The FEBS Journal

FEBS PRESS science publishing by scientists

Characterization of two steroid hydroxylases from different *Streptomyces* spp. and their ligand-bound and -unbound crystal structures

Bikash Dangi[1], Chang Woo Lee[2,3], Ki-Hwa Kim[1], Sun-Ha Park[2], Eun-Ji Yu[1], Chang-Sook Jeong[2,3], Hyun Park[2,3], Jun Hyuck Lee[2,3] and Tae-Jin Oh[1,4,5]

1 Department of Life Science and Biochemical Engineering, Sunmoon University, Asansi, Korea
2 Unit of Polar Genomics, Korea Polar Research Institute, Incheon, Korea
3 Department of Polar Sciences, University of Science and Technology, Incheon, Korea
4 Genome-based BioIT Convergence Institute, Asansi, Korea
5 Department of Pharmaceutical Engineering and Biotechnology, Sunmoon University, Asansi, Korea

Keywords
crystal structure; cytochrome P450; steroid hydroxylase; *Streptomyces*; X-ray crystallography

Correspondence
J. H. Lee, Unit of Polar Genomics, Korea Polar Research Institute, Incheon 21990, Korea
Tel: +82 32 760 5555
E-mail: junhyucklee@kopri.re.kr
and
T.-J. Oh, Department of Pharmaceutical Engineering and Biotechnology, Sunmoon University, Asansi 31460, Korea
Tel: +82 41 530 2677
E-mail: tjoh3782@sunmoon.ac.kr

Bikash Dangi and Chang Woo Lee contributed equally to this work.

(Received 9 July 2018, revised 20 November 2018, accepted 9 December 2018)

doi:10.1111/febs.14729

Bacterial cytochrome P450 (CYP) enzymes are involved in the hydroxylation of various endogenous substrates while using a heme molecule as a cofactor. CYPs have gained biotechnological interest as useful biocatalysts capable of altering chemical structures by adding a hydroxyl group in a regiospecific manner. Here, we identified, purified, and characterized two CYP154C4 proteins from *Streptomyces* sp. W2061 (*St*CYP154C4-1) and *Streptomyces* sp. ATCC 11861 (*St*CYP154C4-2). Activity assays showed that both *St*CYP154C4-1 and *St*CYP154C4-2 can produce 2′-hydroxylated testosterone, which differs from the activity of a previously described *Nf*CYP154C5 from *Nocardia farcinica* in terms of its 16α-hydroxylation of testosterone. To better understand the molecular basis of the regioselectivity of these two CYP154C4 proteins, crystal structures of the ligand-unbound form of *St*CYP154C4-1 and the testosterone-bound form of *St*CYP154C4-2 were determined. Comparison with the previously determined *Nf*CYP154C5 structure revealed differences in the substrate-binding residues, suggesting a likely explanation for the different patterns of testosterone hydroxylation, despite the high sequence similarities between the enzymes (54% identity). These findings provide valuable insights that will enable protein engineering for the development of artificial steroid-related CYPs exhibiting different regiospecificity.

Heme-containing cytochrome P450 enzymes (CYPs) are found in all kingdoms. CYPs are usually termed P450, because they have an absorption-maximum at ~ 450 nm following their reaction with carbon monoxide (CO). CYPs have been frequently studied since their discovery based on their ability to synthesize and metabolize various cellular molecules. Additionally, CYPs are attractive due to their capabilities of hydroxylation, epoxidation, decarboxylation, and dealkylation of endogenous and xenobiotic compounds, with these

Abbreviations
CO, carbon monoxide; CYP, cytochrome P450; Fdr, ferredoxin reductase; Fdx, ferredoxin; LC-MS, liquid chromatography mass spectrometry; PhIO, iodosobenzene; PIDA, diacetoxyiodobenzene.

reactions targeting various molecules, including cholesterol, sterols, alkanes, and fatty acids [1–4]. In humans, CYPs are key enzymes involved in drug metabolism and are targets for drug design; however, investigations tend to focus on bacterial rather than mammalian CYPs due to the insolubility of membrane-bound eukaryotic CYPs, which limits their study. By contrast, soluble bacterial CYPs are easily overexpressed and engineered, and CYP mutagenesis is highly effective at increasing product yield. Therefore, bacterial CYPs have been the basis for studies involving the drug-development industry [5–9]. CYP enzymes are classified into two major groups based on their localization and redox partners. Class I includes mitochondrial and bacterial P450s, which use two separate redox partners: an iron–sulfur protein (ferredoxin and adrenodoxin) and a flavin-containing reductase (ferredoxin reductase and adrenodoxin reductase). Class II enzymes localize to microsomes and receive electrons from NADPH-CYP reductase (CPR) [10]. There are also other minor classes of P450 enzymes, such as P450BM3 (CYP102A1), which does not require separate redox partners due to the fact that a diflavin reductase partner is contained within the polypeptide chain [11]. CYP154 family members have been identified in several Actinobacteria, such as *Streptomyces coelicolor*, *Streptomyces griseus*, and *Nocardia farcinica*. They oxidatively hydroxylate steroid molecules regio- and stereoselectively [12–14].

Steroids are widespread organic compounds that are essential for multiple functions in plants and animals [15]. They are widely used as 'anti-agents' for their anti-inflammatory, antimicrobial, antifungal, and antiviral effects [16]. Modification by hydroxylation can strongly affect the biological activity, polarity, and toxicity of steroids. The position at which a steroid is hydroxylated is of great importance; in particular, the presence of hydroxyl or ketone groups at carbons 3, 11, 17, and 21 correlate with biological activity. Commercially, 11α-hydroxylation of progesterone to yield cortisone and 11β-hydroxylation of 11-deoxycortisol to form hydrocortisone have been established at an industrial scale [17,18]. Steroids are used as a source of nutrients by bacteria, in which steroid hydroxylase is involved in some steroid catabolic pathway reactions for the degradation of an aliphatic side chain or opening of a sterane ring [19]. A number of P450s from diverse bacterial sources are known to be involved in steroid hydroxylation, including CYP106A family (CYP106A1 and CYP106A2), CYP154C family (CYP154C3, CYP154C5, and CYP154C8), CYP260 family (CYP260A1 and CYP260B1), and CYP109 family (CYP109B1 and CYP109E1) members [20]. The mycobacterium P450s CYP125A1 and CYP142A1 are involved in cholesterol degradation via hydroxylation and its sequential oxidation to aldehyde and carboxylic acid [21]. In one particularly interesting study, the screening of 250 bacterial P450s expressed in *Escherichia coli* identified 24 P450s that were able to hydroxylate testosterone at different positions in a regio- and stereoselective manner [22].

In the CYP154 family, two crystal structures (CYP154A1 and CYP154C1) have been reported from *S. coelicolor* A3(2) [12,23]. Additionally, various steroid (pregnenolone, progesterone, androstenedione, and testosterone)-bound forms of CYP154C5 from *N. farcinica* have been solved, with these proteins sharing 93% amino acid sequence identity [24]. Interestingly, CYP154A1 structure analysis revealed that the heme orientation was 180° opposite to that of most other CYP structures, including CYP154C1; however, the reason for this and its biological meaning remain unclear. In steroid-bound CYP154C5 structures, the C16 position of bound steroid molecules faces the heme iron. This structural feature possibly explains the regio- and stereoselectivity of CYP154C5 in the conversion of steroid molecules to yield 16α-hydroxylated steroid products [24,25].

In this study, we report two CYP154C4 crystal structures from *Streptomyces* spp., namely the ligand-unbound form from *Streptomyces* sp. W2061 (*St*CYP154C4-1) and the testosterone-bound form from *Streptomyces* sp. ATCC 11861 (*St*CYP154C4-2). Notably, crystals of the testosterone-bound form of *St*CYP154C4-1 and apo-*St*CYP154C4-2 could not be obtained, despite their high sequence identity (93%). Structural comparison of the two obtained structures revealed different structural rearrangement of the α13-helix region according to testosterone binding. Additionally, biochemical activity assays showed that *St*CYP154C4-1 and *St*CYP154C4-2 yielded a 2′-hydroxylated testosterone product through a mechanism that differs from that of other CYP154 proteins. Furthermore, these structures show different active site residue compositions that might explain variations in the regiospecificity of the testosterone product among different CYP154 family members. Our results provide useful insights for future protein engineering and applications of this enzyme.

Results and Discussion

Expression, purification, and spectrophotometric characterization of *St*CYP154C4s

Both CYPs, *St*CYP154C4-1 and *St*CYP154C4-2, were expressed in their soluble form in *E. coli* BL21(DE3) cells. Each fraction eluted using different concentrations

of imidazole-containing potassium phosphate buffer was confirmed by SDS/PAGE analysis, which showed a band at ~ 65 kDa (Fig. 1A), although the theoretical molecular mass calculated for both CYPs was ~ 45 kDa. The difference in molecular masses was due to the Trx-His-s-enterokinase fusion sequence in the pET32a(+) vector transcribed and translated along with the *St*CYP154C4s sequence. *St*CYP154C4s purity was determined by RZ value, which was calculated as the ratio of absorbance at λ_{max} of the Soret band to the absorbance at 280 nm, with those for *St*CYP154C4-1 and *St*CYP154C4-2 at 1.32 and 1.28, respectively, indicating a high degree of purity for both CYPs [26]. UV-visible absorption spectroscopy of the oxidized forms of *St*CYP154C4-1 and *St*CYP154C4-2 showed a Soret band at 417 nm. Similarly, in their reduced state and CO-bound form, both CYPs showed a peak at 447 nm (Fig. 1B,C), which represents the characteristic signature of CYP heme in its $Fe^{2+}CO$ complex form.

