

Supplementary Information

Sequential resolution of (*S*) and (*R*)-6-fluoro-chroman-2-carboxylic acid by two esterases in turn

Min Jiang^a, Yongjun Liu^a, Guosong Liu, Wei Zhou^a, Yaling Shen^a, Bei Gao^{a*},
Fengqing Wang^{a*}, Dongzhi Wei^{a*}

^a State Key Lab of Bioreactor Engineering, New World Institute of Biotechnology, East China University of Science and Technology, Shanghai, 200237, China

*Corresponding authors. Address: East China University of Science and Technology, P.O.B.311, 130 Meilong Road, Shanghai 200237, China. Fax: +8621 64250068
E-mail address: gaobei@ecust.edu.cn (B. Gao); fqwang@ecust.edu.cn (F. Wang); dzhwei@ecust.edu.cn (D. Wei)

List of contents

Supplementary experimental methods

1. Sequence Analysis
2. Purification of EstS and EstR
3. Determination of kinetic parameters
4. Structural modeling, docking simulation and molecular dynamics simulation
5. Immobilized recombinant cells
6. Analytical methods
7. Reference

Supplementary data

Table S1. The screening of microorganism for enantioselective hydrolysis of racemic MFCC. (S5)

Table S2. Primers used in this study. (S6)

Figure S1. Multiple sequence alignment of EstS and EstR. (S6)

Figure S2. Phylogenetic tree analysis of EstS and EstR. (S7)

Figure S3. SDS-PAGE analysis of EstS and EstR. (S8)

Table S3. Kinetic parameters of EstS and EstR. (S8)

Figure S4. Docking models of (*SR*)-MFCC into EstS. (S9)

Figure S5. Docking models of (*SR*)-MFCC into EstR. (S9)

Figure S6. Overview of the difference between two isomers in EstS and EstR. (S10)

Figure S7. DHG-O2 change as a function of MD simulation time over 50 ns. (S10)

Figure S8. RMSD of complex backbone atoms as a function of the simulation time over 50 ns for (*S*)-MFCC in EstS (A), (*R*)-MFCC in EstS (B), (*S*)-MFCC in EstR (C) and (*R*)-MFCC in EstR (D). (S10)

Table S4 Binding energies and docking score of (*S*)-MFCC in EstS, (*R*)-MFCC in EstS, (*R*)-MFCC in EstR and (*S*)-MFCC in EstR. (S11)

Figure S9 Effect of temperature and buffer pH on the enzyme hydrolysis racemic MFCC in biphasic system. (S12)

Table S5 Resolution results of racemic MFCC by lycEstS and lycEstR in the optimal reaction conditions. (S13)

Figure S10. (A) HPLC analysis of (*S*)-FCCA produced by EstS. (B) HPLC analysis of (*R*)-FCCA produced by EstR. (S14)

Figure S11. ¹³C NMR spectrum of (*S*)-FCCA. (S15)

Figure S12. ¹H NMR spectrum of (*S*)-FCCA. (S15)

Figure S13. ¹³C NMR spectrum of (*R*)-FCCA. (S16)

Figure S14. ¹H NMR spectrum of (*R*)-FCCA. (S16)

Supplementary experimental methods

1. Sequence Analysis

Open reading frames (ORFs) were analyzed using <http://nocardia.nih.go.jp/fp4/>. The deduced amino acid sequences were analyzed using the BLASTP program (<http://blast.ncbi.nlm.nih.gov/>). The putative signal peptide and its cleavage site were predicted using the SignalP 5.0 server. The theoretical values of protein molecular weight and *pI* were estimated using ExPASy (http://web.expasy.org/compute_pi/). Multiple sequence alignments were performed using DNAMAN software and exported using ESPript 3.0 (<https://esprict.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>). Phylogenetic tree analysis was conducted using the MEGA software version 7.

2. Purification of EstS and EstR

Harvested recombinant *E. coli* cells were resuspended in buffer A (50 mM phosphate buffer, 500 mM NaCl, pH7.4). The cells were disrupted by sonication for 30 min, and the cell lysates were centrifuged at 4 °C. Then the supernatant was filtered and loaded onto a Ni-NTA Sepharose column (Qiagen, Dusseldorf, Germany) pre-equilibrated with buffer A containing 20 mM imidazole, and the enzymes were eluted with an increasing gradient of imidazole from 100-500 mM in buffer A. The purified proteins were collected, concentrated by ultrafiltration and detected by SDS-PAGE. Protein concentrations were determined using the BCA Protein Assay Kit (Biotech Well, Shanghai, China) with bovine serum albumin (BSA) as a standard.

3. Determination of kinetic parameters

The kinetic parameters of the purified enzyme toward substrate racemic MFCC were determined by measuring the activities at various racemic MFCC concentrations (5–100 mM) using the standard HPLC method. The Michaelis-Menten constant (K_m) and catalytic number (k_{cat}) of the enzyme were calculated according to the Lineweaver-Burk plots.

4. Analysis of enantioselectivity mechanism through structural modeling, docking simulation and molecular dynamics simulation

First, protein structure determination: the structure of EstS and EstR were determined by trRosetta program (<https://yanglab.nankai.edu.cn/trRosetta/>). The quality of the final model was evaluated with the program PROCHECK. Second, docking of the substrate to the protein: the 3D structure of (*S*)-MFCC and (*R*)-MFCC were drawn using ChemBioOffice 2018 (CambridgeSoft, Cambridge, UK). Molecular docking was carried out through the maestro program. The protein Preparation Wizard module was employed for the preconditioning of the protein structure, including addition of hydrogen atoms, deletion of waters, treatment of overlaps, filling of the missing side

