A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Tracking Down a Novel Steroid Hydroxylating Promiscuous Cytochrome P450, CYP154C8 from Streptomyces sp. W2233-SM

Authors: Bikash Dangi, Ki-Hwa Kim, Sang-Ho Kang, and Tae-Jin Oh

To be cited as: *ChemBioChem* 10.1002/cbic.201800018

Link to VoR: http://dx.doi.org/10.1002/cbic.201800018

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

Tracking Down a Novel Steroid Hydroxylating Promiscuous Cytochrome P450, CYP154C8 from *Streptomyces* sp. W2233-SM

Bikash Dangi[1], Ki-Hwa Kim[1], Sang-Ho Kang[2], and Tae-Jin Oh[1,3,4]*

[1]*Department of Life Science and Biochemical Engineering, SunMoon University, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam 31460, Republic of Korea*

[2]*Genomics Division, National Institute of Agricultural Sciences, RDA, Jeonju, 54874, Republic of Korea*

[3]*Department of Pharmaceutical Engineering and Biotechnology, SunMoon University, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam 31460, Republic of Korea*

[4]*Genome-based BioIT Convergence Institute, 70 Sunmoon-ro 221, Tangjeong-myeon, Asan-si, Chungnam 31460, Republic of Korea*

Author for correspondence:

Prof. Tae-Jin Oh

Tel: +82(41)530-2677;

Fax: +82(41)530-2279;

E-mail: tjoh3782@sunmoon.ac.kr

Running title: Steroid hydroxylase CYP154C8

Abstract: CYP154C8 from *Streptomyces* sp. was identified as a novel cytochrome P450 with substrate flexibility to different sets of steroids. The *in vitro* reaction of these steroids with CYP154C8 revealed interesting product formation patterns with the same group of steroids. Nuclear magnetic resonance study revealed the major product of corticosterone hydroxylated at the C21 position; while progesterone, androstenedione, testosterone, and 11-ketoprogesterone were exclusively hydroxylated at 16α position. However, the 16α-hydroxylated product of progesterone was further hydroxylated to yield dihydroxylated products. 16-hydroxyprogesterone was hydroxylated at two positions yielding 2α,16α-dihydroxyprogesterone and 6β,16α-dihydroxyprogesterone. To our knowledge, this is the first report of generation of such products in enzymatic hydroxylation by CYP450. Considering the importance of modified steroids as pharmaceutical components, CYP154C8 has immense potential to be utilized in bioproduction of the hydroxylated derivative compounds directly employed for pharmaceutical applications.

Introduction

Cytochrome P450s (CYP450s) are versatile hemoproteins present in all forms of life.[1] They catalyze oxyfunctionalization of a broad variety of molecules, including steroids, xenobiotics, alkanes, terpenes, antibiotics, and fatty acids.[2, 3] Many physiological functions of bacterial CYP450s remain unknown. However, their flexibility to a vast range of substrates and potential to introduce hydroxyl groups into inactive carbon of different molecules in a regio- and stereo-selective manner make them fascinating catalysts for pharmaceutical industries.[4-8] Typically, hydroxylating compounds can bring dramatic changes to efficacy, activity, toxicity, and solubility in parent molecules.

Steroids are important compounds with a wide range of therapeutic properties such as anticancer, antiandrogenic, progestational, immunosuppressive, diuretic, anti-inflammatory, and contraceptive activities.[9, 10] The physiological function of steroids depends on their structures, especially determined by type, number, and position of their functional groups (Scheme 1). For example, oxyfunctionalization at 11β is crucial for their anti-inflammatory activity.[11] Industrially, regio- and stereo-selective hydroxylation of steroids has important applications for synthesizing highly valuable steroid hormones such as mineralocorticoids, glucocorticoids, and sex hormones. It has been reported that 11α- and 11β-hydroxyprogesterone play an important role in controlling blood pressure with anti-androgenic activity but with minimal estrogenic and progestational side effects.[7, 12–14] Reichstein S, a derivative of progesterone and a corticosteroid precursor with anti-inflammatory activity, is industrially hydroxylated at 11β position. Pregnenolone is synthesized from ergosterol using recombinant *Saccharomyces cerevisiae* harboring CYP450 and its redox partner adrenodoxin and adrenodoxin reductase.[15]

A number of eukaryotic CYP450s involved in steroid synthesis are known, their characteristic expression in membrane-bound form with low catalytic efficiency has limited their potential uses in biotechnology applications. But bacterial CYP450s overexpressed in soluble form, making them suitable for industrial use to produce valuable compounds.[16, 17] However, only a few bacterial CYP450s capable of hydroxylating steroids have been characterized. CYP106A2 from *Bacillus megaterium* ATCC 13368 and CYP106A1 from *B. megaterium* are known for steroid hydroxylating activity.[3, 14] Similarly, CYP514C3 from *S. griseus* and CYP154C5 from *Nocardia francinica* can oxidize steroids such as progesterone, testosterone, nandrolone, dehydroepiandrostenedione, and androstenedione at 16α position.[16, 18] Recently, it has been shown that CYP260A1 and CYP260B1 from *Sporangium cellulosum* and CYP109E1

from *B. megaterium* DSM319 can also hydroxylate steroids.[8, 19-21] CYP154 group of enzymes are widely present in Actinomycetes, and many of them have been identified in *Streptomyces* species. Some of the CYP154 genes have been expressed and characterized. CYP154C1 has shown an *in vitro* activity towards 14 and 12-membered ring macrolide YC-17 and narbomycin, respectively.[22] Similarly, CYP154A1 also was found to bind narbomycin, however, no catalysis was observed with the same substrate.[23] Interestingly, the same enzyme was shown to catalyze unexpected intramolecular cyclization of dipentaenone to a Paterno`-Bu¨chi-like product, without oxidation/reduction.[24] The crystal structure of CYP154C1 and CYP154A1 from *Streptomyces coelicolor* A3(2) has been already been elucidated. A wide range of substrate screening shows CYP154A8 from *N. francinica* and CYP154E1 from *Thermobifida fusca* can oxidize fatty acids, primary alcohols, secondary alcohols, and terpenoids.[25] CYP154H1 from *T. fusca* oxidizes organic sulfides, different arylaliphatic sulfides, indole and indole derivatives, and small arylaliphatic substrates like ethylbenzene, propylbenzene, and styrene.[26] Recently, CYP154F1 from *T. fusca* has been characterized which preferably accepts small linear molecules with a carbon chain length of 8-10 atoms. However, a mutant form of it was able to convert larger substrates like (E)-stilbene and (+)-nootkatone.[27] In this study, we identified another novel steroid hydroxylating CYP450 from *Streptomyces* sp. W2233-SM belonging to CYP154 family and studied its potential use in hydroxylation of diverse sets of steroids *in vitro* and *in vivo* using *Escherichia coli* cells.

Results and Discussion

Screening of steroid modification by *Streptomyces* strain

Initially, *Streptomyces* sp. W2233-SM was checked for its biotransformation ability to diverse steroids. Ethyl acetate extracts of whole-cell bioconversion showed multiple peaks in HPLC

chromatograms of steroid substrates used for biotransformation (Figure S1). Therefore, reaction mixtures were further analyzed by high-resolution electrospray ionization mass spectrometry (HR-ESI/MS) in positive ion mode. Mass spectra of these new peaks resembled exact masses of monohydroxylated products of substrates used. The substrate progesterone with chemical formula of $C_{21}H_{31}O_2^+$ had exact mass of m/z^+ $[M+H]^+$ 315.2557 and a single product (P) had exact mass of m/z^+ $[M+H]^+$ 331.2279, whose calculated m/z^+ $[M+H]^+$ value is 331.2268 (Figure S1A). In case of androstenedione, two products were obtained whereas P1 was identified as monohydroxylated product with an exact mass of m/z^+ $[M+H]^+$ 303.2019, whose calculated mass for the chemical formula $C_{19}H_{27}O_3^+$ is 303.1955 (Figure S1B). The substrate androstenedione ($C_{19}H_{27}O_2^+$) had an exact mass of 287.2161. Similarly, among the three products peak of nandrolone, P1 was found to be a hydroxylated product that had an exact mass of m/z^+ $[M+H]^+$ 291.1955 for which calculated mass for molecular formula $C_{18}H_{27}O_3^+$ was 291.1949. Nandrolone with chemical formula $C_{18}H_{27}O_2^+$ had the exact mass of 275.2004 (Figure S1C).

Identification of strain

Based on 16S rRNA gene sequence, the strain showed close similarity with the genus *Streptomyces*. A comparison of 16S rRNA gene sequences from the selected strains belonging to genus *Streptomyces* is shown in the supplementary information (Figure S2).