Substrate-binding assay

Optical-absorption spectroscopy has major advantages for the physiochemical characterization of CYP and has been extensively utilized to obtain concentration measurements, monitor substrate and inhibitor binding, and find different kinetics parameters. Substrate binding by CYPs causes the displacement of water as a sixth ligand, which subsequently causes a spin shift of the heme iron from low (415–418 nm) to high (390–394 nm) [27]. *St*CYP154C4s showed this phenomenon upon binding to the testosterone substrate, with Soret absorption at 418 nm shifting to 392 nm upon addition of testosterone (Fig. 1D), indicating that both CYPs were capable of binding testosterone. Therefore, testosterone was selected as a substrate for *in vitro* reactions catalyzed by *St*CYP154C4s. Similarly, based on this, the dissociation constant (K_d) values for CYP154C4s were determined by titrating different concentrations of substrate (testosterone) until saturation

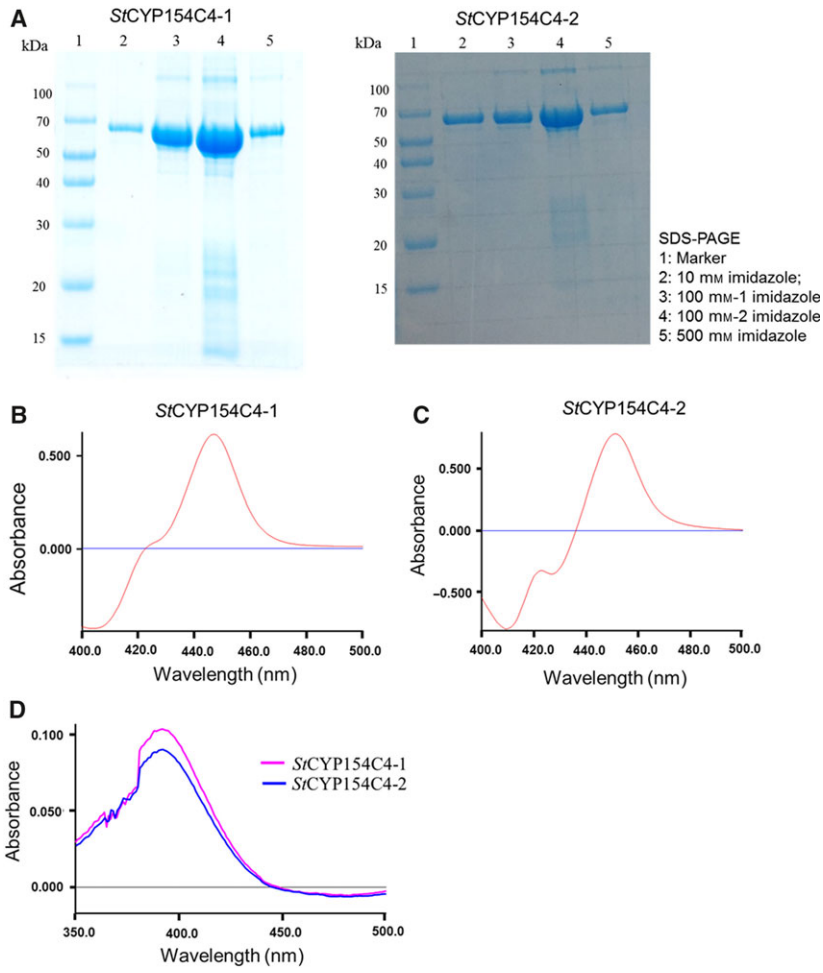

Fig. 1. Spectral analysis of *St*CYP154C4s. (A) SDS/PAGE of *St*CYP154C4-1 and *St*CYP154C4-2. (B,C) The dithionite-reduced CO-bound form of *St*CYP154C4-1 (B) and *St*CYP154C4-2 (C). (D) Substrate-binding shift in the spectra of *St*CYP154C4-1 (pink line) and *St*CYP154C4-2 (blue line). The maximum absorbance (Soret peak) was observed at 392 nm and characterized by a high-spin shift achieved due to substrate binding in the CYP active site.

(Fig. 2). The K_d value for *St*CYP154C4-1 was found to be less than 10 μM, indicating tight binding. *St*CYP154C4-2 had a higher K_d value than CYP154C4-1 (Table 1).

Enzyme activity and product identification

In a reaction mixture (reaction system I) containing 3 μM CYP, 0.025 U ferredoxin reductase (Fdr), 6 μM ferredoxin (Fdx), and 1 U glucose-6-phosphate dehydrogenase, the conversion of all three substrates (progesterone, testosterone, and androstenedione) was less than 10%. HPLC analysis of the progesterone reaction mixture showed a single product peak (Fig. 3A). Liquid chromatography mass spectrometry (LC-MS) analysis of the reaction mixture showed the exact mass of a hydroxylated progesterone (data not shown). The progesterone product was identified as 16α-hydroxyprogesterone by coelution with an authentic standard (Fig. 3A, panel III). The testosterone reaction mixture showed the formation of at least two products by both CYPs (Fig. 3B). Among them, one of the products (P1) was identified as 16α-hydroxytestosterone based on coelution with a standard (Fig. 3B, panel III). Based on NMR structural elucidation, the other product (P2) observed in the testosterone reaction mixture was identified as a mixture of 2α- and 2β-hydroxytestosterone (Fig. 3C).

Having found a very low conversion rate with reaction system I, in another reaction (reaction system II), increases in the concentrations of reaction constituents to 15 μM CYP, 15 μM Fdx, 0.1 U Fdr, and 5 U glucose-6-phosphate dehydrogenase resulted in not only high conversion (> 70%) rates for all three substrates but also variation in product distributions. Among the three products found in the progesterone reaction mixture (Fig. 4A), P2 was identified as monohydroxylated progesterone with an exact mass of $[M+H]^+$ 331.2291. Products P3 and P1 had exact masses of $[M+H]^+$ 329.2115 and $[M+H]^+$ 347.2222, respectively. Both of these products (P1 and P3) may have been obtained from the monohydroxylated product (P2), which was further supported by the time-dependent reaction results, in which the concentration of P2 decreased with an increase in the time interval, while the concentrations of the other two products (P1 and P3) increased with an increase in the time interval (Fig. 4B). P1 may have been derived from sequential hydroxylation of P2, while P3 may have resulted from dehydrogenation of the hydroxyl group, yielding a keto group. In another reaction (reaction system III), we reduced the concentrations used in reaction system II by half, and carried out the reaction catalyzed by *St*CYP154C4-1. This resulted in mostly a single monohydroxylated product (P2) that was found in reaction system II, while the other two products P1 and P3

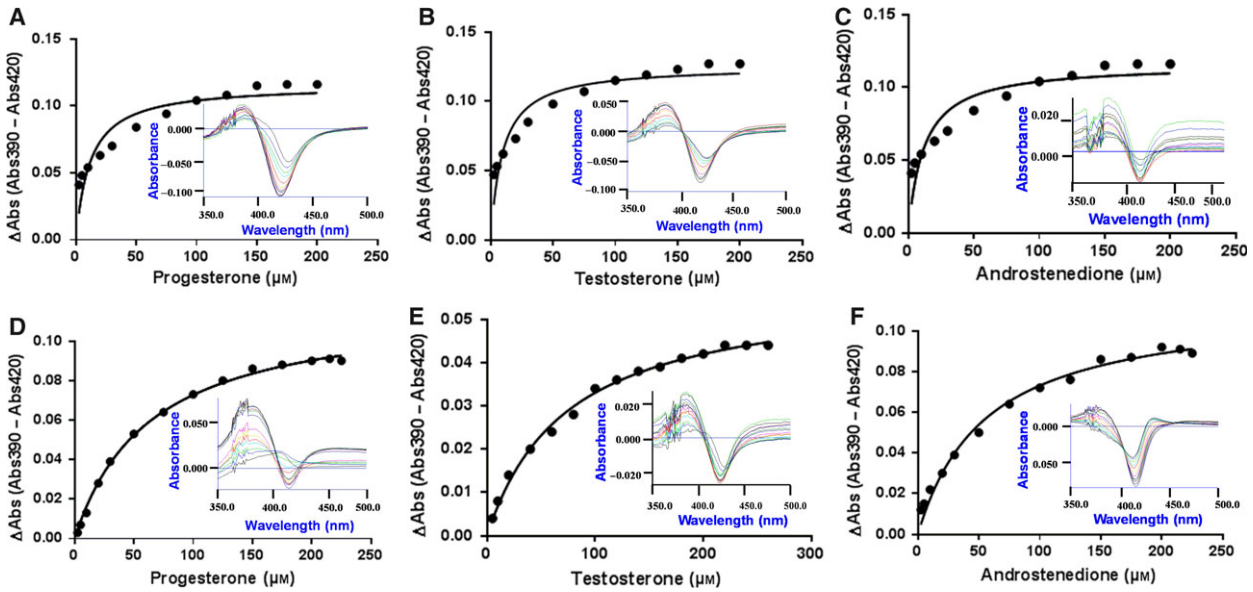

Fig. 2. Spectrometric absorbance of *St*CYP154C4-1 (A–C) and *St*CYP154C4-2 (D–F) bound with different concentrations of testosterone substrate. Dissociation constant (K_d) values were calculated by plotting the peak-to-trough ($Abs_{389} - Abs_{419}$) absorbance against different concentrations of the substrate. Upper panel shows the shift in spin of the heme induced by binding of different concentrations of substrate, and lower panel shows the plot of peak-to-trough absorbance against the corresponding concentration of substrate. (A,D) progesterone, (B,E) testosterone, (C,F) androstenedione.

exhibited very low conversion rates (Fig. 4C). This clearly indicates that P1 and P3 must have resulted from sequential oxidation of P2 by *St*CYP154C4-1. Surprisingly, the progesterone exclusively hydroxylated at the C16α-position (P) found in reaction system I was not detected in reaction system II, based on HPLC comparisons with the standard (16α-hydroxyprogesterone; Fig. 4D). Another substrate, testosterone, was converted into a number of products using reaction system II. P3 was identified as a major product of catalysis by *St*CYP154C4-1 (Fig. 5A, panel II); however, in a reaction catalyzed by *St*CYP154C4-2, P2 was observed as a major product (Fig. 5A, panel I). The P2 produced by reaction system II was identified as a mixture of 2α- and 2β-testosterone based on HPLC retention time comparisons (Fig. 5A, panel III). The possible 16α-hydroxytestosterone (P1) exhibited a very low rate of conversion compared to the other products. Androstenedione was converted into one major monohydroxylated product (P3) and two other minor products (P1 and P2) by *St*CYP154C4-1

Table 1. Dissociation constant (K_d) values of *St*CYP154C4-1 and *St*CYP154C4-2 with steroid substrates.