chains, optimization of hydrogen bonds, and energy minimization. The grid file of the docking site of the receptor was generated using the Receptor Grid Generation module. The size of the docking box was $20 \text{ \AA} \times 20 \text{ \AA} \times 20 \text{ \AA}$. Substrates were treated using the LigPrep module at $\text{pH } 7.0 \pm 2.0$. Finally, molecular docking was performed using the Ligand docking module. All initial structures of the complex ((*S*)-MFCC-EstS, (*R*)-MFCC-EstS, (*S*)-MFCC-EstR, (*R*)-MFCC-EstR) were described using PyMOL, and interactions between proteins and ligands were visualized using BIOVIA Discovery Studio (Accelrys, CA, USA). Third, the binding free energy was calculated by MD simulation. The aforementioned structures of the complex were subjected MD simulations using AMBER18. Proteins were treated by the Amber ff14SB force field, and substrates were dealt with Generalized Amber Force Field (GAFF). The complex was solvated in a TIP3P water box with a buffer distance of 12 \AA , and counterions were added to maintain the electrically neutral system. Energy minimization was carried out in two steps using the steepest descent and conjugate gradient method: in step 1, solvent water molecules were optimized by holding the PEP with a force constant of $500 \text{ kcal/mol/\AA}^2$; and in step 2, the PEP and solvent water molecules minimized energy without any restrictions. The entire system was heated from 0 to 300 K with a $10 \text{ kcal/mol/\AA}^2$ restraint for 300 ps. The SHAKE algorithm was applied to constrain all chemical bonds involving hydrogen atoms. Finally, the MD simulations were performed for 50 ns. The MM-GB/SA methods from Amber were used to quantify the binding free energy of the substrate in both EstS and EstR enzymes. The DHG-O2 was measured using the cpptraj package in amber18.

5. Immobilized recombinant cells

The immobilized cells were prepared using a previously described method with slight modification.¹ Briefly, 4 g polyvinyl alcohol and 1 g sodium alginate were added to 50 mL saline solution, stirred and dissolved in a thermostatic water bath at $80 \text{ }^\circ\text{C}$. After the solution was completely dissolved and cooled to $20 \text{ }^\circ\text{C}$, the solution was evenly mixed with the same volume of bacterial liquid with a mass concentration of 20%. The resulting mixture was dropped by means of an injector into a gently stirred saturated boric acid solution with a mass concentration of 2% calcium chloride and stored for 4 h to form gel beads. The formed particles were washed thrice with physiological saline and then stored in phosphate buffer at $4 \text{ }^\circ\text{C}$ until further use.

6. Analytical methods

Chiral HPLC analysis was performed on a Waters Technologies 2695 Alliance machine and a Chiralpak AD-H column ($25 \text{ cm} \times 4.6 \text{ mm}$, $5 \text{ }\mu\text{m}$). HPLC was operated as follows: mobile phase, eluent isopropanol and n-hexane (10:90); flow rate, 1 mL min^{-1} ; column temperature, $25 \text{ }^\circ\text{C}$; and UV detection, 254 nm . NMR spectra were recorded on a Bruker Ascend 600 spectrometer with CDCl_3 as the solvent.

7. Reference

- 1 L.-S. Zhang, W.-Z. Wu and J.-L. Wang, *J. Environ.Sci.*, 2007, **19**, 1293-1297.

Supplementary data

1. The screening of candidate microorganism

Table S1 The screening of microorganism for enantioselective hydrolysis of racemic MFCC^a.

Strain	Conv. (%)	ee _p (%) / Config
<i>Serratia marcescens</i>	15.6±1.1	21.2±0.9 (R)
<i>Pseudomonas fluorescens</i>	27.3±0.9	65.6±1.6 (S)
<i>Arthrobacter aurescens</i>	10.2±0.8	17.3±0.7 (S)
<i>Arthrobacter arilaitensis</i>	21.7±1.3	34.4±1.5 (S)
<i>Arthrobacter chlorophenolicus</i>	56.8±2.1	63.5±1.8 (S)
<i>Arthrobacter phenanthrenivorans</i>	32.4±1.5	83.1±1.6 (S)
<i>Bacillus subtilis</i> TD7	23.3±1.2	69.2±1.3 (S)
<i>Acinetobacter tandoii</i>	4.6±0.6	6.9±0.2 (R)
<i>Geobacillus kaustophilus</i>	8.7±0.5	75.2±1.4 (S)
<i>Geobacillus thermocatenulatus</i>	36.7±1.2	95.6±1.2 (S)

^a Reaction conditions: the reaction mixture containing 0.1 g lyophilized cells suspended in 8.0 mL phosphate buffer (200 mM, pH7.5) and 2.0 mL toluene containing 0.5 mM MFCC (apparent concentration, 50 mM based on the total volume) was incubated at 220 rpm and 30 °C for 12 h. All experiments were performed in triplicate.

2. Selection, expression, purification and Kinetic parameters

Table S2. Primers used in this study. ^a restriction sites are underlined.

Designation	Primer sequence (5'-3')
Lip1F	CGCGGATCCATGAACGAACAATATCCGGT ^a
Lip1R	CGCAAGCTTCTATTTTCGAATGCTTGGCGA
Lip2F	CGCGGATCCATGAAACGATGGGGATGGTT
Lip2R	CGCAAGCTTTTCATTGCAGCAAACCTCCCTA
Est1F	CGCAAGCTTTCATTGCAGCAAACCTCCCTA
Est1R	CGCAAGCTTTTCATTTCGTGTGATGACTCCT
Est2F	CGCAAGCTTTCATTTCGTGTGATGACTCCT
Est2R	CGCAAGCTTCTATAACGATAAAAACAGCGG
EstSF	CGCGGATCCATGGTCATTGTTGAAACAGA
EstSR	CGCAAGCTTTTACACATGCTCGCGAAACC
EstRF	CGCGGATCCGTGCAAGACCAGTTTTTTTTC
EstRR	CGCAAGCTTTTCATTGTTACCCCTCCTCCG