Selection of CYP450 responsible for steroid hydroxylation

Among the 20 putative CYP450 in the genome of *Streptomyces* sp. W2233-SM, CYP154C8 showed the highest similarity with previously well-known steroid hydroxylases such as CYP154C3 (74%) and CYP154C5 (66%). Phylogenetic tree analysis of CYP154C8 also revealed the closest relation with CYP154C3 and CYP154C5 among the CYP154 group of

CYP450s (Figure 1). Hence, we speculated the whole-cell bioconversion of steroids in wild-type strain might have been catalyzed by CYP154C8.

Cloning, overexpression, purification, and spectral characterization of CYP154C8 and redox partner

CYP154C8-encoding gene was PCR amplified from the *Streptomyces* sp. W2233-SM using a pair of primers as described in materials and methods. The sequence-confirmed gene was cloned into a pET32a(+) vector under control of a T7 promoter. The redox partner, *camA* and *camB* were cloned separately into pET28a(+) and pET32a(+), respectively. To develop the *in vivo* expression system, *camA* and *camB* were also cloned into pAYAC-Duet1 vector.

All proteins were overexpressed and purified using recombinant *E. coli* BL21(DE3) (Figure 2A). Theoretical molecular weight calculated for CYP154C8 was found to be ~47 kDa, however, the SDS-PAGE analysis showed a band at ~70 kDa due to Trx-His-s-enterokinase fusion sequence of pET32a(+) vector transcribed and translated along with the sequence of CYP154C8. The purified form of CYP154C8 showed a Soret peak at 419 nm in an oxidized form. Carbon monoxide-bound and dithionite-reduced form of CYP154C8 exhibited maximum absorption at a wavelength of 449 nm, one of the major spectral characteristics of CYP450 (Figure 2B). The purity of CYP450 can be evaluated by Rz value, which is calculated as the ratio of absorbance (A) at λ_{max} of the soret band to the A at 280 nm. The purified enzyme had the Rz value of 1.45, indicating high purity.[28]

Substrate binding assay

Substrate binding to active CYP450 sites displaces the water molecule at the sixth coordination position of heme iron, thereby causing a shift of heme iron from low spin to high spin.[29]

CYP450 unbound with any substrate will show a maximum absorption at approximately around 420 nm, whereas substrate binding will cause a minimum absorption at 420 nm with a maximum absorption at 390 nm. This is called type I spectral change.[30] CYP154C8 exhibited maximum A at 389 nm at high spin and at 419 nm at low spin for all substrates (Figure S3). Steroids with higher hydrophobicity (group I) showed markedly lower K_d values, indicating tight binding. Progesterone, androstenedione, testosterone, and nandrolone had K_d values of less than 0.5 µM. However, substrates such as prednisone and cortisone had K_d values of 2.50 µM and 4.16 µM, respectively, while prednisolone and hydrocortisone had K_d values of 12.78 µM and 15.80 µM, respectively. 16α-hydroxyprogesterone showed the K_d value of 1.21 ± 0.1. Progesterone as a substrate had the lowest K_d value at 0.083 µM while hydrocortisone had the highest K_d value at 15.80 µM (Table 1).

Corticosterone did not induce a high-spin shift of heme among steroid substrates used in this study. Group I steroids K_d value appeared to be exceptionally small. These steroids had K_d value below 0.5 µM, indicating high affinity, while group II steroids had comparatively large K_d values. This clearly indicates with the increase in the hydrophobicity of the substrates, the binding affinity of CYP154C8 increases. As reported previously, CYP154C3 and CYP154C5 also displayed tight binding (low K_d) with steroids. Tight binding is likely to be the consequence of steroid hydrophobicity which leads to an unfavorable entropy reduction of the water molecules surrounding steroids in an aqueous environment of the active site. Substrate binding to the active site releases the low-entropic water molecule from the solvation shell of steroids, thus increasing the entropy of the system. This result in large hydrophobic interaction interfaces which promotes the overall process of steroids binding to the active site of CYP450.[31]

In vitro assay analysis and product identification

The *in vitro* reaction mixture of CYP154C8 reconstituted with PDX (redoxin) and PDR (reductase) as redox partner, and NADH as cofactor showed a single product peak for the group I, III, and IV steroids (Figure 3A), while group II substrates had multiple products peaks (Figure 3B). LC-MS analysis of the reaction mixtures showed the exact mass of monohydroxylated products of the respective substrate (Table S1). However, progesterone reaction mixture showed three product peaks (P1, P2, and P3), where P1 and P2 were shown to have a very low conversion, and P1 was found to be a major product (Figure 4 inset I). LC-MS analysis of the reaction mixture showed P1 and P2 as dihydroxylated products, and P3 a monohydroxylated product of progesterone (data not shown). We thought the two dihydroxylated products might be due to subsequent hydroxylation of the 16α-hydroxyprogesterone by CYP154C8. To confirm it, 16α-hydroxyprogesterone was purified as described in material and method, and it was used as a substrate for *in vitro* reaction with CYP154C8. Reaction mixture analysis of 16α-hydroxyprogesterone showed two product peaks (P1 and P2) in HPLC (Figure 4 inset II); P1 and P2 were found to be a major and a minor product of 16α-hydroxyprogesterone, respectively. The same two products were confirmed by HPLC (Figure 4 inset II) and LC-MS analyses (data not shown) which were obtained in an *in vitro* reaction with the substrate progesterone. The two dihydroxylated products, P1 and P2 of progesterone characterized by NMR were identified as 2α,16α-dihydroxyprogesterone (major) and 6β,16α-dihydroxyprogesterone (minor), respectively. To our knowledge, this is the first report of such dihydroxylated products formation by any CYP450 enzyme in a two-step oxidation of progesterone. Although, *S. roseochromogenes* NCIB 10984 was shown to hydroxylate exogenous progesterone to 16α-hydroxyprogesterone and subsequently to 2β,16α-dihydroxyprogesterone.[32] Some of the CYP450s like CYP17A1, cytochrome P450scc

(CYP11A1), cytochrome P450BioI (CYP107H1), and MycG have been known to catalyze the multi-step oxidation of their respective substrate in a sequential manner.[33-42] The monohydroxylated products of progesterone, androstenedione, testosterone, and 11-ketoprogesterone analyzed by NMR (^{1}H-NMR, C^{13}-NMR, and 2D-NMR) showed hydroxylation at 16α-position. 16α-hydroxydehydroepiandrosterone and its 3β-sulfate ester, 16α-hydroxydehydroepiandrosterone sulfate, are metabolic intermediates in the biosynthesis of estriol from dehydroepiandrosterone during pregnancy. 16α-hydroxydehydroepiandrosterone has estrogenic activity.[43, 44] 16α-hydroxyprogesterone and 6β-hydroxyprogesterone have shown to inhibit 5α-reductase, and these modified steroids can be good drug target for the prostate cancer.[45]

However, different product formation patterns were observed when substrates (group II) had hydroxyl or carbonyl groups at C11 and C21 positions. Corticosterone was converted into three products (Figure 3B inset II). HPLC chromatogram showed a flat peak at the base of a major product (P2) which was unusual. To further characterize this peak, LC-MS and NMR analyses were carried out. Hydroxylation at the C21 position of corticosterone was found, yielding (11β)-11,21,21-trihydroxypregn-4-ene-3,20-dione (21-hydroxycorticosterone) which was an unusual conversion (Scheme 2). This is the first report of such product formed by any enzymatic conversion reactions. Although the same product showing similar peak characteristic (a flat peak at the base) in HPLC chromatogram has been reported previously,[46] it was a hydrocortisone degradation product synthesized by chemical method. A similar mechanism of forming such corticosteroids has also been reported previously by chemical methods.[47–49] This product was confirmed to be the result of CYP154C8 catalysis in this study by carrying out series of control reactions. 21-hydroxycorticosterone might have resulted due to dehydrogenation of 21-hydroxyl

of corticosterone by CYP154C8 giving rise keto aldehyde form. A hydrated form of the product (21-hydroxycorticosterone) is expected to be more favorable over the keto aldehyde form in aqueous solution,[46] although the keto aldehyde form was never detected in our reactions with corticosterone.