Substrate	*St*CYP154C4-1	*St*CYP154C4-C2
Progesterone	7.12 ± 1.08	62.76 ± 3.35
Androstenedione	11.35 ± 3.14	54.57 ± 6.33
Testosterone	9.11 ± 2.15	73.38 ± 5.21

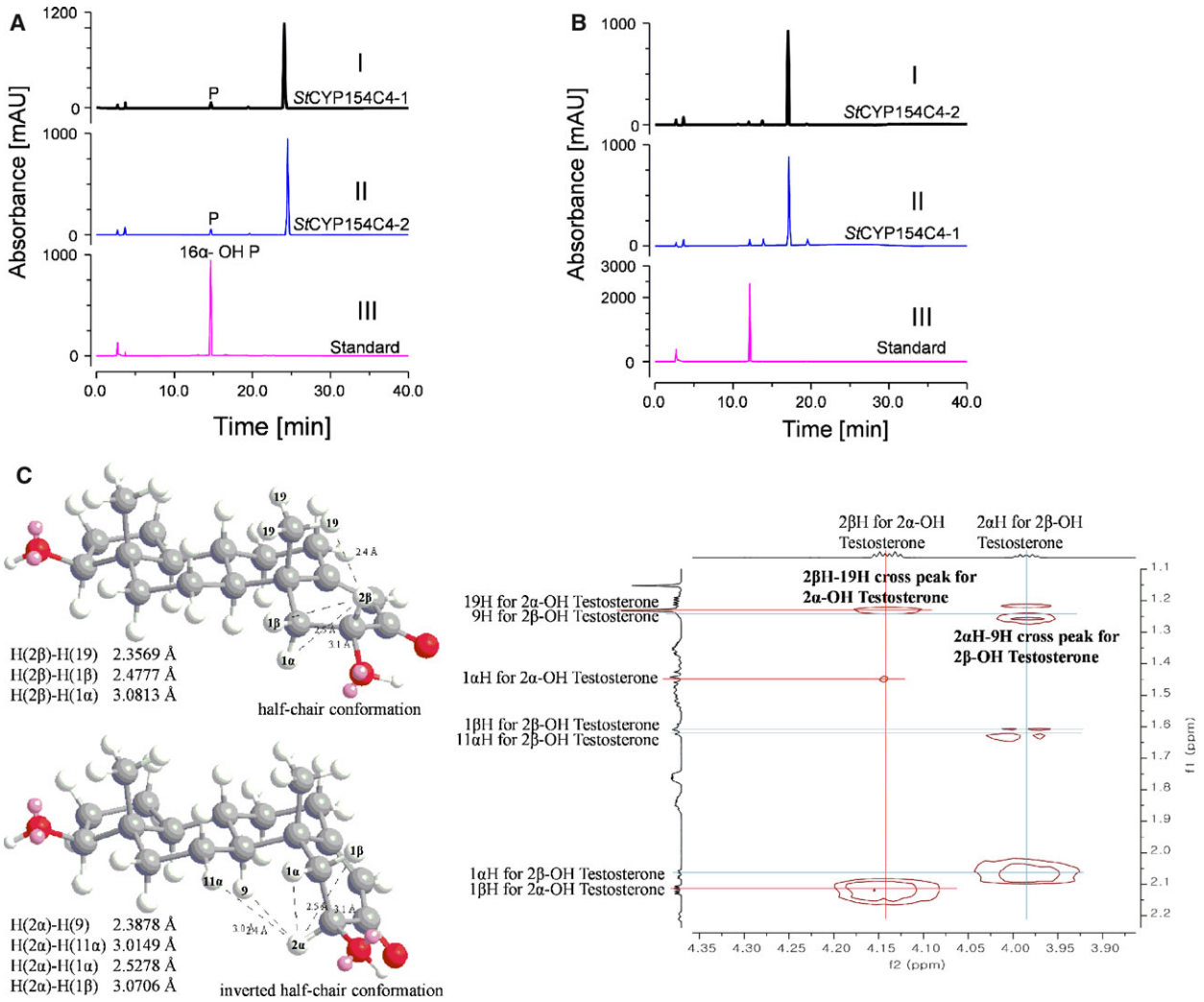

Fig. 3. HPLC chromatogram of the progesterone (A) and testosterone (B) reaction mixture catalyzed by *St*CYP154C4-1 (panel I) and *St*CYP154C4-2 (panel II) supported by reaction system I. P indicates the hydroxylated product of progesterone, which was identified as 16α-hydroxyprogesterone. (C) The ROSEY of the mixture of 2α-OH testosterone and 2β-OH testosterone. The distance between a 2α hydrogen and each hydrogen of 2α- and 2β-OH testosterone by force field MMFF94 optimization [52–56].

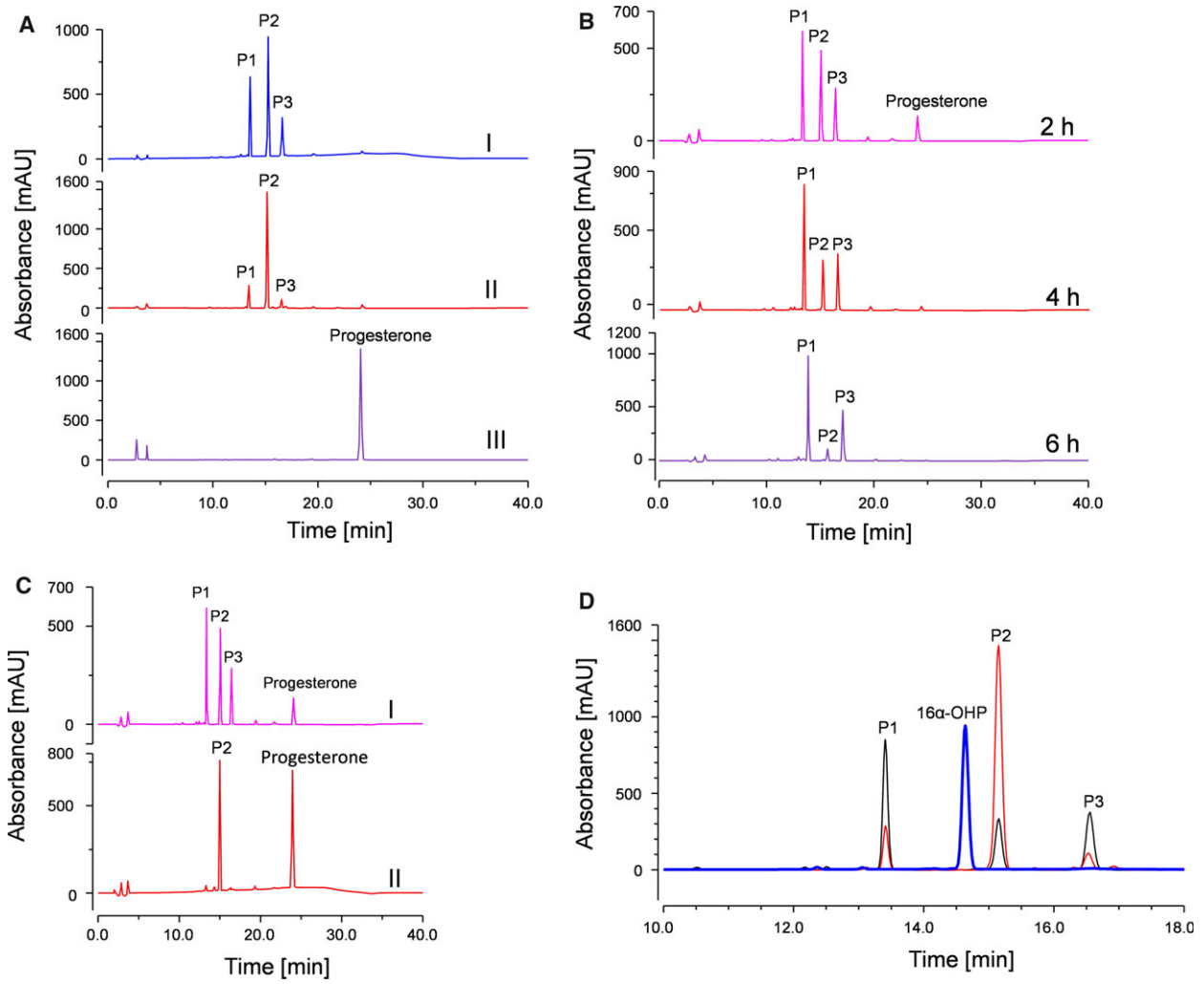

Fig. 4. (A) Catalytic oxidation of progesterone supported by reaction system II, and catalyzed by *St*CYP154C4-1 (panel I) and *St*CYP154C4-2 (panel II). (B) HPLC chromatograms of progesterone reaction mixture catalyzed by *St*CYP154C4-1 for different time intervals supported by reaction system II. The reaction interval is indicated (2, 4, and 6 h). (C) HPLC chromatogram comparison of progesterone reaction mixture catalyzed by *St*CYP154C4-1 supported by reaction system II (panel I) and reaction system III (panel II). (D) HPLC chromatogram of 16α-OH progesterone (authentic standard; indicated by blue color) in the progesterone reaction mixture catalyzed by *St*CYP154C4-1 (indicated by black color) and *St*CYP154C4-2 (indicated by red color) supported by reaction system II. P1, P2, P3, and 16α-OH P indicate product 1, product 2, product 3, and 16α-OH progesterone, respectively. LC-MS analysis of the products showed P1 as a dihydroxylated product ($[M+H]^+$ 347.2222), P2 as a monhydroxylated ($[M+H]^+$ 331.2291), and P3 as a possible oxidation product of the monohydroxylated product ($[M+H]^+$ 329.2115).

(Fig. 5B, panel I). The LC-MS analysis of P1 and P2 showed the exact masses of dihydroxy- and monohydroxyandrostenedione, respectively. In addition to three different androstenedione products formed by catalysis with *St*CYP154C4-1, *St*CYP154C4-2 also resulted in the formation of another monohydroxylated product (P4; Fig. 5B, panel II).

We also carried out an *in vitro* reaction supported by oxygen surrogates diacetoxyiodobenzene (PIDA) and hydrogen peroxide (H_2O_2). Iodosobenzene is a PhIO-activated compound similar to PIDA that has been used in many cases for P450-catalyzed *in vitro* reactions [28–31]. A PIDA-supported reaction catalyzed by both *St*CYP154C4s with a progesterone substrate resulted in major conversion (P) into 16α-hydroxyprogesterone, which was identified based on coelution with the standard (Fig. 5C). The *St*CYP154C4-2-catalyzed reaction with testosterone showed the formation of a single product (P4), which was identified as a mixture of 2α- and 2β-hydroxytestosterone (Fig. 5D, panel I). In a reaction catalyzed by *St*CYP154C4-1, minor conversion into 16α-hydroxytestosterone (P1) and major

