Supporting Information

Rationally engineering an H_2O_2-dependent P450 dihydroxylase for steroid functionalisation

Quanjin Wang,[a†] Mingming Qin,[b†] Qian Wang,[b†‡] Kaiming Wang,[a] and Zhiqi Cong*[b]

[a] University of Jinan, School of Biological Science and Technology, Jinan 250024, China.

[b] Key Laboratory of Photoelectric Conversion and Utilization of Solar Energy, Qingdao New Energy Shandong Laboratory, CAS Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Synthetic Biology, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266101, China.

† These authors contributed equally to this work.

‡ Current address: Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266000, China.

* Corresponding to congzq@qibebt.ac.cn

Experimental Section

Materials

All chemical reagents were purchased from commercial sources (e.g. Aldrich, TCI, Fluka, and Alladin) and used without further purification until otherwise noticed.

Expression and Purification of CYP105D18

The pET-28a (+) vectors containing CYP105D18 and its variants were transformed into *Escherichia coli* BL 21(DE3) cells, and the cells were cultivated in LB medium containing 50 μg/ml kanamycin. The cultures were grown at 37 ºC with vigorous shaking (~200 rpm). When the OD_{600} of the cultures reached 0.8~1.0, the temperature was cooled to 20 ºC, and the expression was induced by the addition of IPTG (1 mM) and δ-aminolevulinic acid hydrochloride (0.5 mM). Following 36 h of expression, the cells were harvested by centrifugation and stored at -20 ºC. Purification was done by Ni-NTA metal-affinity chromatography. Cell pellets were resuspended in ice-cold buffer A (100 mM KP_i, 100 mM NaCl, imidazole (20 mM), pH 7.4) and lysed by sonication. Cell debris was removed by centrifugation for 30 min at 20 000 g, and the crude cell extraction were applied to a 5 mL bed volume column pre-equilibrated with buffer A. Nonspecifically bound proteins were washed from the column with 5 column volumes of buffer A containing 30 mM imidazole. The bound protein was eluted with buffer B (100 mM KPi, 100 mM NaCl, imidazole (200 mM), pH 7.4). The purified protein solution was exchanged with buffer C (100 mM KPi, 100 mM NaCl, pH 7.4), enzyme its variants concentrated by ultrafiltration and frozen in buffer C plus 50% glycerol at -20 ºC. The concentrations of CYP105D18 and its variants were measured by Hemochrome binding assay. A pyridine solution was made by combining pyridine (1.75 mL) and 1 M aqueous of NaOH (0.75 mL). The solution was mixed at room temperature then centrifuged for 30 s at 5000 rpm to remove excess aqueous base. To a cuvette containing 0.75 mL of protein solution in phosphate buffer (0.1 M, pH 8.0), 0.25 mL of the pyridine solution was added followed by 2 mg of sodium dithionite. A UV/vis spectrum was recorded immediately. Hemoprotein concentration was determined from the absorbance of the hemochrome complex using extinction

coefficients of ε418 = 196 $mM^{-1}cm^{-1}$. Absorbance was assigned as the difference between the peak max at 418 nm and the baseline at 420 nm as determined by extrapolating from two points on either side of the hemochrome peak (390 nm and 450 nm).

Mutagenesis

All the mutations were made by PCR based site-directed mutagenesis and verified by DNA sequencing. The primers used were as follows:

Mutation	Sequence
E354A-F	CACGTATTgcaCTGGAAATTGCCCTGGGTACC
E354A-R	TTCCAGtgcAATACGTGCCAGATTCTGGCCCA
F105A-F	ATTCCGAGTgcaACCCTGCGTCGTGCAGAAAG
F105A-R	AGGGTtgcACTCGGAATCAGCATGCGACGCTG
H343A-F	TGGCATTgcaCAGTGTCTGGGCCAGAATCTGG
H343A-R	GACACTGtgcAATGCCAAAACCAAAGGCAACA
L350A-F	CCAGAATgcaGCACGTATTGAACTGGAAATTGC
L350A-R	TACGTGCtgcATTCTGGCCCAGACACTGATGA
G347A-F	GTGTCTGgcaCAGAATCTGGCACGTATTGAACTG
G347A-R	GATTCTGtgcCAGACACTGATGAATGCCAAAACC
F338A-F	TGTTGCCgcaGGTTTTGGCATTCATCAGTGTCT
F338A-R	CAAAACCtgcGGCAACATGATGGCGTGCGCTG
L285A-F	ATTGCCGATGGCgcaGTGCGCGTGGCACGTGAA
L285A-R	ACtgcGCCATCGGCAATACTCAGAAAACGCAG
R287A-F	TGgcaGTGGCACGTGAAGATGTGCCGGTGGGC
R287A-R	TTCACGTGCCACtgcCACCAGGCCATCGGCAAT
I195A-F	CGAACTGgcaGAAACCAAACGTACCGATCCGG
I195A-R	TGGTTTCtgcCAGTTCGCCAAAATAGCCTTCC
I150A-F	TATGGTTgcaTGCGAACTGCTGGGTGTGCCGT
I150A-R	GTTCGCAtgcAACCATACTCGGAACCGGCAGG
V234A-F	TGgcaGCAGGTCATGAAACCACCGCCAATATG
V234A-R	TTCATGACCTGCtgcCAGCAGAACGGTGGCAAACT
H161A-F	GTATGGCGATgcaGAATTTTTCGAAGAACAGAGCCG
H161A-R	ATTCtgcATCGCCATACGGCACACCCAGCAGT
L232A-F	TTTGCCACCGTTgcaCTGGTGGCAGGTCATGAAACC

L232A-R	AGtgcAACGGTGGCAAACTGCACCAGTTCATC
L187A-F	TTTCGTAGCgcaGAAGGCTATTTTGGCGAACTGA
L187A-R	CCTTCtgcGCTACGAAATGCTTTTTCAATATCTT
L346A-F	AGTGTgcaGGCCAGAATCTGGCACGTATTGAA
L346A-R	ATTCTGGCCtgcACACTGATGAATGCCAAAACCA
Q348A-F	TGTCTGGGCgcaAATCTGGCACGTATTGAACTGGA
Q348A-R	AGATTtgcGCCCAGACACTGATGAATGCCAAA
D160A-F	GTATGGCgcaCATGAATTTTTCGAAGAACAGAGC
D160A-R	ATTCATGtgcGCCATACGGCACACCCAGCAGT
L153A-F	TTTGCGAAgcaCTGGGTGTGCCGTATGGCGAT
L153A-R	ACCCAGtgcTTCGCAAATAACCATACTCGGAA
Y190A-F	AAGGCgcaTTTGGCGAACTGATTGAAACCAAA
Y190A-R	TTCGCCAAAtgcGCCTTCCAGGCTACGAAATGC
I281A-F	TTCTGAGTgcaGCCGATGGCCTGGTGCGCGTG
I281A-R	ATCGGCtgcACTCAGAAAACGCAGCAGTTCTT
L194A-F	GGCGAAgcaATTGAAACCAAACGTACCGATCC
L194A-R	GTTTCAATtgcTTCGCCAAAATAGCCTTCCAG
L359A-F	AAATTGCCgcaGGTACCCTGCTGCGCCGTCTG
L359A-R	GGTACCtgcGGCAATTTCCAGTTCAATACGTG
H161A-F	GTATGGCGATgcaGAATTTTTCGAAGAACAGAGCCG
H161A-R	ATTCtgcATCGCCATACGGCACACCCAGCAGT
N349A-F	AGgcaCTGGCACGTATTGAACTGGAAATTGCC
N349A-R	AATACGTGCCAGtgcCTGGCCCAGACACTGATGAATG
V286A-F	TTGCCGATGGCCTGgcaCGCGTGGCACGTGAAGAT
V286A-R	tgcCAGGCCATCGGCAATACTCAGAAAACGCA
T248A-F	TCTGGCAgcaTATACCCTGCTGGAACATCCGG
T248A-R	GGGTATAtgcTGCCAGACTAATCATATTGGCG
P202A-F	TACCGATgcaGGTGAAGGTGTTATTGATGATCTGG
P202A-R	CTTCACCtgcATCGGTACGTTTGGTTTCAATCA
F340A-F	CTTTGGTgcaGGCATTCATCAGTGTCTGGGCC
F340A-R	GAATGCCtgcACCAAAGGCAACATGATGGCGT
D208A-F	TATTgcaGATCTGGTGGCCCGTCAGCGTGAAG
D208A-R	CCACCAGATCtgcAATAACACCTTCACCCGGATCG
T197A-F	TGAAgcaAAACGTACCGATCCGGGTGAAGGTG
T197A-R	CGGTACGTTTtgcTTCAATCAGTTCGCCAAAATAGC

F164A-F	TTgcaGAAGAACAGAGCCGTCGCCTGCTGCGT
F164A-R	GCTCTGTTCTTCtgcAAATTCATGATCGCCATACGG
T200A-F	ACCAAACGTgcaGATCCGGGTGAAGGTGTTATTG
T200A-R	GGATCtgcACGTTTGGTTTCAATCAGTTCGCC
K198A-F	TTGAAACCgcaCGTACCGATCCGGGTGAAGGT
K198A-R	GGTACGtgcGGTTTCAATCAGTTCGCCAAAAT
E196A-F	CTGATTgcaACCAAACGTACCGATCCGGGTGA
E196A-R	CGTTTGGTtgcAATCAGTTCGCCAAAATAGCCT
L154A-F	TTTGCGAACTGgcaGGTGTGCCGTATGGCGATC
L154A-R	ACCtgcCAGTTCGCAAATAACCATACTCGGAA
K199A-F	AACCAAAgcaACCGATCCGGGTGAAGGTGTTA
K199A-R	GATCGGTtgcTTTGGTTTCAATCAGTTCGCCA
I353A-F	CACGTgcaGAACTGGAAATTGCCCTGGGTACC
I353A-R	TTCCAGTTCtgcACGTGCCAGATTCTGGCCCA
V211A-F	TTGATGATCTGgcaGCCCGTCAGCGTGAAGAAG
V211A-R	GGCtgcCAGATCATCAATAACACCTTCACCCG
L233A-F	TTCTGgcaGTGGCAGGTCATGAAACCACCGCC
L233A-R	ACCTGCCACtgcCAGAACGGTGGCAAACTGCA
S104A-F	GATTCCGgcaTTTACCCTGCGTCGTGCAGAAA
S104A-R	GGGTAAAtgcCGGAATCAGCATGCGACGCTGT
F191A-F	TATgccGGCGAACTGATTGAAACCAAACGTAC
F191A-R	ATCAGTTCGCCggcATAGCCTTCCAGGCTACGAAATG
L154G-F	TTTGCGAACTGggtGGTGTGCCGTATGGCGATC
L154G-R	ACCaccCAGTTCGCAAATAACCATACTCGGAA
L154S-F	TTTGCGAACTGagcGGTGTGCCGTATGGCGATC
L154S-R	ACCgctCAGTTCGCAAATAACCATACTCGGAA
L154T-F	TTTGCGAACTGaccGGTGTGCCGTATGGCGATC
L154T-R	ACCggtCAGTTCGCAAATAACCATACTCGGAA
F164G-F	TTggtGAAGAACAGAGCCGTCGCCTGCTGCGT
F164G-R	GCTCTGTTCTTCaccAAATTCATGATCGCCATACGG
F164S-F	TTagcGAAGAACAGAGCCGTCGCCTGCTGCGT
F164S-R	GCTCTGTTCTTCgctAAATTCATGATCGCCATACGG
F164T-F	TTaccGAAGAACAGAGCCGTCGCCTGCTGCGT
F164T-R	GCTCTGTTCTTCggtAAATTCATGATCGCCATACGG
T200G-F	ACCAAACGTggtGATCCGGGTGAAGGTGTTATTG