The ions at m/z 121 and 145 of all three products (Figure S4) indicate the presence of an unmodified steroidal A/B-ring.[50, 51] A possible position of hydroxylation of P1 was estimated at C16 of D-ring based on HPLC analysis, LC-MS fragmentation, and characteristics of CYP154C8. A comparative study of P1 with hydrocortisone (a compound with same molecular weight and chemical formula as P1) showed no similarity in HPLC retention time, while LC-MS fragments were found to be highly similar. It showed P1 was not a C17 hydroxylated product. Comparison of the HPLC retention pattern of P1 with other known hydroxylated products (16α hydroxylated products of progesterone, androstenedione, and testosterone) showed strong evidence of P1 being hydroxylated at C16 position (Figure 3). This was further supported by the characteristic of CYP154C8 which had been shown to catalyze hydroxylation at C16 position. Another product P3 had the mass of m/z^+ $[M+H]^+$ 361.2007 for which the calculated mass of molecular formula $C_{21}H_{29}O_5^+$ was 361.2010. This might be due to the presence of a double bond, a keto group, or epoxide formation as keto groups and epoxides can form from hydroxyl groups.[19] It is possible that P3 might be a product obtained by oxidation of P2 at C21 (C-OH → C=O) to give keto acid. All group II substrates were hydroxylated in multiple positions with the characteristic of peaks similar to those of corticosterone P2. Therefore, those peaks might represent glyoxal hydrate of respective substrates, similar to corticosterone P2 (glyoxal hydrate). Functionalization at C11 position of steroids might play an important role in positioning steroids at the active site of CYP154C8. To determine such possibility, more steroids (group III) with the functional group

(hydroxyl or carbonyl) at C11 position were selected. However, a single product was detected in HPLC (Figure 3A inset IV and VI) and LC-MS analysis (Table S1). Speculating that the presence of C21 hydroxyl group might have a role in multiple product formations, steroids lacking C11 functional group but having C21 hydroxyl group (group IV) were further selected. A single monohydroxylated product was also detected by HPLC (Figure 3A inset V) and LC-MS (Table S1). This clearly indicates that CYP154C8 catalyzes hydroxylation in more than one position when the substrate has a functional group (hydroxyl or carbonyl) at C11 and C21 positions. P2 of group II steroids was a result of such structure. Hence, group II steroids having a hydroxyl or carbonyl functional group at C11 and C21 positions were metabolized into two or more products. All substrates that were hydroxylated at one position were believed to be 16α-hydroxylated products based on HPLC peak analysis, substrate structure, and comparative studies of CYP154C8 with already characterized similar proteins (CYP154C5 and CYP154C3). Although CYP154C8 has a major characteristic of hydroxylating at C16α, the presence of hydroxyl or carbonyl groups at C11 and C21 might have changed the conformation of a substrate at active sites. It has been previously reported that CYP154C5 shows major differences in steroid binding orientation between C17-acetyl-substituted steroids (progesterone and pregnenolone) and C17-hydroxyl/carbonyl substituted steroids (androstenedione and testosterone).[31] Therefore, steroid orientation is highly influenced by C17 substituents. Other substrates with bulky C17 substituent groups such as lanosterol, cholesterol, and desmosterol were also screened in this study. However, no conversion was found *in vitro* (data not shown). Dexamethasone, a C16-methylated steroid, also showed no conversion. Protein multiple sequence alignment of CYP154C8 and CYP154C5 shows a few differences in the active sites (Figure S5). Polar amino acid Gln239 is thought to play an important role in the dynamic process of steroid binding at the

active site of CYP154C5.[31] It was substituted by Lys249 in CYP154C8. Other active site amino acids Val87 and Val291 in CYP154C5 were replaced by Ala98 and Thr301 in CYP154C8, respectively. CYP154C5 has been reported in bioconversion of the only group I like substrates; it would be interesting to see the products of group II substrates catalyzed by it, which has not been previously reported.

Kinetic analysis

Progesterone was found to be the most preferred substrate in *in vitro* conversion assay. Hence, we selected the substrate progesterone to determine the kinetic parameters of CY154C8-catalyzed hydroxylation. The kinetics parameters for CYP154C8 were determined by plotting product formation rate vs substrate concentration. The K_m and k_{cat} values were estimated to be 28.31 ± 3.35 µM and 2.38 ± 0.078 min^{-1}, respectively (Figure 5).

Biotransformation of steroids using recombinant *E. coli*

The *in vivo* bioconversion of steroids by *E. coli* BL21(DE3) cells harboring pAYAC_camAB and pET32a_CYP154C8, and only pET32a_CYP154C8 was found to be similar. Expression level analysis of CYP154C8, PDX, and PDR by SDS-PAGE showed a clear band of CYP154C8 in all system, however, no clear band of PDX and PDR was observed (Figure S6). This showed that the intrinsic redox partner of *E. coli* BL21(DE3) effectively transfer reducing equivalents to CYP154C8. In *E. coli*, the ferredoxin-$NADP^+$ oxidoreductase transfers reducing equivalent between NADPH and iron-sulfur clusters of various ferredoxins, and in some cases NADPH and the FMN containing flavodoxin.[52] As reported previously, bovine cytochrome P450C17 hydroxylase activity is supported by NADPH-flavodoxin (ferredoxin) oxidoreductase in combination with flavodoxin.[53] Another cytochrome P450cin (CYP176A) is believed to have a

flavodoxin as redox partner that might utilize the endogenous *E. coli* flavodoxin/flavodoxin reductase system.[54] This shows endogenous flavodoxin and flavodoxin reductase might have possibly involved in transferring reducing equivalents to CYP154C8. However, the involvement of ferredoxin and ferredoxin reductase cannot be ignored.

In vivo bioconversion of C21 hydroxysteroids (prednisolone, hydrocortisone, prednisone, cortisone, and 11-deoxycortisol) except corticosterone showed a major peak that had longer retention time than the substrate peak (Figure S7 inset I). LC-MS analysis of products revealed the exact mass of acetylated steroid of the respective substrate (Table S2). Cortisone and prednisone were most preferred substrates for formation of such product (Figure S8). A similar phenomenon has been previously reported where chloramphenicol acetyltransferase I (CATI) was found to transfer an acetyl group to C21 hydroxysteroids. All the substrates were selectively transformed into corresponding 21-acetoxy derivatives.[55] while the hydroxylated products from these substrates except 11-deoxycortisol was found to be a minor product (less than 1%). Surprisingly, no acetylated product of corticosterone was detected which is also a C21 hydroxysteroid. The *in vitro* reaction with acetyl-CoA did not show any acetylated products for such substrates. The plasmid vector pAYACDuet-1 used for the cloning of *camA* and *camB* gene as redox partner contains chloramphenicol resistance gene. The activation of the chloramphenicol acetyltransferase in *E. coli* to detoxify the antibiotic chloramphenicol might have involved in the acetylation of C21 steroids.[56] To confirm about it, the pCDFDuet-1 vector which does not contain chloramphenicol resistance gene was selected for the expression of *camA* and *camB* in *E. coli* cells. The *In vivo* bioconversion of such steroids harboring pET32a_cyp154C8 and pCDF_camAB recombinant plasmids showed no possible acetylated product peak in HPLC chromatogram (Figure S7 inset II). This clearly indicates that C21

hydroxysteroids were acetylated by chloramphenicol acetyltransferase. The hydroxylated product yield of C21 hydroxysteroids with the pCDFDuet-1 vector system was not found to have improved, however, the bioconversion of progesterone and androstenedione was increased, comparatively (Figure 6). Progesterone was most favored substrate with overall bioconversion of ~93% that showed a major conversion (~90%) into a monohydroxylated product (16α-hydroxyprogesterone) and a low conversion (~3%) into two dihydroxylated products (2α,16α-dihydroxyprogesterone and 6β,16α-dihydroxyprogesterone). The least favored substrate *in vivo* were group II substrates which had a very low bioconversion into hydroxylated products. The bioconversion of corticosterone was estimated to be 7.2 ± 2.5 %. The group I substrates (progesterone, androstenedione, testosterone, and nandrolone) along with 11-ketoprogesterone were the most preferred substrates by an *in vivo* bioconversion (Figure S9).

Conclusions

We have investigated a new steroid hydroxylating CYP450 by genome mining. CYP154C8 was found to exhibits oxyfunctionalization of broad steroid substrates. The regio- and stereo-specific hydroxylation at C16α position with the steroids lacking functional group at C11 and C21 position, and stereo- and regio-selective hydroxylation of corticosterone at C21 position bearing functional group at both positions is an interesting mechanism of functioning. Along with this, CYP154C8 can subsequently hydroxylate 16α-hydroxyprogesterone to give rise two dihydroxylated products. We believe CYP154C8 can be a suitable model, and biocatalyst for tailoring CYP450 activities to generate steroids with therapeutic values and other high-value metabolites.

