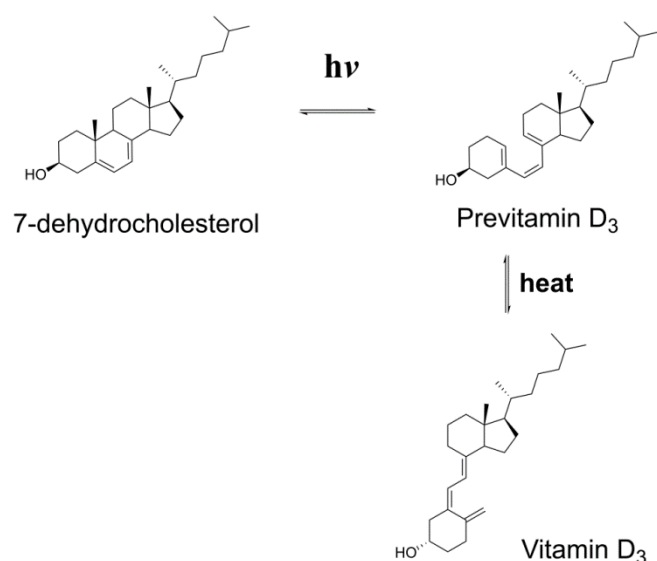


Figure 1.1. Vitamin D₃ synthesis in the skin. The photolytic conversion of 7-dehydrocholesterol into previtamin D₃ is followed by heat-induced isomerization producing vitamin D₃.

The consequences of low vitamin D₃ level in the blood are mainly remarkable in bone formation and mineralization, because vitamin D₃ metabolites are principle factors in calcium and phosphorus homeostasis. In absence of vitamin D₃, only 10 to 15% of dietary calcium and about 60% of phosphorus could be absorbed (DeLuca, 2004; Hossein-nezhad and Holick, 2013), resulting in rickets among children and osteomalacia among adults (Underwood and DeLuca, 1984; Heaney et al., 2003). Furthermore, vitamin D₃ metabolites play a key role in several physiological processes such as cell growth, apoptosis, modulation of immune response and renin-angiotensin system (Abe et al., 1981; Wagner et al., 2003; Li, 2003; Di Rosa et al., 2011). Therefore, worldwide there has been growing interest in the physiological and therapeutic aspects of vitamin D₃. Several studies have linked vitamin D₃

deficiency to the development of different chronic diseases such as cardiovascular and autoimmune diseases, neuromuscular dysfunction, diabetes mellitus, different cancers and gynecological disorders (DeLuca and Cantorna, 2001; Grant, 2002; Holick, 2004).

1.3. Metabolism of vitamin D₃

Apart from the origin of vitamins D₃, this molecule is biologically inert. Two sequential hydroxylation steps, the first one at C-25 and the second at C-1, are required to produce the hormonally active form of vitamin D₃, 1 α ,25-dihydroxyvitamin D₃ (1,25(OH)₂VD₃) (**Figure 1.2**). These reactions are catalyzed by different enzymes of the same family, namely cytochrome P450 (CYP or P450).

Studies on the 25-hydroxylase in human and animals showed that several cytochrome P450 enzymes, such as CYP2R1, CYP27A1, CYP2J2, CYP2D25 and CYP3A4, can act as vitamin D₃ 25-hydroxylases (Prosser and Jones, 2004; Schuster, 2011). However, only CYP2R1 is considered as the physiologically relevant enzyme (Cheng et al., 2004). The 25-hydroxylation of vitamin D₃ occurs primarily in the liver, producing 25-hydroxyvitamin D₃ (25(OH)VD₃), also known as calcidiol. It also has been shown that 25(OH)D₃ is produced in other tissues such as kidney and intestine (Horsting and DeLuca, 1969; Holick et al., 1972; Tucker et al., 1973). 25(OH)VD₃ is the major circulating vitamin D₃ metabolite and, for this reason, its plasma level is considered as the best indicator for vitamin D₃ status. Among experts, there is no consensus on the optimal plasma level of this indicator. However, a 25(OH)VD₃ concentration less than 20 ng/mL (50 nmol/L) is related with vitamin D₃ deficiency (Bischoff-Ferrari et al., 2006; Holick, 2007). The Clinical Practice Guideline of the Endocrine Society distinguishes between vitamin D deficiency (< 20 ng/mL), insufficiency (21 to 29 ng/mL) and sufficiency (30 ng/mL), whereas a concentration between 40 and 60 ng/mL is recommended (Holick et al., 2011). 25(OH)VD₃ is not only interesting for its task as an indicator for vitamin D₃ status, but also as therapeutic agent. It has been found that supplementation of 25(OH)VD₃ has a preventive, therapeutic effect against chronic kidney disease, Crohn's and cholestatic liver disease (Leichtmann et al., 1991; Jean et al., 2008). In addition, anti-proliferative activity of this metabolite has been shown (Munetsuna et al., 2014). 25(OH)VD₃ is also used as supplement in animal food, as it has been found that this metabolite is the most active form of vitamin D₃ capable of

supporting both cellular functions and embryonic development in chickens and turkeys when fed as the sole source of vitamin D₃ (Soares et al., 1995).

The second step in vitamin D₃ activation, 1 α -hydroxylation, occurs mainly in the kidneys by the mitochondrial enzyme CYP27B1 producing the secosteroid hormone 1 α ,25-dihydroxyvitamin D₃ (1 α ,25(OH)₂D₃), also known as calcitriol, which is the most active form of vitamin D₃. Most of vitamin D₃ physiological functions rely on the action of this active metabolite. More detailed information on biological functions of 1 α ,25(OH)₂D₃ can be found in (Brown et al., 2002; DeLuca, 2004; Deluca and Cantorna, 2001; Duso AS. et al., 2004; Sutton and MacDonald, 2003).

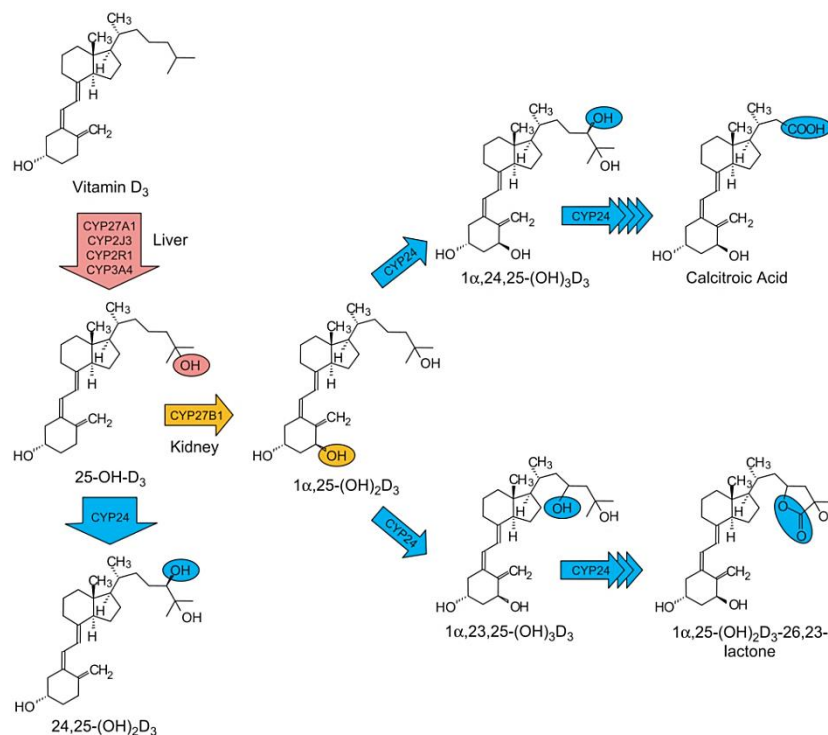


Figure 1.2. Vitamin D₃ activation and inactivation. Liver 25-hydroxylase converts vitamin D₃ into 25(OH)VD₃, the most circulating metabolite of vitamin D₃. The second activation step occurs primarily in the kidney producing the most active metabolite, 1 α .25(OH)₂VD₃. Vitamin D₃ inactivation is performed by 24-hydroxylase (CYP24A1) after a series of oxidation reactions resulting in calcitroic acid or 1 α ,25(OH)₂VD₃-26,23-lactone. The metabolite 24,25(OH)₂VD₃ is produced from 25(OH)VD₃ by CYP24A1. Figure is taken from (Prosser and Jones, 2004).

The inactivation of vitamin D₃ is carried out by the mitochondrial enzyme CYP24A1, which is involved in two catabolic pathways; (i) the carbon-24 oxidation pathway from 1 α ,25(OH)₂VD₃ to calcitroic acid (ii) the lactone pathway from 1 α ,25(OH)₂VD₃ to 1 α ,25(OH)₂VD₃-26,23-lactone (**Figure 1.2**) (Prosser and Jones, 2004). In addition, CYP24A1 has an activating role in vitamin D₃ metabolism as it catalyzes the formation of the metabolite 24,25(OH)₂VD₃, which has been proposed to play a biological role in bone fracture repair or cartilage maturation (Seo et al., 1997).

1.4. Action of the hormone 1 α ,25-dihydroxyvitamin D₃

The secosteroid hormone, 1 α ,25(OH)₂VD₃, accomplishes its biological functions, like other steroid hormone, through intracellular receptors that function as transcription factors (Tsai and O'Malley, 1994). The hormone 1 α ,25(OH)₂VD₃ binds to the nuclear vitamin D receptor (VDR), which is present in most tissues and cells. The complex consisting of 1,25(OH)₂VD₃ and VDR forms then a heterodimer with the retinoic acid X receptor (RXR) that binds to specific short repeating hexameric nucleotide sequences in the genome known as vitamin D response elements (VDREs) (Ozono et al., 1990; Barletta et al., 2002). The VDREs are usually located upstream in the promoter of the target gene (Gamble and Freedman, 2002). As a result of ligand-induced conformational change of the VDR-RXR heterodimer, co-repressor proteins such as the nuclear receptor co-repressor (NCoR) are dissociated from the complex and interactions with different co-activators are created (Haussler et al., 1998). This complex has the ability to recruit a large coregulatory complex into the transcriptional preinitiation complex of the target gene resulting in enhancement or suppression of gene transcription (Pike and Meyer, 2010; Sutton and MacDonald, 2003).

1.5. Cytochrome P450 monooxygenases

Cytochromes P450 (CYPs or P450s) are heme-containing enzymes, which are found throughout the three biological kingdoms (Nelson, 2011). The earliest insight into P450 research began in 1958, when a carbon-monoxide binding pigment was isolated from rat and pig liver microsomes (Garfinkel, 1958; Klingenberg, 1958). Later, Omura and Sato demonstrated the presence of iron-protoporphyrin IX in this pigment, which was named cytochrome P450 (Omura and Sato, 1964). Such name is uncommon for enzymes, which are normally named after their function. The terminology of cytochromes P450 describes the nature and characteristic feature of this enzyme class; cytochrome stands for hemoprotein,

the letter P for ‘Pigment’ and the number 450 relates the characteristic absorption peak at 450 nm when the enzyme is complexed with carbon monoxide (CO).

P450 enzymes are classified into families and subfamilies based on the sequence identity. According to the widely used P450 nomenclature system (Nelson, 2009), members of the same family share an amino acid sequence identity of more than 40%, and those with more than 55% identity are assigned to the same subfamily. The conventional name of **cytochrome P450** enzymes starts with the abbreviation CYP followed by an Arabic number for the P450 family, a letter for the subfamily and a second Arabic number for the individual gene, for instance, CYP109E1 is a member of the family 109 and the subfamily E. The function of P450 enzymes is the activation of molecular oxygen for the metabolism of different chemical structures (e.g. steroids, fatty acids, xenobiotics and terpenes). They can introduce one oxygen atom of the molecular oxygen into an allylic position, double bond and non-activated C-H bond, while the second oxygen atom is reduced to water. The most common reaction of P450s is a hydroxylation reaction (**Figure 1.3**). However, P450s are able to catalyze a variety of reactions including epoxidation, sulfoxidation, N-, O-, S-dealkylation, and C-C bond cleavage. An overview of the reactions catalyzed by P450s is given by (Sono et al., 1996; Bernhardt, 2006; Guengerich and Munro, 2013).

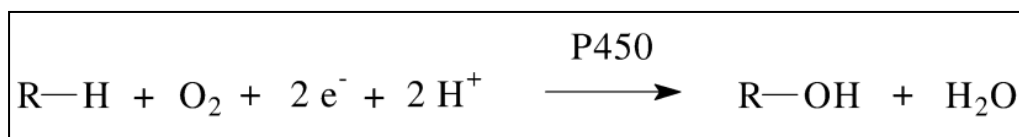


Figure 1.3. General hydroxylation reaction catalyzed by P450 enzymes. Adapted from (Urlacher et al., 2004).

The heme-iron plays a central role in the multi-step catalytic cycle of P450s, which changes between two oxidation states, the Fe III (ferric) and the Fe II (ferrous). At the beginning of each cycle, the heme iron adopts an inactive state, low spin Fe (III), which is six-coordinated by four nitrogen atoms of protoporphyrin IX ring (equatorial), the sulphur atom from cysteine and a water molecule (axial) (**Figure. 1.4**) (1). The substrate binding causes a displacement of the water molecule producing a five coordinated high-spin Fe (III) (2). The first electron delivered by the redox partners is used to reduce the ferric iron into its ferrous form (Fe (II)) (3). In the next step, molecular oxygen is bound, resulting in a ferrous dioxy

species (4). A second reduction followed by a proton delivery results in an iron-hydroperoxo intermediate (5). This complex undergoes heterolytic cleavage, producing the putative iron-oxo ferryl species named “Compound I” (6). According to the oxygen rebound mechanism, compound I abstracts a hydrogen atom from the closely bound substrate forming the intermediate compound II and a substrate radical (7) (Groves, 2006). Through radical recombination, the hydroxylated product is formed and finally released. The enzyme returns to its initial state (low-spin ferric form) and the cycle can start again.

During the catalytic cycle reactive oxygen species (ROS) are produced, which can lead to an inactivation of P450 enzymes. Interestingly, some P450s can oxidize their substrates using peroxides as electron and oxygen donor (shunt pathway) (Nordblom et al., 1976; Bernhardt, 1996). This peroxidase-like mechanism allows P450s to function without the cost-expensive cofactor NADPH and electron transport proteins. The usage of peroxide instead of NADPH results, however, in a decrease of reaction rates (Bernhardt and Urlacher, 2014).

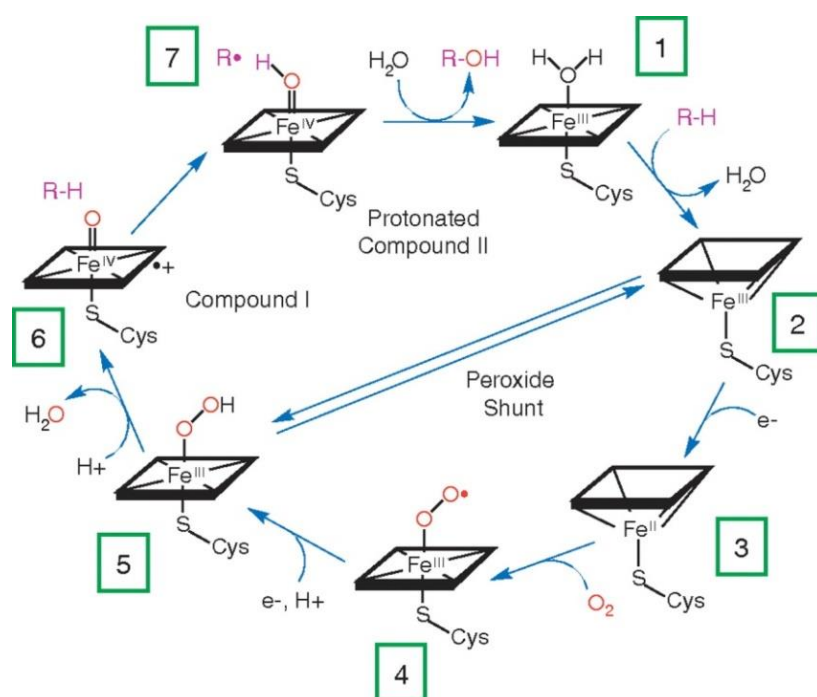


Figure 1.4. Catalytic cycle of cytochromes P450. Figure is taken from (Ener et al., 2010).

1.6. Electron transfer in cytochrome P450 systems

As mentioned above, P450s require sequential delivery of two electrons in order to facilitate their activity, which are usually obtained from the pyridine nucleotide coenzyme NADPH. The transport of these two electrons from NADPH to P450 is performed by means other protein (s), redox partner (s). Concerning the involved redox partners, P450s were classified after their discovery into two main classes, the mitochondrial class I and microsomal class II (Omura et al., 1966; Lu et al., 1969). Since then, several cytochrome P450 systems have been reported, which could be classified into ten classes (Hannemann et al., 2007). However, the most commonly classes are class I and class II (**Figure 1.5**). Most of bacterial and the mitochondrial cytochrome P450 systems belong to class I, consisting of three distinct proteins: (i) a FAD-containing ferredoxin reductase, which is soluble in case of bacteria and associated to the inner mitochondrial membrane in eukaryotes. (ii) a soluble ferredoxin, which receives the electrons one by one from the reductase and shuttles them to the last component of the system, (iii) the cytochrome P450 itself, which is soluble in case of bacteria and bound to the inner mitochondrial membrane in eukaryotes (**Figure 1.5 A, B**). The class II systems are commonly present in eukaryotes and known as microsomal systems, since their two components, the P450 and the NADPH-cytochrome P450 reductase (CPR), are bound to the membrane of the endoplasmic reticulum (**Figure 1.5 C**). CPR includes the prosthetic groups FAD and FMN.

1.7. Structure of P450 enzymes

Despite the low sequence identity among the members of the P450 family (< 20%), P450 enzymes share a common structural fold and topology, which is essential for their function. The conserved overall structure of P450s consists of approximately 13 α -helices (named A-L) and five β -sheets (1-5) with the heme group in a buried substrate binding pocket, forming the so called P450 fold (Graham and Peterson, 1999) (**Figure 1.6**). The most conserved structural elements of the P450 fold include a four-helix bundle (helices D, E, I, and L) and the J- and K-helices. The prosthetic heme group is found between the I- and L-helices binding to the Cys-heme-ligand loop containing the P450 signature sequence FxxGx(H/R)xCxG. Herby, the highly conserved cysteine is bound to the heme iron as fifth ligand, while in other cytochromes like hemoglobin, a histidine residue is coordinating the heme iron (Denisov et al., 2005).

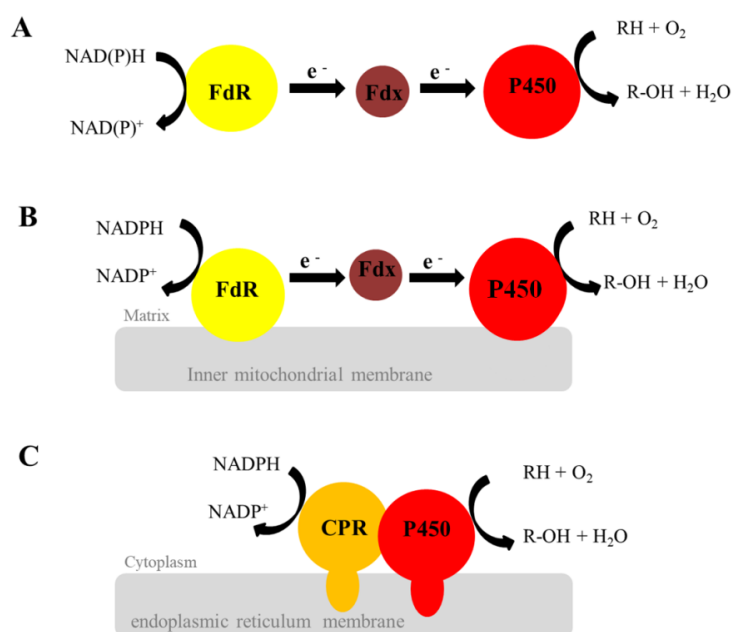


Figure 1.5. Scheme of different cytochrome P450 systems. (A) class I, bacterial system containing three soluble proteins, ferredoxin reductase (FdR), ferredoxin (Fdx) and P450. (B) Class I, mitochondrial system. (C) Class II, microsomal system. Figures modified according to (Hannemann et al., 2007).

The linkage of the heme iron to the cysteine thiolate is the origin of the characteristic Soret absorption at 450 nm when carbon monoxide is bound to the ferrous form of the enzyme. The longest helix (I-helix) is located in the immediate vicinity of the heme and contains a highly conserved threonine and an acidic residue known as acid-alcohol pair, which is supposed to be involved in oxygen activation and proton delivery (Imai et al., 1989; Martinis et al., 1989). Characteristic feature of the I-helix is the local split in the H-bonding pattern of α -helix near the conserved threonine. This helical deformation is critical for the oxygen activation. It is known that the K-helix contains the conserved ExxR motif, which is suggested to be important for stabilizing the core structure.

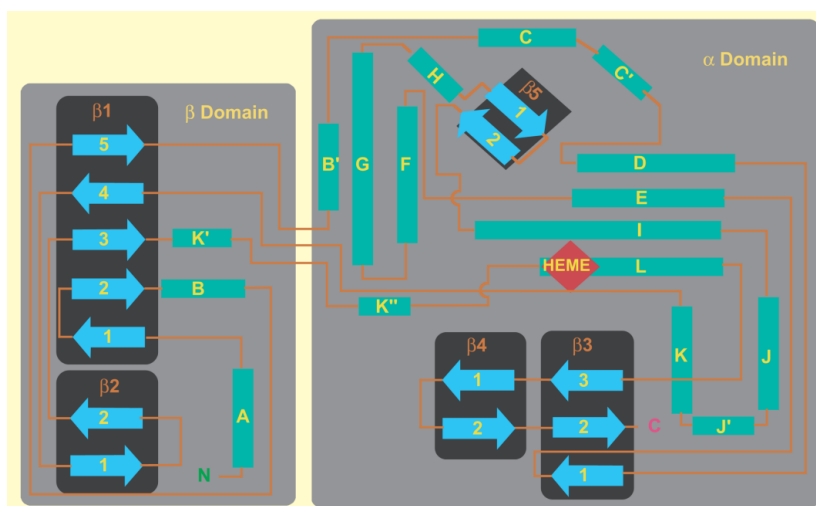


Figure 1.6. Topological illustration of P450 structure. The structural elements α -helices and β -sheets are presented as green tube and blue arrows, respectively. Figure is taken from (Peterson and Graham, 1998).

However, there are variable regions in P450 structures providing the individual properties. Regions in P450s responsible for substrate recognition and binding, known as substrate recognition sites (SRSs), were first identified by Gotoh (Gotoh, 1992), consisting usually of variable residues allowing the conversion of an enormous number of different substrates (Peterson and Graham, 1998). On the basis of sequence or structure comparison, the individual SRSs can be determined for each P450 (Gotoh, 1992; Gricman et al., 2015). Generally, P450 structures contain six SRSs (**Figure 1.7**): SRS1 is located on the BC loop, SRS2, 3 and 4 are located on the F-, G- and I-helices, respectively, and SRS5 is located on $\beta 1$ -4 and SRS6 spans over $\beta 4$ -1 and $\beta 4$ -2. The residues located in the SRSs interact directly with a ligand molecule, or determine the structure or flexibility of the binding site. SRS2 and 3 are located on the entrance channel leading from the bulk to the binding pocket, whereas SRS1, 4, 5, and 6 form the access to the heme.

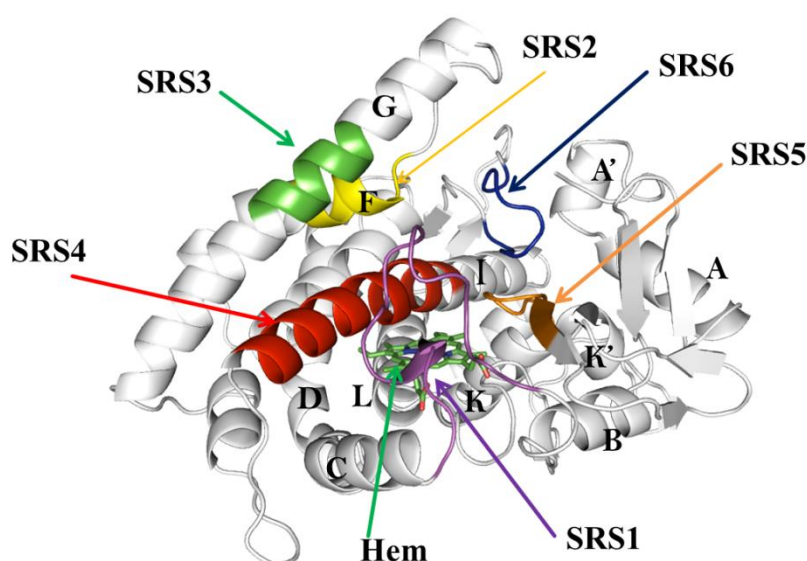


Figure 1.7. The six substrate recognition sites in P450s are shown by arrows (SRS) and represented here for the crystal structure of CYP109A2 (PDB: 5OFQ). The α -helices are labeled following the common P450 nomenclature (A-L).

The number of obtained P450 structures has been increased during the last decades providing significant information about structure-function relationship. It is known that P450s possess flexible structures in order to allow substrate entry followed by closing of the active site. The first indications of the open/closed forms of P450 structures were based on the comparison of BM3 substrate-free and -bound complexes (Li and Poulos, 1997). In the substrate-free form, the F-G loop defines the access channel to the active site, while the F- and G- helices are positioned so that the active site is wide open. Upon substrate binding, remarkable conformational changes occur in the F- and G-helices leading to closure of the active site. Besides the open/closed forms of P450 structures, the active site can adopt different shapes depending on the structure of bound ligands. Such induced fit was firstly observed in CYP119 by the elucidation of its crystal structure with two different ligands; phenylimidazole and imidazole (Yano et al., 2000). The structural flexibility of P450 throws light on the question how the same P450 fold can bind and convert a broad range of different substrates.

1.8. Cytochrome P450s for the production of active vitamin D₃ metabolites

In recent years, P450 biocatalysts have been used for the synthesis of bulk chemicals and pharmaceuticals (Bernhardt, 2006; Bernhardt and Urlacher, 2014; Hazel M Girvan & Andrew W Munro, 2016; Sakaki, 2012; Urlacher et al., 2004). The production of pravastatin from compactin through 6 β -hydroxylation catalyzed by CYP105A3 is one of the most successful applications of P450 enzymes (Watanabe et al., 1995). Another successful industrial application of cytochrome P450 is the bioconversion of vitamin D₃ to 1 α -25-dihydroxyvitamin D₃. Firstly, Sasaki et al. showed that two *Streptomyces* strains are capable of converting 25(OH)VD₃ and 1 α -(OH)VD₃, respectively, to 1 α -25(OH)₂VD₃ (Sasaki et al., 1991). However, the conversion of vitamin D₃ to 25(OH)VD₃ was not detected. Subsequently, the conversion of vitamin D₃ to 1 α -25(OH)₂VD₃ via 25(OH)VD₃ was reported using *Amycolata autotrophica* FERM BP-1573 (now renamed to *Pseudonocardia autotrophica*) (Sasaki et al., 1992). Using this strain, a preparative-scale conversion of vitamin D₃ (200 L tank fermenter) yielded 8.3 mg 25(OH)VD₃/L culture and 0.17 mg 1 α -25(OH)₂VD₃/L culture after 120 h (Sasaki et al., 1992). Afterwards, the identification of vitamin D₃ 25-hydroxylase from this strain was reported, which has been classified as CYP105A2 (Kawauchi et al., 1994). Using recombinant *Streptomyces lividans* cells expressing CYP105A2, the maximum yield of 25(OH)VD₃ was 20 mg/L culture after 72 h cultivation time. In addition, a vitamin D₃ hydroxylase from *P. autotrophica* has been identified and classified as CYP107 (known as Vdh), which can convert vitamin D₃ to 1 α -25(OH)₂VD₃ via 25(OH)VD₃ (Fujii et al., 2009). Another vitamin D₃ hydroxylase from *Streptomyces griseolus* has been identified, namely CYP105A1 (Sawada et al., 2004). CYP107 and CYP105A1 are currently applied in the biotechnological production of active vitamin D₃ metabolites (Sakaki et al., 2011). Using site-directed mutagenesis and directed evolution, the catalytic activity of CYP107 and CYP105A1 towards vitamin D₃ has been enhanced (Hayashi et al., 2008; Yasutake et al., 2013). To the best of our knowledge, only these two P450s are employed in a biotechnological process for the production of the active vitamin D₃ metabolites, 25(OH)VD₃ and/or 1 α -25-dihydroxyvitamin D₃, demonstrating the necessity of further biotechnological alternative.

1.9. Aim of this work

Bacillus megaterium (*B. megaterium*) is considered as an ideal industrial organism (Vary et al., 2007). It has been industrially employed to produce proteins like penicillin acylase (Martín et al., 1995), various amylases (Hebeda et al., 1988) and glucose dehydrogenase (Kittsteiner-Eberle et al., 1989). In addition, it has been used for the production of different biotechnological interesting substances like pyruvate and vitamin B₁₂ (Hollmann and Deckwer, 2004; Raux et al., 1998). Previously, the genome sequence of *B. megaterium* DSM319 has been published and it was shown to contain four open reading frames encoding four P450s; CYP102A1, CYP106A1, CYP109E1 and CYP109A2.

This work focuses on the two members of the family CYP109 from *B. megaterium* DSM319 and their application as vitamin D₃ hydroxylase. The biotechnologically important *B. megaterium* strain MS941 was used as expression and conversion host. The goal of the presented work is the production of active vitamin D₃ metabolites, mainly 25-hydroxyvitamin D₃. For this purpose, the action of these two P450s on vitamin D₃ has to be firstly determined. In order to increase the yield of our target product, 25-hydroxyvitamin D₃, protein engineering methods should be applied to improve the activity and/or regio-selectivity of the selected P450s towards vitamin D₃. Hereby, the crystal structures of CYP109E1 and CYP109A2 have to be elucidated in collaboration with the laboratory of biophysical chemistry at the University of Groningen. By means of molecular docking and structural comparison, functionally important residues of both enzymes have to be predicted. Based on this, site-directed mutagenesis should be performed to create a focused mutant library. The enzyme mutants should be screened for higher product formation.

2. Scientific articles

2.1 (Abdulgughni et al., 2017a)

Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D₃ hydroxylase.

Ammar Abdulmughni, Ilona K. Jóźwik, Natalia Putkaradze, Elisa Brill, Josef Zapp, Andy-Mark W. H. Thunnissen, Frank Hannemann and Rita Bernhardt.

Journal of Biotechnology, 2017 Feb; 243: 38 –47.

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



Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D₃ hydroxylase



Ammar Abdulmughni^a, Ilona K. Jóźwik^b, Natalia Putkaradze^a, Elisa Brill^a, Josef Zapp^c, Andy-Mark W.H. Thunnissen^b, Frank Hannemann^{a,*}, Rita Bernhardt^{a,*}

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

^b Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

^c Pharmaceutical Biology, Campus C2.2, 66123, Saarland University, Saarbrücken, Germany

ARTICLE INFO

Article history:

Received 22 September 2016

Received in revised form

26 December 2016

Accepted 28 December 2016

Available online 30 December 2016

Keywords:

Bacillus megaterium

CYP109E1

Whole-cell conversion

Vitamin D₃

25-Hydroxy-vitamin D₃

Site-directed mutagenesis

ABSTRACT

In this study the ability of CYP109E1 from *Bacillus megaterium* to metabolize vitamin D₃ (VD₃) was investigated. In an *in vitro* system using bovine adrenodoxin reductase (AdR) and adrenodoxin (Adx₄₋₁₀₈), VD₃ was converted by CYP109E1 into several products. Furthermore, a whole-cell system in *B. megaterium* MS941 was established. The new system showed a conversion of 95% after 24 h. By NMR analysis it was found that CYP109E1 catalyzes hydroxylation of VD₃ at carbons C-24 and C-25, resulting in the formation of 24(S)-hydroxyvitamin D₃ (24S(OH)VD₃), 25-hydroxyvitamin D₃ (25(OH)VD₃) and 24S,25-dihydroxyvitamin D₃ (24S,25(OH)₂VD₃). Through time dependent whole-cell conversion of VD₃, we identified that the formation of 24S,25(OH)₂VD₃ by CYP109E1 is derived from VD₃ via the intermediate 24S(OH)VD₃. Moreover, using docking analysis and site-directed mutagenesis, we identified important active site residues capable of determining substrate specificity and regio-selectivity. HPLC analysis of the whole-cell conversion with the I85A-mutant revealed an increased selectivity towards 25-hydroxylation of VD₃ compared with the wild type activity, resulting in an approximately 2-fold increase of 25(OH)VD₃ production (45 mg l⁻¹ day⁻¹) compared to wild type (24.5 mg l⁻¹ day⁻¹).

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Cytochromes P450 (P450s) are heme-containing enzymes found in all domains of life (Nelson, 2011). They are involved in many metabolic processes, including the biosynthesis of steroids and fatty acids, the metabolism of drugs and the detoxification of xenobiotics (Bernhardt, 2006). P450 monooxygenases are gaining importance as enzymes for industrial biotechnology since they have the ability to introduce oxygen into non-activated C–H bonds of various compounds in a regio- and stereo-selective manner under mild conditions (Bernhardt and Urlacher, 2014; Urlacher and Girhard, 2012).

Vitamin D₃ (VD₃) is a fat-soluble prohormone, which is synthesized in the presence of ultraviolet radiation from the precursor 7-dehydrocholesterol (Holick et al., 1979; Kametani and Furuyama, 1987). The activation of VD₃ is achieved by different P450s: the mitochondrial CYP27A1 mediates hydroxylation of VD₃ at carbon

25, producing 25-hydroxyvitamin D₃ (25(OH)VD₃), which is then further hydroxylated by CYP27B1 resulting in the most active form of VD₃, i.e., 1α,25-dihydroxyvitamin D₃ (1α-25(OH)₂VD₃) (Prosser and Jones, 2004; Schuster, 2011). It was found that other P450s such as the microsomal CYP2R1, CYP3A4 and CYP2J3 can also hydroxylate VD₃ at C-25 (Cheng et al., 2014; Gupta et al., 2004).

The active form of VD₃, 1α-25(OH)₂VD₃, is involved in the regulation of the calcium and phosphate metabolism, amongst other physiological processes (Sakaki et al., 2005; Demay, 2006; Jurutka et al., 2007). However, sufficient levels of the precursor 25(OH)VD₃ are required for the regulatory action of 1α-25(OH)₂VD₃ (Di Rosa et al., 2011). Moreover, 25(OH)VD₃ represents the most abundant VD₃ circulating metabolite and, therefore, is used clinically as an indicator for the VD₃ status of patients (Hollis, 2005).

During the last years there has been a growing interest in the biotransformation of VD₃ to its active metabolites, 25(OH)VD₃ and 1α-25(OH)₂VD₃. Thereby, recent research activity focused on microbial P450s (Sakaki et al., 2011). However, only few bacterial P450s are known to produce these active metabolites such as CYP105A1 from *S. griseolus* (Sasaki et al., 1991) and CYP107 (Vdh) from *P. autotrophica* (Fujii et al., 2009). Therefore, the identification

* Corresponding authors.

E-mail addresses: f.hannemann@mx.uni-saarland.de (F. Hannemann), ritabern@mx.uni-saarland.de (R. Bernhardt).

of new microbial P450s with 1- α and/or 25-hydroxylation activity towards VD₃ is of great interest.

Recently, CYP109E1 from *Bacillus megaterium* DSM319 was identified and characterized in our group (Jóźwik et al., 2016). It was shown that CYP109E1 has a 16 β -hydroxylation activity towards testosterone. In addition, the X-ray crystal structures of CYP109E1 were solved for substrate-free protein and in complexes with testosterone or corticosterone. In the absence of bound steroids, CYP109E1 contains a large, open active site pocket at the distal side of the heme. The testosterone-bound CYP109E1 structure shows a different conformation, in which the active site pocket is more narrow (closed state of CYP109E1) (Jóźwik et al., 2016), which likely reflects the protein's functionally relevant state and therefore was applied in this study for vitamin D₃ docking calculations.

In this study, the substrate specificity of CYP109E1 was investigated for VD₃. Herein, for the first time we present the results demonstrating that CYP109E1 from *B. megaterium* exhibits hydroxylation activity towards VD₃. In addition, whole-cell conversion of VD₃ was carried out using *B. megaterium*. Furthermore, site-directed mutagenesis based on docking simulations of CYP109E1 and VD₃ was performed in order to optimize the regio-selectivity of CYP109E1 towards 25-hydroxylation. The effect of the mutations on the conversion of VD₃ was examined in the whole-cell system.

2. Materials and methods

2.1. Chemicals

VD₃, 25(OH)VD₃, 1 α -25(OH)₂VD₃, (2-hydroxypropyl)- β -cyclodextrin and saponin (from quillaja bark) were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). Isopropyl β -D-1-thiogalactopyranoside (IPTG) and 5-aminolevulinic acid were purchased from Carbolution chemicals (Saarbrücken, Germany). Bacterial media were purchased from Becton Dickinson (Heidelberg, Germany). All other chemicals were from standard sources and of highest purity available.

2.2. Bacterial strains and plasmids

Cloning experiments were carried out with *E. coli* Top10 (Invitrogen, San Diego, USA). The *E. coli* strain C43 (DE3) for the heterologous protein expressions was purchased from Lucigen Corporation (Wisconsin, USA). Whole-cell conversions were carried out using *B. megaterium* MS941 (Wittchen and Meinhardt, 1995; Stammen et al., 2010), pET17b (Merck Bioscience, Bad Soden, Germany) and pSMF2.1 (Bleif et al., 2012) were used for expression purposes in *E. coli* and *B. megaterium*, respectively.

2.3. Cloning of CYP109E1

The coding region for CYP109E1 (GenBank GeneID 9119265) was amplified by polymerase chain reaction (PCR) using genomic DNA of *B. megaterium* MS941 as template. The coding region of CYP109E1 was cloned into the *Spe*I and *Kpn*I restriction sites of pSMF2.1, yielding pSMF2.1.CYP109E1. The gene of CYP109E1 with a hexahistidine tag at the C-terminus was cloned into *Nde*I and *Kpn*I restriction sites of pET17b, yielding pET17b.CYP109E1.

2.4. Site-directed mutagenesis

The mutants of CYP109E1 were generated by the QuikChange site-directed mutagenesis method using the plasmid pSMF2.1.CYP109E1 as template and Phusion DNA polymerase (Thermo Fisher Scientific GmbH, Dreieich, Germany). The PCR primers (MWG-Biotech AG, Ebersberg, Germany) were designed

Table 1

Oligonucleotides used in PCR to generate CYP109E1 mutants.

Primer name	oligonucleotides
I85A-for	5'-ACGAGCCTAGCTAATATTGATCCGCTAAG-3'
I85A-rev	5'-CTTAGCGGATCAATATTAGCTAGGCTCGT-3'
I168A-for	5'-TCGGATATTGCCGTAGCCGGTCTTCTAATAACGAACGT-3'
I168A-rev	5'-ACGTTCTGTTATTAGAAGGACCGGTACGGCAATATCCGA-3'
V169A-for	5'-GATATTATCGCAGCCGGTCTTCTAATAACGAACGT-3'
V169A-rev	5'-ACGTTCTGTTATTAGAAGGACCGGTGCGGATAATATC-3'
K187A-for	5'-CTCCAGCAAGAGGCAATGAAAGCAAATGATGAGC-3'
K187A-rev	5'-GCTCATCATTTGCTTTCATTGCTCTTGTCTGGAG-3'
I241A-for	5'-CTATTTTGCTACTGGCTGCTGGAACGAAACCAACAC-3'
I241A-rev	5'-GTGGTGTGTTCTGTTCCAGCAGCAGTAGCAAAATAG-3'

to introduce point mutation at the desired positions. The oligonucleotide primers for mutagenesis are shown in Table 1. The reactions were performed in a 50 μ l volume using a gradient cyclor (PTC-200 DNA Engine cyclor). 20 cycles were carried out as follows: initial denaturation at 95 °C for 30 s, denaturation at 95 °C for 30 s, annealing at 58 °C for 30 s and extension at 72 °C for 4 min. Correct generation of the desired mutations was confirmed by DNA sequencing, carried out by Eurofines-MWG (Ebersberg, Germany).

2.5. Heterologous expression in *E. coli* and purification

To express CYP109E1 and its mutants, *E. coli* C43 (DE3) cells were transformed with the corresponding expression plasmids (pET17b.CYP109E1) and cultured overnight in Luria-Bertani (LB) medium containing ampicillin (100 μ g ml⁻¹) at 37 °C and 140 rpm. The seed culture was used to inoculate a 200 ml Terrific Broth (TB) medium containing ampicillin 100 μ g ml⁻¹ (1:100 dilution) in a 2-l baffled flask. The main culture was grown at 37 °C and 140 rpm. When the OD₆₀₀ reached 0.5, the expression was induced with 1 mM IPTG. 0.5 mM delta-aminolevulinic acid served as heme precursor. The cultures were shifted to 30 °C and 120 rpm for 24 h. The *E. coli* cells were harvested by centrifugation at 4500 rpm for 30 min, and the cell pellets were stored at -20 °C until purification. All purification steps were performed at 4 °C. The cell pellets were resuspended in 50 mM potassium phosphate buffer (pH 7.4) containing 300 mM NaCl and 20% glycerol. Phenylmethylsulphonyl fluoride (PMSF) was added to a final concentration of 1 mM and the suspension was sonicated with a T13-sonotrode for 15 min with an amplitude of 12% and consisted of repeated intervals of 15 s pulse and 15 s pause. Cell free extract was obtained by ultracentrifugation at 30,000 rpm for 30 min. The supernatant was applied to an immobilized metal ion affinity chromatography column (TALON, Takara Bio Europe, Saint-Germain-en-Laye, France) that had been equilibrated with 50 mM potassium phosphate buffer (pH 7.4) containing 300 mM NaCl and 20% glycerol. The column was washed with 5 column volumes of 50 mM potassium phosphate buffer (pH 7.4) containing 300 mM NaCl, 20 mM imidazole and 20% glycerol. The tagged protein was eluted with 50 mM potassium phosphate buffer (pH 7.4) containing 300 mM NaCl, 150 mM imidazole and 20% glycerol. The bovine Adx₄₋₁₀₈ and AdR were expressed and purified as described elsewhere (Sagara et al., 1993; Uhlmann et al., 1994).

2.6. Carbon monoxide (CO) difference spectroscopy

The reduced CO difference spectra of P450 were measured with a double-beam spectrophotometer (UV-2101PC, Shimadzu, Japan). The concentration of P450 was estimated using a molar extinction coefficient of $\epsilon_{450-490} = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ referred to the method of Omura and Sato (1964).

2.7. In vitro conversion of VD_3

A reconstituted *in vitro* system containing CYP109E1 (1 μ M), AdR (3 μ M), Adx₄₋₁₀₈ (20 μ M), MgCl₂ (1 mM), and a cofactor regenerating system with glucose-6-phosphate (5 mM) and glucose-6-phosphate dehydrogenase (1 U) was used in a final volume of 250 μ l in potassium phosphate buffer (20 mM, pH 7.4). The substrate was dissolved in 2-hydroxypropyl- β -cyclodextrin (2.25% w/v) and added to a final concentration of 200 μ M. The reaction was started by addition of 0.5 mM NADPH. After 1 h at 30 °C the reaction was stopped and extracted twice with 2 vol of ethyl acetate. The organic phases were combined, evaporated to dryness and prepared for analysis by high-performance liquid chromatography (HPLC).

2.8. Whole-cell conversion of VD_3

The whole-cell conversions were performed in *B. megaterium* MS941. The cells were transformed with the corresponding pSMF2.1.CYP109E1 plasmid using the polyethylene glycol-mediated protoplast transformation method (Barg et al., 2005). The seed culture was prepared with LB medium (10 μ g/ml tetracycline). The main culture (50 ml TB medium, 10 μ g/ml tetracycline) was inoculated with 500 μ l of seed culture (dilution 1:100) in 300 ml baffled flasks and incubated at 37 °C, 140 rpm. The culture was grown to OD₅₇₈ of 0.4 and recombinant gene expression was induced with xylose (5 mg/ml). The culture was grown further at 30 °C for 24 h with shaking at 140 rpm.

For whole-cell conversion experiments, VD_3 was dissolved in 45% 2-hydroxypropyl- β -cyclodextrin and 4% Quillaja Saponin as membrane solubilizing agent. After 24 h of protein expression, 2.5 ml of the substrate solution were added to the 50 ml culture. A final substrate concentration of 200 μ M was used for all conversion experiments. Afterwards, the conversion was performed for the indicated time at 30 °C and 120 rpm in 300 ml baffled flasks. 500 μ l samples of the cultures were taken after defined time periods, extracted and prepared for HPLC analysis.

Large scale whole-cell conversions for purification of VD_3 metabolites were performed in 21 baffled flasks using 250 ml of main culture. Cultivation of bacteria as well as the whole-cell conversion and extraction were accomplished as mentioned above. After extraction, organic phases were dried using solvent evaporator and stored at –20 °C under protection from UV light until product purification by HPLC.

2.9. High-performance liquid chromatography (HPLC)

The HPLC was carried out on a Jasco system (Pu-980 HPLC pump, AS-950 sampler, UV-975 UV/visible detector, LG-980-02 gradient unit; Jasco, Gross-Umstadt, Germany) equipped with a Nucleodur 100-5 C18 column (125 \times 4 mm; Macherey-Nagel, Düren, Germany). The column temperature was adjusted to 40 °C. The samples were dissolved in 200 μ l acetonitrile. The flow rate was 1 ml/min with a linear gradient of 60–100% aqueous acetonitrile for 15 min followed by 100% acetonitrile for 15 min. The UV detection of the substrate and products was accomplished at 265 nm. The absorption properties of the products did not differ from the substrate and, therefore, the product formation was calculated from the relative peak area (area%) of the HPLC chromatograms, dividing each respective product peak area by the sum of all peak areas.

2.10. Product purification

Purification of the products was carried out with reversed-phase HPLC using a preparative column VP 250/8 NUCLEODUR 100-5 C18ec (Macherey-Nagel, Düren, Germany). First, the dried

extract was dissolved in an acetonitrile/water mixture and filtered through the Rotilabo syringe filters (0.22 μ m, Carl Roth GmbH, Karlsruhe, Germany). For purification of product P2, a linear gradient of 80–100% acetonitrile aqueous solution as a mobile phase for 17 min was applied (UV detection: 265 nm; flow rate: 3.5 – 4 ml/min; column temperature: 40 °C). Products P4 and P5 were purified isocratically using a 65% acetonitrile aqueous solution as a mobile phase for 40 min (UV detection: 265 nm; flow rate: 2.5 ml/min; column temperature: 40 °C). Collected product fractions were combined, evaporated to dryness and analyzed by NMR characterization.

2.11. NMR characterization of the metabolites

The NMR spectra were recorded in CDCl₃ with 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) 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.

2.12. Molecular docking

Vitamin D₃ (ligand) was docked into the active site of CYP109E1 (receptor) in its closed conformation (PDB: 5L94, CYP109E1-TES) with the use of Autodock Vina 1.1.2 (Trott and Olson, 2010). The testosterone molecule and all waters were removed, and the resulting model was used as a template for the docking experiments. Coordinates for the ligand were taken from an available crystal structure (PDB: 3VRM); six bonds were kept rotatable as confirmed by manual inspection in AutoDockTools 1.5.6. Hydrogens and Gasteiger charges were also added in AutoDockTools 1.5.6. The protein was kept rigid during docking. The simulation cell was limited to a grid box centered at the heme iron with sides adjusted to cover the whole distal heme pocket (x:28 Å, y:34 Å, z:48 Å). Docking simulations were done in triplicate and twenty docking poses were generated for each simulation. The binding poses were analyzed according to lowest binding energies and distances of the target carbon atom (C-25) to the heme iron. Ligand binding residues were identified by analysis done with LigPlot+ (Laskowski and Swindells, 2011) and further visualized with ViewDock tool in UCSF Chimera (Pettersen et al., 2004).

3. Results

3.1. Bioconversion of vitamin D₃ by CYP109E1

CYP109E1 from *B. megaterium* was previously cloned and characterized in our laboratory (Jóźwik et al., 2016). In order to identify new substrates for this enzyme, screening of a focused library consisting of different steroids was carried out. Hereby, VD_3 was identified as new substrate for CYP109E1. In an *in vitro* reconstituted system containing CYP109E1, AdR and Adx₄₋₁₀₈, about 90% of 200 μ M VD_3 was converted within 1 h into 7 products with the following distribution: 3%, 15%, 8%, 22%, 42%, 5% and 5% of product 1 to product 7 (P1–P7), respectively (Fig. 1A). Over time, it was observed that the increase in the formation of P2 is related to a corresponding decrease of P5. Therefore, we assumed that P5 is converted by CYP109E1 into P2. Through comparison of the retention time (t_R) of the detected products with that of available authentic standards, P1 (t_R = 4.8 min) and P4 (t_R = 11.6 min) were identified as 1 α -25(OH)₂ VD_3 and 25(OH) VD_3 , respectively (data not shown).

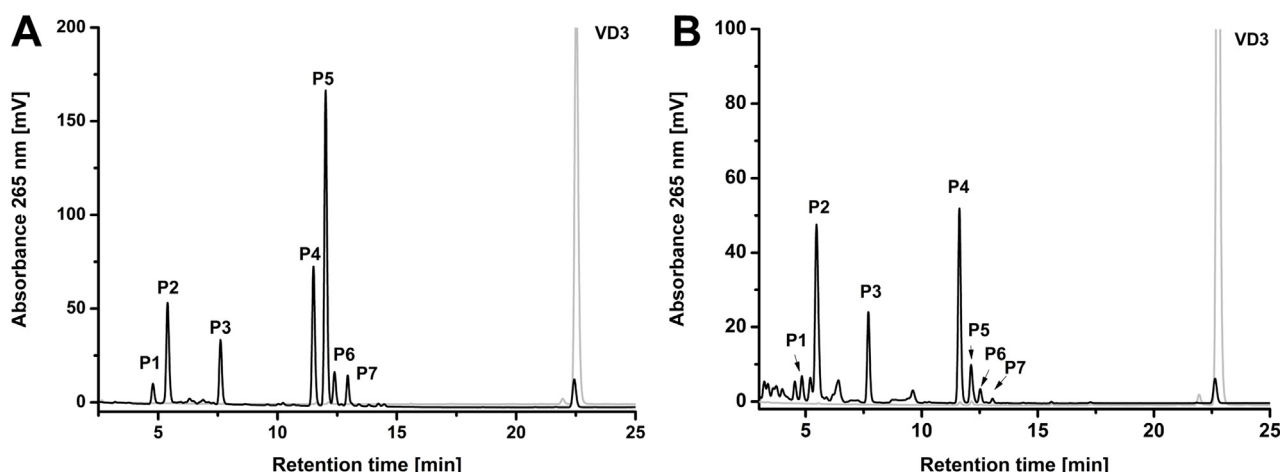


Fig. 1. HPLC chromatogram of the CYP109E1 catalyzed VD_3 conversion. (A) *in vitro* VD_3 conversion, using bovine Adx_{4-108} (20 μM) and AdR (3 μM) and CYP109E1 (1 μM). The reaction was carried out in a final volume of 250 μl at 30 °C for 1 h. (B) CYP109E1-dependent whole-cell conversion of VD_3 in *B. megaterium* MS941. The reaction was carried out in 50 ml TB medium for 24 h at 30 °C. Substrate was added at a final concentration of 200 μM . The peaks of detected products and substrate are labeled with (P1–P7) and VD_3 , respectively. The authentic standard of 200 μM VD_3 (grey) was detected with the same HPLC method.

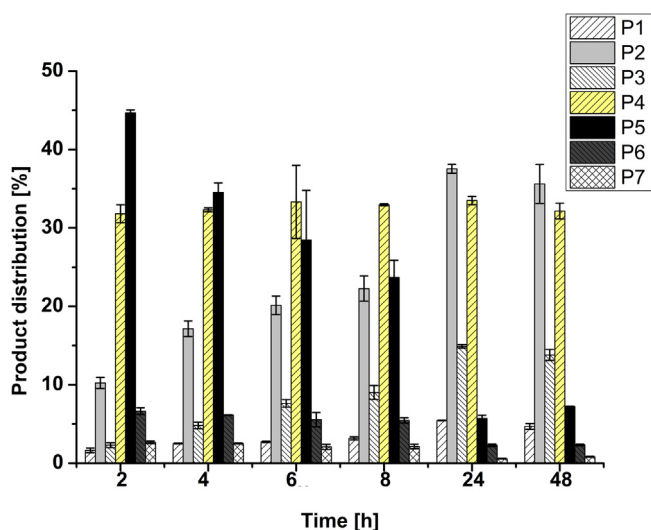


Fig. 2. Product distribution of VD_3 whole-cell conversion in *B. megaterium*. Reactions were carried out in 50 ml TB medium for 48 h at 30 °C. Samples of the cultures were taken after defined time points (2, 4, 6, 8, 24 and 48 h) and analyzed by HPLC as described in “Materials and methods”. Labeling of the products (P1–P7) corresponds to that in Fig. 1. The vertical bars indicate the standard deviation values of the mean from two independent experiments.

3.2. Whole-cell conversion of VD_3 in *B. megaterium*

After successful *in vitro* conversion of VD_3 , a whole-cell conversion system was established. Fig. 1B shows an HPLC profile of VD_3 whole-cell conversion in *B. megaterium* demonstrating a similar pattern as for the *in vitro* conversion. After 24 h, 95% conversion of 200 μM substrate was achieved (Fig. 2). The product distribution after 24 h conversion was as follows: 6%, 37%, 14%, 33%, 6%, 2% and 2% for products P1–P7, respectively.

For further characterization, we analyzed the time course of product formation during whole-cell conversions, which showed that the product distribution changed over time (Fig. 2). In the first 2 h of the reaction, P5 was found to be the main product with 44% of total share (40 $mg l^{-1} day^{-1}$). Afterwards, P5 constituted only 5% of total share (4.3 $mg l^{-1} day^{-1}$). Accompanied by the decrease of P5, an increase of P2 was observed (Fig. 2). The results obtained thus indicate that P2 formation is dependent on the action of CYP109E1

on P5. Therefore, we hypothesized that CYP109E1 has the potency to hydroxylate VD_3 at different positions. To further test, this we decided to identify the main products by nuclear magnetic resonance (NMR) spectroscopy.

3.3. Large scale conversion of VD_3 and product identification

In order to obtain sufficient amounts of VD_3 metabolites for further characterization by NMR spectroscopy, whole-cell conversions with a total culture volume of 750 ml were performed. Three products were obtained with sufficient purity and amounts (5–25 mg) for structural characterization via NMR spectroscopy.

In contrast to vitamin D_3 , its conversion product P4 showed resonances of an additional tertiary hydroxyl group in the ^{13}C NMR spectrum (δ_C 71.15) and the resonances of the methyl groups C-26 and C-27 appeared both as singlets (δ_H 1.19 s, 6H) in the 1H NMR spectrum. This clearly indicated hydroxylation at position C-25 and led to the structure of 25(OH) VD_3 for P4. The data were in accordance with those reported in literature (Helmer et al., 1985; Mizhiritskii et al., 1996):

1H NMR ($CDCl_3$, 500 MHz): δ 0.52 (s, 3xH-18), 0.92 (d, $J=6.5$ Hz, 3xH-21), 1.02 (m, H-22a), 1.06 (m, H-22b), 1.19 (s, 6H, 3xH-26 and 3xH-27), 1.20 (m, H-23a), 1.24 (m, H-16a), 1.28 (m, H-17), 1.29 (m, H-12a), 1.36 (m, H-24a), 1.39 (m, H-23b), 1.45 (m, 2H, H-11a and H-24b), 1.51 (m, H-20), 1.52 (m, 2H, H-11b and H-15a), 1.60 (m, H-2a), 1.63 (m, H-15b), 1.69 (m, H-9a), 1.86 (m, H-16b), 1.94 (m, H-14), 1.96 (m, H-2b), 1.98 (m, H-12b), 2.16 (dd, $J=13.5, 8.5$ and 5.0 Hz, H-1a), 2.26 (dd, $J=13.0$ and 7.5 Hz, H-4a), 2.38 (ddd, $J=13.5, 7.5$ and 4.6 Hz, H-1b), 2.55 (dd, $J=13.0$ and 4.0 Hz, H-4b), 2.82 (m, H-9b), 3.92 (m, H-3), 4.80 (d, $J=2.5$ Hz, H-19a), 5.03 (m, H-19b), 6.01 (d, 11.3 Hz, H-7), 6.21 (d, 11.3 Hz, H-6). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 11.99 (CH₃, C-18), 18.81 (CH₃, C-21), 20.82 (CH₂, C-23), 22.24 (CH₂, C-11), 23.56 (CH₂, C-15), 27.67 (CH₂, C-16), 29.07 (CH₂, C-9), 29.29 (CH₃, C-26), 29.34 (CH₃, C-27), 31.92 (CH₂, C-1), 35.16 (CH₂, C-2), 36.10 (CH₂, C-22), 36.40 (CH, C-20), 40.53 (CH₂, C-12), 44.39 (CH₂, C-24), 45.85 (C, C-13), 45.92 (CH₂, C-4), 56.33 (CH, C-14), 56.53 (CH, C-17), 69.21 (CH, C-3), 71.15 (C, C-25), 112.40 (CH₂, C-19), 117.51 (CH, C-7), 122.45 (CH, C-6), 135.00 (C, C-5), 142.89 (C, C-8), 145.09 (C, C-10).

The NMR spectra of P5 revealed an additional secondary hydroxyl group (δ_H 3.34 m, δ_C 77.41 CH). Its position at C-24 was obvious by vicinal correlations of its proton to the isopropyl proton H-25 (δ_H 1.69 m) in the HHCOSY and to the methyls C-26 (16.69,

CH₃) and C-27 18.93, CH₃) in the HMBC. Comparison of the data with those of an authentic sample (Xi et al., 2014) supported these findings. Especially the chemical shifts of C-1 to C-19 fitted perfectly to each other, whereas the resonances for the side chain slightly differed. This might originate from a different stereochemistry at C-24 in our sample and its epimer from literature. Unfortunately the authors gave no hint to the stereochemistry at C-24 and our sample decomposed within one day in solution. Therefore, the assignment of the absolute configuration at C-24 could not be solved by subsequent NMR measurements, such as Mosher's method. But as was found for closely related 24-hydroxylated steroids the 24(R)- and 24(S)-isomers showed characteristic differences in the ¹³C NMR (Koizumi et al., 1979). Applying these observations to our problem led to the identification of the 24(S)-form for our molecule and to the 24(R)-form for the epimer reported by Xi et al. (2014):

NMR (CDCl₃, 500 MHz): δ 0.57 (s, 3xH-18), 0.92 (d, J = 6.9 Hz, 3xH-26), 0.95 (d, J = 6.9 Hz, 3xH-27), 0.97 (d, J = 6.5 Hz, 3xH-21), 1.08 (m, H-22a), 1.28 (m, H-23a), 1.33 (m, H-16a and H-17), 1.34 (m, H-12a), 1.43 (m, H-20), 1.51 (m, H-11a), 1.56 (m, H-15a), 1.58 (m, H-23b), 1.59 (m, H-11b), 1.64 (m, H-22b), 1.68 (m, H-2a), 1.69 (m, H-25), 1.70 (m, H-15b), 1.72 (m, H-9a), 1.93 (m, H-16b), 1.96 (m, H-2b), 2.01 (m, H-14), 2.03 (m, H-12b), 2.20 (dd, J = 13.5, 8.5 and 5.0 Hz, H-1a), 2.30 (dd, J = 13.0 and 7.5 Hz, H-4a), 2.42 (ddd, J = 13.5, 7.8 and 4.6 Hz, H-1b), 2.60 (dd, J = 13.0 and 4.0 Hz, H-4b), 2.86 (m, H-9b), 3.34 (m, H-24), 3.97 (m, H-3), 4.86 (d, J = 2.5 Hz, H-19a), 5.07 (dt, J = 2.5 and 1.3 Hz, H-19b), 6.06 (d, 11.3 Hz, H-7), 6.26 (d, 11.3 Hz, H-6). ¹³C NMR (CDCl₃, 125 MHz): δ 12.00 (CH₃, C-18), 16.69 (CH₃, C-26), 18.93 (CH₃, C-27), 19.05 (CH₃, C-21), 22.23 (CH₂, C-11), 23.55 (CH₂, C-15), 27.62 (CH₂, C-16), 28.99 (CH₂, C-9), 30.74 (CH₂, C-23), 31.91 (CH₂, C-1), 32.16 (CH₂, C-22), 33.15 (CH, C-25), 35.14 (CH₂, C-2), 36.29 (CH, C-20), 40.50 (CH₂, C-12), 45.84 (C, C-13), 45.90 (CH₂, C-4), 56.30 (CH, C-14), 56.37 (CH, C-17), 69.20 (CH, C-3), 77.41 (CH, C-24), 112.44 (CH₂, C-19), 117.51 (CH, C-7), 122.43 (CH, C-6), 135.07 (C, C-5), 142.21 (C, C-8), 145.05 (C, C-10).