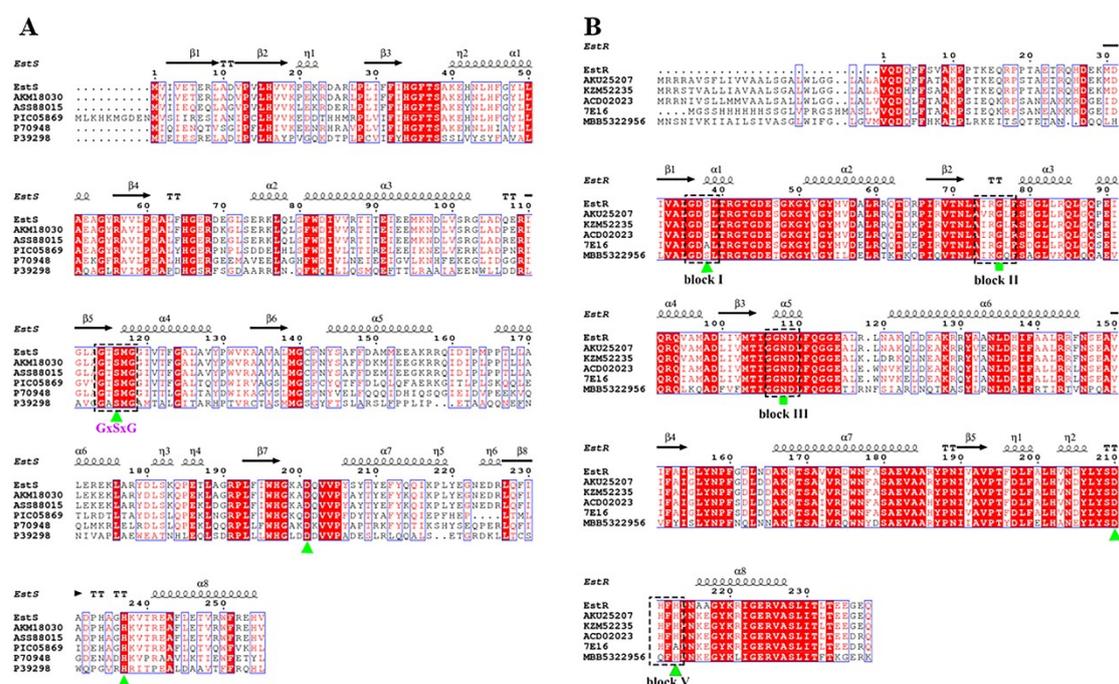


Figure S1. Multiple sequence alignment of EstS and EstR. (A) EstS: the esterase in this

study; ASS88015: esterase from *Geobacillus lituanicus*; AKM18030: esterase from *Geobacillus* sp.; PIC05869: esterase in *Anoxybacillus flavithermus*; P70948: putative esterase *Bacillus subtilis*; P39298: esterase in *Escherichia coli* K-12 ATCC 7954. The catalytic triad is emphasized by green triangles, and the GxSxG motif is boxed. (B) EstR: the esterase in this study; AKU25207: GDSL family lipase from *Geobacillus* sp.; KZM52235: GDSL family lipase from *Geobacillus stearothermophilus*; ACD02023: GDSL family esterase from *Geobacillus thermodenitrificans*; 7E16: GDSL family esterase from *Geobacillus thermodenitrificans*; MBB5322956: lysophospholipase L1-like esterase from *Anoxybacillus tepidamans*. Four conserved blocks of I, II, III, and V are boxed. The catalytic triad is emphasized by green triangles, the oxyanion hole supported Gly76 and Asn108 residues are emphasized by green quadrates.

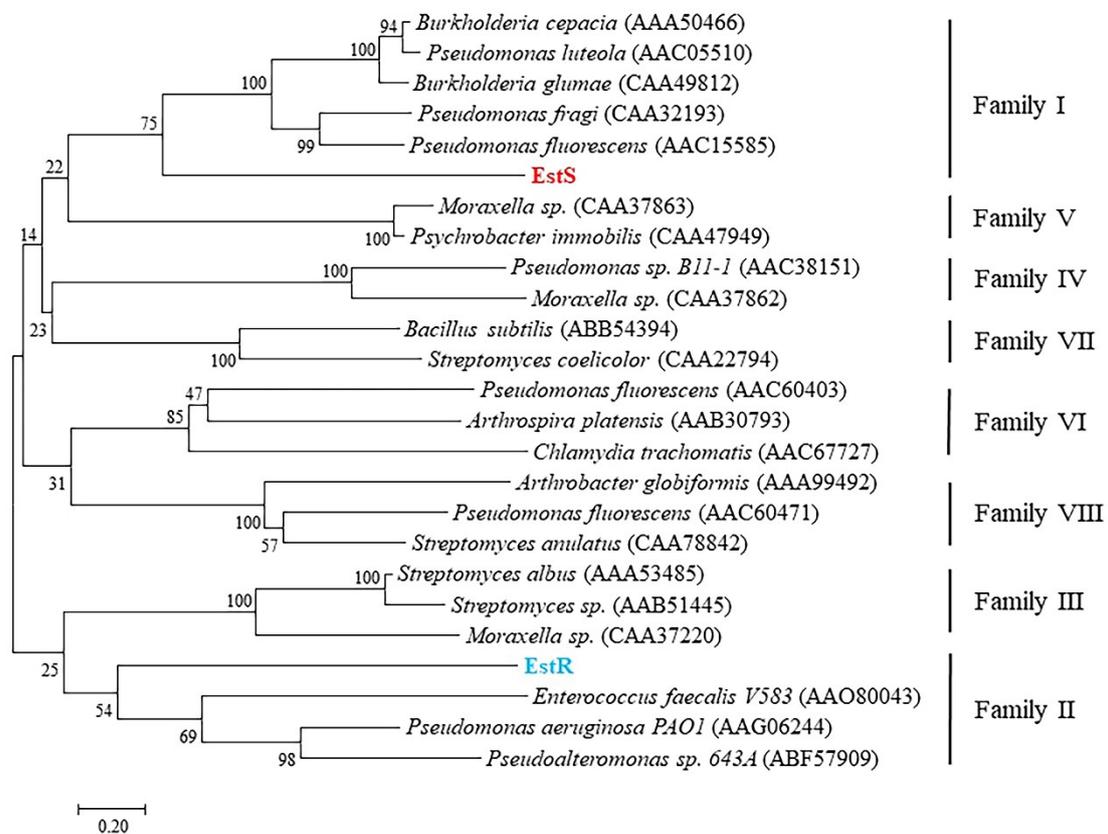


Figure S2. Phylogenetic tree analysis of EstS and EstR. The scale at the bottom represents the number of substitution events.