T200G-R	GGATCaccACGTTTGGTTTCAATCAGTTCGCC
T200S-F	ACCAAACGTagcGATCCGGGTGAAGGTGTTATTG
T200S-R	GGATCgctACGTTTGGTTTCAATCAGTTCGCC
V211G-F	TTGATGATCTGggtGCCCGTCAGCGTGAAGAAG
V211G-R	GGCaccCAGATCATCAATAACACCTTCACCCG
V211S-F	TTGATGATCTGagcGCCCGTCAGCGTGAAGAAG
V211S-R	GGCgctCAGATCATCAATAACACCTTCACCCG
V211T-F	TTGATGATCTGaccGCCCGTCAGCGTGAAGAAG
V211T-R	GGCggtCAGATCATCAATAACACCTTCACCCG
F184G-F	AAAAGCAggtCGTAGCCTGGAAGGCTATTTTG
F184G-R	GGCTACGaccTGCTTTTTCAATATCTTCTGCTGC
F184S-F	AAAAGCAagcCGTAGCCTGGAAGGCTATTTTG
F184S-R	GGCTACGgctTGCTTTTTCAATATCTTCTGCTGC
F184T-F	AAAAGCAaccCGTAGCCTGGAAGGCTATTTTG
F184T-R	GGCTACGggtTGCTTTTTCAATATCTTCTGCTGC
L233G-F	TTCTGggtGTGGCAGGTCATGAAACCACCGCC
L233G-R	ACCTGCCACaccCAGAACGGTGGCAAACTGCA
L233S-F	TTCTGagcGTGGCAGGTCATGAAACCACCGCC
L233S-R	ACCTGCCACgctCAGAACGGTGGCAAACTGCA
L233T-F	TTCTGaccGTGGCAGGTCATGAAACCACCGCC
L233T-R	ACCTGCCACggtCAGAACGGTGGCAAACTGCA
R199G-F	AACCAAAggtACCGATCCGGGTGAAGGTGTTA
R199G-R	GATCGGTaccTTTGGTTTCAATCAGTTCGCCA
R199S-F	AACCAAAagcACCGATCCGGGTGAAGGTGTTA
R199S-R	GATCGGTgctTTTGGTTTCAATCAGTTCGCCA
R199T-F	AACCAAAaccACCGATCCGGGTGAAGGTGTTA
R199T-R	GATCGGTggtTTTGGTTTCAATCAGTTCGCCA

The double mutants were prepared used single mutants as parent templates. Prepare triple mutants using double mutants as parent templates.

Water access tunnel analysis

The wild type CYP105D18 (PDB ID: 7DI3) was used for water access tunnel analysis. The water access tunnels were analyzed by CAVER 3.0.3 tool with following parameters: 0.9 Å of probe radius; 3 Å of shell radius; 4 Å of shell depth; 3.5 of clustering threshold; the start point was the heme cofactor.

General procedure for H_2O_2-dependent hydroxylation of steroids catalyzed by CYP105D18

The CYP105D18 and its variants (1 μM) were transferred to a glass flask containing 0.1 M, pH 7.4 phosphate buffer, steroids (0.2-1 mM, 2% DMSO). The reaction mixture was incubated in water bath at 30 ºC for 2 min and initiated by the addition of H_2O_2 (10-120 mM, dissolved 7.4 phosphate buffer) in water bath at 30 ºC for 60 min. After the reaction is complete, samples (0.5 mL) were extracted using 0.5 mL dichloromethane. The organic phases were dried over anhydrous sodium sulfate. After complete evaporation of the dichloromethane, the residue was dissolved in 0.5mL acetonitrile and analyzed using high-performance liquid chromatography (HPLC). The conversion rate of substrate was calculated based on the unreacted steroids. The turnover number (TON) is defined as the ratio of product concentration to enzyme concentration.

General procedure for calculating TON of testosteroneby H_2O_2-dependent CYP105D18

The CYP105D18 and its variants (1 μM) were transferred to a glass flask containing 0.1 M, pH 7.4 phosphate buffer, Substrates (0.2mM, 2% DMSO). The reaction mixture was incubated in water bath at 30 ºC for 2 min and initiated by the addition of H_2O_2 (10-120 mM, dissolved pH 7.4 phosphate buffer) in water bath at 30 ºC for 60 min. After the reaction is complete, samples (0.5 mL) were extracted using 0.5 mL dichloromethane. The organic phases were dried over anhydrous sodium sulfate. After complete evaporation of the dichloromethane, the residue was dissolved in 0.5 mL acetonitrile. Subsequently, the conversion rate was detected and calculated using high-performance liquid chromatography (HPLC).

Determination of kinetic parameters

To determine the relevant kinetic parameters of testosterone, the initial rate of substrate oxidation was measured in 100 mM PBS buffer (pH 7.4) containing 1 μ M enzyme, 60 mM H_2O_2, and 10-200 μ M testosterone.

MD simulation method

All models were built using Gromacs 2023 with periodic boundary conditions.[1] The protonation of enzymes was calculated at pH 7.4 using the online server H++.[2, 3]The AMBER-ff19SB force field[4] was used for enzymes. Structure optimization and frequency analysis of small molecules were performed using ORCA 5.0.4.[5], the force field parameters of which were then obtained, and topology files were generated using both Multiwfn[6] and AmberTools.[7, 8] The TIP3P water model was used to construct a cubic water box to solvate the systems, and ions were added to maintain electrical neutrality.

All simulations were performed using the GPU version of Gromacs 2023, and pre-equilibrium simulations preceded production phase simulations. In the pre-equilibrium stage, the steepest descent method was used to minimize energy. Under a constant pressure of 1 bar, the temperature was gradually heated from 0 K to 298 K in 1000 ps and given sufficient time for relaxation. The equilibrium was ensured by monitoring the potential energy, temperature and density convergence of the system at each stage. The SETTLE algorithm[9] was used to constrain the structure of water, while the LINCS algorithm[10] was used to constrain the bonds connected to H atoms. The Particle Mesh Ewald (PME) algorithm[11] was used to estimate the long-range electrostatic interactions. The V-rescale algorithm was used to couple the temperature of the system, and the Berendsen algorithm[12] and Parrinello-Rahman algorithm[13] were used to couple the pressure of the system at NPT and the production phase, respectively.

Crystallization of CYP105D18 F191A and F191A/F184A mutants.

Before crystallization, the buffer of CYP105D18 mutants was changed into 10mM Tris pH 8.0 by superdex increase 10/300 GL column. The crystallization experiment of F191A was performed at 20 °C using a hanging drop evaporation method by mixing 1 μL of the protein solution (40 mg/mL protein supplemented with 1.25μM testosterone dissolved by DMSO) and 1 μL of a reservoir solution (0.15M $(NH_4)_2SO_4$, 0.1M Bis-tris PH 6.5, 23% PEG3350). Dimond-like crystals were obtained within 2 weeks and soaked in a cryoprotecting solution containing 0.15M $(NH_4)_2SO_4$, 0.1M Bis-tris PH

6.5, 23% PEG3350, and 20% glycerol for 3 min before they were transferred into liquid nitrogen for data collection. As for F191A/F184A double mutant， the crystallization experiment was performed similar at 20 °C by hanging drop evaporation method by mixing 1 μL of the protein solution (50 mg/mL protein supplemented with 1.25μM testosterone dissolved by DMSO) and 1 μL of a reservoir solution (0.2 M Li_2SO_4, 0.1 M Bis-Tris, pH 6.5, and 23 %(w/v) PEG 3350). Single crystal was obtained within few days and soaked in a cryoprotecting solution for later X-ray diffraction.

Data collection and structure determination

The X-ray diffraction data of CYP105D18 mutants was collected from Shanghai Radiation Facility (SSRF) BL10U2 beamline by the EIGER2 X 16M detector under 100K. The collected dataset was first processed by Aquarium[14] and autoPX[15] for reflection data integration and reduction. Then the crystal structure was solved by molecular replacement (MR) strategy by Phenix[16] Phaser-MR using the template structure (PDB ID:7DI3). The structure model was then built by COOT[17]and future refined by Phinex. Finalized structure quality was evaluated by MolProbility[18] and deposited to Protein Data Bank (PDB ID: 9U7W and 9U7X). The details of the structures were summarized in Table S1. Figures of the structure were generated PyMol(https://www.schrodinger.com).

Table S1 Data collection and refinement statistics. [a]

Protein	CYP105D18 F191A	CYP105D18 F184A/F191A
PDB Entry	**9U7X**	**9U7W**
Data collection		
Space group	*C 1 2 1*	*C 1 2 1*
Unit cell dimensions		
a, b, c (Å)	91.154, 52.109, 87.412	91.28, 52.342, 87.673
α,β,γ (°)	90, 110.72, 90	90, 110.684, 90
Multiplicity	5.7 (3.3)	3.0 (2.0)
Completeness (%)	99.29 (94.45)	98.51 (93.49)
I/σ_I	11.85 (4.31)	11.61 (1.82)
R_{merge}	0.0862 (0.2671)	0.0551 (0.4149)
Wilson B factor (Å^2)	9.48	11.23
$CC_{1/2}$	0.989 (0.91)	0.983 (0.698)
Refinement		
Resolution range (Å)	40.88 - 1.45 (1.502 - 1.45)	36.77 - 1.09 (1.129 - 1.09)
NO. reflections	67656 (6412)	158630 (14920)
R_{work}/R_{free}	0.1510/0.1660	0.1523/0.1692
Number of non-hydrogen atoms	3656	3694
Macromolecules	3057	3024
Ligands/ions	85	91
Solvent	514	579
Average B-factor	19.40	21.16
Macromolecules	17.35	18.45
Ligands	18.78	25.42
Solvent	31.69	34.66
r.m.s.d. for ideal value		
Bond length (Å)	0.007	0.016
Bond angle (°)	0.98	1.70
Ramachandran plot		
Favored/allowed/outliers (%)	97.92/2.08/0	98.43/1.57/0

[a] Values for the highest resolution shell are given in parentheses.