P2 was found to be a dihydroxylated conversion product of vitamin D₃. The NMR spectra revealed resonances for a supplementary secondary (δ_C 79.55 CH) and a tertiary (δ_C 73.15C) hydroxyl function. 2D NMR HHCOSY, HSQCED and HMBC measurements revealed their positions as immediate neighbours at C-24 and C-25. According to P5, C-24 in P2 was expected to be in (S)-configuration. However, we wanted to prove this assumption in an independent manner. Both epimers were known from literature but no comparative NMR studies were available. Therefore, we performed a comparison of the NMR data of their synthetic precursors, the 24(R)- and 24(S)-forms of de-A,B-cholesta-8,24,25-triol (Pérez Sestelo et al., 2002). Analysis of the ¹³C NMR data for the side chain of P2 and comparison with those of the C-24 epimeric de-A,B-cholestanes gave a clear and unambiguous evidence for 24S,25(OH)₂VD₃ as structure for P2:

NMR (CDCl₃, 500 MHz): δ 0.55 (s, 3xH-18), 0.95 (d, J = 6.5 Hz, 3xH-21), 1.04 (m, H-22a), 1.14 (m, H-23a), 1.17 (s, 3xH-26), 1.22 (s, 3xH-27), 1.28 (m, H-17), 1.29 (m, H-16a), 1.31 (m, H-12a), 1.41 (m, H-20), 1.49 (m, 2H, H-11a and H-11b), 1.54 (m, H-15a), 1.57 (m, H-23b), 1.68 (m, H-2a and H-15b), 1.69 (m, H-9a), 1.77 (m, H-22b), 1.89 (m, H-16b), 1.94 (m, H-2b), 1.98 (m, H-14), 2.00 (m, H-12b), 2.21 (m, H-1a), 2.30 (dd, J = 13.0 and 7.5 Hz, H-4a), 2.41 (ddd, J = 13.5, 7.8 and 4.6 Hz, H-1b), 2.58 (dd, J = 13.0 and 4.0 Hz, H-4b), 2.83 (m, H-9b), 3.29 (d, J = 10.1 and 2.0 Hz, H-24), 3.94 (m, H-3), 5.05 (dt, J = 2.5 and 1.3 Hz, H-19b), 4.82 (d, J = 2.5 Hz, H-19a), 6.03 (d, 11.3 Hz, H-7), 6.23 (d, 11.3 Hz, H-6). ¹³C NMR (CDCl₃, 125 MHz): δ 12.00 (CH₃, C-18), 18.94 (CH₃, C-21), 22.23 (CH₂, C-11), 23.17 (CH₃, C-26), 23.54 (CH₂, C-15), 26.51 (CH₃, C-27), 27.61 (CH₂, C-16), 28.34 (CH₂, C-23), 28.99 (CH₂, C-9), 31.95 (CH₂, C-1), 33.22 (CH₂, C-22), 35.19 (CH₂, C-2), 36.26 (CH, C-20), 40.51 (CH₂, C-12), 45.83 (C, C-13), 45.94 (CH₂, C-4), 56.28 (CH, C-14), 56.39 (CH, C-17), 69.20 (CH, C-3), 73.17 (C, C-

25), 79.55 (CH, C-24), 112.42 (CH₂, C-19), 117.56 (CH, C-7), 122.33 (CH, C-6), 135.21 (C, C-5), 142.08 (C, C-8), 145.09 (C, C-10).

These results confirmed our assumption, that CYP109E1 can hydroxylate VD₃ at different positions. As shown above, the ratio of P2 (24S,25(OH)₂VD₃) increased inversely proportional to the ratio of P5 (24S(OH)VD₃) over time, indicating that CYP109E1 has 25-hydroxylation activity towards VD₃ as well as 24S(OH)VD₃. The reaction pathway of VD₃ conversion by CYP109E1 is shown in Fig. 3.

3.4. Molecular docking of VD₃ to CYP109E1

Unfortunately soaking/co-crystallization trials to obtain the structure of the CYP109E1-VD₃ complex proved unsuccessful, therefore docking of the VD₃ molecule was performed to identify potential substrate-binding residues. A structural comparison of CYP109E1 to other P450s converting VD₃ found in the PDB, revealed that the closed conformer of CYP109E1 (PDB: 5L94) is highly similar to the closed state observed for CYP107 (Vdh) crystallized in complex with VD₃ (PDB: 3A50, Yasutake et al., 2010, r.m.s.d of 1.23 Å for 347 Cα atoms). Therefore, the closed conformation of CYP109E1 likely reflects the functionally relevant conformational state of the protein and was chosen in this study for the substrate docking calculations. Among the calculated VD₃ conformations, the one showing a suitable distance of the C-25 atom to the heme iron and the lowest predicted free energy of binding was chosen for further analysis. The docked pose places the aliphatic side chain of VD₃ close to the heme iron, productively for 25-hydroxylation (C-25-Fe distance of 3.7 Å), and is well in agreement with the crystallographically observed VD₃ binding mode in CYP107 (Vdh) (C-25-Fe distance of 4.6 Å). The only hydrophilic group (3β-OH group at the A-ring) of VD₃ is solvent exposed, predicted not to interact with any of the protein side chains; the rest of the VD₃ molecule interacts with hydrophobic residues lining the active site pocket, similarly as in Vdh. Since VD₃ hydroxylation is of outstanding importance for sustainable biotransformation process of the production of active VD₃, we were interested to improve the yield of 25(OH)VD₃. Four amino acids predicted to interact with VD₃ were mutated to alanine to determine their roles in CYP109E1 activity and selectivity towards VD₃: I85 (BC-loop, substrate recognition site 1, SRS1), I168 and V69 (F-helix, SRS2) and I241 (I-helix, SRS4). Additionally, one more residue was chosen for mutagenesis, K187 (SRS3), since its mutation to alanine in CYP109E1 was previously shown to cause a slight decrease in testosterone conversion activity. Thus we considered the possibility that this flexible residue, located at the top of the active site (G helix), might also take part in VD₃ binding (Fig. 4).

3.5. Conversion of VD₃ by CYP109E1 mutants

To investigate the effect of selected mutations on activity and regio-selectivity of CYP109E1 towards VD₃ directly in the whole-cell system, *B. megaterium* cells were transformed with the plasmid pSMF2.1.CYP109E1 containing the corresponding CYP109E1 mutations. The results clearly showed that all CYP109E1 mutants still maintained VD₃ hydroxylation activity. However, activity or/and selectivity of CYP109E1 was affected by the amino acid replacements (Fig. 5).

While the K187A mutant showed the same conversion ratio (97%) as compared with wild type, the activity of CYP109E1 was changed by selected mutations (Fig. 6). All other mutants showed decreased activity in the first 8 h of the reaction. However, the activity in case of I85A, I168A, V169A and I241A mutants was increased afterwards. After 24 h conversion, the I168A mutant exhibited a comparable conversion (96%) as the wild type and the I85A mutant showed a maximum conversion of 87%. On the other hand, a significant decrease of the activity was observed in case of V169A and

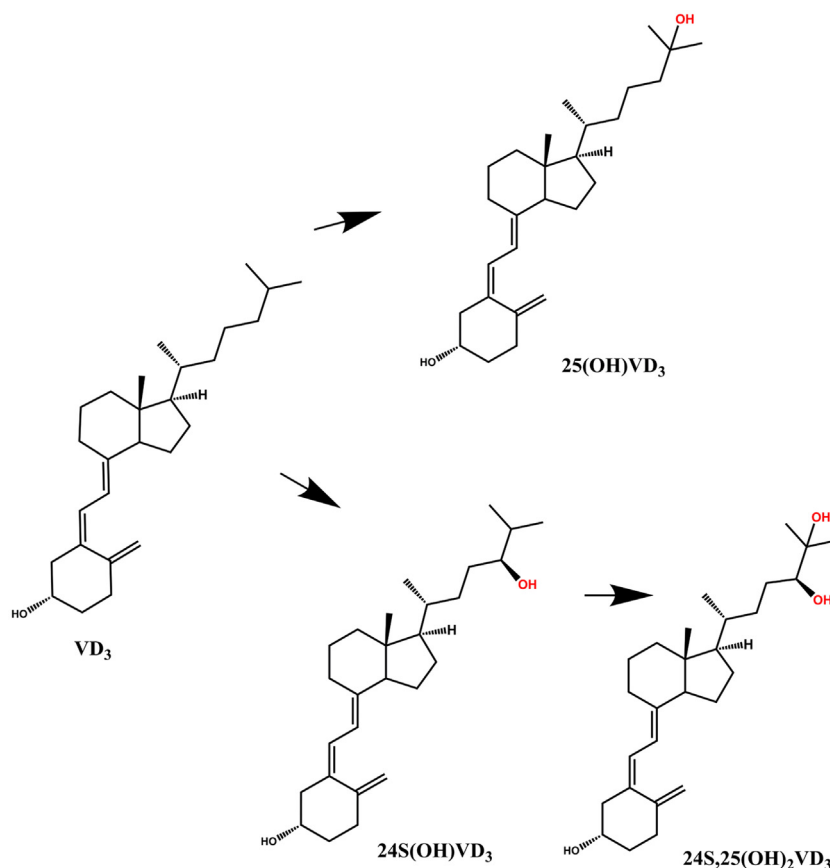


Fig. 3. Reaction pathway of VD₃ by CYP109E1 from *B. megaterium*.

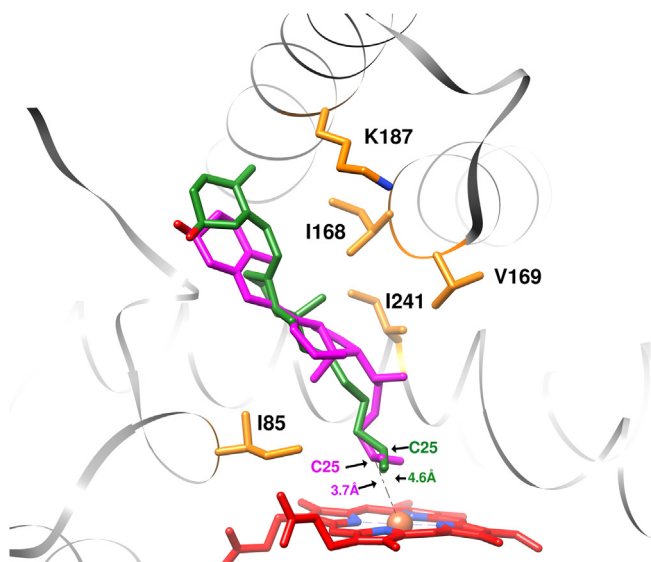


Fig. 4. Docking model of VD₃ in the active site of CYP109E1. The suitable conformation of VD₃ for 25-hydroxylation is shown (in magenta), and compared to the crystallographically observed VD₃ binding mode in CYP107 (Vdh) (in green). Heme is in red coloured sticks. Amino acids selected for site-directed mutagenesis are shown in orange sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

I241A mutants, displaying only 76% and 54% conversion, respectively.

Additionally, the effect of the point mutations on the regio-selectivity of CYP109E1-dependent VD₃ conversion was studied (Table 2). Compared to wild type, the formation of 24S(OH)VD₃

was decreased over time in case of the I85A and I168A mutants, which, as a consequence, led to reduced amounts of the derived product 24S,25(OH)₂VD₃. On the other hand, the 25-hydroxylation activity towards VD₃ was strongly preferred in the reactions catalyzed by the I85A and I168A mutants resulting in 63% and 49% of total products, respectively (Table 2). Moreover, mutant I85A displayed a significant reduction in the number of products. Consequently, an increase of the absolute 25(OH)VD₃ production was observed with these mutants, compared to the wild type (Fig. 7). These results indicate that the side chains of amino acids I85 and I168 are essential for determining the regio-selectivity of CYP109E1 towards VD₃. Compared to wild type, the regio-selectivity of the K187A mutant was also slightly changed towards 25-hydroxylation (Table 2). Furthermore, it was observed that the 24S,25(OH)₂VD₃ production decreased in the whole-cell conversion with the K187A mutant, compared to wild type. In contrast to the decrease in the 24S,25(OH)₂VD₃ production, an accumulation of 24S(OH)VD₃ was observed, suggesting that the K187A mutant has less specificity towards 24S(OH)VD₃, compared to wild type.

Furthermore, significant changes of the specificity and regio-selectivity of CYP109E1 were determined in whole-cell conversions with the V169A and I241A mutants. Compared to the wild type, it was observed that reactions of variant V169A showed a decreased 25-hydroxylation activity of 24S(OH)VD₃, whereas the 25-hydroxylation of VD₃ was enhanced (Table 2). In addition, no formation of P1, P3 and P7 was observed with this mutant. In whole-cell conversions with the V169A mutant, the product distribution was as follows: 24S(OH)VD₃ (40%), 25(OH)VD₃ (47%), 24S,25(OH)₂VD₃ (3%) and P6 (10%). It was further observed that the substitution of I241 to alanine shifts the regio-selectivity of CYP109E1 towards 24-hydroxylation (60% of total products). In

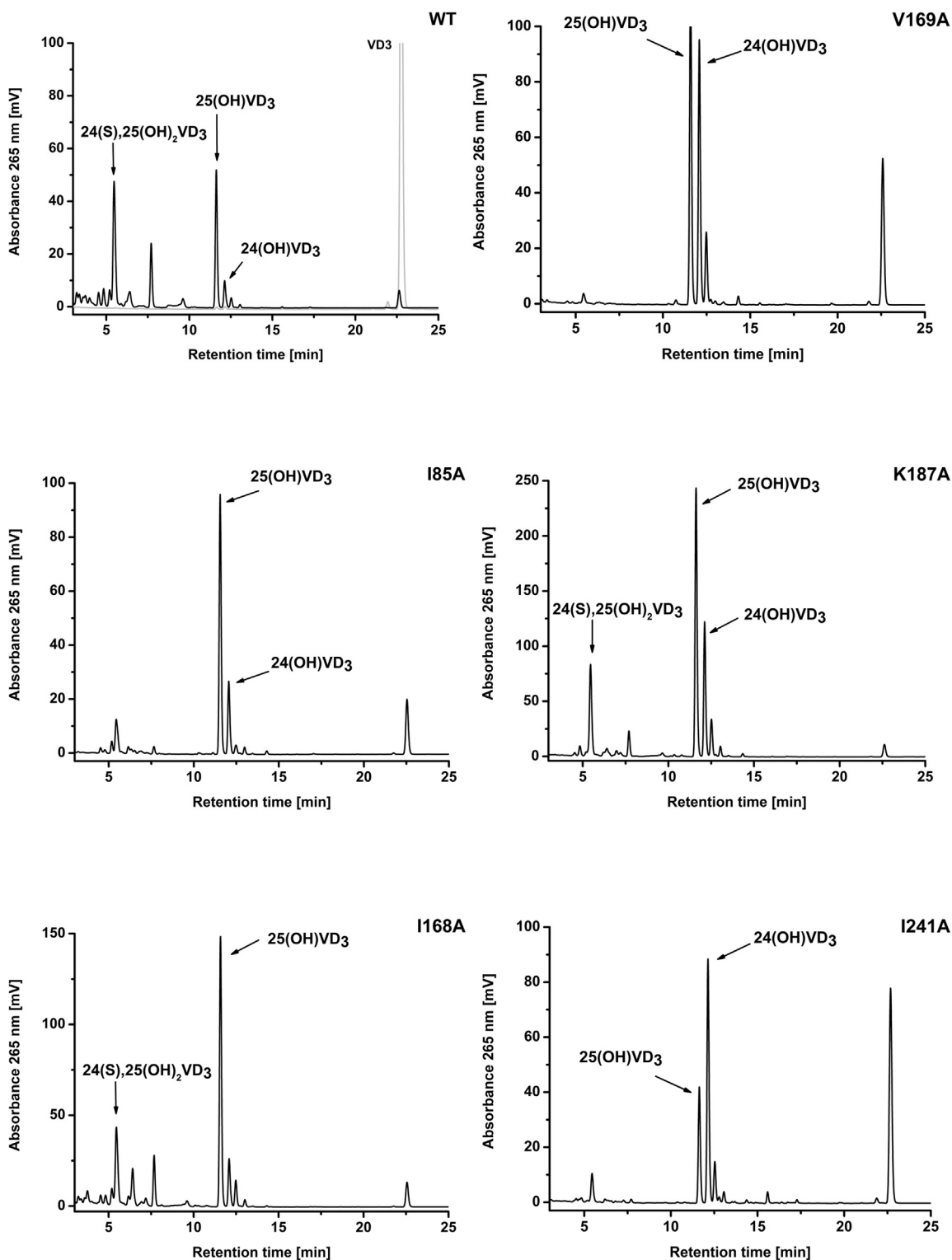


Fig. 5. HPLC chromatograms of VD_3 whole-cell conversions by *B. megaterium* expressing different variants of CYP109E1. The reactions were carried out in 50 ml TB medium for 24 h at 30 °C. The substrate was added at a final concentration of 200 μM . The authentic standard of 200 μM VD_3 (grey) was detected with the same HPLC method.

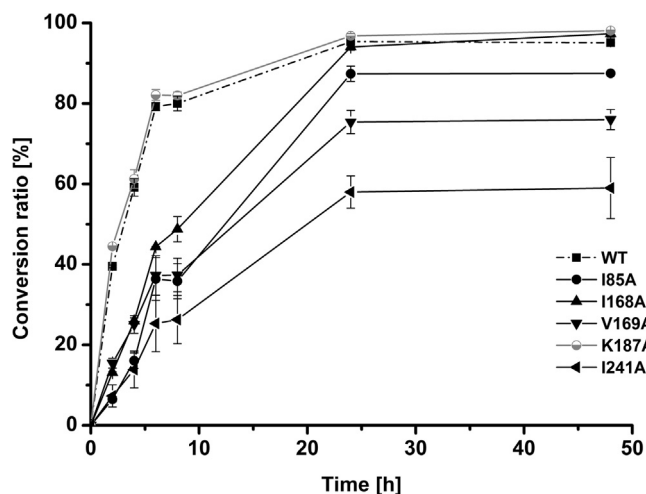
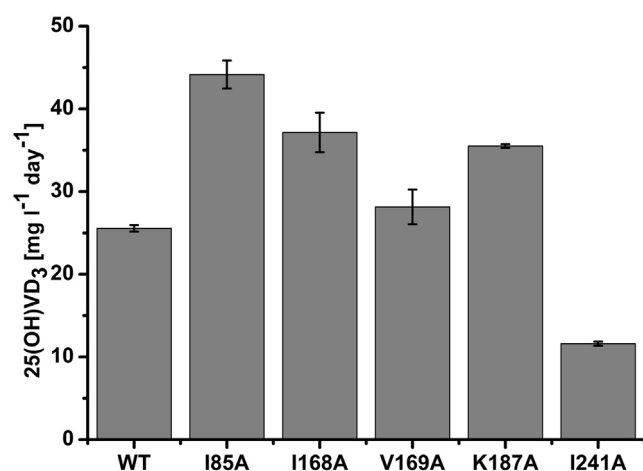
contrast, a decrease of $24\text{S},25(\text{OH})_2\text{VD}_3$ was observed (only 3% of total products) indicating that the 25-hydroxylation activity of this mutant of CYP109E1 towards $24\text{S}(\text{OH})\text{VD}_3$ is dramatically decreased.

4. Discussion

During the past years bioconversion processes, including specific hydroxylations, have gained increasing interest, since chemical synthesis often requires complex procedures and environmentally unfriendly conditions. The ability of P450s to

Table 2Comparison of products distribution of 24 h whole-cell conversions of VD₃ by different variants of CYP109E1.

Enzyme variant	Conversion [%]	Number of products	Major product(s) [%] ^a
Wild type	95	7	24S,25(OH) ₂ VD ₃ (37%) 25(OH)VD ₃ (33%)
I85A	87.3	3	25(OH)VD ₃ (63%)
I168A	94	7	24S,25(OH) ₂ VD ₃ (20%) 25(OH)VD ₃ (49%)
V169A	75.4	3	25(OH)VD ₃ (47%) 24S(OH)VD ₃ (42%)
K187A	96.8	7	25(OH)VD ₃ (45%) 24S(OH)VD ₃ (22%)
I241A	53.8	4	25(OH)VD ₃ (27%) 24S(OH)VD ₃ (58%)

^a Only products with a ratio of ≥20% of total products are defined here as major products.**Fig. 6.** Effect of selected mutations in CYP109E1 on the conversion of VD₃. Reactions were performed using *B. megaterium* MS941 in 50 ml TB medium at 30 °C. The substrate was added at a final concentration of 200 μM. Samples of the cultures were taken after defined time points (2, 4, 6, 8, 24 and 48 h) and analyzed by HPLC as described in “Materials and methods”. The vertical bars indicate the standard deviation values of the mean from three independent whole-cell experiments.**Fig. 7.** Production of 25(OH)VD₃ in *B. megaterium* overexpressing different variants of CYP109E1. Reactions were performed in 50 ml TB medium for 24 h at 30 °C. The substrate was added at a final concentration of 200 μM. The amount of 25(OH)VD₃ was measured by HPLC as described in “Materials and methods”. The vertical bars indicate the standard deviation values of the mean from three independent whole-cell experiments.

hydroxylate a broad range of compounds makes them suitable as versatile biocatalysts (Bernhardt, 2006; Bernhardt and Urlacher, 2014).

It is known that VD₃ is activated in kidneys and liver by different P450s such as mitochondrial cytochromes CYP27A1 and CYP27B1 as well as the microsomal enzymes CYP2R1, CYP3A4 and CYP2J3. However, low activity and stability of mammalian P450s compared with those of bacterial origin are important factors limiting their industrial applications (Julsing et al., 2008). As a result of such limitation, bacterial P450s constitute an attractive alternative for the industrial production of different valuable products, including VD₃ metabolites. Until now, only a few studies have reported the usage of bacterial P450s for the production of VD₃ metabolites such as CYP105A1 and CYP107 (Vdh) (Sakaki et al., 2011). Therefore, the identification of new bacterial P450s with hydroxylation activity towards VD₃ is of a great interest for the industry. In particular, *B. megaterium* offers a great potential, both as a source and expression host, for such new bacterial P450s. It has been used since many decades for the production of different industrial enzymes (Vary et al., 2007; Korneli et al., 2013). It offers an advantage that replicative plasmids are stable and maintained (Stammen et al., 2010). In addition, *B. megaterium* lacks external alkaline proteases and, therefore, is suitable for heterologous protein expression. Moreover, *B. megaterium* can grow on simple media utilizing a broad spectrum of carbon sources (Vary, 1994).

Therefore, the establishment of *B. megaterium*-based whole cell system using a bacterial P450 as biocatalyst for the production of active VD₃ metabolites, such as 25(OH)VD₃, is an interesting alternative for the industrial production. Recently, *B. megaterium* was used in our laboratory for the conversion of the steroid hormone precursor cholesterol to pregnenolone and for the hydroxylation of 11-keto-β-boswellic acid (KBA) (Gerber et al., 2015; Bleif et al., 2012; Brill et al., 2014). Moreover, a CYP27A1-based whole-cell system was established in *B. megaterium* allowing efficient production of valuable pharmaceuticals such as 27-hydroxycholesterol, 26/27-hydroxy-7-dehydrocholesterol, 25(OH)VD₃ and 25-hydroxy-7-dehydrocholesterol (Ehrhardt et al., 2016).

In this study, we demonstrated for the first time the conversion of VD₃ by a member of the CYP109 family, namely CYP109E1 from *B. megaterium* DSM319. We showed that CYP109E1 can convert VD₃ into different products, in both enzyme-based and whole-cell-based assays. The new whole-cell system converts more than 90% of the added substrate (200 μM) within 24 h supporting its promising potential as VD₃ hydroxylase. The two main products of CYP109E1-dependent conversion of VD₃ were identified by NMR analysis as 25(OH)VD₃ (33%) and 24S, 25 (OH)₂VD₃ (37%) with yields of 24.5 mg l⁻¹ day⁻¹ and 28.6 mg l⁻¹ day⁻¹, respectively. In addition, the product P5 was identified by NMR analysis as 24S(OH)VD₃. For further characterization of product formation, time-dependent whole-cell conversions were performed. It was

shown that both, 24S(OH)VD₃ (P5) and 25(OH)VD₃ (P4), are initially detected as major products (Fig. 3). Nevertheless, a decrease in the 24S(OH)VD₃ ratio was observed in correlation with the formation of 24S,25(OH)₂VD₃ (P2). In contrast, the ratio of 25(OH)VD₃ was relatively constant over time (30–33% of total products). Thus we conclude that the formation of 24S,25(OH)₂VD₃ from VD₃ proceeds via the intermediate 24S(OH)VD₃. This is to the best of our knowledge a novel reaction pathway of VD₃ leading to 24S,25(OH)₂VD₃ via 24S(OH)VD₃. In addition, these results demonstrate that CYP109E1 has a 25-hydroxylation activity towards VD₃ as well as 24S(OH)VD₃.

It is known that the hydroxylation of VD₃ at C-25 is the first step of the VD₃ activation (Prosser and Jones, 2004; Sakaki et al., 2005; Schuster, 2011). 25(OH)VD₃ is the best indicator of the nutritional status of VD₃ in the circulation (Hollis, 2005). This compound has been gaining importance in recent years as it was shown to be a better therapeutic agent for several diseases than VD₃ itself, due to its direct biological effect and better intestinal absorption (Jean et al., 2008; Leichtmann et al., 1991). It was found that supplementation of 25(OH)VD₃ has a preventive, therapeutic effect against diseases such as hyperglycemia, chronic kidney disease, Crohn's and cholestatic liver disease (Jean et al., 2008; Leichtmann et al., 1991). In addition to applications of this metabolite for human health, it is used as supplement in animal feed (Soares et al., 1995). Whereas specific biological activities for the natural VD₃ metabolite 24R,25(OH)₂VD₃ were identified (Norman et al., 1983; St-Arnaud and Glorieux, 1998; Yamato et al., 1989), so far no functions are known for its unnatural epimer 24S,25(OH)₂VD₃, which was identified here in the CYP109E1-based whole cell conversion of VD₃.

Our aim in the next step of this study was the enhancement of the 25(OH)VD₃ production yield by CYP109E1. Since the crystal structure of a VD₃-bound CYP109E1 complex is not available, the putative substrate-binding residues were predicted with docking simulations. It is anticipated that VD₃ binds with its aliphatic side chain close to the heme iron and the C-25 atom is in a suitable distance for hydroxylation (Fig. 4). High similarity of the CYP109E1-VD₃ model with the experimental CYP107 (Vdh)-VD₃ crystal structure gives credence to the reliability of our docking results. A similar set of hydrophobic residues is predicted to interact with VD₃ in CYP109E1, including I85, I241, I168 and V169, which correspond to I88, I235, L171, and V172 in Vdh (Yasutake et al., 2010). To test their potential role in VD₃ binding and conversion in CYP109E1, these four residues were targeted for mutation to alanines. In addition, residue K187 was chosen for mutagenesis, to test whether this flexible residue may play a role in substrate binding by interacting with the 3β-OH group of VD₃. Previously, it was shown that mutation of this residue to an alanine slightly lowers the conversion rate of testosterone by CYP109E1 (Jóźwik et al., 2016). However, our results show that considering VD₃ hydroxylation, the mutation of K187 to alanine does not influence CYP109E1 activity (Fig. 6). Thus, residue K187 plays no role in VD₃ binding by CYP109E1, in accordance with the docking results and with the lack of substrate binding interactions by the equivalent residue (K180) in the VD₃-bound crystal structure of Vdh.

Mutagenesis of the other four residues (I85, I168, V169 and I241) ultimately yielded an improved regio-selectivity of CYP109E1 towards VD₃. It has been found that all mutants, except I241A, exhibited a higher selectivity towards 25-hydroxylation than the wild type (Table 2). This enhancement of the regio-selectivity resulted in an increase of the 25(OH)VD₃ production upon whole-cell conversion of VD₃ (Fig. 7). Compared to the wild type, approximately a 2-fold increase of the 25(OH)VD₃ yield was achieved with the I85A mutant, 45 mg l⁻¹ day⁻¹ compared to 24.5 mg l⁻¹ day⁻¹. In the literature, the conversion yield of 25(OH)VD₃ by the most widely investigated P450 for vitamin D₃ hydroxylation, CYP105A1, had been initially very low and was

later significantly improved by protein engineering (Sasaki et al., 1991). The most productive mutant of CYP105A1 (R73A/R84A) is reported to produce 8.3 mg l⁻¹ day⁻¹ of 25(OH)VD₃ (Hayashi et al., 2010; Sakaki et al., 2011). Residue I85 of CYP109E1, as mentioned above, is located in SRS1, in the close vicinity of the heme, and is one of the most investigated residues in P450s. The corresponding position in different P450s was described to interact with the P450s ligands and, therefore, to affect the activity and selectivity of the enzyme (Gricman et al., 2015). The importance of this position was observed in case of CYP102A1 as the substitutions of F87 caused improved selectivity towards propylbenzene and terpene substrates, among others (Gricman et al., 2015; Li et al., 2001; Seifert et al., 2009). In addition, the substitution of S122 to threonine in CYP1A1 (corresponding position to I85 in CYP109E1 and F87 in CYP102A1) improved the 7-methoxy- and 7-ethoxyresorufin O-dealkylase activity (Liu et al., 2004).

Furthermore, our results showed that the substitution of V169 or I241 to alanine residues significantly diminished the conversion of the 24S(OH)VD₃ intermediate into 24S,25(OH)₂VD₃, as compared to reactions with the wild type enzyme. The functional importance of V169 and I241 was also previously determined for testosterone conversion by CYP109E1 (Jóźwik et al., 2016). It was shown that by substituting either residue to an alanine the activity of CYP109E1 towards testosterone is completely abolished. Therefore, we suggest that V169 and I241 of CYP109E1 are substrate specificity-determining residues.

5. Conclusions

This study describes the identification of a new VD₃ hydroxylase from *B. megaterium*, CYP109E1, with 25- and 24-hydroxylation activity. The established *B. megaterium*-based whole-cell system for the conversion of VD₃ has a promising potential for the biotechnological production of the valuable metabolite 25(OH)VD₃.

In addition, further investigations on CYP109E1 were performed in order to increase the production of 25(OH)VD₃. Based on docking studies, site-directed mutagenesis was performed at selected positions in the active site of the enzyme. We were able to demonstrate the importance of selected active site residues for VD₃ conversion. A change of the regio-selectivity and substrate specificity was observed for most of the mutants. A considerable increase of the production of the valuable metabolite 25(OH)VD₃ was achieved with the I85A mutant. In addition, two novel metabolites of VD₃, 24S(OH)VD₃ and 24S,25(OH)₂VD₃, have been identified, whose potential for further drug development needs to be investigated.

Competing interests

The authors declare that they have no competing interest.

Author contributions

AA designed and carried out the experiments, analyzed and interpreted the results and drafted the manuscript. IJ performed the docking simulations and assisted in drafting the manuscript. NP and EB purified the products for NMR analysis. EB carried out the substrate screening. JZ performed the NMR measurements and structure determination of vitamin D₃ metabolites. AT analyzed and interpreted the results, and assisted in drafting the manuscript. FH and RB designed the project, analyzed and interpreted the results, and assisted in drafting the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This work was kindly supported by the German Federation of Industrial Research Associations (AIF/ZIM project FKZ 2214512AJA). The authors would like to thank Birgit Heider-Lips for the purification of AdR and Adx₄₋₁₀₈. Thanks to Mohammed Milhim for proof reading.

References

- Barg, H., Malten, M., Jahn, M., Jahn, D., 2005. Protein and vitamin production in *Bacillus megaterium*. In: Barredo, J.-L. (Ed.), *Microbial Processes and Products, Methods in Biotechnology*. Humana Press, pp. 205–223.
- 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.
- 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., 2014. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1701–1717.
- Cheng, C.Y.S., Slominski, A.T., Tuckey, R.C., 2014. Metabolism of 20-hydroxyvitamin D3 by mouse liver microsomes. *J. Steroid Biochem. Mol. Biol.* 144 (Part B), 286–293.
- Demay, M.B., 2006. Mechanism of vitamin D receptor action. *Ann. N. Y. Acad. Sci.* 1068, 204–213.
- Di Rosa, M., Malaguarnera, M., Nicoletti, F., Malaguarnera, L., 2011. Vitamin D3: a helpful immuno-modulator. *Immunology* 134, 123–139.
- 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.
- Fujii, Y., Kabumoto, H., Nishimura, K., Fujii, T., Yanai, S., Takeda, K., Tamura, N., Arisawa, A., Tamura, T., 2009. Purification, characterization, and directed evolution study of a vitamin D3 hydroxylase from *Pseudonocardia autotrophica*. *Biochem. Biophys. Res. Commun.* 385, 170–175.
- 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.
- Gricman, L., Vogel, C., Pleiss, J., 2015. Identification of universal selectivity-determining positions in cytochrome P450 monooxygenases by systematic sequence-based literature mining. *Proteins Struct. Funct. Bioinf.* 83, 1593–1603.
- Gupta, R.P., Hollis, B.W., Patel, S.B., Patrick, K.S., Bell, N.H., 2004. CYP3A4 is a human microsomal vitamin D 25-hydroxylase. *J. Bone Miner. Res.* 19, 680–688.
- Hayashi, K., Yasuda, K., Sugimoto, H., Ikushiro, S., Kamakura, M., Kittaka, A., Horst, R.L., Chen, T.C., Ohta, M., Shiro, Y., Sakaki, T., 2010. Three-step hydroxylation of vitamin D3 by a genetically engineered CYP105A1. *FEBS J.* 277, 3999–4009.
- Helmer, B., Schnoes, H.K., DeLuca, H.F., 1985. ¹H nuclear magnetic resonance studies of the conformations of vitamin D compounds in various solvents. *Arch. Biochem. Biophys.* 241, 608–615.
- Holick, M.F., Richtand, N.M., McNeill, S.C., Holick, S.A., Frommer, J.E., Henley, J.W., Potts, J.T., 1979. Isolation and identification of previtamin D3 from the skin of rats exposed to ultraviolet irradiation. *Biochemistry (Mosc.)* 18, 1003–1008.
- Hollis, B.W., 2005. Circulating 25-hydroxyvitamin D levels indicative of vitamin D sufficiency: implications for establishing a new effective dietary intake recommendation for vitamin D. *J. Nutr.* 135, 317–322.
- 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.
- Jean, G., Terrat, J.-C., Vanel, T., Hurot, J.-M., Lorriaux, C., Mayor, B., Chazot, C., 2008. Daily oral 25-hydroxycholecalciferol supplementation for vitamin D deficiency in haemodialysis patients: effects on mineral metabolism and bone markers. *Nephrol. Dial. Transplant.* 23, 3670–3676.
- Julsing, M.K., Cornelissen, S., Bühler, B., Schmid, A., 2008. Heme-iron oxygenases: powerful industrial biocatalysts? *Curr. Opin. Chem. Biol.* 12, 177–186.
- Jurutka, P.W., Bartik, L., Whitfield, G.K., Mathern, D.R., Barthel, T.K., Gurevich, M., Hsieh, J.-C., Kaczmarek, M., Haussler, C.A., Haussler, M.R., 2007. Vitamin D receptor: key roles in bone mineral pathophysiology, molecular mechanism of action, and novel nutritional ligands. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* 22 (Suppl. 2), V2–10.
- Kametani, T., Furuyama, H., 1987. Synthesis of vitamin D3 and related compounds. *Med. Res. Rev.* 7, 147–171.
- Koizumi, N., Fujimoto, Y., Takeshita, T., Ikekawa, N., 1979. Carbon-13 nuclear magnetic resonance of 24-substituted steroids. *Chem. Pharm. Bull. (Tokyo)* 27, 38–42.
- Korneli, C., David, F., Biedendieck, R., Jahn, D., Wittmann, C., 2013. Getting the big beast to work—systems biotechnology of *Bacillus megaterium* for novel high-value proteins. *J. Biotechnol.* 163, 87–96.
- Laskowski, R.A., Swindells, M.B., 2011. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* 51, 2778–2786.
- Leichtmann, G.A., Bengoa, J.M., Bolt, M.J., Sitrin, M.D., 1991. Intestinal absorption of cholecalciferol and 25-hydroxycholecalciferol in patients with both Crohn's disease and intestinal resection. *Am. J. Clin. Nutr.* 54, 548–552.
- Li, Q.S., Ogawa, J., Schmid, R.D., Shimizu, S., 2001. Residue size at position 87 of cytochrome P450 BM-3 determines its stereoselectivity in propylbenzene and 3-chlorostyrene oxidation. *FEBS Lett.* 508, 249–252.
- Liu, J., Ericksen, S.S., Sivaneri, M., Besspiata, D., Fisher, C.W., Szklarz, G.D., 2004. The effect of reciprocal active site mutations in human cytochromes P450 1A1 and 1A2 on alkoxyresorufin metabolism. *Arch. Biochem. Biophys.* 424, 33–43.
- Mizhiritskii, M.D., Konstantinovskii, L.E., Vishkauts, R., 1996. 2D NMR study of solution conformations and complete ¹H and ¹³C chemical shifts assignments of vitamin D metabolites and analogs. *Tetrahedron* 52, 1239–1252.
- Nelson, D.R., 2011. Progress in tracing the evolutionary paths of cytochrome P450. *Biochim. Biophys. Acta* 1814, 14–18. *Proteins Proteomics, Cytochrome P450: Structure, biodiversity and potential for application.*
- Norman, A.W., Leathers, V., Bishop, J.E., 1983. Normal egg hatchability requires the simultaneous administration to the hen of 1 α , 25-dihydroxycholecalciferol and 24R,25-dihydroxycholecalciferol. *J. Nutr.* 113, 2505–2515.
- Omura, T., Sato, R., 1964. The carbon monoxide-binding pigment of liver microsomes. *J. Biol. Chem.* 239, 2370–2378.
- Pérez Sestelo, J., Cornella, I., de Uña, O., Mourriño, A., Sarandeses, L.A., 2002. Stereoselective convergent synthesis of 24,25-dihydroxyvitamin D3 metabolites: a practical approach. *Chem. Wein. Bergstr. Ger.* 8, 2747–2752.
- Petersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612.
- Prosser, D.E., Jones, G., 2004. Enzymes involved in the activation and inactivation of vitamin D. *Trends Biochem. Sci.* 29, 664–673.
- 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., Kagawa, N., Yamamoto, K., Inouye, K., 2005. Metabolism of vitamin D3 by cytochromes P450. *Front. Biosci. J. Virtual Lib.* 10, 119–134.
- 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 Proteom.* 1814, 249–256.
- Sasaki, J., Mikami, A., Mizoue, K., Omura, S., 1991. Transformation of 25- and 1 α -hydroxyvitamin D3 to 1 α , 25-dihydroxyvitamin D3 by using *Streptomyces* sp. strains. *Appl. Environ. Microbiol.* 57, 2841–2846.
- Schuster, I., 2011. Cytochromes P450 are essential players in the vitamin D signaling system. *Biochim. Biophys. Acta* 1814, 186–199.
- Seifert, A., Vomund, S., Grohmann, K., Kriening, S., Urlacher, V.B., Laschat, S., Pleiss, J., 2009. Rational design of a minimal and highly enriched CYP102A1 mutant library with improved regio-, stereo- and chemoselectivity. *Chembiochem. Eur. J. Chem. Biol.* 10, 853–861.
- Soares, J.H., Kerr, J.M., Gray, R.W., 1995. 25-hydroxycholecalciferol in poultry nutrition. *Poult. Sci.* 74, 1919–1934.
- St-Arnaud, R., Glorieux, F.H., 1998. Editorial: 24, 25-dihydroxyvitamin D—active metabolite or inactive catabolite? *Endocrinology* 139, 3371–3374.
- Stammen, S., Müller, B.K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E., Jahn, D., 2010. High-yield intra- and extracellular protein production using *Bacillus megaterium*. *Appl. Environ. Microbiol.* 76, 4037–4046.
- Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461.
- Uhlmann, H., Kraft, R., Bernhardt, R., 1994. C-terminal region of adrenodoxin affects its structural integrity and determines differences in its electron transfer function to cytochrome P-450. *J. Biol. Chem.* 269, 22557–22564.
- Urlacher, V.B., Girhard, M., 2012. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends Biotechnol.* 30, 26–36.
- Vary, P.S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W.-D., Jahn, D., 2007. *Bacillus megaterium*—from simple soil bacterium to industrial protein production host. *Appl. Microbiol. Biotechnol.* 76, 957–967.
- Vary, P.S., 1994. Prime time for *Bacillus megaterium*. *Microbiol. Read. Engl.* 140 (Pt. 5), 1001–1013.
- Wittchen, K.D., Meinhardt, F., 1995. Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl. Microbiol. Biotechnol.* 42, 871–877.
- Xi, Z., Zhai, L., Xie, H., Zeng, Z., 2014. Preparation and identification of the related impurities from biotransformed calcifediol. *Jingxi-Huagong* 31, 979–982.
- Yamato, H., Matsumoto, T., Fukumoto, S., Ikeda, K., Ishizuka, S., Ogata, E., 1989. Effect of 24,25-dihydroxyvitamin D3 on 1,25-dihydroxyvitamin D3 [1,25-(OH)₂D₃] metabolism in vitamin D-deficient rats infused with 1,25-(OH)₂D₃. *Endocrinology* 124, 511–517.
- Yasutake, Y., Fujii, Y., Nishioka, T., Cheon, W.-K., Arisawa, A., Tamura, T., 2010. Structural evidence for enhancement of sequential vitamin D3 hydroxylation activities by directed evolution of cytochrome P450 vitamin D3 hydroxylase. *J. Biol. Chem.* 285, 31193–31201.

2.2 (Abdulgugni et al., 2017b)


Biochemical and structural characterization of CYP109A2, a vitamin D₃ 25-hydroxylase from *Bacillus megaterium*.

Ammar Abdulgugni, Ilona K Jóźwik, Elisa Brill, Frank Hannemann, Andy-Mark W. H. Thunnissen and Rita Bernhardt.

FEBS Journal, 2017 Nov; 284: 3881-3894.

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

Biochemical and structural characterization of CYP109A2, a vitamin D₃ 25-hydroxylase from *Bacillus megaterium*

Ammar Abdulmughni^{1,†}, Ilona K. Jóźwik^{2,†} , Elisa Brill¹, Frank Hannemann¹, Andy-Mark W. H. Thunnissen² and Rita Bernhardt¹

¹ Department of Biochemistry, Saarland University, Saarbrücken, Germany

² Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands

Keywords

Bacillus megaterium; biocatalysis; cytochrome P450; regio-selectivity; vitamin D₃

Correspondence

A.-M. W. H. Thunnissen, Molecular Enzymology Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
Fax: +31 50 3634165
Tel: +31 50 3634209
E-mail: a.m.w.h.thunnissen@rug.nl

and
R. Bernhardt, Institute of Biochemistry, Saarland University, Campus B 2.2, 66123 Saarbrücken, Germany
Fax: +49 681 302 4739
Tel: +49 681 302 4241
E-mail: ritabern@mx.uni-saarland.de

[†]These two authors contributed equally to this study

(Received 27 July 2017, revised 7 September 2017, accepted 19 September 2017)

doi:10.1111/febs.14276

Cytochrome P450 enzymes are increasingly investigated due to their potential application as biocatalysts with high regio- and/or stereo-selectivity and under mild conditions. Vitamin D₃ (VD₃) metabolites are of pharmaceutical importance and are applied for the treatment of VD₃ deficiency and other disorders. However, the chemical synthesis of VD₃ derivatives shows low specificity and low yields. In this study, cytochrome P450 CYP109A2 from *Bacillus megaterium* DSM319 was expressed, purified, and shown to oxidize VD₃ with high regio-selectivity. The *in vitro* conversion, using cytochrome P450 reductase (BmCPR) and ferredoxin (Fdx2) from the same strain, showed typical Michaelis–Menten reaction kinetics. A whole-cell system in *B. megaterium* overexpressing CYP109A2 reached $76 \pm 5\%$ conversion after 24 h and allowed to identify the main product by NMR analysis as 25-hydroxylated VD₃. Product yield amounted to $54.9 \text{ mg} \cdot \text{L}^{-1} \cdot \text{day}^{-1}$, rendering the established whole-cell system as a highly promising biocatalytic route for the production of this valuable metabolite. The crystal structure of substrate-free CYP109A2 was determined at 2.7 Å resolution, displaying an open conformation. Structural analysis predicts that CYP109A2 uses a highly similar set of residues for VD₃ binding as the related VD₃ hydroxylases CYP109E1 from *B. megaterium* and CYP107BR1 (Vdh) from *Pseudonocardia autotrophica*. However, the folds and sequences of the BC loops in these three P450s are highly divergent, leading to differences in the shape and apolar/polar surface distribution of their active site pockets, which may account for the observed differences in substrate specificity and the regio-selectivity of VD₃ hydroxylation.

Database

The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession code 5OFQ (substrate-free CYP109A2).

Enzymes

Cytochrome P450 monooxygenase CYP109A2, EC 1.14.14.1, UniProt ID: D5DF88, Ferredoxin, UniProt ID: D5DFQ0, cytochrome P450 reductase, EC 1.8.1.2, UniProt ID: D5DGX1.

Abbreviations

25(OH)VD₃, 25-hydroxyvitamin D₃; BmCPR, *B. megaterium* cytochrome P450 reductase; DLS, dynamic light scattering; P450, cytochrome P450; PDB, Protein Data Bank; TB, Terrific Broth; VD₃, vitamin D₃; δ -ALA, δ -aminolevulinic acid.

Introduction

Vitamin D₃ (VD₃) is a secosteroid that is produced in the human skin from the precursor compound 7-dehydrocholesterol upon exposure to UV light irradiation, resulting in the opening of the precursor B-ring. Afterward, VD₃ is metabolized successively in the liver and kidneys to form two derivatives that are highly important for human health, the 25-hydroxyvitamin D₃ [25(OH)VD₃] and 1 α ,25-dihydroxyvitamin D₃ [1 α ,25(OH)₂VD₃, Fig. 1]. While the latter derivative is the biologically active form of VD₃, the former one is the most abundant circulating metabolite of VD₃ in the blood and is, therefore, considered to be an optimal indicator of the nutritional status concerning VD₃ in patients [1,2]. Maintaining an optimal level of 25(OH)VD₃ is crucial for reducing the risk of several chronic diseases or cancer conditions [3]. Moreover, the 25(OH)VD₃ metabolite is acknowledged for better absorption properties as compared with VD₃ in certain supplement-based treatments [4,5]. Unfortunately, chemical synthesis of VD₃ derivatives is a challenging, multistep procedure resulting in low product yields, limiting their usefulness as therapeutic dietary supplements [6]. Biotechnological approaches involving enzymes capable of specifically oxidizing VD₃ offer an attractive alternative to circumvent such limitations.

In humans, the two consecutive VD₃ hydroxylation steps, at carbons C-25 and C-1, are catalyzed by proteins belonging to the cytochrome P450 superfamily (P450s or CYPs) [7]. The P450s are heme-containing enzymes that catalyze the monooxygenation of a large variety of organic substrates like steroids, fatty acids, and are generally known to be involved in the oxidative metabolism of xenobiotics and drugs. Considering the broad range of P450 substrates, the enzymes are of

benefit to synthesis of pharmaceutically and biotechnologically valuable compounds, in particular when high regio- and stereo-selectivity of hydroxylation is required. Microbial P450s are favored, as they are soluble and more stable than their eukaryotic counterparts [8–10]. Thus far, only a few bacterial P450 enzymes have been reported that hydroxylate VD₃ to produce the desired 25(OH)VD₃ metabolite, like CYP105A1 from *Streptomyces griseolus* [11] or CYP107BR1 (Vdh) from *Pseudonocardia autotrophica* [12]. Most recently, our investigations led to the identification and characterization of another bacterial VD₃-hydroxylase, cytochrome P450 CYP109E1 from *Bacillus megaterium* DSM319 [13]. The improvement of activity and/or regio-selectivity of the three above-mentioned enzymes by protein engineering was supported by the knowledge of their structure–function relationships, due to determination of their crystal structures in both the open, substrate-free and closed, substrate-bound states [14–18].

Here, we describe the identification and analysis of a new cytochrome P450 enzyme from *B. megaterium* DSM319, classified as CYP109A2, which shares 45% amino acid sequence identity with CYP109E1. Similar to CYP109E1, CYP109A2 exhibits 25-hydroxylation activity toward VD₃, however, with significantly higher regio-selectivity. A whole-cell *B. megaterium*-based VD₃ oxidation system was established for CYP109A2 and resulted in higher yields of 25(OH)VD₃ as compared with the previously reported most productive variant of CYP109E1, i.e., mutant I85A [13]. The three-dimensional structure of CYP109A2 was obtained by X-ray crystallography and compared with the structures of other bacterial P450s that convert VD₃, suggesting putative structural features associated with differences in substrate specificity and regio-selectivity.

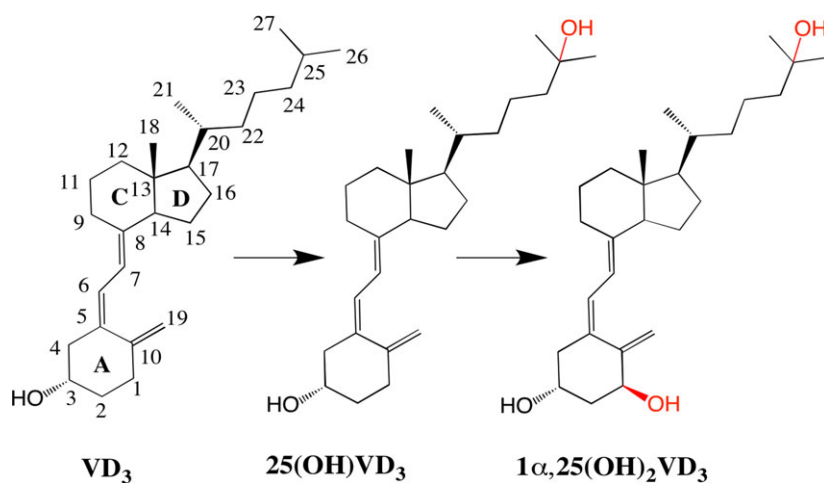


Fig. 1. The activation pathway of VD₃ in mammals. First, VD₃ undergoes a C-25 hydroxylation resulting in the formation of 25(OH)VD₃ (in the liver) followed by the second hydroxylation at C-1 atom producing 1 α ,25(OH)₂VD₃ (in the kidneys).

Results and Discussion

Identification, sequence analysis, and initial characterization of CYP109A2

Bioinformatic mining of the *B. megaterium* DSM319 genome previously led to identification of four genes putatively encoding for cytochrome P450 enzymes (CYP102A1, CYP106A1, CYP109E1, and CYP109A2) [18,19]. Here, we focus on CYP109A2, the remaining P450 that has not been studied in detail yet. Its open-reading frame (KEGG: BMD_2035) encodes for a 403 amino acid long protein containing three characteristic sequence motifs, which support its annotation as a P450 enzyme: (a) heme-binding domain signature proceeding the L helix (FxxGx(H/R)xCxG), (b) the acid-alcohol pair in the I helix (Glu244, Thr245), and (c) ExxR motif in the K helix (Fig. 2). Interestingly, CYP109A2 shares 45% sequence identity with CYP109E1 from the same organism (the sequence identity with CYP106A1 and CYP102A1 is 36% and 24%, respectively). Comparative sequence analysis including proteins from different bacterial sources further shows that CYP109A2 shows significant sequence homology to VD₃ hydroxylating P450s from the CYP107 and CYP105 family [37% identity to CYP107BR1 (Vdh) from *P. autotrophica*, 36% to CYP107CB2 from *Bacillus lehensis* G1, 32% identity to CYP105A1 from *S. griseolus*, and 30% to CYP105A2 from *P. autotrophica*]. Residues in CYP109A2 with the highest degree of conservation have a role in heme binding or belong to substrate recognition site 4 (SRS4) (central part of I helix), whose residues have common roles in P450s associated with substrate binding, oxygen activation, and proton delivery during catalysis [20]. Other regions associated with substrate binding (SRS2, SRS5, and SRS6) are less well conserved, although relatively higher conservation levels may be observed in SRS5 or SRS6 upon pairwise comparison of P450s belonging to the same CYP family, such as CYP109A2 and CYP109E1. The lowest degree of conservation is displayed by the amino acids belonging to SRS1 (the BC loop region) and SRS3 (the G helix) (Fig. 2).

Since a significant sequence similarity between bacterial P450s belonging to the same CYP family may correlate to similarities in substrate scope [21], we set out to establish whether VD₃, being a substrate for CYP109E1, is also converted by CYP109A2. To produce sufficient amounts of pure and functional CYP109A2 enzyme for the intended functional analysis, the *cyp109a2* gene was cloned into the pET17b vector and expressed in *Escherichia coli* C43 (DE3) cells with high yield (4210 nmol·L⁻¹). The protein was purified by affinity chromatography and displayed

characteristic P450 spectroscopic behavior, considering both the UV-visible spectrum of its oxidized ligand-free and reduced CO-bound forms (Fig. 3). Spectroscopic analysis, thus, confirmed that the purified protein contained the correctly incorporated heme group and may be functional as a P450 enzyme.

CYP109A2 catalyzes regio-selective oxidation of vitamin D₃ *in vitro*

Cytochromes P450 are external monooxygenases and thus require electrons for their catalytic activity from different (autologous or heterologous) redox partners [22]. Selection of suitable redox partners to support CYP109A2 activity was based on previous observations that *B. megaterium* cytochrome P450 reductase (BmCPR) and a ferredoxin (Fdx2) were able to support CYP106A1 with much higher efficiency than the commonly used AdR-Adx₄₋₁₀₈ (adrenodoxin reductase – adrenodoxin) redox partners [23]. Indeed, reconstitution of CYP109A2 activity was accomplished using both redox partner systems, but the CYP109A2-catalyzed *in vitro* reactions (200 μM, 1 h) yielded a significantly higher VD₃ conversion ratio when using the BmCPR-Fdx2 redox pair (36 ± 4%) (Fig. 4A) as compared with the AdR-Adx₄₋₁₀₈ pair (20 ± 5%, data not shown). Usage of the different redox pairs (BmCPR-Fdx2 or AdR-Adx₄₋₁₀₈) to reconstitute CYP109A2 activity toward VD₃ resulted in identical product patterns, showing one main product P1 (~ 64% of total products) and one minor product P2 (~ 36% of total products) (Fig. 4A). Determination of the steady-state kinetics of VD₃ conversion by CYP109A2 showed a typical Michaelis-Menten behavior with k_{cat} and K_{M} values of $3.42 \pm 0.15 \text{ min}^{-1}$ and $49.0 \pm 5.8 \text{ μM}$, respectively (Fig. 5). These results demonstrate that the CYP109A2 *in vitro* activity toward VD₃ is lower than that of other previously characterized bacterial VD₃ hydroxylases [24]. In addition, CYP109A2 displays a lower substrate conversion ratio in comparison to its most closely related homolog CYP109E1, which was shown to convert ~ 90% of VD₃ under the same conditions (when supported by the AdR-Adx₄₋₁₀₈ pair) [13]. The affinity for VD₃ is also lower in CYP109A2 as compared with wild-type CYP105A1 or Vdh, which showed K_{M} values of $0.54 \pm 0.09 \text{ μM}$ and $13.5 \pm 3.7 \text{ μM}$, respectively [12,25].

However, the regio-selectivity of CYP109A2, as evident from its product pattern of VD₃ conversion, is superior to that of the other bacterial VD₃ hydroxylases. CYP109A2 converts VD₃ into two products only, whereas previously it was shown that CYP109E1 converts the same substrate to seven products [13]. CYP105A1 and Vdh convert VD₃ to three products

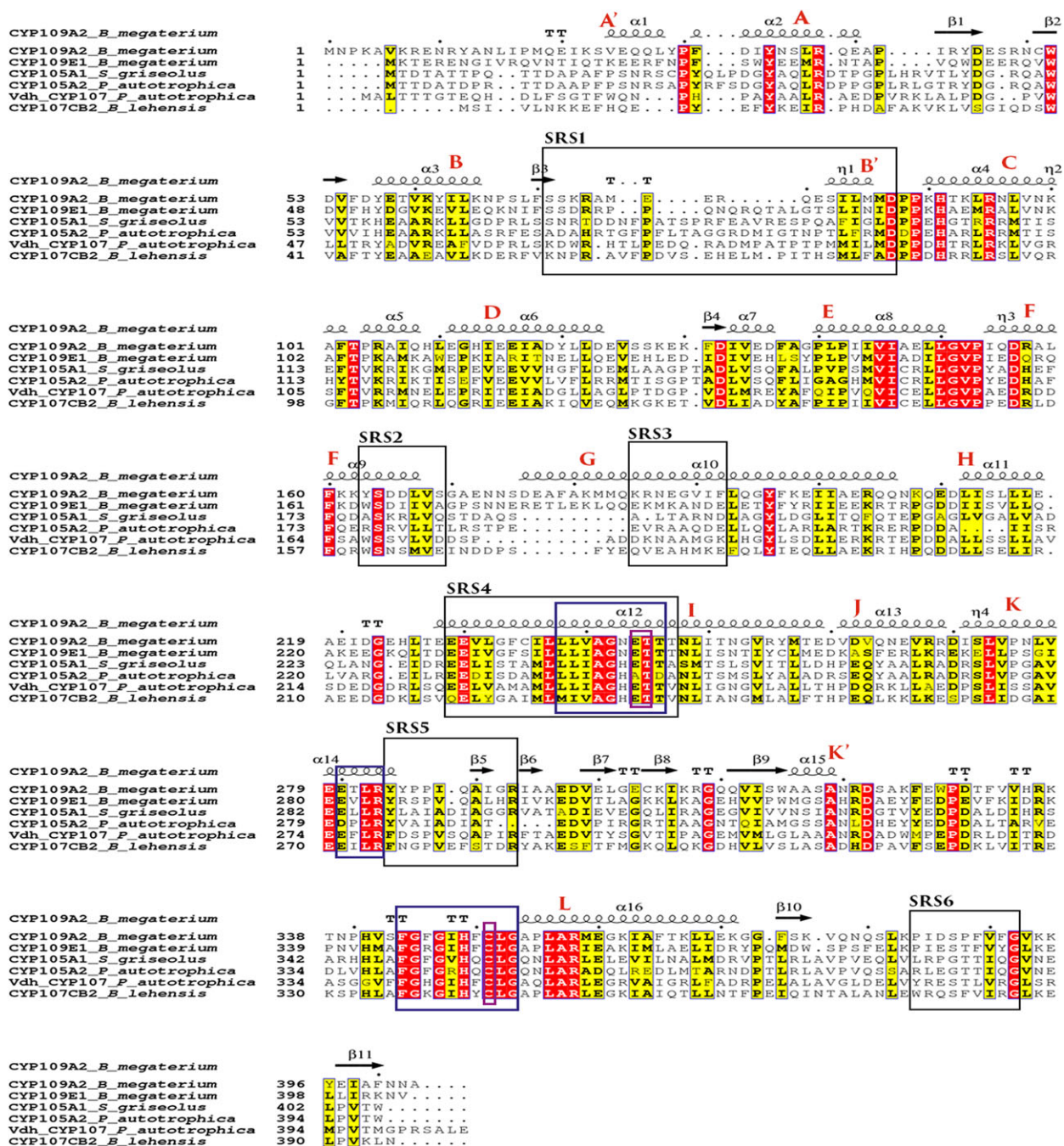


Fig. 2. Multiple sequence alignment of CYP109A2 and other bacterial P450s showing 25-hydroxylation activity with VD₃ or VD₃ derivatives. Secondary structural elements are shown as in the crystal structure of CYP109A2. Conserved and similar residues are highlighted in red and yellow, respectively. Helices are labeled A–L. Residues in the six substrate recognition sites (SRS) are enclosed in black boxes. The characteristic P450 sequence motifs are in blue boxes: the heme-binding domain signature just before the L helix, central part of I helix with the highly conserved ‘acid-alcohol pair’ of residues, and the ‘ExxR’ motif in the K helix.