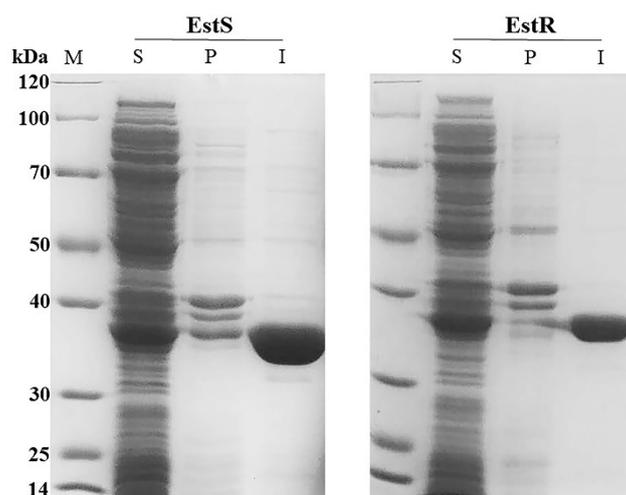


Figure S3. SDS-PAGE analysis of EstS and EstR. EstS, *E. coil* BL21 (DE3) expressing recombinant plasmid pET28a-EstS; EstR, *E. coil* BL21 (DE3) expressing recombinant plasmid pET28a-EstR. lane M: standard marker proteins; lane S: supernatant of cell lysate; lane P: precipitation of cell lysate; lane I: the purified protein.

Table S3. Kinetic parameters of EstS and EstR^a

enzyme	Substrate	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)
EstS	MFCC	23.14	3.50	0.151
EstR	MFCC	33.46	1.73	0.052

^aThe assay was performed using the HPLC method.

3. Docking simulation and molecular dynamics simulation

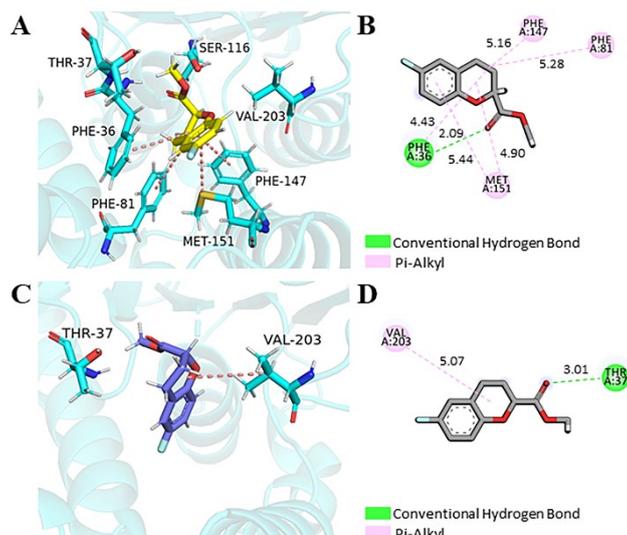


Figure S4. Docking models of (*SR*)-MFCC into EstS. (A) Detailed interactions of the docking (*S*)-MFCC and residues in the active site of EstS. (B) 2D ligand interaction diagram of docking (*S*)-MFCC with the EstS active site. (C) Detailed interactions of the docking (*R*)-MFCC and residues in the active site of EstS. (D) 2D ligand interaction diagram of docking (*R*)-MFCC with the EstS active site. Hydrogen bonds are indicated with green dashed lines; π -interactions (π - π and π -alkyl) are shown with violet dashed lines, yellow and blue sticks denote (*S*)-MFCC and (*R*)-MFCC, respectively.

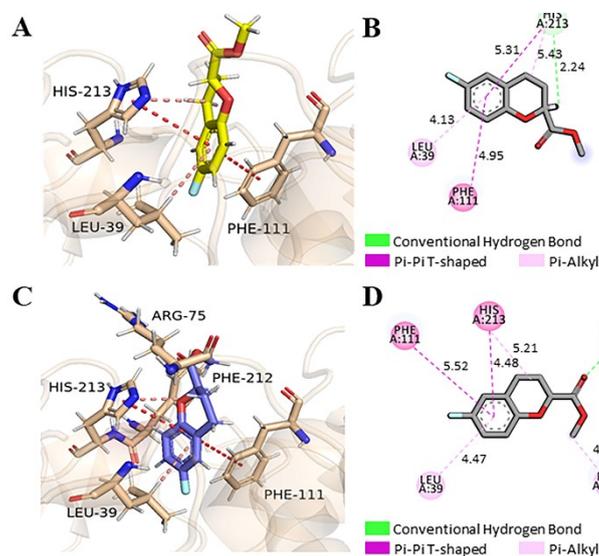


Figure S5. Docking models of (*SR*)-MFCC into EstR. (A) Detailed interactions of the docking (*S*)-MFCC and residues in the active site of EstR. (B) 2D ligand interaction diagram of docking (*S*)-MFCC with the EstR active site. (C) Detailed interactions of the docking (*R*)-MFCC and residues in the active site of EstR. (D) 2D ligand interaction diagram of docking (*R*)-MFCC with the EstR active site. Hydrogen bonds are indicated with green dashed lines; π -interactions (π - π and π -alkyl) are shown with violet dashed lines, yellow and blue sticks denote (*S*)-MFCC and (*R*)-MFCC, respectively.

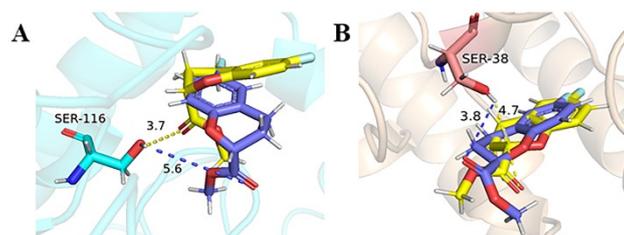


Figure S6. (A) Overview of the difference between two isomers in EstS. (B) Overview of the difference between two isomers in EstR. Yellow and blue sticks denote (*S*)-MFCC and (*R*)-MFCC, respectively.

