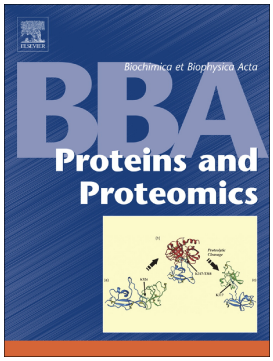


CYP106A2 – a versatile biocatalyst with high potential for biotechnological production of selectively hydroxylated steroid and terpenoid compounds

Daniela Schmitz, Simon Janocha, Flora Marta Kiss, Rita Bernhardt

PII: S1570-9639(17)30166-8
DOI: doi: 10.1016/j.bbapap.2017.07.011
Reference: BBAPAP 39976

To appear in:

Received date: 3 April 2017
Revised date: 14 July 2017
Accepted date: 18 July 2017

Please cite this article as: Daniela Schmitz, Simon Janocha, Flora Marta Kiss, Rita Bernhardt , CYP106A2 – a versatile biocatalyst with high potential for biotechnological production of selectively hydroxylated steroid and terpenoid compounds, (2017), doi: 10.1016/j.bbapap.2017.07.011

CYP106A2 – a versatile biocatalyst with high potential for biotechnological production of selectively hydroxylated steroid and terpenoid compounds

Daniela Schmitz, Simon Janocha, Flora Marta Kiss, Rita Bernhardt*

*corresponding author

Department of Biochemistry, Saarland University, Campus B2.2, 66123 Saarbruecken, Germany

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

Phone: (+49) 681-302-4241

Fax: (+49) 681-302-4739

Abstract

CYP106A2 from *Bacillus megaterium* ATCC13368, was identified in the 1970s as one of the first bacterial steroid hydroxylases responsible for the conversion of progesterone to 15β-hydroxyprogesterone. Later on it has been proven to be a potent hydroxylase of numerous 3-oxo-Δ^4 as well as 3-hydroxy-Δ^5-steroids and has recently also been characterized as a regioselective allylic bacterial diterpene hydroxylase. The main hydroxylation position of CYP106A2 is thought to be influenced by the functional groups at C3 position in the steroid core leading to a favored 15β-hydroxylation of 3-oxo-Δ^4-steroids and 7β-hydroxylation of 3-hydroxy-Δ^5-steroids. However, in some cases the hydroxylation is not strictly selective, resulting in the formation of undesired side-products. To overcome the unspecific hydroxylations or, on the contrary, to gain more of these products in case they are of industrial interest, rational protein design and directed evolution have been successfully performed to shift the stereoselectivity of hydroxylation by CYP106A2. The subsequently obtained hydroxylated steroid and terpene derivatives are especially useful as drug metabolites and drug precursors for the pharmaceutical industry, due to their diverse biological properties and hardship of their chemical synthesis. As a soluble prokaryotic P450 with broad substrate spectrum and hydroxylating capacity, CYP106A2 is an outstanding candidate to establish bioconversion processes. It has been expressed with respectable yields in *Escherichia coli* and *Bacillus megaterium* and was applied for the preparative hydroxylation of several steroids and terpenes. Recently, the application of the enzyme was assessed under process conditions as well, depicting a successfully optimized process development and getting us closer to industrial scale process requirements and a future large scale application.

Background

CYP106A2 is a soluble cytochrome P450 (P450) derived from the soil bacterium *Bacillus megaterium* (*B. meg.*) ATCC13368. As such, it catalyzes the insertion of one oxygen atom into an aliphatic or aromatic substrate and the subsequent reduction of the second oxygen atom to water [1-3]. The P450 superfamily currently counts more than 300,000 representatives, comprising one of the largest and oldest gene families known. They are found among various organisms from bacteria to human (http://drnelson.uthsc.edu/CytochromeP450.html) [4, 5] and their ability to catalyze reactions that do not require oxygen, points to a possible existence even before atmospheric molecular oxygen appeared on the Earth [6]. The vast number of individual proteins reflects the significance of these enzymes in nature [7, 8]. P450s are key players in a variety of anabolic and catabolic processes: in metazoa, P450s are involved in detoxification and biosynthetic processes [9]; in microorganisms, they are vital enzymes in the biosynthesis of secondary metabolites such as polyketides, terpenoids, and others [10]; furthermore, they are also involved in the breakdown of alternative carbon sources as well as in the virulence of harmful bacteria [11]. Their extraordinary versatility makes P450s attractive as biocatalysts, especially for the production of drugs and fine chemicals unattainable from classical chemical reactions [1, 2, 12].

According to their reaction mechanisms, P450s are monooxygenases, requiring electron donors like NADH or NADPH. Most P450s are not able to directly interact with their electron suppliers, hence, additional electron transfer proteins are needed to be involved in the process. According to their electron transfer system, P450s can be grouped into different classes, which were comprehensively reviewed by Hannemann *et al.* and will not be discussed here in detail [13]. CYP106A2 belongs to the group of class I P450 systems, in which the electrons are transferred via an FAD containing reductase and the corresponding iron-sulfur protein from NAD(P)H to the cytochrome.

CYP106A2 - or 15β-hydroxylase, as it was long known owing to its favored hydroxylation position on 3-oxo-Δ^4 steroids - is a ~47 kDa protein consisting of 410 amino acids. It was one of the first steroid hydroxylases found in bacteria, first identified in *B. meg.* ATCC13368 and later characterized as a typical P450 enzyme sharing the common overall fold [14]. So far, only one other P450 was found to which the sequence similarity is more than 55%, namely CYP106A1 from *B. meg.* DSM319 [15]. However, closely related P450s from other subfamilies do exist and might be able to catalyze similar reactions.

Although individual members of the P450 superfamily show a remarkably high similarity in their overall three-dimensional fold, the diversity of substrates that can bind within the active site of P450s is very

high. The reason for this is the high degree of variability in the protein regions lining the active site. There are six sites in total, which are, due to their function, called substrate recognition sites (SRS) [16]. Their identification can be performed via sequence alignment and does not require structural data. The SRS 1 is located in the flexible B-C loop, SRS 2 comprises the C-terminal part of helix F and SRS 3 covers the N-terminal part of helix G. These are secondary sites that mostly come into close contact with the substrate after the closing of the enzyme upon substrate-binding to the active site. In contrast, the regions SRS 4-6 are involved in the initial substrate binding and recognition. SRS 4, comprising the residues of the I-helix in proximity to the heme, SRS 5 (β1-4) as well as SRS 6 (β4-1/β4-2 loop) flank the prosthetic group and build a boundary to the active site (see **Fig. 1**). As the SRSs play a significant role in substrate recognition and orientation, these protein regions are of particular interest for mutagenesis studies with the aim to alter the regio- and/or stereoselectivity of a P450.

Being one of only a few cloned bacterial steroid hydroxylases, CYP106A2 is a highly interesting candidate for establishing biotechnological steroid hydroxylation processes [17]. Furthermore, the enzyme has been proved to be a potent hydroxylase not only of steroids but of also tricyclic diterpene as well as pentacyclic triterpene acids (**Fig. 2**) with similar reaction velocities for all the abovementioned substrates, which to the best of our knowledge makes CYP106A2 unique among bacterial steroid hydroxylases [18-21]. The selective hydroxylation of the aforementioned highly functionalized molecules by chemical means requires complex chemical reactions. The stereo- and regioselective chemical synthesis of hydroxysteroids is often cost-intense, requires protection groups, toxic catalysts as well as harsh reaction conditions, not to mention the undesired side-reactions [22]. In contrast, P450s perform selective biocatalysis under mild reaction conditions: temperatures around 30 °C, atmospheric pressure and aqueous media [19, 23]. The synthetic industrial scale application of P450s is, however, so far limited. Major drawbacks include: low stability and low reactivity under process conditions; stoichiometric consumption of electrons during the reaction, usually applied in the form of the expensive co-factors NADH or NADPH; uncoupling and formation of reactive oxygen species (ROS) [2, 24-26]. The application of whole-cell systems containing P450s provides a suitable alternative to overcome the above-mentioned shortcomings. In case of CYP106A2, several whole-cell systems have been established to date, based mainly on bacterial hosts such as *Escherichia coli* (*E. coli*) and *B. meg.* and harboring diverse expression systems for CYP106A2 in combination with different electron suppliers [18, 27, 28]. A summary of the established whole-cell applications, including the selected host organism, the vectors, the electron transfer proteins and the respective catalyzed reaction is shown in **Table 1**. With these well-studied systems, steroids, terpenoids, and terpene acids have been successfully converted into

hydroxylated products with reasonable amounts approaching industrial process requirements. In this context it should be mentioned that approximately 300 approved steroid drugs are known to date [29], ranking among the most marketed medicinal products next to antibiotics [30]. According to the review of Lundemo and Woodley for high-value drug production, such as hydroxylated steroids, a space-time yield of 2g/L/d should be approached. However, this benchmark can also be lower depending on the value of the obtained product and the specific shortcomings of the chemical process [24].

Although CYP106A2 WT, in contrast to P450BM3 WT- the most intensively studied P450 regarding its potential for white biotechnology - has relatively low activity, its outreaching ability to perform stereo- and regioselective hydroxylations on complex molecules makes it an interesting candidate for biotechnological application.

This review article attempts to summarize the important achievements made in the development of CYP106A2 based biotransformations within the last four decades and to discuss possibilities of further improvement and potential biotechnological application.

CYP106A2-catalyzed steroid hydroxylation

In 1958, McAleer et al. described the 15β-hydroxylation of progesterone using *B. meg.* ATCC13368 [31]. Almost twenty years from that, in 1975, the responsible progesterone converting enzyme, CYP106A2 (P450meg), was partially purified, characterized and identified as a member of the cytochrome P450 family [32]. CYP106A2 was classified as a bacterial class I P450, being part of an enzyme system consisting of two other soluble proteins: a strictly NADPH-dependent FMN-containing flavoprotein (megaredoxin reductase) and an iron-sulfur protein (megaredoxin) [14]. The components of the steroid hydroxylase system were purified to homogeneity from *B. meg.* ATCC13368 not before three years later [33-35], yet reporting no detailed biochemical and biophysical characterization of the redox partners. Nevertheless, still in 1976, Berg *et al.* investigated the CYP106A2 dependent hydroxylation of a variety of 3-oxo-Δ^4-steroids and, except for one, all steroids were hydroxylated specifically at position 15β. Only the conversion of progesterone resulted in an additional 6β-hydroxy compound, produced only to a minor extent [14]. Investigations performed later revealed that CYP106A2 produces 9α- and 11α-hydroxyprogesterone as well [36-38].