[12,25], while for the VD₃ 25-hydroxylase CYP105A2 from *P. autotrophica*, the exact product profile is not yet determined [24]. Interestingly, upon comparison of the CYP109A2 product retention times with the

available authentic standards and HPLC data previously obtained for CYP109E1, it was hypothesized and later shown (see below) that product P1 is 25-hydroxyvitamin D₃. Therefore, to the best of our

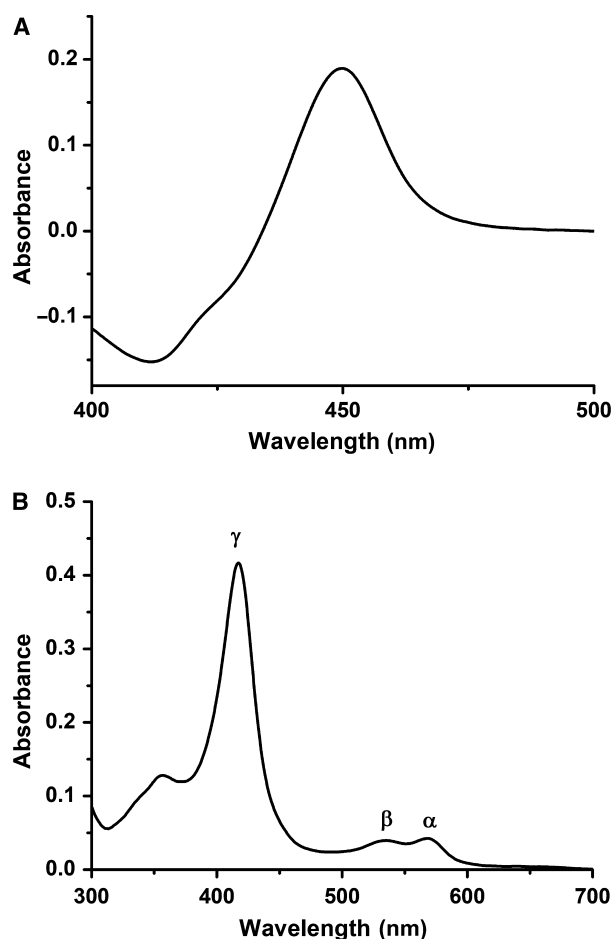


Fig. 3. Spectroscopic properties of CYP109A2. (A) The UV-visible absorbance spectrum of the reduced CO-bound form of CYP109A2 showing the characteristic Soret peak at 450 nm. (B) The UV-visible absorbance spectrum of ligand-free CYP109A2 displaying typical peaks; γ (at 417 nm), α (568 nm), and β (536 nm).

knowledge, the newly identified enzyme, CYP109A2 from *B. megaterium*, is the most regio-selective bacterial VD₃ hydroxylase characterized so far.

Whole-cell-based bioconversion of vitamin D₃ by CYP109A2

Our recent study of CYP109E1 has demonstrated that the *B. megaterium*-based whole-cell system may be an efficient and sustainable route to large-scale production of VD₃ metabolites like 25(OH)VD₃ [13]. *B. megaterium* is a Gram-positive, nonpathogenic bacterium and a well-studied biotechnologically applicable production host. Its use in the pharmaceutical and food industries is preferred due to the ability of this bacterium to grow on a variety of carbon sources and simple media as well as due to its high protein

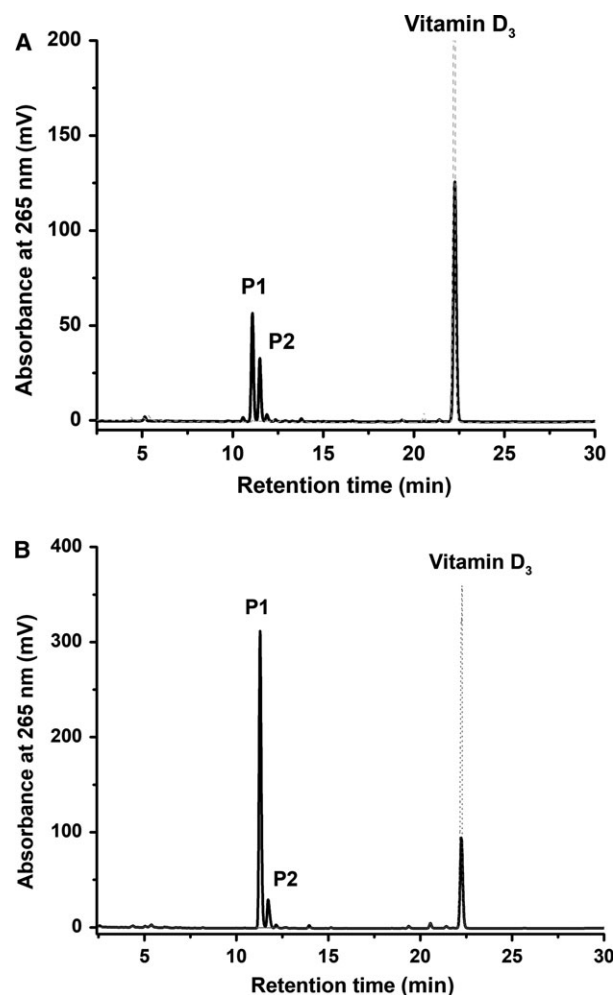


Fig. 4. HPLC analysis of CYP109A2-dependent VD₃ conversion. (A) *In vitro* conversion. (B) Whole-cell conversion in *Bacillus megaterium*. The products are labeled P1-P2. The authentic standard of VD₃ (gray) was detected with the same HPLC method.

production capacity and plasmid stability [26,27]. Importantly, as shown by research results considering various P450 enzymes, the *B. megaterium* offers certain advantages, of being not only a source of new, uncharacterized P450s but also a very good host for P450s and/or their homologous or heterologous redox partners [19,28–30]. Use of the whole-cell expressed P450 (s) in a biotransformation process allows for regeneration of the expensive NADPH cofactor that is necessary for P450 activity.

The *B. megaterium* MS941 cells were transformed with expression plasmid pSMF2.1.CYP109A2, and using such a system $76 \pm 5\%$ of 200 μM VD₃ was successfully converted within 24 h into one main product P1 (90%) and one minor product P2 (10%) (Fig. 4B). In comparison to the profile observed *in vitro*, the

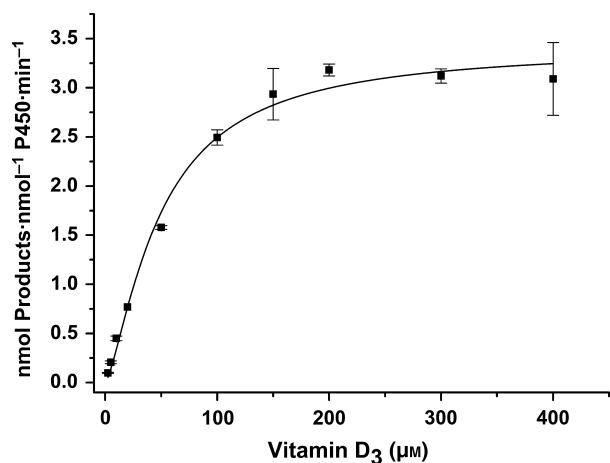


Fig. 5. Determination of the kinetic parameters of VD₃ conversion catalyzed by CYP109A2. Mean values and standard deviations (error bars) were calculated from the results of three independent experiments. The data were fitted to the Michaelis–Menten equation.

product ratio obtained *in vivo* is slightly changed, favoring generation of the main product P1 over product P2. The changes in the product profile are most probably caused by the fact that no heterologous redox partners were expressed along with CYP109A2, leaving only the endogenous *B. megaterium* redox proteins (which remain to be identified) to support its activity. In fact, it was recently observed that varying the redox partners for reconstitution of P450-based reactions can influence the observed product ratios [31,32].

Product isolation and identification by NMR analysis

As for the *in vitro* conversion of VD₃ by CYP109A2, the main hydroxylated product obtained by the *in vivo* conversion showed an identical retention time in HPLC analysis as the authentic standard of 25-hydroxyvitamin D₃. To allow full characterization of the product(s) by NMR spectroscopy, the *in vivo* reactions were up-scaled to 400 mL of total culture volume (200 μM VD₃), and thus 16 mg of the product of interest (P1) was obtained. Unfortunately, the yields of product P2 were not sufficient to allow for the direct product identification by NMR. For the isolated product P1, the ¹H NMR and ¹³C NMR spectra were recorded. It was observed that in comparison to standard VD₃, product P1 showed resonances for an additional tertiary hydroxyl group (at C-25) in the ¹³C NMR spectrum (δ_C 71.15), whereas the resonances of the methyl groups C-26 and C-27 both appeared as singlets (δ_H 1.19 s, 6H) in the ¹H NMR spectrum. This clearly indicated that CYP109A2

catalyzes conversion of VD₃ to 25(OH)VD₃ as the main product (Fig. 6). The NMR data are in accordance with reported literature (Table S1) [33,34].

Efficient biocatalytic production of 25-hydroxyvitamin D₃ by CYP109A2

Based on the results of the *in vivo* conversion of VD₃ performed by the CYP109A2-based whole-cell system, the yield of 25(OH)VD₃ was calculated to be 54.9 mg·L⁻¹·day⁻¹. Hence, it becomes apparent that the production yield of the CYP109A2-based whole-cell system is significantly higher as compared with other bacterial *in vivo* CYP-based systems known to produce the same product. For example, the previously studied CYP109E1 expressed in *B. megaterium* under the same conditions and reaction time produced 24.5 mg·L⁻¹·day⁻¹ of 25(OH)VD₃, whereas with its best mutant I85A, the achieved yield was 45 mg·L⁻¹·day⁻¹ [13]. The yield achieved with the CYP109A2-based system is also higher than the yield reported for the most productive mutant of CYP105A1 (R73A/R84A), which produced 7.8 mg 25(OH)VD₃·L⁻¹·day⁻¹ [14]. Only Vdh expressed in *Amycolata autotrophica* (now referred to as *P. autotrophica*) was reported to give a higher yield than CYP109A2 being 137 mg·L⁻¹ [35]. However, this value was obtained after 72 h, whereas in case of CYP109A2 described here, 54.9 mg·L⁻¹ was obtained already after 24 h. Thus, the *B. megaterium* CYP109A2-based whole-cell system offers a very promising biotechnological alternative to the production of 25-hydroxyvitamin D₃ in comparison to other characterized P450 biocatalysts.

Crystal structure of substrate-free CYP109A2

In order to determine the geometry of its active site pocket and predict the residues responsible for VD₃

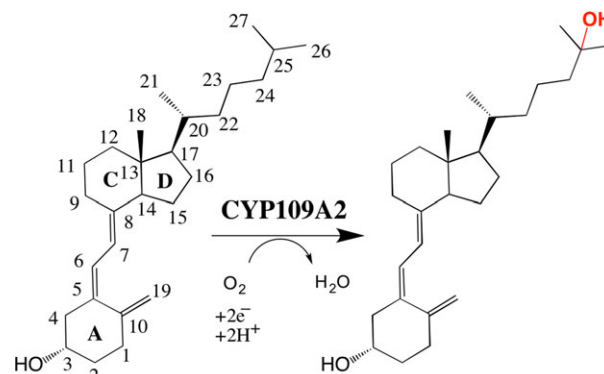


Fig. 6. Scheme of the CYP109A2-catalyzed conversion of VD₃. Hydroxylation site is indicated in red.

binding, the CYP109A2 enzyme was studied by X-ray crystallography. For crystallization, a C-terminally His-tagged protein was purified using the previously described three-step purification procedure, resulting in protein sample of > 98% purity and low polydispersity (12.2%) as confirmed by SDS/PAGE analysis and dynamic light scattering (DLS) measurements (data not shown). The substrate-free CYP109A2 crystallized in an orthorhombic space group ($P2_12_12_1$), with four monomers in the asymmetric unit and a solvent content of 50%. The crystal structure was solved at 2.7 Å resolution by the molecular replacement method, using the previously determined structure of CYP109E1 (sequence similarity 45%, Protein Data Bank [PDB] entry 5L90, [18]) as a search model (see Table 1 for data collection and refinement statistics). All four CYP109A2 molecules have well-defined electron density, except for the 15 or 16 N-terminal residues and the C-terminal His₆-tags, for which no electron density was observed. The central parts of the BC loop, FG loop, and N-terminal end of the G helix show relatively high B-factors, related to high mobility of those regions, as generally observed in P450 crystal structures. The four CYP109A2 molecules found in the asymmetric unit have almost identical conformations [the root mean square deviation (r.m.s.d) of C α -backbone atoms ranges between 0.36 and 0.43 Å]. Therefore, only one of the protein molecules will be described further. The structure of CYP109A2 (residues 16–401) exhibits the typical organization of a P450 domain, containing 15 helices (labeled A–L, A', B', and K') and 10 β -strands (arranged in three β -sheets; in the same structural arrangement as observed in the CYP109E1 crystal structure, Fig. 7A). The centrally located heme group is sandwiched between the I and L helices, similarly as in other P450s. The bound heme is covalently attached to the protein by an iron-sulfur bond with Cys351 and is further stabilized by several residues providing van der Waals and hydrophobic interactions, while its propionate side chains interact with the side chains of His91, Arg95, Arg293, and His349. At the distal side of the heme, no clear electron density was observed for a heme-bound water molecule, in contrast to many other P450 crystal structures determined in the substrate-free state. A small blob of electron density is located at approximately 5 Å distance from the heme-iron, which may correspond to a glycerol molecule from the cryobuffer, but the quality of the electron density did not allow unambiguous ligand identification. In addition, five PEG 400 molecules and three sulfate ions (SO₄²⁻), originating from the crystallization solution, are present in the structure.

Table 1. Crystallographic data collection and refinement statistics of CYP109A2.

	CYP109A2
PDB code	5OFQ
Model statistics	
Monomers in the AU	4
Solvent content (%)	50
Ligands	5 × PEG 400, 3 × sulfate ion
Data collection	
Beamline (ESRF)	ID23-2
Wavelength (Å)	0.873
Resolution range (Å)	52–2.70 (2.78–2.70) ^a
Space group	$P2_12_12_1$
Unit-cell parameters	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	77.12 155.53 158.15
α , β , γ (°)	90, 90, 90
Observed reflections	287 529 (25 387)
Unique reflections	51 368 (4458)
Multiplicity	5.6 (5.7)
Completeness (%)	97.3 (98.5)
$\langle I/\sigma(I) \rangle$	12.4 (1.7)
<i>R</i> _{merge} (%)	0.10 (0.94)
<i>R</i> _{p.i.m.}	0.07 (0.63)
CC (1/2) (%)	99.6 (52.3)
Refinement	
<i>R</i> _{work} (%)	22.4
<i>R</i> _{free} (%)	27.5
R.m.s.d, bond lengths (Å)	0.018
R.m.s.d, bond angles (°)	1.418
Average <i>B</i> -factors (Å) ²	
Overall	50.5
Protein	50.4
Heme	50.0
Ramachandran plot statistics	
Most favored (%)	95.6
Allowed regions (%)	4.4
Disallowed regions (%)	0
Molprobability overall score	1.49

^aValues in parentheses are for the highest resolution shell.

Structural comparison of CYP109A2 overall fold and active site architecture with related vitamin D₃ hydroxylases

CYP109A2 crystallized in an 'open' conformation, which is confirmed by structural alignments against other P450 structures deposited in the PDB, giving the best match to those displaying open active site geometry (Table S2). The highest structural similarity is observed with the 'open' crystal structure of CYP109E1 containing four bound corticosterone molecules in the active site pocket (PDB entry 5L91, r.m.s.d. of 1.27 Å for 368 C α atoms). In addition, a significant structural similarity is observed with the structures of Vdh from *P. autotrophica*, considering

both its wild-type and its mutant forms (L348M and F106V) [36] (e.g. wild-type Vdh, PDB entry 3A4G, r.m.s.d. of 1.44 Å for 349 C α atoms). Visual analysis of the CYP109A2, CYP109E1, and Vdh structures shows that they contain similar, wide-open, funnel-shaped, and solvent-exposed ‘distal’ heme pockets. The main difference concerns the BC loop/SRS1 region (loop connecting the B and C helices, residues 71–87 in CYP109A2), whose length, fold, and sequence vary significantly in the structures of the three different enzymes (Figs 2 and 7B). CYP109A2 contains the shortest BC loop, whereas the longest loop is present in Vdh (Fig. 2). While the BC loop in CYP109A2 folds toward the active site pocket, the loops in CYP109E1 and especially Vdh are folded more outward (Figs 7B and 8A). As a result, the BC loop in CYP109A2 restricts access of substrate to the heme-iron to a larger extent than the BC loops in CYP109E1 and Vdh. The structural differences in the BC loop region further result in a striking difference in the distribution of hydrophobic and polar regions in the active sites of the three P450s (Fig. 8B–D). In the region closest to the heme, the active site pockets in CYP109A2, CYP109E1, and Vdh display a similar, mainly apolar surface due to the presence of conserved hydrophobic residues, i.e., Leu84, Met85 (SRS1); Leu237, Val240, Ala241 (SRS4); Ile288, Ile291 (SRS5); and Phe389, Val390 (SRS6) in CYP109A2. A bit further away from the heme, however, the surfaces show some significant differences. Residues Arg74, Glu78, Arg79, and Glu81 of the BC loop in CYP109A2 form a highly polar surface restricting one side of the active site pocket (Fig. 8B). In the active site pocket of CYP109E1, the BC loop also forms a polar surface, but it is much less pronounced as in CYP109A2 (Fig. 8C), while in Vdh, it is nearly absent (Fig. 8D). The observed differences in BC loop geometry, in the extent of restricted access to the heme-iron and in the apolar/polar surface distribution around the active site pocket, will likely influence the substrate-binding modes in the three different enzymes. In the absence of a closed, VD₃-bound crystal structure of CYP109A2, it remains unknown whether these differences are directly related to the differences in regio-selectivity for VD₃ hydroxylation. Arguably, a significantly more polar BC loop region and a more constricted active site pocket in CYP109A2, in comparison to CYP109E1 or Vdh, allows for lower occurrence of differently stabilized VD₃-binding modes, therefore resulting in an increased regio-selectivity of substrate hydroxylation in comparison to CYP109E1 or Vdh, as observed experimentally. Indeed, the BC loop, a highly variable and flexible

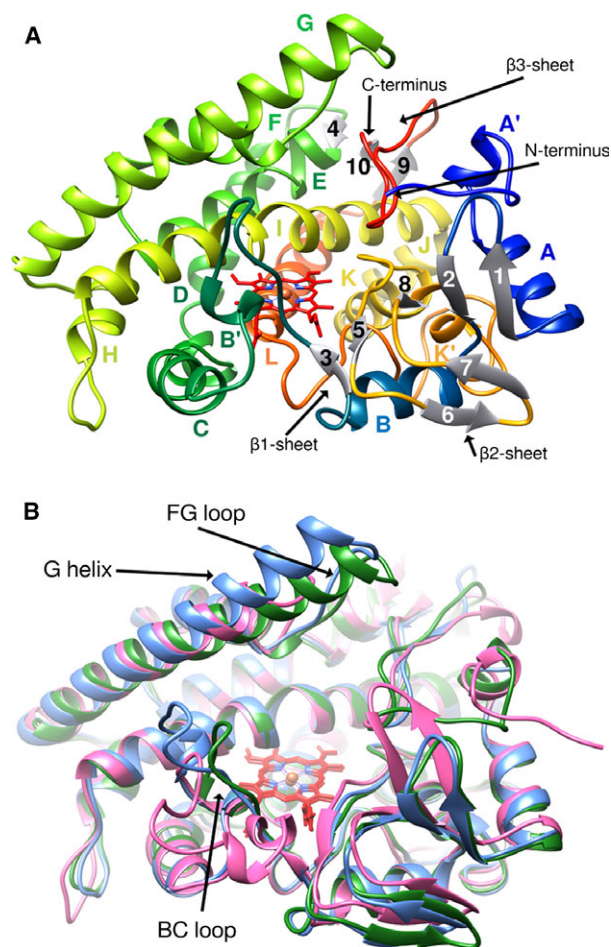


Fig. 7. Crystal structure of substrate-free CYP109A2. (A) Ribbon representation with secondary structure elements labeled following the common P450 nomenclature (rainbow coloring from blue N' terminus to red C' terminus). Helices are labeled A–L following the common P450 nomenclature of secondary structure elements. Heme is shown as red stick model. (B) Superposition of the overall open conformations of CYP109A2, CYP109E1 from *Bacillus megaterium* DSM319 and Vdh from *Pseudonocardia autotrophica*. Cartoon representation of the aligned structures of CYP109A2 (green), CYP109E1 (5L90, blue), and Vdh (3A4G, pink) is shown.

region in P450s, has been previously associated with differences in substrate specificity [37].

Prediction of vitamin D₃-binding residues in CYP109A2

Unfortunately, our experiments to crystalize a structure of CYP109A2 in its closed conformation were not successful, despite extensive efforts to obtain a structure of substrate-bound and/or product-bound CYP109A2.

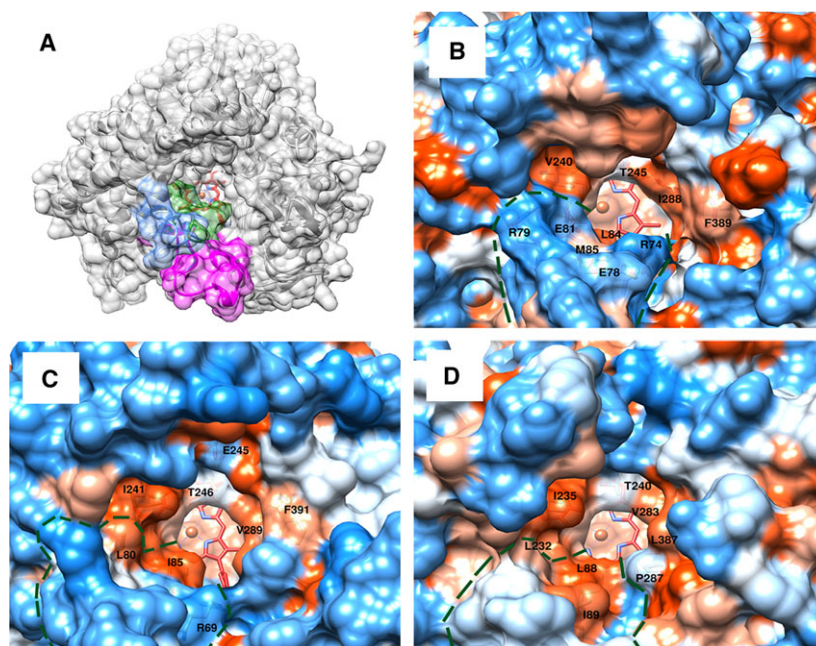


Fig. 8. Comparison of the substrate-binding pockets in CYP109A2, CYP109E1 from *Bacillus megaterium* DSM319 and Vdh from *Pseudonocardia autotrophica*, all presenting an open conformation. (A) Three structures are superimposed and their molecular surfaces are shown in gray. The steric restriction of the active site pocket imposed by the BC loop region is significantly larger in the CYP109A2 structure (surface in green) than in CYP109E1 (5L90, in blue) and Vdh (3A4G, in magenta). (B–D) Surface representation of the ‘distal’ heme substrate-binding pockets in CYP109A2 (B), CYP109E1 (C), and Vdh (D). Apolar residues are colored in orange and polar residues in blue. The different BC loop regions in the three P450s (lower parts of the images) are outlined with a green-dashed line.

The residues known or predicted to bind and stabilize VD₃ in the active sites of CYP109E1 or Vdh were, therefore, compared to their structural equivalents in CYP109A2 (Table 2). Based on this comparison, the following 13 residues of CYP109A2 are predicted to interact with VD₃: Arg74, Leu84, Met85 from SRS1; Leu167, Val168 from SRS2; Arg186, Asn187, Val190 from SRS3; Leu237, Val240, Thr245 from SRS4; Ile291 from SRS5 and Phe389 from SRS6 (Fig. 9). Of these residues, six are identical to those in CYP109E1 or Vdh, while the other residues are relatively similar. This significant degree of conservation further supports the notion that the overall binding mode of VD₃ leading to 25-hydroxylation is similar in these enzymes. Whether the predicted residues in CYP109A2 indeed have important roles in substrate binding and conversion has to await further analysis by site-directed mutagenesis.

In summary, we have shown that cytochrome P450 monooxygenase CYP109A2 from *B. megaterium* is able to catalyze hydroxylation of VD₃ to its 25-hydroxylated derivative, with high regio-selectivity. The activity was reconstituted in the whole-cell-based system, allowing to obtain high yields of the desired 25-hydroxylated product and therefore renders the system as most efficient and regio-selective in comparison to other related bacterial P450s previously shown to oxidize VD₃.

These characteristics highlight the significance of CYP109A2 as a promising, potential industrially applicable biocatalyst. The crystal structure of CYP109A2 allowed for a detailed analysis of the active site geometry and prediction of VD₃ interacting residues. Our study forms the basis to guide future efforts to explore the substrate scope and selectivity of CYP109A2.

Experimental procedures

Materials

VD₃, 25(OH)VD₃, (2-hydroxypropyl)- β -cyclodextrin and saponin (from quillaja bark) were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). IPTG and δ -aminolevulinic acid (δ -ALA) were purchased from Carbolution Chemicals GmbH (Saarbruecken, Germany). Bacterial media were purchased from Becton Dickinson (Heidelberg, Germany). All other chemicals were from standard sources and of highest purity available. The *E. coli* TOP10 strain used for cloning purposes was bought from Invitrogen (San Diego, CA, USA), whereas the *E. coli* C43 (DE3) strain and pET17b vector used for heterologous expression were purchased from Novagen (Merck Bioscience, Bad Soden, Germany). The pSMF2.1 vector was used for protein expression in *B. megaterium* MS94 [38].

Table 2. Identification of residues with predicted VD₃-binding roles in CYP109A2, based on structural comparison to CYP109E1 and Vdh. (A) Residues of CYP109E1 previously predicted to interact with the VD₃ molecule [13] and their structurally equivalent residues in CYP109A2. (B) VD₃-binding residues in the crystal structure of VD₃-bound Vdh (PDB id: 3A50) [16] and their structurally equivalent residues in CYP109A2. Residues of CYP109A2 appearing in both panels are highlighted in blue. The residues of CYP109E1 studied previously by site-directed mutagenesis [13,18] are highlighted in gray. Conserved CYP109A2 residues are shown in bold.

	Predicted VD ₃ -binding residues		VD ₃ - binding residues	
A	CYP109E1	CYP109A2	B	Vdh (3A50)
				CYP109A2
SRS1	Arg69	Arg74	SRS1	Thr84
	Pro71	–		Met86
	Thr78	–		Ile88
	Leu80	–		Leu89
	Ile85	Leu84	SRS2	Leu171
	Asn86	Met85	SRS3	Lys180
SRS2	Ile168	Leu167		Asn181
	Val169	Val168		Met184
SRS3	Asn191	Val190	SRS4	Leu232
SRS4	Leu238	Leu237		Ile235
	Ile241	Val240		Thr240
	Thr246	Thr245	SRS5	Pro287
			SRS6	Leu387
				Phe389

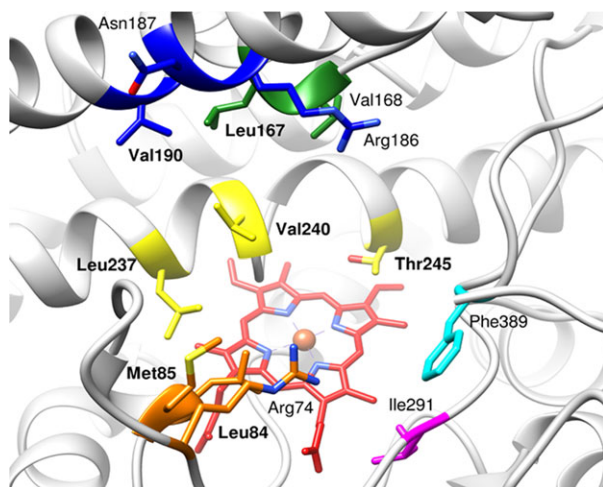


Fig. 9. Predicted VD₃-binding residues in CYP109A2, identified based on structural comparison with CYP109E1 and Vdh (Table 2). The residues are shown as sticks and colored according to the corresponding substrate recognition site (SRS1 orange, SRS2 green, SRS3 blue, SRS4 yellow, SRS5 pink, and SRS6 cyan). Residues overlapping in both predictions are labeled in bold. The heme group is shown as red colored sticks.

Cloning

The *cyp109a2* gene (GenBank GeneID 9117426) was amplified by PCR using genomic DNA of *B. megaterium* MS941 as template (mutant derived from DSM319 strain) [39,40]. To allow expression of CYP109A2 carrying a C-terminal His₆-tag, the *cyp109a2* gene was extended with a sequence at the 3'-end coding for the desired tag and cloned into the expression vector pET17b with the *NdeI/KpnI* restriction sites yielding pET17b.CYP109A2. The primer pair used was 5'-ATATCATATGAATCCAAAAGCAGT-GAAAAGAGAAAATCG-3' and 5'-TTAGGTACCTCAATGGTGGTGGTGATGATG-GGCGTTATTTAAAGCAATTC-3'. The *cyp109a2* gene was also cloned into the pSMF2.1 vector using *PacI/SpeI* restriction sites and including an optimized ribosomal-binding site for *B. megaterium*, yielding pSMF2.1.CYP109A2. The primer pairs used were 5'-CGCTTAATTAATAAATCAAGGAGGTGAATGTACAATGAATCCAAAAGCAGTGAAGAGAG-3' and 5'-TTATCAACTAGTTCAGGCGTTATTTAAAGCAATTCATATA-TTTTACC-3'.

Heterologous gene expression and purification

The *E. coli* C43 (DE3) cells serving as expression host were transformed with the pET17b.CYP109A2 construct and cultured overnight in LB medium containing 100 µg·mL⁻¹ ampicillin, at 37 °C and 140 r.p.m. The overnight culture was used to inoculate a 250 mL main culture [Terrific Broth (TB) medium] containing 100 µg·mL⁻¹ ampicillin. The main culture was grown at 37 °C and 140 r.p.m. until the optical density at 600 nm reached 0.6. Gene expression was induced by addition of 1 mM IPTG and 0.5 mM δ-ALA (heme precursor). The culture was further grown at 30 °C (120 r.p.m.) for 24 h. The *E. coli* cells were harvested by centrifugation at 4500 *g* for 30 min, and cell pellets were stored at -20 °C until purification. For biochemical characterization, the protein was purified using a simplified procedure (one-step immobilized metal ion affinity chromatography). For crystallization, a three-step purification protocol was followed, as described previously (ion exchange chromatography, gel filtration, mixed-mode ion exchange chromatography) [18]. Purification of the redox partner protein pairs (BmCPR-Fdx2 and AdR-Adx₄₋₁₀₈) was performed according to the previously reported procedures [19,23].

Protein quality analysis

Spectroscopic properties of the purified P450 were analyzed using a double-beam spectrophotometer (UV-2101PC; Shimadzu Corporation, Kyoto, Japan). All UV-visible absorbance spectra were recorded from 200 to 700 nm. Concentration of CYP109A2 was estimated based on CO-difference spectroscopy (using an extinction coefficient of

91 mm⁻¹·cm⁻¹), according to the method of Omura and Sato [41]. Protein purity and heterogeneity were analyzed by SDS/PAGE and DLS, respectively. The DLS measurement was performed using a DynaPro NanoStar instrument thermostated at 293 K (Wyatt Instruments, Santa Barbara, CA, USA).

In vitro conversion and kinetic analysis

Enzymatic activity of CYP109A2 was reconstituted in an *in vitro* system containing 1 μM CYP109A2, 3 μM ferredoxin reductase (BmCPR or AdR), 20 μM ferredoxin (Fdx2 or Adx₈₋₁₀₄), 1 mM MgCl₂, and a cofactor regenerating system with glucose-6-phosphate (5 mM) and glucose-6-phosphate dehydrogenase (1 U), all in 50 mM potassium phosphate buffer containing 20% glycerol (pH 7.4) and at 250 μL final volume. VD₃ was dissolved in (2-hydroxypropyl)-β-cyclodextrin (2.25%) and added at 200 μM final concentration. The reaction was started by addition of 0.5 mM NADPH at 30 °C. After 1 h, the reaction was stopped and extracted twice with 500 μL of ethyl acetate. The organic phases were combined, evaporated to dryness, and prepared for analysis by HPLC. For determination of kinetic parameters of VD₃ conversion by CYP109A2, three independent *in vitro* reactions were performed (each continuing for 20 min). VD₃ concentrations ranged from 0 to 400 μM. The *k*_{cat} and *K*_M values were calculated by plotting the product formation rate against increasing substrate concentration. The data were fit to the Michaelis–Menten equation using ORIGIN software from OriginLab Corporation (Northampton, MA, USA).

High-performance liquid chromatography

The HPLC analysis was carried out on a Jasco system (Pu-980 HPLC pump, AS-950 sampler, UV-975 UV-visible detector, LG-980-02 gradient unit; Jasco, Gross-Umstadt, Germany) equipped with a reversed-phase Nucleodur 100-5 C₁₈ column (125 × 4 mm; Macherey-Nagel, Düren, Germany). The column temperature was adjusted to 40 °C. The samples were dissolved in 200 μL of acetonitrile. A linear gradient of 60–100% aqueous acetonitrile was applied for 15 min (1 mL·min⁻¹ flow rate), followed by 100% acetonitrile for 15 min. The UV detection of the substrate and products was accomplished at 265 nm. Assuming that the absorption properties of the products did not differ from the substrate, the product formation was calculated as the relative peak area (area %) of the HPLC chromatograms, dividing each respective product peak area by the sum of products and substrate areas.

Whole-cell conversion in *Bacillus megaterium* MS941

The whole-cell conversions were performed in *B. megaterium* MS941 cells transformed with the

pSMF2.1.CYP109A2 construct using the polyethylene glycol-mediated protoplast transformation method [42]. Initial seed culture was prepared in LB medium (10 μg·mL⁻¹ tetracycline). The main culture (50 mL TB medium, 10 μg·mL⁻¹ tetracycline) was inoculated with 500 μL of the seed culture (dilution 1 : 100) in 300 mL baffled flasks and incubated under shaking at 37 °C (140 r.p.m.). The culture was grown to OD₅₇₈ of 0.4, and at that point, the recombinant gene expression was induced with xylose (5 mg·mL⁻¹). The culture was further grown at 30 °C for 24 h (140 r.p.m.). To initialize the whole-cell conversion experiments, after 24 h of protein expression, the substrate [200 μM VD₃ dissolved in 45% (2-hydroxypropyl)-β-cyclodextrin] and a membrane-solubilizing agent (4% *Quillaja* Saponin) were added to the cultures. Afterward, the conversions were allowed to proceed at 30 °C and 120 r.p.m. for 24 h. Finally, 500 μL samples of the cultures were taken, extracted, and prepared for HPLC analysis.

Product isolation and NMR characterization

In order to obtain sufficient amounts of the reaction product, the whole-cell based cultivation and conversion experiments were scaled up to 400 mL (2 × 200 mL) and performed as described above. The main reaction product (P1) was isolated and purified by reversed-phase HPLC using a preparative column VP Nucleodur 100-5 C₁₈ (8.0 × 250 mm; Macherey-Nagel). The NMR spectra were recorded in CDCl₃ with a Bruker Avance 500 NMR spectrometer at 298 K (Bruker AXS Advanced X-ray Solutions GmbH, Karlsruhe, Germany). The chemical shifts were relative to CHCl₃ at δ 7.26 (¹H NMR) and CDCl₃ at δ 77.00 (¹³C NMR) 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.

Crystallization, data collection, and structure determination

Aliquots with purified CYP109A2 were thawed, buffer exchanged to 20 mM Tris/HCl, pH 8.0, 0.1 mM dithioerythritol and concentrated to 40 mg·mL⁻¹ by ultrafiltration using a 30-kDa cutoff membrane. Crystallization condition screening was done in 96-well plate format. Sitting-drop vapor-diffusion experiments were set up by a Mosquito crystallization robot (TTP LabTech, Melbourne, UK), using commercially available high-throughput crystallization screens. Multiple hits were obtained overnight, and the most promising condition [2.0 M ammonium sulfate, 0.1 M HEPES, pH 7.5, and 2% (v/v) PEG 400] was chosen for further optimization. Crystal optimization was performed manually at 293 K by the sitting-drop vapor diffusion method in 24-well plates using reservoir solution volumes of 300 μL. Protein drops were prepared by mixing 1 μL of

protein and 1 µL of reservoir solution. A single, red plate-like crystal of CYP109A2 appeared after 3 days using a reservoir solution containing 1.9 M ammonium sulfate, 0.1 M HEPES, pH 7.5, and 2% (v/v) PEG 400 and needed about a week to grow to its final dimensions ($0.12 \times 0.04 \times 0.03$ mm³).

For data collection, the crystal was shortly swiped through a cryobuffer droplet [mother liquor supplemented with 30% (v/v) glycerol] and flash-cooled to 100 K in a cold nitrogen gas cryostream. X-ray diffraction data were collected at the ID23-2 beam line of the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. Reflections were indexed and integrated using XDS [43], while scaling and merging of the data was done with AIMLESS from the CCP4 software suite [44]. The structure of CYP109A2 was determined by molecular replacement using PHASER from the PHENIX software package [45], applying the structure of substrate-free CYP109E1 (PDB entry 5L90, [18]) as a search model. The structure was completed with several cycles of model refinement using PHENIX.REFINE [46] alternated with manual model re-building using COOT [47]. At the final stages of refinement and model building, water molecules were added to the model based on positive difference electron density peaks and strict interaction criteria. The final model was validated with MOLPROBITY [48]. The model coordinates and structure factors are deposited in the PDB with accession code 5OFQ.

Sequence and structure analysis

Protein sequences were compared using the Basic Local Alignment Search Tool (BLAST, NCBI). Substrate recognition sites (SRS) in CYP109A2 were identified by alignment with P450_{cam} (WP_032492633) and followed a description defined by Gotoh [49]: SRS1: 71–87 (BC loop), SRS2: 163–169 (C'-terminal part of F helix), SRS3: 185–192 (N'-terminal part of G helix), SRS4: 229–247 (central part of I helix), SRS5: 284–293 (K-β5 connection), and SRS6: 384–392 (β9-β10 turn). Pairwise structural alignments were done using PDBeFold [50]. Multiple sequence alignment was done with CLUSTAL OMEGA [51], and the output figure was prepared with ESPRIPT3 [52]. The GenBank accession codes for analyzed sequences are as follows: CYP109A2 from *B. megaterium* DSM319 (WP_013082916), CYP109E1 from *B. megaterium* DSM319 (WP_013084555), CYP105A1 from *S. griseolus* (BAG50411), CYP105A2 from *P. autotrophica* (BAA05541), Vdh (CYP107BR1) from *P. autotrophica* (C4B644), CYP107CB2 from *B. lehensis* G1 (AIC83164).

Acknowledgements

This work was financially supported by the German Federation of Industrial Research Associations (AIF/ZIM project FKZ 2214512AJA) and the People

Programme (Marie Curie Actions) of the European Union's 7th Framework Programme (FP7/2007–2013) under REA Grant Agreement 289217 (P4FIFTY). Thanks to Dr Josef Zapp for measuring the NMR samples. Thanks to Dr Flora Kiss and Dr Mohammed Milhim for protein purification. Authors also thank the European Synchrotron Radiation Facility (ESRF) and their staff for beam line access and data collection assistance.

Author contributions

AA, IKJ, FH, AMWHT, and RB designed the study. AA performed all biochemical experiments and collected data. EB purified the products for NMR analysis. AA, RB, and FH analyzed and interpreted the biochemical data. IKJ crystallized the protein, collected crystallographic data, and determined the crystal structure. IKJ and AMWHT analyzed and interpreted the structural data. AA and IKJ wrote the manuscript. AMWHT, RB, and FH provided supervision and participated in writing of the manuscript. All authors read and approved the final version of the manuscript.

References

- Pludowski P, Holick MF, Grant WB, Konstantynowicz J, Mascarenhas MR, Haq A, Povoroznyuk V, Balatska N, Barbosa AP, Karonova T *et al.* (2017) Vitamin D supplementation guidelines. *J Steroid Biochem Mol Biol*, In Press. <https://doi.org/10.1016/j.jsbmb.2017.01.021>
- Wacker M & Holick MF (2013) Sunlight and vitamin D: a global perspective for health. *Dermatoendocrinol* **5**, 51–108.
- Hosseini-nezhad A & Holick MF (2013) Vitamin D for health: a global perspective. *Mayo Clin Proc* **88**, 720–755.
- Jean G, Terrat JC, Vanel T, Hurot JM, Lorriaux C, Mayor B & Chazot C (2008) Daily oral 25-hydroxycholecalciferol supplementation for vitamin D deficiency in haemodialysis patients: effects on mineral metabolism and bone markers. *Nephrol Dial Transplant* **23**, 3670–3676.
- Leichtmann GA, Bengoa JM, Bolt MJ & Sitrin MD (1991) Intestinal absorption of cholecalciferol and 25-hydroxycholecalciferol in patients with both Crohn's disease and intestinal resection. *Am J Clin Nutr* **54**, 548–552.
- Zhu G & Okamura WH (1995) Synthesis of vitamin D (calciferol). *Chem Rev* **95**, 1877–1952.
- Jones G, Prosser DE & Kaufmann M (2014) Cytochrome P450-mediated metabolism of vitamin D. *J Lipid Res* **55**, 13–31.
- Bernhardt R (2006) Cytochromes P450 as versatile biocatalysts. *J Biotechnol* **124**, 128–145.

- 9 Bernhardt R & Urlacher VB (2014) Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl Microbiol Biotechnol* **98**, 6185–6203.
- 10 Girvan HM & Munro AW (2016) Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Curr Opin Chem Biol* **31**, 136–145.
- 11 Sasaki J, Mikami A, Mizoue K & Omura S (1991) Transformation of 25- and 1 alpha-hydroxyvitamin D₃ to 1 alpha, 25-dihydroxyvitamin D₃ by using *Streptomyces* sp. strains. *Appl Environ Microbiol* **57**, 2841–2846.
- 12 Fujii Y, Kabumoto H, Nishimura K, Fujii T, Yanai S, Takeda K, Tamura N, Arisawa A & Tamura T (2009) Purification, characterization, and directed evolution study of a vitamin D₃ hydroxylase from *Pseudonocardia autotrophica*. *Biochem Biophys Res Commun* **385**, 170–175.
- 13 Abdulmughni A, Jóźwik IK, Putkaradze N, Brill E, Zapp J, Thunnissen AW, Hannemann F & Bernhardt R (2016) Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D₃ hydroxylase. *J Biotechnol* **243**, 38–47.
- 14 Hayashi K, Yasuda K, Sugimoto H, Ikushiro S, Kamakura M, Kittaka A, Horst RL, Chen TC, Ohta M, Shiro Y *et al.* (2010) Three-step hydroxylation of vitamin D₃ by a genetically engineered CYP105A1: enzymes and catalysis. *FEBS J* **277**, 3999–4009.
- 15 Hayashi K, Sugimoto H, Shinkyo R, Yamada M, Ikeda S, Ikushiro S, Kamakura M, Shiro Y & Sakaki T (2008) Structure-based design of a highly active vitamin D hydroxylase from *Streptomyces griseolus* CYP105A1. *Biochemistry* **47**, 11964–11972.
- 16 Yasutake Y, Fujii Y, Nishioka T, Cheon WK, Arisawa A & Tamura T (2010) Structural evidence for enhancement of sequential vitamin D₃ hydroxylation activities by directed evolution of cytochrome P450 vitamin D₃ hydroxylase. *J Biol Chem* **285**, 31193–31201.
- 17 Yasutake Y, Nishioka T, Imoto N & Tamura T (2013) A single mutation at the ferredoxin binding site of P450 Vdh enables efficient biocatalytic production of 25-hydroxyvitamin D(3). *ChemBioChem* **14**, 2284–2291.
- 18 Jóźwik IK, Kiss FM, Gricman Ł, Abdulmughni A, Brill E, Zapp J, Pleiss J, Bernhardt R & Thunnissen AW (2016) Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J* **283**, 4128–4148.
- 19 Brill E, Hannemann F, Zapp J, Bruning G, Jauch J & Bernhardt R (2014) A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto-beta-boswellic acid (KBA). *Appl Microbiol Biotechnol* **98**, 1701–1717.
- 20 Hamdane D, Zhang H & Hollenberg P (2008) Oxygen activation by cytochrome P450 monooxygenase. *Photosynth Res* **98**, 657–666.
- 21 Kiss FM, Schmitz D, Zapp J, Dier TK, Volmer DA & Bernhardt R (2015) Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* – identification of a novel 11-oxidase activity. *Appl Microbiol Biotechnol* **99**, 8495–8514.
- 22 Hannemann F, Bichet A, Ewen KM & Bernhardt R (2007) Cytochrome P450 systems – biological variations of electron transport chains. *Biochim Biophys Acta* **1770**, 330–344.
- 23 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.
- 24 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* **1814**, 249–256.
- 25 Sawada N, Sakaki T, Yoneda S, Kusudo T, Shinkyo R, Ohta M & Inouye K (2004) Conversion of vitamin D₃ to 1alpha,25-dihydroxyvitamin D₃ by *Streptomyces griseolus* cytochrome P450SU-1. *Biochem Biophys Res Commun* **320**, 156–164.
- 26 Bunk B, Schulz A, Stammen S, Munch R, Warren MJ, Rohde M, Jahn D & Biedendieck R (2010) A short story about a big magic bug. *Bioeng Bugs* **1**, 85–91.
- 27 Vary PS, Biedendieck R, Fuerch T, Meinhardt F, Rohde M, Deckwer WD & Jahn D (2007) *Bacillus megaterium* – from simple soil bacterium to industrial protein production host. *Appl Microbiol Biotechnol* **76**, 957–967.
- 28 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.
- 29 Putkaradze N, Kiss FM, Schmitz D, Zapp J, Hutter MC & Bernhardt R (2017) Biotransformation of prednisone and dexamethasone by cytochrome P450 based systems – identification of new potential drug candidates. *J Biotechnol* **242**, 101–110.
- 30 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.
- 31 Zhang W, Liu Y, Yan J, Cao S, Bai F, Yang Y, Huang S, Yao L, Anzai Y, Kato F *et al.* (2014) New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners. *J Am Chem Soc* **136**, 3640–3646.

- 32 Khatri Y, Schiffrin A & Bernhardt R (2017) Investigating the effect of available redox protein ratios for the conversion of a steroid by a myxobacterial CYP260A1. *FEBS Lett* **591**, 1126–1140.
- 33 Helmer B, Schnoes HK & DeLuca HF (1985) ¹H nuclear magnetic resonance studies of the conformations of vitamin D compounds in various solvents. *Arch Biochem Biophys* **241**, 608–615.
- 34 Mizhiritskii MD, Konstantinovskii LE & Vishkautsan R (1996) 2D NMR study of solution conformations and complete ¹H and ¹³C chemical shifts assignments of vitamin D metabolites and analogs. *Tetrahedron* **52**, 1239–1252.
- 35 Takeda K, Asou T, Matsuda A, Kimura K, Okamura K, Okamoto R, Sasaki J, Adachi T & Omura S (1994) Application of cyclodextrin to microbial transformation of vitamin D₃ to 25-hydroxyvitamin D₃ and 1 α ,25-dihydroxyvitamin D₃. *J Ferment Bioeng* **78**, 380–382.
- 36 Yasutake Y, Kameda T & Tamura T (2017) Structural insights into the mechanism of the drastic changes in enzymatic activity of the cytochrome P450 vitamin D 3 hydroxylase (CYP107BR1) caused by a mutation distant from the active site. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **73**, 266–275.
- 37 Lepesheva GI, Virus C & Waterman MR (2003) Conservation in the CYP51 family. Role of the B' helix/BC loop and helices F and G in enzymatic function. *Biochemistry* **42**, 9091–9101.
- 38 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**, 1135–1146.
- 39 Wittchen KD & Meinhardt F (1995) Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl Microbiol Biotechnol* **42**, 871–877.
- 40 Eppinger M, Bunk B, Johns MA, Edirisinghe JN, Kutumbaka KK, Koenig SS, Creasy HH, Rosovitz MJ, Riley DR, Daugherty S *et al.* (2011) Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J Bacteriol* **193**, 4199–4213.
- 41 Omura T & Sato R (1964) The carbon monoxide-binding pigment of liver microsomes: I. Evidence for its hemoprotein nature. *J Biol Chem* **239**, 2370–2378.
- 42 Barg H, Malten M, Jahn M & Jahn D (2005) Protein and vitamin production in *Bacillus megaterium*. In *Microbial Processes and Products* (Barredo J, ed.), pp. 205–223. Humana Press, New York, NY.
- 43 Kabsch W (2010) XDS. *Acta Crystallogr D* **66**, 125–132.
- 44 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **67**, 235–242.
- 45 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* **66**, 213–221.
- 46 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* **68**, 352–367.
- 47 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132.
- 48 Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **66**, 12–21.
- 49 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.
- 50 Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D* **60**, 2256–2268.
- 51 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.
- 52 Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–W324.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. NMR data for the main product of CYP109A2-catalyzed VD₃ conversion.

Table S2. Highest ranked structural homologs of CYP109A2 found in the PDB.

Supporting Information

Biochemical and structural characterization of CYP109A2, a vitamin D₃ 25-hydroxylase from *Bacillus megaterium*

Ammar Abdulmughni, Ilona K. Jóźwik, Elisa Brill, Frank Hannemann, Andy-Mark W.H. Thunnissen and Rita Bernhardt

Table S1. NMR data for the main product of CYP109A2-catalyzed VD₃ conversion, namely the 25-hydroxylated vitamin D₃.

25-hydroxyvitamin D₃ (Product P1)
<p>¹H NMR (CDCl₃, 500 MHz): δ 0.52 (s, 3xH-18), 0.92 (d, J=6.5 Hz, 3xH-21), 1.02 (m, H-22a), 1.06 (m, H-22b), 1.19 (s, 6H, 3xH-26 and 3xH-27), 1.20 (m, H-23a), 1.24 (m, H-16a), 1.28 (m, H-17), 1.29 (m, H-12a), 1.36 (m, H-24a), 1.39 (m, H-23b), 1.45 (m, 2H, H-11a and H-24b), 1.51 (m, H-20), 1.52 (m, 2H, H-11b and H-15a), 1.60 (m, H-2a), 1.63 (m, H-15b), 1.69 (m, H-9a), 1.86 (m, H-16b), 1.94 (m, H-14), 1.96 (m, H-2b), 1.98 (m, H-12b), 2.16 (dd, J=13.5, 8.5 and 5.0 Hz, H-1a), 2.26 (dd, J=13.0 and 7.5 Hz, H-4a), 2.38 (ddd, J=13.5, 7.5 and 4.6 Hz, H-1b), 2.55 (dd, J=13.0 and 4.0 Hz, H-4b), 2.82 (m, H-9b), 3.92 (m, H-3), 4.80 (d, J= 2.5 Hz, H-19a), 5.03 (m, H-19b), 6.01 (d, 11.3 Hz, H-7), 6.21 (d, 11.3 Hz, H-6).</p>
<p>¹³C NMR (CDCl₃, 125 MHz): δ 11.99 (CH₃, C-18), 18.81 (CH₃, C-21), 20.82 (CH₂, C-23), 22.24 (CH₂, C-11), 23.56 (CH₂, C-15), 27.67 (CH₂, C-16), 29.07 (CH₂, C-9), 29.29 (CH₃, C-26), 29.34 (CH₃, C-27), 31.92 (CH₂, C-1), 35.16 (CH₂, C-2), 36.10 (CH₂, C-22), 36.40 (CH, C-20), 40.53 (CH₂, C-12), 44.39 (CH₂, C-24), 45.85 (C, C-13), 45.92 (CH₂, C-4), 56.33 (CH, C-14), 56.53 (CH, C-17), 69.21 (CH, C-3), 71.15 (C, C-25), 112.40 (CH₂, C-19), 117.51 (CH, C-7), 122.45 (CH, C-6), 135.00 (C, C-5), 142.89 (C, C-8), 145.09 (C, C-10).</p>

Table S2. Highest ranked structural homologues of CYP109A2 found in the PDB. The search was performed with PDBeFold [49].

CYP109A2 (open, substrate-free, chain C)						
	PDB code and chain id	Q score	RMSD C α [Å]	N _{align}	% Sequence	P450 and organism
1	5L91:B	0.77	1.27	368	45	CYP109E1 <i>B. megaterium</i> DSM319
2	4RM4:A	0.67	1.44	339	47	CYP109B1 <i>B. subtilis</i>
3	5GNM:B	0.66	1.42	354	34	Vdh (CYP107BR1)-L348M <i>P. autotrophica</i>
4	5GNL:A	0.66	1.43	351	35	Vdh (CYP107BR1)-F106V <i>P. autotrophica</i>
5	4YZR:A	0.65	1.58	354	37	PksS <i>B. subtilis</i>
6	3A4G:A	0.65	1.44	349	35	Vdh (CYP107BR1) <i>P. autotrophica</i>
7	2BVJ:B	0.64	1.36	340	37	PikC (CYP107L1) <i>S. venezuelae</i>
8	2WIO:A	0.63	1.41	342	35	EryK (CYP113A1) <i>S. erythraea</i>
9	3EJE:D	0.62	1.85	357	35	BioI (CYP107H1) <i>B. subtilis</i>
10	5IKI:A	0.62	1.81	352	38	CYP106A2 <i>B. megaterium</i>

2.3 (Jóźwik et al., 2016)

Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*.

Ilona K. Jóźwik, Flora M. Kiss, Łukasz Gricman, **Ammar Abdulmughni** , Elisa Brill, Josef Zapp, Jürgen Pleiss, Rita Bernhardt and Andy-Mark W. H. Thunnissen.

FEBS Journal, 2016 Nov; 283: 4128–4148.

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

Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*

Ilona K. Jóźwik¹, Flora M. Kiss², Łukasz Gricman³, Ammar Abdulmughni², Elisa Brill², Josef Zapp⁴, Juergen Pleiss³, Rita Bernhardt² and Andy-Mark W. H. Thunnissen¹

¹ Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands

² Institute of Biochemistry, Saarland University, Saarbrücken, Germany

³ Institute of Technical Biochemistry, University of Stuttgart, Germany

⁴ Pharmaceutical Biology, Saarland University, Saarbrücken, Germany

Keywords

crystallography; cytochrome P450; steroid; structure-function; testosterone

Correspondence

R. Bernhardt, Institute of Biochemistry, Campus B2.2, 66123, Saarland University, Saarbrücken, Germany

Fax: +49 (0) 681 302 4739

Tel: +49 (0) 681 302 4241

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

and

A.-M. W. H. Thunnissen, University of Groningen, Laboratory of Biophysical Chemistry, Nijenborgh 7, 9747 AG Groningen, The Netherlands

Fax: +31 50 3634800

Tel: +31 50 3634380

E-mail: a.m.w.h.thunnissen@rug.nl

(Received 15 June 2016, revised 15 September 2016, accepted 27 September 2016)

doi:10.1111/febs.13911

Cytochrome P450 monooxygenases (P450s) are attractive enzymes for the pharmaceutical industry, in particular, for applications in steroidal drug synthesis. Here, we report a comprehensive functional and structural characterization of CYP109E1, a novel steroid-converting cytochrome P450 enzyme identified from the genome of *Bacillus megaterium* DSM319. *In vitro* and whole-cell *in vivo* turnover experiments, combined with binding assays, revealed that CYP109E1 is able to hydroxylate testosterone at position 16 β . Related steroids with bulky substituents at carbon C17, like corticosterone, bind to the enzyme without being converted. High-resolution X-ray structures were solved of a steroid-free form of CYP109E1 and of complexes with testosterone and corticosterone. The structural analysis revealed a highly dynamic active site at the distal side of the heme, which is wide open in the absence of steroids, can bind four ordered corticosterone molecules simultaneously, and undergoes substantial narrowing upon binding of single steroid molecules. In the crystal structures, the single bound steroids adopt unproductive binding modes coordinating the heme-iron with their C3-keto oxygen. Molecular dynamics (MD) simulations suggest that the steroids may also bind in $\sim 180^\circ$ reversed orientations with the C16 carbon and C17-substituents pointing toward the heme, leading to productive binding of testosterone explaining the observed regio- and stereoselectivity. The X-ray structures and MD simulations further identify several residues with important roles in steroid binding and conversion, which could be confirmed by site-directed mutagenesis. Taken together, our results provide unique insights into the CYP109E1 activity, substrate specificity, and regio/stereoselectivity.

Database

The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes [5L90](#) (steroid-free CYP109E1), [5L91](#) (CYP109E1-COR4), [5L94](#) (CYP109E1-TES), and [5L92](#) (CYP109E1-COR).

Enzymes

Cytochrome P450 monooxygenase CYP109E1, [EC 1.14.14.1](#), UniProt ID: [D5DKI8](#), Adrenodoxin reductase [EC 1.18.1.6](#).

Abbreviations

COR, corticosterone; P450, cytochrome P450 monooxygenase; SRS, substrate recognition site; TES, testosterone.

Introduction

Cytochrome P450 monooxygenases (P450s) constitute a large superfamily of enzymes, found in all domains of life, which catalyze oxygen-mediated hydroxylation of a wide variety of aromatic and aliphatic compounds. In nature, these enzymes play essential roles in metabolic processes like steroid biosynthesis, fatty acid metabolism, or biotransformation of drugs and other xenobiotics. In the laboratory and biotechnological industry, they are considered as high potential biocatalysts, owing to their ability to selectively oxidize unreactive C-H bonds at mild conditions [1]. Their application toward cost-effective and environmentally friendly production of steroid derivatives is of particular interest, considering the wide use of such compounds as therapeutic agents [2,3]. Currently more than 300 steroid drugs are authorized, making them the most marketed group of products in the pharmaceutical industry [4]. The ability of P450s to perform regioselective and stereoselective hydroxylation of steroids has led to an ongoing search and characterization of new enzymes, in particular from prokaryotic sources, as these are more amenable to industrial application than their eukaryotic counterparts [5].

To efficiently handle P450s and to design improved variants for synthetic applications, it is crucial to understand the structure-function relationships governing their substrate specificity and regio- and stereoselectivity [6,7]. Although a wealth of functional and structural data are available for these enzymes, it has proven difficult to pinpoint the molecular determinants of their different specificities. Crystal structures of P450 enzymes from mammalian and bacterial sources have revealed a high conformational variability, despite their common overall fold [8–10]. In particular, a few characteristic, highly flexible regions around the P450 distal heme pocket (the BC loop, F and G helices together with the FG loop) are associated with substrate access, substrate binding and product release. Understanding substrate binding and conversion by P450s requires the analysis of crystallographic ‘snapshots’ of their open (substrate-free) and closed (substrate-bound) states, which frequently are not available for individual P450s. Combining findings from X-ray crystallography, computational docking, and MD simulations has significantly increased the understanding of mammalian and bacterial P450s [11,12]. Nevertheless, for new, soluble, prokaryotic P450s, structural characterization is indispensable for adapting them to specific biotechnological applications.