Instruments and analytical conditions

HPLC

Dissolve the dry residue of biotransformation containing testosterone in acetonitrile and inject it into a high-performance liquid chromatography (HPLC) system equipped with Shim pack GIST C18-AQ, 5um, 4.6 x 250 column (Shimadzu, Japan) at 30 ℃ using Beckmann Coulter. Acetonitrile: Water (60:40) was used as the mobile phase, with a flow rate of 1 mL/min. Detect substrates and their respective products by UV absorption at 243nm. The substrate was eluted at 7.15 minutes, while the product of 2β-hydroxytestosterone was detected at 4.66 minutes. The products of 16α-hydroxytestosterone were eluted in 3.29 minutes. The products of 2β,15α-dihydroxytestosterone were eluted in 2.9 minutes.

LC-MS

The preliminary product identification was carried out using a liquid chromatography-Q-TOF high-resolution mass spectrometer (Bruke Maxis UHR TOF), and Acetylotrile: Water (60:40) was used as the mobile phase, with a flow rate of 1 mL/min.

Preparation, separation, and purification of 2β-hydroxytestosterone and 2β,15α-dihydroxytestosterone

The preparation of 2β-hydroxytestosterone was respectively carried out in 500mL pH 6.0 PBS buffer at 30 ℃. The initial reaction includes 0.2mM substrate (2% Ethyl alcohol solubilized), 20mM H_2O_2, and 0.5 μM F184A/F191A/E196A. After the reaction is complete, extract three times with equal volumes of dichloromethane, combine the organic phases, dry with anhydrous sodium sulfate, and remove the solvent under reduced pressure. Hydroxylation products were separated using silica gel column chromatography (200 - 300 mesh) with ethyl acetate/hexane (5:1) as the eluent. The organic phase was concentrated under reduced pressure to remove the solvent, yielding a white solid product: 7 mg, 2 β -hydroxytestosterone, 24.3% yield.

The preparation of 2β,15α-dihydroxytestosterone was respectively carried out in 500mL pH 6.0 PBS buffer at 30 ℃. The initial reaction includes 26.2mg 2β-hydroxytestosterone (2% Ethyl alcohol solubilized), 60mM H_2O_2, and 6 μM F184A/F191A/E196A. After the reaction is complete, extract three times with equal volumes of dichloromethane, combine the organic phases, dry with anhydrous sodium sulfate, and remove the solvent under reduced pressure. 2β,15α-dihydroxytestosterone products were separated using silica gel column chromatography (200 - 300 mesh) with ethyl acetate/carbinol/hexane (5:1:3) as the eluent. The organic phase was concentrated under reduced pressure to remove the solvent, yielding a white solid product:, 16.2 mg, 2β,15α-dihydroxytestosterone, 61.8% yield.

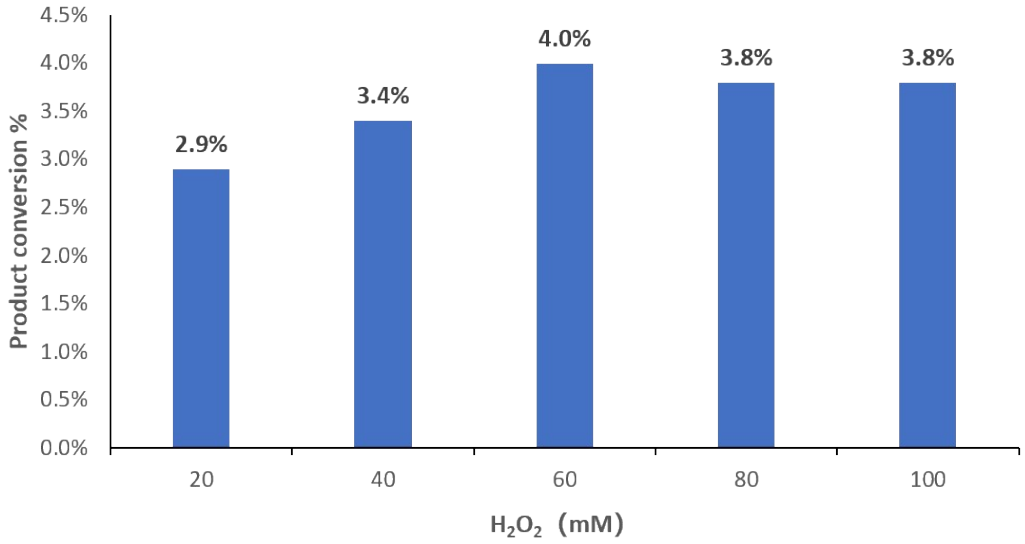


Figure S1. CYP105D18 heme domain using H_2O_2 concentrations ranging from 20–100 mM. The reaction was carried out in 500 µL of 100 mM PBS (pH 7.4) containing 1 µM enzyme and 200 µM testosterone, with different concentrations of H_2O_2(20-100mM), at 30°C for 1 hour.

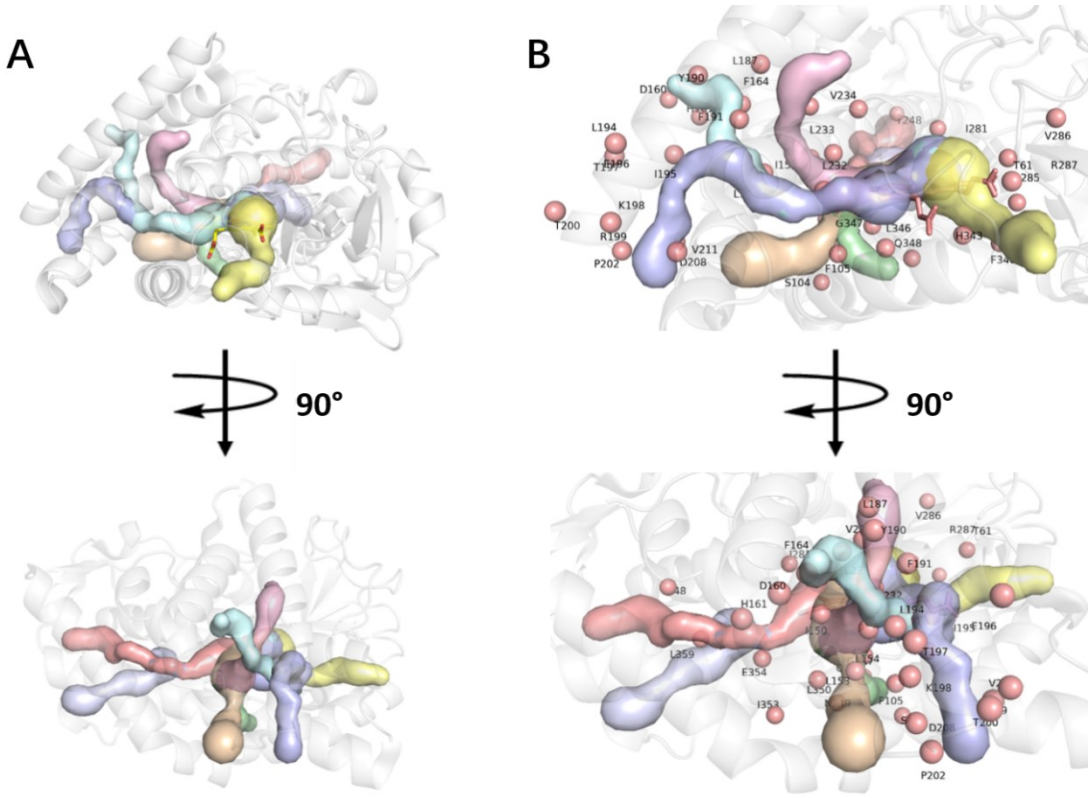


Figure S2. (A) All water tunnels identified in the crystal structure of CYP105D18 WT. (B) Key amino acid sites.

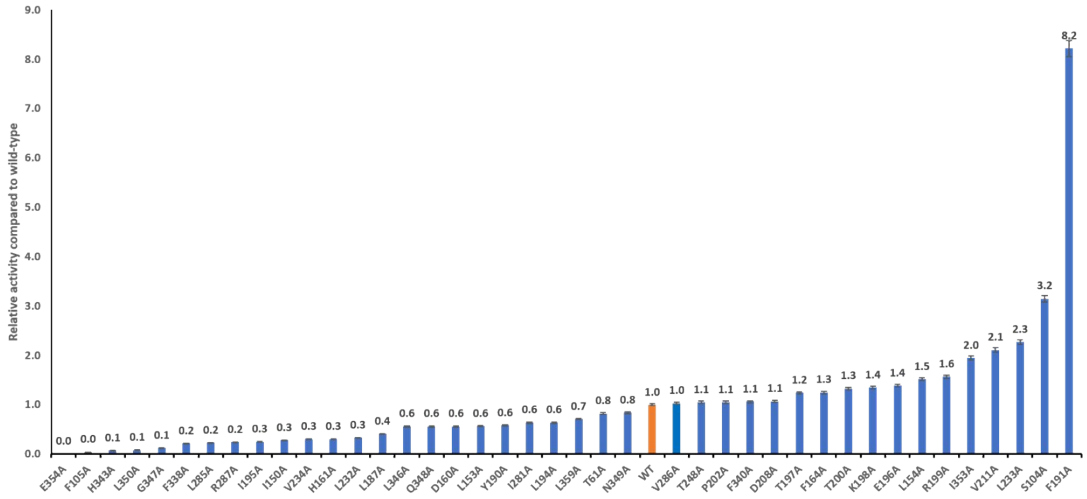


Figure S3. First - round HTE alanine scanning relative to WT activity.The reaction was carried out in 500 μL of 100 mM PBS (pH 7.4) containing 1 μM enzyme and 200 μM testosterone, with 60 mM H_2O_2, at 30°C for 1 hour.