Although CYP106A2 shows high specificity towards steroids, its role in the metabolism of *B. meg.* remains elusive, since the metabolome of this organism has not been studied in detail yet. In order to identify potential substrates of CYP106A2, Berg and Rafter studied the inducibility and substrate

specificity of the system and published their results in 1981 [35]. In total, 20 different compounds were tested, including fatty acids, sterols, cholic acids, drugs and human metabolites (included in **Table 2**). Except for aniline, none of these compounds were successfully converted into an oxygenated product. Furthermore, the authors pointed out that aniline is probably oxidized through the oxidizing character of the heme itself rather than due to a selective interaction with CYP106A2.

Although CYP106A2 was proven to be highly selective towards steroids, no alteration of the spin-state was shown with any of the substrates tested before 2004, when Lisurek *et al.* identified cholestenone as the first type I shift inducing ligand. Before the first type I ligand was found, FTIR-spectroscopy measurements in the presence of the known substrate 11-deoxycorticosterone (DOC) gave insight to the CYP106A2-substrate interaction, proving that DOC is bound to the active site of CYP106A2 above the heme, similar to other P450-substrate interactions [39]. After the identification of cholestenone as a high-spin ligand, its reaction with CYP106A2 was verified using LC-MS/MS analysis, which revealed the product 15β-cholestenone [36].

Lisurek *et al.* managed to identify cholestenone as a high-spin ligand by difference spectroscopy using tandem cuvettes in a double beam photometer for analysis, which provides a much higher sensitivity compared with absolute spectra and thus is the method-of-choice for small alterations of the spin state. It has been shown that the substrate induced alterations in the absorbance - high-spin shift from 417 towards lower wavelengths - are usually less than 15 % in CYP106A2 [40]. This behavior has also been observed in previous studies with some of the drug converting P450s [41]. The low percentage of the high-spin heme of the substrate bound complex could be reconciled with the high flexibility of the CYP106A2's active site and substrate entry channel, which could also give an explanation, why substrate conversions without any detection of high-spin heme are regularly observed with CYP106A2 [20].

CYP106A2-catalyzed terpene hydroxylation

Encouraged by the identification of cholestenone as a type I substrate, the substrate spectrum of CYP106A2 was investigated to identify novel substrates using high-throughput screening based on difference spectroscopy. A library containing more than 16,000 synthetic compounds was screened, and dihydroquinone pimaric acid was recognized as another type I ligand [21]. Using this scaffold for further investigations, one of the main terpenoic acids of colophony, abietic acid, was also identified as a CYP106A2-substrate (type 0, no spin-state alteration), and its reaction products were characterized as 12α- and 12β-abietic acid. These results led to the recognition of the enzyme as the first regioselective bacterial allylic diterpene hydroxylase, with dihydroquinone pimaric acid and abietic acid as its first

discovered terpene substrates. The abovementioned terpenes both lack the 3-oxo-Δ^4 moieties, which until then was assumed as an essential prerequisite for CYP106A2 based hydroxylations. Because terpenoic acids are highly interesting from a pharmaceutical and, therefore, also from a biotechnological point of view, further terpenoic acids with larger nuclei were tested for CYP106A2 substrates, leading to the discovery of a highly selective hydroxylation of the pentacyclic triterpene 11-keto-β-boswellic acid at position 15β [18]. Taken together, CYP106A2 has not only been demonstrated as a potent steroid hydroxylase, but also a regio-selective diterpene and triterpene hydroxylase.

Screening for novel CYP106A2 substrates

To further expand the substrate spectrum of CYP106A2, with special attention to potential natural substrates, the screening of a compound library - entailing 502 bioactive natural products of various origin and diverse structures - was carried out. Substrates of cytochromes P450 usually induce a shift of the heme iron to the high-spin state, resulting in a high-spin shift for so-called type I substrates [42], which is characterized by an increase of the absorption at 390 nm and a corresponding decrease at 420 nm. For the screening assay, the well-established difference spectroscopy method was adapted to a 96-well plate format. Screening of the natural product library eventually led to the identification of twelve additional type I ligands for CYP106A2: betulinic acid, oleanolic acid, glycyrrhetinic acid, ursolic acid, betuline, dipterocarpol, alpha-apo-oxytetracycline, noscapine, panaxadiol, geraldol, sarsasapogenin and rhamnetin [27]. Two of these twelve hits –

dipterocarpol and betuline – were also identified as new substrates of CYP106A2 in our follow-up studies. Surprisingly, dipterocarpol was not hydroxylated at position 15β but at positions 7β and 7β,11α respectively. Dipterocarpol belongs to the group of dammarane triterpenoids (C30). Dammaranes comprise the aglycones of many sapogenins, especially from ginseng plants. Dammaranes share a gonane skeleton with an eight carbon atom side chain attached to C-17, a double methylated C4 and another methyl group at C14 (**Fig. 2)**. Therefore, these compounds are highly similar to C27 steroids, yet lack the 3-oxo-Δ^4 moiety and other functionalizations. Nevertheless, dipterocarpol is hydroxylated by CYP106A2, furthermore, shifts the heme-iron into its high-spin form upon binding. This information stimulated further investigations on the CYP106A2 enzyme's hydroxylation activity towards 3-hydroxy-Δ^5, A-ring reduced and aromatic steroids such as dehydroepiandrosterone (DHEA), pregnenolone, digitoxigenin, prednisone and dexamethasone, later identified as non-3-oxo-Δ^4 substrates [43]. The oxidation state of the A-ring apparently influences enzyme's stereo-selectivity: the change from 3-oxo-Δ^4 to 3-hydroxy-Δ^5 moiety in otherwise identical steroids alters the main hydroxylation position from C15 to

C7. Thus, in the case of DHEA and pregnenolone, the main hydroxylation position was shown to shift from 15β to 7β and, for DHEA a minor hydroxylation at position 7α was also observed. Side product formation was similarly detected during pregnenolone conversion, and although its structure was not solved, it is assumed to be the 7α-derivative as well.

In total, as described in this chapter and shown in **Table 2**, 76 different substances have been tested as potential substrates for CYP106A2, 38 of them were identified as substrates, out of which 26 induce a high-spin shift of the heme iron. A detailed list of all substances tested for hydroxylation with CYP106A2 is given in **Table 2** and the skeletons of the identified substrate structures are depicted in **Fig. 3.** All CYP106A2 substrates share the isoprenoid structure (except aniline and imipramine); however, the list is so far constricted to di- and triterpenes including steroids. The substrate spectrum of CYP106A2 seems to be rather narrow, yet some of the CYP106A2 substrates, like the steroids, are highly interesting pharmaceuticals [30] making CYP106A2 an attractive candidate for biotechnological applications.

Development of whole-cell systems using CYP106A2

Hydroxylation of steroids is of utmost interest in the so-called white biotechnology, because the mode of action of steroids directly depends on their structure, i.e. the number, type, regio- and stereo- position of functional groups linked to the steroid core as well as the oxidation state of the rings [29]. This analogously applies for the bioactivity of terpenes. As an example, boswellic acids comprise the major component of the resin of Frankincense (*Boswellia sacra*) and are used as anti-inflammatory herbal drugs since centuries. However, they are highly lipophilic compounds, exhibiting bad pharmacokinetics, meaning poor intestinal absorption and distribution in the body, resulting in low serum levels [44]. The insertion of polar groups, e.g. by hydroxylation, increases the polarity and facilitates further modifications of the compound. Abietic acid and boswellic acid depict the first non-steroidal natural products being efficiently hydroxylated by CYP106A2, both *in vitro* and *in vivo.* Since the scale-up and industrial use of *in vitro* CYP106A2-systems is not feasible due to the need of redox partners and NAD(P)H, whole-cell biotransformation is currently the best method of choice. In 2006, Hannemann *et al.* designed an *E. coli*-based whole-cell bioconversion system with CYP106A2 and the mitochondrial electron transfer system from bovine adrenals using a two-plasmid strategy (**Table 1**) [45]. Conversion rates of DOC were achieved up to 1 mM/d, resulting in a theoretical space-time yield of 0.33 g/L/d. As NADPH is needed in stoichiometric amounts to accomplish P450 reactions, the intracellular NADPH content was thought to be one possible limiting factor in biocatalysis. This whole-cell system was then augmented with a cofactor regeneration system by implementing the alcohol dehydrogenase, LbADH,

from *Lactobacillus brevis* [46]. Based on this construct, preparative scale conversions were established for progesterone and testosterone hydroxylation in *E. coli*. However, the intracellular NADPH content was shown not to be the major limiting factor in this investigated CYP106A2–dependent steroid hydroxylation. Cell density and cell permeability are additional drawbacks known for whole-cell bioconversions. High cell densities usually result in a higher amount of catalyst, yet no increase of product yield is described using high cell densities for CYP106A2 bioconversions. Nevertheless, as shown by Zehentgruber *et al.*, steroid conversion seems to be drastically improved by enhancing cell permeability using CCEs (crude cell extracts): a 6-fold increase in the reactivity was measured with progesterone and testosterone, prompting the assumption that the substrate import through the cell membrane is the primary limitation of the bioconversion system [46]. Together with optimization of the reaction conditions (e.g. pH, temperature, propan-2-ol and initial substrate concentration) the productivity was eventually increased by a factor of 18, resulting in a theoretical space-time productivity, calculated from the initial reaction rates, of 3.7 g/L/d for progesterone and 5.5 g/L/d for testosterone [46]. Furthermore, side product formation was reduced by using resting cells (non-dividing cells, due to the exchange of growth medium to carbon, energy and nitrogen depleted reaction medium, usually Tris-HCl or potassium phosphate buffer) instead of growing cells.

In order to increase the yield of a biotechnological process with cytochrome P450 enzymes, the redox partners are as important as the P450 itself, since the electron transfers are often the rate-limiting steps in the catalytic cycle of P450s [47]. Because for long time the natural electron transfer partners of CYP106A2 have not been cloned, heterologous redox partners were commonly used to reconstitute the system's electron supply. The best-characterized system consists of the electron transfer proteins from bovine adrenals, with adrenodoxin reductase (AdR) and the native adrenodoxin (Adx), replaced by the truncated form (Adx_{4-108}), which has shown a significantly higher efficiency in transferring electrons to CYP106A2 [37, 39]. Besides, endogenous redox partners from *Pseudomonas putida* S12 have also shown substantial interaction with CYP106A2 as well as the ferredoxin cloned from *Bacillus subtilis* [48]. More recently, novel redox partners from *B. meg.* DSM319 have been identified, cloned and used to reconstitute the *in-vitro* activity of CYP106A1, a close homolog of CYP106A2. One flavodoxin and 4 ferredoxins have been investigated and all 5 proteins were able to support CYP106A1 activity. Unfortunately, the corresponding reductase was not identified in this strain. The close homology between the CYP106A subfamily-members suggests that the identified redox partners may be applicable to reconstitute the CYP106A2-activity, yet, further investigations are recommended to assess their potential.