To date, only a few bacterial steroid-specific P450s have been functionally and structurally characterized. Recently published examples include CYP106A2 from *Bacillus megaterium* ATCC 13368 [13], CYP154C5 from *Nocardia farcinica* [14], CYP154C3 from *Streptomyces griseus* [15], and CYP109B1 from *Bacillus subtilis* [16]. Of these, CYP154C5 is of particular interest, as it exhibits an exceptionally high regio- and stereoselectivity to various pregnanes and androstanes, including pregnenolone, progesterone, testosterone, and androstenedione, yielding only 16 α -hydroxylated steroids. Crystal structures of steroid-bound CYP154C5 reveal a narrow, nearly closed active site, mostly hydrophobic with two opposing polar regions, perfectly matching the size and polarity distribution of the steroids. The apolar/polar shape complementarity leads to a highly ordered binding mode of the steroids, with the α -face of their C16 carbon at a suitable distance from the heme-iron for allowing a radical attack by the reactive iron-oxo species (compound I) of the catalytic cycle, thus explaining the high regio- and stereoselectivity of the enzyme. Of the other two enzymes, CYP109B1 is of interest as it represents the first member of the CYP109 family of which a crystal structure has been determined. It was crystallized in an ‘open’ conformation with no ligands bound in the active site. Sequence alignment with different CYP109 family members, and members of related P450 families, suggested that variations in the so called BC loop, one of the common P450 regions involved in substrate binding, may primarily account for the diverse substrate specificities among these enzymes.

Recently, the genomes of *Bacillus megaterium* DSM319 [17] and ATCC 13368 were sequenced, allowing exploration of their cytochrome P450 complement. Its analysis has resulted in the identification and characterization of several novel steroid-converting P450s [18–21]. Here, we report the functional and structural properties of CYP109E1 from *B. megaterium* DSM319. Using *in vitro* and *in vivo* turnover experiments, we demonstrate that the enzyme converts testosterone to 16 β -hydroxytestosterone with a high stereo- and regioselectivity. Crystal structures of CYP109E1 were determined both in steroid-free and steroid-bound states allowing a detailed analysis of the interactions and conformational changes associated with steroid binding. In addition, we employed MD simulations to test putative productive steroid-binding modes in the active site pocket of CYP109E1, leading to a better understanding of the structural determinants of the enzyme’s activity.

Results

Identification and bioinformatic analysis of CYP109E1

The *cyp109e1* gene was identified in the genome of *B. megaterium* DSM319 by the same bioinformatic search strategy used previously for identifying *cyp106a1* [20]. A multiple amino acid sequence alignment of CYP109E1 (UniProtKB entry [d5dki8](#)) with its closest homologs is presented in Fig. 1. Classification of the enzyme into the CYP109 family was based on the conventional P450 nomenclature system [22]. Indeed, CYP109E1 shares close similarity with other characterized CYP109 family members, that is, CYP109B1 from *B. subtilis* strain 168 (47% sequence identity), CYP109A1 from *B. subtilis* strain W23 (43% sequence identity), and the three fatty acid-oxidizing proteins CYP109C1, CYP109C2, and CYP109D1 from *Sorangium cellulosum* strain So ce56 (39%, 41%, and 33% sequence identity, respectively) [16,23,24]. However, the phylogenetic tree (Fig. 2) indicates that CYP109E1 shares the closest similarity with CYP106A1 from the same organism (42% amino acid sequence identity). This close relationship, which has been noted also for other CYP109 and CYP106 family members from *Bacillus* species [25], may indicate that these enzymes convert similar substrates.

Expression, purification, and spectroscopic characterization of CYP109E1

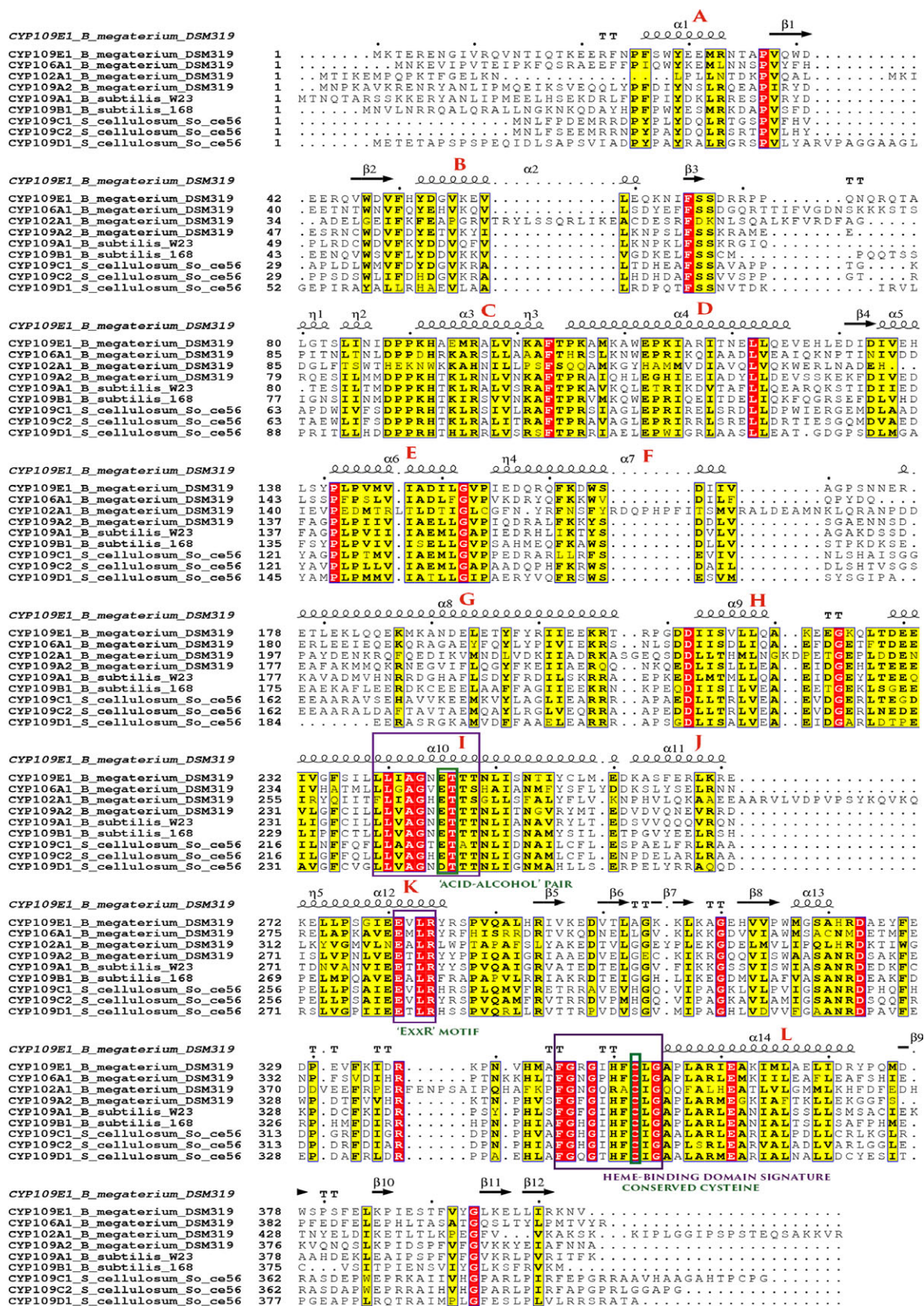
CYP109E1 was expressed in *E. coli* as a soluble protein, containing the gene-encoded residues 1–404, fused to a C-terminal polyhistidine tag. Purification was accomplished by employing a three-step purification strategy, including anion-exchange, size-exclusion, and mixed mode ion-exchange chromatography. Immobilized metal ion affinity chromatography was on purpose excluded from the purification protocol to eliminate the risk of imidazole affecting the functional assays and crystallization experiments. Typically, 25 mg of pure protein was obtained from 0.5 L of expression culture, with an estimated purity of > 95% as judged by SDS/PAGE. Analysis of UV-visible absorption and reduced CO-difference

spectra of reduced CYP109E1 revealed all characteristic P450 peaks, confirming the structural integrity of the protein with correct incorporation of the heme cofactor. The purified protein thus obtained was used for subsequent functional and structural characterization.

Substrate screening, conversion, and product identification

Considering the close phylogenetic distance between CYP109E1 and CYP106A1, it was proposed that steroids being substrates of CYP106A1 could also be suitable ligands for CYP109E1. The potential of CYP109E1 as a steroid hydroxylase was, therefore, tested in ligand-binding experiments using a library of 13 steroids (Table 1, see Table S2 for chemical structures). Six steroids induced a type I spectral shift upon titration to CYP109E1: androstenedione (1), corticosterone (4), deoxycorticosterone (5), dexamethasone (7), testosterone (11), and testosterone acetate (12), indicating displacement of the axial heme water and thus marking them as potential CYP109E1 substrates. However, several studies have shown that, on one hand, not all compounds inducing type I shift are substrates of P450s [19,26,27] while, on the other hand, substrates do not always induce a type I spectral shift [18]. Consequently, all 13 steroids were further subjected to an *in vitro* CYP109E1-dependent enzymatic conversion assay, using bovine adrenodoxin and adrenodoxin reductase (Adx_{4–108} and AdR) as surrogate redox partners for the reconstitution of P450 activity. Most steroids were not converted by CYP109E1 (including all steroids which did not induce a type I spectral shift) or showed only very low to negligible conversion rates (< 10%). Only testosterone (11) showed substantial turnover by CYP109E1, and analysis of the reaction mixture by HPLC revealed one main (69%) and one minor (14%) product (Fig. 3), with a catalytic activity of 4.35 and 0.93 nmol product·(nmol P450)^{−1}·min^{−1}, respectively. Structure determination of these products by NMR spectroscopy required a substantial increase in their amounts, which could not easily be accomplished with the *in vitro* activity assay. This was the incentive for developing a *B. megaterium* whole-cell system overexpressing

Fig. 1. Multiple sequence alignment of CYP109E1 with P450s identified in *B. megaterium* DSM319 and other known CYP109 family members. Secondary structural elements are shown as in the substrate-free CYP109E1 crystal structure (helices labeled A–L). Conserved and similar residues are highlighted in red and yellow, respectively. Highly conserved, functionally relevant regions (central part of I helix, EXXR motif and heme-binding domain signature) are shown with violet frames. For all sequences shown, the UniProtKB accession numbers are the same as used in the phylogenetic tree (Figure 2) (for CYP102A1 only residues 1–472 of the heme domain are shown).



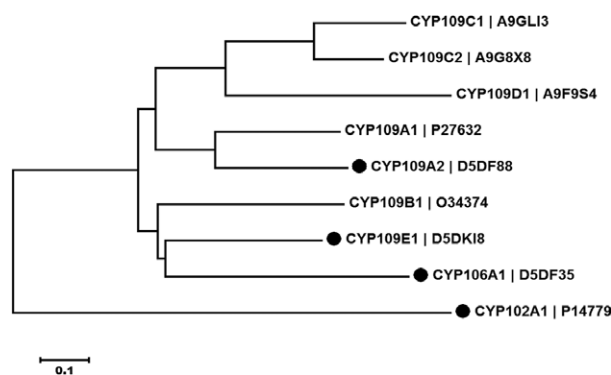


Fig. 2. Phylogenetic tree of the CYP109 family members and related P450s from *B. megaterium* DSM319. P450s from *B. megaterium* DSM319 are indicated with a closed circle (•). CYP109C1, CYP109C2, and CYP109D1 are from *Sorangium cellulosum* So ce56; CYP109A1 and CYP109B1 are from *B. subtilis* W23 and 168, respectively. The UniProtKB accession numbers are given next to the associated CYP names. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

CYP109E1, thus allowing the conversion of substrates under more optimal *in vivo* conditions and without the need of surrogate P450 redox partners. Testosterone turnover by CYP109E1 was successfully reproduced in the *B. megaterium* whole-cell system, enabling conversion of 26 mg of testosterone (300 μ M) within 6 h to 19 mg and 5 mg of the main and minor product, respectively. Structure elucidation by NMR identified the main product as 16 β -hydroxytestosterone and the side product as androstenedione (Table S3). The formation of androstenedione as a side product may be the result of a weak hydroxylation activity toward the C17 atom of testosterone, yielding a 17,17-gem-diol intermediate that subsequently dehydrates to the corresponding oxocarbon.

To investigate whether the differences in the conversion of compounds inducing a type I spectral shift are due to major differences in binding to CYP109E1, we compared the affinities of testosterone (converted substrate) and corticosterone (bound, but not converted). Quantitative analysis of the spectral shift titration data revealed that both compounds displayed very similar binding affinities to CYP109E1 with a dissociation constant (K_D) of 105 ± 10 μ M and 91 ± 8 μ M, respectively (Fig. 4). Thus, the highly selective activity of CYP109E1 toward testosterone, versus no conversion of corticosterone, is not correlated with a difference in steroid-binding affinity. In summary, the functional experiments revealed that CYP109E1 displays a rather narrow steroid-binding specificity, with relatively weak-binding affinities, and is able to hydroxylate

Table 1. Steroid binding and *in vitro* conversion activity of CYP109E1 toward selected steroids. Chemical structures of the steroids are shown in supplementary Table S2. The + and – indicate positive and negative outcomes of what is stated in the column headers.

Compound	Type I shift	<i>In vitro</i> conversion
Androstenedione (1)	+	+ (< 10%)
Cortisol (2)	–	–
Cortisone (3)	–	–
Corticosterone (4)	+	– (< 2%)
Deoxycorticosterone (5)	+	+ (< 7%)
11-deoxycortisol (6)	–	–
Dexamethasone (7)	+	–
Prednisolone (8)	–	–
Prednisone (9)	–	–
Progesterone (10)	–	–
Testosterone (11)	+	+
Testosterone acetate (12)	+	+ (< 10%)
19-nortestosterone (13)	–	–

testosterone to 16 β -hydroxytestosterone with very high regio- and stereoselectivity.

Overall structure of steroid-free CYP109E1

To evaluate the structural basis for the selectivity of steroid binding and conversion, crystallization of CYP109E1 in the presence and absence of a substrate was performed. Purified CYP109E1 was successfully crystallized in the absence of steroids, and its crystal structure was solved and refined to a 2.55 Å resolution (see Table 2 for the relevant data collection and refinement statistics). The CYP109E1 crystals contain two protein molecules in the asymmetric unit. Both polypeptide chains, and their associated heme cofactors, are well defined in the electron density, except for residues 1–20 and the C-terminal His₆-tags. Since the two crystallographically independent CYP109E1 molecules have nearly identical conformations [the root mean square deviation (RMSD) of C α -backbone atoms is 0.34 Å for 374 aligned residues], we will restrict the description to only one of them. CYP109E1 adopts the characteristic triangular ‘P450-fold’ and contains 13 α -helices, two $_3_{10}$ helices, and 10 β -strands (arranged in a five-stranded sheet, three-stranded sheet, and a β -finger, Fig. 5A). The overall fold of the P450 enzyme is well conserved: the closest structural homolog of CYP109E1 is substrate-free CYP109B1 from *B. subtilis* (PDB entry 4rm4, RMSD of C α -backbone atoms of 1.19 Å for 347 common residues). A unique feature of the CYP109E1 structure is presented by its G helix (following nomenclature of secondary

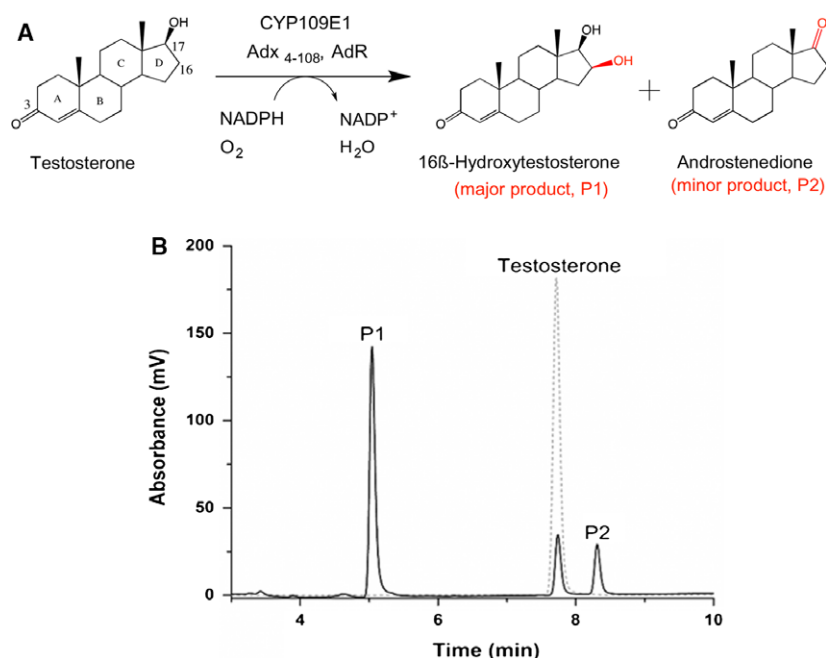


Fig. 3. *In vitro* conversion of testosterone by purified CYP109E1. (A) Schematic representation of the reaction and (B) HPLC chromatogram.

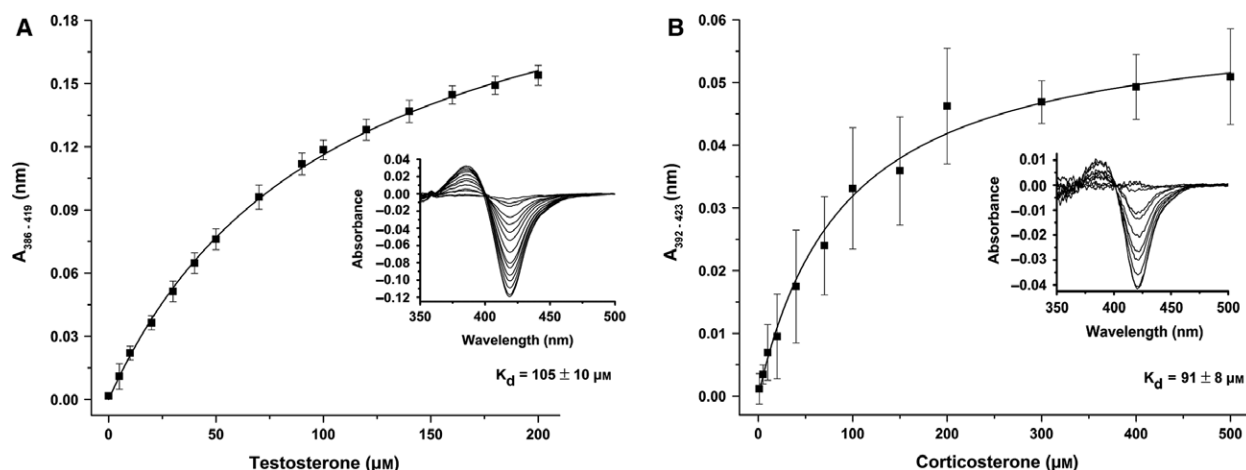


Fig. 4. Type I spectral shifts induced by steroid binding to CYP109E1 and the derived binding curves. (A) Binding curve of testosterone. (B) Binding curve of corticosterone. The spectral shifts upon titration of the steroids are shown as insets. The enzyme solution (10 μM , 50 mM potassium phosphate buffer, pH 7.4) was titrated with increasing amounts of substrates dissolved in DMSO. The peak-to-through absorbance differences were plotted against the increasing concentrations of the substrate. Plotted data points are the mean values from three independent measurements. Error bars represent the standard deviations. The curves were fitted by hyperbolic regression.

structural elements established by Poulos *et al.* [28]), which is significantly longer than in other bacterial P450 structures. A structural comparison with other P450s further reveals that steroid-free CYP109E1 adopts an open conformation with a spacious active site pocket at the distal side of the heme, which is freely accessible to solvent (Fig. 5B). The heme cofactor is bound between the I and L helices, with the thiolate group of Cys352 serving as the fifth ligand of the heme-iron. No electron density was

observed for a water molecule coordinating the heme-iron as a sixth ligand. Instead, a small blob of electron density was visible at about 4 Å distance from the heme, but the density was of insufficient quality to allow identification of the bound ligand. The noncovalent interactions between CYP109E1 and its heme are highly similar as in other class I P450s. Stabilization is provided by van der Waals interactions between the heme core and several apolar protein side chains, and by ion pairs formed between

Table 2. Data collection, refinement, and model statistics of CYP109E1.

	Substrate-free CYP109E1	CYP109E1-COR4	CYP109E1-TES	CYP109E1-COR
Conformation	Open	Open	Closed & open	Closed & open
PDB code	5L90	5L91	5L94	5L92
Model statistics				
Monomers in the AU	2	2	2	2
Solvent content (%)	61.7	60.2	49.1	49.1
Ligands	n/a	Four corticosterone molecules/monomer	Testosterone (TES) (in chain A)	Corticosterone (COR), malonic acid (MLA) (in chain A)
Data collection				
X-ray source (ESRF/in-house)	ID29	ID23-2	In-house	ID23-1
Wavelength (Å)	1.04541	0.87261	1.5418	0.97241
Resolution range (Å)	60–2.55 (2.65–2.55) ^a	49–2.20 (2.26–2.20)	56–2.25 (2.32–2.25)	48–2.10 (2.16–2.10)
Space group	P 3 ₂ 2 1	P 3 ₂ 2 1	P 1 2 ₁ 1	P 1 2 ₁ 1
Unit-cell parameters				
a, b, c (Å)	121.3, 121.3, 144.2	120.4, 120.4, 140.8	60.4, 134.9, 61.9	60.5, 135.7, 61.6
α, β, γ, °	90, 90, 120	90, 90, 120	90, 113.8, 90	90, 114.4, 90
Observed reflections	293 843 (30 950)	355 244 (26 012)	293 196 (26 695)	156 293 (12 345)
Unique reflections	40 454 (4508)	60 153 (4384)	42 810 (3945)	51 292 (12 345)
Multiplicity	7.3 (6.9)	5.9 (5.9)	6.8 (6.8)	3.0 (2.9)
Completeness (%)	99.9 (99.9)	99.7 (100)	99.7 (99.1)	97.4 (98.8)
<I/σ(I)>	11.6 (2.6)	15.9 (2.3)	12.2 (2.7)	23.3 (4.1)
R _{merge} (%)	9.6 (80.2)	8.7 (66.9)	14.2 (73.9)	2.8 (22.2)
CC (1/2) (%)	99.8 (84.8)	99.8 (80.1)	99.6 (83.7)	99.9 (91.4)
Refinement				
R _{work} (%)	21.44	18.48	21.90	20.09
R _{free} (%)	25.88	23.21	26.28	24.31
R.m.s.deviation, bond lengths (Å)	0.017	0.014	0.014	0.014
R.m.s.deviation, bond angles (°)	1.679	1.322	1.252	1.299
Average B-factors (Å) ²				
Overall	73.65	46.81	40.16	42.65
Protein	73.96	46.49	40.45	42.92
Heme	48.29	28.30	27.39	28.22
Testosterone (TES)	n/a	n/a	35.29	n/a
Corticosterone (COR)	n/a	n/a	n/a	43.43
COR-1	n/a	48.66	n/a	n/a
COR-2	n/a	77.89	n/a	n/a
COR-3	n/a	86.81	n/a	n/a
COR-4	n/a	56.60	n/a	n/a
Malonic acid (MLA)	n/a	n/a	n/a	42.14
Ramachandran plot statistics				
Most favored (%)	94.9	95.6	96.1	96.6
Allowed regions (%)	4.8	4.3	3.9	3.4
Disallowed regions (%)	0.3	0.1	0.00	0.00
Molprobrity overall score	1.62	1.54	1.22	1.14

^a Values in parentheses refer to data in the highest resolution shells.

the heme propionates and the side chains of highly conserved residues His92, Arg96, Arg294, and His350.

Active site pocket of steroid-free CYP109E1

The geometry of the active site pocket in the open form of CYP109E1 was analyzed in more detail. As

shown in Fig. 5B, it resembles a funnel, wide open at the entrance but becoming more constricted as it leads toward the heme. The structural regions surrounding the active site pocket are equivalent to the six substrate recognition sites commonly found in P450s [29]: the BC loop (substrate recognition site 1 or SRS1), parts of the F and G helices (SRS2 and SRS3, respectively), the central segment of the I helix (SRS4), the

region connecting helix K and strand $\beta 5$ (SRS5) and the $\beta 9$ – $\beta 10$ turn (SRS6). Close to the heme, the active site is very hydrophobic and its surface is roughly built of two rings of residues. The first ring (closest to the heme) contains residues Ile85 (BC loop); Leu238, Ala242, and Thr246 (I helix); Pro288, Val289, and Leu292 (K– $\beta 5$ loop). Many of these residues are highly to moderately conserved among P450s, for example, Thr246, Ala242, or Val289, in agreement with common roles in catalysis and substrate binding [30]. The second ring contains residues Arg69, Leu80, and Asn86 (BC loop); Ile168 and Val169 (F helix); Ile241 and Glu245 (I helix); Phe391 and Val392 ($\beta 9$ – $\beta 10$ turn). Residues Arg69, Asn86, and Glu245 form two polar-charged surface patches located at opposite sides of the active site pocket. Further away from the heme, near the entrance of the active site pocket, residues are mostly polar charged. Active site residues in the second ring, and residues at the entrance of the active site pocket in CYP109E1, show substantial divergence with respect to other P450s, including the testosterone-hydroxylating enzymes like CYP109B1 from *B. subtilis* or CYP154C5 from *N. farcinica*.

Structures of CYP109E1 with single bound corticosterone or testosterone

To characterize how steroids interact with CYP109E1, crystal structures of CYP109E1 with bound corticosterone (type I shift, but no substrate) or testosterone

(type I shift, substrate) (CYP109E1-COR and CYP109E1-TES) were obtained by cocrystallization and refined at 2.25–2.1 Å resolution (see Table 2 for details). Like the steroid-free crystals, the steroid-bound cocrystals contain two polypeptide chains per asymmetric unit (solvent content of 49%), yet they belong to a different space group and show a different crystal packing geometry. Only one of the two unique CYP109E1 protein molecules in each cocrystal has a bound steroid. The protein molecules lacking a bound steroid adopt an open conformation similar to that of steroid-free CYP109E1. No electron density is present for the BC loop, FG loop, and several residues at the end of the G helix, indicating that these regions are significantly disordered. In the protein molecules with a bound steroid, a remarkable narrowing and partial closure of the active site pocket is observed, allowing most substrate-binding regions to closely approach and interact with the steroids. The F and G helices, together with the FG loop, show the largest displacement, while the BC loop and $\beta 9$ – $\beta 10$ turn show smaller readjustments (Fig. 6A). The central part of the BC loop (residues 71–75) is highly disordered and could not be resolved in the electron density map. The conformational change in the FG helices is accompanied by reorientations of the H helix and HI loop. In addition, a local widening of the I helix is observed at residues 242–246, creating a groove which allows close packing of the steroids near the heme. A similar groove in the I helix has also been observed in other

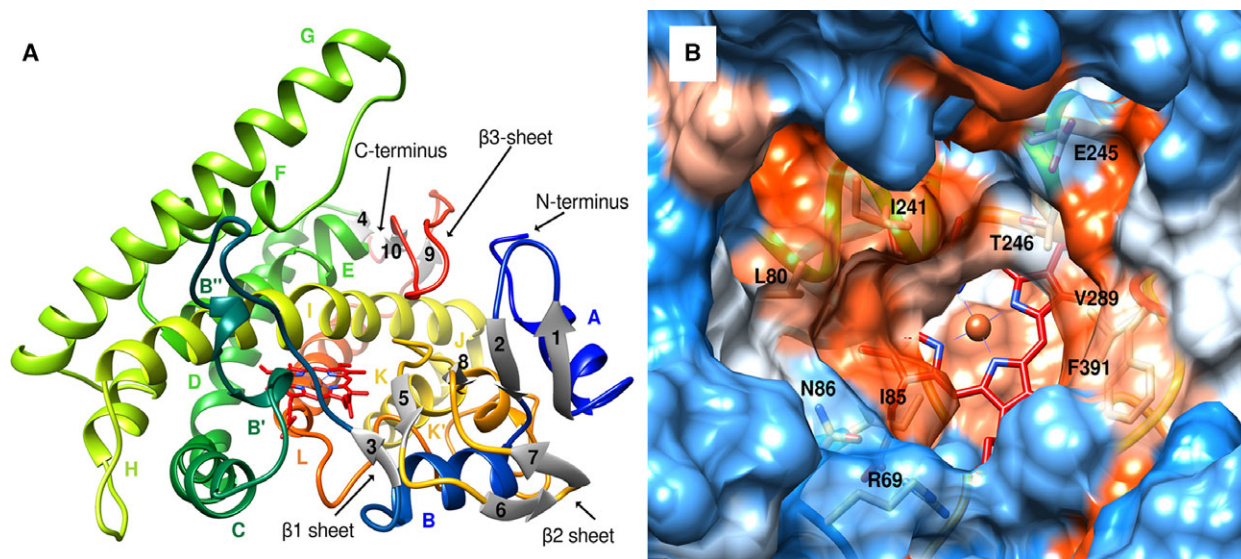


Fig. 5. Overall structure and the active site of CYP109E1. (A) Overall structure shown in ribbon representation, with rainbow coloring from blue N' terminus to red C' terminus. The heme cofactor is shown in red and beta strands in gray. Secondary structural elements are labeled following the common P450 nomenclature. (B) Active site pocket of substrate-free CYP109E1. Coloring apolar–polar, orange–blue.

substrate-bound bacterial P450 structures, and its substrate-induced conformation is believed to play an important role in oxygen activation and proton delivery [31]. Interestingly, four ordered water molecules are found at the interface of the I and E helix, two of which are located in the groove of the I helix (Fig. 7). The four waters form a continuous hydrogen-bonded network, which includes the main-chain carbonyls of Ala242 and Gly243, the side-chain hydroxyl groups of Thr246 and Thr247, and the main-chain amide of Thr248. Although the network is not directly connected to bulk solvent (the distance with the nearest water bound at the protein surface is 5.3 Å), it correlates well with the so called 'solvent channel' found in many P450s, located between the F, E, and I helices, which is believed to serve as a water access channel and/or proton delivery network [32].

Electron density of the bound steroids was of sufficient quality to allow a clear characterization of their binding modes (Fig. 6B, C). Both steroids show an unproductive binding mode: the corticosterone and testosterone molecules are bound in roughly perpendicular orientations relative to the heme plane, with their C3-keto oxygen atoms coordinating the heme-iron at the sixth axial position. The β -faces of the steroids are oriented toward the I helix, while the C17 substituents are pointing away from the heme toward the entrance of the active site pocket. The two steroids are mainly bound by van der Waals and hydrophobic interactions. Residues in the active site pocket making hydrophobic interactions with the steroids are Leu80 (BC loop, SRS1), Ile168, Val169 (F helix, SRS2), Leu238, Ile241, Ala242, Thr246 (I helix, SRS4), Val289, Ala291, His293 (K/ β 5 connecting region, SRS5), and Phe391, Val392 (β 9– β 10 turn, SRS6). No direct hydrogen bonds are observed between the protein and the two steroid molecules. In the structure with bound corticosterone, a malonic acid molecule is present in the active site forming van der Waals contacts with ring B of the steroid (Fig. 6C). The malonic acid molecule is further stabilized by hydrogen bonds with His293 (SRS5) and one of the heme-propionate groups, and by van der Waals contacts with residues Arg69, Ile85 (SRS1), Leu292 (SRS5), and Phe391 (SRS6). Its presence is most likely a crystallographic artifact, as malonic acid is the major component of tacsimate, the reagent used in the crystallization of CYP109E1-COR.

Multiple corticosterone-bound CYP109E1 structure

An additional structure of corticosterone-bound CYP109E1 was obtained at 2.2 Å resolution by soaking

the steroid into pregrown, steroid-free crystals (CYP109E1-COR4, Table 2). Interestingly, the corticosterone-soaked crystals display a water as the sixth axial ligand of the heme-iron and multiple steroid binding (four corticosterones, Fig. 8). For clarity, we abbreviate the ligands as COR-1, COR-2, COR-3, and COR-4. COR-1 is located closest to the heme-iron, while COR-4 is most distant. The protein molecules in CYP109E1-COR4 adopt an open state, with only a minor repositioning of the G helix toward the bound steroids. The limited conformational changes compared to steroid-free CYP109E1 are not surprising, since crystal packing interactions between the neighboring protein molecules lock the F and G helices, FG loop, and BC loop in an 'open' position (Fig. 8A). It is remarkable, however, that the open state allows ordered binding of multiple steroids.

The four steroids are bound in different orientations and show substantial intermolecular van der Waals contacts between their steroid rings. COR-1 is pointing with its bulky C17-substituent toward the heme and its C21-hydroxyl group forms a hydrogen bond with a water molecule coordinating the heme-iron (Fig. 8B). Additional interactions of this steroid at the active site pocket include two direct hydrogen bonds, between the C3-keto oxygen and the side chain of Lys187 (SRS3) and between the C11-hydroxyl group and the main-chain carbonyl oxygen of Ile241 (SRS4), two water-mediated hydrogen bonds, between the C11-hydroxyl group and the side chain of Glu245 and between the C21-hydroxyl group and the main-chain carbonyl group of Ala242 (SRS4), as well as several hydrophobic contacts with residues Ile85 (SRS1), Ile168, Val169 (SRS2), Thr246 (SRS4), Val289 (SRS5), and Val392 (SRS6). The other three corticosterones also make specific contacts with the protein (interactions described in detail in Table S4).

Modeling of regio- and stereoselectivity of CYP109E1

In the CYP109E1-TES crystal structure, the shape of the electron density for the bound steroid suggested that testosterone may perhaps adopt an alternative, reversed binding mode, in which the C16 and C17 carbons are located close to the Fe atom (~ 4 Å), consistent with the observed activity of CYP109E1. Addition of such a putative productive binding mode in the refined CYP109E1-TES structure resulted in a poor fit to the observed electron density, thus, we assume that testosterone predominantly adopts the unproductive binding mode in the protein crystals. Therefore, to explore the molecular basis of its selective steroid conversion and regio/stereoselectivity, MD simulations of

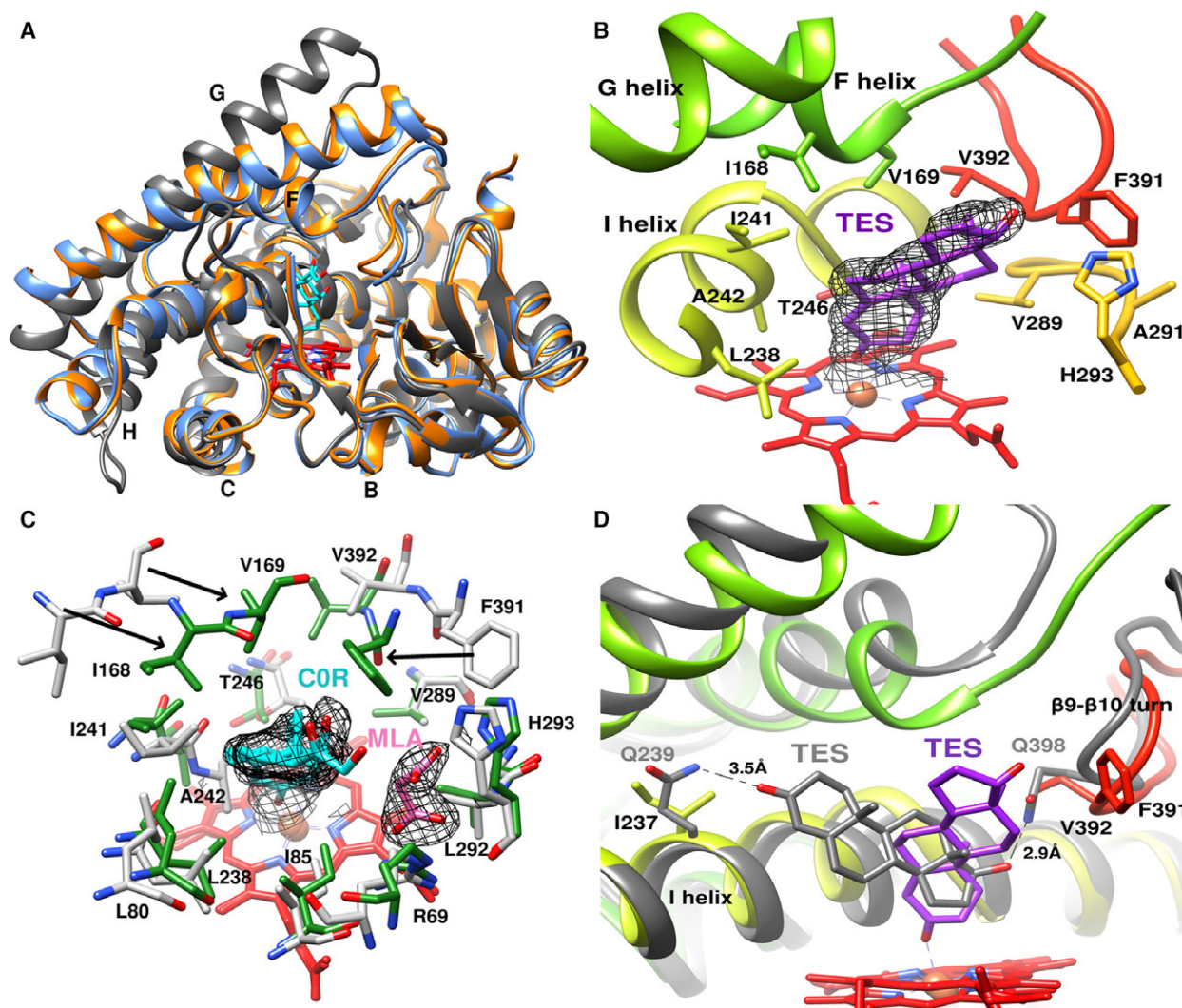


Fig. 6. Conformational changes and steroid-binding modes in CYP109E1. (A) Superposition of substrate-free CYP109E1 (gray), CYP109E1-TES (light blue), and CYP109E1-COR (orange) showing the open–closed conformational changes. COR is shown as cyan stick model. (B) Amino acid residues involved in testosterone binding (CYP109E1-TES) are shown and TES is shown as violet stick model. Black mesh is the composite omit $2F_o - F_c$ electron density map, calculated at 2.25 Å resolution and contoured at 1σ . (C) Residues binding single corticosterone and malonic acid in the active site of CYP109E1-COR (green) in comparison to substrate-free structure (gray), with COR shown in cyan and MLA in light pink. Black mesh is the composite omit $2F_o - F_c$ electron density map, calculated at 2.1 Å resolution and contoured at 1σ . (D) Comparison of testosterone binding modes observed in CYP109E1 (violet stick model) and CYP154C5 from *N. farcinica* (in gray, PDB entry 4j6d). Amino acid residues providing stabilizing interactions in CYP154C5 are shown and compared to their structural homologues in CYP109E1. Distances are given in ångströms next to the black dashed lines.

the CYP109E1-oxoferryl species (compound I) in complex with corticosterone or testosterone were performed. Starting models for the MD simulations were based on the CYP109E1-TES structure in which the crystallographically observed testosterone molecule was replaced by docked steroids (corticosterone or testosterone) in reversed orientations, with the C16 carbon atom or C17 substituents oriented toward the heme-iron. Steroid conversion and regio/stereoselectivity of the CYP109E/steroid complexes were modeled

by assigning each frame of the MD simulation as being in either a near-attack conformation or in a non-productive conformation, on the basis of distance and angle cutoffs, following previously published procedures [33–35]. The number of near-attack conformations was the highest for MD simulations with testosterone, where 16% of the simulation frames revealed near-attack conformations suggesting formation of 16 β -hydroxytestosterone (pro-16 β , 99%) or androstenedione (pro-17 α , 1%). No pro-16 α near-

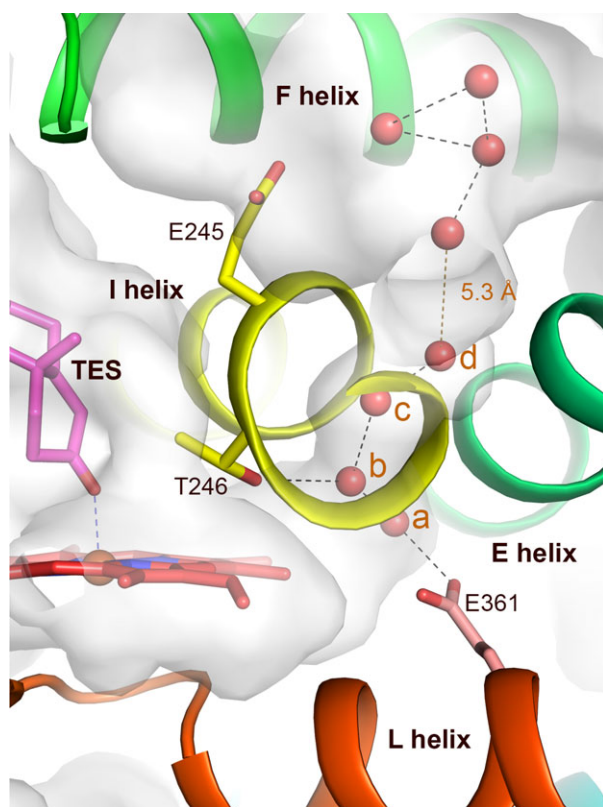


Fig. 7. Putative proton relay network in CYP109E1. Local distortion in the central part of I helix, upon binding of single testosterone or corticosterone (CYP109E1-TES is depicted here) allowed for binding of four ordered water molecules in the back of I helix (labeled a, b, c, and d). The putative proton delivery pathway from conserved Thr246 to bulk solvent is shown, involving water molecules found in the ‘solvent channel’ between the F, E and I helices. The solvent accessible surface is shown (calculated after removing all waters, testosterone and the heme from the structure). For clarity, only a selection of the hydrogen bonds of the waters with protein atoms are shown. Residue Glu245 does not participate in the hydrogen bond network.

attack conformations occurred in the simulations, in accordance with the observed stereoselectivity of the enzyme. Interactions of one of the most frequently visited testosterone-binding poses were analyzed in detail (Fig. 9). The testosterone molecule is positioned optimally for abstraction of the C16 β -hydrogen, and forms hydrophobic interactions with the same residues as in the CYP109E1-TES crystal structure, that is, Leu80 (SRS1), Ile168 (SRS2), Ile241 and Thr246 (SRS4), Val289 (SRS5), Phe391 and Val392 (SRS6). In addition, the testosterone C3-keto oxygen is in hydrogen-bonding distance to Lys187 (SRS3). In the case of corticosterone, only 0.8% of the MD simulation frames showed near-attack conformations (suggesting generation of a 16 α -hydroxylated product). The results

are in concordance with the lack of experimentally observed corticosterone conversion by CYP109E1.

Site-directed mutagenesis

To substantiate the crystallographic and modeling results, single alanine mutations were prepared of a few selected residues (Val169, Lys187, Ile241, Glu245, and Thr246) in the active site pocket of CYP109E1 and their effect on CYP109E1-catalyzed conversion of testosterone was analyzed using the *in vitro* activity assay (Table 3). Replacement of Val169 and Ile241 by alanine resulted in almost complete abolishment of 16 β -hydroxytestosterone production, confirming the importance of these residues for productive steroid binding. The K187A mutation also caused a decrease in activity compared to the wild-type enzyme, but the effect is much smaller than for the V169 and I241 mutations. Thus, the hydrogen bond of the testosterone C3-keto group with the side chain of Lys187, as observed in the MD simulations, is not a crucial interaction for productive binding of testosterone. Interestingly, alanine mutations of Glu245 and Thr246 (the conserved ‘acid-alcohol’ pair) led to opposite effects on CYP109E1 activity toward testosterone. While the T246A mutation resulted in a drastic decrease in 16 β -hydroxytestosterone production, the E245A mutation did not significantly affect the CYP109E1 activity. These results support the relevance of the water channel observed in the single steroid-bound CYP109E1 structures, and implicates a role for Thr246, but not Glu245, in proton delivery and oxygen activation within the active site.

Discussion

Only a few bacterial P450s have been characterized that are able to hydroxylate testosterone at different positions in the steroid skeleton, with high regio- and stereoselectivity [36]. Recent examples include CYP109B1 from *B. subtilis* and CYP154C5 from *N. farcinica*, which produce 15 β - and 16 α -hydroxylated testosterone, respectively [37,38]. CYP109E1 from *B. megaterium* is the first example of a wild-type bacterial P450 showing highly selective 16 β -hydroxylase activity toward testosterone. Production of 16 β -hydroxytestosterone has previously been observed with specific P450 BM3 mutants, obtained by protein engineering, but these mutants do not display the same high level of regio- and stereoselectivity as CYP109E1 [39–41]. Similarly, various mammalian P450s are able to hydroxylate testosterone at position 16 β , but they

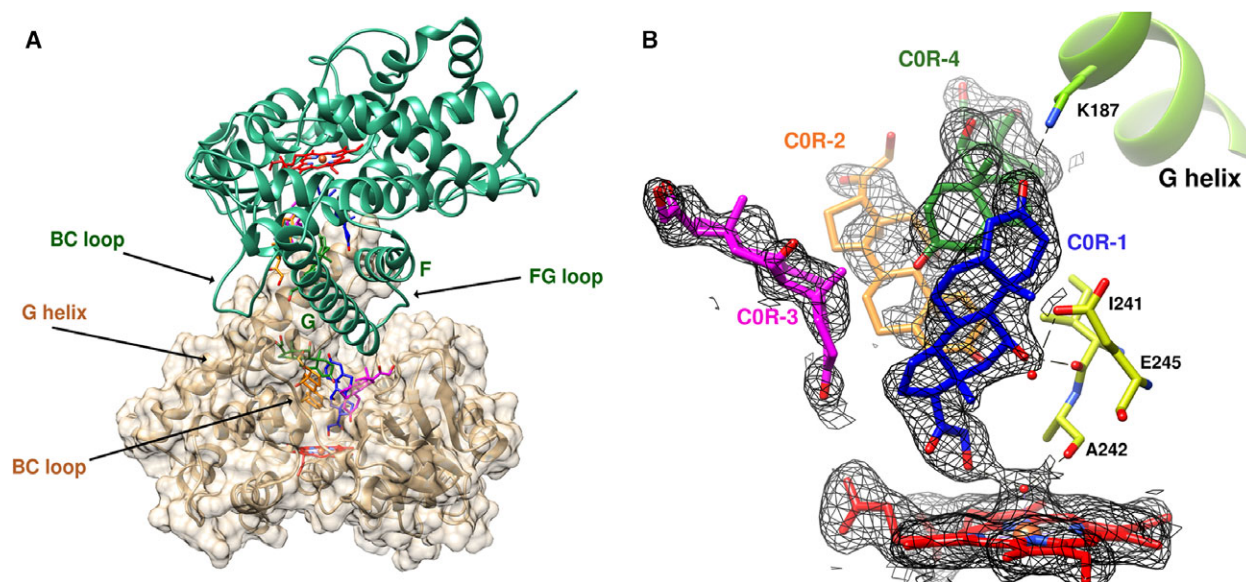


Fig. 8. Multiple corticosterone binding in CYP109E1. (A) Representation of intermolecular packing interactions in the crystal lattice of CYP109E1-COR4. The mobility of the F and G helices, FG and BC-loop is restricted, thus locking the CYP109E1 molecules in an open state. (B) Close up on multiple corticosterone binding orientations in CYP109E1 and recognition of COR-1. Hydrogen-bonding interactions of COR-1 are shown as black dashed lines. Black mesh is the composite omit $2F_o - F_c$ electron density map, calculated at 2.2 Å resolution and contoured at 1σ .

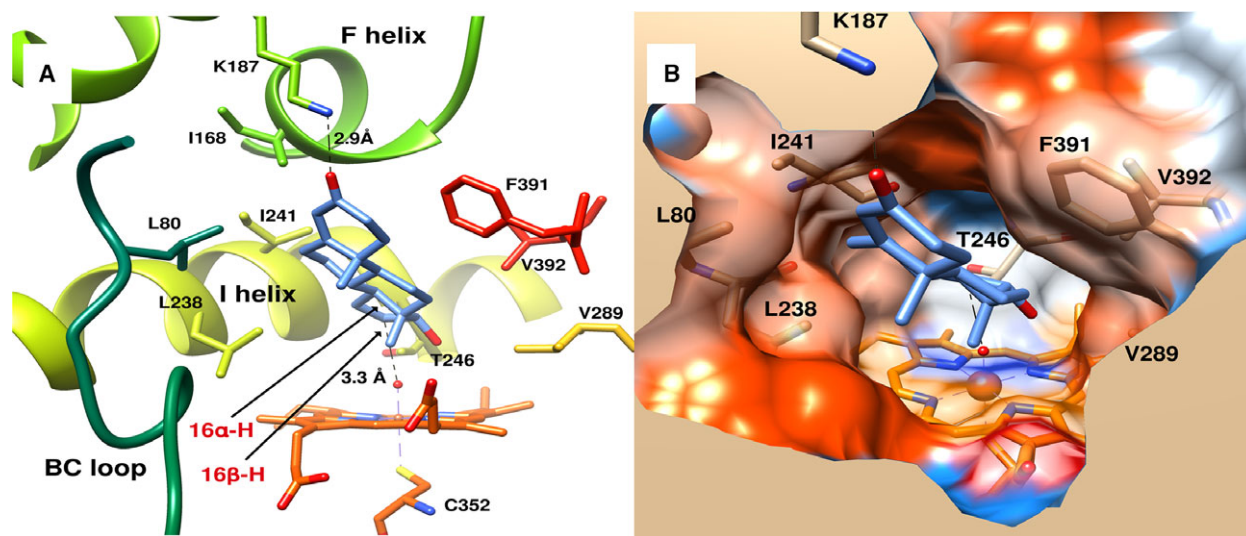


Fig. 9. Main productive testosterone binding mode in CYP109E1. (A) Representation of the most frequently visited testosterone conformation in the hydrophobic active site of CYP109E1 during molecular dynamics simulations. The C16 atom of testosterone is at an appropriate distance and angle from the oxoferryl oxygen atom to allow abstraction of its 16β -hydrogen, in accordance with the formation of the main observed turnover product, 16β -hydroxytestosterone. Testosterone is drawn in light blue-colored sticks and the hypothetical heme oxoferryl moiety ($\text{Fe(IV)} = \text{O}$ or compound I) is in orange. The locations of the two C16-hydrogens targeted in predictions by MD simulations are shown by arrows. Distances are given in Å next to the black dashed lines. (B) Surface representation of the same binding mode as depicted in panel A, coloring apolar-polar, orange-blue.

lack sufficient regio- and stereoselectivity. Low solubility and generally low expression levels of mammalian P450s further hinder their biotechnological use [42]. In

contrast, the successful development of a CYP109E1-catalyzed whole-cell system for the production of 16β -hydroxylated testosterone, as reported here,

Table 3. Effect of introduced mutations on production of 16 β -hydroxytestosterone (16 β -OH-TES) by CYP109E1 enzyme variants. The *in vitro* reactions were carried out in 50 mM potassium phosphate buffer with 2% glycerol, pH 7.4, at 30 °C for 30 min. Bovine Adx₄₋₁₀₈ and AdR were used as redox partners and 200 μ M of substrate dissolved in DMSO was added. Shown are the mean values and standard deviations (\pm SD) of three independent measurements.

CYP109E1 variant	Catalytic activity for the formation of 16 β -OH-TES, nmol·(nmol P450) ⁻¹ ·min ⁻¹
WT	4.35 \pm 0.02
K187A	1.43 \pm 0.06
V169A	0.07 \pm 0.02
I241A	0.01 \pm 0.01
E245A	4.69 \pm 0.04
T246A	0.12 \pm 0.04

demonstrates the potential of this bacterial enzyme for biotechnological applications.

To further improve CYP109E1 for biotechnological purposes and to better understand its structure-function relationships, crystal structures of CYP109E1 with and without steroids were obtained, revealing interesting features related to steroid binding and protein conformational dynamics. The structures confirm the general view that P450s possess a highly dynamic active site, which exists in a primarily open state in the absence of bound substrate, but changes toward a more closed state when the substrate is bound. Unexpectedly, and for the first time, crystal structures have revealed a P450 distal pocket with either four ligand molecules or only a single ligand molecule bound in a distinct way. Arguably, the presence of four bound corticosterone molecules in the open active site pocket of CYP109E1 is a crystallographic artifact, as it is influenced by the high concentration of the steroid in the crystal-binding experiment. Furthermore, crystal-packing interactions in the CYP109E1-COR4 structure effectively lock CYP109E1 in an open conformation, resisting the conversion into the closed structure expected upon substrate binding. On the other hand, the crystallographically observed binding of multiple corticosterone steroids to the open form of CYP109E1 may represent a snapshot of the initial substrate recognition and binding events. Multiple substrate binding occurs in other P450s, and has been implicated in the mechanisms of homo- and heterotopic cooperativity observed in mammalian P450s enzymes [43,44]. There are no indications, however, for cooperative steroid binding by CYP109E1, as evident from the hyperbolic binding curves derived from the P450 spectral shift titration assays (no sigmoidal fit).

The single steroid-bound structures of CYP109E1 in the closed conformation, complemented with the MD simulations and site-directed mutagenesis results, provide clear insights on how this enzyme accomplishes selective 16 β -hydroxylation of testosterone. The narrow shape of the active site pocket in the closed form of the enzyme, and the almost exclusively hydrophobic surface of the active site walls near the heme, restrict the binding modes of the steroids to an orientation in which their longitudinal axis is roughly perpendicular to the heme plane. Two main binding modes are possible for 3-oxo- Δ^4 -steroids like testosterone and corticosterone, with either the C3-keto oxygen down (toward heme) and C17-substituent up (away from heme) or with the C17-substituent down and C3-keto oxygen up. The first binding mode may have a possible inhibitory effect, by preventing oxygen binding to the heme-iron. Although it should be noted that coordination of the keto C3-group is less likely to occur with reduced iron in the ferrous state. The second binding mode may lead to 16 β -hydroxylation for steroids carrying a small C17 substituent, like testosterone, as supported by the MD simulations which show a substantial number of pro-16 β productive conformations. Bulky polar C17 substituents (as in corticosterone) most probably cause a steric hindrance, blocking an optimal approach of the C16 carbon to the heme-iron, and result in steroids being bound in unproductive conformations (as supported by our MD simulations with corticosterone). For testosterone, a third, minor productive binding mode positions the α -face of the C17 carbon close to the heme-iron, (in MD simulations a few pro-17 α near-attack conformations were observed), so that a second hydroxylation at this position can occur followed by removal of a water molecule, explaining the production of androstenedione as a side product. A similar oxidation reaction has very recently been observed with CYP106A1 where an 11-oxidase activity toward 11-hydroxy steroids was demonstrated [19,45].

The molecular basis for the stereoselectivity of CYP109E1 toward 16 β -hydroxylation of testosterone is further clarified by a structural comparison with CYP154C5, which converts testosterone to 16 α -hydroxytestosterone. The shape and volume of the active site pockets in the testosterone-bound crystal structures of these two enzymes are significantly different, which is primarily due to the highly variable BC loop. This loop is much longer in CYP154C5 and functions as a lid that almost completely locks the active site pocket. A closure to this extent is not observed in the testosterone-bound CYP109E1 structure. Although near to the heme, the active site pocket in testosterone-bound CYP109E1 is more restricted

compared to CYP154C5, near its entrance, it is more open and accessible to solvent. The different shapes of the active site pockets are coupled to different orientations of the bound steroids relative to the heme plane: roughly perpendicular in the CYP109E1-TES structure, while more-or-less parallel in the CYP154C5-TES complex. While the perpendicular binding mode of testosterone is optimal for C-H abstraction from the β -face of the C16 carbon, the parallel binding orientation in CYP154C5-TES positions the α -face of the testosterone C16 carbon close to the heme-iron, explaining the difference in stereoselectivity of these enzymes [14]. The parallel binding orientation of testosterone in CYP154C5 is stabilized by interactions of the apolar steroid ring system with hydrophobic residues from two opposite regions (BC loop and I helix) in the active site pocket. The polar C17 substituent forms a hydrogen bond with Gln398 (from SRS6), while the C3 substituent is bound near to Gln239 in a solvent-accessible pocket. In CYP109E1, the equivalent residues to Gln239 and Gln398 of CYP154C5 are Ile237 and Val392, respectively (Fig. 6D). Lack of polar residues in the active site pocket near the heme thus prevents a similar binding orientation of testosterone in CYP109E1 as in CYP154C5. In addition, the side chains of Leu80 (BC loop) and Ile241 (I helix) in CYP109E1 pose a steric hindrance to the incoming steroid molecule, favoring a perpendicular over a parallel binding orientation. Thus, the difference in stereoselectivity between CYP109E1 and CYP154C5 toward C16-hydroxylation of testosterone is a result of the differences in the shape and the apolar/polar surface distribution of their active site pocket, leading to a different binding orientation of the steroid relative to the heme.

In conclusion, we have identified CYP109E1 as a novel steroid-hydroxylating cytochrome P450 enzyme from *Bacillus megaterium* DSM319, with the ability to selectively convert testosterone to 16 β -hydroxytestosterone. Our combined structural, biochemical, and molecular modeling studies provided several insights into the molecular basis of steroid binding by CYP109E1 and of its regio- and stereoselectivity toward testosterone. First, binding of single steroid molecules like testosterone and corticosterone stabilizes a change in the active site pocket toward a more closed and narrow conformation. The steroid-induced structural changes include the local widening of the central I helix, which is coupled with the formation of a water channel believed to function as a water access channel and/or proton delivery network during catalysis. Secondly, the steroids may bind in two opposite orientations, with either their C3-keto oxygen or their

C16 carbon and C17 substituent directed toward the heme. The first orientation leads to nonproductive binding. The second orientation results in productive binding for testosterone, and nicely explains the high selectivity toward 16 β -hydroxylation. The larger C17 substituent of corticosterone, and the presence of its C21-hydroxyl group, prohibit productive binding of this steroid. Our results will facilitate future protein engineering experiments to improve this enzyme for biotechnological applications.

Experimental procedures

Materials

The steroid compounds used in this study were obtained from Sigma-Aldrich (Steinheim, Germany). All other chemicals were of highest grade available.

Bioinformatics analysis

Identification of close homologs and comparison of protein sequences were performed using the Basic Local Alignment Search Tool (BLAST, NCBI). Multiple sequence alignment was done with Clustal Omega [46] and visualized with ESPript3 [47]. Evolutionary analysis was carried out using Molecular Evolutionary Genetics Analysis (MEGA) version 6.0 software [48]. The phylogenetic tree was constructed using the Neighbor-joining method [49] and the evolutionary distances were computed using the Poisson correction method [50].

Cloning of wild-type enzyme

The gene encoding CYP109E1 (GenBank GeneID 9119265) was amplified by the PCR using genomic DNA of *B. megaterium* MS941, a mutant form of *B. megaterium*, derived from the DSM319 strain [51]. The PCR primers were designed (MWG-Biotech AG, Ebersberg, Germany) to introduce an *Nde*I restriction site at the 5' end of the fragment and a *Kpn*I restriction site with a 6-histidine tag at the 3' end. Following the amplification of the *cyp109e1* gene, the PCR product was cloned into the pCR4BI-TOPO vector (Invitrogen, San Diego, CA, USA) and the vector was further digested with the restriction enzymes cutting at the above-mentioned sites and ligated into the expression vector pET17b (Merck Bioscience, Bad Soden, Germany), creating the pET17b_109E1 vector. The *cyp109e1* gene was further amplified by PCR from the previously constructed pET17b expression vector using PCR primers designed to include a *Spe*I restriction site at the 5' end and a *Kpn*I restriction site at the 3' end of the fragment. The resulting PCR product was subcloned in the pCR4-TOPO vector and digested with the corresponding restriction enzymes.

The fragment was then ligated to the previously linearized pSMF2.1 vector [52], yielding the pSMF2.1E construct. Sequences of the designed primers are given in supplementary Table S1. Sequences of all created vectors were verified by DNA sequencing, carried out by Eurofins-MWG (Ebersberg, Germany).

Site-directed mutagenesis

The mutants of CYP109E1 were generated by the Quik-Change site-directed mutagenesis method using the plasmid pET17b_109E1 as template and Phusion polymerase for DNA replication (Thermo Fisher Scientific GmbH, Dreieich, Germany). The reactions were performed in 50 μ L, using a gradient cyler (PTC-200 DNA Engine cyler). Twenty cycles were carried out as follows: initial denaturation at 95 °C for 30 sec, denaturation at 95 °C for 30 sec, annealing at 58 °C for 30 sec, and extension at 72 °C for 4 min. The oligonucleotide primers for mutagenesis are shown in Table S1. Correct generation of the desired mutations was confirmed by DNA sequencing.