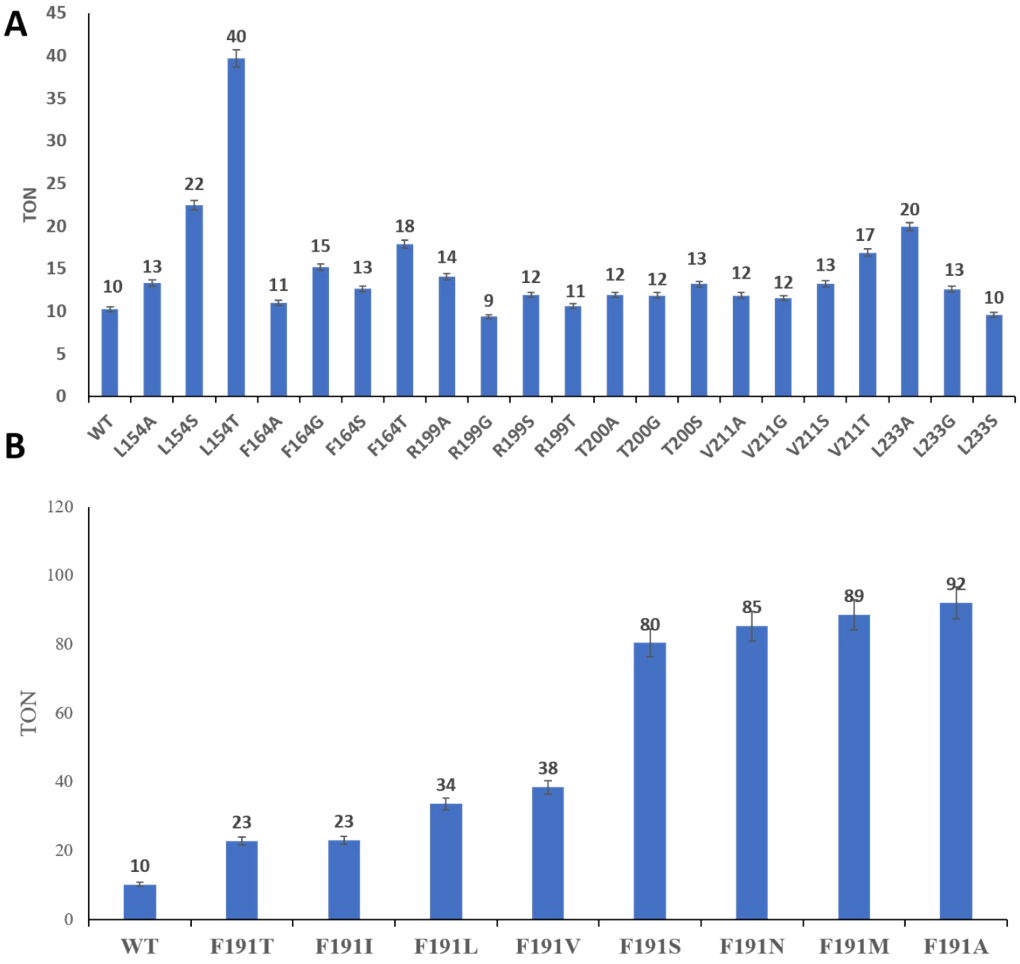


Figure S4. The TON of superior mutation site was further screened. The reaction was carried out in 500 μL of 100 mM PBS (pH 7.4) containing 1 μM enzyme and 200 μM testosterone, with 60 mM H_2O_2, at 30°C for 1 hour.

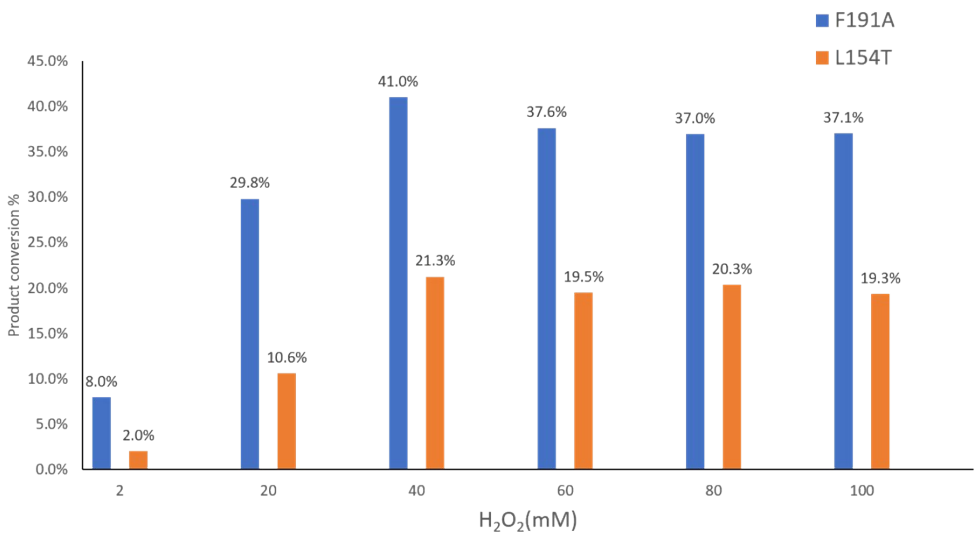


Figure S5. F191A and L154T heme domain using H_2O_2 concentrations ranging from 2–100 mM.The reaction was carried out in 500 µL of 100 mM PBS (pH 7.4) containing 1 µM enzyme and 200 µM testosterone, with different concentrations of H_2O_2(2-100mM), at 30°C for 1 hour.

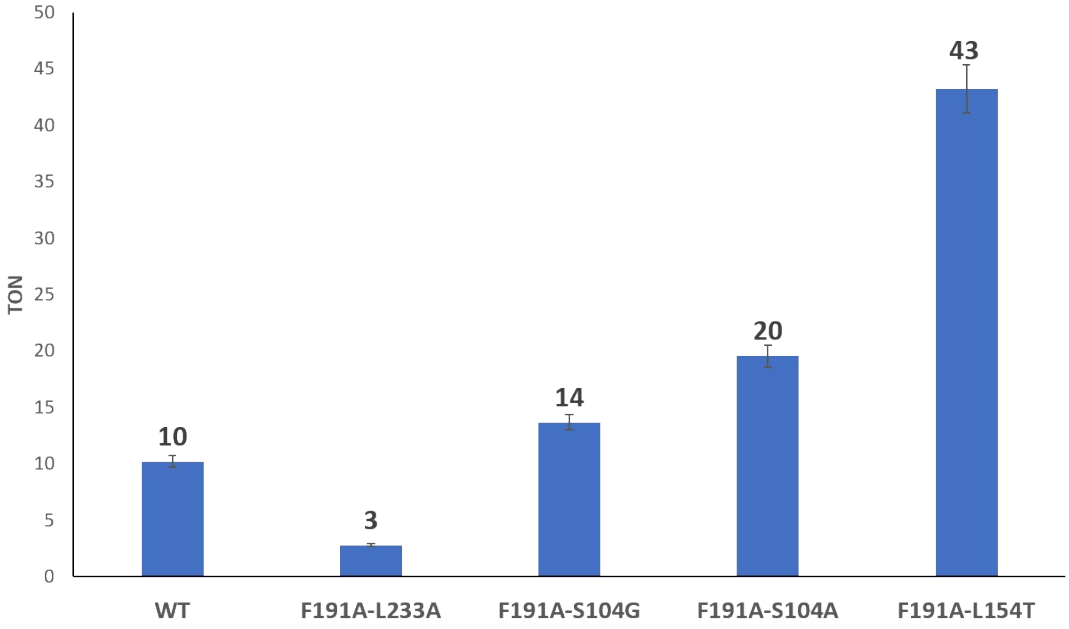


Figure S6. The TON of double mutant. The reaction was carried out in 500 μL of 100 mM PBS (pH 7.4) containing 1 μM enzyme and 200 μM testosterone, with 60 mM H_2O_2, at 30°C for 1 hour.

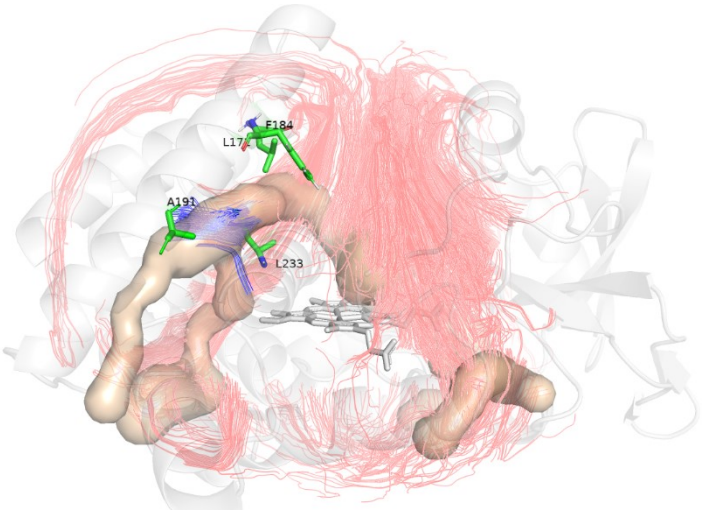


Figure S7. MD simulation results of 191A mutant and prediction results of CAVER 3.0.3 water tunnel.

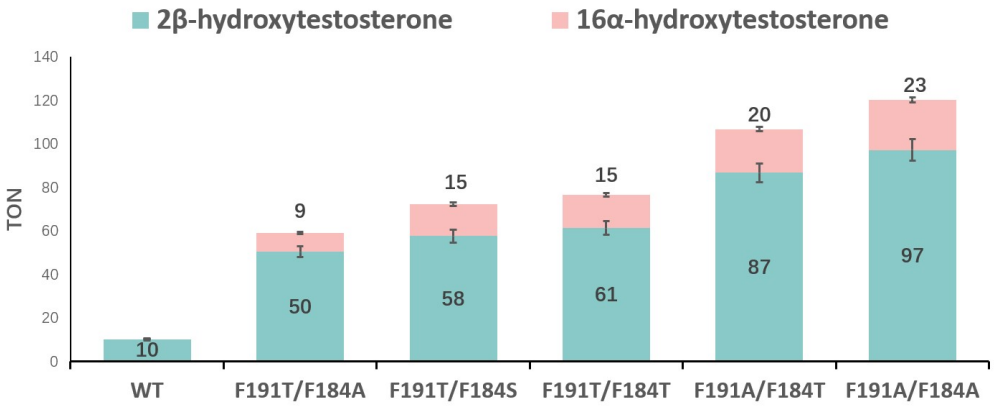


Figure S8. F184 and F191 combined double mutation.The reaction was carried out in 500 μL of 100 mM PBS (pH 7.4) containing 1 μM enzyme and 200 μM testosterone, with 60 mM H_2O_2, at 30°C for 1 hour.

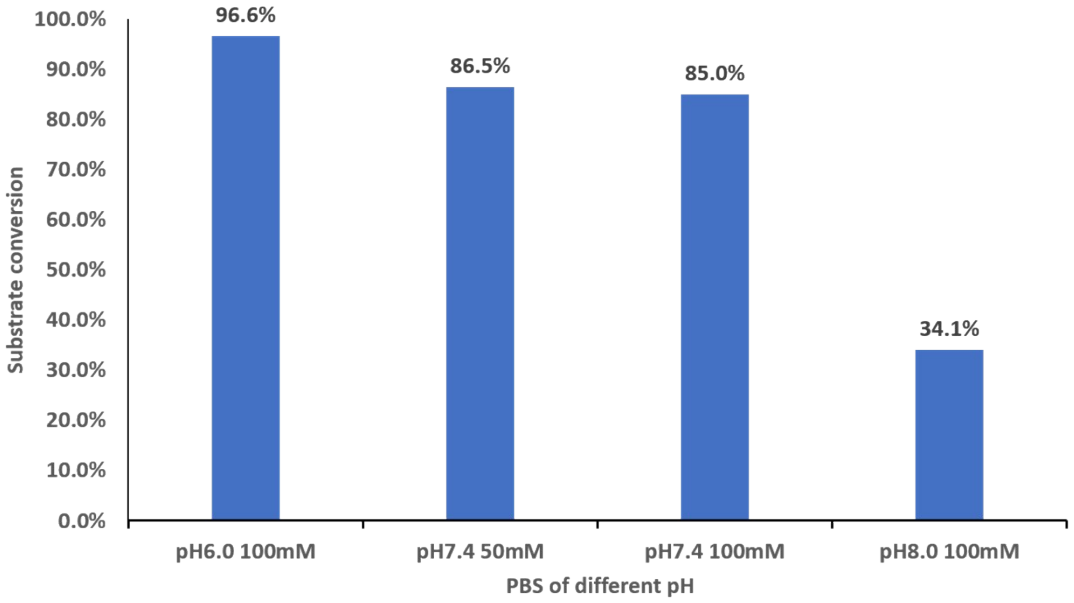


Figure S9. Optimization of reaction conditions for AAA mutant. The reaction was carried out in 500 μL of PBS with different pH values, containing 1 μM AAA and 200 μM testosterone, with 60 mM H_2O_2, at 30°C for 1 hour.