Experimental Section

Chemicals and reagents: All steroid substrates and sodium formate were purchased from Tokyo Chemical Industry Co., Ltd. (Korea). Isopropyl-1-thio-β-$_D$-galactopyranoside (IPTG), 1,4-dithiothreitol (DTT), and kanamycin were obtained from Duchefa Bohemie (Korea). Ampicillin (Amp), chloramphenicol (Cm), α-aminolevulinic acid (ALA), nicotinamide adenine dinucleotide (NADH), cytochrome C, catalase, and formate dehydrogenase were purchased from Sigma-Aldrich (Korea). Restriction enzymes were obtained from Takara Clontech (Korea). T4 DNA ligase, DNA polymerase, and dNTPs were from Takara Bio (Japan). All other chemicals were high-grade products obtained from commercially available sources.

Sequence accession numbers and bioinformatics analysis: Gene responsible for encoding CYP450 was found based heme-binding domain signature, FxxGx(H/R)xCxG, as query sequence from *Streptomyces* sp. W2233-SM genome (unpublished). The nucleotide sequence of CYP154C8 and 16S ribosomal RNA gene sequence of the strain has been deposited into the GenBank under the accession number MF398962 and MG198705, respectively. Microbial identification was carried out based on 16S ribosomal gene sequence. The name of CYP450 was assigned by Dr. David Nelson (http://drnelson.utmem.edu/CytochromeP450.html). The evolutionary history was inferred using the Neighbor-Joining method.[57] The optimal tree with the sum of branch length = 6.82869639 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown below the branches.[58] The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method[59] and are in the units of the number of amino acid substitutions per site. The analysis involved 25 amino acid sequences. All positions

containing gaps and missing data were eliminated. There were a total of 289 positions in the final dataset. Evolutionary analyses were conducted in MEGA6.[60]

Strains, media, and conditions: Wild-type *Streptomyces* sp. W2233-SM isolated from Korean soil was cultivated in Bacto™ Tryptic Soy Broth (TSB; 17.0 g L^{-1} pancreatic digest of casein, 3.0 g L^{-1} papaic digest of soybean, 5 g L^{-1} sodium chloride, 2.5 g L^{-1} dextrose, and 2.5 g L^{-1} dipotassium phosphate, media was adjusted to pH 7.3) at 28°C with shaking (200 rpm). After 120 h of culture, the cells were collected by centrifugation followed by washing twice with Tris buffer (pH 7.4). The washed cells were resuspended in potassium phosphate buffer (pH 7.4) and stored in -20°C until use. *E. coli* XL1 Blue (Stratagene, La Jolla, CA, USA), pMD20-T (Takara), pET28a(+), pET32a(+), and pAYAC-Duet1 (Novagen, Germany) were used for sub-cloning and DNA manipulation. *E. coli* BL21(DE3) (Stratagene) was used for recombinant protein expression and whole-cell biotransformation. All *E. coli* strains were grown at 37°C in Luria-Bertani (LB) media or plates supplemented with ampicillin (100 µg/ml), streptomycin (50 µg/ml), and chloramphenicol (34 µg/ml) when required. X-gal and IPTG were used for colony screening and protein expression.

Whole-cell biotransformation by wild-type strain: *Streptomyces* sp. W2233-SM were screened for biotransformation capability of steroids. The bioconversion assay was performed in potassium phosphate buffer (pH 7.4) with substrates (progesterone, androstenedione, and testosterone) at concentration of 1 mM at 28°C (200 rpm). After 24 h of bioconversions, the reaction mixture was extracted with ethyl acetate. The ethyl acetate fractions were collected, dried, and reconstituted in 500 ml of acetonitrile. High-performance liquid chromatography photo-diode array and liquid chromatography-mass spectrometry (LC-MS) analysis were used for further characterization.

Cloning and overexpression of CYP154C8 and redox partner: The CYP154C8 encoding sequence (1,266 bp, 421 amino acids, accession number MF398962) was amplified from the genomic DNA of *Streptomyces* sp. W2233-SM. The PCR primers used for amplification of the gene were designed as 5' - GAA TTC ATG AAC GGT CAG TCA GCG A - 3' (*Eco*RI) as forward and 5' - AAG CTT TCA GCT GCC GTG GAG CA - 3' reverse primer (*Hin*dIII). The letters underlined indicate the restriction site for an endonuclease in bracket, respectively. The PCR product obtained was cloned into pMD20-T vector using *E. coli* XL1-Blue and the nucleotide sequence was confirmed by automated sequencing (Macrogen, Korea). Further, the gene product was ligated into pET32a(+) vector to create pET32aCYP154C8 construct. The resulting construct encoding N-terminal His6-tag protein under the control of a T7 promoter was transformed into chemically competent *E. coli* Bl21 (DE3) and plated on LB agar containing 100 µg/ml ampicillin. A single colony from the plate was selected and grown overnight at 37°C. 1 ml of overnight grown seed culture was added to 100 ml of LB-medium supplemented with 100 µg/ml ampicillin and incubated at orbital shaker (180 rpm) at 37°C until cell density was about 0.6-0.8 at OD_{600nm}. Cultures were induced with 0.5 mM of IPTG, and supplemented with 1 mM of 5-aminolevulinic acid hydrochloride (5-ALA) and 0.5 mM $FeCl_3$ to support heme synthesis. The cells were incubated for 48 h at 20°C. The cell pellets were harvested by centrifugation (3,500 rpm) for 20 min at 4°C and washed twice with 50 mM Tris-HCl buffer (pH 7.4) containing 10% glycerol, 100 mM NaCl, and 1 mM DTT. The *in vitro* reconstituted system was supported by redox partners PDX (*camA*) and PDR (*camB*).[61] They were expressed as His-tagged proteins in *E. coli* BL21(DE3) using plasmid constructs pET28a(+) and pET32a(+) as described previously.[62] For the *in vivo* system, the *camA*-harboring plasmid pET28a_PDX was digested with *Nde*I and *Xho*I restriction enzymes to obtain *camA*. The *camB*-carrying plasmid

pET32a_PDR was digested with *Eco*RV and *Hin*dIII to obtain *camB*. These genes were subsequently ligated into respective sites of vector pAYAC-Duet to create plasmid pAYAC_camAB. Similarly, *camA* and *camB* genes previously cloned into pCDF-duet1 vector was also used separately for *in vivo* conversion of steroid substrates.[63]

CYP154C8 and redox partner purification: Cell pellets were harvested after protein overexpression and lysed by ultra-sonication followed by centrifugation at 15,000 rpm for 40 min at 4°C. Soluble fractions of these proteins were purified with Ni^{2+} affinity chromatography using TALON His-tag; soluble protein extracts were mixed and agitated with TALON His-tag resin on ice for 60 min. Protein-bound resins were pre-equilibrated with two-column volumes of equilibrium buffer (potassium phosphate, pH 7.4). Bound proteins were then eluted with elution buffer (potassium phosphate, pH 7.4, 10% glycerol, and 100 mM NaCl) containing 20 mM, 100 mM, and 250 mM imidazole. Fractions containing concerned proteins were pooled and concentrated by ultra-filtration using Amicon centrifugal filters (Millipore) with molecular weight cut-off (MWCO) of 30 kDa for CYP154C8 and PDR and MWCO of 10 kDa for PDX.

Determination of enzyme concentrations: CYP154C8 concentration was determined based on CO A-difference spectra as described previously.[64] Proteins (2 ml) were diluted with potassium phosphate buffer (50 mM, pH 7.4) and separated into two cuvettes. A pinch of sodium dithionite was added into both cuvettes. One cuvette of samples was bubbled with CO. Both samples were then scanned using a Biochrome Libra S35PC UV/Visible Spectrophotometer (England) between wavelength of 400 and 500 nm repeatedly until the reading at 449 nm was lower than the highest reading. CYP154C8 amount was calculated based on $\varepsilon_{449-489}$ = 91 $mM^{-1}cm^{-1}$. PDR concentration was determined as the average of concentrations calculated from wavelengths 378, 454, and 480 nm using extinction coefficients (ε) of 9.7, 10.0, and 8.5 $mM^{-1}cm^{-1}$, respectively, and PDX

concentration was also determined as the average of concentrations calculated from wavelengths 415 and 454 nm using extinction coefficients of 11.1 and 10.4 $mM^{-1}cm^{-1}$, respectively.[65] PDX and PDR activities were determined by monitoring cytochrome C reduction at 550 nm using extinction coefficient of 19.1 $mM^{-1}cm^{-1}$.[66]

Substrate binding assay: CYP154C8 dissociation constants (K_d) were determined by titrating substrate concentrations until they were saturated. CYP154C8 enzyme was prepared at concentration of 1 µM in 50 mM phosphate buffer (pH 7.4). Steroid substrates were prepared in DMSO. A values of all samples were measured at wavelength of 350–500 nm using Biochrome Libra S35PC spectrophotometer. K_d of different steroids were obtained by fitting plots of resulting A differences ($Abs_{389\ nm}$ - $Abs_{419\ nm}$) against substrate concentrations using tight binding equation as described previously.[67] Graphpad prism 6 software was used to fit titration data points to the following quadratic equation[68]: $A_{obs}=A_{max}\{([S]+[E_t]+K_D)-(([S]+[E_t]+K_D)\text{^}2-(4[S][E_t]))\text{^}0.5\}/2[E]$, where A_{obs} was the A difference at any ligand concentration, A_{max} was the A difference at ligand saturation, S was substrate concentration, E_t was enzyme concentration, and K_d was dissociation constant for enzyme-ligand complex. The K_d values were calculated as average of three different experiments.