Expression and purification

A 30-mL preculture of *E. coli* C43 (DE3) cells carrying the pET17b_109E1 vector was grown overnight in LB medium containing 100 μ g·mL⁻¹ ampicillin at 37 °C (150 rpm). This culture was used to inoculate a 1.2-L production culture, divided over four 2-L baffled flasks, in Terrific Broth (TB) medium containing 100 μ g·mL⁻¹ ampicillin. Cultivation was continued at 37 °C (150 rpm) until the OD₆₀₀ reached 0.5, after which 1 mM IPTG and 0.5 mM δ -amino-levalulinic acid were added to start protein expression and support heme synthesis, respectively. After 24 h of incubation at 30 °C, 100 rpm, the cells were harvested by centrifugation (4500 *g* for 35 min) and the cell pellet was stored at -20 °C until further use. All purification steps were performed at 4 °C. For crystallization and spectral characterization of wild-type CYP109E1, a three-step purification procedure was applied, starting with resuspension of the cell pellet in 100 mL cold lysis buffer containing 50 mM Tris/HCl, pH 8.0, 1 mM EDTA, 20 mM NaCl, and 0.1 mM dithioerythritol, followed by the addition of 50 μ g·mL⁻¹ PMSF. The mixture was sonicated for 15 min (15 sec on, 15 sec off) on ice and, subsequently, the same amount of PMSF was added. Cell-free extract was obtained by ultracentrifugation at 40 000 *g* for 35 min at 4 °C. The supernatant containing CYP109E1 was loaded onto a 50-mL SOURCE 30Q anion-exchange column (GE Healthcare, Solingen, Germany) equilibrated with three column volumes of 20 mM Tris/HCl, pH 7.4, 0.1 mM dithioerythritol. The column was washed with the same buffer before elution of CYP109E1 with a linear gradient of 0–500 mM NaCl. Fractions with the highest A₄₁₈/A₂₈₀ ratio were

combined and concentrated by ultrafiltration using a 30-kDa cutoff membrane (Amicon Ultra/Millipore). The protein concentrate was then manually loaded onto a Superdex 75 (200 mL) gel filtration column (GE Healthcare) and CYP109E1 was eluted with 50 mM potassium phosphate buffer, pH 7.4, 0.1 mM dithioerythritol. The fractions with the highest A₄₁₈/A₂₈₀ ratio were pooled, diluted 1:5 with 5 mM potassium phosphate buffer, pH 7.4 and 0.05 mM dithioerythritol, before loading them onto a hydroxyapatite column (50 mL, Bio-Rad, Hercules, CA, USA). The column was washed with 10 mM potassium phosphate buffer, pH 7.4, 0.1 mM dithioerythritol, and CYP109E1 was eluted with a buffer concentration gradient ranging from 10 to 100 mM. Fractions containing purified CYP109E1 with an A₄₁₈/A₂₈₀ ratio larger than 1.6 were collected, concentrated by ultrafiltration using a 30-kDa cutoff membrane, and stored at -80 °C after flash-freezing in liquid nitrogen. For *in vitro* conversion experiments, the wild-type protein and its mutants were purified with an alternative one-step purification procedure. The cell pellets were resuspended in 50 mM potassium phosphate buffer, pH 7.4, containing 300 mM NaCl and 20% glycerol. Then, PMSF was added to a final concentration of 1 mM and the suspension was sonicated for 15 min (15 sec on, 15 sec off) on ice. Cell-free extract was obtained by ultracentrifugation at 30 000 rpm for 30 min, at 4 °C. The supernatant was applied to immobilized metal ion affinity chromatography column (TALON, Takara Bio Europe, Saint-Germain-en-Laye, France) equilibrated with 50 mM potassium phosphate buffer, pH 7.4, containing 300 mM NaCl and 20% glycerol. The column was washed with 5 column volumes of the same buffer containing 20 mM imidazole, and the tagged protein was eluted with a buffer containing 150 mM imidazole.

Enzyme analysis

The UV/Vis spectra of the purified protein were recorded using a double-beam spectrophotometer (UV-2101PC, Shimadzu, Japan) from 200 to 700 nm. CYP109E1 concentrations were determined by CO-difference spectroscopy of the reduced protein, following the method of Omura and Sato [53] and using an extinction coefficient of 91 mm⁻¹·cm⁻¹. Protein purity was analyzed by SDS/PAGE.

Difference spectroscopy and determination of dissociation constants (*K_D*)

Substrate-induced spin-shift states were studied via a double-beam spectrophotometer (UV-2101PC, Shimadzu, Japan), using tandem quartz cuvettes, according to the method of Schenkman and Jansson [54]. One of the cuvette chambers contained the purified CYP109E1 protein solution (10 μ M) in 50 mM potassium phosphate buffer, pH 7.4, while the other chamber was filled with the corresponding

buffer only. The steroids were dissolved in DMSO (2.5–20 mM stock solutions) and the enzyme solution was titrated with increasing amounts of testosterone or corticosterone, in the range of 0–200 μM or until saturation was reached, while the spectrum was recorded between 350 and 500 nm. All titrations were carried out in triplicates. The data were analyzed by plotting the peak-to-trough differences ($\Delta A: A_{\text{max}} - A_{\text{min}}$) against the steroid concentrations. The subsequent hyperbolic fitting was performed using ORIGIN software (OriginLab Corporation, Northampton, MA, USA), and the equilibrium dissociation constants (K_D) were determined with a regression coefficient of $R^2 = 0.99$.

In vitro substrate turnover

The *in vitro* turnover of tested steroids was performed in a final volume of 250 μL , using 50 mM potassium phosphate buffer with 2% glycerol, pH 7.4, at 30 °C for 30 min. The reconstituted system contained CYP109E1 (1 μM), bovine Adx₄₋₁₀₈ (20 μM), AdR (2 μM), a NADPH-regenerating system (1 mM MgCl_2 , 5 mM glucose-6-phosphate, 1 U glucose-6-phosphate dehydrogenase) and 200 μM of the corresponding steroid dissolved in DMSO. The reactions were initiated by the addition of NADPH to a final concentration of 1 mM, then stopped and extracted twice with the addition of 250 μL of ethyl acetate. The samples were centrifuged (10 000 rpm, 10 min), the organic phases were combined and evaporated until complete dryness and, after resuspension, analyzed by HPLC. Since, the absorption properties of the products did not differ from the respective substrates, the product formation was calculated from the relative peak area (area %) of the HPLC chromatograms, dividing each respective product peak area by the sum of all peak areas.

Whole-cell conversion with *Bacillus megaterium* MS941

The *in vivo* conversions were done with the *B. megaterium* strain MS941 [51]. The MS941 cells were transformed with the pSMF2.1E vector, using the polyethylene glycol-mediated protoplast transformation method [55]. For cultivation, a complex medium was used (24 g·L⁻¹ yeast extract, 12 g·L⁻¹ soytone, 2.31 g·L⁻¹ KH_2PO_4 , and 1.25 g·L⁻¹ KH_2PO_4) supplemented with 10 $\mu\text{g}\cdot\text{mL}^{-1}$ tetracycline at 30 °C, 180 rpm. First, a 50-mL overnight culture was prepared, inoculated from a –80 °C glycerol stock of the transformed MS941 cells. This culture was then used to inoculate the main culture in a 300-mL baffled shake flask filled with 50 mL complex medium. The main culture was incubated, until the OD₅₇₈ reached 0.4, when the protein expression was induced by the addition of xylose at a final concentration of 5 mg·mL⁻¹. After 24 h, the cultures were harvested by centrifugation (15 min, 10 000 g, 4 °C) and the pellets resuspended in 50 mL of 50 mM potassium phosphate buffer, pH 7.4. The substrates, dissolved in DMSO, were added at a final concentration of

200 μM . The 500 μL samples were taken from the cultures at fixed time points, followed by extraction and HPLC analysis. To obtain sufficient product quantities for structural analysis by NMR spectroscopy (mg range), the volume of the main culture was increased to 750 mL (3 \times 250 mL) in 2-L baffled flasks. After a 24-h expression period, the cultures were harvested by centrifugation, resuspended in 375 mL of 50 mM potassium phosphate buffer, pH 7.4, and the corresponding substrates were added at a final concentration of 300 μM . Following the conversion, the cultures were extracted twice with ethyl acetate, the organic phases were combined and evaporated (Rotavapor R-114; BÜCHI Labortechnik AG, Flawil, Switzerland). The purification and isolation of products was carried out by preparative HPLC.

HPLC analysis and product isolation

The HPLC analyses were carried out with a Jasco system (Pu-980 HPLC pump, AS-950 sampler, UV-975 UV/visible detector, LG-980-02 gradient unit; Jasco, Gross-Umstadt, Germany), using a reversed-phase ec MN Nucleodur C₁₈ (3 μM , 4.0 \times 125 mm) column (Macherey-Nagel, Bethlehem, PA, USA) kept at an oven temperature of 40 °C. The steroids were eluted using a gradient method, starting with a mobile phase consisting of acetonitrile:H₂O in a ratio of 1:9, increasing it to 1:1. The flow rate was 1 mL·min⁻¹ and the UV detection of the substrate and product was accomplished at 240 or 254 nm. For product isolation, the dissolved extracts were filtered (Acrodisc® PTFE syringe filter, 0.45 μm , PALL) and injected to a reversed-phase ec MN Nucleodur C₁₈ (5 μM , 8.0 \times 250 mm) column (Macherey-Nagel, Bethlehem, PA, USA). The HPLC run was carried out analogously to the analytical HPLC conditions with an increased flow rate of 2.5 mL·min⁻¹. Fractions were collected with an Advantec CHF122 SB fraction collector, combined, and evaporated on a rotary evaporator (Rotavapor R-114 from BÜCHI Labortechnik AG).

NMR spectroscopy

The NMR spectra were recorded in deuterated chloroform (CDCl_3) with a Bruker DRX 500 or a Bruker Avance 500 NMR spectrometer at 300 K. The chemical shifts were relative to CDCl_3 at δ 24.7 (¹H NMR) and 77.00 (¹³C NMR), 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.

Crystallization

After thawing, aliquots of purified CYP109E1 were buffer exchanged to 20 mM Tris/HCl, pH 8.0, 0.1 mM

dithioerythritol using a PD-10 desalting column (GE Healthcare) and concentrated to 40 mg·mL⁻¹. Screening for crystallization growth conditions was done at 293 K by the sitting-drop vapor-diffusion method, using 96-well crystallization plates, a Mosquito crystallization robot, (TTP LabTech, Melbourn, UK) and a few commercially available screens. Lead crystallization conditions were optimized manually in 24-well plates at 293 K. Initial crystals grew with a reservoir solution containing 25% poly(ethylene glycol) 3350, 0.1 M Bis-tris, pH 6.5, and 4% tacsimate reagent (pH 6.0) (Hampton Research, Aliso Viejo, CA, USA). Diffraction quality crystals were obtained by applying a streak-seeding protocol. First, protein drops were prepared by mixing 2 µL of concentrated CYP109E1 (40 mg·mL⁻¹) with an equal volume of reservoir solution, containing 20% poly(ethylene glycol) 3350, 0.1 M Bis-tris, pH 6.5, and 4% tacsimate reagent pH 6.0. Droplets were equilibrated against 500 µL of the reservoir solution for 1 h, and then streak-seeded. Red-colored, cubic-shaped crystals of CYP109E1 (space group P3₂21) grew overnight and reached a final average size of approximately 0.15 × 0.15 × 0.15 mm³ in about 7 days. Corticosterone-bound CYP109E1 crystals were initially obtained by crystal soaking experiments. For this, corticosterone was added as a solid powder directly into crystallization drops containing native CYP109E1 crystals, followed by equilibration against the original crystallization solution for 3–4 weeks. Alternatively, cocrystallization screens were employed with corticosterone and testosterone (in a 1:5 molar ratio of protein:steroid, using stocks solutions of the steroids in DMSO), resulting in large cuboid-shaped crystals (space group P2₁) grown from 9% poly(ethylene glycol) 3350 and 8% tacsimate pH 5.0 (Hampton Research). Prior to data collection, crystals were briefly soaked in a cryoprotectant solution containing mother liquor, supplemented with 20% (v/v) glycerol. Cryoprotection of crystals obtained by cocrystallization was accomplished by raising the poly(ethylene glycol) 3350 concentration to 35%. Subsequently, all crystals were flash-cooled at 100 K in the cold nitrogen gas stream of the camera's cryostat.

Data collection and structure determination

X-ray diffraction data were collected at the ID29, ID23-1, and ID23-2 beam lines of the European Synchrotron Radiation Facility (ESRF), Grenoble, all equipped with Pilatus detectors. Additional data were recorded using an in-house rotating anode X-ray source (Bruker MicroSTAR) and an image plate detector (*mar345*TM). Single crystals of CYP109E1 were used to obtain diffraction datasets in the 2.55–2.1 Å resolution range. Reflections were indexed and integrated using iMosflm [56] or XDS [57], while scaling and merging of the data was done with AIMLESS from the CCP4 software suite [58]. The structure of native CYP109E1 was solved by molecular replacement with

Phaser from the PHENIX suite [59], using the structure of HmtT from *Streptomyces himastatinicus* (PDB ID 4ggv, 39% sequence identity) as a search model. Two protein molecules were located in the asymmetric unit, consistent with Matthews coefficient calculations indicating a solvent content of 61%. Automatic model rebuilding using routines in PHENIX yielded approximately 80% of the polypeptide model. The model was completed by iterative cycles of model building with Coot [60] and structure refinement with Phenix.refine [61]. The structure of substrate-free CYP109E1 then served as a starting point to solve the CYP109E1 steroid-bound structures. At the final stages of refinement, water molecules were added to the structures based on peaks in the electron density maps and using strict interaction criteria. The quality of the refined protein models was validated using MolProbity [62].

Structure analysis

Pairwise comparison of the obtained structures with other P450 structures and RMSD calculations were done with the PDBeFold engine [63]. Protein–ligand interactions were analyzed by LigPlot+ [64]. Substrate recognition sites (SRS) in CYP109E1 were identified based on alignment with P450_{cam} and the description provided by Gotoh [29]. The respective residue ranges in CYP109E1 are: SRS1 67–89, SRS2 165–171, SRS3 187–194, SRS4 230–249, SRS5 286–295, and SRS6 388–395.

Molecular dynamics simulations

Molecular dynamics simulations were performed using the CYP109E1-TES crystal structure as a starting model, following the same procedure as previously described [33]. First, testosterone and all water molecules were removed from the crystal structure, and the heme and its cysteine ligand were replaced by a model of heme compound I covalently bound to cysteine. The tested ligand (testosterone or corticosterone) was then manually placed back in the binding pocket using PyMOL (Schrödinger, Cambridge, MA, USA). Testosterone was placed in 10 different starting orientations where the distance between the C16β hydrogen and the ferryl oxygen atom of heme compound I was less than 4 Å. Corticosterone was placed in starting orientations similar to the COR-1 steroid orientation in the CYP109E1-COR4 crystal structure. To ensure extensive conformational sampling, ten 100-ns MD simulations with different initial orientations of the steroid were performed for each CYP109E1-steroid system, and the AMBER03 force field [65] in GROMACS version 5.0.4 was used [66]. Force fields for testosterone and corticosterone were generated by use of the structures derived from the PubChem database [67] and subsequent energy minimization in YASARA (www.YASARA.org). The RESP partial charges of the molecules were calculated using the R.E.D webserver with

the RESP-A1B charge model [68]. The force field was built with the ANTECHAMBER module of AMBER 10 [69] and converted into the GROMACS topology format. The force field of the cysteine-heme compound I complex was used as previously described [70]. An octahedral water box of SPC/E water with periodic boundaries at least 1.2 nm from the protein was used. Simulations were run at 300 K and 1 bar. Pressure coupling was performed with a Parrinello–Rahman barostat [71]. The Nose–Hoover coupling scheme was used to maintain the temperature, with coupling constant of 0.5 ps [72]. Initial velocities were randomly assigned. The LINCS algorithm was applied to constrain all bonds containing hydrogen atoms [73]. Seventeen sodium counter ions were added to maintain the neutral charge of the systems. Long-range electrostatic interactions were treated by using the particle-mesh Ewald method [74]. Energy minimization was performed using the steepest descent method with positional restraints on protein heavy atoms and a maximum allowed force of 1000 kJ·mol⁻¹·nm⁻¹. A 2-fs time step was used and coordinates were saved every 1000 steps (2 ps). To determine near-attack conformations the distance between the 16 α or 16 β hydrogen and the ferryl oxygen of compound I ferryl oxygen was measured. In case of simulations with testosterone, additionally the distance between the 17 α hydrogen and the ferryl oxygen was measured. Angles between the C16 or C17 carbon, their hydrogens and the ferryl oxygen were also measured. Based on the previously described cut-offs [33–35], ligand conformations were considered near to attack, if at least one of the measured distances was < 3.5 Å and the angle was 180 \pm 45°. All other conformations were considered nonproductive. The number of frames in MD trajectories suggesting pro-16 α , pro-16 β , pro-17 α , and nonproductive conformations were counted and expressed as percentages.

Acknowledgements

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's 7th Framework Programme (FP7/2007–2013) under REA Grant Agreement 289217 (ITN P4FIFTY). We thank the beam-line scientists of ID29, ID23-1, and ID23-2 (ESRF, European Synchrotron Radiation Facility) for assistance. We thank the high performance computing center Stuttgart (HLRS) for their support and for supplying the computational resources.

Author contributions

IKJ, FMK, LG, RB, and AMWHT designed the study. IKJ, FMK, LG, AA, EB, and JZ performed the experiments. IKJ, FMK, LG, EB, and AA analyzed

the data. IKJ, FMK, LG, AA, JP, RB, and AMWHT wrote the manuscript. JP, RB, and AMWHT provided supervision.

References

- Bernhardt R & Urlacher VB (2014) Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl Microbiol Biotechnol* **98**, 6185–6203.
- Bhatti HN & Khera RA (2012) Biological transformations of steroidal compounds: a review. *Steroids* **77**, 1267–1290.
- Donova MV & Egorova OV (2012) Microbial steroid transformations: current state and prospects. *Appl Microbiol Biotechnol* **94**, 1423–1447.
- Tong WY & Dong X (2009) Microbial biotransformation: recent developments on steroid drugs. *Recent Pat Biotechnol* **3**, 141–153.
- Grogan G (2011) Cytochromes P450: exploiting diversity and enabling application as biocatalysts. *Curr Opin Chem Biol* **15**, 241–248.
- Ost TW, Miles CS, Murdoch J, Cheung Y, Reid GA, Chapman SK & Munro AW (2000) Rational re-design of the substrate binding site of flavocytochrome P450 BM3. *FEBS Lett* **486**, 173–177.
- Noble MA, Miles CS, Chapman SK, Lysek DA, MacKay AC, Reid GA, Hanzlik RP & Munro AW (1999) Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem J* **339** (Pt 2), 371–379.
- Denisov IG, Shih AY & Sligar SG (2012) Structural differences between soluble and membrane bound cytochrome P450s. *J Inorg Biochem* **108**, 150–158.
- Johnson EF & Stout CD (2005) Structural diversity of human xenobiotic-metabolizing cytochrome P450 monooxygenases. *Biochem Biophys Res Commun* **338**, 331–336.
- Podust LM & Sherman DH (2012) Diversity of P450 enzymes in the biosynthesis of natural products. *Nat Prod Rep* **29**, 1251–1266.
- Otyepka M, Berka K & Anzenbacher P (2012) Is there a relationship between the substrate preferences and structural flexibility of cytochromes P450? *Curr Drug Metab* **13**, 130–142.
- McLean KJ, Lafite P, Levy C, Cheesman MR, Mast N, Pikuleva IA, Leys D & Munro AW (2009) The structure of *Mycobacterium tuberculosis* CYP125: molecular basis for cholesterol binding in a P450 needed for host infection. *J Biol Chem* **284**, 35524–35533.
- Janocha S, Carius Y, Hutter M, Lancaster CR & Bernhardt R (2016) Crystal structure of CYP106A2 in substrate-free and substrate-bound form. *ChemBioChem* **9**, 852–860.

- 14 Herzog K, Bracco P, Onoda A, Hayashi T, Hoffmann K & Schallmeyer A (2014) Enzyme-substrate complex structures of CYP154C5 shed light on its mode of highly selective steroid hydroxylation. *Acta Crystallogr D Biol Crystallogr* **70**, 2875–2889.
- 15 Makino T, Katsuyama Y, Otomatsu T, Misawa N & Ohnishi Y (2014) Regio- and stereospecific hydroxylation of various steroids at the 16 α position of the D ring by the *Streptomyces griseus* cytochrome P450 CYP154C3. *Appl Environ Microbiol* **80**, 1371–1379.
- 16 Zhang A, Zhang T, Hall EA, Hutchinson S, Cryle MJ, Wong LL, Zhou W & Bell SG (2015) The crystal structure of the versatile cytochrome P450 enzyme CYP109B1 from *Bacillus subtilis*. *Mol Biosyst* **3**, 869–881.
- 17 Eppinger M, Bunk B, Johns MA, Edirisinghe JN, Kutumbaka KK, Koenig SS, Creasy HH, Rosovitz MJ, Riley DR, Daugherty S *et al.* (2011) Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J Bacteriol* **193**, 4199–4213.
- 18 Schmitz D, Zapp J & Bernhardt R (2014) Steroid conversion with CYP106A2 – production of pharmaceutically interesting DHEA metabolites. *Microb Cell Fact* **13**, 81.
- 19 Kiss FM, 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**, 2320–2326.
- 20 Brill E, Hannemann F, Zapp J, Bruning G, Jauch J & Bernhardt R (2014) A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto-beta-boswellic acid (KBA). *Appl Microbiol Biotechnol* **98**, 1701–1717.
- 21 Lee GY, Kim DH, Kim D, Ahn T & Yun CH (2015) Functional characterization of steroid hydroxylase CYP106A1 derived from *Bacillus megaterium*. *Arch Pharm Res* **38**, 98–107.
- 22 Nelson DR (2009) The cytochrome P450 homepage. *Human Genomics* **4**, 59–65.
- 23 Khatri Y, Hannemann F, Ewen KM, Pistorius D, Perlova O, Kagawa N, Brachmann AO, Muller R & Bernhardt R (2010) The CYPome of *Sorangium cellulosum* So ce56 and identification of CYP109D1 as a new fatty acid hydroxylase. *Chem Biol* **17**, 1295–1305.
- 24 Khatri Y, Hannemann F, Girhard M, Kappl R, Meme A, Ringle M, Janocha S, Leize-Wagner E, Urlacher VB & Bernhardt R (2013) Novel family members of CYP109 from *Sorangium cellulosum* So ce56 exhibit characteristic biochemical and biophysical properties. *Biotechnol Appl Biochem* **60**, 18–29.
- 25 Furuya T, Shibata D & Kino K (2009) Phylogenetic analysis of *Bacillus* P450 monooxygenases and evaluation of their activity towards steroids. *Steroids* **74**, 906–912.
- 26 Khatri Y, Ringle M, Lisurek M, von Kries JP, Zapp J & Bernhardt R (2016) Substrate hunting for the myxobacterial CYP260A1 revealed new 1 α -hydroxylated products from C-19 steroids. *ChemBioChem* **17**, 90–101.
- 27 Schmitz D, Zapp J & Bernhardt R (2012) Hydroxylation of the triterpenoid dipterocarpol with CYP106A2 from *Bacillus megaterium*. *FEBS J* **279**, 1663–1674.
- 28 Poulos TL, Finzel BC, Gunsalus IC, Wagner GC & Kraut J (1985) The 2.6-Å crystal structure of *Pseudomonas putida* cytochrome P-450. *J Biol Chem* **260**, 16122–16130.
- 29 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.
- 30 Gricman L, Vogel C & Pleiss J (2015) Identification of universal selectivity-determining positions in cytochrome P450 monooxygenases by systematic sequence-based literature mining. *Proteins* **83**, 1593–1603.
- 31 Hamdane D, Zhang H & Hollenberg P (2008) Oxygen activation by cytochrome P450 monooxygenase. *Photosynth Res* **98**, 657–666.
- 32 Cojocaru V, Winn PJ & Wade RC (2007) The ins and outs of cytochrome P450s. *Biochim Biophys Acta* **1770**, 390–401.
- 33 Eichler A, Gricman L, Herter S, Kelly P, Turner N, Pleiss J & Flitsch S (2015) Enantioselective benzylic hydroxylation catalysed by P450 monooxygenases: characterisation of a P450cam mutant library and molecular modelling. *ChemBioChem* **5**, 426–432.
- 34 Paulsen MD & Ornstein RL (1992) Predicting the product specificity and coupling of cytochrome P450cam. *J Comput Aided Mol Des* **6**, 449–460.
- 35 Filipovic D, Paulsen MD, Loida PJ, Sligar SG & Ornstein RL (1992) Ethylbenzene hydroxylation by cytochrome P450cam. *Biochem Biophys Res Commun* **189**, 488–495.
- 36 Agematu H, Matsumoto N, Fujii Y, Kabumoto H, Doi S, Machida K, Ishikawa J & Arisawa A (2006) Hydroxylation of testosterone by bacterial cytochromes P450 using the *Escherichia coli* expression system. *Biosci Biotechnol Biochem* **70**, 307–311.
- 37 Girhard M, Klaus T, Khatri Y, Bernhardt R & Urlacher VB (2010) Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Appl Microbiol Biotechnol* **87**, 595–607.
- 38 Bracco P, Janssen DB & Schallmeyer A (2013) Selective steroid oxyfunctionalisation by CYP154C5, a bacterial cytochrome P450. *Microb Cell Fact* **12**, 95.

- 39 van Vugt-Lussenburg BM, Damsten MC, Maasdijk DM, Vermeulen NP & Commandeur JN (2006) Heterotropic and homotropic cooperativity by a drug-metabolising mutant of cytochrome P450 BM3. *Biochem Biophys Res Commun* **346**, 810–818.
- 40 Rea V, Kolkman AJ, Vottero E, Stronks EJ, Ampt KA, Honing M, Vermeulen NP, Wijmenga SS & Commandeur JN (2012) Active site substitution A82W improves the regioselectivity of steroid hydroxylation by cytochrome P450 BM3 mutants as rationalized by spin relaxation nuclear magnetic resonance studies. *Biochemistry* **51**, 750–760.
- 41 Venkataraman H, Beer SB, Bergen LA, Essen N, Geerke DP, Vermeulen NP & Commandeur JN (2012) A single active site mutation inverts stereoselectivity of 16-hydroxylation of testosterone catalyzed by engineered cytochrome P450 BM3. *ChemBioChem* **13**, 520–523.
- 42 Niwa T, Murayama N, Imagawa Y & Yamazaki H (2015) Regioselective hydroxylation of steroid hormones by human cytochromes P450. *Drug Metab Rev* **47**, 89–110.
- 43 Davydov DR & Halpert JR (2008) Allosteric P450 mechanisms: multiple binding sites, multiple conformers or both? *Expert Opin Drug Metab Toxicol* **4**, 1523–1535.
- 44 Zhao B, Lei L, Kagawa N, Sundaramoorthy M, Banerjee S, Nagy LD, Guengerich FP & Waterman MR (2012) Three-dimensional structure of steroid 21-hydroxylase (cytochrome P450 21A2) with two substrates reveals locations of disease-associated variants. *J Biol Chem* **287**, 10613–10622.
- 45 Kiss FM, Schmitz D, Zapp J, Dier TK, Volmer DA & Bernhardt R (2015) Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* – identification of a novel 11-oxidase activity. *Appl Microbiol Biotechnol* **99**, 8495–8514.
- 46 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.
- 47 Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–W324.
- 48 Tamura K, Stecher G, Peterson D, Filipowski A & Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**, 2725–2729.
- 49 Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- 50 Zuckerkandl E & Pauling L (1965) Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins* (Bryson V & Vogel HJ, eds), pp. 97–166. Academic Press, Cambridge, MA.
- 51 Wittchen KD & Meinhardt F (1995) Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. *Appl Microbiol Biotechnol* **42**, 871–877.
- 52 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**, 1135–1146.
- 53 Omura T & Sato R (1964) The carbon monoxide-binding pigment of liver microsomes: I. Evidence for its hemoprotein nature. *J Biol Chem* **239**, 2370–2378.
- 54 Schenkman JB & Jansson I (1998) Spectral analyses of cytochromes P450. *Methods Mol Biol* **107**, 25–33.
- 55 Barg H, Malten M, Jahn M & Jahn D (2005) Protein and vitamin production in *Bacillus megaterium*. In *Microbial Processes and Products* (Barredo J, ed.), pp. 205–223. Humana Press, New York, NY.
- 56 Battye TG, Kontogiannis L, Johnson O, Powell HR & Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281.
- 57 Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
- 58 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242.
- 59 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.
- 60 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
- 61 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352–367.
- 62 Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21.
- 63 Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256–2268.

- 64 Laskowski RA & Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778–2786.
- 65 Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T *et al.* (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* **24**, 1999–2012.
- 66 Hess B, Kutzner C, van der Spoel D & Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* **4**, 435–447.
- 67 Wang Y, Xiao J, Suzek TO, Zhang J, Wang J & Bryant SH (2009) PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res* **37**, W623–W633.
- 68 Vanquelef E, Simon S, Marquant G, Garcia E, Klimerak G, Delepine JC, Cieplak P & Dupradeau FY (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res* **39**, W511–W517.
- 69 Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W *et al.* (2008) *Amber 10*. University of California.
- 70 Seifert A, Tatzel S, Schmid RD & Pleiss J (2006) Multiple molecular dynamics simulations of human p450 monooxygenase CYP2C9: the molecular basis of substrate binding and regioselectivity toward warfarin. *Proteins* **64**, 147–155.
- 71 Parrinello M & Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys* **52**, 7182–7190.
- 72 Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys* **52**, 255–268.
- 73 Hess B, Bekker H, Berendsen HJC & Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* **18**, 1463–1472.
- 74 Darden T, York D & Pedersen L (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089–10092.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. Oligonucleotide primers used in this work.

Table S2. Chemical structures of compounds used for CYP109E1 substrate screening.

Table S3. Structural NMR data of 16 β -hydroxytestosterone and androstendione in CDCl₃.

Table S4. Binding interactions of COR-2, COR-3, and COR-4 in CYP109E1-COR4 crystal structure.

Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*

Ilona K. Jóźwik, Flora M. Kiss, Łukasz Gricman, Ammar Abdulmughni, Elisa Brill, Josef Zapp, Juergen Pleiss, Rita Bernhardt and Andy-Mark W. H. Thunnissen

DOI: 10.1111/febs.13911

Supplementary Information

for

Jóźwik *et al.* (2016)

Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1

from *Bacillus megaterium*

Table S1 Oligonucleotide primers used in this work. The restriction sites for the respective restriction enzymes are shown in *italics*, and the sequence encoding the His₆-tag is underlined.

Name	Description	Nucleotide sequence (5' → 3')
205-for-NdeI	Forward primer for the amplification of CYP109E1 with NdeI restriction site	<i>CATATG</i> AAAACAGAAAGAGAAAACGG
205-rev-KpnI	Reverse primer for the amplification of CYP109E1 with His ₆ -tag and KpnI restriction site	<u>GGTACCTTAATGGTGATGGTGATGGTG</u> TACGTTTTTACGAATCAATAATT
205-SpeI-for	Forward primer for the amplification of CYP109E1 with SpeI restriction site and RBS	<i>ACTAGTAAATCAAGGAGGTGAATATACAATG</i> AAAACAGAAAGAGAAAACGG
205-KpnI-rev	Reverse primer for the amplification of CYP109E1 with His ₆ -tag and KpnI restriction site	<u>GGTACCTTATACGTTTTTACGAATCAATAATT</u>
205-RT-for	RT-Forward primer for the amplification of CYP109E1	AAACGGAATTGTCCGTCAG
205-RT-rev	RT-Reverse primer for the amplification of CYP109E1	CACTTCTTTGACCCCGTCAT
Mutation	Oligonucleotides	
K187A-109E1-for	CTCCAGCAAGAG <u>GCA</u> ATGAAAGCAAATGATGAGC	
K187A-109E1-rev	GCTCATCATTTGCTTTTCAT <u>TGC</u> CTCTTGCTGGAG	
V169A-109E1-for	GATATTATC <u>GCA</u> GCCGGTCCTTCTAATAACGAACGT	
V196A-109E1-rev	ACGTTTCGTTATTAGAAGGACCGGCT <u>TGC</u> GATAATATC	
I241A-109E1-for	CTATTTTGCTACTG <u>GCT</u> GCTGGAAACGAAACAACCAC	
I241A-109E1-rev	GTGGTTGTTTCGTTTCCAGC <u>AGC</u> CAGTAGCAAAATAG	
E245A-109E1-for	CTGATTGCTGGAAAC <u>GCA</u> ACAACCACAAATTTAATTTC	
E245A-109E1-rev	GAAATTAAATTTGTGGTTGT <u>TGC</u> GTTTCCAGCAATCAG	
T246A-109E1-for	CTGATTGCTGGAAACGA <u>GCA</u> ACCACAAATTTAATTTC	
T246A-109E1-rev	GAAATTAAATTTGTGGT <u>TGCT</u> TCGTTTCCAGCAATCAG	

Table S2 Chemical structures of compounds used for CYP109E1 substrate screening.

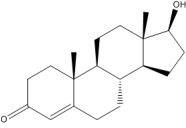
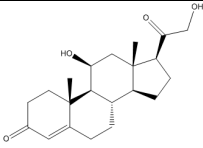
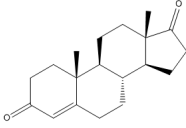
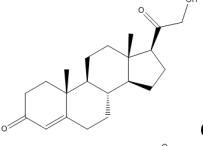
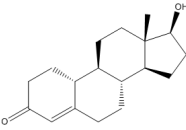
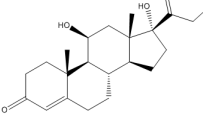
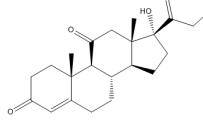
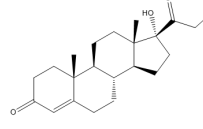
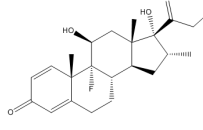
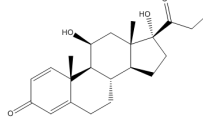
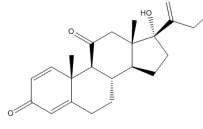
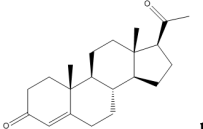
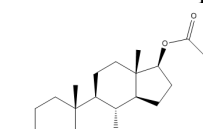
Short C17 substituent at steroid D ring	Large C17 substituent at steroid D ring
 <p>testosterone (11)</p>	 <p>corticosterone (4)</p>
 <p>androstenedione (1)</p>	 <p>deoxycorticosterone (5)</p>
 <p>19- nortestosterone (13)</p>	 <p>cortisol (2)</p>
	 <p>cortisone (3)</p>
	 <p>11-deoxycortisol (6)</p>
	 <p>dexamethasone (7)</p>
	 <p>prednisolone (8)</p>
	 <p>prednisone (9)</p>
	 <p>progesterone (10)</p>
	 <p>testosterone acetate (12)</p>

Table S3 Structural NMR data of 16 β -hydroxytestosterone and androstendione in CDCl₃.

16β-Hydroxytestosterone			Androstendione	
	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
1	35.63	2.00 <i>ddd</i> (13.5, 5.0, 3.5) 1.66 <i>ddd</i> (14.0, 13.5, 5.0)	35.69	2.02 <i>ddd</i> (13.5, 5.0, 3.2) 1.70 <i>m</i>
2	33.88	2.38 <i>m</i> 2.31 <i>m</i>	33.89	2.40 <i>m</i> 2.35 <i>m</i>
3	199.68	-----	199.29	-----
4	123.84	5.70 <i>s</i>	124.15	5.73 <i>brs</i>
5	171.20	-----	170.27	-----
6	32.69	2.37 <i>m</i> 2.25 <i>m</i>	32.55	2.42 <i>m</i> 2.31 <i>m</i>
7	31.66	1.84 <i>m</i> 0.98 <i>m</i>	30.74	1.96 <i>m</i> 1.11 <i>m</i>
8	34.99	1.62 <i>m</i>	35.15	1.72 <i>m</i>
9	54.01	0.92 <i>m</i>	53.82	0.97 <i>ddd</i> (12.5, 10.5, 4.0)
10	38.66	-----	38.63	-----
11	20.32	1.57 <i>m</i> 1.43 <i>m</i>	20.31	1.68 <i>m</i> 1.44 <i>m</i>
12	37.01	1.87 <i>m</i> 1.08 <i>ddd</i> (13.0, 13.0, 4.5)	35.73	2.45 <i>ddd</i> (19.5, 9.0, 1.0) 2.09 <i>ddd</i> (19.5, 9.0, 9.0)
13	42.33	-----	47.49	-----
14	46.96	0.78 <i>ddd</i> (13.5, 11.0, 6.7)	50.84	1.28 <i>m</i>
15	34.91	2.20 <i>m</i> 1.28 <i>ddd</i> (13.5, 13.5, 5.0)	21.74	1.96 <i>m</i> 1.56 <i>m</i>
16	69.83	4.14 <i>ddd</i> (7.7, 7.7, 5.0)	31.27	1.85 <i>ddd</i> (13.0, 4.0, 2.5) 1.27 <i>m</i>
17	80.55	3.35 <i>d</i> (7.7)	220.30	-----
18	11.83	0.83 <i>s</i> (3H)	13.69	0.90 <i>s</i> (3H)
19	17.37	1.17 <i>s</i> (3H)	17.37	1.20 <i>s</i> (3H)

Table S4 Binding interactions of COR-2, COR-3 and COR-4 in CYP109E1-COR4 crystal structure. Hydrogen bonding and hydrophobic interactions are described as identified by LigPlot+.

Ligand	Hydrogen bonding	Hydrophobic interactions
COR-2	Corticosterone C21-hydroxyl with Gln75 (SRS1) Corticosterone C11-hydroxyl with Thr78 (SRS1) (water-mediated bond)	Arg69 (SRS1) Pro71 (SRS1) Gln75 (SRS1) Leu80 (SRS1) Gly81 (SRS1) Ile85 (SRS1) Asn86 (SRS1) Ile241 (SRS4) COR-1 COR-4
COR-3	Corticosterone C3-keto oxygen with heme propionate (water-mediated bond) Corticosterone C3-keto oxygen with amide nitrogen of His293 (SRS5) Corticosterone C21-hydroxyl group with backbone carbonyl of Arg44	Val46 Leu292 (SRS5) His312 Phe391 (SRS6)
COR-4	Corticosterone C21-hydroxyl and C20-keto oxygen with side chain of Arg76 (SRS1)	Thr78 (SRS1) Lys187 (SRS3) Met188 (SRS3) Asn191 (SRS3) Ile241 (SRS4) COR-1 COR-2 COR-3

2.4 (Putkaradze et al., 2017)

CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*.

Putkaradze N, Litzenburger M, **Ammar Abdumughni**, Mohammed Milhim, Elisa Brill, Frank Hannemann and Rita Bernhardt.

Applied Microbiology Biotechnology, 2017 Dec; 101: 8379-8393.

Reprinted with permission of the Applied Microbiology Biotechnology. All rights reserved.

CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*

Natalia Putkaradze¹ · Martin Litzenburger¹ · Ammar Abdulmughni¹ · Mohammed Milhim¹ · Elisa Brill¹ · Frank Hannemann¹ · Rita Bernhardt¹

Received: 31 May 2017 / Revised: 22 August 2017 / Accepted: 26 September 2017 / Published online: 11 October 2017
© Springer-Verlag GmbH Germany 2017

Abstract CYP109E1 is a cytochrome P450 monooxygenase from *Bacillus megaterium* with a hydroxylation activity for testosterone and vitamin D3. This study reports the screening of a focused library of statins, terpene-derived and steroidal compounds to explore the substrate spectrum of this enzyme. Catalytic activity of CYP109E1 towards the statin drug-precursor compactin and the prodrugs lovastatin and simvastatin as well as biotechnologically relevant terpene compounds including ionones, nootkatone, isolongifolen-9-one, damascones, and β -damascenone was found in vitro. The novel substrates induced a type I spin-shift upon binding to P450 and thus permitted to determine dissociation constants. For the identification of conversion products by NMR spectroscopy, a *B. megaterium* whole-cell system was applied. NMR analysis revealed for the first time the ability of CYP109E1 to catalyze an industrially highly important reaction, the production of pravastatin from compactin, as well as regioselective oxidations generating drug metabolites (6' β -hydroxy-lovastatin, 3' α -hydroxy-simvastatin, and 4''-hydroxy-simvastatin) and valuable terpene derivatives (3-hydroxy- α -ionone, 4-hydroxy- β -ionone, 11,12-epoxy-nootkatone, 4(*R*)-hydroxy-isolongifolen-9-one, 3-hydroxy- α -damascone, 4-hydroxy- β -damascone, and 3,4-epoxy- β -damascone). Besides that, a novel compound, 2-hydroxy- β -damascenone, produced by CYP109E1 was identified. Docking calculations using the crystal structure of

CYP109E1 rationalized the experimentally observed regioselective hydroxylation and identified important amino acid residues for statin and terpene binding.

Keywords *Bacillus megaterium* · CYP109E1 · Biotransformation · Pravastatin · Statins · Terpenes

Introduction

Bacillus megaterium is a gram-positive soil bacterium with industrial importance. The ability of growth on different low-priced media, the absence of alkaline proteases and endotoxins, the stable maintenance of plasmid vectors, and high protein expression capacity make this microorganism an excellent biotechnological production host (Rygus and Hillen 1992; Vary et al. 2007). It is used for the production of industrially important enzymes as well as polyhydroxybutyrate and vitamin B12 (Biedendieck et al. 2010; Korneli et al. 2013; Kulprecha et al. 2009; Malten et al. 2005; Stammen et al. 2010). Besides that, the full genome sequencing of two *B. megaterium* strains (Eppinger et al. 2011) gave rise to the identification of several biotechnologically interesting proteins from this bacterium such as cytochromes P450 (P450s) (Abdulmughni et al. 2017; Brill et al. 2014; Jóźwik et al. 2016; Kiss et al. 2015; Milhim et al. 2016).

P450s are heme-containing external monooxygenases. Besides their essential role in steroid hormone biosynthesis and drug metabolism, several more applications of this enzyme class such as bioremediation and implementation of a great variety of chemical reactions are also described (Sono et al. 1996). The broad spectrum of substrates and the ability to perform diverse synthetically challenging reactions under mild conditions provide a high biotechnological potential of these enzymes (Bernhardt 2006; Bernhardt and Urlacher

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00253-017-8552-6>) contains supplementary material, which is available to authorized users.

✉ Rita Bernhardt
ritabern@mx.uni-saarland.de

¹ Institute of Biochemistry, Saarland University,
66123 Saarbruecken, Germany

2014). P450 family members are widely distributed among different species of life. In contrast to mammalian and plant P450s, microbial ones have the advantages that they are soluble proteins and exhibit higher stability simplifying their handling (Fulco 1991; Urlacher et al. 2004). Three P450s from *B. megaterium*, the self-sufficient fatty acid hydroxylase CYP102A1 (BM3) and two steroid hydroxylases of the CYP106 family, CYP106A1 and CYP106A2, have been extensively studied and described as attractive biocatalysts for potential biotechnological applications (Berg and Rafter 1981; Bleif et al. 2011; Brill et al. 2014; Janocha et al. 2016; Kiss et al. 2015; Putkaradze et al. 2017; Narhi and Fulco 1987; Schmitz et al. 2014; Urlacher et al. 2006; Virus et al. 2006).

CYP109E1 is a new member of *Bacillus* P450s, recently identified by our laboratory, but is not fully characterized so far. The crystal structure of CYP109E1 in substrate-free and substrate-bound state has provided first insights into the structural background of substrate binding and activity of this enzyme (Jóźwik et al. 2016). Due to the close phylogenetic distance of CYP109E1 to the steroid hydroxylase CYP106A1, its potential as steroidogenic P450 has been investigated. It has been found that CYP109E1 possesses very low or no catalytic activity towards steroidal substrates except for testosterone and vitamin D3 (Abdulgugni et al. 2017; Jóźwik et al. 2016). To extend the substrate range of CYP109E1, a focused library was screened. The library consisted of steroids, statins, and terpenoids reported to be substrates of some *Bacillus* P450s of the 109 family (Furuya et al. 2009; Girhard et al. 2010) (Scheme 1).

Statins are a class of powerful, widely used drugs against cardiovascular diseases. They effectively reduce plasma LDL-cholesterol levels and coronary heart disease risk by inhibiting the key enzyme 3-hydroxy-3-methylglutaryl-coenzyme-A (HMG-CoA) reductase in the mevalonate pathway (Brown 2007; Endo and Hasumi 1993; Lamon-Fava 2013). The first discovered statin, the natural product compactin (mevastatin) undergoes a stereoselective hydroxylation at the position C6' resulting in the formation of pravastatin (6'β-hydroxycompactin). Pravastatin as well as the naturally occurring statin lovastatin and its semi-synthetic derivative simvastatin are efficient drugs. The bioconversion of compactin to pravastatin is one of the successful biotechnological applications of P450-based systems with only few examples of P450s described in the literature capable of performing this reaction (Bernhardt and Urlacher 2014; Matsuoka et al. 1989; McLean et al. 2015; Milhim et al. 2016; Sakaki 2012).

Terpenes and terpenoids are the most diverse class of chemicals occurring mainly in plants. These compounds and their derivatives are important for biotechnology since they are applied in different fields, such as chemical, pharmaceutical, or flavor and fragrance industry (Janocha et al. 2015). The regio- and stereoselective oxyfunctionalization of

terpenes and terpenoids is of high interest but remains to be a challenging task for synthetic chemistry. Microbial P450 enzymes represent an effective alternative to chemical synthesis since they possess the ability to oxidize diverse biotechnologically important terpene and terpenoid compounds in a highly stereo- and regioselective manner. However, only a limited number of microbial P450s have been identified as terpene hydroxylases and epoxidases until now (Janocha et al. 2015). The well-studied and effective bacterial terpene hydroxylating-P450s include CYP101A1, CYP108A1, and CYP111A1 from *Pseudomonas* species having camphor, terpineol, and linalool as their natural substrates, respectively (Katagiri et al. 1968; Peterson et al. 1992; Ullah et al. 1990), as well as enzymes from *Novosphingobium aromaticivorans* such as CYP101B1 (Bell and Wong 2007; Hall and Bell 2014). Two P450s from *B. megaterium* (CYP106 family) have been shown to hydroxylate di- and triterpenes (Bleif et al. 2011; Brill et al. 2014; Schmitz et al. 2014).

Here, we report highly regioselective oxidations of three statins (compactin **1**, lovastatin **2**, and simvastatin **3**) as well as seven terpenes, including α-ionone **4**, β-ionone **5**, nootkatone **6**, isolongifolen-9-one **7**, α-damascone **8**, β-damascone **9**, and β-damascenone **10** by CYP109E1. Reconstituted P450 and *B. megaterium* whole-cell systems were successfully applied for the biotransformation of these substrates. Characterization of the reaction products indicated hydroxylase and epoxidase activity of CYP109E1, producing an important drug, pravastatin, as well as known and novel statin drug metabolites and terpene derivatives.

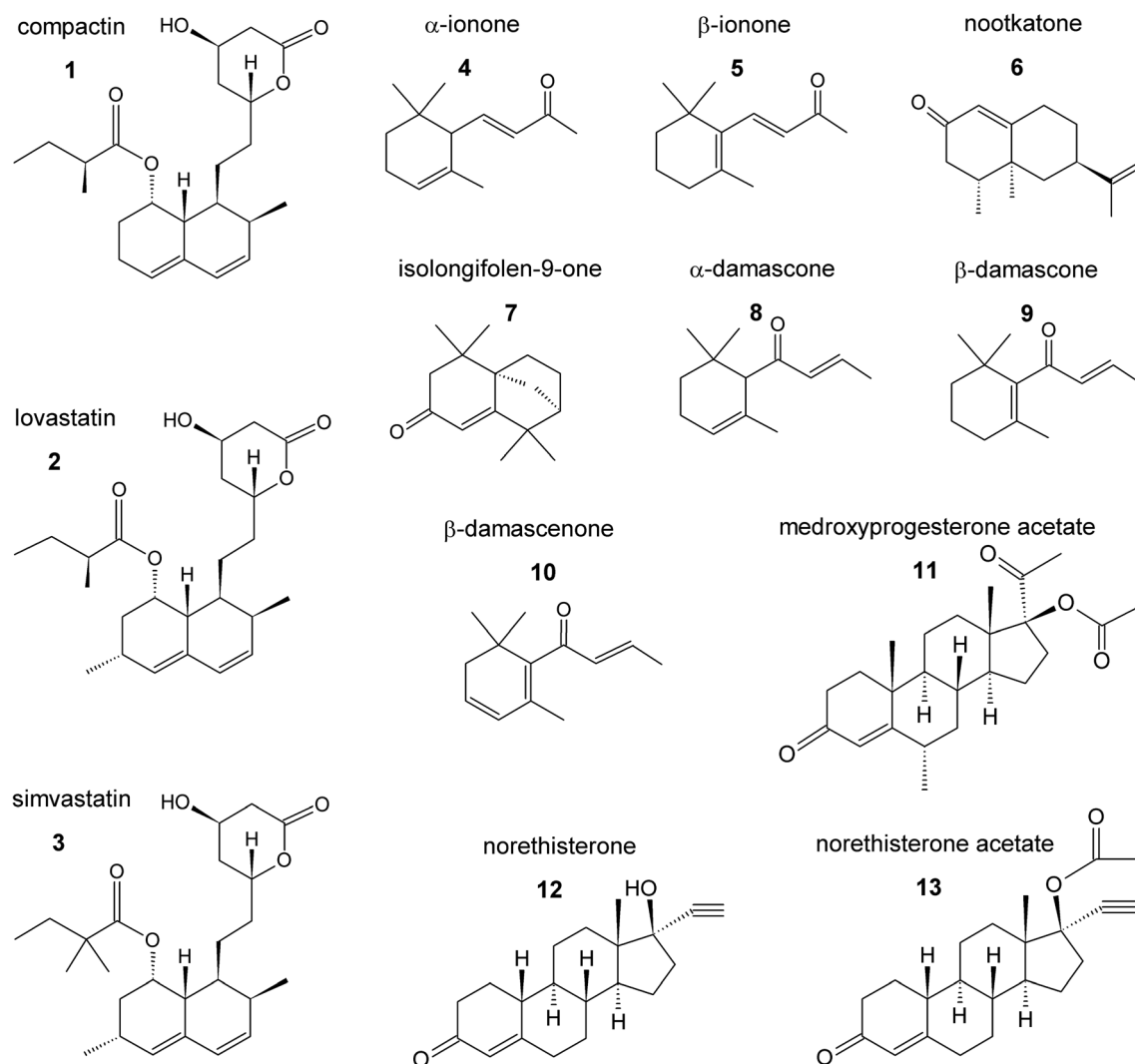
Materials and methods

Reagents and chemicals

Compactin, lovastatin, simvastatin, medroxyprogesterone acetate, norethisterone, and norethisterone acetate were purchased from TCI chemicals (Eschborn, Germany). (+)-Nootkatone, (–)-isolongifolen-9-one, and α- and β-ionone were purchased from Sigma-Aldrich (St. Louis, MI, USA). α-Damascone, β-damascone, and β-damascenone were provided from Bell Flavors & Fragrances (Leipzig, Germany). Bacterial culture media were purchased from Becton, Dickinson and Company (Franklin Lakes, NJ, USA). All other chemicals and solvents were obtained from standard sources and were of the highest purity available.

Strains and plasmids

The *E. coli* C43 (DE3) cells (Lucigen, Middleton, WI, USA) transformed with the plasmid pET17b.CYP109E1 (Abdulgugni et al. 2017) were used for the expression of CYP109E1. *B. megaterium* MS941 cells (Wittchen and



Scheme 1 Chemical structures of compounds investigated with CYP109E1

Meinhardt 1995) transformed either with the plasmid pSMF2.1.CYP109E1 (Abdulmughni et al. 2017) or the control plasmid pSMF2.1 (Bleif et al. 2011) were used for whole-cell conversions.

Enzyme purification and spectral measurements

E. coli cells were cultured overnight in Luria-Bertani (LB) medium containing 100 µg/mL ampicillin at 37 °C and 150 rpm in a rotary shaker. Five milliliters of overnight culture was used to inoculate 500 mL Terrific Broth (TB) medium containing 100 µg/mL ampicillin in a 2-L baffled shake flask, and the culture was incubated at 37 °C and 100 rpm. At OD₆₀₀ of 0.5, the expression of CYP109E1 was induced with 1 mM of isopropyl-thio-β-D-galactopyranoside (IPTG). Simultaneously, 0.5 mM of δ-aminolevulinic acid (δ-ALA) was added to support the heme synthesis. After a 24-h expression, the cells were harvested by centrifugation at 4500×g and stored at −20 °C until purification of the enzyme. The purification of CYP109E1

was carried out using immobilized metal ion affinity chromatography (TALON™ Resin, Clontech) as described elsewhere (Abdulmughni et al. 2017). Fractions containing CYP109E1 were collected, concentrated, and further purified using size exclusion chromatography (Superdex 75 column, GE Healthcare Life Sciences) in 50 mM potassium phosphate buffer (pH 7.4) with a flow rate of 0.1 mL/min. The purified protein fractions were collected, concentrated by ultrafiltration, and stored at −20 °C. The concentration of purified CYP109E1 was estimated by carbon monoxide difference spectroscopy using an extinction coefficient (450–490 nm) of 91 mM^{−1} cm^{−1} according to the method of Omura and Sato (1964).

Bovine adrenodoxin reductase (AdR) and a truncated form of bovine adrenodoxin (Adx_{4–108}) were expressed and purified as described elsewhere (Sagara et al. 1993; Uhlmann et al. 1992). Protein concentrations were measured spectroscopically using corresponding molar extinction coefficients as described elsewhere (Lisurek et al. 2004).

Spin-state shift and estimation of dissociation constant (K_d)

Spin-state shifts were investigated using a double beam spectrophotometer (UV-2101PC, Shimadzu, Japan) and two tandem quartz cuvettes. One chamber of each cuvette contained 10 μ M CYP109E1 in 50 mM potassium phosphate buffer (pH 7.4), while the other chamber was filled with buffer alone. The substrates were dissolved in DMSO (5 mM stock solutions) and were added into the chamber with CYP109E1 solution of the sample cuvette while an equal amount of each substrate was also added into the buffer containing a chamber of the reference cuvette. Difference spectra were recorded between 350 and 500 nm. Dissociation constant (K_d) was estimated by titrating each substrate (0–200 μ M) until saturation. The data were analyzed by plotting the peak-to-trough differences against the substrate concentrations and fitting them (except for lovastatin and simvastatin) with the hyperbolic function $\Delta A = \Delta A_{\max} \times [S] / (K_d + [S])$. Lovastatin and simvastatin exhibited tight binding ($K_d < 5[E]$), and for these compounds, the data were fitted to the tight binding quadratic equation: $\Delta A = (\Delta A_{\max}/2[E]) \times \{(K_d + [E] + [S]) - \{(K_d + [E] + [S])^2 - 4[E][S]\}^{1/2}\}$ (Williams and Morrison 1979). ΔA represents the peak-to-trough absorbance difference, ΔA_{\max} is the maximum absorbance difference, $[E]$ is the enzyme concentration, and $[S]$ is the substrate concentration. The data processing was done with OriginPro 9.0G software (OriginLab, MA, USA). All K_d values represent the mean of three independent measurements with the coefficient of determination (R^2) of 0.99. Spin-state shifts were calculated (to approximately $\pm 5\%$) using the ΔA_{\max} value of each substrate as a percentage of the maximum expected shift for 10 μ M enzyme estimated as described by Luthra et al. (2011).

Measurement of CYP109E1 activity in vitro

The catalytic activity of CYP109E1 towards the selected compounds was investigated using a reconstituted in vitro system consisting of 0.5 or 1 μ M CYP109E1, bovine AdR, and Adx_{4–108} with a molar ratio of 1:2:20. The reactions were carried out in 250 μ L of 50 mM potassium phosphate buffer (pH 7.4) with 10% glycerol. For the sufficient electron supply, a NADPH regeneration system containing glucose-6-phosphate-dehydrogenase (1 U), glucose-6-phosphate (5 mM), and MgCl₂ (1 mM) was used. The in vitro reactions with 100 μ M statin **1–3** (20 mM stock solution in DMSO), 200 μ M steroidal **11–13** (20 mM stock solution in DMSO), and terpene **4–10** (50 mM stock solution in ethanol) compounds were started by addition of 1 mM NADPH in 1.5-mL Eppendorf tubes under mixing at 30 °C and 700 rpm and stopped after 15 min by addition of 250 μ L of ethyl acetate or chloroform. The reaction mixtures were extracted twice with 500 μ L of the organic solvent. The organic phases

were combined, evaporated to dryness, and stored at -20 °C until analysis via either high-performance liquid chromatography (HPLC) or gas chromatography-mass spectrometry (GC-MS).

Whole-cell conversion with CYP109E1 in *B. megaterium* MS941

The whole-cell conversions were performed in plasmidless *B. megaterium* MS941 strain transformed with pSMF2.1.CYP109E1 shuttle vector by the method of Barg et al. (2005). As a control of the CYP109E1-based whole-cell biotransformation, MS941 cells transformed with the empty vector pSMF2.1 were used. The main cultures were prepared by inoculating 50 mL of complex medium (24 g/L yeast extract, 12 g/L soytone, 0.5% glycerol (v/v), 2.31 g/L KH₂PO₄ and 12.5 g/L K₂HPO₄) with overnight culture (1% of the main culture volume) in 300 mL baffled shake flask and incubated at 37 °C and 160 rpm in a rotary shaker until an OD₅₇₈ of 0.5 was reached. At this time point, the expression of CYP109E1 was induced by adding 5 g/L xylose. After 24 h expression at 30 °C, the cultures were harvested by centrifugation (4500 \times g, 4 °C), washed, and resuspended in 50 mM potassium phosphate buffer (pH 7.4) with 2% glycerol. Whole-cell conversions of 100 μ M of statins **1–3** and 200 μ M of terpenes **4–10** were carried out in 25 mL volume in a rotary shaker at 30 °C and 150 rpm. The conversions of nootkatone **6**, isolongifolen-9-one **7**, α -ionone **4**, β -ionone **5**, α -damascone **8**, β -damascone **9**, and β -damascenone **10** were performed in sealed baffled shake flasks. Substrate stock solutions were prepared by dissolving them in DMSO or ethanol. The biotransformation of each substrate was monitored within 2 h by taking 500- μ L culture samples. The samples were extracted with double volume of ethyl acetate, dried, and stored at -20 °C until HPLC analysis. To obtain sufficient amounts (2–10 mg) of conversion products for the structure elucidation via NMR, the whole-cell reactions were scaled up to 0.5–1 L culture volume. The reactions were stopped after 2 or 4 h and extracted with double volume of ethyl acetate. The organic phases were dried over anhydrous MgSO₄, concentrated to dryness using a rotivapor (Büchi R-114), and stored at -20 °C until purification via HPLC.

Conversion analysis and product isolation via HPLC

The conversion analysis was performed via reversed phase HPLC technique using a Jasco system (a Pu-980 HPLC pump, an AS-950 sampler, an UV-975 UV/Vis detector, a LG-980–02 gradient unit; Jasco, Gross-Umstadt, Germany) and an ec MN Nucleodur C18 (5 μ m, 4.0 \times 125 mm) column (Macherey-Nagel, Bethlehem, PA, USA). For the purification of conversion products, a preparative ec MN Nucleodur C18

VP (5 μm , 8.0×250 mm) column (Macherey-Nagel, Bethlehem, PA, USA) was used. The mobile phase consisted of 10% acetonitrile in water (solvent A) and pure acetonitrile (solvent B). A gradient from 20 to 80% of solvent B was used for the separation of α -ionone **4**, β -ionone **5**, α -damascone **8**, β -damascone **9**, β -damascenone **10**, compactin **1**, and lovastatin **2** conversions, from 20 to 100% for the analysis of nootkatone **6** and isolongifolen-9-one **7** conversions and from 0 to 100% for simvastatin **3**, medroxyprogesterone acetate **11**, norethisterone **12**, and norethisterone acetate **13** conversions. A flow rate of either 1 mL/min (conversion measurements) or 3.5 mL/min (product isolation) and a temperature of 40 °C were set during the analysis of all compounds. The UV detection was accomplished at 236 (statins **1–3**), 240 (steroids **11–13**, nootkatone **6**, and isolongifolen-9-one **7**), 228 (α -ionone **4**, α -damascone **8**, and β -damascone **9**), 232 (β -damascenone **10**), or 296 nm (β -ionone **5**).