Since the membrane-associated mammalian ferredoxin reductase (AdR) showed low expression yields in *E. coli*, the question arose, whether this might be a possible limiting factor within the whole-cell system. To investigate this problem, Ringle *et al.* developed an alternative electron transfer system based on the *E. coli* flavoprotein reductase (FpR) together with the highly active bovine Adx_{4-108} supporting the provision of electrons to CYP106A2. The Fpr-Adx-P450 system showed a 33% increased initial reaction rate compared to the AdR-Adx-P450-system in *E. coli* [49]. As the product yield did not increase using FpR and Adx for electron supply, this step did not seem to be limiting for the CYP106A2-dependent steroid or terpene hydroxylation capacity in *E. coli* whole-cells under the investigated conditions.

Although the abovementioned *E. coli* whole-cell systems were successfully applied for preparative scale steroid hydroxylations, terpenoic acid conversions have not been accomplished by this recombinant host, most probably due to the organization of the outer cell membrane (OM) of the gram-negative microbe. The OM of *E. coli* is coated with lipopolysaccharides, creating a polar environment which is supposed to block the entry of terpene acids into the cell [50]. This hindrance can be met by permeabilizing the cells using detergents or pore forming agents such as polymyxins. This approach has successfully been established for the whole-cell conversion of abietic and other diterpene acids using CYP105A1 from *Streptomyces griseolus* in *E. coli* [51]. To avoid the permeabilization of the cells, which may also have harmful effects on cell viability, gram-positive bacteria can be used due to their lack of the OM and hence the lipopolysaccharides. Bleif *et al.* showed that the natural host of CYP106A2, *B. meg.* ATCC13368 is suitable to convert even preparative amounts of abietic acid in a whole-cell conversion system [21]. Additionally, they succeeded in creating the first recombinant expression system for P450s in *B. meg.*, using the MS941 strain - an alkaline protease deficient derivative of the plasmidless DSM319 strain - to successfully perform the whole-cell conversion of 11β-keto-boswellic acid [18]. *B. meg.* was subsequently also used for the preparative scale conversions of dipterocarpol. Using the ATCC13368 strain and a discontinuous substrate-feeding strategy in a lab-scale batch fermentation process, a space-time productivity of 150 mg/L/d was achieved [27]. With the above described system, the steroid hormone precursor DHEA was also successfully converted into its 7β-hydroxylated derivative, achieving space time yields up to 400 µM/d. It is noteworthy that the described conversion rates were reached using the native host of CYP106A2, constitutively expressing the enzyme in the stationary phase of growth. However, using *B. meg.* MS941_CAA (co-expressing the P450 and the bovine adrenal redox partners) Schmitz *et al.* observed significantly higher conversion rates alongside with high regioselectivities. The 7β-hydroxylation of DHEA, for instance, was completed with a stereoselectivity of 90 % (with the MS941_CAA strain) and a space-time productivity of 115 mg/L/h (2.7 g/L/d), reflecting an

exceptionally high initial reaction velocity compared to other P450 whole-cell systems [43]. In addition, as it has been already observed with *E. coli* based biotransformation, the side product formation was decreased when conversions were performed in buffer instead of a complex medium (unpublished data). It should also be mentioned here that the expression of CYP106A2 was also accomplished in a solvent tolerant *P. putida* strain in order to overcome the poor solubility of its substrates. However, production rates were not given in this report [48].

The stability and retained activity of a catalyst under process conditions is crucial to accomplish a successful industrial process. The feasibility to apply CYP106A2 under process conditions was recently investigated, using the enzyme overexpressed in the *B. meg.* MS941 strain (without heterologous redox partners) [52]. The approach started with systematically identifying the limitations of substrate conversion and addressing the upcoming problems one-by-one during the upscale of the reaction from shake-flask to a laboratory scale bioreactor. The low overall solubility of both, substrate and product, and the limited P450 stability, in combination with product inhibition, were recognized as main bottlenecks. By using an alternative substrate-dissolution method (2-hydroxypropyl-β-cyclodextrine) to circumvent the solubility and product inhibition issues, a nearly complete conversion of the substrate, cyproterone acetate, was achieved resulting in a product formation of 0.43 g/L at a 400 mL scale, getting closer to a future industrial-scale application.

Despite the success in process development, there are still remaining challenges such as the limited stability of the P450 in the bioreactor environment (reaction mixing, aeration etc.) to be overcome. This was observed by the decrease of the correctly folded P450 (CO-difference spectroscopy) after 4 h reaction time, leading to a more than 50 % drop in the product formation during the 24 h timeframe of the biotransformation process. The correlation between process conditions and enzyme stability was confirmed by storing the same batch of resting cells expressing CYP106A2 at 4°C in potassium phosphate buffer, in which case the enzyme activity remained unaffected within 7 days [52].

The loss of enzyme activity has been reported previously for CYP106A2 within 24 hours [43] and has also been observed with other P450 whole-cell conversions [24]. However, detailed investigations on the factors leading to the fast degradation of active enzyme during the bioconversions are missing. Recent results indicate that high substrate concentrations could indeed negatively influence the *in vivo*-stability of the P450 as no drop in activity was observed when cells were cultivated for up to 72 hours prior to bioconversion [43]. High substrate concentrations lead to high P450 turnovers, likely resulting in

metabolic stress and increased production of reactive oxygen species; however, this assumption remains to be investigated in detail.

In summary, CYP106A2 whole-cell systems have efficiently been used for the production of several hydroxylated steroids and terpenoids. Using steroids as substrates, productivities up to 5.5 g/L/d were achieved, while with terpenoids, like 11β-keto-boswellic acid or dipterocarpol, the productivities were lower, yet, several hundred mg/L/d of the hydroxyl product are produced (**Table 1**).

Engineering CYP106A2

Over the past 50 years, CYP106A2 has emerged as a potent biocatalyst for the conversion of steroids and terpenoids with the help of efficient expression systems, well-established on a preparative scale. Nevertheless, further approaches to improve the enzyme's performance and to change the stereo- and/or regio-selectivity of hydroxylation for a desired product pattern is still of high biotechnological interest. To generate an enzyme with new or improved functions, protein engineering is a popular method, involving two major strategies: rational protein design, based on identifying amino acid exchanges leading to improved performance with the help of the crystal structure or a homology model of the enzyme, and directed evolution, which involves random mutagenesis, followed by high-throughput screening of the generated mutant libraries to identify improved species.

CYP106A2 targets highly interesting positions within steroid molecules, such as 6β, 7β, 7α, 9α and 11α, but the principal hydroxylation of 3-oxo-Δ^4-steroids still remains to be at the 15β position. The often observed minor side-products displaying 6β-, 9α- and 11α-hydroxylation are, however, more interesting from a pharmaceutical point of view. 11α-OH progesterone shows anti-androgenic activity with minimal estrogenic and progestational side effects [53, 54] and is an important intermediate for the synthesis of cortisone [55]. 9α-OH progesterone is a useful intermediate in the synthesis of physiologically active substances with glucocorticoid and progestational activity [56] and 6β-OH progesterone can be used as an intermediate for the synthesis of 6β,14α-dihydroxyandrost-4-ene-3,17-dione, an inhibitor of the growth of breast cancer cells [57]. Being only side reactions, the 6β-, 9α- and 11α-hydroxylations of CYP106A2 were targeted by protein engineering in prospect of improvement. Progesterone, the most intensively investigated CYP106A2 substrate was taken as a model substrate for rational protein design/site-directed mutagenesis and directed evolution approaches.

CYP106A2 hydroxylates progesterone (P) at positions 15β, 11α, 6β, and 9α [36, 38]. In order to alter the enzyme's activity and stereoselectivity in favor of the 11α position, a homology model was compiled to

predict positions for mutagenesis [37]. As all P450s share highly conserved structural motifs like the SRS regions and the substrate entry channel, these regions were chosen as hot spots for mutagenesis. Two sets of mutants were generated, one with mutations located within the substrate entry region including the B'-helix and the F/G-loop and the other, containing mutations in the SRS6. In order to increase the 11α-hydroxylation activity of CYP106A2, site-directed mutagenesis was performed. Since at the time of these experiments [37] the crystal structure of CYP106A2 was not available, the human 11-hydroxylase (CYP11B1) was used as a template to provide an idea about potential target residues for mutagenesis. Residues involved in 11-hydroxylation of the human CYP11B1 were identified, and the respective residues in CYP106A2, were transformed into those from CYP11B1: S394I, A395L, T396R, G397P, and Q398S. As expected, the strongest effects were observed with mutations placed in the center of an SRS, namely SRS6: A395L and G397P. For mutant A395L an increase of the 11α-OH-P formation from 9 to 39 % was found, while the 9α-OH-P production improved from 4 to 20 %, and the 6β-OH-P formation remained comparable to the wild type (WT) P450. Thus, the relative 11α-hydroxylation activity was increased fourfold, resulting in an absolute formation of 60 nmol 11α-OH-P min^{-1} per nmol P450. With mutant G397P, the relative formation of 11α-OH-P is increased to 41%, followed by 6β-OH-P (21 %) and 9α-OH-P (9 %). The overall activity of G397P was, however, low with only 11 nmol 15β-OH-P min^{-1} per nmol P450 which is approximately 7 % compared with the WT, leading to a lower absolute product formation of 11α-OH-P than the WT CYP106A2. Nevertheless, the mutagenesis results revealed that the positions A395 and G397 are critical in progesterone 11α-hydroxylation [37] and the identified mutants A395L and G397P were used in further experiments as starting points for an improved 11α-hydroxylation by means of directed evolution.