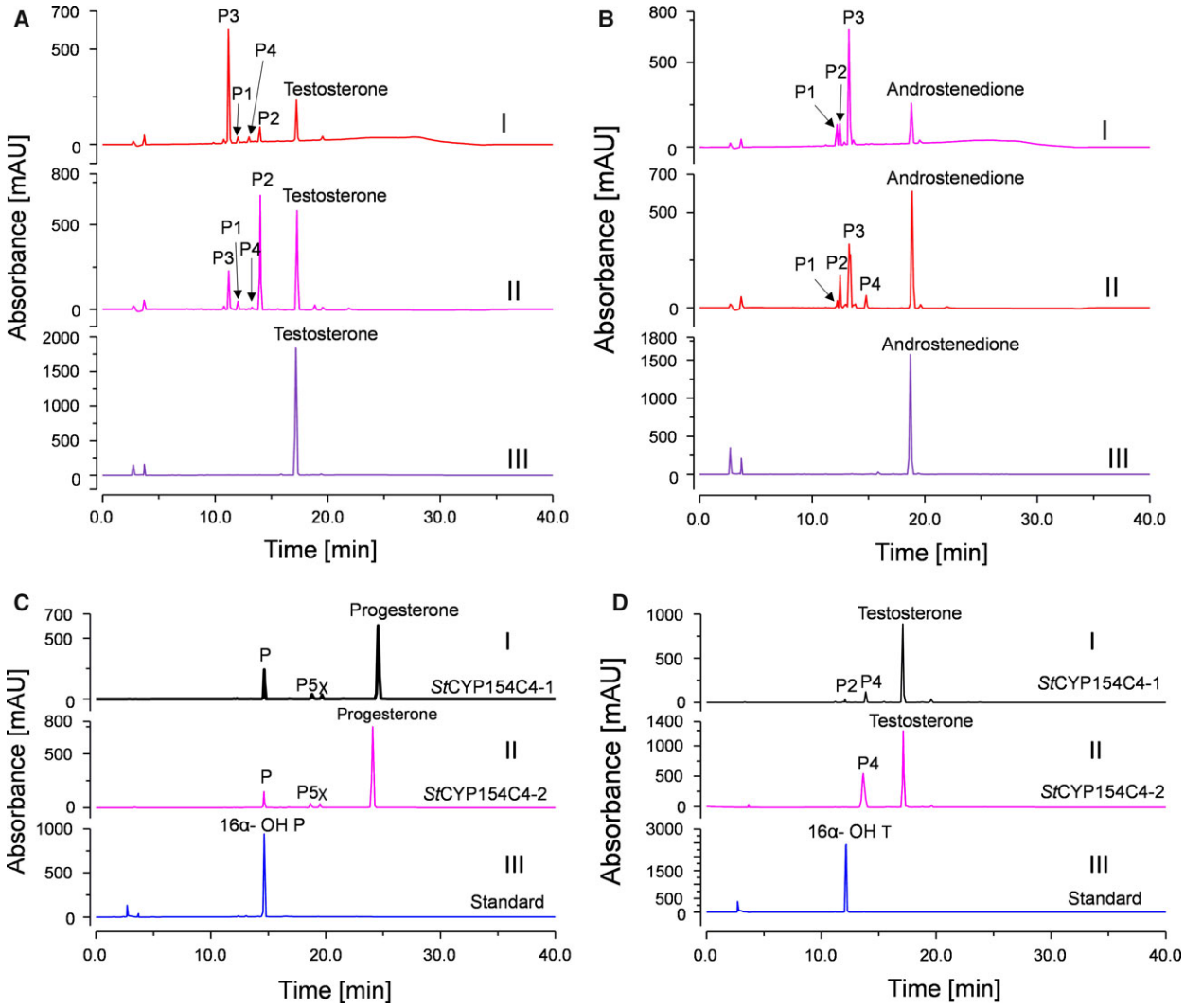

Fig. 5. (A) Catalytic oxidation of testosterone supported by reaction system II and catalyzed by *St*CYP154C4-1 (panel III) and *St*CYP154C4-2 (panel I). Panel II shows the chromatograms of the authentic standard (16α-hydroxytestosterone). P1, P2, P3, and P4 indicate product 1, product 2, product 3, and product 4 of testosterone, respectively. (B) HPLC chromatogram of androstenedione reaction mixture catalyzed by *St*CYP154C4-1 (panel I) and *St*CYP154C4-2 (panel II) supported by reaction system II. Panel III shows the chromatogram of the control reaction, carried out in the absence of NADPH. P1, P2, P3, and P4 indicate product 1, product 2, product 3, and product 4 of androstenedione, respectively. (C) HPLC chromatogram of the progesterone reaction mixture catalyzed by *St*CYP154C4-1 (panel I) and *St*CYP154C4-2 (panel II) supported by PIDA. Panel III shows the chromatogram of the authentic standard 16α-OH progesterone (16α-OH P). P indicates product which was identified as 16α-OH progesterone. (D) HPLC chromatogram of the testosterone reaction mixture catalyzed by *St*CYP154C4-1 (panel I) and *St*CYP154C4-2 (panel II) supported by PIDA. Panel III shows the chromatogram of the authentic standard 16α-OH progesterone (16α-OH P). P2 and P4 are two different product peaks of testosterone.

conversion into 2α- and 2β-hydroxytestosterone were observed (Fig. 5D, panel II). PhIO- and NADPH-supported reactions by P450 can be mechanistically distinct, leading to differences in product profiles [31]. A similar difference in product profiles was recently reported by us with CYP154C8-catalyzed reactions supported by NADPH or PhIO [32]. In addition to 16α-hydroxylation, a PIDA (PhIO system)-supported reaction by CYP154C8 also catalyzed the 6β-hydroxylation of progesterone and 11-ketoprogesterone, neither of which was observed when the reactions were supported by NADPH or NADH systems. In another report, CYP17A1 produced different products in the presence of iodosobenzene and NADPH-P450 reductase to form compound I [30]. MycG, a P450 enzyme, was also reported to display altered catalytic activity when

it was partnered with an alternative surrogate redox partner [33]. Recently, CYP106A2, when partnered with 11 different redox partners in separate reactions, exhibited differences in product patterns caused by differences in the binding mode of the redox partners [34]. The reaction in the presence of a high concentration (40 mM) of H_2O_2 showed low conversion rates (~ 5%) of testosterone and progesterone (data not shown). The major conversion via 2-hydroxylation was observed with testosterone, while the progesterone reaction mixture showed a single product peak corresponding to 16α-hydroxylation. Androstenedione was not converted to any product following catalysis with both CYPs in the presence of H_2O_2.

To the best of our knowledge, *St*CYP154C4 is the first reported wild-type CYP enzyme that hydroxylates testosterone to yield isomers of hydroxytestosterone. A similar pattern of product formation was observed with CYP106A2, with dehydroepiandrosterone hydroxylated at the 7α and 7β positions to form the stereoisomer [35]. In a similar manner, CYP106A2 was found to hydroxylate abietic acid, a diterpene, at the 12α and 12β positions [36]. Not many wild-type P450s are known to hydroxylate the C2 positions, although the best-characterized P450BM3 (CYP102A1) has been engineered to obtain 2-hydroxylated testosterone [37–39]. *St*CYP154C4s may represent important biocatalysts in the future, with the aim of developing steroids with different regiospecificities by protein engineering. Recently, a steroidogenic P450 CYP260A1 from *Sorangium cellulosum* strain So ce56 was engineered for stereo- and regioselective hydroxylation of progesterone at the very important 1α and 17α positions [40]. In another study involving the protein engineering of P450BM3 variant M01 A82W S72I, which is capable of producing 16α-progesterone, researchers inverted the stereoselectivity to 16β using a single active site mutation [41]. Other P450s of bacterial origin have been demonstrated to hydroxylate testosterone at different positions [13,22,25]. However, there have been no reports of any wild-type CYP that hydroxylates at both positions 2α and 2β.

*St*CYP154C4s represent a suitable model for understanding the function of CYP154C subclass P450 enzymes. Our results show that not only does the choice of redox partners affect the type and selectivity of P450 reactions, but also the varied concentrations of P450s and auxiliary proteins (redox partners) can alter the reaction type catalyzed by P450 enzymes. Such results provide a unique opportunity to further exploit the chemistries behind altered activity and product distribution. The lack of suitable redox partners for support for *in vivo* conversion of the substrates catalyzed by both *St*CYP154C4s was a major problem for identifying the products by NMR. The *in vivo* bioconversion carried out in *E. coli* cells, which contained a plasmid harboring the genes encoding *St*CYP154C4s, putidaredoxin reductase (Pdr) and putidaredoxin (Pdx), did not result in the conversion of the three substrates. However, traces of products were detected when the reaction was supported by Pdx and Pdr *in vitro*. Considering the interesting functions of both P450s, further research is needed to find suitable redox partners to support *in vivo* conversion, leading to the elucidation of the structures of the modified products. Mutagenesis of both enzymes should also be carried out to identify the specific residues involved in these functions.

Overall structure

The crystal structures of CYP154C4 from *Streptomyces* sp. W2061 (*St*CYP154C4-1) and testosterone-bound CYP154C4 from *Streptomyces* sp. ATCC 11861 (*St*CYP154C4-2) were determined at 2.2 and 2.7 Å, respectively (Fig. 6A,B). Those two CYP154C4s share 93% sequence identity, with 97% positives and only 29 different residues. Both crystals belonged to space group $I2_12_12_1$ and contained a single molecule in each asymmetric unit. *St*CYP154C4-1 was solved by molecular replacement using CYP154C1 from *S. coelicolor* (PDB code 1GWI; 73% sequence identity) as a search model [12]. The *St*CYP154C4-1 structure was subsequently used for molecular replacement to solve the *St*CYP154C4-2 structure. Both crystal structures showed typical characteristics of a CYP, with the heme-binding site located at the center of an α-helix-rich region. *St*CYP154C4-1 comprises 21 α-helices and six β-strands (Fig. 6A), whereas testosterone-bound *St*CYP154C4-2 comprises 19 α-helices and six β-strands, with two small α-helices missing as compared with the *St*CYP154C4-1 structure (Fig. 6B). Structural alignments using the Dali server (http://ekhidna2.biocenter.helsinki.fi/dali/) confirmed that *St*CYP154C4 has the highest structural similarity with CYP154C5 from *N. farcinca* (PDB code 4J6B), with a Z-score of 58.7 [24]. Additionally, CYP154C1 from *S. coelicolor* (PDB code 1GWI) and CYP154A1 from *S. coelicolor* (PDB code 1ODO) also showed high levels of structural similarity [12,23].