Conversion analysis via GC-MS and LC-MS

GC-MS measurements were performed with a system consisting of an AI/AS 3000 autosampler, a DSQ II 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 conversions were analyzed as described elsewhere (Litzenburger and Bernhardt 2016).

Analytical HPLC-MS was performed on a Thermo Dionex UltiMate 3000 HPLC using NUCLEOSHELL RP 18plus (100 \times 4.6 mm, 2.7 μm) column (Macherey-Nagel, Bethlehem, PA, USA) coupled to a Bruker AmaZon SL ESI-MS system (Bruker, Billerica, MA, USA). The mobile phase consisted of water with 0.1% formic acid (solvent A) and pure acetonitrile with 0.1% formic acid (solvent B). A gradient from 5 to 95% of solvent B was used for the separation. Mass spectra were acquired in positive mode ranging from 100 to 600 m/z using UltraScan mode.

Structure elucidation by NMR spectroscopy

NMR spectra were recorded on a Bruker (Rheinstetten, Germany) DRX 500 NMR spectrometer. A combination of ^1H , ^{13}C , ^1H , ^1H -COSY, HSQC, and HMBC experiments was used for structure elucidation. All chemical shifts are relative to CHCl_3 ($\delta = 7.24$) or CDCl_3 ($\delta = 77.00$) using the standard δ notion in parts per million (ppm).

Molecular docking

For the docking experiments, the recently solved crystal structure of CYP109E1 was used (Jóźwik et al. 2016). Two structurally different substrates, the statin compactin **1** and the norisoprenoid compound α -ionone **4**, were docked into the

active site of CYP109E1 in its closed conformation (PDB: 5L94) 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 ligands (Sanner 1999). A partial charge of + 0.400e was assigned manually to the heme iron, which corresponds to Fe(II) that was compensated by adjusting the partial charges of the ligating nitrogen atoms to $-0.348e$. While the protein was kept as rigid, the flexible bonds of the ligands were assigned automatically and verified by manual inspection. A cubic grid box with a size of $52 \text{ \AA} \times 50 \text{ \AA} \times 48 \text{ \AA}$ was fixed above the heme moiety to cover the whole active site of the enzyme. The spacing between grid points was 0.375 \AA . Two hundred docking runs (for each substrate) were carried out applying the Lamarckian genetic algorithm using default parameter settings. The binding conformations were analyzed according to estimated binding energies and distances of the target carbon atom to the heme iron.

Results

Substrate binding to CYP109E1

To screen CYP109E1 for potential new substrates, a library of 13 compounds was used. The library contained the steroids norethisterone **12** and medroxyprogesterone acetate **11**, compactin **1**, and the terpenes α -ionone **4**, β -ionone **5**, and nootkatone **6**, which previously have been identified as substrates of the CYP109 family (Furuya et al. 2009; Girhard et al. 2010), as well as related compounds with biotechnological or pharmaceutical impact (Scheme 1). First, the potential substrates were investigated for their ability to bind to CYP109E1. This was done by monitoring the high-spin shift induction using difference spectroscopy. When a P450 substrate displaces the axial water ligand, it changes the spin-state of the iron atom from low-spin to high-spin (Schenkman and Jansson 1998). Spectroscopically, this transition is reflected by a shift in the absorption maximum from around 420 to 390 nm (type I difference spectrum). Previously, it has been reported that the spin-shift is not necessary for the catalytic activity of a P450 (Girhard et al. 2010; Simgen et al. 2000). On the other hand, a very large type I spin-shift observed for many bacterial P450s (Bell and Wong 2007) indicating the ability of its substrates to displace almost all water molecules in the active site results in an effective electron transfer and high catalytic activities by substrate-gating mechanism (Sligar and Gunsalus 1976; Fisher and Sligar 1985; Honeychurch et al. 1999; Poulos and Raag 1992). Therefore, this spectroscopic method remains to be well suitable for screening of potential P450 substrates as well as inhibitors (Khatri et al. 2016; Schmitz et al. 2012). All here tested compounds, except for steroids **11–13**, were able to induce a type I spin-shift of

CYP109E1, which allowed us to investigate the substrate binding spectroscopically (Fig. 1 and S1) and to determine dissociation constants (K_d). Statin compounds **1–3** showed tighter binding to P450 than terpene substrates **4–10** (Table 1). Our results indicate that simvastatin **3** binds tighter to CYP109E1 than the other statins. The K_d of simvastatin **3** [$K_d = 4.7 \mu\text{M}$] was shown to be approximately 18 times lower than that of compactin **1** [$K_d = 84 \mu\text{M}$] and two times lower than that of lovastatin **2** [$K_d = 9.6 \mu\text{M}$] (Fig. 1 and S1). The strongest binding to CYP109E1 among terpenes **4–10** was observed for β -ionone **5** [$K_d = 91 \mu\text{M}$] and the weakest for isolongifolen-9-one **7** [$K_d = 216 \mu\text{M}$] (Fig. S1). Interestingly, α -ionone **4** (51%) and simvastatin **3** (43%) were able to induce the highest spin-state shifts among the investigated substrates, whereas lovastatin **2** shifted the state to a lesser extent (11%).

Biotransformation of medroxyprogesterone acetate, norethisterone, and norethisterone acetate

Although the tested steroids **11–13** did not show any spectral shift when incubated with CYP109E1, their potential biotransformation was investigated in vitro. As observed previously with deoxycorticosterone and testosterone as substrates of CYP106A2 and CYP109B1, respectively, a type I spectral shift is not a necessary prerequisite for bioconversion (Girhard et al. 2010; Simgen et al. 2000). The steroidal substrates medroxyprogesterone acetate **11**, norethisterone **12**, and norethisterone acetate **13** were thus investigated using a reconstituted P450 system with AdR and Adx_{4–108} as redox partners. CYP109E1 did not show any activity towards these compounds (data not shown), and they were, therefore, not further tested with the CYP109E1 based whole-cell system in vivo.

Biotransformation of statin substrates

The three statins, compactin **1**, lovastatin **2**, and simvastatin **3**, identified as CYP109E1 ligands by difference spectroscopy, were further tested as potential substrates for CYP109E1 in vitro. The activity of CYP109E1 was reconstituted with bovine AdR and truncated Adx_{4–108} as described in the “Material and methods” section. CYP109E1 was found to convert all three statin compounds **1–3** resulting in one major reaction product **14**, **15** for compactin **1** and lovastatin **2**, respectively (Fig. 2a, c) and two products **16**, **17** for simvastatin **3** conversion (Fig. 2e). The selectivity of compactin and lovastatin conversions (77 and 94%) was found to be remarkably higher compared to that of simvastatin (Table S1). Interestingly, the conversion ratio of CYP109E1 towards compactin **1** was much lower compared with the other statins resulting in only 11% conversion after 15 min compared with lovastatin **2** (63%) and simvastatin **3** (72%) (Fig. 2a, c, e and S2).

To prepare and isolate higher amounts of compactin **1**, lovastatin **2**, and simvastatin **3** conversion products for the analysis via NMR spectroscopy, *B. megaterium* MS941 whole cells transformed with pSMF2.1.CYP109E1 were used. After in vivo conversion of statins, the main products of the compactin **1** and lovastatin **2** conversion (**14** and **15**, Fig. 2b, d) as well as two products of simvastatin **3** conversion (**16** and **17**, Fig. 2f) have been purified in sufficient amounts. NMR analysis (data S7) revealed that CYP109E1 is able to hydroxylate compactin **1** and lovastatin **2** at position C6 β resulting in the formation of pravastatin **14** and 6 β -hydroxy-lovastatin **15**, respectively (Scheme 2). The simvastatin **3** metabolites were identified as 3 α -hydroxy-simvastatin **16** and 4 α -hydroxy-simvastatin **17**. The *B. megaterium* MS941 strain containing the pSMF2.1 vector, which was used as a negative

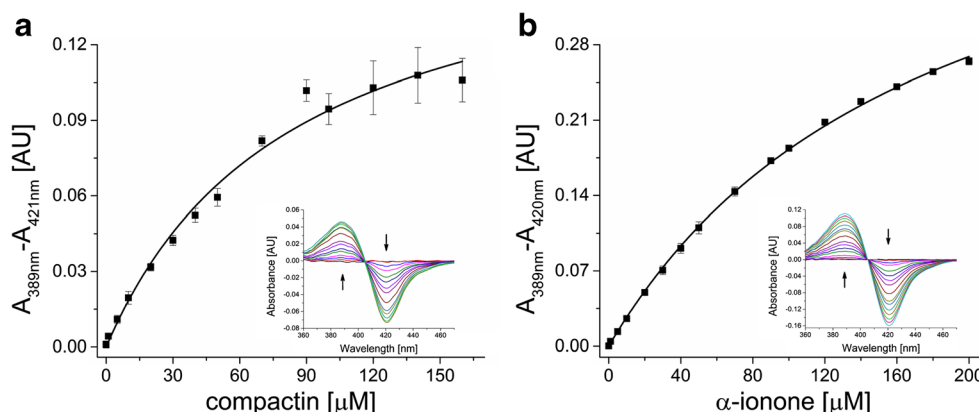


Fig. 1 The absorbance changes plotted against the corresponding concentrations of compactin **1** (a) and α -ionone **4** (b) titrated to CYP109E1 as described in the “Material and methods” section. The mean values were fitted by hyperbolic regression, and the K_d value was

calculated. The inset shows type I spectral shifts induced by the binding of the increasing amount of the substrate to CYP109E1. Arrows show the direction of spectral changes at increasing substrate concentrations

Table 1 Binding of different compounds to CYP109E1

Compound	Spin-state shift [%] ^a	K_d [μ M] ^b
Compactin	17	84 \pm 13
Lovastatin	11	9.6 \pm 3.5
Simvastatin	43	4.7 \pm 1.8
α -Ionone	51	178 \pm 6
β -Ionone	23	91 \pm 9
Nootkatone	22	129 \pm 8
Isolongifolen-9-one	24	216 \pm 18
α -Damascone	27	161 \pm 8
β -Damascone	20	167 \pm 7
β -Damascenone	17	154 \pm 7
Medroxyprogesterone acetate	–	–
Norethisterone	–	–
Norethisterone acetate	–	–

K_d dissociation constant, – no spin-state shift observed

^a Estimated to approximately \pm 5%

^b Estimated with a coefficient of determination (R^2) \geq 0.99

control for the CYP109E1-dependent conversions, showed no activity towards the investigated substrates (data not shown).

Biotransformation of terpene substrates

All tested terpenes **4–10** showed type I binding spectra, and the reactions of CYP109E1 with these compounds were further characterized. Similar to statins **1–3**, the activity of CYP109E1 towards the potential terpene substrates α -ionone **4**, β -ionone **5**, nootkatone **6**, isolongifolen-9-one **7**, α -damascone **8**, β -damascone **9**, and β -damascenone **10** was initially tested using the reconstituted system. CYP109E1 showed high conversion within 15 min under in vitro conditions (95% of α -ionone **4**, 96% of β -ionone **5**, 70% of nootkatone **6**, 91% of isolongifolen-9-one **7**, 72% of α -damascone **8**, 70% of β -damascone **9** and β -damascenone **10**) (Fig. S2). The reaction selectivity except for β -damascenone conversion was high (77–93%) yielding one main product (Table S1). All these substrates were also successfully converted by *B. megaterium* cells overexpressing CYP109E1, and the product patterns and selectivities were similar compared to those in vitro. The main products of α -ionone **4**, β -ionone **5**, α -damascone **8**, and β -damascone **9** were compared with authentic standards by GC-MS (Litzenburger and Bernhardt 2016) and identified as 3-hydroxy- α -ionone **18**, 4-hydroxy- β -ionone **19**, 3-hydroxy- α -damascone **23**, and 4-hydroxy- β -damascone **24**, respectively (Fig. 3a, b, e, f and S3–S6). Moreover, 3-hydroxy- α -ionone **18** and 3-hydroxy- α -damascone **23** were found to be 3,6-trans-products. The conversion products of nootkatone **6**, isolongifolen-9-one **7**, and β -damascenone **10** were isolated via HPLC and elucidated by NMR

spectroscopy. The products of nootkatone **6** and isolongifolen-9-one **7** conversion were identified as 11(*R*),12-epoxy-nootkatone **20**, 11(*S*),12-epoxy-nootkatone **21**, and 4(*R*)-hydroxy-isolongifolen-9-one **22**, respectively (Fig. 3c, d). Products of in vivo conversion of β -damascenone **10** were found to be 3,4-dihydroxy- β -damascone **27**, 2-hydroxy- β -damascenone **25**, and 3,4-epoxy- β -damascone **26** (Fig. 4a). The control strain *B. megaterium* MS941 containing only the pSMF2.1 vector showed low conversion of the investigated terpene substrates **4–10** after 2 h (data not shown), most probably due to the activity of CYP109E1 encoded in the bacterial genome.

Interestingly, the formation of 3,4-dihydroxy- β -damascone **27** was observed in vivo (Fig. 4a), whereas no corresponding peak (Fig. 4a, t_R = 2.8 min) was detectable in vitro using the reconstituted CYP109E1-based system (data not shown). For further characterization, the in vivo conversion was monitored over time. The HPLC results showed significant changes in product distribution within 4 h. After 1 h of conversion, 3,4-epoxy- β -damascone **26** and 3,4-dihydroxy- β -damascone **27** were found to be the major and minor products with 34 and 8% of total shares, respectively. The amount of 3,4-dihydroxy- β -damascone **27** increased up to 45% after 4 h of conversion, and the increase was correlated with a decrease of 3,4-epoxy- β -damascone **26** (Fig. 4b). Thus, the time-dependent data suggested that the double hydroxylated product is formed from 3,4-epoxy- β -damascone **26** by a CYP109E1-independent reaction.

Molecular docking

The crystal structure of CYP109E1 in its closed form was used to investigate the enzyme-substrate interaction and observed regioselectivity of hydroxylation for two of the novel substrates, compactin **1** and α -ionone **4**, by docking of these structurally different substances into the active site of CYP109E1. Both substrates appeared in docking positions allowing hydroxylation at experimentally identified positions, C6' (distance \sim 4.4 Å) for compactin **1** and C3 (distance \sim 4 Å) for α -ionone **4**, respectively. Eight amino acid residues including Ile85, Leu238, Ile241, Ala242, Thr246, Ala291, Leu292, and His293 were found to form the binding pocket of CYP109E1 with compactin **1** whereas only six (Arg69, Ile85, Ala242, Val289, Leu292, His293) were predicted to interact with the smaller substrate α -ionone **4** (Fig. 5). The results predicted that both substrates were bound by van der Waals forces and hydrophobic interactions. Active site residues with predicted hydrophobic interactions with both substrates are Ile85 (BC-loop, SRS1), Ala242 (I-helix, SRS4), and Leu292 (K- β 5-loop, SRS5). Considering the electrostatic interactions with the substrates, only His293 was predicted to form a hydrogen bond with compactin **1** and α -ionone **4** (Fig. 5).

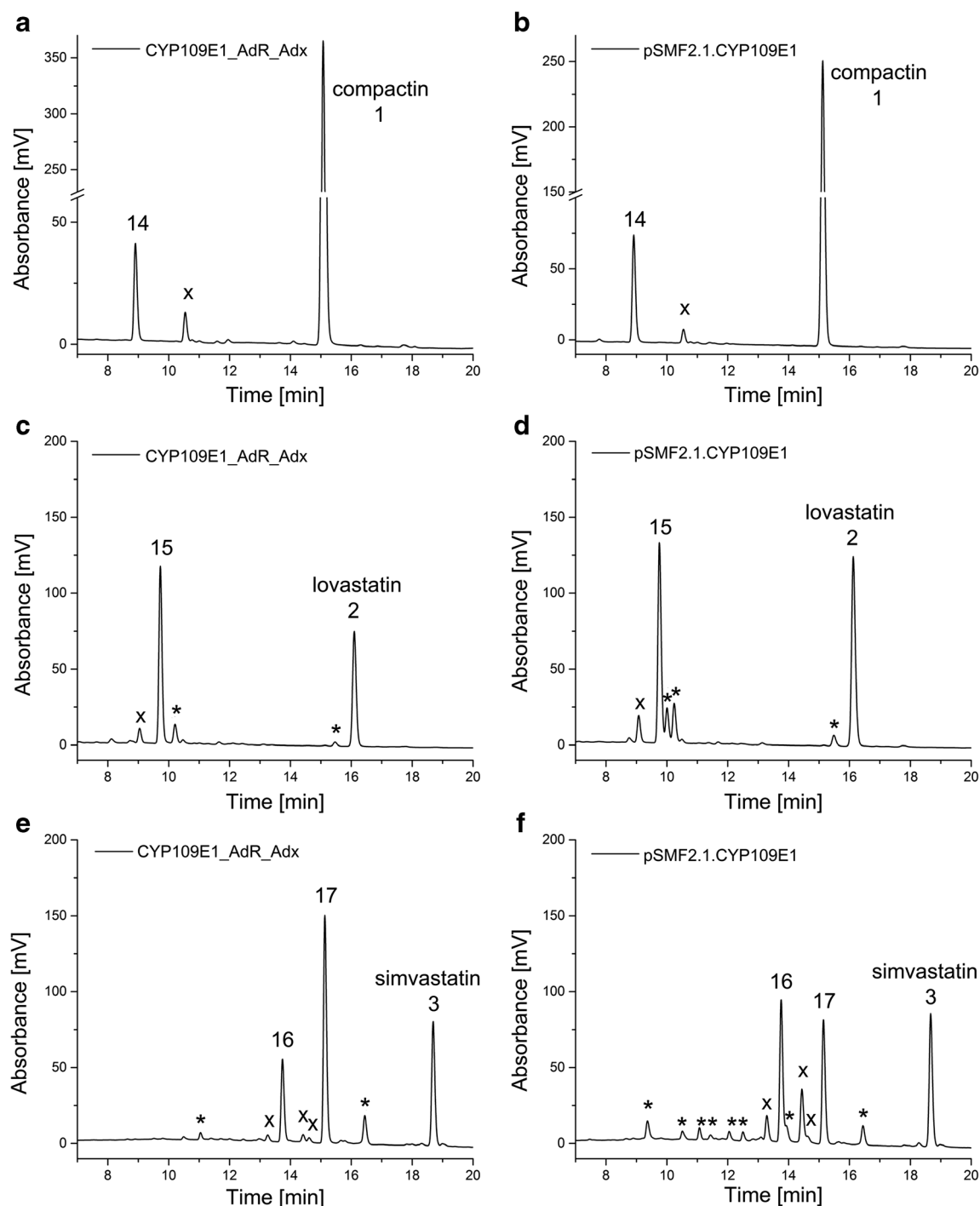


Fig. 2 HPLC chromatograms of in vitro (a, c, e) and in vivo (b, d, f) conversion of statins 1–3. 14, 15, 16, and 17 represent products of compactin 1, lovastatin 2, and simvastatin 3 conversion by CYP109E1,

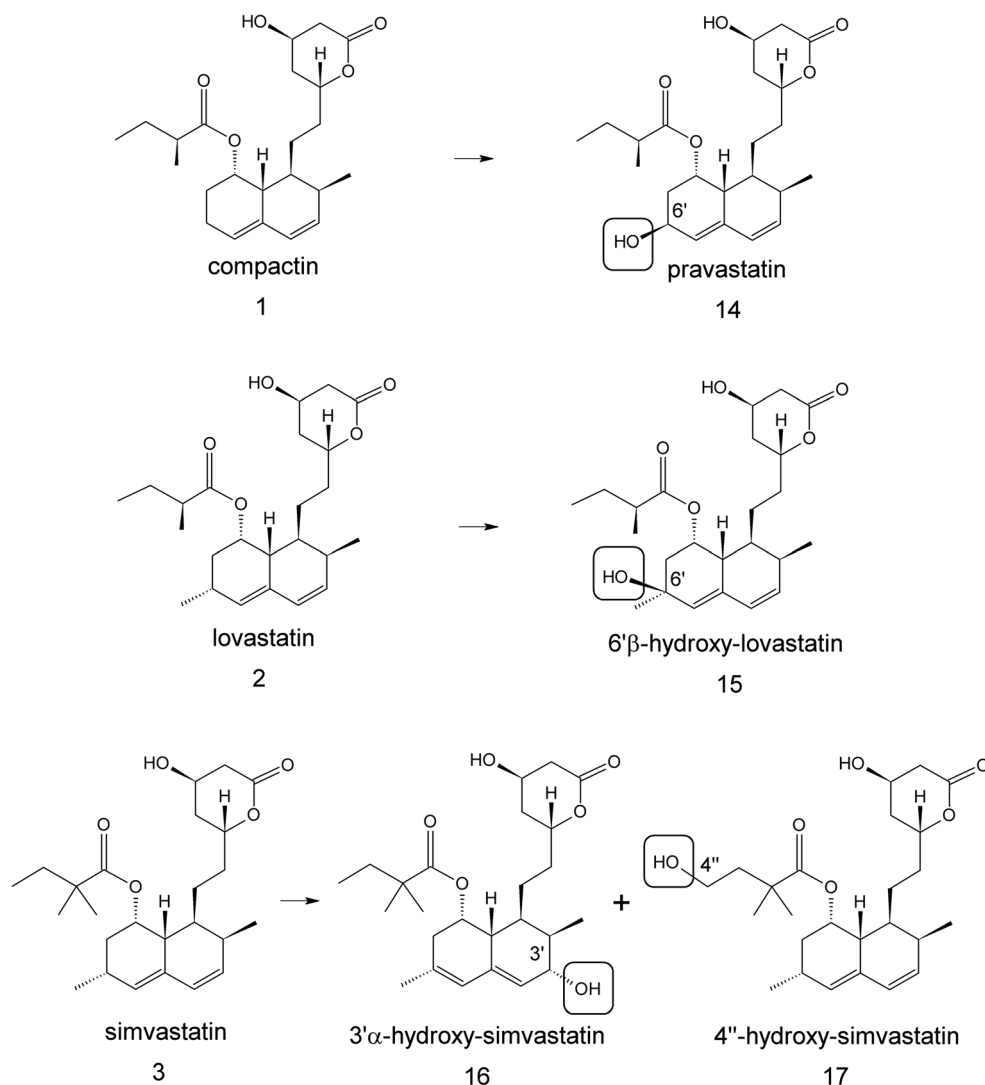
respectively, which have been isolated and characterized. The minor products are labeled with “x” (masses are provided in Table S1) and impurities with asterisk

Discussion

Various P450s from *B. megaterium* attract interest due to their ability to perform the biotransformation of a broad spectrum of highly interesting compounds such as steroids (Berg and Rafter 1981; Kiss et al. 2015; Putkaradze et al. 2017; Rauschenbach et al. 1993), di- and triterpenes (Bleif et al. 2011; Brill et al. 2014;

Schmitz et al. 2014), as well as sesquiterpenes (Sowden et al. 2005), fatty acids, amides, and alcohols (Miura and Fulco 1975). CYP109E1 is a newly identified member of this group, and therefore, elucidation of its substrate range is of great interest. During the initial study on CYP109E1, the crystal structure was solved revealing a highly dynamic active site (Jóźwik et al. 2016). Phylogenetically, CYP109E1 was found to be related to

Scheme 2 Chemical structures of the statin conversion products 14–17 identified via NMR



the steroid hydroxylase CYP106A1 and, therefore, a focused library of important steroids was tested as its potential substrates. The study revealed that among the 13 tested steroidal compounds, only testosterone was converted by CYP109E1 resulting in the corresponding 16β-hydroxy product (Jóźwik et al. 2016). Another important compound which was recently identified as substrate of CYP109E1 is the secosteroid vitamin D3 (Abdulgħni et al. 2017). The enzyme showed 25- and 24-hydroxylase activities towards vitamin D3 generating the valuable compound 25-hydroxy vitamin D3 and two new metabolites hydroxylated at position C24(*S*) and C25 (Abdulgħni et al. 2017).

In order to further characterize CYP109E1, we aimed to extend the substrate spectrum by investigating its activity towards a focused library of biotechnologically valuable compounds. The enzyme showed no activity for the three tested steroidal compounds 11–13, whereas several statins 1–3 and terpenes 4–10 were successfully converted (Scheme 1). Compactin 1, lovastatin 2, and simvastatin 3, identified as novel substrates of CYP109E1

during our study, belong to the group of statins, effective pharmaceutical agents widely used against lipid disorders. In several studies, a broad range of pleiotropic effects of statins was also described proposing their application for the treatment of inflammatory and neurological diseases as well as tumors (Gazzerro et al. 2012). All three compounds were able to shift the heme iron of CYP109E1 into the high-spin state. Dissociation constants determined by substrate titrations showed strongest binding for simvastatin 3 to CYP109E1, with a K_d of $4.7 \pm 1.8 \mu\text{M}$ compared to $9.6 \pm 3.5 \mu\text{M}$ and $84 \pm 13 \mu\text{M}$ for lovastatin 2 and compactin 1, respectively. Investigated statin compounds 1–3 were successfully converted by CYP109E1 into one main product, and the activity of the P450 towards lovastatin 2 and simvastatin 3 was higher than that towards compactin 1 using the heterologous redox partners AdR and Adx_{4–108} (Fig. 2a, c, e). The CYP109E1-based reactions were further investigated in vivo in a mutant of *B. megaterium* DSM319, MS941, extensively investigated by our group to generate pharmaceutically important metabolites (Abdulgħni et al. 2017; Kiss et al.

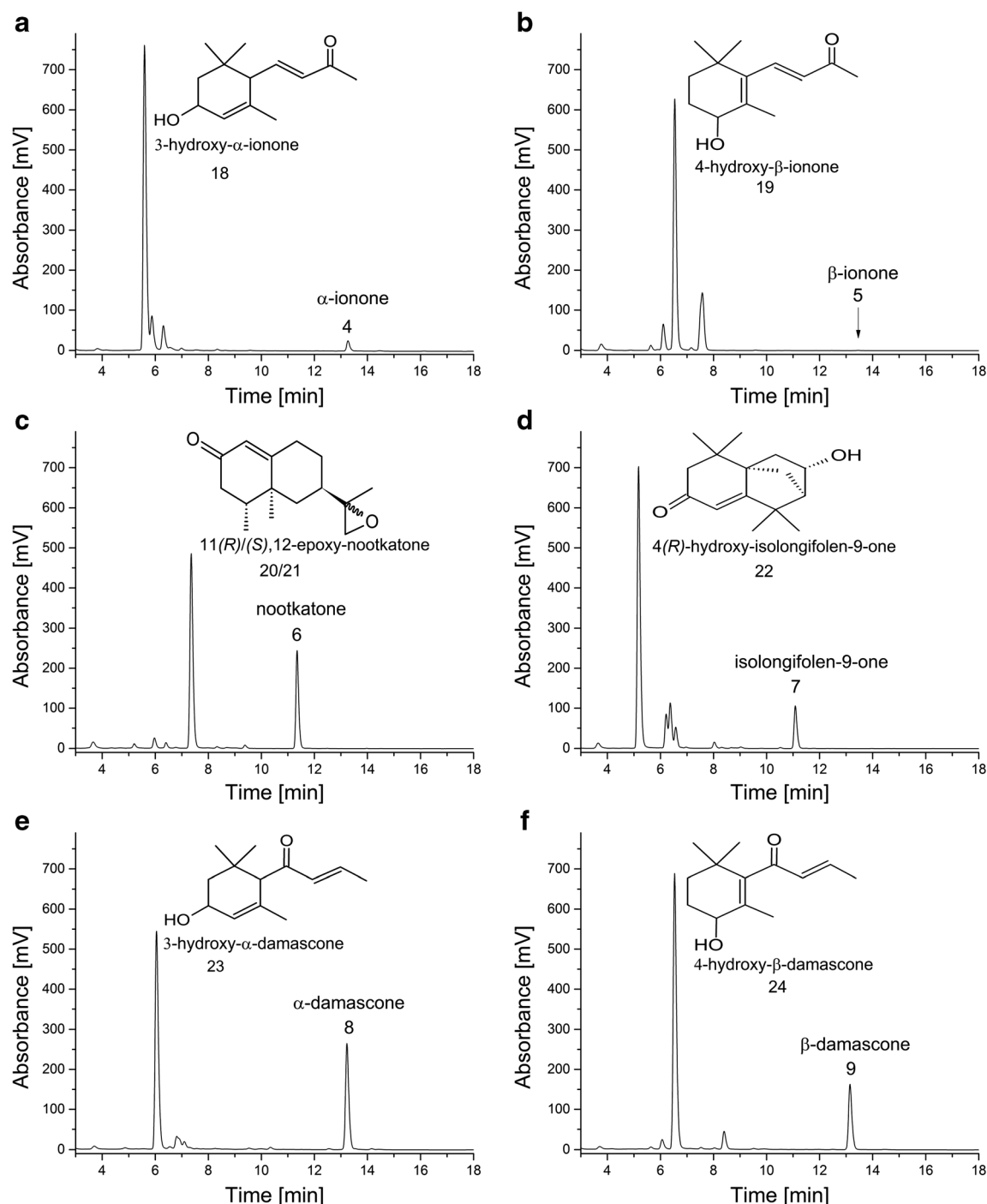


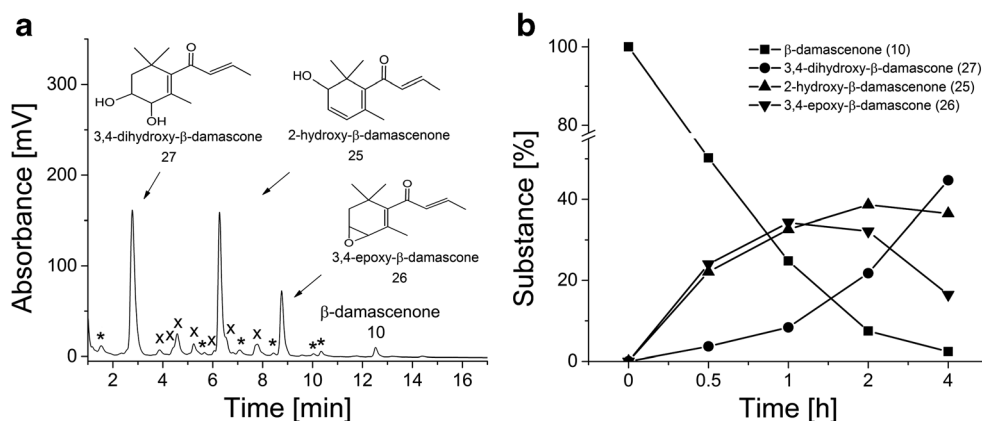
Fig. 3 HPLC chromatograms of in vivo conversions of α -ionone **4** (a), β -ionone **5** (b), nootkatone **6** (c), isolongifolen-9-one **7** (d), α -damascone **8** (e), and β -damascone **9** (f) with the identified products shown close to

the peaks. Masses of the minor products detected in vitro as well as in vivo are provided in Table S1

2015; Putkaradze et al. 2017). The observed product patterns of statins in vitro and in vivo (Fig. 2) were very similar and allowed us to scale up the whole-cell reactions to isolate sufficient amounts of the main products for structure identification via NMR spectroscopy. The data revealed that CYP109E1 hydroxylates compactin **1** with a high stereoselectivity at the C6' position forming the biologically active form of the widely used

pharmaceutical pravastatin **14** and this way characterizing CYP109E1 as a novel pravastatin synthase. To the best of our knowledge, so far only two other non-mutated P450 monooxygenases are known being able to convert compactin **1** to active pravastatin variant. The CYP109E1-based whole-cell system produced up to 14 mg/L pravastatin **14** after 4 h of bio-transformation which is comparable to the previously described

Fig. 4 HPLC chromatogram of β -damascenone **10** conversion by the *B. megaterium* MS941 cells overexpressing CYP109E1 (a) and time-dependent in vivo conversion within 4 h (b). The minor products are labeled with “x” (masses are provided in Table S1) and impurities with asterisk

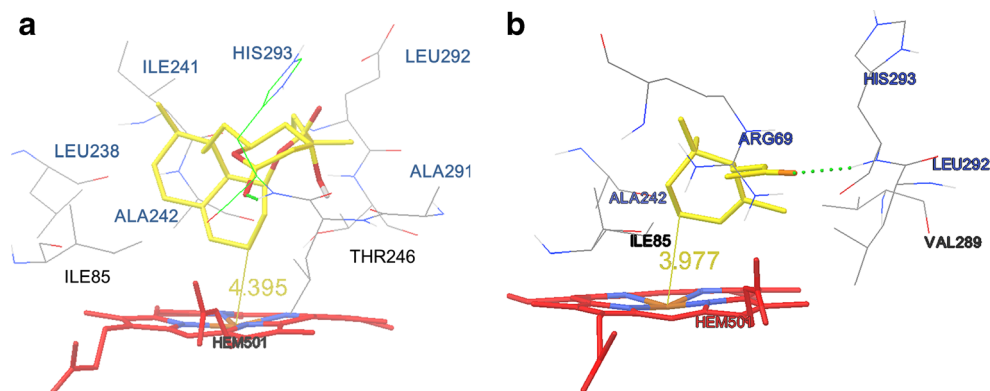


E. coli systems with P450sca-2 (12.9 mg/L within 21 h) (Ba et al. 2013) and CYP107DY1 (13.2 mg/L within 20 h) (Milhim et al. 2016) but provides much higher space-time yield. The activity of this industrially relevant P450 system for pravastatin **14** production might be further optimized by identification of a more effective redox chain, by more suitable production hosts, and by a rational or semi-rational design of the enzyme (Ba et al. 2013; McLean et al. 2015; Milhim et al. 2016). Moreover, to overcome known limitations of cytochrome P450 whole-cell systems on industrial scale, such as substrate solubility, toxicity, and uptake from the cells, expression of CYP109E1 by compactin-producing *Penicillium chrysogenum* similar to the expression of P450_{Prava} (McLean et al. 2015) might be promising. Besides compactin **1**, lovastatin **2** was also converted into the C6' β -hydroxylated product by CYP109E1 while simvastatin **3** was found to be hydroxylated at the C3' α and C4'' positions (Scheme 2). 6' β -Hydroxy-lovastatin **15** and 3' α -hydroxy-simvastatin **16** are human liver metabolites of the statin drugs formed by P450 enzymes, whereas 4''-hydroxy-simvastatin **17** is a synthetic drug derivative. None of these compounds are commercially available although they are highly required in larger quantities for activity and safety studies, particularly due to the reported improved pharmacokinetic properties of some statin metabolites (Kandel et al. 2014) and the known drug-drug interactions of statin pharmaceuticals (Kellick et al. 2014). There are few reports describing the generation of these metabolites, such as chemical

synthesis of 3' α -hydroxy-simvastatin **16**, electrochemical oxidation, and biotransformation of lovastatin **2** and simvastatin **3** with CYP3A4 or CYP102A1 mutants (Khera and Hu 2013; Kim et al. 2011; Stokker 1994). However, CYP109E1 seems to be a perfect candidate for the production of these metabolites through mild and cost-effective biocatalysis with much higher selectivities and higher product yields compared to previously established methods. Moreover, CYP109E1 represents the first enzyme that is able to introduce a hydroxyl group at the C4'' position into the simvastatin molecule.

In addition to statins **1–3**, seven terpene compounds **4–10** were identified as novel substrates of CYP109E1 and their biotransformation was further investigated. Terpene and terpenoid hydroxylation and epoxidation by microbial P450 enzymes have been reported in the literature (Çelik et al. 2005; Hall and Bell 2014; Litzemberger and Bernhardt 2016; Schiffrin et al. 2015). However, the high demand of valuable terpene derivatives and their mostly unselective and low-yield production via chemical synthesis makes exploration of new effective biocatalysts for terpene oxyfunctionalization highly important. Here, α -ionone **4**, β -ionone **5**, α -damascone **8**, β -damascone **9**, β -damascenone **10**, nootkatone **6**, and isolongifolen-9-one **7** were found to serve as substrates for CYP109E1 identifying this P450 as a novel terpene hydroxylase and epoxidase. All compounds were able to bind to the enzyme and induce a type I difference spectrum of CYP109E1. Concerning dissociation constants, terpenes **4–10**

Fig. 5 Docking orientations of compactin **1** (a) and α -ionone **4** (b) shown in yellow in the active site of CYP109E1 capable of hydroxylation at C6' and C3, respectively. The distance of the corresponding carbon atom from the heme iron is given in angstrom. Amino acids forming the active site in the presence of each substrate are shown and named



showed weaker binding to the active site of CYP109E1 than statin substrates **1–3**. However, they have been converted by CYP109E1 with high efficiencies and regioselectivities in vitro as well as in vivo. The reaction products of α -ionone **4**, β -ionone **5**, α -damascone **8**, and β -damascone **9** were identified as 3-hydroxy- α -ionone **18**, 4-hydroxy- β -ionone **19**, 3-hydroxy- α -damascone **23**, and 4-hydroxy- β -damascone **24**, respectively. Thus, CYP109E1 hydroxylates preferably at allylic positions which are energetically favorable and also preferred by other P450s converting these compounds (Çelik et al. 2005; Hall and Bell 2014; Khatri et al. 2010; Litzenburger and Bernhardt 2016). The resulting hydroxylated ionones and damascenes are important compounds which might be used as precursors or building blocks in chemical synthesis of fragrance constituents due to their floral and fruity scents as well as in bioassays due to their bioactive properties. For example, the 4-hydroxy derivative of β -ionone **5** has been reported to be a versatile synthon (More and Bhat 2013) whereas 4-hydroxy- β -damascone **24** was found to possess very strong antifeedant properties against lesser mealworm *Alphitobius diaperinus* (Gliszczynska et al. 2016). The CYP109E1-based *B. megaterium* system used in our study is very suitable for the production of these metabolites. To compare it with the previously established P450-based production systems, whole-cell experiments using higher substrate concentrations and longer incubation times are necessary. However, based on the observed fast conversion of up to 200 μ M substrate within 2 h under non-optimized conditions, the whole-cell system used in our study seems to be efficient for the production of these valuable metabolites. The NMR data of nootkatone **6** and β -damascenone **10** conversion products revealed for the first time the epoxidation activity of CYP109E1 resulting in the formation of 11(*R*),12-epoxy-nootkatone **20**, 11(*S*),12-epoxy-nootkatone **21**, and 3,4-epoxy- β -damascone **26**, respectively. 3,4-Epoxy- β -damascone **26** has been previously reported to be a conversion product of β -damascone **9** by CYP101C1 from *N. aromaticivorans* DSM12444 (Ma et al. 2011), whereas 11,12-epoxy-nootkatone has been described to be produced by CYP102A1 mutants as well as unknown fungal proteins. It showed antiproliferative activity against leukemia cell line HL-60 (Gliszczynska et al. 2011; Sowden et al. 2005). In our studies, 3,4-epoxy- β -damascone **26** was further metabolized in vivo to 3,4-dihydroxy- β -damascone **27** by a CYP109E1-independent reaction proposed to be catalyzed by an unknown epoxide hydrolase from the DSM319 as described for other *B. megaterium* strains (Michaels et al. 1980; Tang et al. 2001; Zhang et al. 2010). Besides 3,4-epoxy- β -damascone **26**, a novel compound, 2-hydroxy- β -damascenone **25**, was identified as conversion product of CYP109E1. Its properties are unknown and need to be investigated. Isolongifolen-9-one **7** was hydroxylated regio- and stereoselectively at C4 by CYP109E1 yielding 4(*R*)-hydroxy-isolongifolen-9-one **22**. This compound has been described in the literature as conversion product of isolongifolen-9-one **7** by four fungal cultures, and it has been shown to have an

inhibitory activity on tyrosinase, the key enzyme for melanin biosynthesis (Choudhary et al. 2003). In another study, the suppressive effect of 4(*R*)-hydroxy-isolongifolen-9-one **22** against chemical mutagen-induced SOS response in *Salmonella typhimurium* TA1535/pSK1002 has been reported (Sakata et al. 2010).

In addition to the in vitro and in vivo characterization of CYP109E1, in silico experiments were performed using the crystal structure of this enzyme in order to predict residues responsible for substrate binding and to understand the structural basis of the observed hydroxylation regioselectivity. The molecular docking of the two selected structurally different novel substrates, compactin **1** and α -ionone **4**, into CYP109E1 indicated the presence of common as well as specific residues interacting with each substrate. The obtained docking orientations of these substrates in the active site supported experimental results, 6' β -hydroxylation of compactin **1** and 3-hydroxylation in the cyclohexene ring of α -ionone **4** (Fig. 5).

Taken together, we were able to characterize CYP109E1 from *B. megaterium* for the first time as a novel and highly regioselective statin and terpene hydroxylase as well as terpene epoxidase. Thus, CYP109E1 might be used to generate pharmaceutically and biotechnologically interesting compounds, such as the drug pravastatin **14** and the human statin drug metabolites 6' β -hydroxy-lovastatin **15** and 3' α -hydroxy-simvastatin **16** as well as several valuable terpene derivatives. Besides that, 4''-hydroxy-simvastatin **17** and a novel compound, 2-hydroxy- β -damascenone **25**, were obtained using CYP109E1 as biocatalyst, and are now available for further investigations. Finally, our *B. megaterium* whole-cell system was successfully utilized for the production of the statin and terpene metabolites in milligram scale. The established CYP109E1 system is a good candidate for improvement towards industrial scale.

Acknowledgements The authors thank Birgit Heider-Lips for the purification of AdR and Adx₄₋₁₀₈, Dr. Josef Zapp for the NMR measurements, and Ghamdan Beshr from the Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) for LC-MS measurements.

Conflict of interest The authors declare that they have no competing interests.

Compliance with ethical standards The article does not contain any studies with human participants or animals performed by any of the authors.

References

- Abdalmughni A, Jóźwik IK, Putkaradze N, Brill E, Zapp J, Thunnissen AMWH, Hannemann F, Bernhardt R (2017) Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D3 hydroxylase. J Biotechnol 243:38–47. <https://doi.org/10.1016/j.jbiotec.2016.12.023>

- Ba L, Li P, Zhang H, Duan Y, Lin Z (2013) Semi-rational engineering of cytochrome P450_{scsA} in a hybrid system for enhanced catalytic activity: insights into the important role of electron transfer. *Biotechnol Bioeng* 110:2815–2825. <https://doi.org/10.1002/bit.24960>
- Barg H, Malten M, Jahn M, Jahn D (2005) Protein and vitamin production in *Bacillus megaterium*. In: Barredo JL (ed) *Microbial processes and products*. Humana Press, New York, pp 205–223
- Bell SG, Wong LL (2007) P450 enzymes from the bacterium *Novosphingobium aromaticivorans*. *Biochem Biophys Res Commun* 360(3):666–672. <https://doi.org/10.1016/j.bbrc.2007.06.119>
- Berg A, Rafter JJ (1981) Studies on the substrate specificity and inducibility of cytochrome P-450_{mev}. *Biochem J* 196:781–786
- Bernhardt R (2006) Cytochromes P450 as versatile biocatalysts. *J Biotechnol* 124:128–145. <https://doi.org/10.1016/j.jbiotec.2006.01.026>
- Bernhardt R, Urlacher VB (2014) Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl Microbiol Biotechnol* 98:6185–6203. <https://doi.org/10.1007/s00253-014-5767-7>
- Biedendieck R, Malten M, Barg H, Bunk B, Martens JH, Deery E, Leech H, Warren MJ, Jahn D (2010) Metabolic engineering of cobalamin (vitamin B12) production in *Bacillus megaterium*. *Microb Biotechnol* 3:24–37. <https://doi.org/10.1111/j.1751-7915.2009.00125.x>
- Bleif S, Hannemann F, Zapp J, Hartmann D, Jauch J, Bernhardt R (2011) 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. <https://doi.org/10.1007/s00253-011-3467-0>
- Brill E, Hannemann F, Zapp J, Brüning G, Jauch J, Bernhardt R (2014) A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl Microbiol Biotechnol* 98:1701–1717. <https://doi.org/10.1007/s00253-013-5029-0>
- Brown AJ (2007) Cholesterol, statins and cancer. *Clin Exp Pharmacol Physiol* 34:135–141. <https://doi.org/10.1111/j.1440-1681.2007.04565.x>
- Çelik A, Flitsch SL, Turner NJ (2005) Efficient terpene hydroxylation catalysts based upon P450 enzymes derived from *Actinomyces*. *Org Biomol Chem* 3:2930–2934. <https://doi.org/10.1039/b506159h>
- Choudhary MI, Musharraf SG, Khan MTH, Abdelrahman D, Parvez M, Shaheen F, Rahman A-u (2003) Microbial transformation of isolongifolen-4-one. *Helv Chim Acta* 86:3450–3460. <https://doi.org/10.1002/hlca.200390289>
- Endo A, Hasumi K (1993) HMG-CoA reductase inhibitors. *Nat Prod Rep* 10:541–550. <https://doi.org/10.1039/NP9931000541>
- Eppinger M, Bunk B, Johns MA, Edirisinghe JN, Kutumbaka KK, Koenig SSK, Creasy HH, Rosovitz MJ, Riley DR, Daugherty S, Martin M, Elbourne LD, 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 PS (2011) Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *J Bacteriol* 193:4199–4213. <https://doi.org/10.1128/JB.00449-11>
- Fisher MT, Sligar SG (1985) Control of heme protein redox potential and reduction rate: linear free energy relation between potential and ferric spin state equilibrium. *J Am Chem Soc* 107(17):5018–5019
- Fulco AJ (1991) P450_{BM-3} and other inducible bacterial P450 cytochromes: biochemistry and regulation. *Annu Rev Pharmacol Toxicol* 31:177–203. <https://doi.org/10.1146/annurev.pa.31.040191.001141>
- Furuya T, Shibata D, Kino K (2009) Phylogenetic analysis of *Bacillus* P450 monooxygenases and evaluation of their activity towards steroids. *Steroids* 74:906–912. <https://doi.org/10.1016/j.steroids.2009.06.005>
- Gazzerro P, Proto MC, Gangemi G, Malfitano AM, Ciaglia E, Pisanti S, Santoro A, Laezza C, Bifulco M (2012) Pharmacological actions of statins: a critical appraisal in the management of cancer. *Pharmacol Rev* 64:102–146. <https://doi.org/10.1124/pr.111.004994>
- Girhard M, Klaus T, Khatri Y, Bernhardt R, Urlacher VB (2010) Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Appl Microbiol Biotechnol* 87:595–607. <https://doi.org/10.1007/s00253-010-2472-z>
- Gliszczyńska A, Lysek A, Janeczko T, Światłowska M, Wietrzyk J, Wawrzęńczyk C (2011) Microbial transformation of (+)-nootkatone and the antiproliferative activity of its metabolites. *Bioorg Med Chem* 19:2464–2469. <https://doi.org/10.1016/j.bmc.2011.01.062>
- Gliszczyńska A, Gładkowski W, Danciewicz K, Gabrys B, Szczepanik M (2016) Transformation of β -damascone to (+)-(S)-4-hydroxy- β -damascone by fungal strains and its evaluation as a potential insecticide against aphids *Myzus persicae* and lesser mealworm *Alphitobius diaperinus* Panzer. *Catal Commun* 80:39–43. <https://doi.org/10.1016/j.catcom.2016.03.018>
- Hall EA, Bell SG (2014) The efficient and selective biocatalytic oxidation of norisoprenoid and aromatic substrates by CYP101B1 from *Novosphingobium aromaticivorans* DSM12444. *RSC Adv* 5: 5762–5773. <https://doi.org/10.1039/c4ra14010a>
- Honeychurch MJ, Hill AO, Wong LL (1999) The thermodynamics and kinetics of electron transfer in the cytochrome P450cam enzyme system. *FEBS Lett* 451(3):351–353
- Janocha S, Schmitz D, Bernhardt R (2015) Terpene hydroxylation with microbial cytochrome P450 monooxygenases. *Adv Biochem Eng Biotechnol* 148:215–250. https://doi.org/10.1007/10_2014_296
- Janocha S, Carius Y, Hutter M, Lancaster CRD, Bernhardt R (2016) Crystal structure of CYP106A2 in substrate-free and substrate-bound form. *Chembiochem* 17(9):852–860. <https://doi.org/10.1002/cbic.201500524>
- Jóźwik IK, Kiss FM, Gricman Ł, Abdulmughni A, Brill E, Zapp J, Pleiss J, Bernhardt R, Thunnissen AWH (2016) Structural basis of steroid binding and oxidation by the cytochrome P450 CYP109E1 from *Bacillus megaterium*. *FEBS J* 283:4128–4148. <https://doi.org/10.1111/febs.13911>
- Kandel SE, Wienkers LC, Lampe JN (2014) Cytochrome P450 enzyme metabolites in lead discovery and development. *Annu Rep Med Chem* 49:347–359. <https://doi.org/10.1016/B978-0-12-800167-7.00022-5>
- Katagiri M, Ganguli BN, Gunsalus IC (1968) A soluble cytochrome P-450 functional in methylene hydroxylation. *J Biol Chem* 243(12): 3543–3546
- Kellick KA, Bottorff M, Toth PP (2014) A clinician's guide to statin drug-drug interactions. *J Clin Lipidol* 8:30–46. <https://doi.org/10.1016/j.jacl.2014.02.010>
- Khatri Y, Girhard M, Romankiewicz A, Ringle M, Hannemann F, Urlacher VB, Hutter MC, Bernhardt R (2010) Regioselective hydroxylation of norisoprenoids by CYP109D1 from *Sorangium cellulosum* So ce56. *Appl Microbiol Biotechnol* 88:485–495. <https://doi.org/10.1007/s00253-010-2756-3>
- Khatri Y, Ringle M, Lisurek M, von Kries JP, Zapp J, Bernhardt R (2016) Substrate hunting for the myxobacterial CYP260A1 revealed new 1 α -hydroxylated products from C-19 steroids. *Chembiochem* 17(1): 90–101. <https://doi.org/10.1002/cbic.201500420>
- Khera S, Hu N (2013) Generation of statin drug metabolites through electrochemical and enzymatic oxidations. *Anal Bioanal Chem* 405:6009–6018. <https://doi.org/10.1007/s00216-013-7021-z>
- Kim KH, Kang JY, Kim DH, Park SH, Park SH, Kim D, Park KD, Lee YJ, Jung HC, Pan JG, Ahn T, Yun CH (2011) Generation of human chiral metabolites of simvastatin and lovastatin by bacterial CYP102A1 mutants. *Drug Metab Dispos* 39(1):140–150. <https://doi.org/10.1124/dmd.110.036392>

- Kiss FM, Schmitz D, Zapp J, Dier TKF, Volmer DA, Bernhardt R (2015) Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium*—identification of a novel 11-oxidase activity. *Appl Microbiol Biotechnol* 99:8495–8514. <https://doi.org/10.1007/s00253-015-6563-8>
- Korneli C, Biedendieck R, David F, Jahn D, Wittmann C (2013) High yield production of extracellular recombinant levansucrase by *Bacillus megaterium*. *Appl Microbiol Biotechnol* 97:3343–3353. <https://doi.org/10.1007/s00253-012-4567-1>
- Kulprecha S, Boonruangthavorn A, Meksiriporn B, Thongchul N (2009) Inexpensive fed-batch cultivation for high poly(3-hydroxybutyrate) production by a new isolate of *Bacillus megaterium*. *J Biosci Bioeng* 107:240–245. <https://doi.org/10.1016/j.jbiosc.2008.10.006>
- Lamon-Fava S (2013) Statins and lipid metabolism: an update. *Curr Opin Lipidol* 24:221–226. <https://doi.org/10.1097/MOL.0b013e3283613b8b>
- Lisurek M, Kang MJ, Hartmann RW, Bernhardt R (2004) Identification of monohydroxy progesterones produced by CYP106A2 using comparative HPLC and electrospray ionisation collision-induced dissociation mass spectrometry. *Biochem Biophys Res Commun* 319:677–682
- Litzenburger M, Bernhardt R (2016) Selective oxidation of carotenoid-derived aroma compounds by CYP260B1 and CYP267B1 from *Sorangium cellulosum* So ce56. *Appl Microbiol Biotechnol* 100:4447–4457. <https://doi.org/10.1007/s00253-015-7269-7>
- Luthra A, Denisov IG, Sligar SG (2011) Spectroscopic features of cytochrome P450 reaction intermediates. *Arch Biochem Biophys* 507(1):26–35. <https://doi.org/10.1016/j.abb.2010.12.008>
- Ma M, Bell SG, Yang W, Hao Y, Rees NH, Bartlam M, Zhou W, Wong LL, Rao Z (2011) Structural analysis of CYP101C1 from *Novosphingobium aromaticivorans* DSM12444. *Chembiochem* 12:88–99
- Malten M, Hollmann R, Deckwer WD, Jahn D (2005) Production and secretion of recombinant *Leuconostoc mesenteroides* dextranucrase DsrS in *Bacillus megaterium*. *Biotechnol Bioeng* 89:206–218. <https://doi.org/10.1002/bit.20341>
- Matsuoka T, Miyakoshi S, Tanzawa K, Nakahara K, Hosobuchi M, Serizawa N (1989) Purification and characterization of cytochrome P-450sca from *Streptomyces carbophilus*. *Eur J Biochem* 184:707–713. <https://doi.org/10.1111/j.1432-1033.1989.tb15070.x>
- McLean KJ, Hans M, Meijrink B, van Scheppingen WB, Vollebregt A, Tee KL, van der Laan JM, Leys D, Munro AW, van den Berg MA (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. <https://doi.org/10.1073/pnas.1419028112>
- Michaels BC, Ruettinger RT, Fulco AJ (1980) Hydration of 9,10-epoxypalmitic acid by a soluble Enzyme from *Bacillus megaterium*. *Biochem Biophys Res Commun* 92(4):1189–1195
- Milhim M, Putkaradze N, Abdulmughni A, Kern F, Hartz P, Bernhardt R (2016) Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. *J Biotechnol* 240:68–75. <https://doi.org/10.1016/j.jbiotec.2016.11.002>
- Miura Y, Fulco AJ (1975) ω -1, ω -2 and ω -3 hydroxylation of long-chain fatty acids, amides and alcohols by a soluble enzyme system from *Bacillus megaterium*. *Biochim Biophys Acta* 388:305–317
- More GP, Bhat SV (2013) Facile lipase catalysed syntheses of (S)-(+)-4-hydroxy- β -ionone and (S)-(+)-4-hydroxy- β -damascone: chiral flavorants and synthons. *Tetrahedron Lett* 54:4148–4149. <https://doi.org/10.1016/j.tetlet.2013.05.089>
- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 30(16):2785–2791. <https://doi.org/10.1002/jcc.21256>
- Narhi LO, Fulco AJ (1987) Identification and characterization of two functional domains in cytochrome P-450BM-3, a catalytically self-sufficient monooxygenase induced by barbiturates in *Bacillus megaterium*. *J Biol Chem* 262:6683–6690
- Omura T, Sato R (1964) The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem* 239:2370–2378
- Peterson JA, Lu JY, Geisselsoder J, Graham-Lorence S, Carmona C, Witney F, Lorence MC (1992) Cytochrome P-450terp. Isolation and purification of the protein and cloning and sequencing of its operon. *J Biol Chem* 267(20):14193–14203
- Poulos TL, Raag R (1992) Cytochrome P450cam: crystallography, oxygen activation, and electron transfer. *FASEB J* 6(2):674–679
- Putkaradze N, Kiss FM, Schmitz D, Zapp J, Hutter MC, Bernhardt R (2017) Biotransformation of prednisone and dexamethasone by cytochrome P450 based systems—identification of new potential drug candidates. *J Biotechnol* 242:101–110. <https://doi.org/10.1016/j.jbiotec.2016.12.011>
- Rauschenbach R, Isernhagen M, Noeske-Jungblut C, Boidol W, Siewert G (1993) Cloning sequencing and expression of the gene for cytochrome P450meg, the steroid-15 β -monooxygenase from *Bacillus megaterium* ATCC 13368. *Mol Gen Genet* 241:170–176
- Rygu T, Hillen W (1992) Catabolite repression of the xyl operon in *Bacillus megaterium*. *J Bacteriol* 174:3049–3055
- Sagara Y, Wada A, Takata Y, Waterman MR, 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 (2012) Practical application of cytochrome P450. *Biol Pharm Bull* 35:844–849
- Sakata K, Oda Y, Miyazawa M (2010) Suppression of SOS-inducing activity of chemical mutagens by metabolites from microbial transformation of (–)-isolongifolene. *J Agric Food Chem* 58:2164–2167. <https://doi.org/10.1021/jf903651c>
- Sanner MF (1999) Python: a programming language for software integration and development. *J Mol Graph Modell* 17(1):57–61
- Schenkman JB, Jansson I (1998) Spectral analyses of cytochromes P450. In: Phillips I, Shephard EA (eds) *Cytochrome P450 protocols*. Humana Press, New York, pp 25–34
- Schiffrin A, Litzenburger M, Ringle M, Ly TTB, Bernhardt R (2015) New sesquiterpene oxidations with CYP260A1 and CYP264B1 from *Sorangium cellulosum* So ce56. *Chembiochem* 16:2624–2632. <https://doi.org/10.1002/cbic.201500417>
- Schmitz D, Zapp J, Bernhardt R (2012) Hydroxylation of the triterpenoid dipterocarpol with CYP106A2 from *Bacillus megaterium*. *FEBS J* 279:1663–1674. <https://doi.org/10.1111/j.1742-4658.2012.08503.x>
- Schmitz D, Zapp J, Bernhardt R (2014) Steroid conversion with CYP106A2—production of pharmaceutically interesting DHEA metabolites. *Microb Cell Factories* 13:81. <https://doi.org/10.1186/1475-2859-13-81>
- Simgen B, Contzen J, Schwarzer R, Bernhardt R, Jung C (2000) Substrate binding to 15 β -hydroxylase (CYP106A2) probed by FT infrared spectroscopic studies of the iron ligand CO stretch vibration. *Biochem Biophys Res Commun* 269:737–742. <https://doi.org/10.1006/bbrc.2000.2348>
- Sligar SG, Gunsalus IC (1976) A thermodynamic model of regulation: modulation of redox equilibria in camphor monooxygenase. *Proc Natl Acad Sci* 73(4):1078–1082
- Sono M, Roach MP, Coulter ED, Dawson JH (1996) Heme-containing oxygenases. *Chem Rev* 96:2841–2888
- Sowden RJ, Yasmin S, Rees NH, Bell SG, Wong LL (2005) Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450cam and P450BM-3. *Org Biomol Chem* 3:57–64. <https://doi.org/10.1039/b413068e>
- Stammen S, Müller BK, Korneli C, Biedendieck R, Gamer M, Franco-Lara E, Jahn D (2010) High-yield intra- and extracellular protein