An important step for the directed evolution of CYP106A2 was the development of an effective screening technique to identify the mutants exhibiting higher steroid hydroxylation activity, since the bottleneck is not the number of mutants generated but the throughput and sensitivity of the screening techniques applied [58]. In order to directly measure steroid hydroxylation activity of the CYP106A2 mutants in *E. coli* whole-cells, a 96-well plate-based high-throughput screening was established, based on a fluorescence assay developed by Appel *et al.* [45, 59]. The technique was applied to screen 2100 CYP106A2-mutants towards higher 11-deoxycortisol and 1900 mutants towards more elevated progesterone hydroxylating activities [60]. In these studies, the A106T/R409L mutant was selected as the most promising one and used to generate a second generation of mutants which finally revealed the quadruple mutant A106T/Q109K/T399S/R409L with a four-fold increase in 11-deoxycortisol hydroxylation activity. Based on these achievements, Nguyen *et al.* applied an additional round of

directed evolution to generate improved mutants favoring the 11α-position of progesterone and giving comparable or even better activities than the wildtype [61]. For this purpose, they conducted saturation mutagenesis on the mutants of positions A395 and G397 using the abovementioned screening approach and identified mutants with 8.9 and 11.5 times higher 11α-hydroxylase activity out of approximately 13100 transformants tested. In order to further increase the activity of these mutants, site-directed mutagenesis was performed (selection of positions for mutations, see [61]) resulting in the identification of the mutant T89N/A395I with the highest progesterone 11α-hydroxylase activity. This mutant was applied to the *E. coli* based whole-cell conversion system. In these experiments the production of 15β-hydroxyprogesterone decreased from 50.4% (WT) to 4.8% and that of 11α-hydroxy progesterone increased from 27.7% (WT) to 80.9% thus demonstrating a dramatic improvement in the regioselectivity. The double mutant T89N/A395I produced 8 mg/L 11α-hydroxy progesterone with 80.9% selectivity towards 11α-hydroxy progesterone using *E. coli* as host. Using the created mutant libraries and approach, studies to change the selectivity towards the 9α- and 6β-position were also performed. Very recent data demonstrates that a triple mutant (F165L/A395E/G397V) can be constructed which showed a ten-fold increase in 9α-hydroxy progesterone compared with the wild type. Moreover, using rational protein design based on the recently obtained 3D structure of CYP106A2 [20] and docking of the steroid substrate, residue A243 was identified as potentially critical for 6β-hydroxylation. According to this result, mutant A243S was constructed and demonstrated 82.7% 6β-hydroxy progesterone compared with only 8.7% produced by the wild type [62].

Outlook

CYP106A2 emerged as a potent biocatalyst for biotechnological applications due to the enzyme's ability to selectively hydroxylate a number of isoprenoids, primary steroids, di- and triterpenoids with reasonable activities. Yields in the range of several g/L/d (**Table 3**) have been obtained and pave the way for further optimizations of the protein itself via various mutagenesis techniques and the application of the enzyme in whole-cell biotransformation processes. The highest efficiencies (k_{cat}/K_m) of substrate hydroxylation were obtained with mutants of CYP106A2 and progesterone as substrate ranging up to 200 x 10^3 M^{-1} s^{-1} [61]. V_{max} values for steroids were in the range of up to 400 nmol/min/nmol P450 [43, 60], whereas those for terpenes reached about 100 nmol/min/nmol P450 [21]. It has to be taken into account, however, that so far no mutants have been created to improve the conversion of terpenes and, therefore, further advances in terpene hydroxylation are to be expected.

To profit from the enzyme's unique selective steroid and terpene hydroxylase activity and to exploit the competence of CYP106A2 in white biotechnology, studies were designed and succesfully applied to improve and/or change the regioselectivity of the enzyme using directed evolution and rational protein design. The results confirmed that CYP106A2 is an evolvable protein, which can undergo larger functional modifications to meet the current pharmaceutical interests. Thus, successful re-design of CYP106A2 from a 15β- to a 6β or even 11α-hydroxylase, changing the hydroxylation position from one side of the steroid molecule to the opposite one, has been performed. The very recently published 3D structure of CYP106A2 in substrate-free and substrate-bound form [20] (**Fig. 4**) will certainly contribute significantly to the design of novel activities and selectivities as well as to the improvement of the efficiency of product formation.

Despite the fact that natural electron transfer partners of CYP106A2 have not been cloned so far, the electron supply to CYP106A2 was accomplished with several heterologous redox systems. This has been an undoubtable advantage for the development of whole-cell systems for preparative scale transformations. Based on the profound background about *E. coli* and *B. meg.* whole-cell applications and our broadened knowledge concerning CYP106A2, the resting whole-cell approach is the desired strategy for the synthesis of novel hydroxy-compounds on an industrial level. This way, the hydroxylation of complex molecules can be selectively performed on a larger scale with attractive space-time yields using simply glucose for cofactor regeneration, without the need for time consuming preliminary protein purification and the use of expensive cofactors. Nonetheless, these whole-cell catalytic systems should be further improved by meticulously identifying and adressing the limiting factors of each one. This could entail improved substrate and product dissolution methods, cell permeablization for improved substrate uptake, coexpression of a cofactor regeneration system, increasing enzyme stability and activity by protein engineering etc.

The achievements of CYP106A2-research summarized herein, strengthen the enzyme's role as an important biocatalyst in the production of pharmaceutically relevant drugs and/or drug precursors. The gathered knowledge will hopefully lead to the development of a straightforward, low-cost fermentation, ultimately replacing synthetic chemical steps in industrial catalysis with bioconversion towards cheaper and greener processes.

Acknowledgements

The authors thank Dr. Martin Litzenburger for the help with Figures 2 and 3.

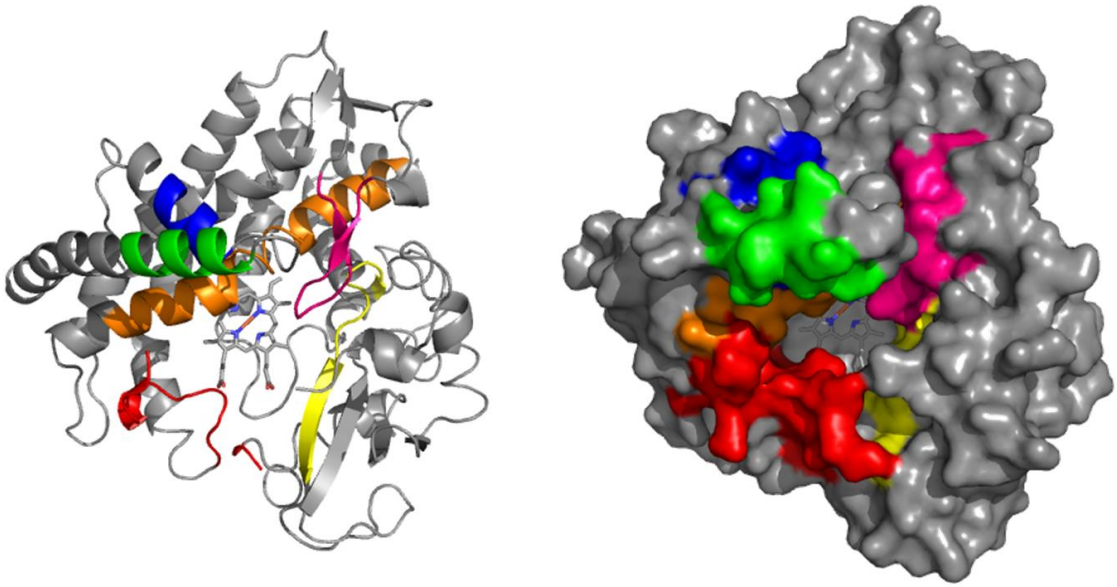


Figure 1: Substrate recognition sites of CYP106A2

This distal view into the substrate binding pocket of CYP106A2 shows the position of the substrate recognition sites. Red: SRS 1 (BC-loop including helix B'). Blue: C-terminal part of helix F. Green: N-terminal part of helix G. Orange: central part of helix I. Yellow: Region around β1-4. Pink: loop between β4-1 and β4-2.

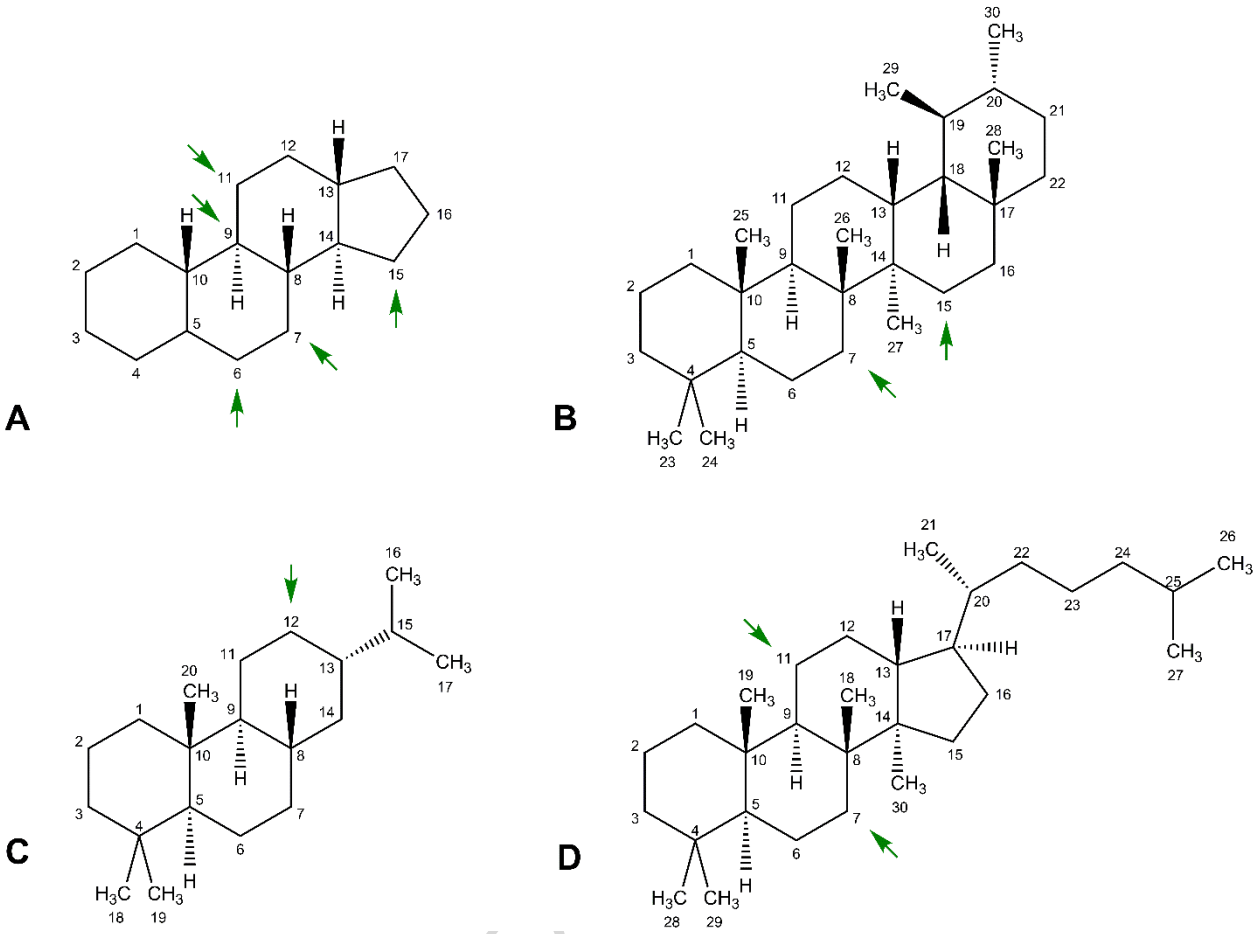
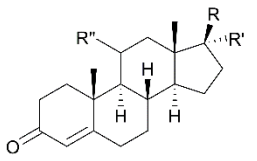
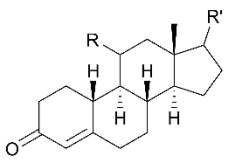


Figure 2: The skeletons of CYP106A2 substrates

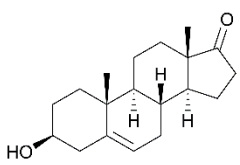
The gonane (A), ursane (B), abietane (C) and dammarane (D) terpenoid cores are shown. Arrows indicate the sites within the molecules known to be hydroxylated by CYP106A2.