The active sites of the *St*CYP154C4s

The cofactor heme is tightly bound in the center of *St*CYP154C4 enzymes. Carboxyl groups of heme form salt bridges with the conserved residues Lys57, His94,

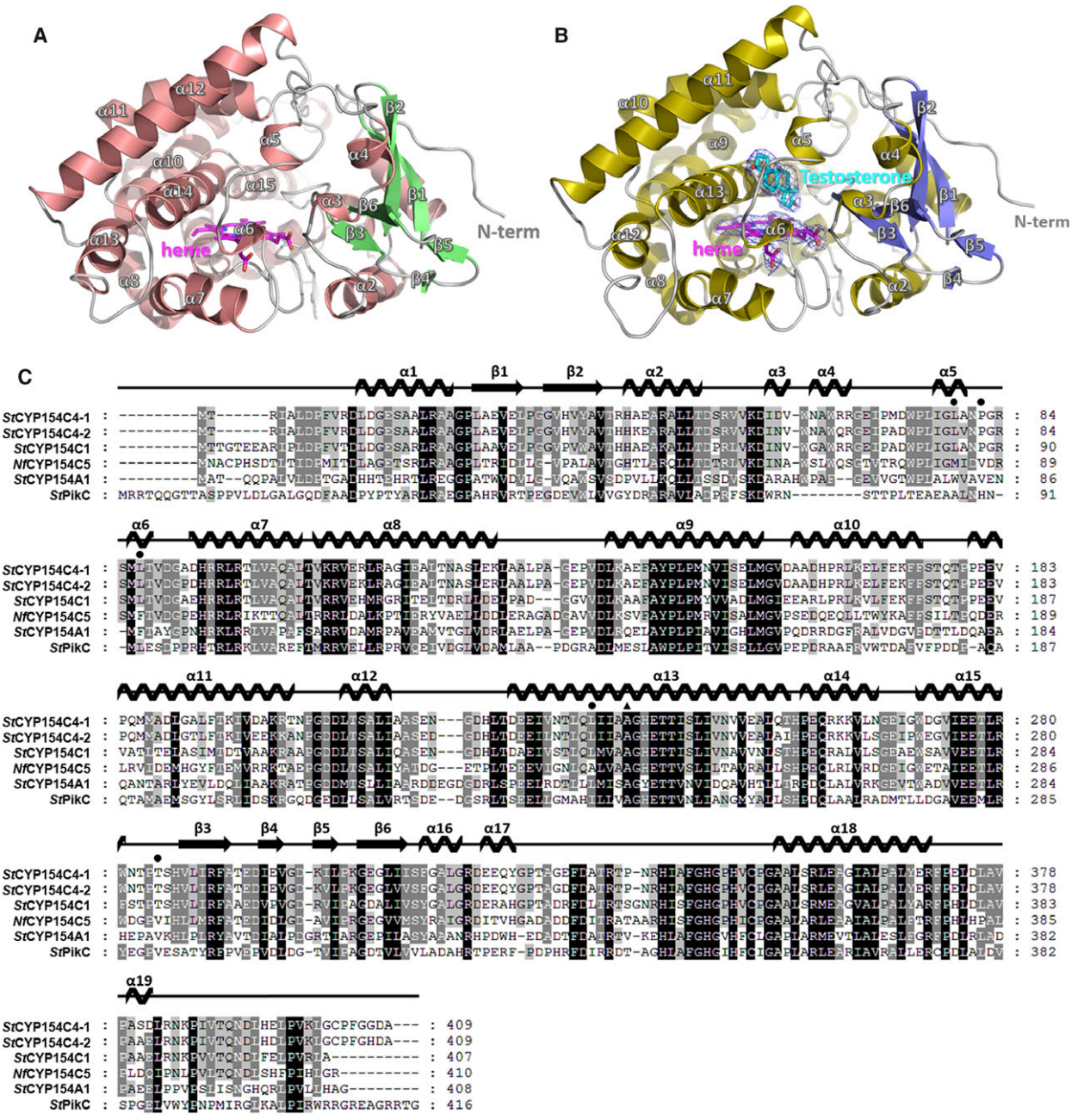

Fig. 6. Crystal structures and multiple sequence alignment of *St*CYP154C4-1 and testosterone-bound *St*CYP154C4-2. (A) The crystal structure of *St*CYP154C4-1. The α-helices and β-strands are represented as ribbons colored in salmon and lime, respectively. The bound heme molecule is shown as a magenta stick model. (B) The overall structure of testosterone-bound *St*CYP154C4-2 is represented as a ribbon, with the α-helices and β-strands colored in olive and slate, respectively. The bound testosterone is shown as a cyan stick model with a blue 2Fo–Fc electron-density map (contoured at 1δ). (C) Structure-based multiple sequence alignment of *St*CYP154C4s with CYP154C1 from *Streptomyces coelicolor* A3(2; PDB code 1GWI; UniProtKB code: Q9L142), CYP154C5 from *Nocardia farcinica* (PDB code 4J6D; UniProtKB code: Q5YNS8), CYP154A1 from *Streptomyces coelicolor* (PDB code 1ODO; UniProtKB code: Q9KZR7), and P450 monooxygenase PikC from *Streptomyces venezuelae* (PDB code 2WHW; UniProtKB code: O87605). The secondary structures of testosterone-bound *St*CYP154C4-2 are shown above the multiple aligned sequences. The Ala238, which altered its position following testosterone binding, is indicated with a black triangle. Residues making specific interactions with testosterone are denoted by a black circle above the residues.

Arg98, Arg291, and His348. The side chain of Cys350 coordinates the central iron of heme on the proximal side. The non-polar region of the heme molecule is surrounded by hydrophobic residues (Fig. 7). Above the bound heme, there is a hydrophobic substrate-binding site, with Pro82, Leu87, Phe173, Phe174, Leu234, Ala237, Ala238, and Val288 residues comprising the non-polar region of the substrate-binding pocket. Additionally, several polar residues, including Gln233, Thr242, Thr285, and Gln391, are also located in the substrate-binding pocket (Fig. 8). They may be responsible for hydrogen bonding with the substrate. All of these residues are shared between *St*CYP154C4-1 and *St*CYP154C4-2 suggesting that both enzymes likely share similar substrate specificity, as well as regiospecificity (Fig. 8A,B). As expected, this prediction was confirmed by the results of biochemical assays. Structural comparison of *St*CYP154C4-1 and *St*CYP154C4-2 showed a conformational change in the α13 helix following testosterone binding (Fig. 8C). The Ala238 located on the α13-helix moved to near the testosterone substrate and formed a hydrophobic interaction. This movement induces rearrangement of the α13 helix. The orientations of the other residues forming the substrate-binding pocket are also slightly altered by testosterone binding. Specifically, interactions between *St*CYP154C4-2 and testosterone involve the O3 and O17 atoms of testosterone interacting with the OG1 of Thr285 and the OE1 of Gln233, respectively. It is possible that these two hydrogen bonds are important in determining the orientation of testosterone binding. Moreover, the C18 and C19 atoms of testosterone are oriented on the opposite side of the α13 helix and form non-polar interactions with Leu79, Pro82, and Leu87. The C2 atom of testosterone faces the Fe of the heme molecule and lies within 4.5 Å in a position that enables 2′-hydroxylation. Activity assays confirmed that *St*CYP154C4 is capable of catalyzing both 2α-hydroxylation and 2β-hydroxylation of testosterone.

The testosterone-binding mode of the *St*CYP154C4s

Structural superposition of testosterone-bound *St*CYP154C4-2 and testosterone-bound *Nf*CYP154C5 (PDB code 4J6D; 52% sequence identity) revealed that the two enzymes have different testosterone-binding modes, as well as different residue compositions in the substrate-binding pocket (Fig. 9) [24]. The hydrophobic residues Leu79, Pro82, Leu87, and Leu234, which comprise the non-polar region of the binding pocket in *St*CYP154C4-2, are represented by Met84, Val88, Phe92, and Ala240, respectively, in *Nf*CYP154C5. These differences form a unique cavity shape in *St*CYP154C4 that might influence ligand binding and regiospecificity. Specifically, Thr285 in *St*CYP154C4 interacts with the O3 atom of testosterone, but is represented by Val291 at the same positon in *Nf*CYP154C5. By contrast, testosterone binding in the *Nf*CYP154C5 structure shows that Gln398 interacts with the O3 atom of testosterone. In the *St*CYP154C4-2 structure, the O3 atom of testosterone forms a hydrogen bond with Thr285 located at a point deeper within the binding pocket than *Nf*CYP154C5 Gln398. These differences in the substrate-binding site and individual side chain interactions likely enable testosterone to bind in the opposite direction in *St*CYP154C4. Moreover, structural information verified the results obtained from activity assays, which showed that *St*CYP154C4 exhibits both 2′-hydroxylation and 16α-hydroxylation activity on testosterone, whereas *Nf*CYP154C5 exhibits only 16α-hydroxylation activity.

In conclusion, we identified broad and flexible substrate-hydroxylation activity on the part of both

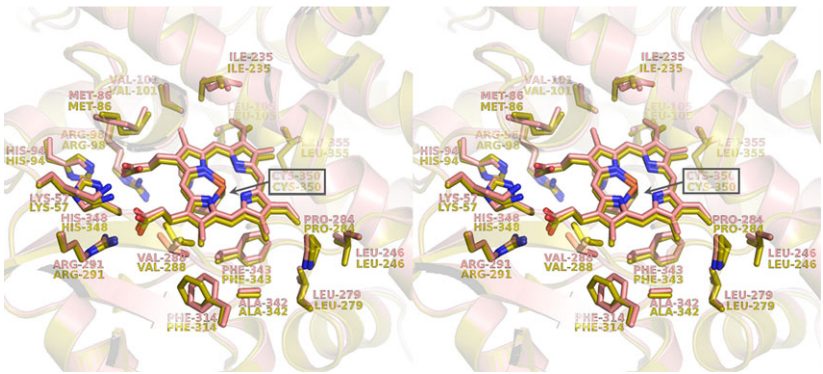

Fig. 7. Heme binding modes of *St*CYP154C4-1 and *St*CYP154C4-2. Stereo view of superposed *St*CYP154C4-1 (salmon) and *St*CYP154C4-2 (olive) shows that the iron of the heme is coordinated with the side chain of Cys350. Carboxyl groups of the heme interact with Lys57, His94, Arg98, and Arg291, while the non-polar region of the heme forms hydrophobic interactions with Met86, Val101, Leu105, Ile235, Leu246, Leu279, Pro284, Val288, Phe314, Ala342, and Phe343.