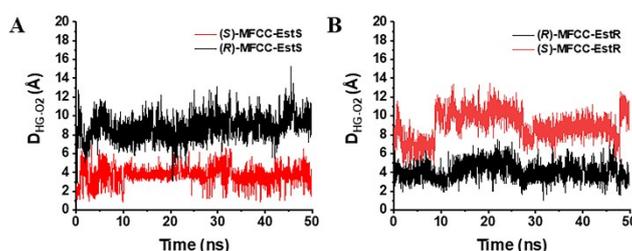


Figure S7. (A) DHG-O2 change as a function of MD simulation time over 50 ns for (*S*)-MFCC-EstS, (*R*)-MFCC-EstS, (B) DHG-O2 change as a function of MD simulation time over 50 ns for (*R*)-MFCC-EstR and (*S*)-MFCC-EstR.

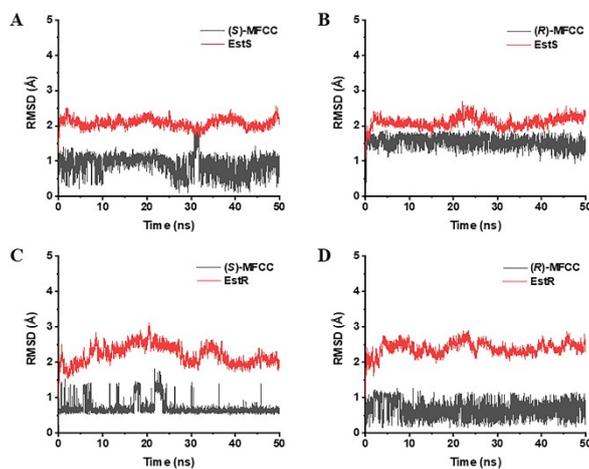


Figure S8. RMSD of complex backbone atoms as a function of the simulation time over 50 ns for (*S*)-MFCC in EstS (A), (*R*)-MFCC in EstS (B), (*S*)-MFCC in EstR (C) and (*R*)-MFCC in EstR (D).

Table S4 Binding energies and docking score of (*S*)-MFCC in EstS, (*R*)-MFCC in EstS, (*R*)-MFCC in EstR and (*S*)-MFCC in EstR.

Energy (kcal/mol)	(<i>S</i>)-MFCC-EstS	(<i>R</i>)-MFCC-EstS	(<i>R</i>)-MFCC-EstR	(<i>S</i>)-MFCC-EstR
ΔE_{vdW}	-25.4380	-24.4963	-32.4710	-29.8013
ΔE_{elec}	-1.9871	-4.0998	-7.1294	-3.6453
$\Delta E_{GB,elec}$	11.6867	13.686	19.2079	15.3171
ΔE_{surf}	-2.6020	-2.4669	-2.8811	-2.6901
ΔG_{bind}	-18.3405	-17.4344	-23.2735	-20.8195
docking score	-5.5030	-4.6450	-5.7870	-5.0350

4. Improving the performances of lycEstS and lycEstR on the resolution of FCCAs

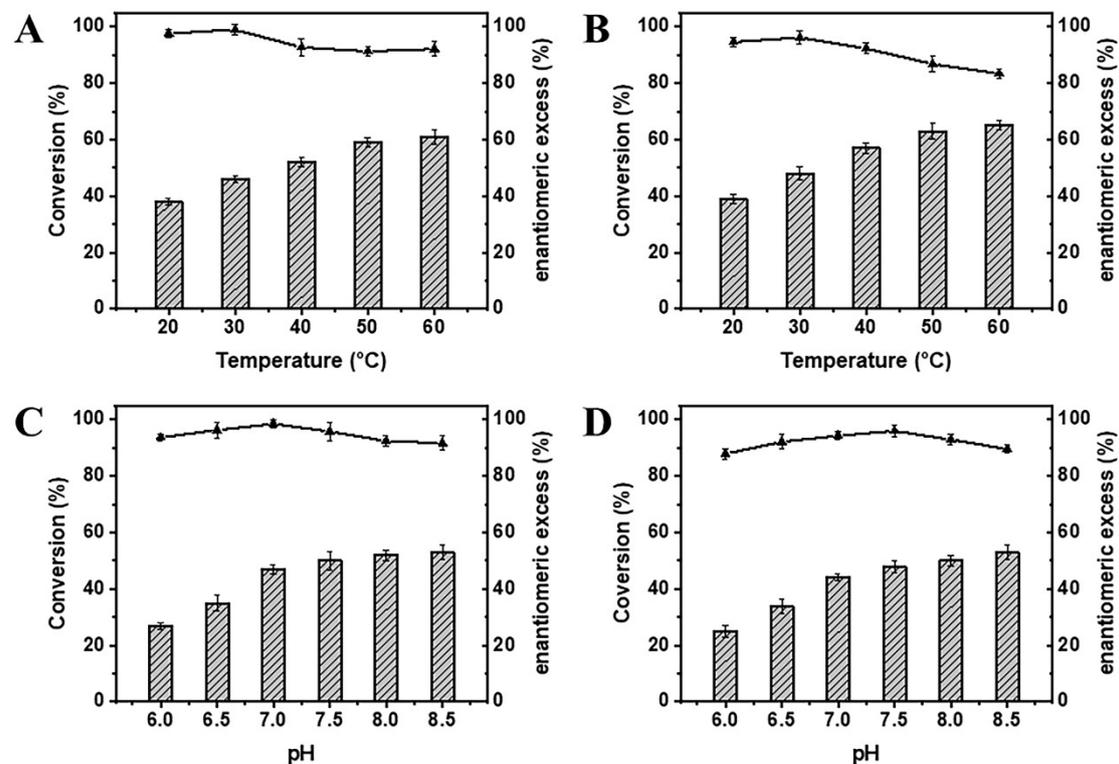


Figure S9 Effect of temperature and buffer pH on the enzyme hydrolysis of racemic MFCC in a biphasic system. (A) Effect of temperature on the lycEstS activity; (B) Effect of temperature on the lycEstR activity; (C) Effect of pH on the lycEstS activity; (D) Effect of pH on the lycEstR activity. Symbols: (▨) conversion; (▲) enantiomeric excess.