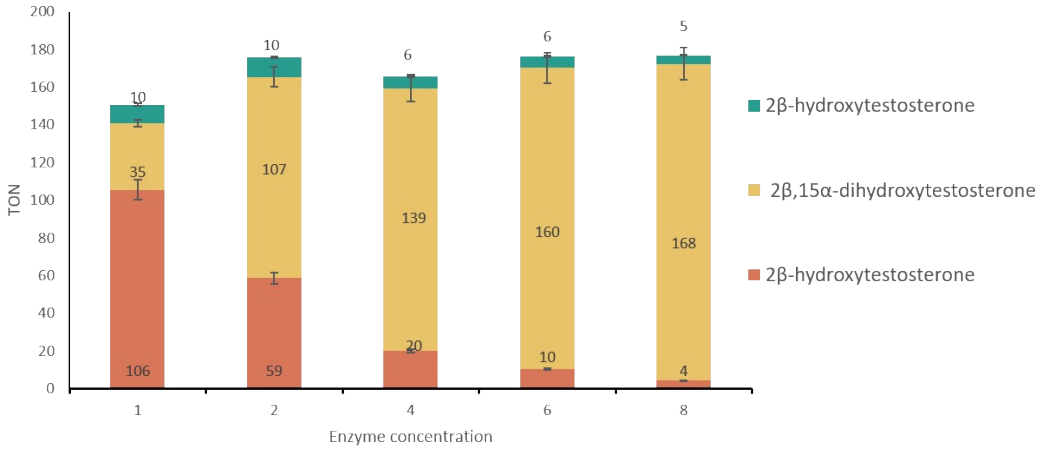


Figure S10. Conversion of 2β-hydroxytestosterone to 2β, 15α-dihydroxytestosterone.The reaction was carried out in 500 μL of 100 mM PBS (pH 6.0), containing 200 μM testosterone and 60 mM H_2O_2, with different concentrations of AAA (1–8 μM), at 30°C for 1 hour.

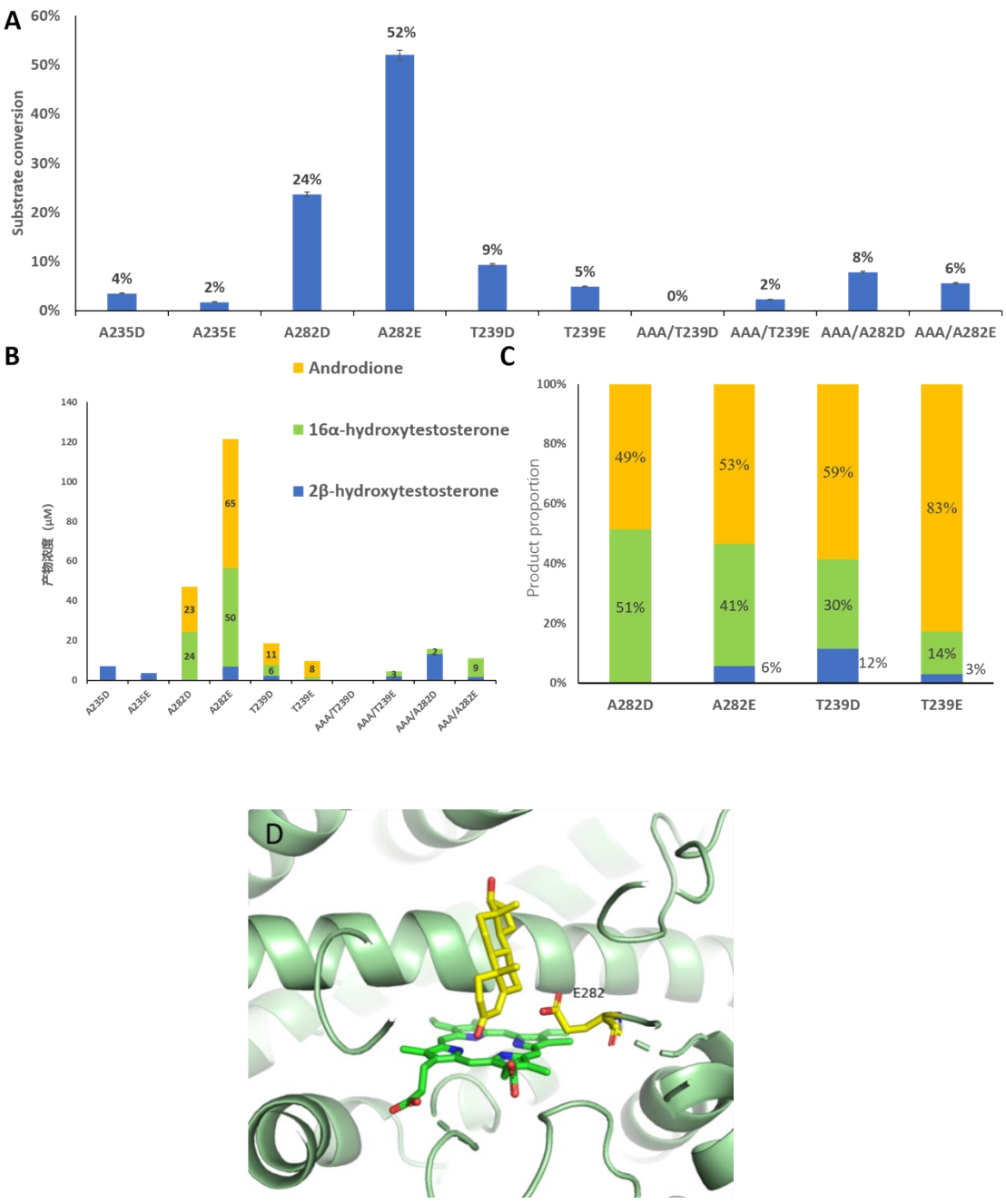


Figure S11. Catalytic residue substrate conversion rate(A) and product distributionand(B-C). The reaction was carried out in 500 μL of 100 mM PBS (pH 6.0), containing 200 μM testosterone, 60 mM H_2O_2, and 1 μM enzyme, at 30°C for 1 hour. (D) Molecular docking of the A282E mutant with testosterone.

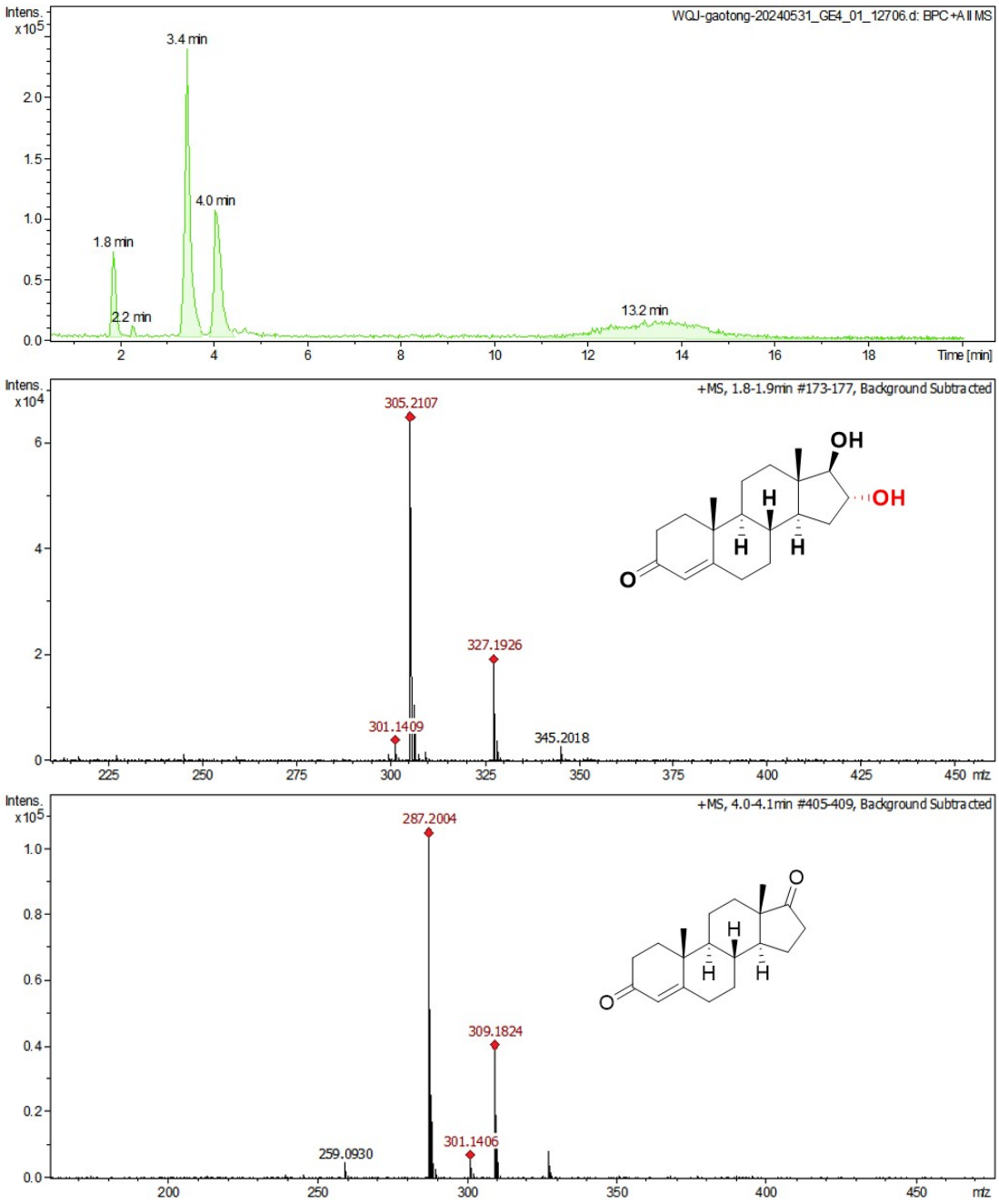


Figure S12. Major products observed in testosterone hydroxylation by the A282E mutant (Top).

LC-MS spectrum of 16α-hydroxytestosterone LCMS (ESI): m/z [M+H]+: calcd. for C19H29O3: 305.4318; found: 305.2107 (Medium).

LC-MS spectrum of androstenedione LCMS (ESI): m/z [M+H]+: calcd. for C19H27O2: 287.3983; found: 287.2004 (Bottom).