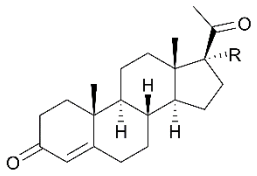
1. R, R' = O, R'' = H
2. R = OH, R' = H, R'' = H
3. R=OH, R' = CH_3, R'' = H
4. R, R' = O, R'' = O



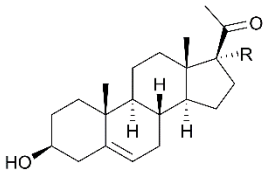
5. R = H, R' = OH



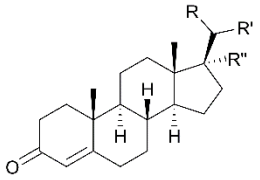
6.



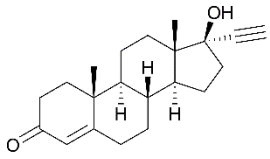
7. R = H
8. R = OH



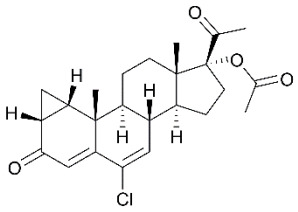
9. R = H
10. R = OH



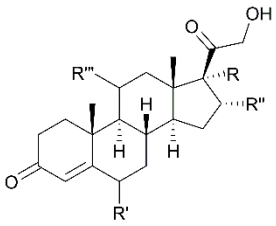
11. R = H, R' = OH, R'' = H
12. R = H, R' = OH, R'' = H
13. R = CH_2OH, R' = OH, R'' = OH



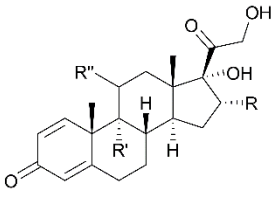
14.



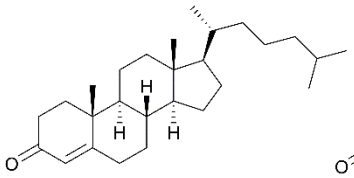
15.



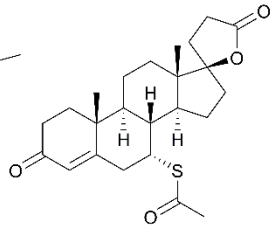
16. R = H, R' = H, R'' = H, R''' = H
17. R = H, R' = H, R'' = H, R''' = OH
18. R = OH, R' = H, R'' = H, R''' = H
19. R = OH, R' = H, R'' = H, R''' = OH
20. R = OH, R' = H, R'' = H, R''' = O
21. R = H, R' = F, R'' = CH_3, R''' = H



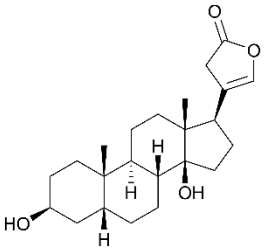
22. R = H, R' = H, R'' = OH
23. R = H, R' = H, R'' = O
24. R = CH_3, R' = F, R'' = OH



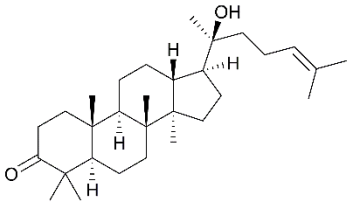
25.



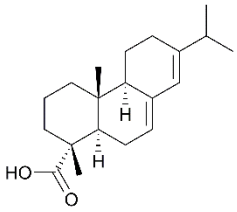
26.



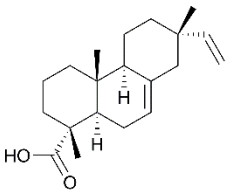
27.



28.



29.



30.

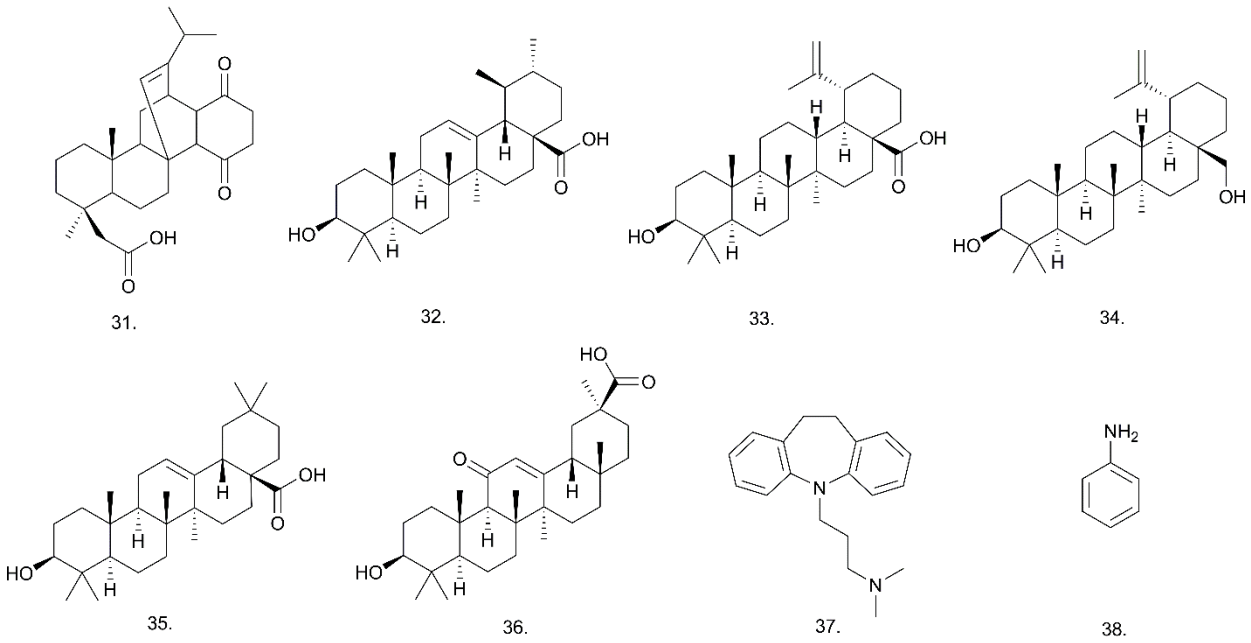


Figure 3: Structures of known CYP106A2 substrates

(1) Androstenedione, (2)Testosterone, (3) 17α-methyltestosterone, (4) Adrenosterone, (5) Nortestosterone , (6) Dehydroepiandrosterone, (7) Progesterone, (8) 17α-hydroxyprogesterone, (9) Pregnenolone, (10) 17α-OH-Pregenenolone, (11) 4-Pregnen-20α-ol-3-one, (12) 4-Pregnen-20β-ol-3-one, (13) 4-Pregnen-17-20α-21-triol-3-one, (14) Ethisterone, (15) Cyproterone acetate, (16) 11-Deoxycorticosterone, (17) Corticosterone, (18) 11-Deoxycortisol, (19) Cortisol, (20) Cortisone, (21) 6-Fluor-16-methyl-deoxycorticosterone, (22) Prednisolone, (23) Prednisone, (24) Dexamethasone, (25) Cholestenone, (26) Spironolacton, (27) Digitoxigenin, (28) Dipterocarpol, (29) Abietic acid, (30) Isopimaric acid, (31) Dihydroquinone pimaric acid, (32) Ursolic acid, (33) Betulinic acid (34) Betulin, (35) Oleanolic acid (36) Glycyrrhetenic acid, (37) Imipramine, (38) Aniline

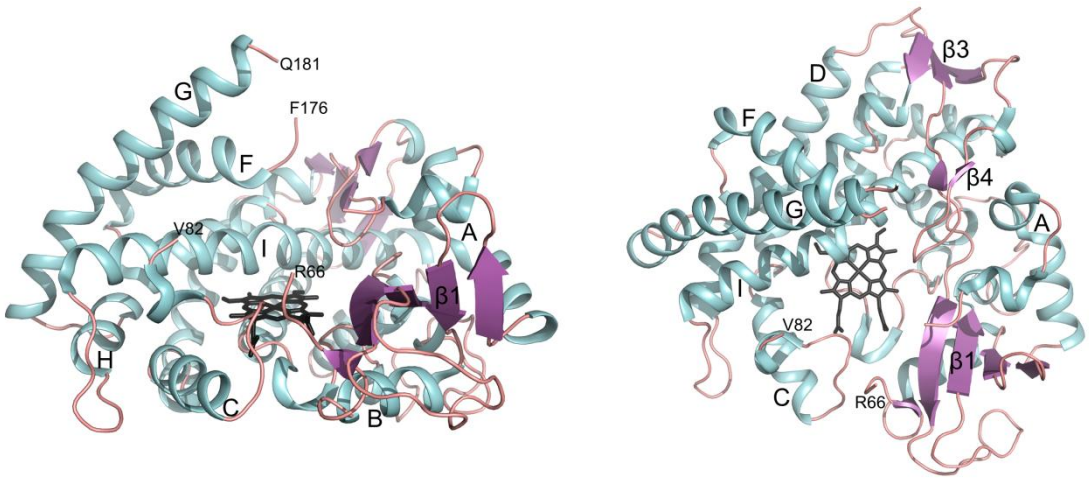


Figure 4: Crystal structure of CYP106A2 at 1.8 Å (PDB code 4YT3)

The secondary structure elements are named according to the CYP-nomenclature [63]. The first and last amino acids behind the flexible loops that could not be modeled are marked. The heme cofactor is shown in black. Left: View from the side onto the central helix I. Right: View from above into the active site.