***In vitro* assay:** The reaction mixture contained 3 µM CYP154C8, 6 µM PDR , 24 µM PDX , 1 U formate dehydrogenase, 100 µg/ml catalase, 1 mM $MgCl_2$, 150 mM sodium formate, and 500 µM substrate in 50 mM potassium phosphate buffer (pH 7.4; 10% glycerol and 100 mM NaCl). The reaction was initiated by adding 250 µM of NADH to make a final mixture volume of 500 µL. The reaction mixture was incubated at 30°C with shaking (200 rpm) for 2 h. All the steroid substrates were prepared in DMSO at stock concentration of 20 mM. Reactions were then extracted twice with 500 µl of ethyl acetate. Organic phases were collected and dried for analysis.

Kinetic analysis: NADH consumption rates were spectrophotometrically measured at 25°C in phosphate buffer (50 mM, pH 7.4). The reaction mixture contained 1 µM CYP154C8, 1 µM PDR, 18 µM PDX, 0.5 mM substrate, and 250 µM NADH. NADH consumption by enzymes was measured at wavelength of 340 nm (ε= 6.22 $mM^{-1}cm^{-1}$) with a UV-Vis scanning photometer (UV-1601PC, Shimadzu, Tokyo, Japan). NADH consumption rates with different substrate concentrations (10–180 µM) were measured to determine K_m. All reaction mixtures were extracted with ethyl acetate after all NADH was consumed. Product formation was calculated using the area of the product formed from HPLC. Coupling efficiency was calculated using amounts of product formed at different substrate concentrations. Coupling efficiency (%) was determined as the amount of product formed (µM)/total NADH consumed (µM) × 100. K_m and k_{cat} values were calculated by plotting product formation rate vs. concentration.

Whole-cell bioconversion by *E. coli* BL21(DE3) and product purification for analysis: Frozen *E. coli* cells stored at -20°C overexpressing CYP154C8, PDX, and PDR were resuspended in potassium phosphate buffer (pH 7.4) to make the desired final OD_{600} value of 50. Bioconversion were carried out in 1 ml and 100 ml scale for analytical and preparative purpose, respectively, at 30°C for 24 h and shaking 200 rpm with the addition of glucose (0.5 mg/ml). All the steroids stocks were prepared in DMSO at the concentration of 100 mM. From these prepared stocks, respective amounts of steroid were added to reaction mixtures to achieve final steroid concentration of 1 mM. After 24 h of incubation, the bioconversions and the control were extracted with ethyl acetate. 1 ml of bioconversions was extracted twice with 1 ml of ethyl acetate while 100 ml of bioconversions was extracted with 300 ml of ethyl acetate. The supernatant was dried under pressure and dissolved in HPLC-grade methanol. The samples were filtered through 0.45 µm pore sized PTFE filter, and subjected to preparative HPLC (Shimadzu)

with a C_{18} column (Mightysil RP-18 GP, 150 × 4.6 mm, 5 µm, Kanto Chemical, Japan). To estimate the amount of product formed, λ_{max} (absorption maximum) of products and substrates were identified. The peak area of each product and substrate were calculated by using their λ_{max} and conversion percentage were calculated. The hydroxylated product of corticosterone for NMR analysis was prepared similar to an *in vitro* conversion assay described previously. A total volume of 100 ml reaction mixture was distributed into 0.5 ml and incubated at 30°C and 200 rpm for 2 h. Further, the samples for purification of the hydroxylated product were prepared as described above.

Analytical methods: Samples were injected into UHPLC and separated with Mightysil Reverse phase C_{18} GP column (4.6 × 250 mm, 5 µm). Water (A) and acetonitrile (B) were used as mobile phases for separation. The reaction mixtures were analyzed with a gradient system of B at 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. Substrates and their products were detected by UV-A at 242 and 245 nm. Reaction mixtures were analyzed with SYNAPT G2-S/ACUITY UPLC liquid chromatography quadrupole time-of-light/electrospray ionization mass spectrometry (Waters, Milford, MA, USA) in positive ion mode.

The purified hydroxylated products were dissolved in dimethyl-sulfoxide (DMSO-d_6) and subjected to NMR analyses on 800 MHz using Varain Unity INOVA spectrometer (Varian, Palo Alto, CA, USA). One dimensional NMR (^{1}H-NMR and ^{13}C-NMR) was performed followed by two-dimensional NMR, heteronuclear multiple bond correlation, correlation spectroscopy, rotating-frame overhauser effect spectroscopy (ROESY), and heteronuclear single quantum coherence spectroscopy (HSQC) to elucidate exact structures when appropriate.

16α-Hydroxyprogesterone. ^{1}H NMR (800 MHz, DMSO-d6) data for 16α-hydroxyprogesterone were: δ 5.64 (s, 1H), 4.73 (d, J = 4.8 Hz, 1H), 4.53 (dt, J = 7.5, 5.3 Hz, 1H), 2.44 (d, J = 6.5 Hz, 1H), 2.43 – 2.38 (m, 1H), 2.41 – 2.36 (m, 1H), 2.25 (dt, J = 13.0, 2.4 Hz, 1H), 2.17 (dt, J = 16.9, 3.8Hz, 1H), 2.10 (s, 2H), 1.97 (ddd, J = 13.4, 5.1, 3.1 Hz, 1H), 1.90 (dt, J = 12.5, 3.4 Hz, 1H), 1.75 (ddt, J = 12.3, 5.4, 2.7 Hz, 1H), 1.65 – 1.58 (m, 2H), 1.54 (dq, J = 13.0, 3.6 Hz, 1H), 1.51 (dd, J = 10.8, 3.4 Hz, 1H), 1.50 – 1.46 (m, 1H), 1.49 – 1.45 (m, 2H), 1.35 (qd, J = 13.1, 4.0 Hz, 1H), 1.14 (s, 3H), 1.01 – 0.95 (m, 2H), and 0.58 (s, 3H). ^{13}C NMR (201 MHz) data for 16α-hydroxyprogesterone were: δ 208.55 (C20), 198.52 (C3), 171.30 (C5), 123.66 (C4), 73.36 (C17), 70.89 (C16) 53.41 (C9), 53.44 (C14), 44.75 (C13), 38.63 (C10), 38.39 (C12), 35.46 (C1), 35.95 (C15), 34.89 (C8), 34.05 (C2), 32.39 (C6), 32.19 (C21), 32.01 (C6), 20.66 (C11), 17.32 (C19), and 14.67 (C18).

16α-Hydroxyandrostenedione. ^{1}H NMR (800 MHz, DMSO-d6) data for 16α-hydroxyandrostenedione were: δ 5.64 (s, 1H), 4.73 (d, J = 4.8 Hz, 1H), 4.53 (dt, J = 7.5, 5.3 Hz, 1H), 2.44 (d, J = 6.5 Hz, 1H), 2.43 – 2.38 (m, 1H), 2.41 – 2.36 (m, 1H), 2.25 (dt, J = 13.0, 2.4 Hz, 1H), 2.17 (dt, J = 16.9, 3.8 Hz, 1H), 2.10 (s, 2H), 1.97 (ddd, J = 13.4, 5.1, 3.1 Hz, 1H), 1.90 (dt, J = 12.5, 3.4 Hz, 1H), 1.75 (ddt, J = 12.3, 5.4, 2.7 Hz, 1H), 1.65 – 1.58 (m, 2H), 1.54 (dq, J = 13.0, 3.6 Hz, 1H), 1.51 (dd, J = 10.8, 3.4 Hz, 1H), 1.50 – 1.46 (m, 1H), 1.49 – 1.45 (m, 2H), 1.35 (qd, J = 13.1, 4.0 Hz, 1H), 1.14 (s, 3H), 1.01 – 0.95 (m, 2H), and 0.58 (s, 3H). ^{13}C NMR (201 MHz) data for 16α-hydroxyandrostenedione were: δ 208.55 (C20), 198.52 (C3), 171.30 (C5), 123.66 (C4), 73.36 (C17), 70.89 (C16), 53.41 (C9), 53.44 (C14), 44.75 (C13), 38.63 (C10), 38.39 (C12), 35.46 (C1), 35.95 (C15), 34.89 (C8), 34.05 (C2), 32.39 (C6), 32.19 (C21), 32.01 (C6), 20.66 (C11), 17.32 (C19), and 14.67 (C18).