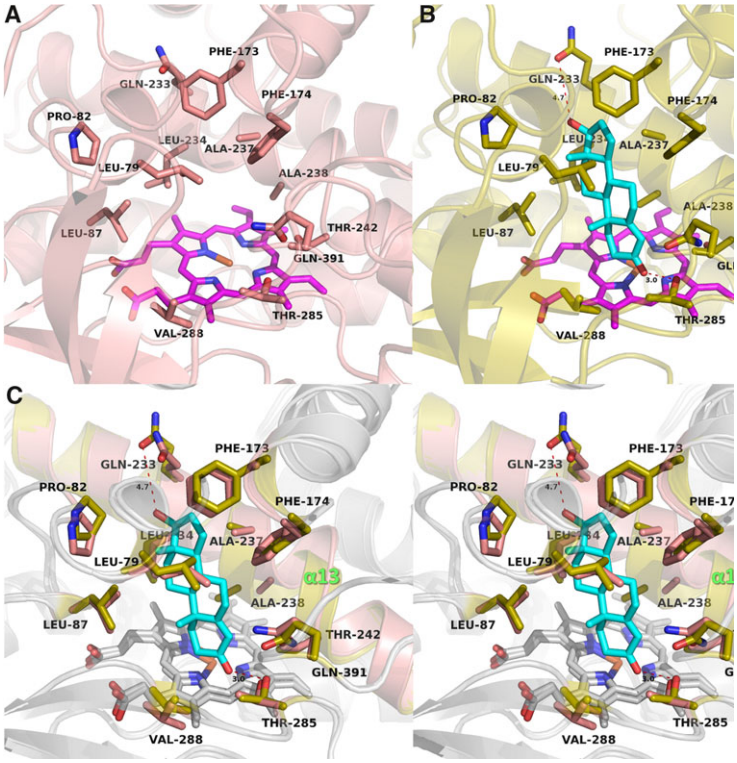

Fig. 8. Substrate-binding site of the *St*CYP154C4s. The substrate-binding site of (A) *St*CYP154C4-1 and (B) testosterone-bound *St*CYP154C4-2 are colored in salmon and olive, respectively. The residues in the substrate-binding site, and the bound heme molecule (magenta) and testosterone (cyan) are shown as stick models. Hydrogen bonds formed between the bound testosterone and *St*CYP154C4-2 are shown with red dashed lines. (C) Stereo view of *St*CYP154C4-1 superposed with testosterone-bound *St*CYP154C4-2. The α13 helix undergoes conformational changes during testosterone binding.

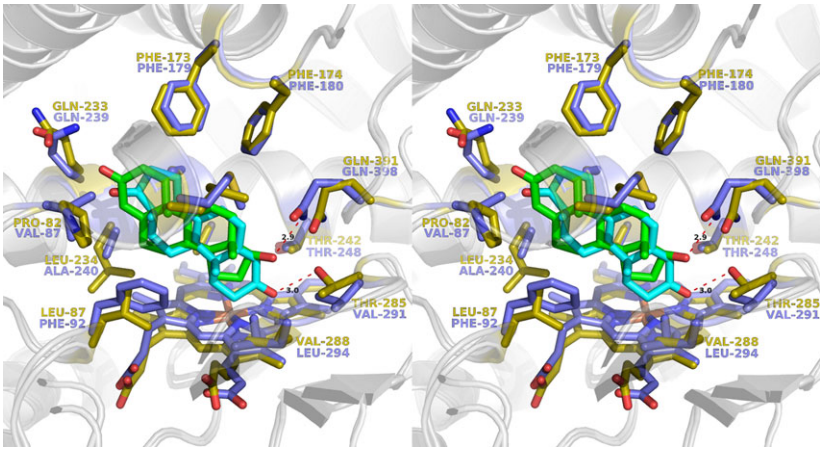

Fig. 9. Stereo view of the superimposed structures of testosterone (cyan)-bound *St*CYP154C4-2 (olive) and testosterone (green)-bound *Nf*CYP154C5 (PDB code 4J6D; slate). The residues in the vicinity of the bound testosterone and testosterone molecules are indicated by sticks. Structural comparisons of the two enzymes show different testosterone-binding modes.

*St*CYP154C4-1 and *St*CYP154C4-2 enzymes as compared with the previously characterized *Nf*CYP154C5 enzyme. Our findings suggest that differences in residue positioning in the active site might account for the altered substrate position and catalytic activity observed in the CYP154C4 enzyme. Collectively, these results provide useful insight for subsequent protein engineering of CYP154C4 family members in order to allow production of modified steroid compounds.

Materials and methods

Materials

Testosterone was purchased from Tokyo Chemical Industry Co., Ltd (Seoul, Korea). NADPH, glucose-6-phosphate, glucose-6-phosphate dehydrogenase, spinach ferredoxin, spinach ferredoxin reductase, hydrogen peroxide, PIDA, catalase, ampicillin, chloramphenicol, and α-aminolevulinic acid were purchased from Sigma-Aldrich (Yongin, Korea).

DNA polymerase, T4 DNA ligase, and dNTPs were obtained from Takara Bio (Shiga, Japan). IPTG and DTT were obtained from Duchefa Bohemie (Seoul, Korea). All other chemicals were high-grade products obtained from commercially available sources. Restriction enzymes (*Eco*RI and *Hin*dIII) were obtained from Takara Clontech (Seoul, Korea).

Sequence accession numbers

The gene responsible for encoding the CYP enzyme was searched based on the heme-binding domain signature [FxxGx(H/R)xCxG] from the sequences of *Streptomyces* sp. W2061 and *Streptomyces* sp. ATCC11861. The name of the CYP enzyme (CYP154C4) was assigned by Dr David Nelson (http://drnelson.uthsc.edu/cytochromeP450.html). The nucleotide sequences of *St*CYP154C4-1 and *St*CYP154C4-2 have been deposited in GenBank under accession numbers MH444361 and MH444362, respectively.

Cloning, expression, and purification of the *St*CYP154C4s

The *St*CYP154C4-1 and *St*CYP154C4-2 genes were amplified from genomic DNA of *Streptomyces* sp. W2061 and *Streptomyces* sp. ATCC 11861, respectively. The *St*CYP154C-1 sequence was amplified using 5′-GAA TTC ATG ACC CGT ATC GCG CTC-3′ (*Eco*RI site is underlined) as a forward primer and 5′-AAG CTT TCA GGC GTC GCC GCC G-3′ (*Hin*dIII site is underlined) as a reverse primer. Similarly, amplification of *St*CYP154C4-2 was performed using specific forward 5′-GAA TTC ATG ACC CGT ATC GCG CTC G-3′ (*Eco*RI site is underlined) and reverse 5′-AAG CTT TCA GGC GTC GTG GCC GAA-3′ (*Hin*dIII site is underlined) primers. The target genes were cloned into the pET32a(+) vector and introduced into *E. coli* BL21(DE3) cells. Transformed cells were grown overnight at 37 °C for seed culture and then inoculated into LB medium supplemented with 100 $\mu g \cdot mL^{-1}$ ampicillin, followed by incubation (180 r.p.m.) at 37 °C until cell density ranged from 0.6 to 0.8 according to the attenuance at 600 nm. Cell cultures were induced with 0.3 mM IPTG and supplemented with 100 $\mu g \cdot mL^{-1}$ 5-aminolevulinic acid hydrochloride and 0.5 mM $FeCl_3$. Protein expression was performed at 20 °C for 48 h, followed harvesting of cells by centrifugation. The cell pellets were further washed and resuspended in 50 mM Tris/HCl buffer (pH 7.4) containing 10% glycerol, 100 mM NaCl, and 1 mM DTT. For purification, the cells were lysed by ultrasonication, followed by centrifugation at 24 650 *g* for 30 min at 4 °C. The soluble fraction of the cell lysate was mixed and purified by Ni^{2+}-affinity chromatography using a TALON His-tag. Protein-bound resins were pre-equilibrated with two-column volumes of equilibration buffer (potassium phosphate, pH 7.4), and bound proteins were eluted with elution buffer [potassium phosphate (pH 7.4), 10% glycerol, and 100 mM NaCl] containing 20, 100, and 300 mM imidazole, respectively. Fractions of eluted proteins were analyzed for purity by SDS/PAGE, and pure fractions were concentrated by ultrafiltration using Amicon centrifugal filters (Millipore, Billerica, MA, USA) with a molecular mass cut-off of 30 kDa.

UV-visible absorption spectroscopy

UV-visible spectra of *St*CYP154C4s were measured at room temperature on a Biochrome Libra S35PC UV/visible spectrophotometer (Biochrome Libra, Cambridge, UK). Absorption spectroscopy was used to determine CYP concentration, purity, and substrate spin shift. Concentration of both CYPs was determined based on CO-difference spectra using the extinction coefficient $\varepsilon_{450-490}$ = 91 000 mM^{-1} cm^{-1} [4]. The proteins were diluted to 2 mL with potassium phosphate buffer (pH 7.4) and separated into two cuvettes (sample and reference). The sample was bubbled with CO for 1 min, followed by addition of sodium dithionite in both cuvettes. The cuvettes were further scanned using the spectrophotometer at between 400 and 500 nm repeatedly until the reading at ~ 450 nm deceased relative to the highest value of the readings. To determine the purity (RZ value) of both CYPs, the sample containing the CYP enzyme was scanned in the range of 200–500 nm, whereas the reference contained only buffer. Substrate-binding shift was observed in a sample containing 1 μM CYP and 20 μM substrate in potassium phosphate buffer (pH 7.4), with the reference containing only and equal concentration of substrate. The sample and reference were scanned at between 350 and 500 nm. Dissociation constant (K_d) values for *St*CYP154C4s (1 μM) were determined by titrating the substrate (testosterone) until the CYPs were saturated. K_d values were determined as described previously [42].

Enzyme activity assay

In vitro activity of both *St*CYP154C4s was determined in the presence of NADPH, PIDA, and H_2O_2, separately. The substrate (testosterone) and PIDA were dissolved in DMSO and 100% ethanol, making the stock concentrations 20 and 50 mM, respectively. Reaction system I contained 3 μM *St*CYP154C4s, spinach Fdx (6 μM), spinach Fdr (0.025 U), glucose-6-phosphate dehydrogenase (1 U), glucose-6-phosphate (10 mM), $MgCl_2$ (5 mM), catalase (100 $\mu g \cdot mL^{-1}$), and substrate (500 μM) in 500 μL potassium phosphate buffer (pH 7.4). The reactions were initiated with 500 μM NADPH. The reaction mixtures were incubated at 30 °C for 2 h, followed by extraction with ethyl acetate (1000 μL), which was dried and analyzed by HPLC and LC-MS. Reaction system II contained *St*CYP154C4s (15 μM), Fdx (15 μM), Fdr (0.1 U), glucose-6-phosphate (100 mM), glucose-6-phosphate dehydrogenase (5 U), $MgCl_2$ (10 mM), catalase (100 $\mu g \cdot mL^{-1}$), substrate (500 μM), and 1 mM NADPH. The reaction mixtures

were incubated at 30 °C for different time intervals ranging from 2 to 6 h. The reactions mixtures were extracted and analyzed as described for reaction system I. The other two reactions, supported with PIDA (2 μM) and H_2O_2 (40 mM), contained 3 μM *St*CYP154C4, with the two *St*CYP154C4s being analyzed separately. All reactions mixtures were incubated at 30 °C for 2 h and extracted and analyzed as described for reaction system I.