5. The MFCC remained after the first reaction with lycEstS or lycEstR has been recovered and determined

Table S5 Resolution results of racemic MFCC by lycEstS and lycEstR in the optimal reaction conditions^a.

enzyme	Conv. (%)	ee _s ^b (%) / Config.
lycEstS	49.1 ± 0.9	95.8 ± 0.9 (<i>R</i>)
lycEstR	50.5 ± 1.4	89.7 ± 1.8 (<i>S</i>)

^a Reaction conditions: the reaction mixture containing 1 g lyophilized cells suspended in 75 mL phosphate buffer (200 mM, pH7.5) and 25 mL toluene containing 20 mM MFCC (apparent concentration, 200 mM based on the total volume) was incubated at 220 rpm and 30 °C for 12 h. All experiments were performed in triplicate.

^b ee_s indicates enantiomeric excess of unreacted substrate.

6. HPLC chromatograms and NMR spectra of (*S*) and (*R*)-FCCA

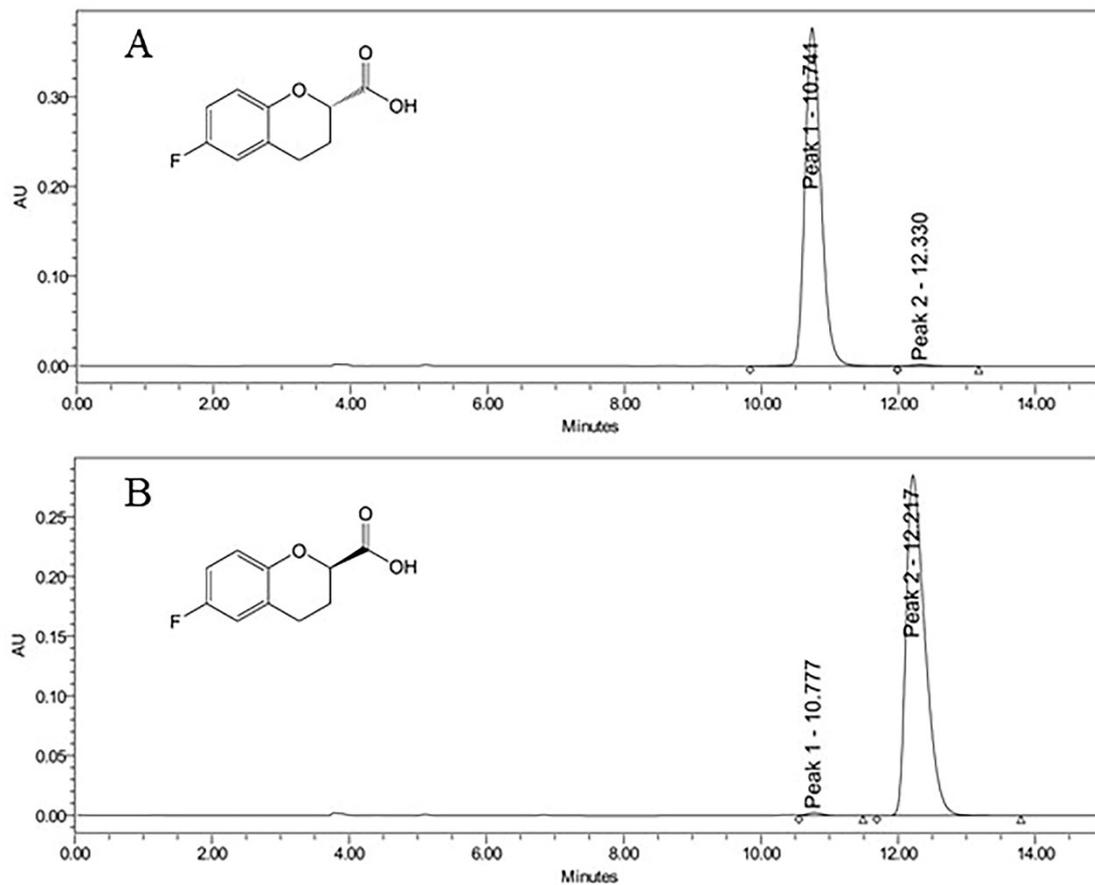


Figure S10. (A) HPLC analysis of (*S*)-FCCA produced by EstS. (B) HPLC analysis of (*R*)-FCCA produced by EstR.