16α-Hydroxytestosterone. ^{1}H NMR (900 MHz, DMSO-d6) data for 16α-hydroxytestosterone were: δ 5.63 (d, J = 1.8 Hz, 1H), 4.67 (d, J = 4.8 Hz, 1H), 4.61 (d, J = 5.1 Hz, 1H), 3.82 (dt, J = 9.2, 3.8 Hz, 1H), 3.21 (t, J = 4.8 Hz, 1H), 2.39 (ddd, J = 16.5, 14.8, 5.1 Hz, 1H), 2.38 (tdd, J = 14.3, 5.4, 2.0 Hz, 2H), 2.24 (ddd, J = 14.6, 4.2, 2.4 Hz, 1H), 2.16 (dt, J = 16.7, 3.8 Hz, 1H), 1.96 (ddd, J = 13.3, 5.1, 3.1 Hz, 1H), 1.74 (ddt, J = 12.0, 5.7, 2.8 Hz, 1H), 1.70 (dt, J = 12.6, 3.4 Hz, 1H), 1.64 (dt, J = 13.0, 9.3 Hz, 1H), 1.60 (dt, J = 14.1, 4.4 Hz, 1H), 1.53 – 1.47 (m, 2H), 1.35 (ddd, J = 13.2, 7.5, 2.1 Hz, 1H), 1.34 (qd, J = 13.1, 4.2 Hz, 1H), 1.20 (ddd, J = 12.5, 10.9, 7.5 Hz, 1H), 1.14 (s, 3H), 1.04 (dt, J = 12.9, 4.2 Hz, 1H), 0.96 – 0.89 (m, 1H), 0.93 – 0.88 (m, 1H), and 0.68 (s, 3H). ^{13}C NMR (226 MHz) data for 16α-hydroxytestosterone were: δ 198.58 (C3), 171.59 (C5), 123.60 (C4), 88.97 (C17), 77.04 (C16), 53.83

(C9), 48.22 (C14), 43.36 (C13), 38.72 (C10), 36.82 (C12), 35.49 (C1), 34.73 (C15), 34.07 (C2), 32.44 (C6), 31.76 (C7), 20.31 (C11), 17.37 (C19), and 12.82 (C18).

(11β)-11,21,21-Trihydroxypregn-4-ene-3,20-dione. ^{1}H NMR (900 MHz, DMSO-d6) data for (11β)-11,21,21-trihydroxypregn-4-ene-3,20-dione were: δ 5.57 (d, J = 1.8 Hz, 1H), 4.89 (s, 1H), 4.19 (q, J = 3.3 Hz, 1H), 2.86 (t, J = 9.2 Hz, 1H), 2.45 (tdd, J = 14.0, 5.6, 1.9 Hz, 1H), 2.39 (ddd, J = 16.4, 13.7, 5.0 Hz, 1H), 2.22 – 2.16 (m, 2H), 2.13 – 2.07 (m, 1H), 2.08 (dd, J = 13.5, 2.5 Hz, 1H), 2.03 (dddd, J = 13.8, 11.7, 9.0, 2.9 Hz, 1H), 1.92 (ddt, J = 17.2, 7.0, 3.6 Hz, 1H), 1.88 (ddd, J = 22.0, 10.9, 3.9 Hz, 1H), 1.82 – 1.76 (m, 1H), 1.68 (dddd, J = 12.2, 9.8, 7.1, 2.9 Hz, 1H), 1.60 (dddd, J = 12.4, 9.3, 6.7, 3.7 Hz, 1H), 1.48 (dd, J = 13.6, 3.4 Hz, 1H), 1.37 (s, 3H), 1.31 – 1.23 (m, 1H), 1.12 (ddd, J = 12.6, 10.5, 7.1 Hz, 1H), 1.03 – 0.96 (m, 1H), 0.95 – 0.92 (m, 1H), and 0.79 (s, 3H). ^{13}C NMR (226 MHz) data for (11β)-11,21,21-trihydroxypregn-4-ene-3,20-dione were: δ 208.70 (C20), 198.58 (C3), 172.83 (C5), 121.94 (C4), 90.68 (C21), 66.70 (C11), 57.48 (C14), 57.18 (C17), 55.86 (C9), 47.34 (C12), 44.15 (C13), 39.31 (C10), 34.46 (C1), 33.93 (C2), 32.96 (C7), 31.81 (C6), 31.61 (C8), 24.67 (C15), 23.30 (C16), 20.85 (C19), and 16.31 (C18).

11-keto,16α-Hydroxyprogesterone. ^{1}H NMR (900 MHz, DMSO-*d*6) δ 5.65 (d, *J* = 1.8 Hz, 1H), 4.97 (d, *J* = 4.8 Hz, 1H), 4.58 – 4.54 (m, 1H), 2.75 (d, *J* =12.2 Hz, 1H), 2.67 (d, *J* = 6.7 Hz, 1H), 2.56 (ddd, *J* = 13.6, 5.1, 3.3 Hz, 1H), 2.46 – 2.40 (m, 2H), 2.45 – 2.38 (m, 2H),2.35 (d, *J* = 12.1 Hz, 1H), 2.27 (ddd, *J* = 14.7, 4.1, 2.3 Hz, 1H), 2.22 (d, *J* = 11.2 Hz, 1H), 2.18 – 2.12 (m, 1H), 2.13(dt, *J* = 16.8, 3.7 Hz, 1H), 2.10 (s, 3H), 1.88 – 1.83 (m, 2H), 1.73 (td, *J* = 13.0, 8.7 Hz, 1H), 1.64 (td, *J* = 14.2, 4.4 Hz,1H), 1.60 (ddd, *J* = 13.4, 7.7, 1.7 Hz, 1H), 1.32 (s, 3H), 1.21 (qd, *J* = 13.2, 4.3 Hz, 1H), 0.49 (s, 2H). ^{13}C NMR (226 MHz, DMSO-d6) data for 11-keto,16α-hydroxyprogesterone were: δ 209.31 (C11), 208.16 (C20), 198.75 (C3), 169.56 (C5), 124.15 (C4), 71.43 (C17), 71.12 (C16), 61.79 (C9), 56.26 (C12), 51.73 (C14), 47.24 (C13), 38.19 (C10), 36.47 (C8), 35.46 (C15), 34.39 (C1), 33.76 (C2), 32.04 (C7), 31.97 (C6), 31.95 (C21), 17.18 (C19), 15.59 (C18).

2α,16α-Dihydroxyprogesterone. ^{1}H NMR (700 MHz, DMSO-d6) δ 5.63 (d, J = 1.3 Hz, 1H), 5.03 (d, J =4.4 Hz, 1H), 4.71 (d, J = 4.9 Hz, 1H), 4.52 (dddd, J = 7.9, 6.3, 4.8, 1.0Hz, 1H), 4.14 (ddd, J = 13.7, 5.3, 4.7 Hz, 1H), 3.32 (s, 1H), 2.42 (d, J =6.5 Hz, 1H), 2.35 (td, J = 13.9, 4.5 Hz, 1H), 2.27 (dt, J = 14.3, 3.8 Hz,1H), 2.11 (dd, J = 12.7, 5.5 Hz, 1H), 2.09 (s, 3H), 1.92 – 1.87 (m, 1H),1.74 (ddt, J = 12.8, 5.7, 2.4 Hz, 2H), 1.60 (td, J = 15.1, 8.8 Hz, 1H), 1.52– 1.45 (m, 2H), 1.47 (dd, J = 12.9, 5.1 Hz, 3H), 1.47 – 1.43 (m, 1H), 1.39– 1.31 (m, 1H), 1.21 (s, 3H), 1.01

– 0.94 (m, 2H), 0.56 (s, 3H). ^{13}C NMR (176 MHz, DMSO-d6) data for 2α,16α-dihydroxyprogesterone δ 208.11 (C20), 199.53 (C3), 170.39 (C5), 121.10 (C4), 72.90 (C17), 70.44 (C16), 68.80 (C2), 53.48 (C9), 52.89 (C14), 44.34 (C13), 44.30 (C1), 40.08 (C10), 37.87 (C12), 35.45 (C15), 33.93 (C8), 31.76 (C21), 31.55 (C7) , 31.49 (C6), 20.09 (C11), 17.55 (C19), 14.22 (C18).