- production using *Bacillus megaterium*. Appl Environ Microbiol 76: 4037–4046. <https://doi.org/10.1128/AEM.00431-10>
- Stokker GE (1994) Synthesis of the 3'(S)-hydroxy derivative of simvastatin. Bioorg Med Chem Lett 4:1767–1770. [https://doi.org/10.1016/S0960-894X\(00\)80377-7](https://doi.org/10.1016/S0960-894X(00)80377-7)
- Tang YF, Xu JH, Ye Q, Schulze B (2001) Biocatalytic preparation of (S)-phenyl glycidyl ether using newly isolated *Bacillus megaterium* ECU1001. J Mol Catal B Enzym 13:61–68. [https://doi.org/10.1016/S1381-1177\(00\)00230-7](https://doi.org/10.1016/S1381-1177(00)00230-7)
- Uhlmann H, Beckert V, Schwarz D, Bernhardt R (1992) Expression of bovine adrenodoxin in *E. coli* and site-directed mutagenesis of 2FE-2S/ cluster ligands. Biochem Biophys Res Commun 188:1131–1138
- Ullah AJ, Murray RI, Bhattacharyya PK, Wagner GC, Gunsalus IC (1990) Protein components of a cytochrome P-450 linalool 8-methyl hydroxylase. J Biol Chem 265(3):1345–1351
- Urlacher VB, Lutz-Wahl S, Schmid RD (2004) Microbial P450 enzymes in biotechnology. Appl Microbiol Biotechnol 64:317–325. <https://doi.org/10.1007/s00253-003-1514-1>
- Urlacher VB, Makhsumkhanov A, Schmid RD (2006) Biotransformation of beta-ionone by engineered cytochrome P450 BM-3. Appl Microbiol Biotechnol 70:53–59. <https://doi.org/10.1007/s00253-005-0028-4>
- Vary PS, Biedendieck R, Fuerch T, Meinhardt F, Rohde M, Deckwer WD, Jahn D (2007) *Bacillus megaterium*—from simple soil bacterium to industrial protein production host. Appl Microbiol Biotechnol 76: 957–967. <https://doi.org/10.1007/s00253-007-1089-3>
- Virus C, Lisurek M, Simgen B, Hannemann F, Bernhardt R (2006) Function and engineering of the 15 β -hydroxylase CYP106A2. Biochem Soc Trans 34:1215–1218. <https://doi.org/10.1042/BST0341215>
- Williams JW, Morrison JF (1979) The kinetics of reversible tight-binding inhibition. Methods Enzymol 63:437–467
- Wittchen KD, Meinhardt F (1995) Inactivation of the major extracellular protease from *Bacillus megaterium* DSM319 by gene replacement. Appl Microbiol Biotechnol 42:871–877. <https://doi.org/10.1007/BF00191184>
- Zhang Z, Sheng Y, Jiang K, Wang Z, Zheng Y, Zhu Q (2010) Bio-resolution of glycidyl (o, m, p)-methylphenyl ethers by *Bacillus megaterium*. Biotechnol Lett 32:513–516. <https://doi.org/10.1007/s10529-009-0181-4>

Supplementary material

Applied Microbiology and Biotechnology

CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*

Natalia Putkaradze^a, Martin Litzenburger^a, Ammar Abdulmughni^a, Mohammed Milhim^a, Elisa Brill^a, Frank Hannemann^a, Rita Bernhardt^a

^a Institute of Biochemistry, Saarland University, D-66123 Saarbruecken, Germany

Corresponding author:

Prof. Dr. Rita Bernhardt

Department of Biochemistry, Saarland University, Campus, B 2.2

D-66123 Saarbruecken, Germany

Tel.: +49 (0)681 302 4241

Fax: +49 (0)681 302 4739

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

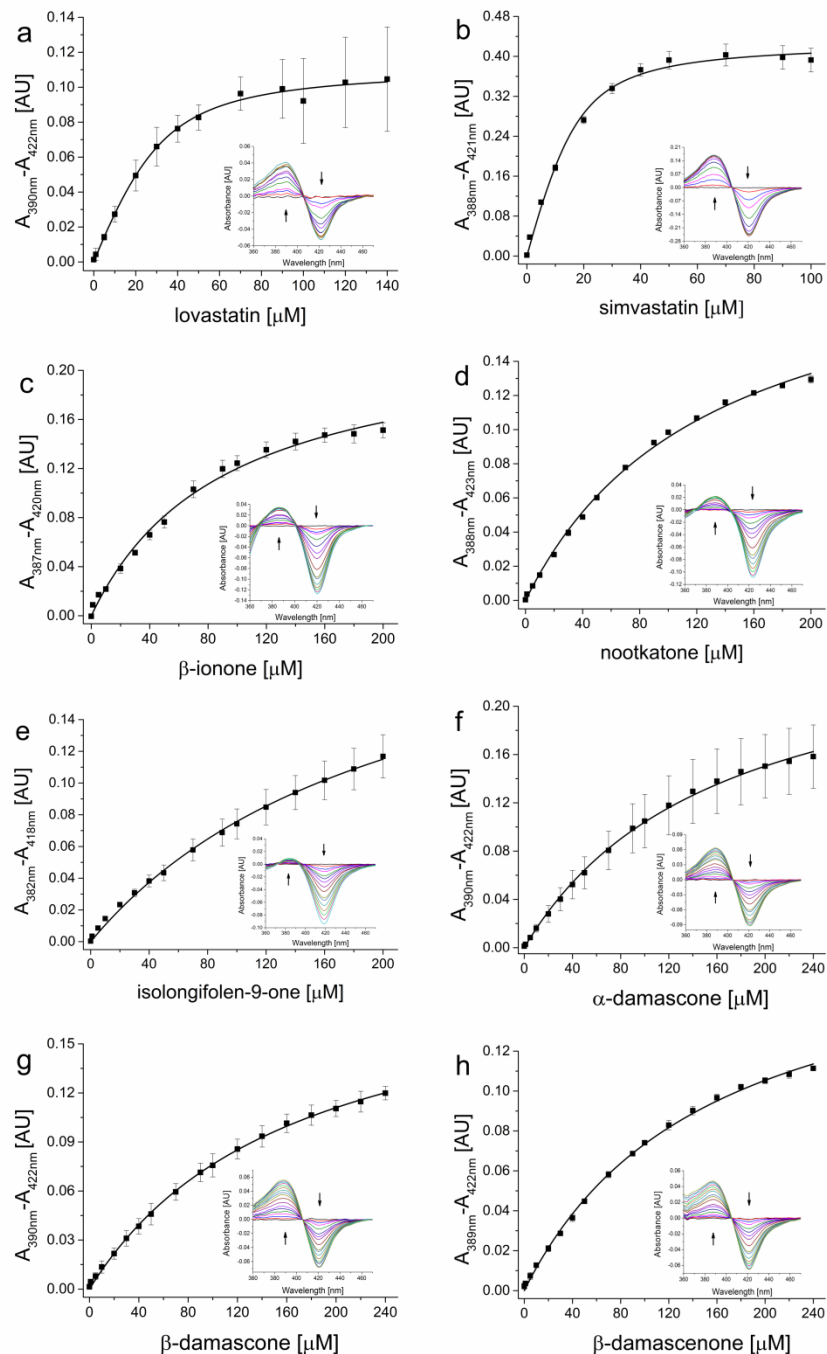


Fig. S1 The absorbance changes plotted against the corresponding concentrations of lovastatin **2** (a), simvastatin **3** (b), β -ionone **5** (c), nootkatone **6** (d), isolongifolen-9-one **7** (e), α -damascone **8** (f), β -damascone **9** (g) and β -damascenone **10** (h) titrated to CYP109E1 as described in "Material and methods". The mean values were fitted by either tight binding equation or hyperbolic regression and the K_d value was calculated. The *inset* shows type I spectral shifts induced by the binding of the increasing

amount of the substrate to CYP109E1. *Arrows* show the direction of spectral changes at increasing substrate concentrations

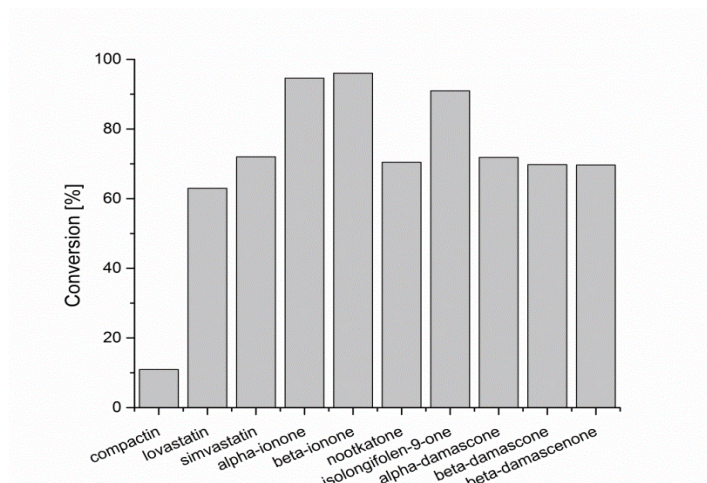


Fig. S2 Conversion ratio of 100 μ M compactin **1**, lovastatin **2** and simvastatin **3** and 200 μ M of other substrates **4-10** after 15 min *in vitro* in the presence of CYP109E1 using AdR and Adx₄₋₁₀₈ as redox partners

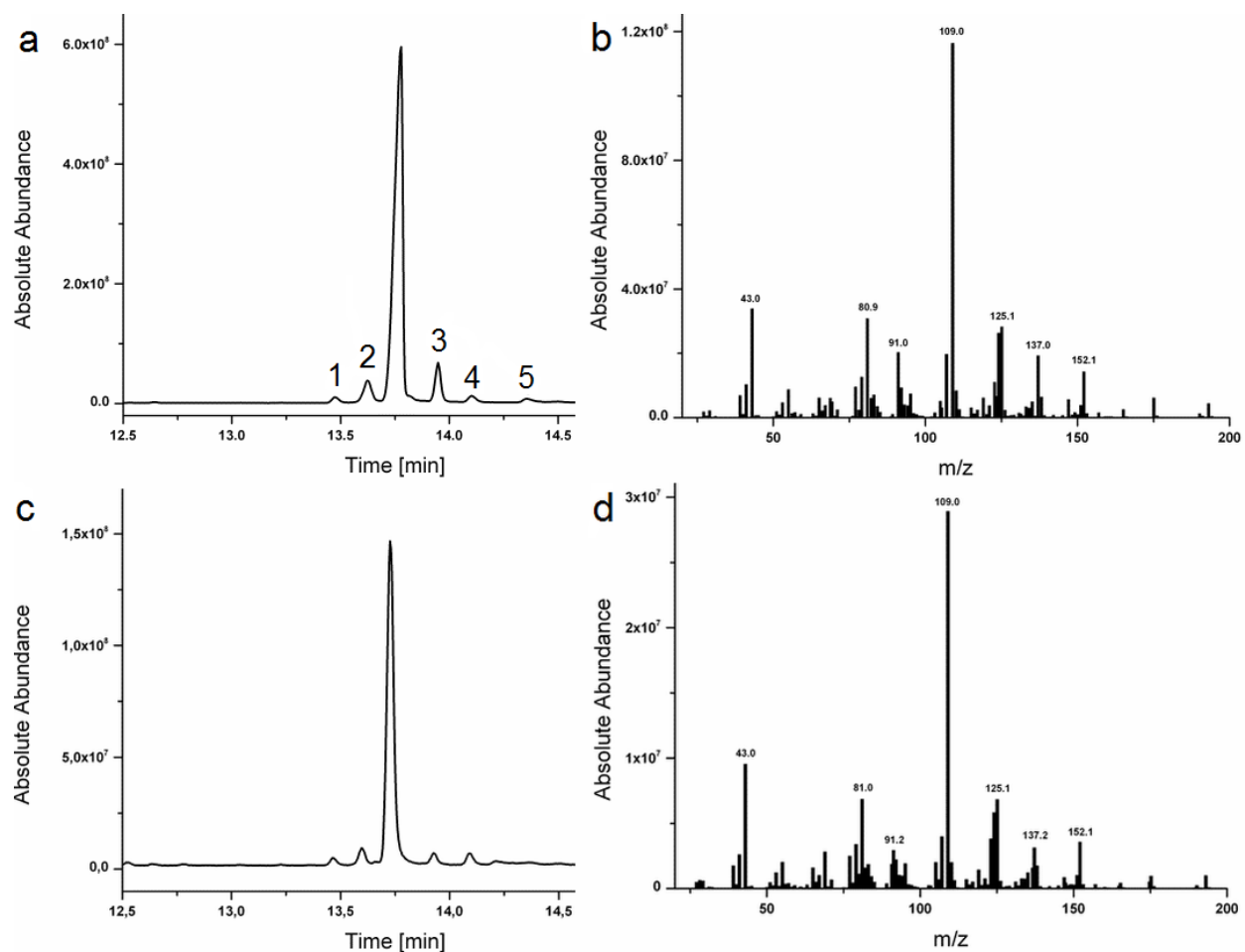


Fig. S3 GC-MS data of the α -ionone **4** conversion product by CYP109E1 (a and b) and by CYP267B1 (c and d) identified as 3-hydroxy- α -ionone **18** (Litzenburger et al. 2016). 1-5 are the minor products with known masses (Table S1)

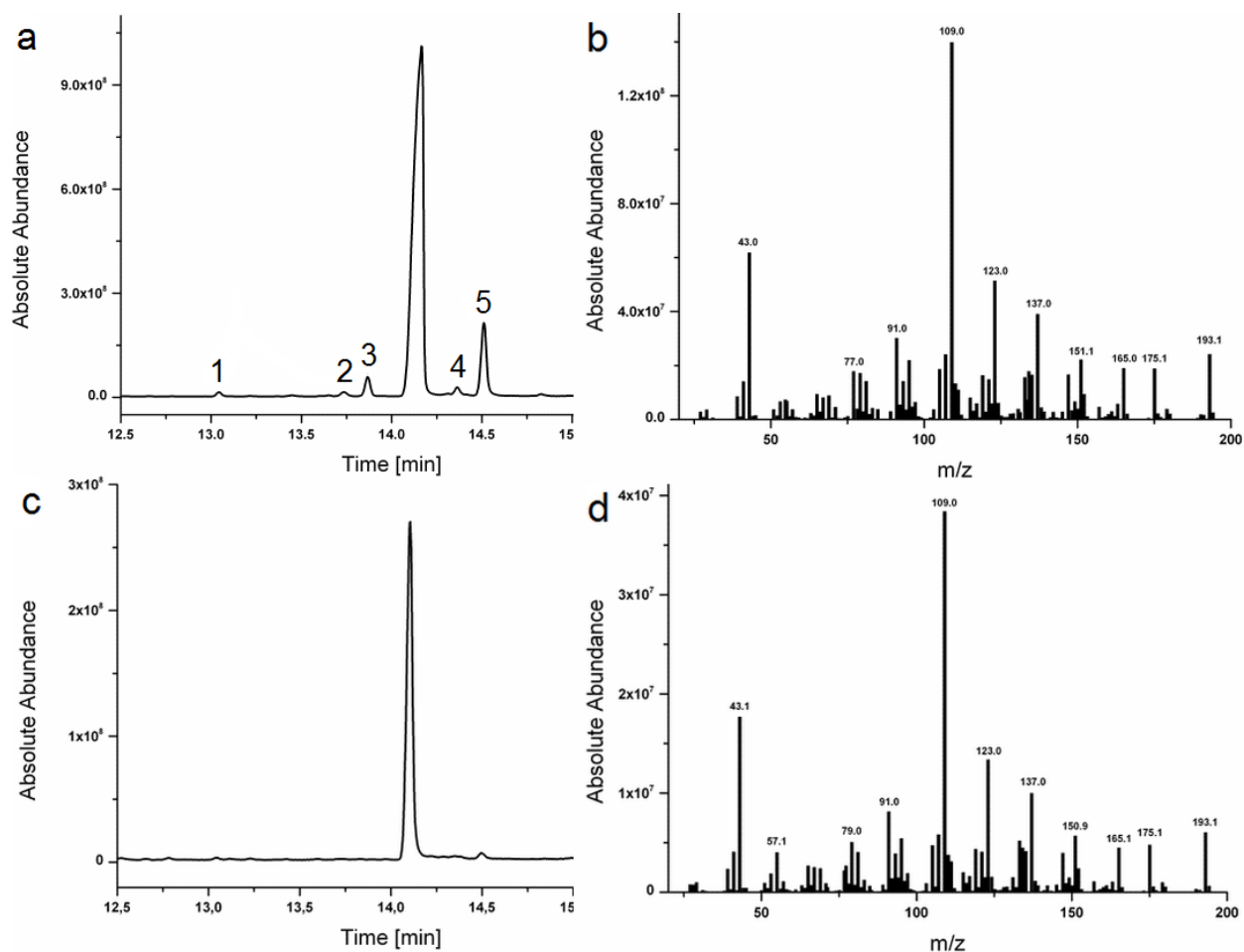


Fig. S4 GC-MS data of the β -ionone **5** conversion product by CYP109E1 (a and b) and by CYP267B1 (c and d) identified as 4-hydroxy- β -ionone **19** (Litzenburger et al. 2016). 1-5 are the minor products with known masses (Table S1)

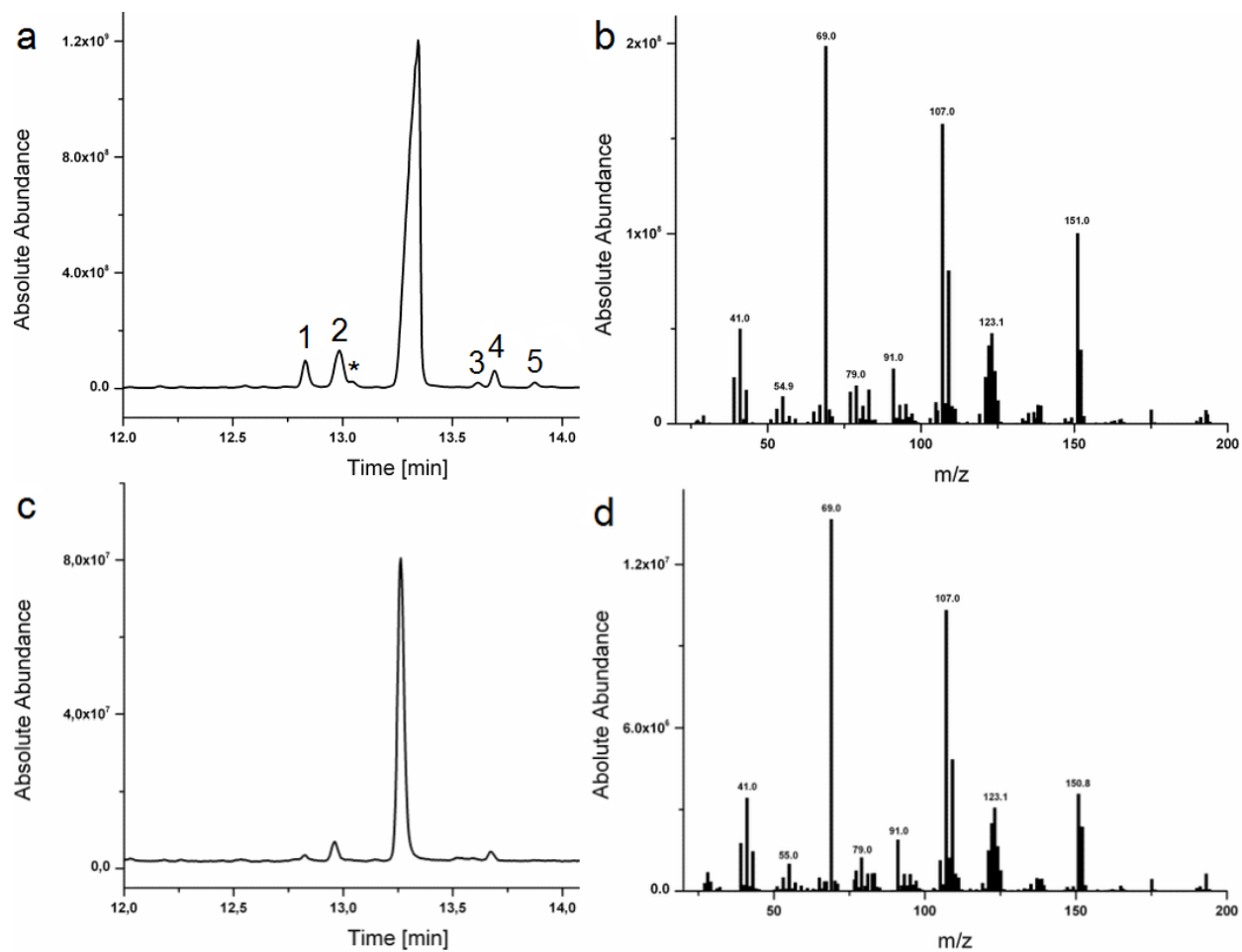


Fig. S5 GC-MS data of the α -damascone **8** conversion product by CYP109E1 (a and b) and by CYP267B1 (c and d) identified as 3-hydroxy- α -damascone **23** (Litzenburger et al. 2016). 1-5 are the minor products with known masses (Table S1). Impurity is labelled with “*”

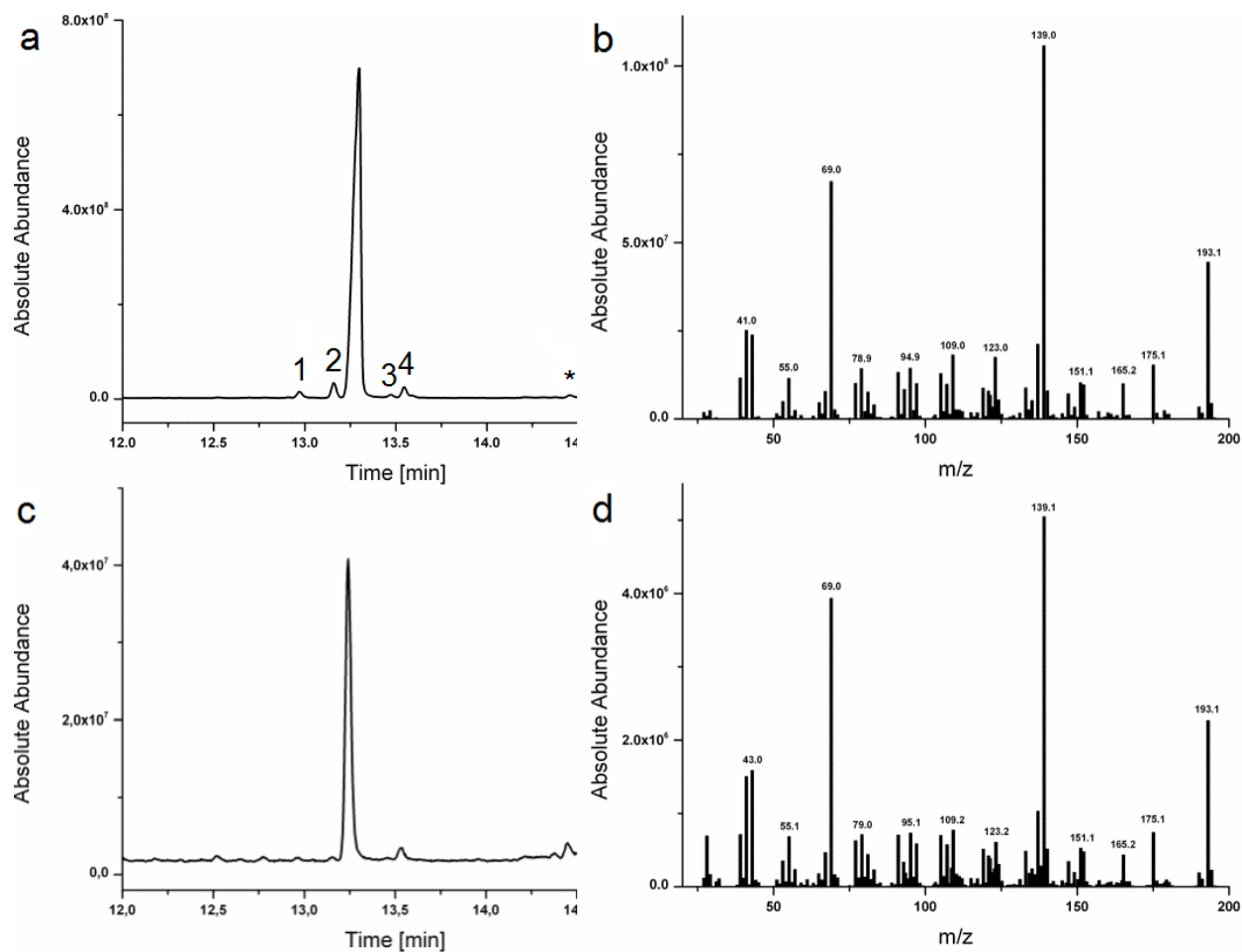
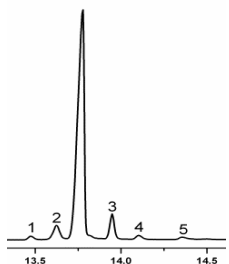
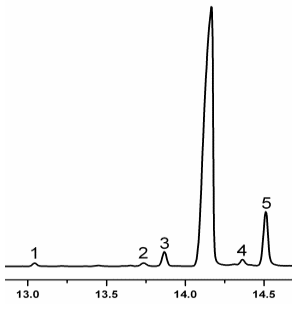
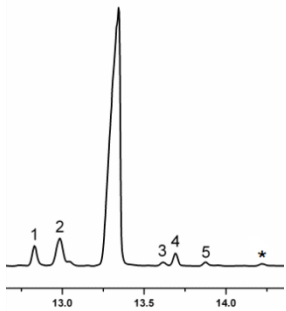
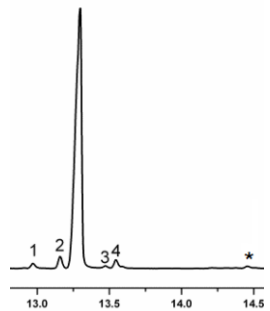
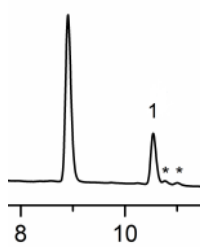
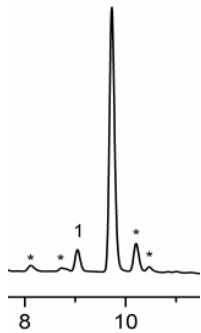
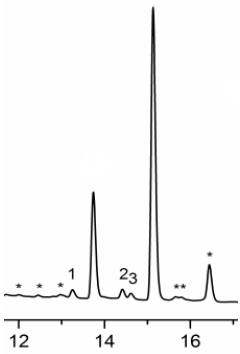
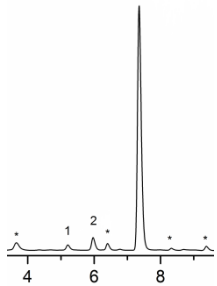
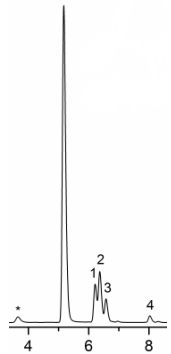


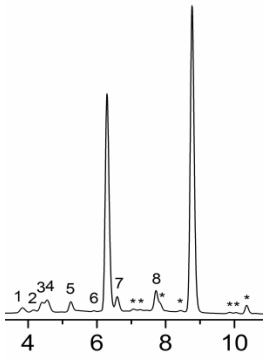
Fig. S6 GC-MS data of the β -damascone **9** conversion product by CYP109E1 (a and b) and by CYP267B1 (c and d) identified as 4-hydroxy- β -damascone **24** (Litzenburger et al. 2016). 1-4 are the minor products with known masses (Table S1). Impurity is labelled with “*”

Table S1 Selectivity of the CYP109E1-catalyzed reactions *in vitro* and masses of the reaction products. Minor products are numbered and labelled. “*” indicates impurities

Compound		product	t_R [min]	$\Delta m/z$	Selectivity [%]
α -ionone					
		1	13.5	+16	1.1
		2	13.6	+16	5.2
		3	13.9	+14	6.2
		4	14.1	+16	1.4
		5	14.4	+16	1.1
		3-hydroxy- α -ionone	13.7	+16	85

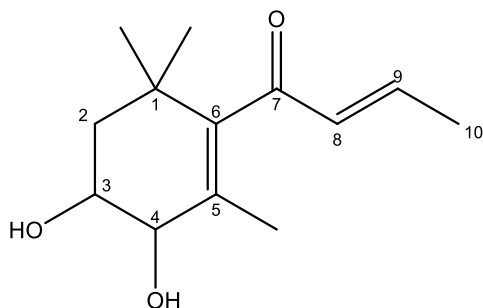
β -ionone		1	13.0	+16	0.6
		2	13.7	+16	0.7
		3	13.9	+16	2.6
		4	14.4	+16	1.3
		5	14.5	+16	10.2
		4-hydroxy- β -ionone	14.2	+16	84.5
α -damascone		1	12.8	+16	3.6
		2	13.0	+16	7.4
		3	13.6	+16	0.6
		4	13.7	+14	2.1
		5	13.9	+16	0.8
		3-hydroxy- α -damascone	13.3	+16	85.4
β -damascone		1	13.0	+16	1.6
		2	13.2	+14	2.7
		3	13.5	+16	0.4
		4	13.6	+16	2.4
		4-hydroxy- β -damascone	13.3	+16	92.8
compactin		1	10.6	+16	23
		pravastatin	8.9	+16	77

lovastatin					
		1	9.1	+16	6
		6'β-hydroxy-lovastatin	9.8	+16	94
simvastatin					
		1	13.3	+16	4.7
		2	14.4	+16	5
		3	14.6	+14	3
		3'α-hydroxy-simvastatin	13.8	+16	23
nootkatone					
		1	5.2	+16	1.2
		2	5.9	+16	6.2
		11,12-epoxy-nootkatone	7.4	+16	92.6
isolongifolen-9-one					
		1	6.3	+16	6.5
		2	6.4	+16	10.6
		3	6.6	+16	4.4
		4	8.1	+14	1.3
		4(R)-hydroxy-isolongifolen-9-one	5.3	+16	77.2

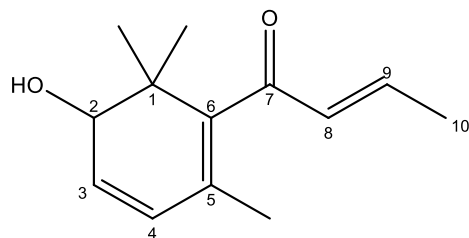
β -damascenone					
		1	3.8	+32	1.6
		2	4.1	+32	0.8
		3	4.4	+32	2
		4	4.5	+32	3
		5	5.3	+18	1.6
		6	5.9	+14	0.5
		7	6.6	+16	3
		8	7.7	+16	3.6
		2-hydroxy- β -damascenone	6.4	+16	35.3
		3,4-epoxy β -damascone	8.7	+16	48.7

Data S7: NMR data of investigated compounds

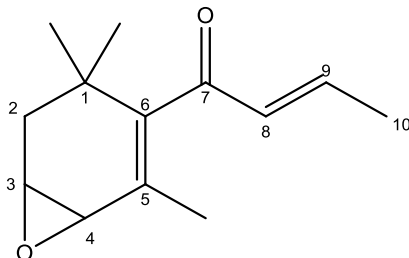
3,4-dihydroxy- β -damascone 27: ^1H NMR (CDCl_3 , 500 MHz): 6.74 (dq, 1H, $J = 15.8, 6.9$ Hz, H9), 6.13 (dq, 1H, $J = 15.8, 1.6$ Hz, H8), 3.94 (d, 1H, $J = 7.9$ Hz, H4), 3.87 (ddd, 1H, $J = 12.2, 8.1, 3.8$ Hz, H3), 1.91 (dd, 3H, $J = 6.9, 1.6$ Hz, H10), 1.74-1.63 (m, 2H, H2a and H2b), 1.60 (d, 3H, $J = 1.0$ Hz, C5-Me), 1.17 (s, 3H, C1-Me), 0.96 (s, 3H, C1-Me); ^{13}C NMR (CDCl_3 , 125 MHz): 200.23 (C7), 146.77 (C9), 142.14 (C6), 133.81 (C8), 129.89 (C5), 76.88 (C4), 71.08 (C3), 44.88 (C2), 36.15 (C1), 29.85 (C1-Me), 28.70 (C1-Me), 18.46 (C10), 15.88 (C5-Me).



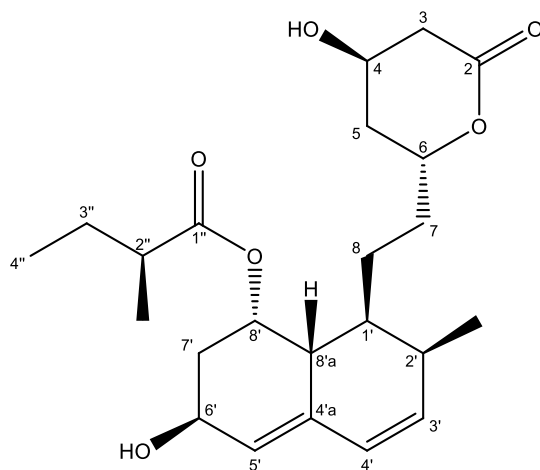
2-hydroxy- β -damascenone 25: ^1H NMR (CDCl_3 , 500 MHz): 6.81 (dq, 1H, $J = 15.7, 6.9$ Hz, H9), 6.17 (dq, 1H, $J = 15.7, 1.6$ Hz, H8), 5.97 (dd, 1H, $J = 9.5$ Hz, 4.7 Hz, H3), 5.87 (d, 1H, $J = 9.5$ Hz, H4), 3.80 (d, 1H, $J = 4.8$ Hz, H2), 1.92 (dd, 3H, $J = 6.9$ Hz, 1.6 Hz, H10), 1.65 (s, 3H, C5-Me), 1.08 (s, 6H, C1-Me (2x)); ^{13}C NMR (CDCl_3 , 125 MHz): 200.34 (C7), 146.98 (C9), 140.05 (C6), 134.28 (C8), 129.15 (C3), 129.06 (C4), 126.80 (C5), 73.79 (H2), 39.21 (C1), 24.69 (C1-Me), 19.37 (C1-Me), 19.24 (C5-Me), 18.93 (C10).



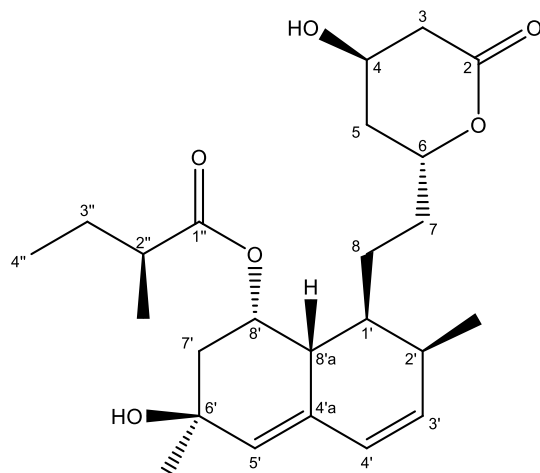
3,4-epoxy β -damascone 26: ^1H NMR (CDCl_3 , 500 MHz): 6.73 (dq, 1H, J = 15.7, 6.9 Hz, H9), 6.13 (dq, 1H, J = 15.7, 1.6 Hz, H8), 3.50 (ddd, J =6.6, 4.3, 2.3 Hz, H3), 3.13 (d, 1H, J = 4.3 Hz, H4), 1.98 (dd, J = 14.6, 2.5 Hz, 1H, H2a), 1.91 (dd, 3H, J = 7.0, 1.5 Hz, H10), 1.74-1.63 (m, 1H, H2b), 1.61 (d, 3H, J = 1.0 Hz, C5-Me), 1.16 (s, 3H, C1-Me), 0.96 (s, 3H, C1-Me); ^{13}C NMR (CDCl_3 , 125 MHz): 200.14 (C7), 147.59 (C9), 144.74 (C6), 133.89 (C8), 127.03 (C5), 54.24 (C4), 51.20 (C3), 37.90 (C2), 33.95 (C1), 32.12 (C1-Me), 29.55 (C1-Me), 19.83 (C10), 18.50 (C5-Me).



Pravastatin 14: ^1H NMR (500 MHz, CDCl_3): 5.98 (d, 1H, J =9.7 Hz, H4'), 5.87 (dd, 1H, J = 9.7, 5.9 Hz, H3'), 5.55 (br s, 1H, H5'), 5.39 (br s, 1H, H8'), 4.62-4.56 (m, 1H, H6), 4.42-4.34 (m, 2H, H6', H4), 2.72 (dd, 1H, J = 17.6, 5.1 Hz, H3a), 2.60 (ddd, J = 17.6, 3.7, 1.6 Hz, H3b), 2.55 (m, 1H, H7), 2.40-2.30 (m, 3H, H2', H8'a, H2''), 1.93-1.90 (m, 1H, H5a), 1.85 (m, 1H, H7a), 1.70-1.62 (m, 3H, H1', H5b, H3''a), 1.44-1.37 (m, 3H, H8, H3''b), 1.28-1.24 (m, 1H, H7b), 1.09 (d, 3H, J =7.0 Hz, C2''-Me), 0.88 (d, 3H, J = 7.0 Hz, C2'-Me), 0.87 (t, 3H, J =7.4 Hz, H4''); ^{13}C NMR (125 MHz, CDCl_3): 176.29 (C1'), 169.94 (C2), 135.79 (C3'), 135.29 (C4'a), 127.39 (C4'), 126.12 (C5'), 75.93 (C6), 69.04 (C8'), 65.04 (C6'), 62.75 (C4), 41.58 (C2''), 38.59 (C3), 37.60 (C8'a), 36.69 (C7'), 36.60 (C1'), 36.18 (C5), 32.75 (C7), 30.95 (C2'), 26.64 (C3''), 23.76 (C8), 16.80 (C2''-Me), 13.56 (C2'-Me), 11.76 (C4'').

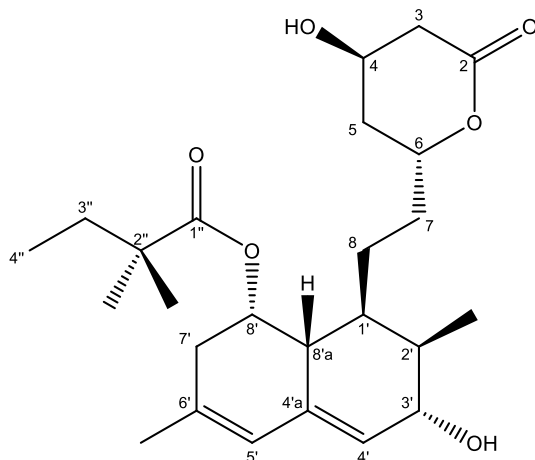


6'β-hydroxy-lovastatin 15: ^1H NMR (500 MHz, CDCl_3): 5.96 (d, 1H, $J = 10.0$ Hz, $\text{H}_{4'}$), 5.89 (dd, 1H, $J = 10.0, 6.1$ Hz, $\text{H}_{3'}$), 5.43 (br s, 1H, $\text{H}_{5'}$), 5.39 (m, 1H, $\text{H}_{8'}$), 4.62-4.55 (m, 1H, $\text{H}_{6'}$), 4.37-4.32 (m, 1H, H_4), 2.70 (dd, 1H, $J = 17.6, 5.1$ Hz, H_{3a}), 2.60 (ddd, 1H, $J = 17.6, 3.8, 1.7$ Hz, H_{3b}), 2.42-2.29 (m, 5H, $\text{H}_{7'}$, $\text{H}_{2'}$, $\text{H}_{8'a}$, $\text{H}_{2''}$), 1.96-1.80 (m, 2H, H_{5a} , H_{7a}), 1.70-1.60 (m, 3H, $\text{H}_{1'}$, H_{5b} , $\text{H}_{3''a}$), 1.44-1.38 (m, 3H, H_8 , $\text{H}_{3''b}$), 1.32 (s, 3H, $\text{C}_{6'}$ -Me), 1.29-1.24 (m, 1H, H_{7b}), 1.10 (d, 3H, $J = 6.9$ Hz, $\text{C}_{2''}$ -Me), 0.87 (d, 3H, $J = 7.2$ Hz, $\text{C}_{2'}$ -Me), 0.86 (t, 3H, $J = 7.4$ Hz, $\text{H}_{4''}$); ^{13}C NMR (125 MHz, CDCl_3): 176.41 ($\text{C}_{1''}$), 170.12 (C_2), 135.88 ($\text{C}_{3'}$), 133.48 ($\text{C}_{4'a}$), 129.64 ($\text{C}_{5'}$), 127.58 ($\text{C}_{4'}$), 76.05 (C_6), 68.92 ($\text{C}_{8'a}$), 68.77 ($\text{C}_{6'}$), 62.71 (C_4), 42.18 ($\text{C}_{7'}$), 41.41 ($\text{C}_{2''}$), 38.54 (C_3), 37.41 ($\text{C}_{8'a}$), 36.34 (C_5), 36.17 (C_6), 32.69 (C_7), 30.75 ($\text{C}_{6'}$ -Me), 30.55 ($\text{C}_{2'}$), 26.82 ($\text{C}_{3''}$), 24.00 (C_8), 16.25 ($\text{C}_{2''}$ -Me), 13.55 ($\text{C}_{2'}$ -Me), 11.71 ($\text{C}_{4''}$).

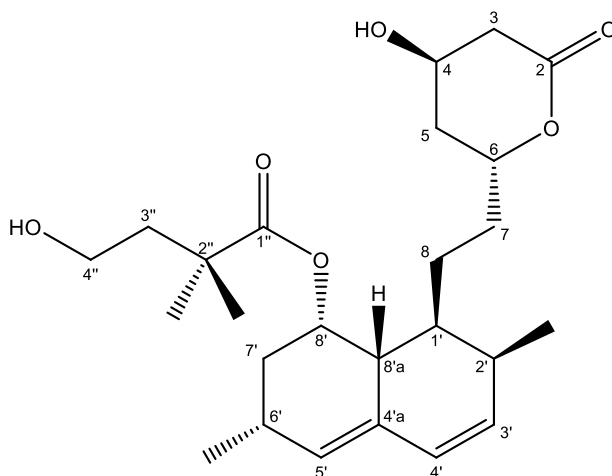


3'α-hydroxy-simvastatin 16: ^1H NMR (500 MHz, CDCl_3): 5.87 (s, 1H, $\text{H}_{5'}$), 5.60 (d, 1H, $J = 4.7$ Hz, $\text{H}_{4'}$), 5.37 (d, 1H, $J = 2.6$ Hz, $\text{H}_{8'}$), 4.67-4.60 (m, 1H, H_6), 4.38-4.34 (m, 1H, H_4), 3.92 (d, 1H, $J = 4.7$ Hz, $\text{H}_{3'}$), 2.72 (dd, 1H, $J = 17.6, 5.1$ Hz, H_{3a}), 2.60 (ddd, 1H, $J = 17.6, 3.8, 1.6$ Hz, H_{3b}), 2.38 (d, 1H, $J = 17.1$

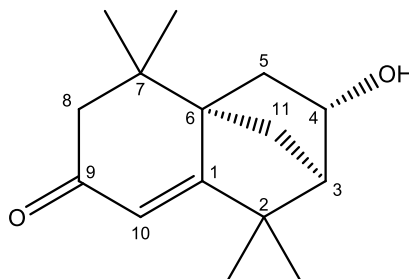
Hz, H7'a), 2.21 (dd, 1H, J= 18.9, 1.9 Hz, H7'b), 2.02-1.91 (m, 3H, H5a, H2' H8'a), 1.80-1.65 (m, 4H, H5b, H1', H8a, H7a), 1.71 (s, 3H, C6'-Me), 1.54-1.44 (m, 4H, H7b, H8b, H3''), 1.07 (s, 3H, C2''-Me), 1.06 (s, 3H, C2''-Me), 0.76 (t, 3H, J=7.5 Hz, C4''), 0.74 (d, 3H, J=7.2 Hz, C2'-Me); ¹³C NMR (125 MHz, CDCl₃): 177.75 (C1''), 170.24 (C2), 137.52 (C6'), 134.06 (C4'a), 124.24 (C5'), 122.38 (C4'), 76.12 (C6), 70.11 (C3'), 67.46 (C8'), 62.76 (C4), 42.93 (C2''), 39.69 (C8'a), 38.58 (C3), 36.37 (C5), 36.33 (C7'), 35.48 (C2'), 33.23 (C7), 33.19 (C3''), 31.33 (C1'), 24.93 (C2''-Me), 24.93 (C8) 24.38 (C2''-Me), 23.29 (C6'-Me), 10.57 (C2'-Me), 9.26 (C4'').



4''-hydroxy-simvastatin 17: ¹H NMR (500 MHz, CDCl₃): 5.97 (d, 1H, J= 9.7 Hz, H4'), 5.76 (dd, 1H, J= 9.6, 6.1 Hz, H3'), 5.49 (br s, 1H, H5'), 5.33 (q, 1H, J= 3.3, 2.9 Hz, H8'), 4.65-4.59 (m, 1H, H6), 4.33-4.30 (m, 1H, H4), 3.70-3.65 (m, 1H, H4''), 3.64- 3.59 (m, 1H, H4''), 2.74 (dd, 1H, J= 17.7, 5.4 Hz, H3a), 2.58 (ddd, 1H, J=17.7, 4.1, 1.5 Hz, H3b), 2.43-2.39 (m, 1H, H6'), 2.37-2.32 (m, 1H, H2'), 2.25 (dd, 1H, J= 11.8, 2.6 Hz, H8'a), 1.98-1.89 (m, 4H, H7', H5a, H3''a), 1.96-1.89 (m, 1H, H7a), 1.79-1.66 (m, 3H, H5b, H1', H3''b), 1.48 -1.39 (m, 1H, H8a), 1.38-1.34 (m, 2 H, H8b, H7b), 1.18 (s, 3H, C2''-Me), 1.16 (s, 3H, C2''-Me), 1.06 (d, 3H, J= 7.4 Hz, C6'-Me), 0.86 (d, 3H, J= 7.0, C2'-Me); ¹³C NMR (125 MHz, CDCl₃): 178.58 (C1''), 170.64 (C2), 132.93 (C3'), 131.39 (C4'a), 129.62 (C5'), 128.32 (C4'), 76.29 (C6), 68.87 (C8'), 62.44 (C4), 59.42 (C4''), 42.54 (C3''), 41.53 (C2''), 38.46 (C3), 37.39 (C8'a), 36.88 (C1'), 35.88 (C5), 32.59 (C7'), 32.56 (C7), 30.61 (C2'), 27.23 (C6'), 26.23 (C2''-Me), 24.93 (C2''-Me), 24.19 (C8), 23.08 (C6'-Me), 13.81 (C2'-Me).



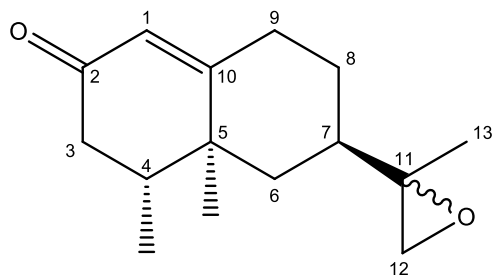
4(R)-hydroxy-isolongifolen-9-one 22: ^1H NMR (CDCl_3 , 500 MHz): 5.70 (s, 1H, H10), 4.31 (brd, 1H, $J=6.6$ Hz, H4), 2.31 (d, 1H, $J=16.2$ Hz, H8a), 2.06 (dd, 1H, $J=16.2, 0.8$ Hz, H8b), 1.95 (brs, 1H, H3), 1.80 (ddd, 1H, $J=13.0, 6.6, 2.6$ Hz, H5a), 1.73 (ddd, 1H, $J=13.0, 2.1, 1.0$ Hz, H5b), 1.59 (m, 2H, H11), 1.11 (s, 3H, C2-Me), 1.09 (s, 3H, C2-Me), 1.06 (s, 3H, C7-Me), 0.97 (s, 3H, C7-Me); ^{13}C NMR (CDCl_3 , 125 MHz): 199.81 (C9), 181.71 (C1), 117.58 (C10), 70.30 (C4), 57.96 (C6), 54.42 (C3), 49.81 (C8), 42.10 (C2), 40.16 (C5), 34.10 (C7), 32.82 (C11), 27.20 (C2-Me), 25.90 (C7-Me), 25.28 (C7-Me), 24.25 (C2-Me).



11(R), 12-epoxy-nootkatone 20: ^1H NMR (CDCl_3 , 500 MHz): 5.73 (brs, 1H, H1), 2.62 (m, 1H, H12a), 2.56 (m, 1H, H12b), 2.42 (m, 1H, H9a), 2.35 (m, 1H, H9b), 2.23 (m, 2H, H3), 2.02 (m, 1H, H6a), 1.99 (m, 1H, H7), 1.88 (m, 1H, H8a), 1.60 (m, 1H, H4), 1.22 (m, 1H, H8b), 1.22 (s, 3H, H13), 1.10 (m, 1H, H6b), 1.04 (s, 3H, C5-Me), 0.96 (d, 3H, $J=6.7$ Hz, C4-Me); ^{13}C NMR (CDCl_3 , 125 MHz): 199.50 (C2), 168.80 (C10), 124.87 (C1), 58.91 (C11), 53.33 (C12), 42.01 (C3), 40.50 (C7), 40.43 (C6), 39.60 (C4), 39.01 (C5), 32.43 (C9), 28.79 (C8), 17.73 (C13), 16.74 (C5-Me), 14.89 (C4-Me).

11(S), 12-epoxy-nootkatone 21: ^1H NMR (CDCl_3 , 500 MHz): 5.73 (brs, 1H, H1), 2.63 (m, 1H, H12a), 2.57 (m, 1H, H12b), 2.42 (m, 1H, H9a), 2.35 (m, 1H, H9b), 2.23 (m, 2H, H3), 1.99 (m, 1H, H7), 1.97 (m, 1H, H8a), 1.90 (m, 1H, H6a), 1.58 (m, 1H, H4), 1.29 (m, 1H, H8b), 1.24 (s, 3H, H13), 1.04 (s, 3H, C5-

Me), 1.00 (m, 1H, H6b), 0.94 (d, 3H, J=6.7 Hz, C4-Me); ^{13}C NMR (CDCl_3 , 125 MHz): 199.40 (C2), 168.87 (C10), 124.87 (C1), 59.13 (C11), 53.65 (C12), 42.01 (C3), 40.74 (C6), 40.44 (C7), 39.42 (C4) 38.97 (C5), 32.38 (C9), 28.37 (C8), 18.04 (C13), 16.74 (C5-Me), 14.89 (C4-Me).



3. General discussion

Cytochrome P450 enzymes (P450) are present in all domains of life (Nelson, 2011). They are involved in important metabolic processes, including biosynthesis of steroids and fatty acids, drug metabolism and detoxification (Bernhardt, 2006). In recent years, there has been recognition of the P450s potential for the industrial production of bulk chemicals and pharmaceuticals. Therefore, there is an increasing interest in the identification, characterization and optimization of P450 enzymes in order to use them for industrial applications. Previously, the genome of *Bacillus megaterium* strain DSM319 was sequenced. This strain has a 5.1 Mbp genome with 5300 open reading frames (Eppinger et al., 2011). Besides the well-studied CYP102A1 (BM3), three genes encoding for cytochrome P450 enzymes have been reported (CYP106A2, CYP109E1 and CYP109A2) (Eppinger et al., 2011; Brill et al., 2014). The focus of the present work lies on CYP109E1 and CYP109A2.

Concerning the substrate spectrum of CYP109E1 and CYP109A2, it has been previously shown that testosterone and nootkatone are converted by CYP109E1 (Brill, 2013). In addition, the catalytic activity towards vitamin D₃ was described for both enzymes (Brill, 2013). The substrate spectrum of these P450s has been further extended in this work. As described in publication 2.4, CYP109E1 shows a catalytic activity towards statins (compactin, lovastatin and simvastatin) and terpenes (α - and β -ionone, isolongifolen-9-one as well as α - and β - damascone), producing pharmaceutically interesting compounds, such as the LDL-cholesterol-lowering drug pravastatin and different statin drug metabolites as well as valuable terpene derivatives (Putkaradze et al., 2017). In addition, the potential of CYP109A2 as steroid hydroxylase has been investigated. In an *in vitro* reconstituted system containing CYP109A2, diflavin reductase (BmCPR) and ferredoxin (Fdx2), the substrate testosterone was converted into one major (P2) and two minor products (P1 and P3) (**Fig. S1**). The product profile of CYP109A2-dependent testosterone conversion is comparable to that of CYP109E1, showing identical retention times of the major product (P2) and one of the minor products (P3), that have been previously identified by NMR spectroscopy as 16 β -hydroxytestosterone and androstenedione (publication 2.3), respectively. However, the formation of P1 was not detectable in case of CYP109E1 reaction (**Fig. S1**).

Among the identified substrates for both enzymes, our focus was on vitamin D₃. Vitamin D₃ plays a substantial role in human health and, therefore, a variety of diseases are associated with vitamin D₃ deficiency. Because of that, vitamin D₃ is used for the treatment of different

health issues, such as chronic renal failure, different cancers, diabetes and neuromuscular dysfunction. The pharmaceutical relevance of vitamin D₃ is based on its active metabolites, where the most active metabolite is a hydroxylated metabolite at C-25 and C-1 (1 α ,25-dihydroxyvitamin D₃). The chemical synthesis of hydroxylated vitamin D₃ is a challenging, multistep procedure with low yields (Kametani and Furuyama, 1987), limiting their application in industrial production and, therefore, there is extraordinary interest in biotechnological production of vitamin D₃ hydroxylated products.

The usage of mammalian vitamin D₃ hydroxylases in biotechnological production is limited due to their membrane-binding nature, low activity and expression level as well as their instability (Kumar, 2010). Therefore, there is a great interest in bacterial vitamin D₃ hydroxylases. Nevertheless, only a few bacterial P450 enzymes are employed in the bioconversion of vitamin D₃ (Sakaki et al., 2011). The most investigated bacterial P450s acting as vitamin D₃ hydroxylase are CYP105A1 from *Streptomyces griseolus* (Sawada et al., 2004) and CYP107 (known as Vdh) from *Pseudonocardia autotrophica* (Fujii et al., 2009). The results of this work extend the tool box of bacterial vitamin D₃ hydroxylases by a new P450 family, namely CYP109 family. Using protein engineering, the yield of specific products has been further optimized, in particular 25(OH)VD₃. In the next sections, the potential of CYP109E1 and CYP109A2 as vitamin D₃ hydroxylase is discussed and compared with that of the most investigated bacterial P450 vitamin D₃ hydroxylases, CYP105A1 and CYP107.

3.1 Bioconversion of vitamin D₃ by CYP109E1 and CYP109A2

It has been found that both enzymes have similar K_m values, indicating their similar affinity for vitamin D₃. Through comparison with the known bacterial vitamin D₃ hydroxylase (CYP105A1 and Vdh), it becomes clear that CYP109E1 and CYP109A2 have a rather low affinity to vitamin D₃ (**Table 3.1**). On the other hand, the two CYP109 enzymes showed a higher turnover number (k_{cat}), in particular CYP109E1, compared to CYP105A1 and Vdh. As a result, the catalytic efficiencies (k_{cat}/K_m) of CYP109 enzymes towards vitamin D₃ conversion are significantly higher in comparison to CYP105A2 or Vdh (**Table 3.1**), demonstrating the potential of these two P450s as vitamin D₃ hydroxylase.

In terms of the regio-selectivity, our results demonstrate the superiority of CYP109A2 over CYP109E1. It has been shown that CYP109A2 converts vitamin D₃ into two products; the metabolite 25(OH)VD₃ (90%) and one minor product (10%) (publication 2.2). On the other

hand, the product profile of CYP109E1-dependent vitamin D₃ conversion consists of seven different products (publication 2.1). Among the other known bacterial vitamin D₃ hydroxylase (CYP105A1 and CYP107), CYP109A2 shows the highest regio-selectivity towards vitamin D₃ 25-hydroxylation, demonstrating the advantage of this enzyme to act as a vitamin D₃ 25-hydroxylase in biotechnological processes.

By analyzing the product profile of CYP109E1, we were able to confirm the hydroxylation of vitamin D₃ at C-24 and C-25, resulting in the formation of 24S(OH)VD₃ and 25(OH)VD₃ (publication 2.1). Furthermore, it was shown that 24S(OH)VD₃ is produced during the reaction as intermediate, which is further converted by CYP109E1 to produce 24S,25(OH)₂VD₃, confirming the 25-hydroxylation activity of CYP109E1 towards vitamin D₃ as well as 24S(OH)VD₃. To the best of our knowledge the formation of 24S,25(OH)₂VD₃ from vitamin D₃ via 24S(OH)VD₃ is a novel pathway.

Table 3.1. Kinetic parameters of known bacterial vitamin D3 hydroxylase.

CPY	K _m [μM]	k _{cat} [min ⁻¹]	k _{cat} /K _m	Reference
CYP109E1	45.3	23	0.5	(Brill, 2013)
CYP109A2	49	3.4	0.06	Publication 2.2
CYP105A1	0.54	0.016	0.029	(Sawada et al., 2004)
CYP107 (Vdh)	13.5	0.29	0.021	(Fujii et al., 2009)

Although the reactions catalyzed by P450s are diverse, their application for commercial purposes is still limited due to some properties of these enzymes (Bernhardt, 2006; Bernhardt and Urlacher, 2014). The reaction cycle of P450s is dependent on an equimolar amount of the expensive cofactor NAD(P)H and electron carrier proteins. This limitation can be overcome by using of whole-cell systems.

Since many decades, *Bacillus megaterium* has been used in industrial processes. Many of its industrial applications have been reviewed in detail (Vary, 1992, 1994). It offers a great potential as a source and expression host (Vary et al., 2007). In contrast to *E. coli*, *B. megaterium* does not produce endotoxins and, therefore, is a safe host for applications in the pharmaceutical and food industry. In addition, the ability of *B. megaterium* cells to metabolize various cheap carbon sources and to maintain the stability of recombinant

plasmids make them of great interest for biotechnological processes (Biedendieck, 2016; Stammen et al., 2010). Previously, a whole-cell system in *B. megaterium* has been used in our laboratory for the bioconversion of steroid hormones and valuable pharmaceuticals (Bleif et al., 2012; Brill et al., 2014; Ehrhardt et al., 2016; Kiss et al., 2015). In this work, the whole-cell conversions were performed using *B. megaterium* MS941.

The obtained results show clearly that CYP109A2-based whole-cell system using *B. megaterium* is an efficient route for the production of 25(OH)VD₃ in comparison to other known P450-based whole-cell systems (**Table 3.2**). In the literature, only Vdh expressed in *P. autotrophica* was reported to provide a higher 25(OH)VD₃ yield than CYP109A2 being 137 mg/L. However, it should be noted that this value was obtained after 72 h, whereas in case of CYP109A2, 54.9 mg/L were obtained already after 24 h. A comparable yield to that of CYP109A2 was obtained using CYP109E1 mutants, I85A or I85W (**Table 3.2**). More details about these mutants are given in section 'Engineering of CYP109E1'. CYP109E1-based system, in addition, offers the opportunity to produce high amounts of the natural product 24S,25(OH)₂VD₃ (28.6 mg/L/day). Although the biological activity of 24(S),25(OH)₂VD₃ is, in general, significantly below that of its isomer 24R,25(OH)₂VD₃ (Kato et al., 1998), its utility as an antitumor agent is patented by Kureha Chemical Industry CO (Seino et al., 1997). In this context, the production of this metabolite using CYP109E1-based system could be of great interest for the pharmaceutical industry. In addition, the V169A- and I241A-based systems showed an optimization in the production of 24S(OH)VD₃, producing 23 and 25 mg/L/day, respectively (see publication 2.1). It should be noted that 24S(OH)VD₃ can be used as precursor for different active vitamin D₃ metabolites such as 20S,24S(OH)₂VD₃ and 1 α ,20S,24S(OH)₃VD₃, which exhibit a therapeutic potential as antiproliferative agents (Lin et al., 2015).

Table 3.2. The yield of 25(OH)VD₃ achieved by different P450-based whole-cell systems.

CYP	Source	Expression	Yield of 25(OH)VD ₃ [mg/L]	Reaction time [h]	Ref.
CYP109A2	<i>B. megaterium</i>	<i>B. megaterium</i>	54.9	24	Publication 2.2
CYP109E1	<i>B. megaterium</i>	<i>B. megaterium</i>	24.5	24	Publication 2.1
CYP109E1 (I85A)	<i>B. megaterium</i>	<i>B. megaterium</i>	45	24	Publication 2.1
CYP109E1 (I85W)	<i>B. megaterium</i>	<i>B. megaterium</i>	53	24	This work
CYP105A2	<i>P. autotrophica</i>	<i>P. autotrophica</i>	8.3	120	(Sasaki et al., 1992)
CYP105A2	<i>P. autotrophica</i>	<i>S. lividans</i>	20	72	(Kawauchi et al., 1994)
CYP105A1 (R73V/R84A)	<i>S. griseolus</i>	<i>S. lividans</i>	7.8	24	(Hayashi et al., 2010)
CYP107BR1 (Vdh)	<i>P. autotrophica</i>	<i>P. autotrophica</i>	137	72	(Takeda et al., 1994)

Despite the promising results obtained with the whole-cell systems presented in this work, further optimization can be obtained. It should be noticed that no heterologous redox partners were co-expressed with the P450s in *B. megaterium*, leaving only the so far unknown endogenous *B. megaterium* redox proteins to support the activity of CYP109E1 and CYP109A2 (and their mutants). Co-expression of redox partners has been previously shown to increase the production of hydrocortisone in CYP11B1 expressing fission yeast *Schizosaccharomyces pombe* (Hakki et al., 2008).