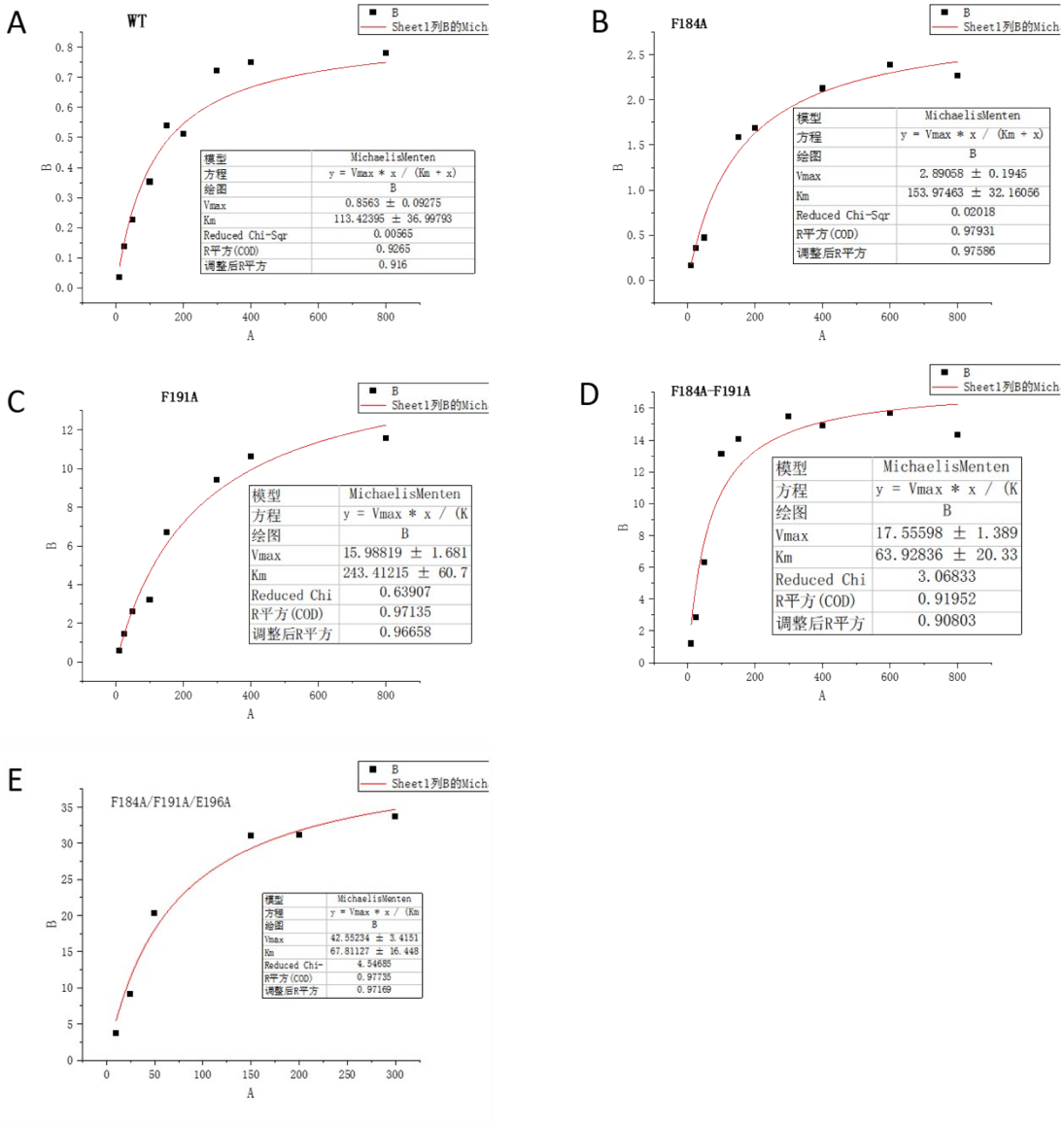


Figure S13. Kinetic parameters of CYP105D18 WT, F184A, F191A, F184A/F191A, and F184A/F191A/E196A for testosterone (A-E). Kinetic data for testosterone were measured in a system containing 60 mM H_2O_2 with an enzyme concentration of 1 μM.

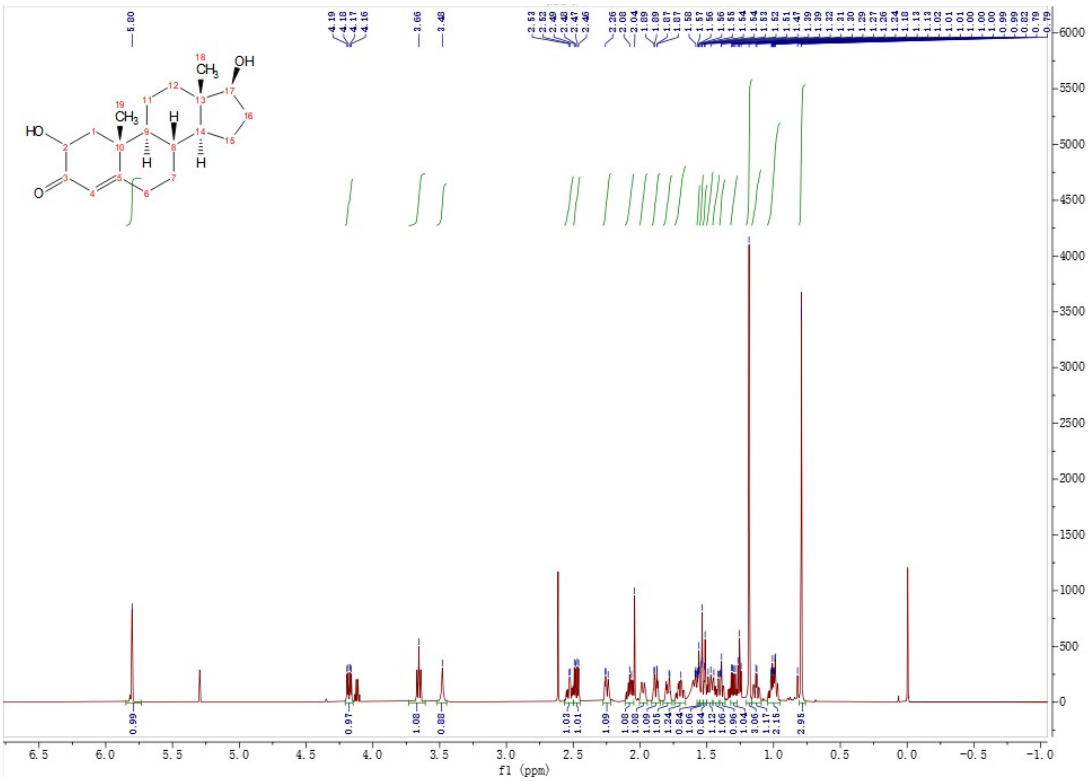


Figure S14 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 5.80 (s, 1H, H-4), 4.17 (dd, 1H, H-2), 3.66（t，1H,H-17）,2.48(dd，1H，H-1α)，2.07（m，1H，H-16α），1.51（H-1β），1.45（H-16β），1.18（s，3H，H-19），0.79（s，3H，H-18）

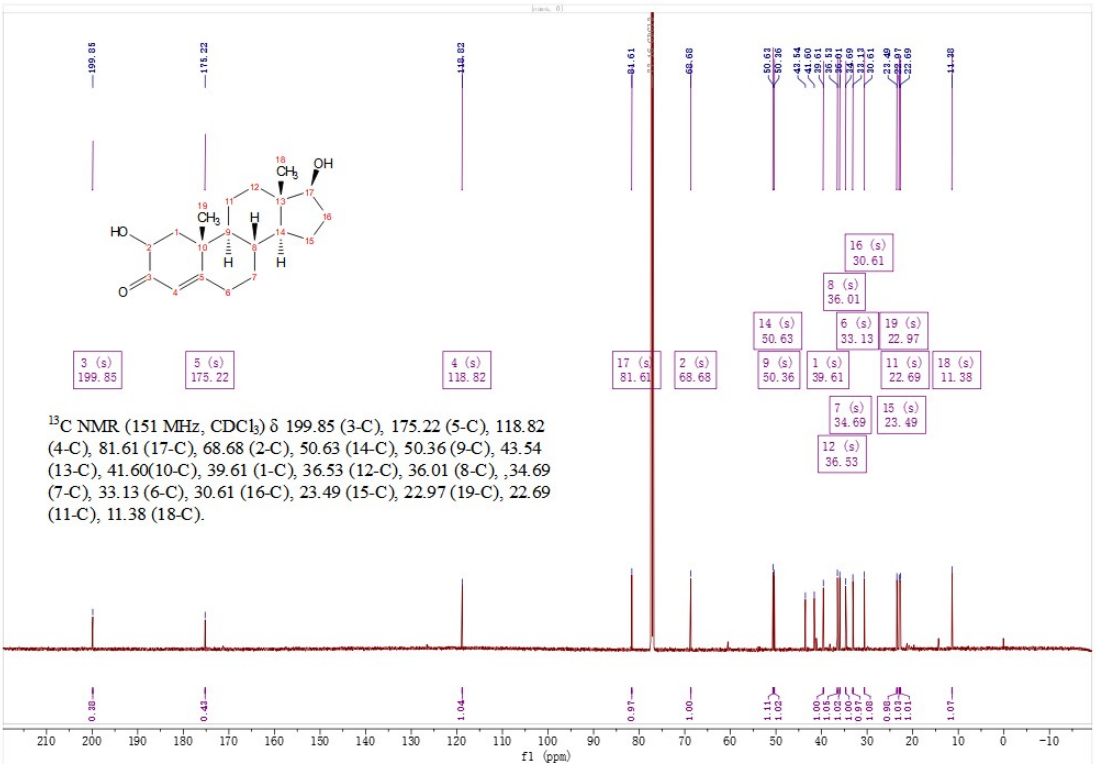


Figure S15 ^{13}C NMR (151 MHz, Chloroform-*d*) δ 199.85 (3-C), 175.22 (5-C), 118.82 (4-C), 81.61 (17-C), 68.68 (2-C), 50.63 (14-C), 50.36 (9-C), 43.54 (13-C), 41.60(10-C), 39.61 (1-C), 36.53 (12-C), 36.01 (8-C), ,34.69 (7-C), 33.13 (6-C), 30.61 (16-C), 23.49 (15-C), 22.97 (19-C), 22.69 (11-C), 11.38 (18-C).