Product purification and characterization

The product (P2) submitted for NMR analysis was prepared by an up-scaled (300 mL) *in vitro* reaction catalyzed by *St*CYP154C-1. The reactions were performed separately in a 10-mL volume in potassium phosphate buffer (pH 7.4) containing 250 μM substrate. The reactions were initiated with the addition of 2 mM PIDA and incubated at 30 °C for 2 h. Each reaction was extracted with 10 mL ethyl acetate, dried *in vacuo*, and dissolved in HPLC-grade methanol. The sample was filtered through a 0.45-μm pore polytetrafluoroethylene filter and subjected to preparative HPLC (Shimadzu, Kyoto, Japan) with a C_{18} column (Mightysil RP-18 GP; 150 × 4.6 mm, 5 μm; Kanto Chemical, Tokyo, Japan). The purity of the product (P2) was reconfirmed by HPLC. Fractions containing the product (P2) were dried and dissolved in DMSO-d_6, and the sample was further subjected to NMR analyses (^{1}H-NMR, ^{13}C-NMR, ROSEY, COSY, HMBC, and HSQC) at 800 MHz using a Varian Unity INOVA spectrometer (Varian, Palo Alto, CA, USA).

2α-Hydroxytestosterone

^{1}H NMR (800 MHz, DMSO-*d*6): δ 5.63 (d, *J* = 1.6 Hz, 1H), 5.02 (d, *J* = 4.4 Hz, 1H), 4.48 (d, *J* = 4.8 Hz, 1H), 4.14 (dt, *J* = 13.7, 4.9 Hz, 2H), 3.44 (dt, *J* = 8.5, 5.0 Hz, 1H), 2.35 (tdd, *J* = 14.0, 5.3, 1.8 Hz, 1H), 2.26 (ddd, *J* = 14.4, 4.3, 2.4 Hz, 1H), 2.12 (dd, *J* = 12.6, 5.5 Hz, 1H), 1.87–1.80 (m, 1H), 1.79–1.73 (m, 2H), 1.52 (qd, *J* = 11.0, 3.6 Hz, 1H), 1.53–1.45 (m, 1H), 1.52–1.44 (m, 1H), 1.44 (t, *J* = 13.3 Hz, 1H), 1.36 (ddd, *J* = 25.9,13.1, 3.9 Hz, 1H), 1.39–1.31 (m, 1H), 1.23 (s, 3H), 1.20 (dp, *J* = 18.1, 5.8 Hz, 1H), 0.98 (td, *J* = 12.9, 4.2 Hz, 1H), 0.95–0.87 (m, 1H), 0.89 (s, 1H), 0.86 (td, *J* = 11.5, 11.1, 4.2 Hz, 1H), 0.68 (s, 3H). ^{13}C NMR (201 MHz, DMSO-*d*6): δ 199.98 (C3), 171.15 (C5), 121.45 (C4), 80.29 (C17), 69.25 (C2), 54.45 (C9), 50.37 (C14), 44.96 (C1), 42.84 (C13), 40.61 (C10), 36.67 (C12), 35.04 (C8), 31.98 (C6), 31.75 (C7), 30.27 (C16), 23.43 (C15), 20.59 (C11), 18.06 (C19), 11.67 (C18).

2β-Hydroxytestosterone

^{1}H NMR (800 MHz, DMSO-*d*6): δ 5.65 (d, *J* = 1.3 Hz, 1H), 5.15 (d, *J* = 4.2 Hz, 1H), 4.48 (d, *J* = 4.8 Hz, 1H), 3.99 (dt, *J* = 10.9, 4.5 Hz, 1H), 3.43 (dt, *J* = 8.6, 4.9, 4.8 Hz, 1H), 2.45 (tdd, *J* = 13.0, 5.0, 1.0 Hz, 1H), 2.20 (ddd, *J* = 12.6, 4.3, 2.5 Hz, 1H), 2.07 (dd, *J* = 13.8,4.8 Hz, 1H), 1.88–1.83 (m, 1H), 1.87–1.80 (m, 1H), 1.79–1.73 (m, 1H), 1.63–1.61 (m,1H), 1.61 (t, *J* = 5.2 Hz, 1H), 1.63–1.59 (m, 1H), 1.53–1.45 (m, 1H), 1.41 (dd, *J* = 13.2,3.8 Hz, 1H), 1.39–1.31 (m, 1H), 1.26–1.21 (m, 1H), 1.20 (dp, *J* = 18.1, 5.8 Hz, 1H),1.15 (s, 3H), 1.04 (td, *J* = 12.9, 4.2 Hz, 1H), 0.95–0.87 (m, 1H), 0.95–0.84 (m, 1H), 0.68 (s, 3H). ^{13}C NMR (201 MHz, DMSO-d6): δ 199.41 (C3), 172.90 (C5), 120.29 (C4), 80.26 (C17), 68.61 (C2), 51.21 (C9), 50.39 (C14), 43.30 (C13), 41.28 (C1), 40.69 (C10), 36.70 (C12), 35.62 (C8), 34.11 (C7), 32.50 (C6), 30.33 (C16), 23.48 (C15), 22.44 (C19), 22.14 (C11), 11.81 (C18).

Analytical methods

The reaction mixtures were analyzed by ultra-HPLC (UHPLC) using an SYNAPT G2-S/ACUITY UPLC quadrupole time-of-flight/electrospray ionization mass spectrometry (Waters, Milford, MA, USA) in positive ion mode. The sample was injected into the UHPLC system and separated using a Mightysil reversed phase C_{18} column (4.6 × 250 mm, 5 μm; Kanto Chemical). Acetonitrile (B) and water (A) were used as mobile phases in a gradient system of B (15% for 0–10 min, 50% for 10–20 min, 70% for 20–25 min, and 15% for 25–40 min) at a flow rate of 1 mL·min^{-1}. Detection of substrates and their product was performed by UV absorbance at 242 nm.

Crystallization and data collection

Initial crystallization screenings of both *St*CYP154C4-1 and testosterone-bound *St*CYP154C4-2 were performed using a crystallization robot (Mosquito; TTP Labtech, Cambridge, MA, USA) with the sitting-drop vapor-diffusion method at 293 K in 96-well crystallization plates (Emerald Bio, Bainbridge Island, WA, USA). Commercially available screening kits were used for screening, including the MCSG suite I–IV (Microlytic, Burlington, MA, USA) and SaltRx (Hampton Research, Aliso Viejo, CA, USA). Co-crystallization of testosterone-bound *St*CYP154C4-2 was prepared by adding 10 mM testosterone. The drops were mixed with 0.3 μL of protein solution and 0.3 μL of reservoir solution, and equilibrated against 80 μL reservoir solution at 293 K. *St*CYP154C4-1 crystals were grown under conditions of 25% (w/v) PEG3350, 0.2 M lithium sulfate, and 0.1 M Bis-Tris-HCl (pH 6.5; MCSG2 no. B11). Crystals of testosterone-bound *St*CYP154C4-2 were obtained from conditions of 1 M sodium citrate tribasic and 0.1 M sodium cacodylate:HCl (pH 6.5; MCSG3 no. A1). Before mounting the crystals, those of testosterone-bound *St*CYP154C4-2 were briefly soaked in paratone-N oil used as a cryoprotectant, whereas *St*CYP154C4-1 crystals were directly mounted. X-ray diffraction data for both crystals were collected at the BL-5C beam line of the Pohang Accelerator Laboratory

Table 2. X-ray diffraction data collection and refinement statistics.

Data set	*St*CYP154C4-1	Testosterone-bound *St*CYP154C4-2
X-ray source	PAL 5C beam line	PAL 5C beam line
Space group	$I2_12_12_1$	$I2_12_12_1$
Unit-cell parameters (Å, °)	a = 83.755, b = 115.032, c = 127.807 $\alpha = \beta = \gamma = 90$	a = 90.381, b = 113.94, c = 128.361 $\alpha = \beta = \gamma = 90$
Wavelength (Å)	0.9795	0.9795
Resolution (Å)	50.00–2.20 (2.24–2.20)	50.00–2.70 (2.75–2.70)
Total reflections	204 234	95 298
Unique reflections	31 349 (1506)	17 614 (872)
Average I/σ (I)	42.9 (7.9)	31.9 (5.8)
R_{merge}[a]	0.087 (0.445)	0.080 (0.419)
Redundancy	6.5 (6.3)	5.4 (5.0)
Completeness (%)	98.0 (94.7)	96.5 (96.8)
Refinement		
Resolution range (Å)	46.47–2.19 (2.25–2.19)	47.55–2.71 (2.78–2.71)
No. of reflections of working set	29 794 (2081)	16 737 (1059)
No. of reflections of test set	1554 (108)	872 (47)
No. of amino acid residues	409	408
No. of water molecules	268	8
R_{cryst}[b]	0.173 (0.200)	0.217 (0.561)
R_{free}[c]	0.222 (0.296)	0.266 (0.594)
r.m.s. bond length (Å)	0.019	0.013
r.m.s. bond angle (°)	2.002	1.766
Average B value (Å^2; protein)	43.956	87.640
Average B value (Å^2; solvent)	47.231	70.805

[a] $R_{merge} = \sum|<I>-I|/\sum<I>$.
[b] $R_{cryst} = \sum||Fo|-|Fc||/\sum|Fo|$.
[c] R_{free} was calculated using 5% of all reflections, which were excluded from the refinement stages using high-resolution data. Values in parentheses refer to the highest resolution shells.

(PAL; Pohang, Korea) at 100 K. The 2.2-Å resolution of the diffracted data for *St*CYP154C4-1 contained 180 images, and the 2.7-Å resolution of the diffracted data for testosterone-bound *St*CYP154C4-2 contained 150 images. Image data were collected at a 1° oscillation and an exposure time of 1 s per image. All diffraction data were indexed, processed, and scaled with HKL-2000 [43]. The statistics associated with the X-ray diffraction data are presented in Table 2.

Structure determination and refinement

Molecular replacement of *St*CYP154C4-1 was successfully performed using the MOLREP program from the CCP4i package [44,45]. The crystal structure of CYP154C1 from *S. coelicolor* (PDB code 1GWI; 73% sequence identity) was used as the template model [12]. One molecule was predicted in an asymmetric unit with a Matthews coefficient of 3.45 Å^3 Da^{-1}, corresponding to 64.50% solvent [46]. The resulting structure was manually rebuilt and refined using COOT and REFMAC5 [47–49]. The final model of *St*CYP154C4-1 showed R_{work} and R_{free} values of 17.24% and 22.17%, respectively. Testosterone-bound *St*CYP154C4-2 was solved by molecular replacement using *St*CYP154C4-1 as a template model. Rebuilding and refinement of the structure were performed using same methods as those used for *St*CYP154C4-1. The final model of testosterone-bound *St*CYP154C4 had R_{work} and R_{free} values of 20.31% and 25.10%, respectively. The qualities of both final models were checked using MOLPROBITY and PROCHECK [50,51]. Structure coordinates and structure-factor data for *St*CYP154C4-1 and testosterone-bound *St*CYP154C4-2 have been deposited in the Protein Data Bank with accession codes 6A7I and 6A7J, respectively.