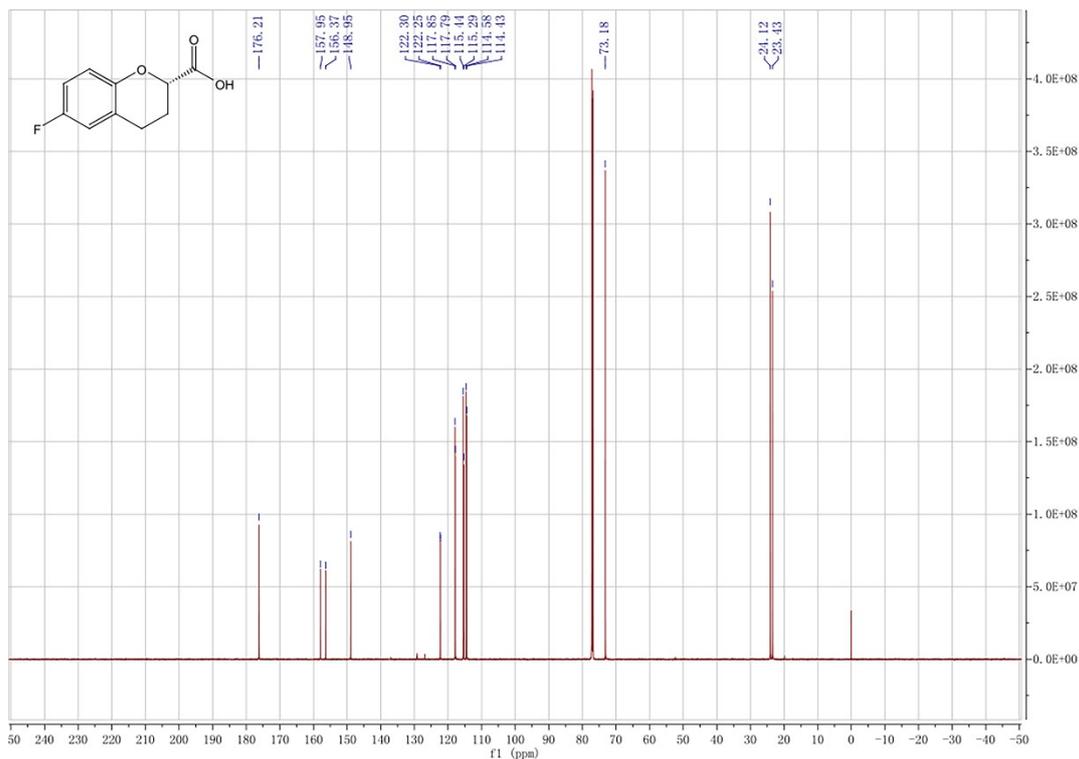


Figure S11. ^{13}C NMR spectrum of (*S*)-FCCA. ^{13}C NMR (151 MHz, CDCl_3) δ 176.29, 157.95, 156.37, 148.95, 122.30, 122.24, 117.85, 117.80, 115.44, 115.29, 114.58, 114.43, 73.18, 24.12, 23.43.

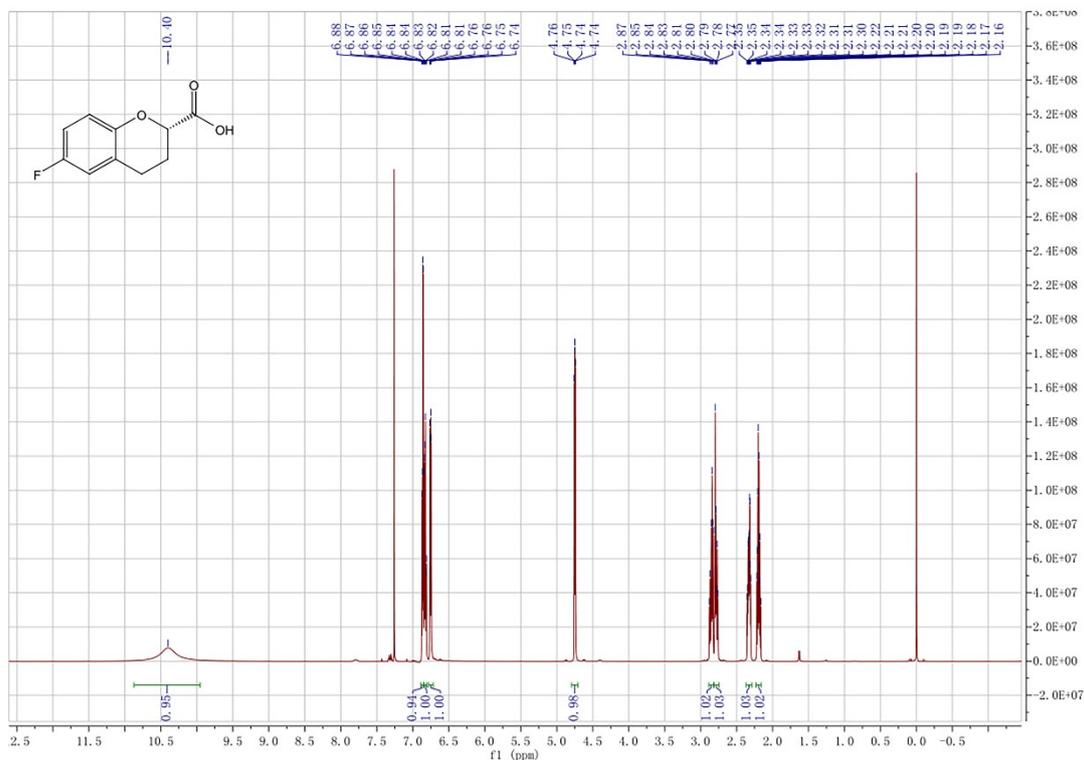


Figure S12. ^1H NMR spectrum of (*S*)-FCCA. ^1H NMR (600 MHz, CDCl_3) δ 10.40 (s, 1H), 6.86 (dd, $J = 9.0, 4.9$ Hz, 1H), 6.83 (td, $J = 8.4, 2.9$ Hz, 1H), 6.75 (dd, $J = 8.7, 2.8$ Hz, 1H), 4.75 (dd, $J = 7.8, 3.6$ Hz, 1H), 2.89 – 2.82 (m, 1H), 2.78 (dt, $J = 16.8, 6.1$ Hz, 1H), 2.33 (dtd, $J = 9.9, 6.3, 3.7$ Hz, 1H), 2.19 (dtd, $J = 13.7, 7.9, 5.8$ Hz, 1H).