6β,16α-Dihydroxyprogesterone. ^{1}H NMR (700 MHz, DMSO-d6) δ 5.67 (s, 0H), 5.11 (d, J = 2.7 Hz, 0H), 4.73 (dd, J = 4.8, 1.4 Hz, 1H), 4.54 (dt, J = 6.8, 3.4 Hz, 1H), 4.16 (dd, J = 5.4, 2.9 Hz, 0H), 2.49 – 2.41 (m, 1H), 2.44 (d, J = 6.5 Hz, 1H), 2.20 (dt, J = 17.8, 3.1 Hz, 1H), 2.09(s, 1H), 1.94 (ddd, J = 12.9, 4.7, 2.6 Hz, 1H), 1.90 (s, 1H), 1.86 – 1.80 (m, 0H), 1.77 (dt, J = 13.6, 2.8 Hz, 0H), 1.65 – 1.57 (m, 1H), 1.57 (d, J = 8.7 Hz, 0H), 1.55– 1.50 (m, 1H), 1.50 – 1.43 (m, 1H), 1.38 (td, J = 12.9, 3.9 Hz, 0H), 1.27 (s, 1H),1.16 – 1.13 (m, 0H), 0.97 – 0.92 (m, 1H), 0.58 (s, 1H). ^{13}C NMR (176 MHz, DMSO-d6) data for 6β,16α-dihydroxyprogesterone δ 208.17 (C20), 199.31 (C3), 168.94 (C5), 125.23 (C4), 73.01 (C17), 71.01 (C6), 70.47 (C16), 52.96 (C14), 52.90 (C9), 44.41 (C13), 38.66 (C7), 37.96 (C12), 37.62 (C10), 36.52 (C1), 35.58 (C15), 33.90 (C2), 31.78 (C21), 28.96 (C8), 20.21 (C11), 18.95 (C19), 14.25 (C18).

Acknowledgements

This research was supported by a grant (NRF-2016R1D1A3B03933814) of the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology, Republic of Korea. In addition, this work was carried out with the support of “Cooperative Research Program for Agriculture Science & Technology Development (Project title: National Agricultural Genome Program, Project No. PJ01045705)” Rural Development Administration, Republic of Korea. We would like to thank Division of Magnetic Resonance, Korea Basic Science Institute, Ochang, Chungbuk, Korea for NMR analyses.

Keywords: Cytochrome P450, CYP154C8, electron transport partner, steroid hydroxylation, whole-cell biotransformation.

References

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

[2] L. M. Podust, D. H. Sherman, *Nat. Prod. Rep.* **2012**, *29*, 1251–1266.

[3] R. Bernhardt, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 6185–6203.

[4] S. B. Mahato, A. Mukherjee, *J. Steroid Biochem.* **1984**, *21*, 341–342.

[5] K. E. Smith, S. Latif, D. N. Kirk, *J. Steroid Biochem.* **1990**, *35*, 115–120.

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

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

[8] I. K. Jóźwik, F. M. Kiss, Ł. Gricman, A. Abdulmughni, E. Brill, J. Zapp, J. Pleiss, R. Bernhardt, A. W. Thunnissen, *FEBS J.* **2016**, *283*, 4128–4148.

[9] V. Luu-The, *J. Steroid Biochem. Mol. Biol.* **2013**, *137*, 176–182.

[10] W. L. Miller, *Mol. Cell Endocrinol.* **2013**, *371*, 5–14.

[11] M. Bureik, R. Bernhardt, Wiley-VCH Verlag GmbH & Co. KGaA. **2007**, Weinheim, Germany.

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

[13] K. T. Nguyen, C. Virus, N. Günnewich, F. Hannemann, R. Bernhardt, *ChemBioChem* **2012**, *13*, 1161–1166.

[14] G. Y. Lee, D. H. Kim, D. Kim, T. Ahn, C. H. Yun, *Arch. Pharm. Res.* **2015**, *38*, 98–107.

576 [15] V. B. Urlacher, S. Lutz-Wahl, R. D. Schmid, *Appl. Microbiol. Biotechnol.* **2004**, *64*, 317–
577 325.
578 [16] P. Bracco, D. B. Janssen, A. Schallmey, *Microb. Cell Fact.* **2013**, *12*, 95.
579 [17] F. M. Kiss, D. Schmitz, J. Zapp, T. K. Dier, D. A. Volmer, R. Bernhardt, *Appl. Microbiol.*
580 *Biotechnol.* **2015**, *99*, 8495–8514.
581 [18] T. Makino, Y. Katsuyama, T. Otomatsu, N. Misawa, Y. Ohnishi, *Appl. Environ.*
582 *Microbiol.* **2014**, *80*, 1371–1379.
583 [19] Y. Khatri, Y. Carius, M. Ringle, C. R. Lancaster, R. Bernhardt, *FEBS Lett.* **2016**, *590*,
584 4638–4648.
585 [20] Y. Khatri, M. Ringle, M. Lisurek, J. P. von Kries, J. Zapp, R. Bernhardt, *ChemBioChem*
586 **2016**, *17*, 90–101.
587 [21] S. G. Salamanca-Pinzon, Y. Khatri, Y. Carius, L. Keller, R. Müller, C. R. Lancaster, R.
588 Bernhardt, *FEBS Lett.* **2016**, *590*, 1838–1851.
589 [22] L. M. Podust, Y. Kim, M. Arase, B. A. Neely, B. J. Beck, H. Bach, D. H. Sherman, D. C.
590 Lamb, S. L. Kelly, M. R. Waterman, *J. Biol. Chem.* **2003**, *278*, 12214–12221.
591 [23] L. M. Podust, H. Bach, Y. Kim, D. C. Lamb, M. Arase, D. H. Sherman, S. L. Kelly, M. R.
592 Waterman, *Protein Sci*. **2004**, *13*, 255–268.
593 [24] Q. Cheng, D. C. Lamb, S. L. Kelly, L. Lei, F. P. Guengerich, *J. Am. Chem. Soc*. **2010**,
594 *132*, 15173–15175.
595 [25] B. C. Von, P. Le-Huu, V. B. Urlacher, *ChemBioChem* **2013**, *4*, 2189–2198.
596 [26] A. Schallmey, G. den Besten, I. G. Teune, R. F. Kembaren, D. B. Janssen, *Appl.*
597 *Microbiol. Biotechnol*. **2011**, *9*, 1475–1485.
598 [27] A. Rühlmann, G. Groth, V. B. Urlacher, *ChemBioChem* **2018**, *19*, 1–9.

[28] Z. Su, J. H. Horner, M. Newcomb, *Chemistry* **2012**, doi:10.1002/chem.201202254.

[29] J. Schenkman, S. Sligar, D. Cinti, *Parmacol. Ther.* **1981**, *12*, 43–71.

[30] T. T. B. Ly, Y. Khatri, J. Zapp, M. C. Hutter, R. Bernhardt, *Appl. Microbiol. Biotechnol.* **2012**, *95*, 123–133.

[31] K. Herzog, P. Bracco, A. Onoda, T. Hayashi, K. Hoffmann, A. Schallmey, *Acta Crystallogr. D Biol. Crystallography.* **2014**, *70*, 2875–2889.

[32] J. R. Berrie, R. A. Williams, K. E. Smith, *J. Steroid Biochem. Mol. Biol.* **1999**, *71*, 153–165.

[33] B. C. Chung, J. Picado-Leonard, M. Haniu, M. Bienkowski, P. F. Hall, J. E. Shively, W. L. Miller, *Proc. Natl. Acad. Sci.* **1987**, *84*, 407–411.

[34] F. K. Yoshimoto, R. Auchus, *J. Steroid Biochem. Mol. Biol.* **2015**, *151*, 52–65.

[35] P. Swart, A. C. Swart, M. R. Waterman, R. W. Estabrook, J. I. Mason, *J. Clin. Endocrinol. Metab.* **1993**, *77*, 98–102.

[36] F. K. Yoshimoto, E. Gonzalez, R. J. Auchus, F. P. Guengerich, *J. Biol. Chem*. **2016**, *291*, 17143–17164.