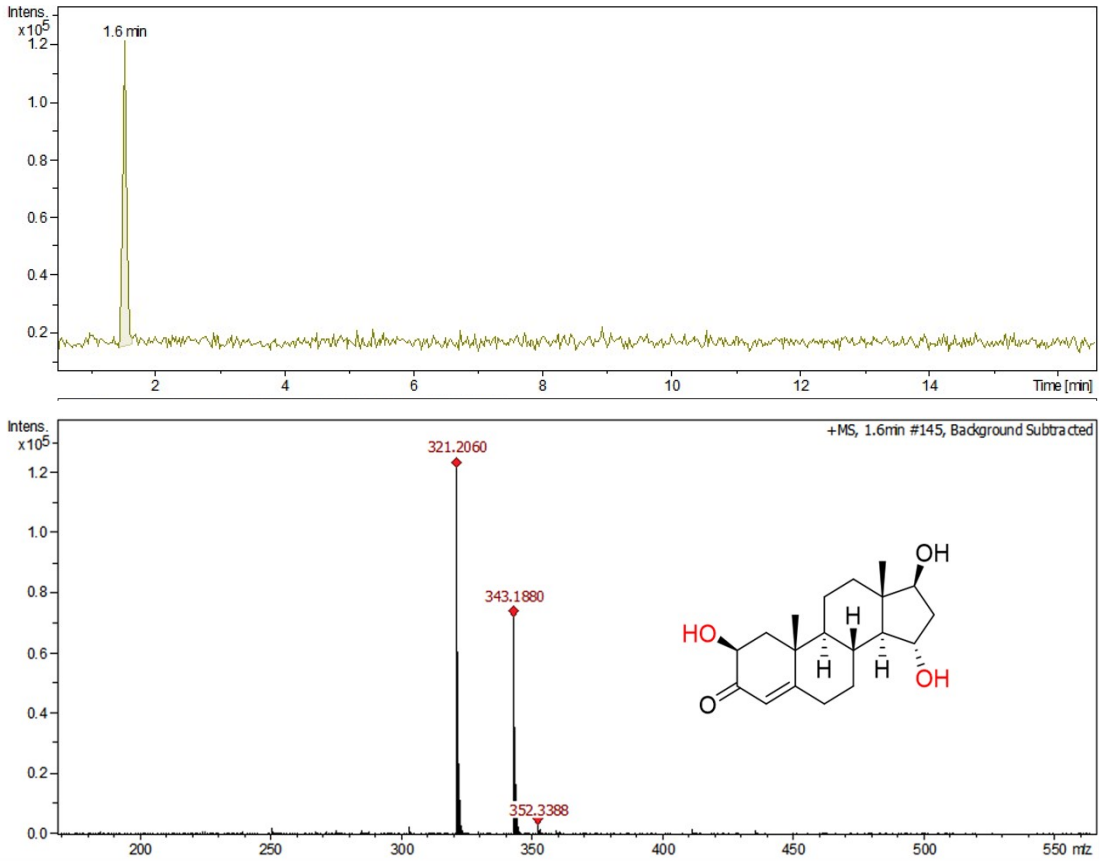
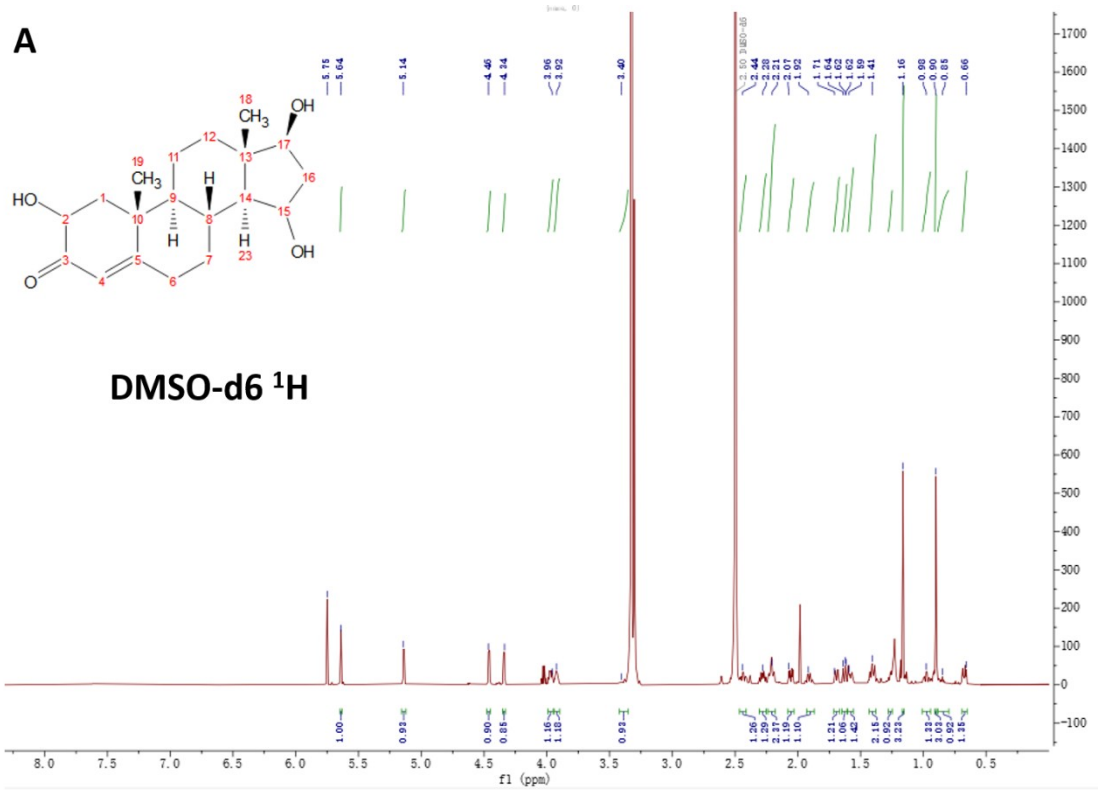
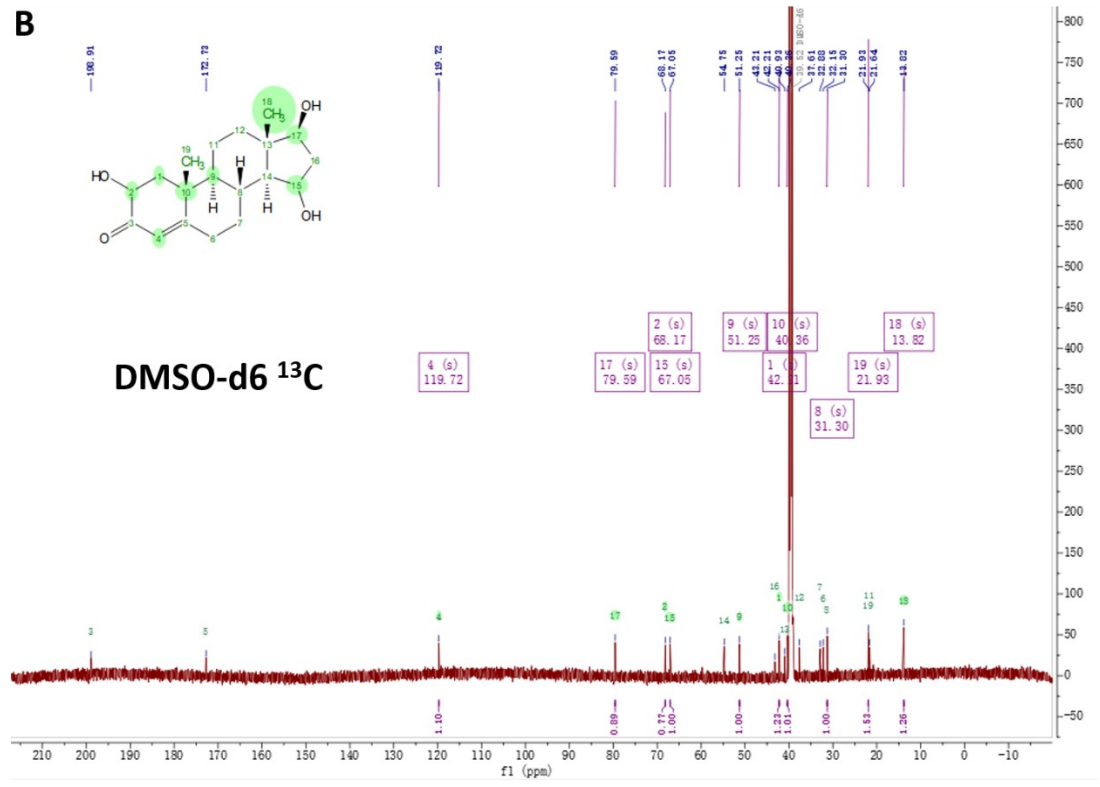
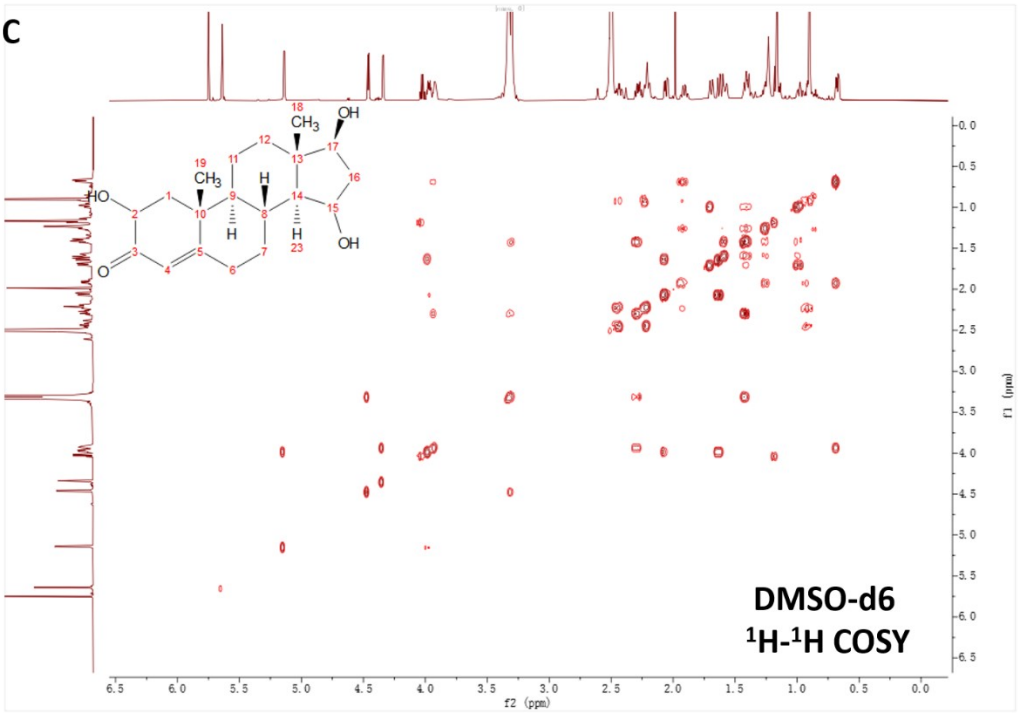
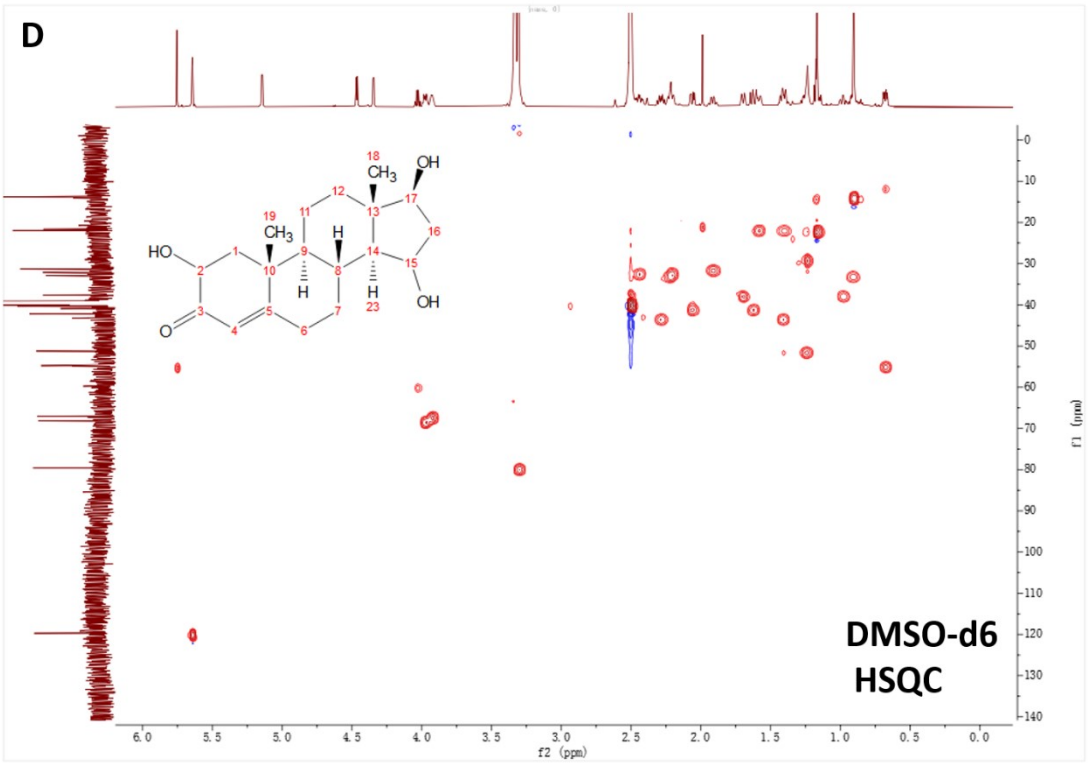