Table 1*:* Whole-cell systems established for CYP106A2

In all cases, substrates were added as ethanolic solution (except for cyproterone acetate, where cyclodextrine was used) and conversions were monitored by either TLC or HPLC techniques. Some data are extrapolated from the actual time of measurement to 24 hours to make the data more compatible.

n.d. not determined

Organism (Expression plasmids)	Reaction conditions	Substrate	TTN	Conversion	Productivity	Selectivity	Reference
E. coli (pBarTwin + pACYC_FHH2.8)	Whole-cell conversion in complex medium	Deoxycorticosterone	n.d	n.d	0.33 g/l/d	n.d	(Virus et al., 2008)
E. coli BL21 (pBarTriple + pACYC_FHH2.8)	Lyophilized cells in buffer	Progesterone	1323	85%	3.7 g/l/d	86%	(Zehentgruber et al., 2010)
E. coli BL21 (pBarTriple + pACYC_FHH2.8)	Lyophilized cells in buffer	Testosterone	608	96%	5.5 g/l/d	79%	(Zehentgruber et al., 2010)
B. megaterium MS941 (pSMF2.1_CAA)	Whole-cells in complex medium	11β-Keto-boswellic acid	n.d	100%	0.56 g/l/d	80%	(Bleif et al., 2012)
B. megaterium ATCC 13368	Whole-cells in complex medium	11β-Keto-boswellic acid	n.d	100%	0.37 g/l/d	80%	(Bleif et al., 2012)
B. megaterium ATCC 13368	Whole-cells in complex medium	Dipterocarpol	n.d	50%	0.15 g/L/d	n.d.	(Schmitz et al., 2012)
B. megaterium ATCC 13368	Whole-cells (resting cells)	Dehydro-epiandrosterone	n.d	100%	2.7 g/l/d	70%	(Schmitz et al., 2014)
B. megaterium MS941 (pSMF2.1_CAA)	Whole-cells (resting cells)	Dehydro-epiandrosterone	n.d.	100%	2.7 g/l/d	90%	(Schmitz et al., 2014)
B. megaterium MS941 (pSMF2.1_C)	Whole-cells in complex medium (resting cells - bioreactor)	Cyproterone acetate	n.d.	98 %	1.3 g/l/d	n.d.	(Kiss et al., 2015b)

Table 2: Substrates tested for *in vitro* conversion with CYP106A2

* = product formation not detected

n.d. = hydroxylation position not determined

	Compound tested	Hydroxylation position (product formation observed)	High-spin-Shift	Reference
1.	deoxycorticosterone	15β, 7β	yes	(Berg et al., 1976), (Kiss et al., 2015c)
2.	11-deoxycortisol	15β	yes	(Kiss et al., 2015c)
3.	cortisol	n.d.	yes	(Kiss et al., 2015c)
4.	cortisone	15β	yes	(Kiss et al., 2015c)
5.	6-fluoric-16-methyl-deoxycorticosterone	15β	yes	[17]
6.	corticosterone	15β	yes	(Berg et al., 1976), (Kiss et al., 2015c)
7.	cholestenone	15β	yes	(Lisurek et al., 2004)
8.	progesterone	15β (11α,6β,9α)	yes	(Berg et al., 1976), (Kiss et al., 2015c)
9.	17α-hydroxyprogesterone	15β	yes	(Berg et al., 1976), (Kiss et al., 2015c)
10.	20α-dihydroprogesterone	15β	no	(Berg et al., 1976)
11.	androstenedione	15β, 7β	yes	(Berg et al., 1976), (Kiss et al., 2015c)
12.	testosterone	15β	yes	(Berg et al., 1976)
13.	17α-methyltestosterone	15β	yes	[64], (Kiss et al., 2015c)
14.	19-nortestosterone	n.d.	no	(Kiss et al., 2015c)
15.	spironolacton	n.d	no	(Lisurek et al., 2004)
16.	deoxycholic acid	n.d.	no	(Schmitz, 2014, unpublished data)
17.	4-pregnen-20β-ol-3-one	n.d.	no	(Bleif, 2012)
18.	ethisterone	n.d.	yes	(Kiss et al., 2015c)
19.	digitoxigenin	n.d	no	(Schmitz et al., 2014)
20.	prednisone	n.d.	yes	(Kiss et al., 2015c)
21.	prednisolone	n.d.	yes	[65]
22.	dexamethasone	n.d	yes	(Kiss et al., 2015a)
23.	dehydroepiandrosterone	7β(7α)	yes	(Schmitz et al., 2014), (Kiss et al., 2015c)
24.	pregnenolone	7β	no	(Schmitz et al., 2014), (Kiss et al., 2015c)
25.	17α-hydroxypregnenolone	7β(7α)	yes	(Schmitz et al., 2014), (Kiss et al., 2015c)
26.	cyproterone acetate	15β	no	(Kiss et al., 2015b)
27.	19-hydroxy-4-androsten-3,17-dione	not metabolized	yes	(Bleif, 2012), (Kiss et al., 2015c)
28.	4-androstene-11β-ol-3,17-dione	n.d.	yes	(Kiss et al., 2015a)
29.	adrenosterone	not metabolized	yes	(Kiss et al., 2015c)
30.	stigmasterol	not metabolized	n.d	(Schmitz et al., 2014)
31.	cholesterol	not metabolized*	no	(Schmitz et al., 2014)
32.	estrone	not metabolized	no	(Schmitz et al., 2014), (Kiss et al., 2015c)

33.	estradiol	not metabolized	no	(Schmitz et al., 2014), (Kiss et al., 2015c)
34.	bufalin	not metabolized*	no	[66]
35.	cinobufagin	not metabolized*	no	[66]
36.	ecdysone	not metabolized*	no	[66]
37.	gitoxigenin	not metabolized*	no	[66]
38.	abietic acid	12α,12β	no	(Bleif et al., 2011)
39.	dihydroquinone pimaric acid	n.d.	yes	(Bleif et al., 2011)
40.	11-keto-β-boswellic acid	15α	yes	(Bleif et al., 2012)
41.	betulinic acid	n.d.	yes	(Bleif, 2012)
42.	oleanolic acid	n.d.	yes	(Bleif, 2012)
43.	ursolic acid	n.d.	yes	(Bleif, 2012)
44.	glycyrrhetinic acid	7β	yes	(Bleif, 2012)
45.	betuline	n.d.	yes	(Schmitz et al., 2012)
46.	melengestrolacetate	n.d	no	(Bleif, 2012)
47.	4-pregnen-17alpha-20alpha-21-triol-3-one	n.d	no	(Bleif, 2012)
48.	sarsasapogenin	not metabolized	yes	(Schmitz, 2013)
49.	panaxadiol	not metabolized*	yes	(Schmitz, 2013)
50.	rhamnetin	not metabolized	yes	(Schmitz, 2013)
51.	geraldol	not metabolized	yes	(Schmitz, 2013)
52.	noscapine	not metabolized	yes	(Schmitz, 2013)
53.	alpha-apo-oxytetracyclin	not metabolized	yes	(Schmitz, 2013)
54.	3-acetyl-11-keto-β-boswellic acid	not metabolized	yes	(Bleif, 2012)
55.	methyl-dihydrochinonpimate	not metabolized	no	(Bleif, 2012)
56.	nootkatone	not metabolized	no	Schmitz unpublished data
57.	flavone	not metabolized	no	Schmitz unpublished data
58.	santonine	not metabolized	no	Schmitz unpublished data
59.	nile red	not metabolized	no	Schmitz unpublished data
60.	nile blue	not metabolized	no	Schmitz unpublished data
61.	coumarin	not metabolized	no	(Bleif, 2012)
62.	7-methoxy-4-trifluormethylcoumarin	not metabolized	no	(Bleif, 2012)
63.	7-methoxy-4-methylcoumarin	not metabolized	no	(Bleif, 2012)
64.	4-methyl-3-phenylcoumarin	not metabolized	no	(Bleif, 2012)
65.	mifepristone	not metabolized	no	(Bleif, 2012)
66.	alpha-naphtoflavone	not metabolized	no	(Schmitz, 2013)
67.	friedelin	not metabolized*	no	(Schmitz, 2013)
68.	solanidine	not metabolized*	no	(Schmitz, 2013)
69.	solasodine	not metabolized*	no	(Schmitz, 2013)
70.	ethoxyresorufin	Not metabolized	no	(Berg and Rafter, 1981)

71.	lidocaine	Not metabolized	no	(Berg and Rafter, 1981)
72.	biphenyl	Not metabolized	no	(Berg and Rafter, 1981)
73.	octadecanoic acid	Not metabolized	no	(Berg and Rafter, 1981)
74.	hexadecanoic acid	Not metabolized	no	(Berg and Rafter, 1981)
75.	aniline	n.d.	no	(Berg and Rafter, 1981)
76.	imipramine	deamination reaction	no	(Berg and Rafter, 1981)

Table 3: Overview of all substrates tested *in-vitro* and in whole-cell systems using CYP106A2 and its mutants

Reaction	Test system	Reaction temperature	Specific activity (per mg protein)	K_M (µM)	V_{max} (nmol/min /mol P450) or k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$) x 10^3	Reference
Abietic acid → 12α-Abietic acid	CYP106A2*	30 °C	n.d.	106.9	13.5	n.d.	(Bleif et al., 2011)
Abietic acid → 12β-Abietic acid	CYP106A2*	30 °C	n.d.	132.3	8.0	n.d.	(Bleif et al., 2011)
4-Androstene-3,17-dione → 15β-Hydroxyandrostenedione	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	0.7 nmol/ 30 min	625	n.d.	n.d.	(Berg et al., 1976)
Aniline → p-Aminophenol	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	n.d.	68	0.43	n.d.	(Berg et al., 1981)
Corticosterone → 15β-Hydroxycorticosterone	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	0.006 nmol /30 min	95	n.d.	n.d.	(Berg et al., 1976)
Cyproterone acetate → 15β-Hydroxycyproterone acetate	CYP106A2 WT*	30 °C	n.d.	103.1	61.65	n.d.	(Kiss et al., 2015)
Dehydroepiandrosterone → 3β,7β-Dihydroxyandrost-4-ene-17-one	CYP106A2*	30 °C	n.d.	168	71	n.d.	(Schmitz et al., 2014)
Dehydroepiandrosterone → 3β,7α-Dihydroxyandrost-4-ene-17-one	CYP106A2*	30 °C	n.d.	98	41	n.d.	(Schmitz et al., 2014)
Deoxycorticosterone → 15β-Hydroxydeoxycorticosterone	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	2.0 nmol/ 30 min	60	n.d.	n.d.	(Berg 1976 #67)
	CYP106A2 WT*	30 °C	n.d.	1.5	246	n.d.	(Lisurek et al., 2008)
	CYP106A2 E90V/D185G*	30 °C	n.d.	1.9	208	n.d.	(Lisurek et al., 2008)
	CYP106A2 K27R/I71T/K215T*	30 °C	n.d.	2.4	265	n.d.	(Lisurek et al., 2008)