[37] R. Davydov, N. Strushkevich, D. Smil, A. Yantsevich, A. Gilep, S. Usanov, B. M. Hoffman, *Biochemistry* **2015**, *54*, 7089–7097.

[38] R. C. Tuckey, *Placenta*. **2005**, *26*, 273–281.

[39] R. C. Tuckey, W. Li, J. K. Zjawiony, M. A. Zmijewski, M. N. Nguyen, T. Sweatman, A. Slominski, *FEBS J.* **2008**, *275*, 2585–2596.

[40] M. J. Cryle, J. J. De Voss, *Chem. Commun (Camb).* **2004**, *1*, 86–87.

[41] Y. Anzai, S. Li, M. R. Chaulagain, K. Kinoshita, F. Kato, J. Montgomery, D. H. Sherman, *Chem. Biol.* **2008**, *15*, 950-959.

[42] W. Zhang, Y. Liu, J. Yan, S. Cao, F. Bai, Y. Yang, S. Huang, L. Yao, Y. Anzai, F. Kato, L. M. Podust, D. H. Sherman, S. Li, *J. Am. Chem. Soc*. **2014**, *136*, 3640–3646.

[43] U. Raju, H. L. Bradlow, M. Levitz, *Ann. N. Y. Acad. Sci*. **1990**, *586*, 83–87.

[44] J. Marcos, W. Y. Craig, G. E. Palomaki, E. M. Kloza, J. E. Haddow, M. Roberson, L. A. Bradley, C. H. Shackleton, *Prenat. Diagn.* **2009**, *29*, 771–780.

[45] A. Yamada, M. Yamada, Y. Fujita, T. Nishigami, K. Nakasho, K. Uematsu, *J. Biol. Chem.* **2001**, *276*, 4604–4610.

[46] F. Zhang, J. Zhou, Y. Shi, P. Tavlarakis, K. Karaisz, *J. Pharm. Biomed. Anal.* **2016**, *128*, 333–341.

[47] M. Li, B. Chen, S. Monteiro, A. M. Rustum, *Tetrahedron Lett.* **2009**, *50*, 4575–4581.

[48] M. Li, B. Chen, M. Lin, T. M. Chan, X. Fu, A. Rustum, *Tetrahedron Lett.* **2007**, *48*, 3901–3905.

[49] J. Hansen, H. Bundgaard, *Int. J. Pharmaceut.* **1980**, *6*, 307–319.

[50] M. Thevis, W. Schänzer, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1660–1669.

[51] H. Li, L. P. Sheng, B. Wang, Z. L. Yang, S. Y. Liu, *J. Chromatogr. Sci.* **2015**, *53*, 285–294.

[52] C. M. Jenkins, M. R. Waterman, *Biochemistry* **1998**, *37*, 6106–6113.

[53] C. M. Jenkins, M. R. Waterman, *J. Biol. Chem.* **1994**, *269*, 27401–27408.

[54] D. B. Hawkes, G. W. Adams, A. L. Burlingame, P. R. Ortiz de Montellano, J. J. De Voss, *J. Biol. Chem.* **2002**, *277*, 27725–27732.

[55] A. Mosa, M. C. Hutter, J. Zapp, R. Bernhardt, F. Hannemann, *ChemBioChem* **2015**, *16*, 1670–1679.

644 [56] W. V. Shaw, L. C. Packman, B. D. Burleigh, A. Dell, H. R. Morris, B. S. Hartley, *Nature*
645 **1979**, *282*, 870–872.
646 [57] N. Saitou, M. Nei, *Mol. Biol. Evol.* **1987**, *4*, 406–425.
647 [58] J. Felsenstein, *Evolution* **1985**, *39*, 783–791.
648 [59] E. Zuckerkandl, L. Pauling, **1965**, pp. 97-166. Academic Press, New York.
649 [60] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, *Mol. Biol. Evol.* **2013**, *30*,
650 2725–2729.
651 [61] J. A. Peterson, M. C. Lorence, B. Amarneh. *J. Biol. Chem.* **1990**, *265*, 6066–6073.
652 [62] S. Bhattarai, K. Liou, T. J. Oh, *Arch. Biochem. Biophys.* **2013**, *539*, 63–69.
653 [63] N. P. Niraula, S. Bhattarai, N. R. Lee, J. K. Sohng, T. J. Oh, *J. Microbiol. Biotechnol.*
654 **2012**, *22*, 1059–1065.
655 [64] T. Omura, R. Sato, *J. Biol. Chem.* **1964**, *239*, 2370–2378.
656 [65] M. M. Purdy, L. S. Koo, P. R. Ortiz de Montellano, J. P. Klinman, *Biochemistry* **2004**, *43*,
657 271–281.
658 [66] P. W. Jr Roome, J. C. Philley, J. A. Peterson, *J. Biol. Chem.* **1983**, *258*, 2593–2598.
659 [67] J. W. Williams, J. F. Morrison, *Methods Enzymol.* **1979**, *63*, 437–467.
660 [68] J. B. Johnston, H. Ouellet, P. R. Ortiz de Montellano, *J. Biol. Chem.* **2010**, *285*, 36352–
661 36360.
662

Figure legends

Figure 1. Phylogenetic tree of all available CYP154 sequences. Alignment and tree building were performed by ClustalW and MEGA6, respectively. CYP154C8 indicated with the closed square.

Figure 2. (**A**) Dithionite-reduced and CO-bound form spectra of CYP154C8. (**B**) SDS-PAGE of purified CYP154C8 (lane 1) and mid-range pre-stained marker (lane M).

Figure 3. HPLC chromatogram of an *in vitro* reaction assay. **A.** Progesterone (I), testosterone (II), nandrolone (III), androstenedione (IV), 11-hydroxyprogesterone (V), 11-deoxycortisol (VI), and 11-ketoprogesterone (VII). **B.** Cortisone (I), corticosterone (II), hydrocortisone (III), prednisone (IV), and prednisolone (V). P, P1, P2, and P3 are product peaks of respective substrates (S). All *in vitro* reactions were carried out in presence of 500 µM substrate.

Figure 4. HPLC chromatogram of 16α-hydroxyprogesterone *in vitro* reaction assay. S, P1, and P2 indicate the peaks of substrate, product 1, and product 2, respectively.

Figure 5. Hyperbolic fit of 16α-hydroxyprogesterone. Values indicate the mean of three independent experiments with standard deviations.

Figure 6. Bioconversion of steroid substrates into hydroxylated products by *E. coli* cells harboring the recombinant plasmids; pET32a_cyp154C8 & pAYAC_camAB, and pET32a_cyp154C8 & pCDF_camAB, separately. Bioconversion were carried out in 1 ml scale and substrate concentration of 1 mM. Values indicate the means from four independent experiments with standard deviations. The X-axis and Y-axis show the percentage of product formed and steroid substrate, respectively.

Scheme legends

Scheme 1. Steroids used to screen CYP154C8 in this study.

Scheme 2. Steroid hydroxylation by CYP154C8. **A.** Progesterone is hydroxylated at 16α position and subsequently hydroxylated at 2α and 6β position. **B, C,** and **D.** Androstenedione, testosterone, and 11-ketoprogesterone are hydroxylated at 16α position. **E.** Corticosterone is hydroxylated at 21 position (major product, P2).

10.1002/cbic.201800018

693 **TABLE 1.** Dissociation constant (K_d) values of CYP154C8 with steroid substrates.
694 Corticosterone did not show any high spin shift upon binding to CYP154C8. ND indicates not
695 detect.

Substrate	K_d (μM)	Induction of high spin shift
Progesterone	0.083 ± 0.030	+
Testosterone	0.162 ± 0.040	+
Androstenedione	0.216 ± 0.050	+
Nandrolone	0.181 ± 0.030	+
Corticosterone	ND	–
Prednisone	2.500 ± 0.350	+
Prednisolone	12.780 ± 1.700	+
Hydrocortisone	15.800 ± 0.560	+
Cortisone	4.160 ± 0.310	+
16α-hydroxyprogesterone	1.210 ± 0.930	+

696

Text for the table of content graphic

An insight into CYP154: Oxyfunctionalization of steroids by CYP154C8 is novel and interesting. Multistep hydroxylation of progesterone and 21-hydroxylation of corticosterone are special characteristic of CYP154C8 in CYP154 family. The generation of 2α,16α-dihydroxyprogesterone, 6β,16α-dihydroxyprogesterone, and 21-hydroxycorticosterone metabolites catalyzed by CYP154C8 is the first report by enzymatic method.