Figure S16. LC-MS spectrum of 2β,15α-dihydroxytestosterone. LCMS (ESI): m/z [M+H]+: calcd. for $C_{19}H_{29}O_4$: 321.41826; found: 321.2060.


A
DMSO-d6 1H
f1 (ppm)


B
DMSO-d6 13C
f1 (ppm)


C
DMSO-d6
^{1}H-^{1}H COSY
f2 (ppm)
f1 (ppm)


D
DMSO-d6
HSQC
f2 (ppm)
f1 (ppm)

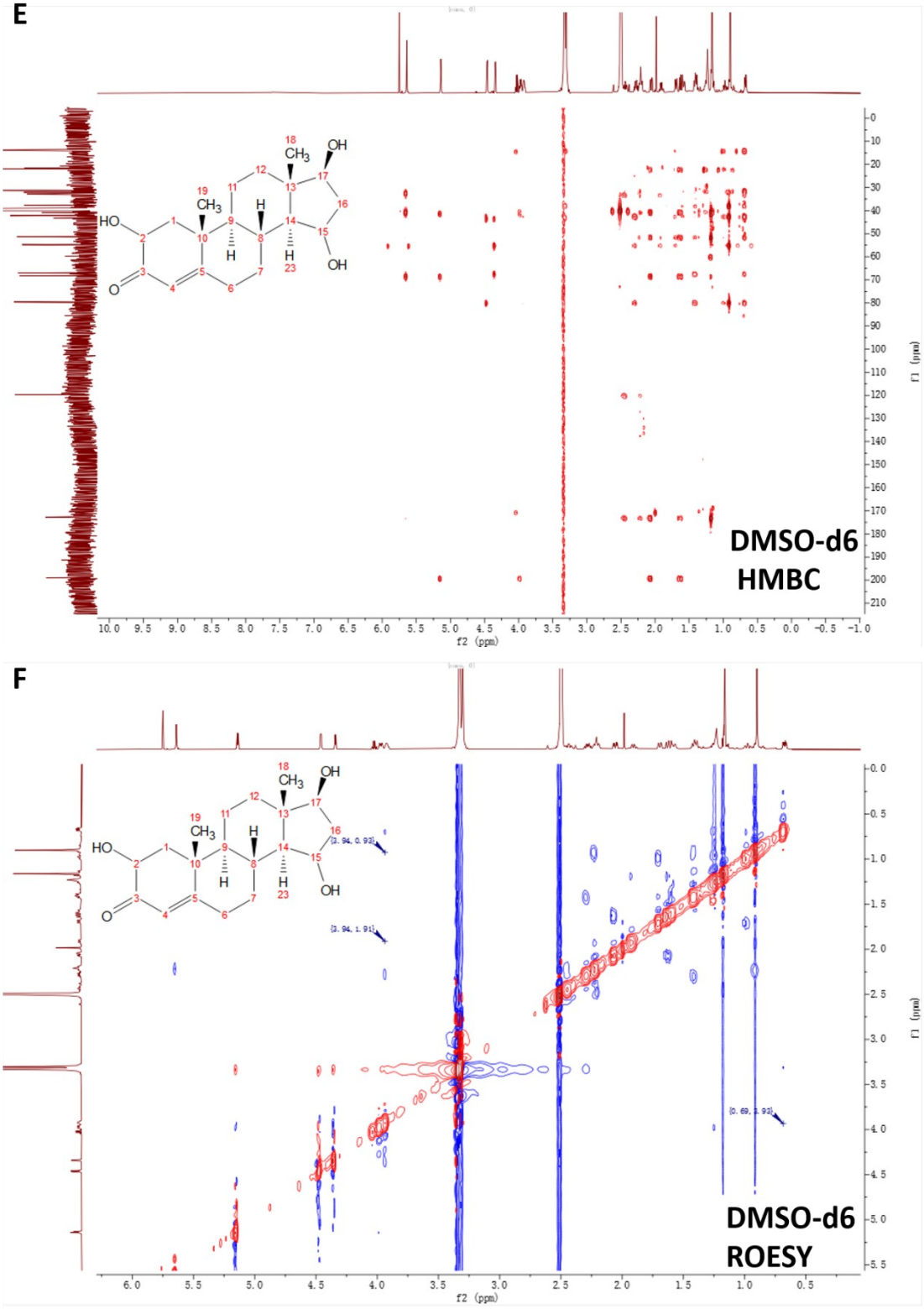
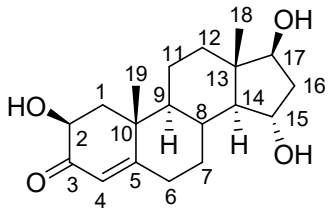
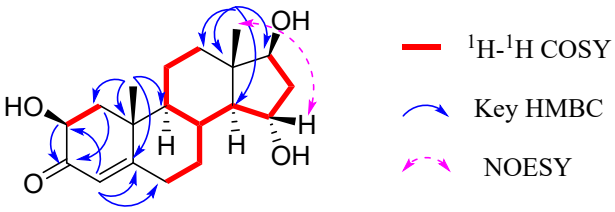


Figure S17. NMR spectrum of 2β,15α-dihydroxytestosterone in DMSO-D6.

Table S2. ^{1}H and ^{13}C NMR data for 2β,15α-dihydroxytestosterone in DMSO-D6



1



1

No	δ_C	δ_H (mult., *J* in Hz)	No	δ_C	δ_H (mult., *J* in Hz)
1	40.9	2.06, dd (13.8, 4.8) 1.62, dd (13.8, 11.0)	11	21.6	1.58, dd (13.5, 3.2) 1.41
2	68.2	3.97, ddd (11.6, 8.1, 3.6)	12	37.6	1.69, m 0.98, dt (12.8, 3.8)
3	198.9	-	13	42.2	-
4	119.7	5.64, brs	14	54.7	0.68, m
5	172.7	-	15	67.1	3.93, m
6	32.1	2.44, m 2.21, m	16	43.2	2.28, dt (8.0, 8.0, 14.2) 1.41, m
7	32.9	2.21, m 0.92, m	17	79.6	3.31
8	31.3	1.91, ddd (22.4, 11.3, 3.4)	18	13.8	0.90, s
9	51.2	1.23, m	19	21.9	1.17, s
10	40.3	-			

Table S3. HMBC,^{1}H-^{1}H COSY and ROESY data for 2β,15α-dihydroxytestosterone in DMSO-D6

No	HMBC	^{1}H-^{1}H COSY	ROESY
1	2，4，19	1，2	
2	1，4	1	
3	1，2		
4	6	6	
5	1，4，6，19		
6	4，7	4，6，7	
7	6，8	6，8，9	
8	7，9，14	7，9，11，14	
9	1，8，11	7，8，11	
10	1，4，6，9		
11	9，12	2，9，11，12，16，17	
12	11，14，18	11，12，16	
13	12，18		
14	8，16，18	8，15	
15	14，16	14，16	18
16	17，18	15，16，17	
17	16，18	16	
18	12，14，17		15
19	1，9		

Reference

1. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. Berendsen, *J. Comput. Chem.* 2005**, 26 (16)**, 1701-1718.
2. J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, A. Onufriev, *Nucleic Acids Res.* 2005**, 33** , W368-W371.
3. R. Anandakrishnan, B. Aguilar, A. V. Onufriev, *Nucleic Acids Res.* 2012, **40**, W537-W541.
4. C. Tian, K. Kasavajhala, K. A. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, C. Simmerling, *J. Chem. Theory. Comput.* 2020, **16 (1)**, 528-552.
5. F. Neese, *WIREs. Comput. Mol. Sci.* 2022, **12 (5)**, e1606.
6. T. Lu, F. Chen, *J. Comput. Chem.* 2012, **33 (5)**, 580-592.
7. D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* 2005, **26 (16)**, 1668-1688.
8. R. Salomon-Ferrer, D. A. Case, R. C.Walker, *WIRES. Comput. Mol. Sci.* 2012, **3 (2)**, 198-210.
9. S. Miyamoto, P. A. Kollman, *J. Comput. Chem.* 1992, **13 (8)**, 952-962.
10. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* 1997, **18 (12)**, 1463-1472.
11. T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* 1993, **98 (12)**, 10089-10092.
12. A. S. Lemak, N. K. Balabaev, *Mol. Simul.* 1994, **13 (3)**, 177-187.
13. M. Parrinello, A. Rahman, *J. Appl. Phys.* 1981, **52 (12)**, 7182-7190.
14. F. Yu, Q. Wang, M. Li, H. Zhou, K. Liu, K. Zhang, Z. Wang, Q. Xu, C. Xu, Q. Pan, J. He, *J. Appl. Crystallogr.* 2019, **52 (2)**, 472–477.
15. L. Wang, Y. Yun, Z. Zhu, L. Niu, *Acta Crystallogr. D Struct. Biol. Crystallogr.* 2022, **78 (7)**, 890–902.
16. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D-Biol. Crystallogr.* 2010, **66**, 213−221.
17. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 2010, **66**, 486−501.
18. V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr., Sect. D: Struct. Biol. Crystallogr.* 2010, **66**, 12−21.