	CYP106A2 I86T*	30 °C	n.d.	1.2	254	n.d.	(Lisurek et al., 2008)
	CYP106A2 S394I*	30 °C	n.d.	4.7	46	n.d.	(Lisurek et al., 2008)
	CYP106A2 A395I*	30 °C	n.d.	4.2	104	n.d.	(Lisurek et al., 2008)
	CYP106A2 G397P*	30 °C	n.d.	4.6	5.3	n.d.	(Lisurek et al., 2008)
	CYP106A2 Q398S*	30 °C	n.d.	5.6	75	n.d.	(Lisurek et al., 2008)
11-Deoxycortisol → 15β-Hydroxydeoxycortisol	CYP106A2 WT*	30 °C	n.d.	93.4	172.1	30.7	(Virus et al., 2008)
	CYP106A2 D153V/121F*	30 °C	n.d.	37.3	181.8	81.2	(Virus et al., 2008)
	CYP106A2 A106T/R409L*	30 °C	n.d.	67.5	374.1	92.4	(Virus et al., 2008)
	CYP106A2 D217V*	30 °C	n.d.	65.9	325.2	82.3	(Virus et al., 2008)
	CYP106A2 A106T*	30 °C	n.d.	58.8	172	30.7	(Virus et al., 2008)
20α-Dihydroprogesterone → 15β-Hydroxy-20α-dihydroprogesterone	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	1.4 nmol /30 min	450	n.d.	n.d.	(Berg 1976 #67)
Dipterocarpol → 7β,11α-Dihdyroxydipterocarpol	CYP106A2*	30 °C	n.d	75.4	18.7	n.d.	(Schmitz et al., 2014)
6α-fluor-16α-Methyl-deoxycorticosterone → 15β-Hydroxy-6α-fluor-16α-methyl-deoxycorticosterone	Cell-free extracts: *E. coli* expressing CYP106A2 + *B. megaterium* ΔCYP106A2	RT	560 pmol/h	n.d.	n.d.	n.d.	(Rauschenbach et al., 1993)
	Cell-free extracts: *E.coli* expressing CYP106A2 + *B. subtilis* 168	RT	220 pmol/h	n.d.	n.d.	n.d.	(Rauschenbach et al., 1993)
	Cell-free extracts: *B. subtilis* expressing CYP106A2	RT	95 pmol/h	n.d.	n.d.	n.d.	(Rauschenbach et al., 1993)

17α-Hydroxyprogsterone → 15β,17α-Dihydroxyprogesterone	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	1.4 nmol/30 min	560	n.d.	n.d.	(Berg et al., 1976)
11β-Keto-boswellic acid → 15α-Hydroxy-11β-keto-boswellic acid (and side products)	CYP106A2*	30 °C	n.d	106.3	97.4	n.d.	(Bleif et al., 2011)
Progesterone → 6β-Hydroxyprogesterone	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	0.4 nmol/30 min	--	n.d.	n.d	(Berg et al., 1976)
Progesterone → 11α-Hydroxyprogesterone	CYP106A2 WT*	30 °C	n.d.	203	0.3 s^{-1}	1.4	(Nguyen et al., 2012)
	CYP106A2 A395I*	30 °C	n.d.	32	1.3 s^{-1}	42	(Nguyen et al., 2012)
	CYP106A2 A106T/A395I*	30 °C	n.d.	91	5 s^{-1}	55	(Nguyen et al., 2012)
	CYP106A2 A106T/A395I/R409L*	30 °C	n.d.	46	6.9 s^{-1}	151.4	(Nguyen et al., 2012)
	CYP106A2 T89N/A395I*	30 °C	n.d.	84	2.9 s^{-1}	34.2	(Nguyen et al., 2012)
	CYP106A2 T89N/A106T/A395I/R409L*	30 °C	n.d.	109	2.5 s^{-1}	23.2	(Nguyen et al., 2012)
	CYP106A2 M155I/F165L/A395I*	30 °C	n.d.	32	6.4 s^{-1}	201.8	(Nguyen et al., 2012)
	CYP106A2 D217V/A395I*	30 °C	n.d.	28	4.3 s^{-1}	151	(Nguyen et al., 2012)
	CYP106A2 A243V/A395I*	30 °C	n.d.	20	1.1 s^{-1}	53.1	(Nguyen et al., 2012)
	CYP106A2 T247V/A395I*	30 °C	n.d.	38	1.9 s^{-1}	49.9	(Nguyen et al., 2012)
	CYP106A2 T248V/A395I*	30 °C	n.d.	52	3.5 s^{-1}	66.8	(Nguyen et al., 2012)
	CYP106A2 A395I/G397K*	30 °C	n.d.	55	1.2 s^{-1}	21.7	(Nguyen et al., 2012)
	CYP106A2 A106T/A295W/G379K/R409L*	30 °C	n.d.	48	4.5 s^{-1}	93.4	(Nguyen et al., 2012)
	CYP106A2 A243V/A395W/G379K*	30 °C	n.d.	26	0.9 s^{-1}	34.5	(Nguyen et al., 2012)

	CYP106A2 T247W/A395W/G379K*	30 °C	n.d.	39	2.7 s^{-1}	70.6	(Nguyen et al., 2012)
Progesterone → 15β - Hydroxyprogesterone	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	1.6 nmol/30 min	143	n.d.	n.d.	(Berg et al., 1976)
	CYP106A2**	22 °C	n.d	n.d.	0.8	n.d.	(Berg et al., 1979)
	CYP106A2**	22 °C	n.d	n.d.	0.12	n.d.	(Berg et al., 1979)
	CYP106A2**	22 °C	n.d	n.d.	0.4	n.d.	(Berg et al., 1979)
	CYP106A2 WT*	30 °C	n.d.	1.5	233	n.d.	(Lisurek et al., 2008)
	CYP106A2 E90V/D185G*	30 °C	n.d.	1.0	205	n.d.	(Lisurek et al., 2008)
	CYP106A2 K27R/I71T/K215T*	30 °C	n.d.	0.7	276	n.d.	(Lisurek et al., 2008)
	CYP106A2 I86T*	30 °C	n.d.	0.9	280	n.d.	(Lisurek et al., 2008)
	CYP106A2 S394I*	30 °C	n.d.	4.3	89	n.d.	(Lisurek et al., 2008)
	CYP106A2 A395I*	30 °C	n.d.	2.4	186	n.d.	(Lisurek et al., 2008)
	CYP106A2 G397P*	30 °C	n.d.	5.0	15	n.d.	(Lisurek et al., 2008)
	CYP106A2 Q398S*	30 °C	n.d.	4.9	155	n.d.	(Lisurek et al., 2008)
Testosterone → 15β-Hydroxytestosterone	Cell-free extracts of *B. megaterium* ATCC13368	22 °C	0.4 nmol/30 min	1110	n.d.	n.d.	(Berg et al., 1976)

*reconstituted *in vitro* test system consisting of bovine adrenodoxin reductase, bovine adrenodoxin$_{4-108}$ and the respective CYP106A2 variant.
**reconstituted *in vitro* test system consisting of *B. megaterium* ferredoxin reductase, *B. megaterium* ferredoxin and CYP106A2.

References

[1] R. Bernhardt, Cytochromes P450 as versatile biocatalysts, J. Biotechnol., 124 (2006) 128-145.
[2] R. Bernhardt, V.B. Urlacher, Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations, Appl. Microbiol. Biotechnol., 98 (2014) 6185-6203.
[3] P. Hlavica, Models and mechanisms of O-O bond activation by cytochrome P450, Eur. J. Biochem., 271 (2004) 4335-4360.
[4] D. Nelson, The cytochrome P450 homepage, Hum. Genomics, 4 (2009) 59-65.
[5] D.R. Nelson, Cytochrome P450 diversity in the tree of life, BBA - Proteins and Proteomics, (2017). In press.
[6] S.L. Kelly, D.E. Kelly, Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us?, Philos. Trans. R. Soc. London, Ser. B., 368 (2013) 20120476.
[7] T. Omura, Contribution of cytochrome P450 to the diversification of eukaryotic organisms, Biotechnol. Appl. Biochem., 60 (2013) 4-8.
[8] T. Omura, Structural diversity of cytochrome P450 enzyme system, J. Biochem., 147 (2010) 297-306.
[9] A. Kawashima, Y. Satta, Substrate-Dependent Evolution of Cytochrome P450: Rapid Turnover of the Detoxification-Type and Conservation of the Biosynthesis-Type, PloS one, 9 (2014) e100059.
[10] L. Podust, D. Sherman, Diversity of P450 enzymes in the biosynthesis of natural products, Nat. Prod. Rep., 29 (2012) 1251-1266.
[11] S. Sinha, X. Shen, F. Gallazzi, Q. Li, J.W. Zmijewski, J.R. Lancaster, K.S. Gates, Generation of Reactive Oxygen Species Mediated by 1-Hydroxyphenazine, a Virulence Factor of *Pseudomonas aeruginosa*, Chem. Res. Toxicol., 28 (2015) 175-181.
[12] V.B. Urlacher, S. Eiben, Cytochrome P450 monooxygenases: perspectives for synthetic application, Trends Biotechnol., 24 (2006) 324-330.
[13] F. Hannemann, A. Bichet, K.M. Ewen, R. Bernhardt, Cytochrome P450 systems--biological variations of electron transport chains, Biochim. Biophys. Acta., 1770 (2007) 330-344.
[14] A. Berg, J. Gustafsson, M. Ingelman-Sundberg, Characterization of a cytochrome P-450-dependent steroid hydroxylase system present in *Bacillus megaterium*, J. Biol. Chem., 251 (1976) 2831-2838.
[15] E. Brill, F. Hannemann, J. Zapp, G. Bruning, J. Jauch, R. Bernhardt, A new cytochrome P450 system from *Bacillus megaterium* DSM319 for the hydroxylation of 11-keto-beta-boswellic acid (KBA), Appl. Microbiol. Biotechnol., 98 (2014) 1701-1717.
[16] O. Gotoh, 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 (1992) 83-90.
[17] R. Rauschenbach, M. Isernhagen, C. Noeske-Jungblut, W. Boidol, G. Siewert, Cloning sequencing and expression of the gene for cytochrome P450meg, the steroid-15β-monooxygenase from *Bacillus megaterium* ATCC 13368, Molec. Gen. Genet., 241 (1993) 170-176.
[18] S. Bleif, F. Hannemann, J. Zapp, D. Hartmann, J. Jauch, R. Bernhardt, 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 (2012) 1135-1146.
[19] S. Janocha, D. Schmitz, R. Bernhardt, Terpene hydroxylation with microbial cytochrome P450 monooxygenases, Adv. Biochem. Eng. Biotechnol., 2015, pp. 148:215-150.
[20] S. Janocha, Y. Carius, M. Hutter, C.R.D. Lancaster, R. Bernhardt, Crystal Structure of CYP106A2 in Substrate-Free and Substrate-Bound Form, Chembiochem, 17 (2016) 852-860.
[21] S. Bleif, F. Hannemann, M. Lisurek, J. Kries, J. Zapp, M. Dietzen, I. Antes, R. Bernhardt, Identification of CYP106A2 as a regioselective allylic bacterial diterpene hydroxylase, Chembiochem, 12 (2011) 576-582.
[22] J.A. Labinger, Selective alkane oxidation: Hot and cold approaches to a hot problem, J. Mol. Catal., 220 (2004) 27-35.