Acknowledgements

We thank the staff at the X-ray core facility of the Korea Basic Science Institute (Ochang, Korea) and BL-5C of the Pohang Accelerator Laboratory (Pohang, Korea) for their assistance with data collection. This work was supported by the Korea Polar Research Institute (grant nos. PE18210 and PE19210). Additionally, we would like to thank the Division of Magnetic Resonance, Korea Basic Science Institute, Ochang, Chungbuk, Korea, for their help with NMR analysis.

Author contributions

HP, JHL, and T-JO designed and supervised the project. BD, K-HK, and E-JY performed the cloning,

expression, and protein purification experiments. CWL and S-HP performed crystallization, collected X-ray data, and solved the protein structures. BD, K-HK, C-SJ, and E-JY performed the enzyme activity experiment. BD and K-HK performed the HPLC and NMR analyses. BD, CWL, JHL, and T-JO wrote the initial manuscript, and HP contributed to revisions. All authors discussed the results, commented on the manuscript, and approved the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

References

1 Nebert DW & Gonzalez FJ (1987) P450 genes: structure, evolution, and regulation. *Annu Rev Biochem* **56**, 945–993.

2 Guengerich FP (2002) Cytochrome P450 enzymes in the generation of commercial products. *Nat Rev Drug Discov* **1**, 359–366.

3 Ortiz de Montellano PR (2005) Cytochrome P450: Structure, Mechanism, and Biochemistry, 3rd edn. Kluwer Academic/Plenum Publishers, New York.

4 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.

5 Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D *et al.* (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147.

6 Lamb DC, Guengerich FP, Kelly SL & Waterman MR (2006) Exploiting *Streptomyces coelicolor* A3(2) P450s as a model for application in drug discovery. *Expert Opin Drug Metab Toxicol* **2**, 27–40.

7 Lamb DC, Waterman MR, Kelly SL & Guengerich FP (2007) Cytochromes P450 and drug discovery. *Curr Opin Biotechnol* **18**, 504–512.

8 Purnapatre K, Khattar SK & Saini KS (2008) Cytochrome P450s in the development of target-based anticancer drugs. *Cancer Lett* **259**, 1–15.

9 Di Nardo G & Gilardi G (2012) Optimization of the bacterial cytochrome P450 BM3 system for the production of human drug metabolites. *Int J Mol Sci* **13**, 15901–15924.

10 Waskell L & Kim JJ (2015) Electron transfer partners of cytochrome P450. In *Cytochrome P450: Structure, Mechanism, and Biochemistry* (Ortiz de Montellano PR, ed), pp. 33–68. Springer International Publishing, Cham, Switzerland.

11 Hannemann F, Bichet A, Even KM & Bernhardt R (2007) Cytochrome P450 systems—biological variations of electron transport chains. *Biochim Biophys Acta* **1770**, 330–334.

12 Podust LM, Kim Y, Arase M, Neely BA, Beck BJ, Bach H, Sherman DH, Lamb DC, Kelly SL & Waterman MR (2003) The 1.92-Å structure of *Streptomyces coelicolor* A3(2) CYP154C1. A new monooxygenase that functionalizes macrolide ring systems. *J Biol Chem* **278**, 12214–12221.

13 Makino T, Katsuyama Y, Otomatsu T, Misawa N & Ohnishi Y (2014) Regio- and stereospecific hydroxylation of various steroids at the 16α position of the D ring by the *Streptomyces griseus* cytochrome P450 CYP154C3. *Appl Environ Microbiol* **80**, 1371–1379.

14 Choi KY, Park HY & Kim BG (2010) Characterization of bi-functional CYP154 from *Nocardia farcinica* IFM10152 in the O-dealkylation and *ortho*-hydroxylation of formononetin. *Enzyme Microb Technol* **47**, 327–334.

15 Baker ME (2011) Origin and diversification of steroids: co-evolution of enzymes and nuclear receptors. *Mol Cell Endocrinol* **334**, 14–20.

16 Donova MV & Egorova OV (2012) Microbial steroid transformations: current state and prospects. *Appl Microbiol Biotechnol* **94**, 1423–1447.

17 Peterson DH, Murray HC, Eppstein SH, Reineke LM, Weintraub A, Meister PD & Leigh HM (1952) Microbiological transfromations of steroids. 1. Introduction of oxygen at carbon-11 of progesterone. *J Am Chem Soc* **74**, 5933–5936.

18 Petzoldt K, Annen K, Laurent H & Wiechert R (1982) Process for the Preparation of 11-Beta-Hydroxy Steroids. Patent 4,353,985, Schering Aktiengesellschaft, Berlin, Germany.

19 Galán B, Garcia-Fernández J, Felpeto-Santero C, Fernández-Cabezón L & Garcia JL (2017) Bacterial metabolism of steroids. In *Aerobic Utilization of Hydrocarbons, Oils and Lipids. Handbook of Hydrocarbon and Lipid Microbiology* (Rojo F, ed), pp. 1–22. Springer, Cham, Switzerland.

20 Szaleniec M, Wojtkiewicz AM, Bernhard R, Borowski T & Donova M (2018) Bacterial steroid hydroxylases: enzyme classes, their functions and comparison of their catalytic mechanisms. *Appl Microbiol Biotechnol* **102**, 8153–8171.

21 Ouellet H, Johnston JB & Ortiz de Montellano PR (2010) The *Mycobacterium tuberculosis* cytochrome P450 system. *Arch Biochem Biophys* **493**, 82–95.

22 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.

23 Podust LM, Bach H, Kim Y, Lamb DC, Arase M, Sherman DH, Kelly SL & Waterman MR (2004)

Comparison of the 1.85 Å structure of CYP154A1 from *Streptomyces coelicolor* A3(2) with the closely related CYP154C1 and CYPs from antibiotic biosynthetic pathways. *Protein Sci* **13**, 255–268.

24 Herzog K, Bracco P, Onoda A, Hayashi T, Hoffmann K & Schallmey 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.

25 Bracco P, Janssen DB & Schallmey A (2013) Selective steroid oxyfunctionalisation by CYP154C5, a bacterial cytochrome P450. *Microb Cell Fact* **12**, 95.

26 Su Z, Horner JH & Newcomb M (2012) Cytochrome P450 119 compounds I formed by chemical oxidation and photooxidation are the same species. *Chemistry*. https://doi.org/10.1002/chem.201202254.

27 Luthra A, Denisov IG & Sligar SG (2011) Spectroscopic features of cytochrome P450 reaction intermediates. *Arch Biochem Biophys* **507**, 26–35.

28 Hrycay EG & Bandiera SM (2012) The monooxygenase, peroxidase, and peroxygenase properties of cytochrome P450. *Arch Biochem Biophys* **522**, 71–89.

29 Gelb MH, Heimbrook DC, Mälkönen P & Sligar SG (1982) Stereochemistry and deuterium isotope effects in camphor hydroxylation by the cytochrome P450cam monoxygenase system. *Biochemistry* **21**, 370–377.

30 Yoshimoto FK, Gonzalez E, Auchus RJ & Guengerich FP (2016) Mechanism of 17α,20-lyase and new hydroxylation reactions of human cytochrome P450 17A1: ^{18}O labeling and oxygen surrogate evidence for a role of a perferryl oxygen. *J Biol Chem* **291**, 17143–17164.

31 Bhakta MN, Hollenberg PF & Wimalasena K (2005) P_{450}/NADPH/O_2- and P_{450}/PhIO-catalyzed N-dealkylations are mechanistically distinct. *J Am Chem Soc* **127**, 1376–1377.

32 Dangi B, Park H & Oh TJ (2018) Effect of alternative redox partners and oxidizing agents in CYP154C8 catalytic activity and product distribution. *ChemBioChem* **19**, 2273–2282.

33 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.

34 Sagadin T, Riehm JL, Milhim M, Hutter MC & Bernhard R (2018) Binding modes of CYP106A2 redox partners determine differences in progesterone hydroxylation product patterns. *Commun Biol* **1**, 99.

35 Schmitz D, Zapp J & Bernhardt R (2014) Steroid conversion with CYP106A2 - production of pharmaceutically interesting DHEA metabolites. *Microb Cell Fact* **13**, 81.

36 Bleif S, Hannemann F, Lisurek M, von Kries JP, Zapp J, Dietzen M, Antes I & Bernhardt R (2011) Identification of CYP106A2 as a regioselective allylic bacterial diterpene hydroxylase. *ChemBioChem* **12**, 576–582.

37 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.

38 Vottero E, Rea V, Lastdrager J, Honing M, Vermeulen NP & Commandeur JN (2011) Role of residue 87 in substrate selectivity and regioselectivity of drug-metabolizing cytochrome P450 CYP102A1 M11. *J Biol Inorg Chem* **16**, 899–912.

39 Kille S, Zilly FE, Acevedo JP & Reetz MT (2011) Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nat Chem* **3**, 738–743.

40 Khatri Y, Jóźwik IK, Ringle M, Ionescu IA, Litzenburger M, Hutter MC, Thunnissen AWH & Bernhardt R (2018) Structure-based engineering of steroidogenic CYP260A1 for stereo- and regioselective hydroxylation of progesterone. *ACS Chem Biol* **13**, 1021–1028.

41 Venkataraman H, Beer SB, Bergen LA, Essen NV, 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 Dangi B, Kim KH, Kang SH & Oh TJ (2018) Tracking down a new steroid-hydroxylating promiscuous cytochrome P450: CYP154C8 from *Streptomyces* sp. W2233-SM. *ChemBioChem* **19**, 1066–1077.

43 Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.

44 Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25.

45 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.

46 Kantardjieff KA & Rupp B (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* **12**, 1865–1871.

47 Adams PD, Afonine PV, Bunkóczi 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.

48 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
49 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355–367.
50 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.
51 Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R & Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477–486.
52 Kuriyama K, Kondo E & Tori K (1963) Conformation of ring A in some 2β-hydroxy- and 2β-acetoxy-Δ^4-3-ketosteroids. *Tetrahedron Lett* **4**, 1485–1491.
53 Duax WL, Eger C, Pokrywiecki S & Osawa Y (1971) Crystalline molecular structures of two derivatives of 2β-hydroxytestosterone having an unusual A-ring conformation. *J Med Chem* **14**, 295–300.
54 Osawa Y & Gardner J (1971) Synthesis and conformation of 2β-hydroxytestosterone. *J Org Chem* **36**, 3246–3247.
55 Burnett RD & Kirk DN (1974) An improved synthesis for 2β-hydroxytestosterone. *J Chem Soc Perkin Trans* **1** (2), 284–285.
56 Duax WL, Cody V, Griffin J, Hazel J & Weeks CM (1978) Steroid structure and function—II. Conformational transmission and receptor binding of medroxyprogesterone acetate. *J Steroid Biochem* **9**, 901–907.