Therefore, we suggest that the identification and co-expression of the endogenous (or suitable heterologous) redox partners in *B. megaterium* can enhance the conversion of vitamin D₃. Further enhancement can be achieved by feeding the culture with the heme precursor δ -aminolevulinic acid (δ -ALA). In our laboratory, it has been observed that the conversion of progesterone in CYP17 expressing *B. megaterium* is further optimized by addition of δ -ALA (data not published).

3.2. Structural aspects of CYP109E1 and CYP109A2

The crystal structures of CYP109E1 and CYP109A2 have been determined in collaboration with the laboratory of biophysical chemistry at the University of Groningen, which allowed the analysis of the active site geometry, providing the basis for future efforts to investigate these two enzymes (see publications 2.2 and 2.3). The structural comparison showed differences between these two enzymes, especially in functionally relevant regions such as BC loop and access channel, as reported in publication 2.2. The polar BC loop and the restricted access channel of CYP109A2, compared to CYP109E1, might decrease the number of stabilized conformation of the hydrophobic vitamin D₃ in the active site resulting in the experimentally observed higher regio-selectivity of CYP109A2.

Unfortunately, the complex of CYP109E1 and CYP109A2 with vitamin D₃ as substrate has not been obtained. In the absence of the vitamin D₃-bound crystal structure, it remains unknown how is vitamin D₃ binding and which residues interact with in the active site of these two P450s. Molecular docking is a practical alternative for the study of substrate-enzyme complexes, providing significant information on the positioning/orientation and interaction of the substrate within the catalytic pocket. Fortunately, the closed conformation of CYP109E1 was determined in complex with the substrate testosterone (publication 2.3), providing the functionally active form of this enzyme and, therefore, this structure was used for docking experiments. The structurally different substrates (vitamin D₃, compactin and α -ionone) were docked into the active site of CYP109E1. The results obtained thus revealed that vitamin D₃ exhibits a conformation in the CYP109E1 active site similar to the experimentally observed one in CYP107 (Vdh), as described in publication 2.1. The obtained results combined with site-directed mutagenesis enabled us to create CYP109E1 variants with optimized regio-selectivity as well as to identify functionally relevant residues responsible for the catalytic activity and substrate specificity (see publications 2.1, 2.3 and section ‘Engineering of CYP109E1’).

Furthermore, our docking calculations predicted the orientation of compactin and α -ionone in CYP109E1 active site, confirming the experimentally observed regio-selectivity of the enzyme concerning the identified substrates (publication 2.4). In addition, the docking experiments allowed us to predict which active site residues interact with compactin and/or α -ionone, whose functional relevance should be investigated in future attempts.

In case of CYP109A2, the docking experiments were not performed, as the functionally relevant crystal structure (closed conformation) of the enzyme is still unknown. However, the comparison of vitamin D₃ binding residues in the active site of CYP109E1 or Vdh to their structural equivalents in CYP109A2 revealed a significant degree of conservation, supporting the suggestion of overall binding mode of vitamin D₃ suitable for 25-hydroxylation in these three enzymes, as reported in publication 2.2.

The observed structural similarity between CYP109E1 and CYP109A2 gives hints for similar functions and, therefore, it is recommended to investigate if CYP109A2 can also convert the other substrates of CYP109E1 (compactin, lovastatin simvastatin, α - and β -ionone, isolongifolen-9-one as well as α - and β - damascone).

3.3. Engineering of CYP109E1

Some enzyme properties have to be improved before enzymes can be employed in industrial biotransformation. From an economic point of view, the formation of side or undesirable products reduces the added value of the biotechnological process, especially when these undesirable products are less valuable than the target product. Moreover, the separation of the target product from other side products (downstream process) drives the costs up and is both time-consuming and technically difficult. Therefore, enzyme specificity and selectivity are crucial properties that affect the whole process. In general, the usage of P450s is often limited due to their poor regio- and stereo-selectivities, requiring the employment of protein engineering methods to overcome such limitations.

As mentioned above, CYP109E1 catalyzes a multiple hydroxylation of vitamin D₃ producing 7 different products. Among them, our interest was focused on calcidiol (25(OH)VD₃), which is produced with a ratio of 37% of total products. Therefore, our next aim was the creation of mutants with a higher regio-selectivity towards 25-hydroxylation. Using docking experiments and site-directed mutagenesis, we determined the functional importance of selected residues in the active site of the enzyme (described in publications 2.1 and 2.3).

Concerning our target product (25(OH)VD₃), an approximately 2-fold increase was achieved by the mutant I85A, compared to the wild type. This enhancement of 25(OH)VD₃ yield is based on the improvement of the regio-selectivity towards 25-hydroxylation of vitamin D₃ (see publication 2.1). The residue I85 of CYP109E1 is located closest to the heme on the BC loop (SRS1), which is one of the most flexible regions in P450s. Previous studies have shown the critical role of the corresponding position (standard position 87 in P450 amino acid sequences (Gricman et al., 2015)) in terms of enzyme selectivity (Li et al., 2001; Liu et al., 2004). The impact of this position in CYP109E1 was further examined by semi-saturated mutagenesis. In total, 10 different residues were introduced at position 85, creating the following CYP109E1 variants (CYP109E1-I85X, where X refers to: A, L, V, M, W, F, H, E, D or R). Interestingly, it has been found that all substitutions at position 85 yielded an improved regio-selectivity towards vitamin D₃ compared to the wild type enzyme (**Figure 3.1 and Table 3.3**). Hereby, 25(OH)VD₃ was produced as the major product. Furthermore, it has been found that the replacement of the hydrophobic residue isoleucine at position 85 by other hydrophobic residues (W, A, M, L, V, and F), results in comparable or slightly lower conversion of vitamin D₃ compared to the wild type. On the other hand, the introduction of charged residues (R, E, D and H) causes a significant decrease in vitamin D₃ conversion (**Table 3.3**), demonstrating the important role of a hydrophobic side chain at position 85 for vitamin D₃ conversion. It can be supposed that substitution at this position causes conformational changes, promoting a higher regio-selectivity towards vitamin D₃ 25-hydroxylation and, therefore, the elucidation of the 3D-structures of CYP109E1 mutants is of great interest. In particular I85W mutant, which shows the highest regio-selectivity towards 25-hydroxylation (81%) (**Table 3.3**).

In addition, it has been shown that the active site residues V169 (SRS2) and I241 (SRS4) are substrate-specificity determining residues as their substitution with alanine led to an abolishment of enzyme activity towards testosterone, while the activity towards vitamin D₃ is still maintained (publication 2.1 and 2.2). Vitamin D₃ conversion by these two mutants has been, however, changed compared to the wild type enzyme. It was observed that these two mutants, in contrast to the wild type, show a significant decrease in 25-hydroxylation activity towards the intermediate product 24S(OH)VD₃, resulting in the accumulation of this metabolite, in particular when using mutant V169A. Furthermore, differences between these two mutants have been observed in terms of their regio-selectivity.

While the mutant V169A showed the same ratio for 24- and 25-hydroxylation, the 24-hydroxylation of vitamin D₃ was preferred in the reaction of the mutant I241A (publication 2.1). These results indicate the functional key role of these two active site residues and, therefore, further investigations of position 169 and 241 in CYP109E1 are highly recommended.

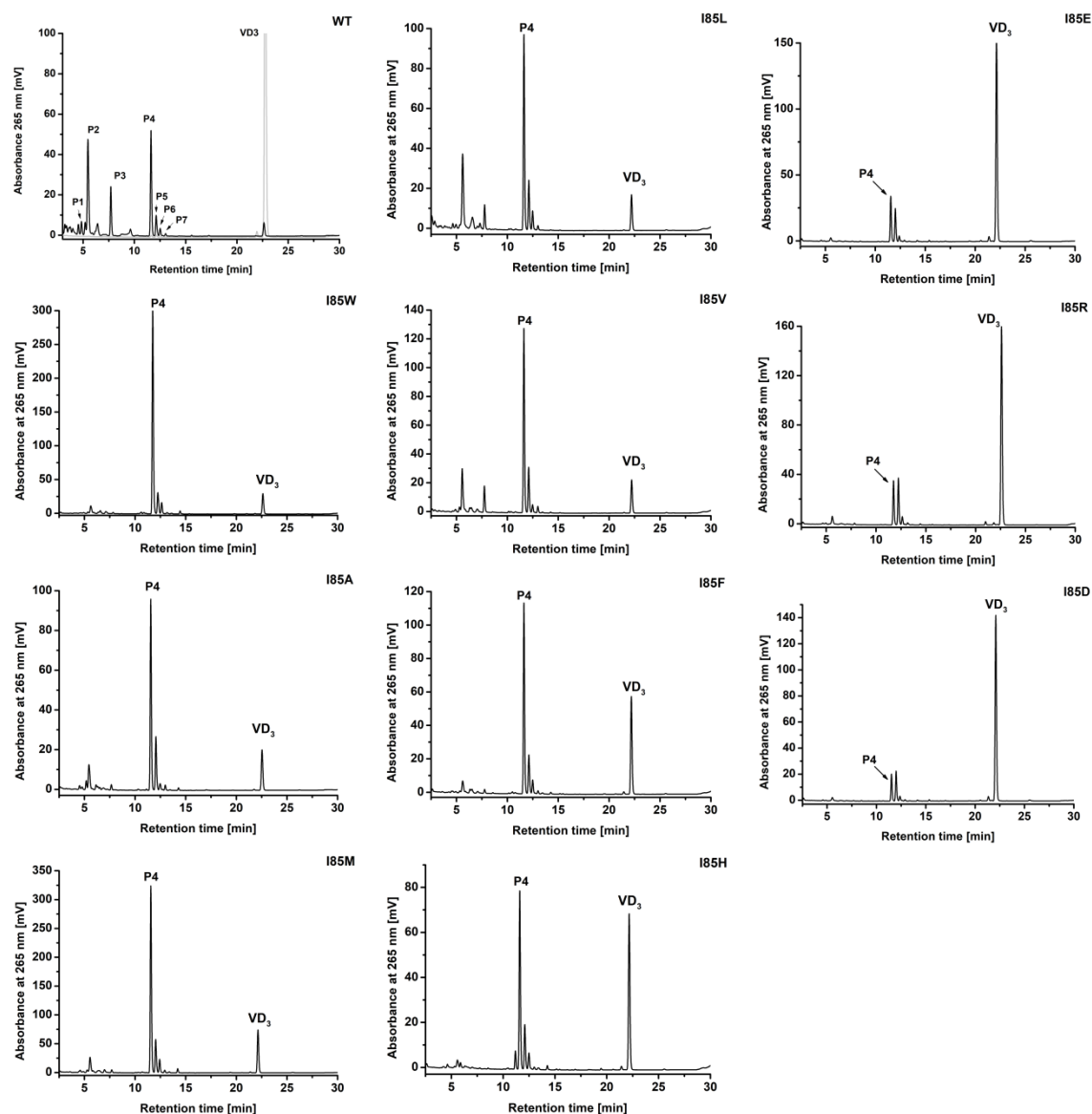


Figure 3.1. HPLC chromatograms of 24 h whole-cell conversion of vitamin D₃ by CYP109E1 and its mutants with one point mutation at position 85. The reactions were performed in *B. megaterium* MS941 overexpressing the corresponding P450 variants. Vitamin D₃ and its derivatives are labeled here by VD₃ and P1-P7, respectively. P4 refers here to as 25(OH)VD₃.

Table 3.3. Vitamin D₃ conversion by different variant of CYP109E1.

Enzyme variant	Conversion [%]	Selectivity towards 25-hydroxylation [%]
WT	95 ± 3	36± 2
I85W	92 ± 2	81± 3
I85A	87 ± 4	64± 2
I85M	83 ± 4	73± 1
I85L	90 ± 2	48± 4
I85V	89 ± 3	57± 1
I85F	69 ± 5	71± 4
I85H	59 ± 7	64± 1
I85E	27 ± 4	48± 3
I85R	31 ± 6	40± 4
I85D	23 ± 8	40± 3

3.4. Engineering of CYP109A2

Previously, CYP109A2 was cloned and expressed based on an open reading frame (ORF) producing a protein of 426 amino acids (Brill, 2013). Using the genome data and depending on ribosomal binding site (RBS) analysis, it has been found in our recent studies that the correct start codon of CYP109A2 is located 69 nucleotides downstream the previously used start codon (Brill, 2013). The new open reading frame was cloned and expressed in *E. coli* leading to a protein with 403 amino acids. Spectral characterization of the purified “truncated” form of CYP109A2 (see publication 2.2) showed the characteristic peak at 450 nm identical with the previously identified form of CYP109A2. The activity of the “truncated” form towards vitamin D₃ was compared with that of the full-length form. Interestingly, a higher *in vitro* conversion of vitamin D₃ (36 ± 4 %) was reached with the new form of CYP109A2 (see publication 2.2), which is significantly higher than the conversion using the previously investigated form of the enzyme (conversion of <5%) (Brill, 2013). It should be mentioned that the previously (as full-length) identified form, in contrast to the new form (truncated), was not able to convert testosterone (Brill, 2013). Therefore, further

investigations of this P450 were performed with the “truncated” form, which is for the simplicity named as CYP109A2.

CYP109A2 was studied by X-ray crystallography providing the typical organization of P450 domains (Publication 2.2). Unfortunately, only the substrate-free structure of CYP109A2 was obtained so far. Hereby, structural comparison can be useful to predict the substrate binding residues and functionally relevant regions. As reported in publication 2.2, CYP109A2 shows the highest structural similarity with the open form of CYP109E1 (PDB: 5L91) and the structure of Vdh from *P. autotrophica* (PDB: 3A4G). It should be membered that the conversion of vitamin D₃ by CYP109A2 is lower than that of CYP109E1 and CYP107. Through structural comparison, differences in the amino acid composition were observed in the active sites of these three enzymes, which lead to different charge distributions in the corresponding active sites. In comparison to CYP109E1 and CYP107, the charged amino acids R74, R78 and E81 (BC loop) of CYP109A2 form a highly polar surface, providing no optimal environment for the binding of highly hydrophobic substrates like vitamin D₃. Therefore, these charged amino acids in CYP109A2 were mutated in order to get a less polar BC loop (more hydrophobic). R74V, R78V and E81L mutants have been created and the whole-cell conversion of vitamin D₃ with these mutants was performed. Unfortunately, the achieved conversion of vitamin D₃ was not as good as expected. The E78V and E81L mutants showed a decreased conversion (44% and 52%) compared to wild type (75%). More interesting is the almost complete abolishment of the catalytic activity of the R74V mutant towards vitamin D₃ (below 15% of added substrate were converted after 24h), suggesting the importance of the R74 residue for the catalytic activity towards vitamin D₃ (**Figure S2**). Conversely, the whole-cell conversion of testosterone using the mutant R74V was comparable with that of the wild type enzyme. Hence, further investigations of this mutant were performed. The R74V mutant was successfully expressed and purified in *E. coli*, showing the characteristic peak at 450 nm. In an *in vitro* reconstituted system containing the mutant R74V, the diflavin reductase BmCPR and the ferredoxin Fdx2, no activity towards testosterone or vitamin D₃ was observed (data not shown). Since a sufficient sequential delivery of electrons from NAD(P)H via redox partners is needed for the catalytic activity of P450 enzymes, we aimed to examine the functional interaction between the electron transfer carriers and this mutant by recording the NADPH reduced CO-complex at peak 450 nm, as described previously (Milhim et al., 2016).

While CYP109A2 wild type produced a typical reduced CO spectrum with a Soret peak at 450 nm, the mutant R74V did not show this Soret peak, demonstrating the disruption of the electron transfer to the mutant enzyme.

This result explains the missing *in vitro* activity of R74V mutant when BmCPR and Fdx2 are used as electron transfer proteins. It is known that R74 is not a part of the redox partner-binding site of CYP109A2. However, we suggest that on one site the substitution of this residue might cause conformational changes in the whole structure that may affect the interaction between electron transfer protein and CYP109A2. On the other hand, it is also possible that the redox potential of this mutant has been changed compared to that of the wild type and, therefore, the electron transfer from the redox partner to this mutant becomes thermodynamically unfavorable. Summarizing, these results suggest that the mutant R74V strictly needs its endogenous redox partner to perform its catalytic activity, while this requirement is relaxed for the wild type enzyme. It should be noted that the observed *in vivo* activities of CYP109A2 and its mutants, including R74V mutant, are supported by the endogenous redox partners, which are still unknown.

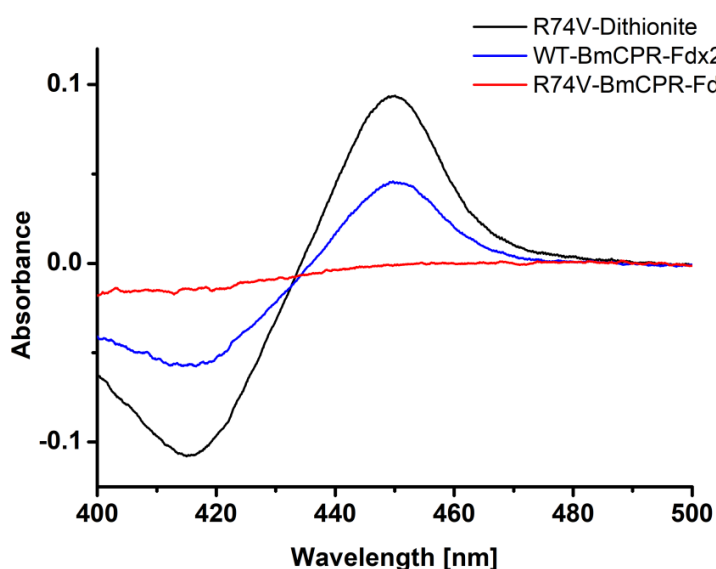


Figure 3.2. Reduction of CYP109A2 wild type (WT) and its mutant R74V using the autologous electron transfer protein BmCPR and Fdx2. NADPH reduced CO-difference spectra of CYP109A2 wild type (in blue) and R74V mutant (in red). The dithionite reduced CO-difference spectrum of the mutant R74V (black). The NADPH (1 mM) reduced CO-difference spectra were recorded in a 400 μ L mixture of CYP/ferredoxin/reductase with a 1:20:3 molar ratio in 50 mM potassium phosphate buffer pH 7.4 containing 20% glycerol.

Further attempts have been made in this work to increase vitamin D₃ conversion by CYP109A2. It has been previously shown that a single mutation at the ferredoxin binding site of Vdh results in the highly active mutant T107A (Yasutake et al., 2013). The residue T103 of CYP109A2 was identified as corresponding residue to T107 of Vdh and, therefore, we aimed to investigate the role of T103 in CYP109A2. Our results showed that the substitution T103A lead to a slight increase of vitamin D₃ conversion (85%), compared to wild type (75%) (**Figure 3.3**). Furthermore, the role of the neighboring residues of T103 (P104, R105 and R106) were tested, as it was observed that both CYP109A2 and Vdh have a certain similarity in this region. P104A, R105A, R105K and A106R mutants of CYP109A2 were created and tested in terms of their activity towards vitamin D₃. Our results showed that the substitution P104A does not affect the activity of CYP109A2 towards vitamin D₃, while the substitution R105A led to a decrease in vitamin D₃ conversion. Interestingly, the substitution of R105 with the positively charged residue lysine showed the same conversion as the wild type enzyme, indicating the relevance of positive charged residue at position of 105 in CYP109A2. In contrary, replacement of A106 with arginine led to a decrease in the vitamin D₃ conversion. Based on these results, we suggest that the positions 103-106 in CYP109A2 form a functionally important region, which should be investigated in more details using single and a combination of two or more mutations. Previously, it has been shown that double and triple mutants of P450s (for example in case of Vdh and CYP105A1) showed a collective effect on enzymatic activity compared to single mutants (Fujii et al., 2009; Hayashi et al., 2010).

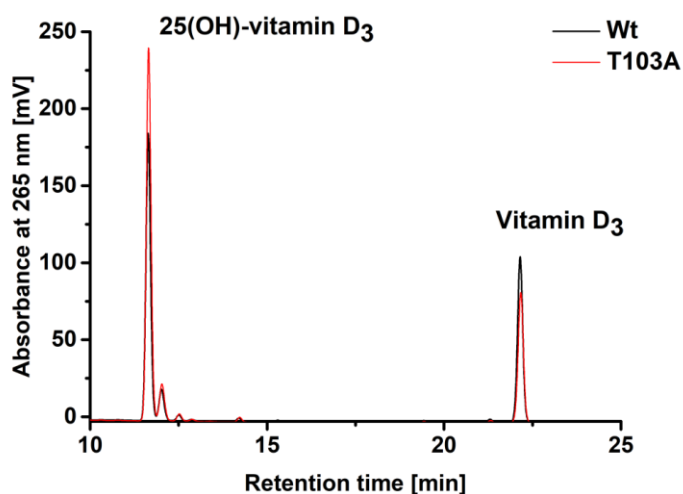


Figure 3.3. HPLC-chromatogram of vitamin D₃ conversion in *B. megaterium* MS941 overexpressing the wild type of CYP109A2 (WT) and its mutant T103A. The reactions were performed as described in publication 2.1 and 2.2.

Taken together, the high potential of CYP109E1 and CYP109A2 as vitamin D₃ hydroxylases has been shown in this work. In comparison to the known whole-cell systems used for the bioconversion of vitamin D₃, *B. megaterium* cells overexpressing CYP109E1 or CYP109A2 provide a high promising route to produce the valuable metabolite 25(OH)VD₃. The elucidation of both, CYP109E1 and CYP109A2, crystal structures offers the first insights into the geometry of these two P450s. Furthermore, protein engineering has been successfully applied in this work to optimize the regio-selectivity and/or activity of CYP109E1 and CYP109A2. Besides the generation of enzyme mutants with improved properties compared to the wild type, the functional role of some residues in CYP109E1 and CYP109A2 has been determined using site-directed mutagenesis.

4. Outlook

During the last years, the enzymatic oxyfunctionalization of unactivated carbons in organic compounds has gained interest, since chemical synthesis often consists of multistep-procedures with environmentally unfriendly conditions. P450 monooxygenases can be the ideal biocatalysts for interesting industrial reactions and, therefore, the identification and characterization of novel P450 enzymes was and is still a main target for scientist.

This work demonstrates the optimization and application of two members of CYP109 family for the conversion of vitamin D₃. The obtained results demonstrate clearly the promising potential of CYP109A2- and CYP109E1-based whole-cell systems, providing a competitive offer for the production of 25(OH)VD₃. Further enhancement of the system's productivity, however, has to be achieved by optimization of the electron transfer, substrate solubility and substrate uptake. Hereby, site-directed mutagenesis performed in this work offers starting points for further mutations in order to optimize the redox-binding sites, in particular for CYP109A2 as the surface residues R105 and R106 seem to be implicated in the catalytic activity of this P450. In addition, the identification of the natural redox partner proteins of CYP109E1 and CYP109A2 is highly important and meaningful. Another point that can be further optimized is the poor solubility of vitamin D₃ in aqueous solutions. Hereby, cyclodextrins have been shown to enhance enzymatic conversions of lipophilic substrates due to its lipophilic interior and hydrophilic outer surface (Harper et al., 2000). Previously, the Enhancement of androstadienedione production from progesterone was achieved by complexation of progesterone with 2-hydroxypropyl-β-cyclodextrin (HPCD) (Manosroi et al., 2008). Although HPCD has been successfully used in the present work to increase the solubility of the hydrophobic vitamin D₃, previous studies have shown that the usage of partially-methylated-β-cyclodextrin (PMCD) is superior among other cyclodextrins derivatives for the bioconversion of vitamin D₃ (Takeda et al., 1994). In this context, the effect of PMCD on vitamin D₃ conversion should be investigated in future efforts with the established whole-cell system. In addition, to overcome the cellular membrane boundaries is an important problem that should be considered. To increase cell membrane permeability, the permeabilizing detergent saponin was used in this work to perform vitamin D₃ *in vivo* conversion. However, other permeabilizing agents should be tested, too. Hereby, it has been reported that the antibiotic Nisin enhances vitamin D₃ bioconversion in *Rhodococcus*

erythropolis through pore formation in the membrane and the disruption of cell wall biosynthesis (Imoto et al., 2011).

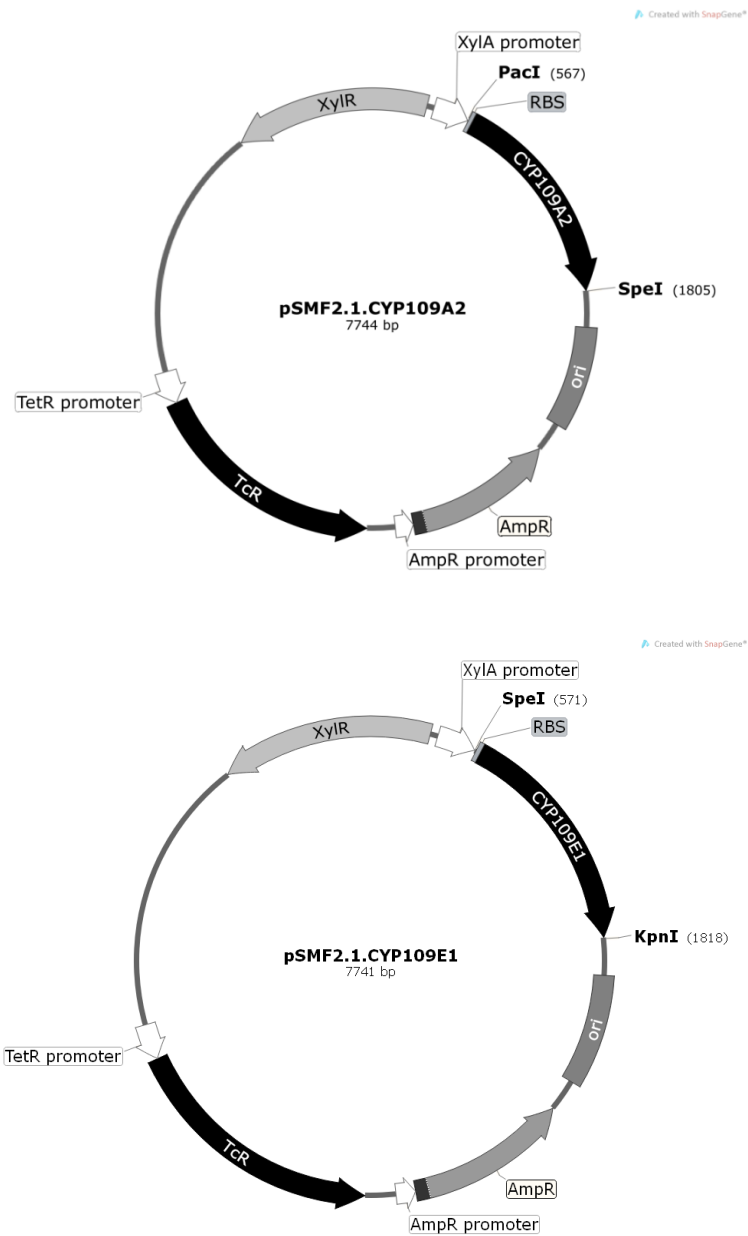
The elucidation of the crystal structure of CYP109E1 in its closed conformation laid the basics for docking experiments and the subsequent site-directed mutagenesis investigations, which led to the creation of mutants with optimized regio-selectivity towards vitamin D₃ 25-hydroxylation. Since only the open form of CYP109A2 was obtained, it is highly recommended to crystallize this P450 in its functionally active form (closed form), if possible in complex with vitamin D₃. Moreover, the interesting mutants of CYP109E1 (I85W) and CYP109A2 (R74V) deserve to be examined by X-ray protein crystallography. We believe that such structural information will improve our understanding for both, CYP109A2 and CYP109E1, and guide their further optimization processes.

We were able to identify two novel vitamin D₃ metabolites (24S(OH)VD₃ and 24S,25(OH)₂VD₃), whose physiological function has not been fully clarified and, therefore, relevant laboratory experiments should be performed to explore their physiological relevance.

Finally, we suggest that the substrate spectrum of both enzymes, in particular of CYP109A2, can be further expanded and, thus, it would be beneficial to apply high-throughput methods to identify new substrates such as drugs and fatty acids. Some mutants of CYP109E1 and CYP109A2 show different substrate specificity compared to the corresponding wild type, hence, these mutants should be also taken into account for substrate screening.

5. Appendix

5.1. Plasmids



5.2. Proteins

> CYP109E1 (BMD_3874)|

MKTERENGIVRQVNTIQTKEERFNPFSWYEEMRNTAPVQWDEERQVWDVFHYDGVKEVLEQKNIFS
SDRRPPQNQRQTALGTSLINIDPPKHAEMRALVNKAFTPKAMKAWEPKIARITNELLQEVEHLEDIDIV
EHLSPVPLPVMVIADILGVPIEDQRQFKDWSDIIVAGPSNNERETLEKLQQEKMKANDELETYFYRIIEE
KRTRPGDDIISVLLQAKEEGKQLTDEEIVGFSILLIAGNETTTNLISNTIYCLMEDKASFERLKREKELL
PSGIEEVLRYRSPVQALHRIVKEDVTLAGKKLKAGEHVVPWMGSAHRDAEYFEDPEVFKIDRKPNVH
MAFGRGIHFCLGAPLARIEAKIMLAELIDRYPQMDWSPSFELKPIESTFVYGLKELLIRKNV

> CYP109A2 (BMD_2035)

MNPKAVKRENRYANLIPMQEIKSVEQQLYPFDIYNLSRQEAPIRYDESRNCWDVFDYETVKYILKNPS
LFSSKRAMEERQESILMMDPPKHTKLRNLVNKAFTPRAIQHLEGHIEEIADYLLDEVSSKEKFDIVEDF
AGPLPIIVIAELLGVPIQDRALFKKYSDDLVSGAENNSDEAFKMMQKRNEGVIQGYFKEIIAERQQ
NKQEDLISLLLEAEIDGEHLTEEEVLGFCILLVAGNETTTNLITNGVRYMTEDVDVQNEVRRDISLVP
NLVEETLRYYPPIQAIGRIAAEDVELGECKIKRGQQVISWAASANRDSAKFEWPDFTFVHRKTNPHVS
FGFGIHFCLGAPLARMEGKIAFTKLEKGGFSKVQNQSLKPIDSPFVFGVKKYEIAFNNA

5.3. Figures

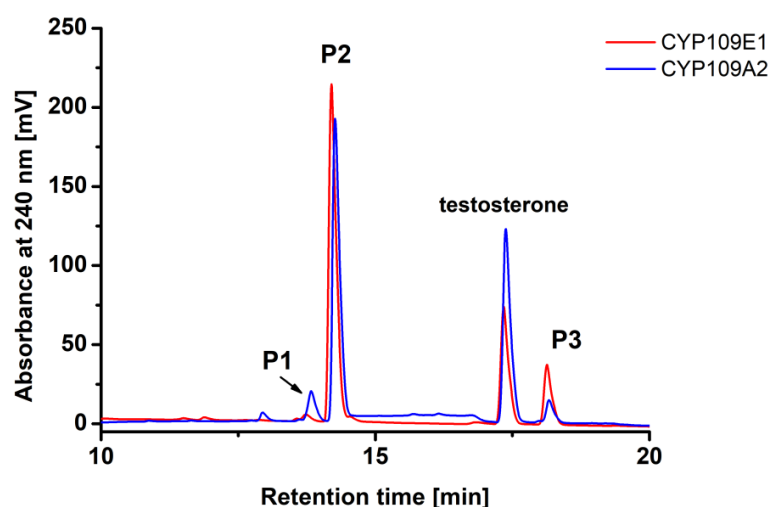


Figure S1. HPLC chromatogram of testosterone conversion by CYP109A2 compared to CYP109E1-dependent conversion. Diflavin reductase (BmCPR) and ferredoxin (Fdx2) were used as redox partners. Testosterone and P450s were added in end concentrations of 200 μ M and 1 μ M, respectively. The reactions were performed in potassium phosphate buffer at 30 $^{\circ}$ C for 1h.

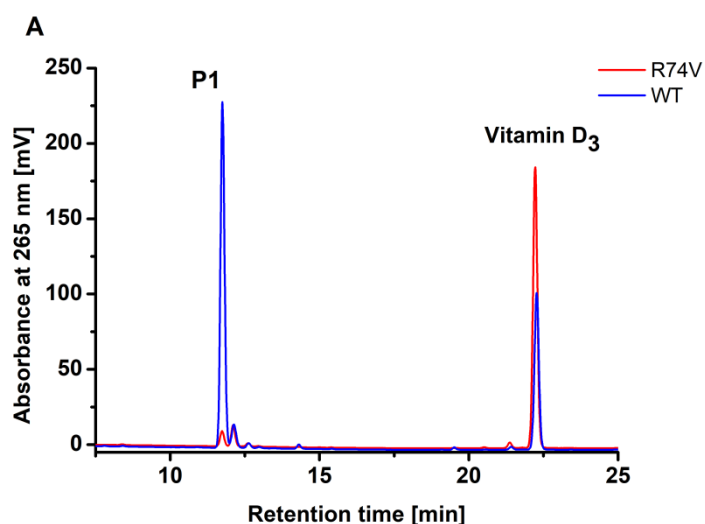


Figure S2. HPLC chromatogram of vitamin D₃ whole-cell conversion using *B. megaterium* cells overexpressing CYP109A2 wild type and CYP109A2 mutant R74V.

6. Abbreviations

AdR	Adrenodoxin reductase
Adx	Adrenodoxin
<i>B. megaterium</i>	<i>Bacillus megaterium</i>
BmCPR	<i>Bacillus megaterium</i> cytochrome P450 reductase
CYP/P450	Cytochrome P450
δ -aminolevulinic acid	δ -ALA
<i>E. coli</i>	<i>Escherichia coli</i>
FAD	Flavine adenine dinucleotide
Fdx2	Ferredoxin 2 from <i>B. megaterium</i> DSM319
FMN	Flavine mononucleotide
HPLC	High performance liquid chromatography
HPCD	2-hydroxypropyl- β -cyclodextrin
L	Liter
M	Molar
mg	Milligram
μ M	Micromolar
NAD(P)H	Nicotinamide adenine dinucleotide (phosphate)
ORF	Open reading frame
<i>P. autotrophica</i>	<i>Pseudonocardia autotrophica</i>
<i>S. griseolus</i>	<i>Streptomyces griseolus</i>
<i>S. lividans</i>	<i>Streptomyces lividans</i>
SRS	Substrate recognition site
Vdh	Vitamin D ₃ hydroxylase CYP107 from <i>P. autotrophica</i>
VD ₃	Vitamin D ₃
VDREs	Vitamin D ₃ response elements
25(OH)VD ₃	25-hydroxyvitamin D ₃
1 α ,25(OH) ₂ VD ₃	1 α ,25-dihydroxyvitamin D ₃
UV/Vis	Ultraviolet–visible spectroscopy
WT	Wild type

7. References

- Abdulgugni, A., Józwik, I.K., Putkaradze, N., Brill, E., Zapp, J., Thunnissen, A.-M.W.H., Hannemann, F., Bernhardt, R., 2017a. Characterization of cytochrome P450 CYP109E1 from *Bacillus megaterium* as a novel vitamin D₃ hydroxylase. *J. Biotechnol.* 243, 38–47.
- Abdulgugni, A., Józwik, I.K., Brill, E., Hannemann, F., Thunnissen, A.-M.W.H., Bernhardt, R., 2017b. Biochemical and structural characterization of CYP109A2, a vitamin D₃ 25-hydroxylase from *Bacillus megaterium*. *FEBS J.* 284, 3881–3894.
- Abe, E., Miyaura, C., Sakagami, H., Takeda, M., Konno, K., Yamazaki, T., Yoshiki, S., Suda, T., 1981. Differentiation of mouse myeloid leukemia cells induced by 1 α ,25-dihydroxyvitamin D₃. *Proc. Natl. Acad. Sci. U. S. A.* 78, 4990–4994.
- Arnold, F.H., 2017. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem. Int. Ed Engl.*
- Barletta, F., Freedman, L.P., Christakos, S., 2002. Enhancement of VDR-mediated transcription by phosphorylation: correlation with increased interaction between the VDR and DRIP205, a subunit of the VDR-interacting protein coactivator complex. *Mol. Endocrinol. Baltim. Md* 16, 301–314.
- Bernhardt, R., 1996. Cytochrome P450: structure, function, and generation of reactive oxygen species. *Rev. Physiol. Biochem. Pharmacol.* 127, 137–221.
- Bernhardt, R., 2006. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* 124, 128–145.
- Bernhardt, R., Urlacher, V.B., 2014. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl. Microbiol. Biotechnol.* 98, 6185–6203.
- Biedendieck, R., 2016. A *Bacillus megaterium* System for the Production of Recombinant Proteins and Protein Complexes. *Adv. Exp. Med. Biol.* 896, 97–113.
- Bischoff-Ferrari, H.A., Giovannucci, E., Willett, W.C., Dietrich, T., Dawson-Hughes, B., 2006. Estimation of optimal serum concentrations of 25-hydroxyvitamin D for multiple health outcomes. *Am. J. Clin. Nutr.* 84, 18–28.
- 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.
- Böttcher, D., Bornscheuer, U.T., 2010. Protein engineering of microbial enzymes. *Curr. Opin. Microbiol., Ecology and industrial microbiology • Special section: Systems biology* 13, 274–282.
- Brill, E., 2013. Identifizierung und Charakterisierung neuer Cytochrom P450 Systeme aus *Bacillus megaterium* DSM319. Universität des Saarlandes.
- Brill, E., Hannemann, F., Zapp, J., Brüning, G., Jauch, J., Bernhardt, R., 2014. A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto- β -boswellic acid (KBA). *Appl. Microbiol. Biotechnol.* 98, 1701–1717.
- Brown, A.J., Dusso, A.S., Slatopolsky, E., 2002. Renal failure and vitamin D. *Clin. Calcium* 12, 711–723.
- Buchholz, K., Kasche, V., Bornscheuer, U.T., 2005. Biocatalysts and enzyme technology. Wiley-VCH, Weinheim.
- Calvo, M.S., Whiting, S.J., Barton, C.N., 2005. Vitamin D intake: a global perspective of current status. *J. Nutr.* 135, 310–316.
- Cheng, J.B., Levine, M.A., Bell, N.H., Mangelsdorf, D.J., Russell, D.W., 2004. Genetic evidence that the human CYP2R1 enzyme is a key vitamin D 25-hydroxylase. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7711–7715.

- Declerck, N., Machius, M., Wiegand, G., Huber, R., Gaillardin, C., 2000. Probing structural determinants specifying high thermostability in *Bacillus licheniformis* α -amylase. Edited by A. R. Fersht. *J. Mol. Biol.* 301, 1041–1057.
- DeLuca, H.F., 2004. Overview of general physiologic features and functions of vitamin D. *Am. J. Clin. Nutr.* 80, 1689S–1696S.
- DeLuca, H.F., Cantorna, M.T., 2001. Vitamin D: its role and uses in immunology. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 15, 2579–2585.
- Denisov, I.G., Makris, T.M., Sligar, S.G., Schlichting, I., 2005. Structure and chemistry of cytochrome P450. *Chem. Rev.* 105, 2253–2277.
- Di Rosa, M., Malaguarnera, M., Nicoletti, F., Malaguarnera, L., 2011. Vitamin D3: a helpful immunomodulator. *Immunology* 134, 123–139.
- Duso, A., Brown A.J., Slatopolsky E., 2004. Vitamin D. - PubMed - NCBI. *Am. Physiol. Soc.* 289.
- 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.
- Ener, M.E., Lee, Y.-T., Winkler, J.R., Gray, H.B., Cheruzel, L., 2010. Photooxidation of cytochrome P450-BM3. *Proc. Natl. Acad. Sci.* 107, 18783–18786.
- 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.
- Fujii, Y., Kabumoto, H., Nishimura, K., Fujii, T., Yanai, S., Takeda, K., Tamura, N., Arisawa, A., Tamura, T., 2009. Purification, characterization, and directed evolution study of a vitamin D3 hydroxylase from *Pseudonocardia autotrophica*. *Biochem. Biophys. Res. Commun.* 385, 170–175.
- Gamble, M.J., Freedman, L.P., 2002. A coactivator code for transcription. *Trends Biochem. Sci.* 27, 165–167.
- Garfinkel, D., 1958. Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. *Arch. Biochem. Biophys.* 77, 493–509.
- Gillam, E.M.J., 2008. Engineering cytochrome p450 enzymes. *Chem. Res. Toxicol.* 21, 220–231.
- 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, S.E., Peterson, J.A., 1999. How similar are P450s and what can their differences teach us? *Arch. Biochem. Biophys.* 369, 24–29.
- Grant, W.B., 2002. An estimate of premature cancer mortality in the U.S. due to inadequate doses of solar ultraviolet-B radiation. *Cancer* 94, 1867–1875.
- Gricman, Ł., Vogel, C., Pleiss, J., 2015. Identification of universal selectivity-determining positions in cytochrome P450 monooxygenases by systematic sequence-based literature mining. *Proteins Struct. Funct. Bioinforma.* 83, 1593–1603.
- Groves, J.T., 2006. High-valent iron in chemical and biological oxidations. *J. Inorg. Biochem.* 100, 434–447.
- Guengerich, F.P., Munro, A.W., 2013. Unusual cytochrome p450 enzymes and reactions. *J. Biol. Chem.* 288, 17065–17073.

- 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 1770, 330–344.
- Harper, J.B., Easton, C.J., Lincoln, S.F., 2000. Cyclodextrins to Increase the Utility of Enzymes in Organic Synthesis | BenthamScience. Curr. Org. Chem. 4, 429–454.
- Haussler, M.R., Whitfield, G.K., Haussler, C.A., Hsieh, J.C., Thompson, P.D., Selznick, S.H., Dominguez, C.E., Jurutka, P.W., 1998. The nuclear vitamin D receptor: biological and molecular regulatory properties revealed. J. Bone Miner. Res. 13, 325–349.
- Hayashi, K., Sugimoto, H., Shinkyo, R., Yamada, M., Ikeda, S., Ikushiro, S., Kamakura, M., Shiro, Y., Sakaki, T., 2008. Structure-based design of a highly active vitamin D hydroxylase from *Streptomyces griseolus* CYP105A1. Biochemistry (Mosc.) 47, 11964–11972.
- Hayashi, K., Yasuda, K., Sugimoto, H., Ikushiro, S., Kamakura, M., Kittaka, A., Horst, R.L., Chen, T.C., Ohta, M., Shiro, Y., Sakaki, T., 2010. Three-step hydroxylation of vitamin D3 by a genetically engineered CYP105A1. FEBS J. 277, 3999–4009.
- Hazel M Girvan & Andrew W Munro, 2016. Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology 31, 136–145.
- Heaney, R.P., Dowell, M.S., Hale, C.A., Bendich, A., 2003. Calcium absorption varies within the reference range for serum 25-hydroxyvitamin D. J. Am. Coll. Nutr. 22, 142–146.
- Hebeda, R.E., Styrlund, C.R., Teague, W.M., 1988. Benefits of *Bacillus megaterium* Amylase in Dextrose Production. Starch - Stärke 40, 33–36.
- Holick, M.F., 2007. Vitamin D deficiency. N. Engl. J. Med. 357, 266–281.
- Holick, M.F., 2004. Vitamin D: importance in the prevention of cancers, type 1 diabetes, heart disease, and osteoporosis. Am. J. Clin. Nutr. 79, 362–371.
- Holick, M.F., Binkley, N.C., Bischoff-Ferrari, H.A., Gordon, C.M., Hanley, D.A., Heaney, R.P., Murad, M.H., Weaver, C.M., Endocrine Society, 2011. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. J. Clin. Endocrinol. Metab. 96, 1911–1930.
- Holick, M.F., DeLuca, H.F., Avioli, L.V., 1972. Isolation and identification of 25-hydroxycholecalciferol from human plasma. Arch. Intern. Med. 129, 56–61.
- Holick, M.F., Frommer, J.E., McNeill, S.C., Richtand, N.M., Henley, J.W., Potts, J.T., 1977. Photometabolism of 7-dehydrocholesterol to previtamin D3 in skin. Biochem. Biophys. Res. Commun. 76, 107–114.
- Holick, M.F., MacLaughlin, J.A., Clark, M.B., Holick, S.A., Potts, J.T., Anderson, R.R., Blank, I.H., Parrish, J.A., Elias, P., 1980. Photosynthesis of previtamin D3 in human skin and the physiologic consequences. Science 210, 203–205.
- Hollmann, R., Deckwer, W.-D., 2004. Pyruvate formation and suppression in recombinant *Bacillus megaterium* cultivation. J. Biotechnol. 111, 89–96.
- Horsting, M., DeLuca, H.F., 1969. In vitro production of 25-hydroxycholecalciferol. Biochem. Biophys. Res. Commun. 36, 251–256.
- Hosseini-nezhad, A., Holick, M.F., 2013. Vitamin D for Health: A Global Perspective. Mayo Clin. Proc. Mayo Clin. 88, 720–755.
- Illanes, A., Cauerhff, A., Wilson, L., Castro, G.R., 2012. Recent trends in biocatalysis engineering. Bioresour. Technol., Biocatalysis 115, 48–57.
- Imai, M., Shimada, H., Watanabe, Y., Matsushima-Hibiya, Y., Makino, R., Koga, H., Horiuchi, T., Ishimura, Y., 1989. Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: possible role of the hydroxy amino acid in oxygen activation. Proc. Natl. Acad. Sci. U. S. A. 86, 7823–7827.

- Imoto, N., Nishioka, T., Tamura, T., 2011. Permeabilization induced by lipid II-targeting lantibiotic nisin and its effect on the bioconversion of vitamin D₃ to 25-hydroxyvitamin D₃ by *Rhodococcus erythropolis*. *Biochem. Biophys. Res. Commun.* 405, 393–398.
- Jean, G., Terrat, J.-C., Vanel, T., Hurot, J.-M., Lorriaux, C., Mayor, B., Chazot, C., 2008. Daily oral 25-hydroxycholecalciferol supplementation for vitamin D deficiency in haemodialysis patients: effects on mineral metabolism and bone markers. *Nephrol. Dial. Transplant.* 23, 3670–3676.
- Julsing, M.K., Cornelissen, S., Bühler, B., Schmid, A., 2008. Heme-iron oxygenases: powerful industrial biocatalysts? *Curr. Opin. Chem. Biol.* 12, 177–186.
- Kametani, T., Furuyama, H., 1987. Synthesis of vitamin D₃ and related compounds. *Med. Res. Rev.* 7, 147–171.
- Kawauchi, H., Sasaki, J., Adachi, T., Hanada, K., Beppu, T., Horinouchi, S., 1994. Cloning and nucleotide sequence of a bacterial cytochrome P-450VD25 gene encoding vitamin D-3 25-hydroxylase. *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* 1219, 179–183.
- 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, 2320–2326.
- Kittsteiner-Eberle, R., Ogbomo, I., Schmidt, H.-L., 1989. Biosensing devices for the semi-automated control of dehydrogenase substrates in fermentations. *Biosensors* 4, 75–85.
- Klingenberg, M., 1958. Pigments of rat liver microsomes. *Arch. Biochem. Biophys.* 75, 376–386.
- Kumar, S., 2010. Engineering Cytochrome P450 Biocatalysts for Biotechnology, Medicine, and Bioremediation. *Expert Opin. Drug Metab. Toxicol.* 6, 115–131.
- Leichtmann, G.A., Bengoa, J.M., Bolt, M.J., Sitrin, M.D., 1991. Intestinal absorption of cholecalciferol and 25-hydroxycholecalciferol in patients with both Crohn's disease and intestinal resection. *Am. J. Clin. Nutr.* 54, 548–552.
- Li, H., Poulos, T.L., 1997. The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nat. Struct. Biol.* 4, 140–146.
- Li, Q.S., Ogawa, J., Schmid, R.D., Shimizu, S., 2001. Residue size at position 87 of cytochrome P450 BM-3 determines its stereoselectivity in propylbenzene and 3-chlorostyrene oxidation. *FEBS Lett.* 508, 249–252.
- Li, Y., Cirino, P.C., 2014. Recent advances in engineering proteins for biocatalysis. *Biotechnol. Bioeng.* 111, 1273–1287.
- Li, Y.C., 2003. Vitamin D regulation of the renin-angiotensin system. *J. Cell. Biochem.* 88, 327–331.
- Lin, Z., Marepally, S.R., Ma, D., Myers, L.K., Postlethwaite, A.E., Tuckey, R.C., Cheng, C.Y.S., Kim, T.-K., Yue, J., Slominski, A.T., Miller, D.D., Li, W., 2015. Chemical Synthesis and Biological Activities of 20S,24S/R-Dihydroxyvitamin D₃ Epimers and Their 1 α -Hydroxyl Derivatives. *J. Med. Chem.* 58, 7881–7887.
- Liu, J., Ericksen, S.S., Sivaneri, M., Besspiata, D., Fisher, C.W., Szklarz, G.D., 2004. The effect of reciprocal active site mutations in human cytochromes P450 1A1 and 1A2 on alkoxyresorufin metabolism. *Arch. Biochem. Biophys.* 424, 33–43.
- Lu, A.Y., Junk, K.W., Coon, M.J., 1969. Resolution of the cytochrome P-450-containing omega-hydroxylation system of liver microsomes into three components. *J. Biol. Chem.* 244, 3714–3721.
- Manosroi, A., Saowakhon, S., Manosroi, J., 2008. Enhancement of androstadienedione production from progesterone by biotransformation using the hydroxypropyl- β -cyclodextrin complexation technique. *J. Steroid Biochem. Mol. Biol.* 108, 132–136.
- Martín, L., Prieto, M.A., Cortés, E., García, J.L., 1995. Cloning and sequencing of the pac gene encoding the penicillin G acylase of *Bacillus megaterium* ATCC 14945. *FEMS Microbiol. Lett.* 125, 287–292.

- Martinis, S.A., Atkins, W.M., Stayton, P.S., Sligar, S.G., 1989. A conserved residue of cytochrome P-450 is involved in heme-oxygen stability and activation. *J. Am. Chem. Soc.* 111, 9252–9253.
- McIntosh, J.A., Farwell, C.C., Arnold, F.H., 2014. Expanding P450 catalytic reaction space through evolution and engineering. *Curr. Opin. Chem. Biol.* 19, 126–134.
- Milhim, M., Putkaradze, N., Abdulmughni, A., Kern, F., Hartz, P., Bernhardt, R., 2016. Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. *J. Biotechnol.* 240, 68–75.
- Munetsuna, E., Kawanami, R., Nishikawa, M., Ikeda, S., Nakabayashi, S., Yasuda, K., Ohta, M., Kamakura, M., Ikushiro, S., Sakaki, T., 2014. Anti-proliferative activity of 25-hydroxyvitamin D3 in human prostate cells. *Mol. Cell. Endocrinol.* 382, 960–970.
- Nelson, D.R., 2011. Progress in tracing the evolutionary paths of cytochrome P450. *Biochim. Biophys. Acta BBA - Proteins Proteomics, Cytochrome P450: Structure, biodiversity and potential for application* 1814, 14–18.
- Nelson, D.R., 2009. The Cytochrome P450 Homepage. *Hum. Genomics* 4, 59–65.
- Nordblom, G.D., White, R.E., Coon, M.J., 1976. Studies on hydroperoxide-dependent substrate hydroxylation by purified liver microsomal cytochrome P-450. *Arch. Biochem. Biophys.* 175, 524–533.
- Okano, T., Yasumura, M., Mizuno, K., Kobayashi, T., 1977. Photochemical conversion of 7-dehydrocholesterol into vitamin D3 in rat skins. *J. Nutr. Sci. Vitaminol. (Tokyo)* 23, 165–168.
- Oliver, C.F., Modi, S., Sutcliffe, M.J., Primrose, W.U., Lian, L.Y., Roberts, G.C., 1997. A single mutation in cytochrome P450 BM3 changes substrate orientation in a catalytic intermediate and the regiospecificity of hydroxylation. *Biochemistry (Mosc.)* 36, 1567–1572.
- 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 P-450 reductase. *Arch. Biochem. Biophys.* 117, 660–673.
- Omura, T., Sato, R., 1964. The carbon Monoxide-binding Pigment of liver Microsomes. *J. Biol. Chem.* 239, 2370–2378.
- Ozono, K., Liao, J., Kerner, S.A., Scott, R.A., Pike, J.W., 1990. The vitamin D-responsive element in the human osteocalcin gene. Association with a nuclear proto-oncogene enhancer. *J. Biol. Chem.* 265, 21881–21888.
- Pike, J.W., Meyer, M.B., 2010. The Vitamin D Receptor: New Paradigms for the Regulation of Gene Expression by 1,25-Dihydroxyvitamin D3. *Endocrinol. Metab. Clin. North Am.* 39, 255–269.
- Prosser, D.E., Jones, G., 2004. Enzymes involved in the activation and inactivation of vitamin D. *Trends Biochem. Sci.* 29, 664–673.
- Putkaradze, N., Litzenburger, M., Abdulmughni, A., Milhim, M., Brill, E., Hannemann, F., Bernhardt, R., 2017. CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*. *Appl. Microbiol. Biotechnol.*
- 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 (Pt 1), 159–166.
- Rojkova, A.M., Galkin, A.G., Kulakova, L.B., Serov, A.E., Savitsky, P.A., Fedorchuk, V.V., Tishkov, V.I., 1999. Bacterial formate dehydrogenase. Increasing the enzyme thermal stability by hydrophobization of alpha-helices. *FEBS Lett.* 445, 183–188.
- Röthlisberger, D., Khersonsky, O., Wollacott, A.M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J.L., Althoff, E.A., Zanghellini, A., Dym, O., Albeck, S., Houk, K.N., Tawfik, D.S., Baker, D., 2008. Kemp elimination catalysts by computational enzyme design. *Nature* 453, 190–195.
- Sakaki, T., 2012. Practical application of cytochrome P450. *Biol. Pharm. Bull.* 35, 844–849.

- 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.
- Sasaki, J., Mikami, A., Mizoue, K., Omura, S., 1991. Transformation of 25- and 1 α -hydroxyvitamin D₃ to 1 α , 25-dihydroxyvitamin D₃ by using *Streptomyces* sp. strains. *Appl. Environ. Microbiol.* 57, 2841–2846.
- Sasaki, J., Miyazaki, A., Saito, M., Adachi, T., Mizoue, K., Hanada, K., Omura, S., 1992. Transformation of vitamin D₃ to 1 α ,25-dihydroxyvitamin D₃ via 25-hydroxyvitamin D₃ using *Amycolata* sp. strains. *Appl. Microbiol. Biotechnol.* 38, 152–157.
- Sawada, N., Sakaki, T., Yoneda, S., Kusudo, T., Shinkyo, R., Ohta, M., Inouye, K., 2004. Conversion of vitamin D₃ to 1 α ,25-dihydroxyvitamin D₃ by *Streptomyces griseolus* cytochrome P450SU-1. *Biochem. Biophys. Res. Commun.* 320, 156–164.
- Schuster, I., 2011. Cytochromes P450 are essential players in the vitamin D signaling system. *Biochim. Biophys. Acta* 1814, 186–199.
- Seino, Y., Tanaka, H., Nagai, Y., 1997. Use of 24,25-dihydroxyvitamin D₃ for the manufacture of a medicament for the treatment of rickets. EP0575974.
- Seo, E.G., Einhorn, T.A., Norman, A.W., 1997. 24R,25-dihydroxyvitamin D₃: an essential vitamin D₃ metabolite for both normal bone integrity and healing of tibial fracture in chicks. *Endocrinology* 138, 3864–3872.
- Soares, J.H., Kerr, J.M., Gray, R.W., 1995. 25-hydroxycholecalciferol in poultry nutrition. *Poult. Sci.* 74, 1919–1934.
- Sono, M., Roach, M.P., Coulter, E.D., Dawson, J.H., 1996. Heme-Containing Oxygenases. *Chem. Rev.* 96, 2841–2888.
- Spiro, A., Buttriss, J.L., 2014. Vitamin D: An overview of vitamin D status and intake in Europe. *Nutr. Bull. Bnf* 39, 322–350.
- Stammen, S., Müller, B.K., Korneli, C., Biedendieck, R., Gamer, M., Franco-Lara, E., Jahn, D., 2010. High-Yield Intra- and Extracellular Protein Production Using *Bacillus megaterium*. *Appl. Environ. Microbiol.* 76, 4037–4046.
- Sutton, A.L.M., MacDonald, P.N., 2003. Vitamin D: more than a “bone-a-fide” hormone. *Mol. Endocrinol. Baltim. Md* 17, 777–791.
- Takeda, K., Asou, T., Matsuda, A., Kimura, K., Okamura, K., Okamoto, R., Sasaki, J., Adachi, T., Omura, S., 1994. Application of cyclodextrin to microbial transformation of vitamin D₃ to 25-hydroxyvitamin D₃ and 1 α ,25-dihydroxyvitamin D₃. *J. Ferment. Bioeng.* 78, 380–382.
- Tsai, M.J., O'Malley, B.W., 1994. Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu. Rev. Biochem.* 63, 451–486.
- Tucker, G., Gagnon, R.E., Haussler, M.R., 1973. Vitamin D₃-25-hydroxylase: Tissue occurrence and apparent lack of regulation. *Arch. Biochem. Biophys.* 155, 47–57.
- Underwood, J.L., DeLuca, H.F., 1984. Vitamin D is not directly necessary for bone growth and mineralization. *Am. J. Physiol.* 246, E493–498.
- Urlacher, V.B., Lutz-Wahl, S., Schmid, R.D., 2004. Microbial P450 enzymes in biotechnology. *Appl. Microbiol. Biotechnol.* 64, 317–325.
- Vary, P., 1992. Development of genetic engineering in *Bacillus megaterium*. *Biotechnol. Read. Mass* 22, 251–310.
- Vary, P.S., 1994. Prime time for *Bacillus megaterium*. *Microbiol. Read. Engl.* 140 (Pt 5), 1001–1013.
- Vary, P.S., Biedendieck, R., Fuerch, T., Meinhardt, F., Rohde, M., Deckwer, W.-D., Jahn, D., 2007. *Bacillus megaterium*—from simple soil bacterium to industrial protein production host. *Appl. Microbiol. Biotechnol.* 76, 957–967.

- Wagner, N., Wagner, K.-D., Schley, G., Badiali, L., Theres, H., Scholz, H., 2003. 1,25-dihydroxyvitamin D₃-induced apoptosis of retinoblastoma cells is associated with reciprocal changes of Bcl-2 and bax. *Exp. Eye Res.* 77, 1–9.
- Watanabe, I., Nara, F., Serizawa, N., 1995. Cloning, characterization and expression of the gene encoding cytochrome P-450sca-in2 from *Streptomyces carbophilus* involved in production of pravastatin, a specific HMG-CoA reductase inhibitor. *Gene* 163, 81–85.
- Whitehouse, C.J.C., Bell, S.G., Wong, L.-L., 2012. P450 (BM3) (CYP102A1): connecting the dots. *Chem. Soc. Rev.* 41, 1218–1260.
- Yano, J.K., Koo, L.S., Schuller, D.J., Li, H., Ortiz de Montellano, P.R., Poulos, T.L., 2000. Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*. *J. Biol. Chem.* 275, 31086–31092.
- Yasutake, Y., Nishioka, T., Imoto, N., Tamura, T., 2013. A single mutation at the ferredoxin binding site of P450 Vdh enables efficient biocatalytic production of 25-hydroxyvitamin D(3). *Chembiochem Eur. J. Chem. Biol.* 14, 2284–2291.
- Yasuda, K., Sugimoto, H., Hayashi, K., Takita, T., Yasukawa, K., Ohta, M., Kamakura, M., Ikushiro, S., Shiro, Y., Sakaki, T., 2018. Protein engineering of CYP105s for their industrial uses. *Biochim. Biophys. Acta* 1866, 23–31.
- Zhao, H., Chockalingam, K., Chen, Z., 2002. Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr. Opin. Biotechnol.* 13, 104–110.