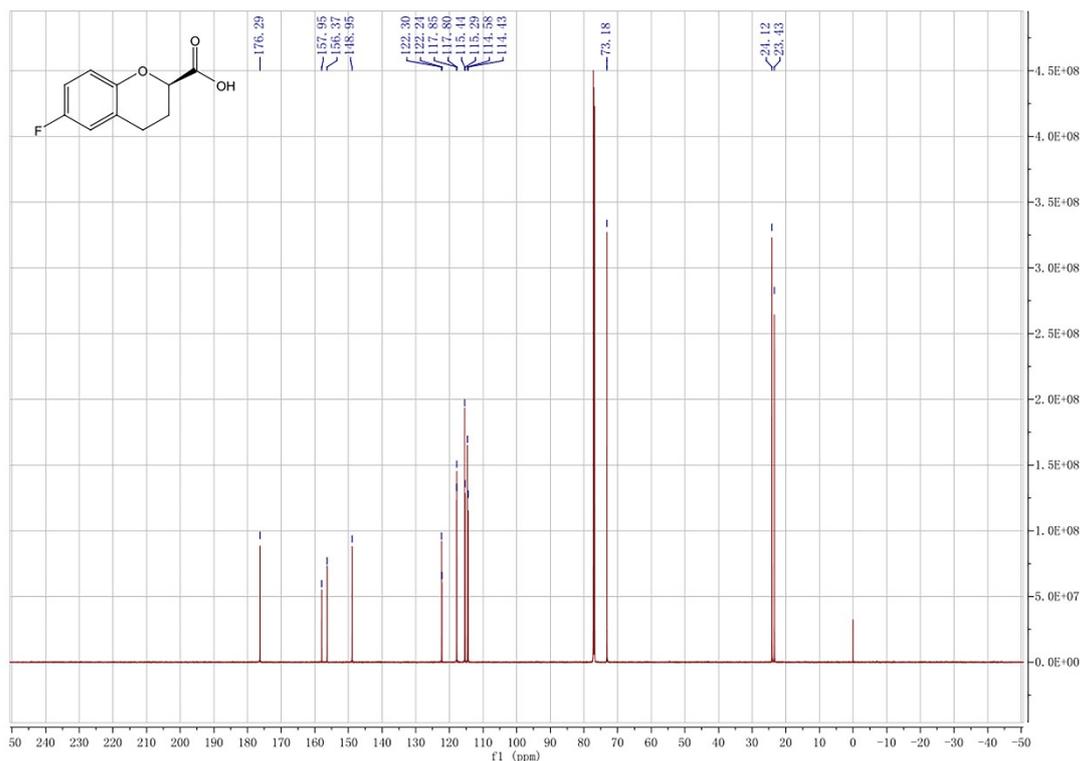


Figure S13. ¹³C NMR spectrum of (*R*)-FCCA. ¹³C NMR (151 MHz, CDCl₃) δ 176.29, 157.95, 156.37, 148.95, 122.30, 122.24, 117.85, 117.80, 115.44, 115.29, 114.58, 114.43, 73.18, 24.12, 23.43.

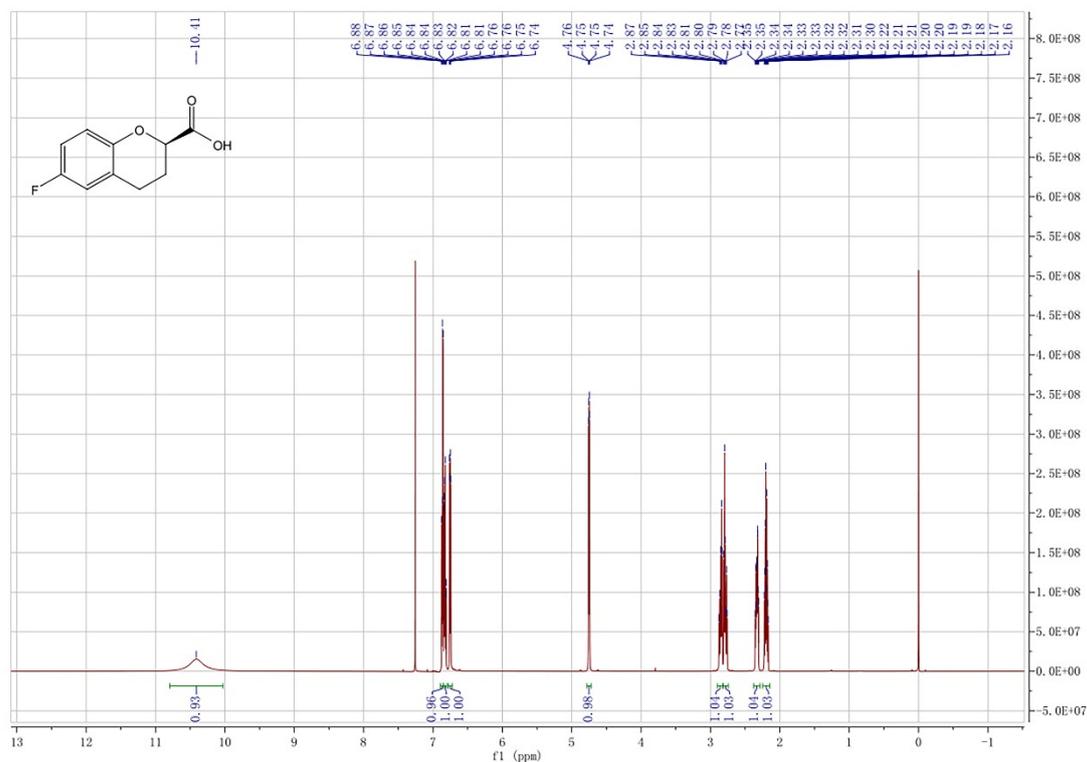


Figure S14. ¹H NMR spectrum of (*R*)-FCCA. ¹H NMR (600 MHz, CDCl₃) δ 10.41 (s, 1H), 6.87 (dd, *J* = 9.0, 4.9 Hz, 1H), 6.83 (td, *J* = 8.5, 2.9 Hz, 1H), 6.75 (dt, *J* = 8.7, 2.8 Hz, 1H), 4.75 (dd, *J* = 7.7, 3.6 Hz, 1H), 2.90 – 2.82 (m, 1H), 2.78 (dt, *J* = 16.8, 6.1 Hz, 1H), 2.33 (dtd, *J* = 9.9, 6.3, 3.7 Hz, 1H), 2.19 (dtd, *J* = 13.7, 7.9, 5.8 Hz, 1H).