[23] S. Schulz, M. Girhard, V.B. Urlacher, Biocatalysis: Key to Selective Oxidations, ChemCatChem, 4 (2012) 1889-1895.
[24] M.T. Lundemo, J.M. Woodley, Guidelines for development and implementation of biocatalytic P450 processes, Appl. Microbiol. Biotechnol., 99 (2015) 2465-2483.
[25] V.B. Urlacher, M. Girhard, Cytochrome P450 monooxygenases: an update on perspectives for synthetic application, Trends Biotechnol., 30 (2012) 26-36.
[26] V.B. Urlacher, S. Lutz-Wahl, R.D. Schmid, Microbial P450 enzymes in biotechnology, Appl. Microbiol. Biotechnol., 64 (2004) 317-325.
[27] D. Schmitz, J. Zapp, R. Bernhardt, Hydroxylation of the triterpenoid dipterocarpol with CYP106A2 from *Bacillus megaterium*, FEBS J., 279 (2012) 1663-1674.
[28] C. Virus, M. Lisurek, B. Simgen, F. Hannemann, R. Bernhardt, Function and engineering of the 15beta-hydroxylase CYP106A2, Biochem. Soc. Trans., 34 (2006) 1215-1218.
[29] M. Donova, O. Egorova, Microbial steroid transformations: current state and prospects, Appl. Microbiol. Biotechnol., 94 (2012) 1423-1447.
[30] W.Y. Tong, X. Dong, Microbial biotransformation: recent developments on steroid drugs, Recent Pat. Biotechnol., 3 (2009) 141-153.
[31] W.J. McAleer, T.A. Jacob, L.B. Turnbull, E.F. Schoenewaldt, T.H. Stoudt, Hydroxylation of progesterone by *Bacillus cereus* and *Bacillus megaterium*, Arch. Biochem. Biophys., 73 (1958) 127-130.
[32] A. Berg, K. Carlström, J.-Å. Gustafsson, M. Ingelman-Sundberg, Demonstration of a cytochrome P-450-dependent steroid 15β-hydroxylase in *Bacillus megaterium*, Biochem. Biophys. Res. Commun., 66 (1975) 1414-1423.
[33] A. Berg, M. Ingelman-Sundberg, J. Gustafsson, Isolation and characterization of cytochrome P-450meg, Acta. Biol. Med. Ger., 38 (1979) 333-344.
[34] A. Berg, M. Ingelman-Sundberg, J.A. Gustafsson, Purification and characterization of cytochrome P-450meg, J. Biol. Chem., 254 (1979) 5264-5271.
[35] A. Berg, J.J. Rafter, Studies on the substrate specificity and inducibility of cytochrome P-450meg, Biochem J., 196 (1981) 781-786.
[36] M. Lisurek, M. Kang, R. Hartmann, R. Bernhardt, Identification of monohydroxy progesterones produced by CYP106A2 using comparative HPLC and electrospray ionisation collision-induced dissociation mass spectrometry, Biochem. Biophys. Res. Commun., 319 (2004) 677-682.
[37] M. Lisurek, B. Simgen, I. Antes, R. Bernhardt, Theoretical and experimental evaluation of a CYP106A2 low homology model and production of mutants with changed activity and selectivity of hydroxylation, Chembiochem, 9 (2008) 1439-1449.
[38] M. Kang, M. Lisurek, R. Bernhardt, R. Hartmann, Use of high-performance liquid chromatography/electrospray ionization collision-induced dissociation mass spectrometry for structural identification of monohydroxylated progesterones, Rapid Commun. Mass. Spectrom., 18 (2004) 2795-2800.
[39] B. Simgen, J. Contzen, R. Schwarzer, R. Bernhardt, C. Jung, Substrate binding to 15beta-hydroxylase (CYP106A2) probed by FT infrared spectroscopic studies of the iron ligand CO stretch vibration, Biochem. Biophys. Res. Commun., 269 (2000) 737-742.
[40] F.M. Kiss, D. Schmitz, J. Zapp, T.F. Dier, D. Volmer, R. Bernhardt, Comparison of CYP106A1 and CYP106A2 from *Bacillus megaterium* – identification of a novel 11-oxidase activity, Appl. Microbiol. Biotechnol., (2015) 1-20.
[41] J. Schenkman, I. Jansson, Spectral analyses of cytochromes P450, Meth. Mol. Biol, 107 (1998) 25-33.
[42] J. Schenkman, S. Sligar, D. Cinti, Substrate interaction with cytochrome P-450., Pharmacol. Ther., 12 (1981) 43-71.
[43] D. Schmitz, J. Zapp, R. Bernhardt, Steroid conversion with CYP106A2 - production of pharmaceutically interesting DHEA metabolites, Microb. Cell Fact., 13 (2014) 81.

[44] P. Bagul, K.S. Khomane, A.K. Bansal, Investigating permeability related hurdles in oral delivery of 11-keto-β-boswellic acid, Int. J. Pharm., 464 (2014) 104-110.
[45] F. Hannemann, C. Virus, R. Bernhardt, Design of an *Escherichia coli* system for whole cell mediated steroid synthesis and molecular evolution of steroid hydroxylases, J. Biotechnol., 124 (2006) 172-181.
[46] D. Zehentgruber, F. Hannemann, S. Bleif, R. Bernhardt, S. Lütz, Towards preparative scale steroid hydroxylation with cytochrome P450 monooxygenase CYP106A2, Chembiochem, 11 (2010) 713-721.
[47] R. Bernhardt, The role of adrenodoxin in adrenal steroidogenesis, Curr. Opin. Endocrinol., 7(2000) 109–115.
[48] H.J. Ruijssenaars, E.M.G.M. Sperling, P.H.G. Wiegerinck, F.T.L. Brands, J. Wery, J.A.M. de Bont, Testosterone 15β-hydroxylation by solvent tolerant *Pseudomonas putida* S12, J. Biotechnol., 131 (2007) 205-208.
[49] M. Ringle, Y. Khatri, J. Zapp, F. Hannemann, R. Bernhardt, Application of a new versatile electron transfer system for cytochrome P450-based *Escherichia coli* whole-cell bioconversions, Appl. Microbiol. Biotechnol., 97 (2013) 7741-7754.
[50] P. Sperandeo, G. Dehò, A. Polissi, The lipopolysaccharide transport system of Gram-negative bacteria, BBA - Molecular and Cell Biology of Lipids, 1791 (2009) 594-602.
[51] S. Janocha, J. Zapp, M. Hutter, M. Kleser, J. Bohlmann, R. Bernhardt, Resin acid conversion with CYP105A1: an enzyme with potential for the production of pharmaceutically relevant diterpenoids, Chembiochem, 14 (2013) 467-473.
[52] F.M. Kiss, M.T. Lundemo, J. Zapp, J.M. Woodley, R. Bernhardt, Process development for the production of 15β-hydroxycyproterone acetate using *Bacillus megaterium* expressing CYP106A2 as whole-cell biocatalyst, Microb. Cell Fact., 14 (2015) 1-13.
[53] J. Tamm, M. Seckelmann, U. Volkwein, E. Ludwig, The effect of the antiandrogen IIα-hydroxyprogesterone on sebum production and cholesterol concentration of sebum, Br. J. Dermatol., 107 (1982) 63-70.
[54] A.H. van der Willigen, J.D. Peereboom-Wynia, T. van Joost, E. Stolz, A preliminary study of the effect of 11a-hydroxyprogesterone on the hair growth in men suffering from androgenetic alopecia., Acta Derm. Venereol., 67 (1987) 82-85.
[55] J.A. Hogg, P.F. Beal, A.H. Nathan, F.H. Lincoln, W.P. Schneider, B.J. Magerlein, A.R. Hanze, R.W. Jackson, The adrenal hormones and related compounds I. A “direct” synthesis of hydrocortisone acetate and cortisone acetate from 11α-hydroxyprogesterone, J. Am. Chem. Soc., 77 (1955) 4436-4438.
[56] D. Perlman, J.D. Dutcher, J. Fried, E.O. Titus, 9alpha-hydroxyprogesterone, U.S. patent no. 2,840,578., 1958.
[57] H. Yoshioka, S. Asada, S. Fujita, Process for production of 6b, 14ahydroxyandrost-4-ene-3,17-dione, European Patent 5: 99, 658., 1994.
[58] M.T. Reetz, Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future, J. Am. Chem. Soc., 135 (2013) 12480-12496.
[59] D. Appel, R. Schmid, C.-A. Dragan, M. Bureik, V. Urlacher, A fluorimetric assay for cortisol, Anal. Bioanal. Chem., 383 (2005) 182-186.
[60] C. Virus, R. Bernhardt, Molecular evolution of a steroid hydroxylating cytochrome P450 using a versatile steroid detection system for screening, Lipids, 43 (2008) 1133-1141.
[61] K. Nguyen, C. Virus, N. Gunnewich, F. Hannemann, R. Bernhardt, Changing the regioselectivity of a P450 from C15 to C11 hydroxylation of progesterone, Chembiochem, 13 (2012) 1161-1166.
[62] J. Nikolaus, K.T. Nguyen, C. Virus, J.L. Riehm, M. Hutter, R. Bernhardt, Engineering of CYP106A2 for steroid 9α- and 6β-hydroxylation, Steroids, 120 (2017) 41-48.
[63] D.R. Nelson, Cytochrome P450 Nomenclature, 2004, in: I.R. Phillips, E.A. Shephard (Eds.) Cytochrome P450 Protocols, Humana Press, 2006, pp. 1-10.

[64] S. Bleif, Von der Identifizierung neuer Substrate der 15beta-Hydroxylase (CYP106A2) zur Cytochrom P450 Ganzzellkatalyse in *Bacillus megaterium*, Saarländische Universitäts- und Landesbibliothek, Saarbrücken, 2012.
[65] F.M. Kiss, Y. Khatri, J. Zapp, R. Bernhardt, Identification of new substrates for the CYP106A1-mediated 11-oxidation and investigation of the reaction mechanism, FEBS Lett., 589 (2015) 2320-2326.
[66] D. Schmitz, Hydroxylierung bioaktiver Naturstoffe mit CYP106A2 aus *Bacillus megaterium* - Screening, Biokatalyse und Produktcharakterisierung, Saarländische Universitäts- und Landesbibliothek, Saarbrücken, 2013.