Electronic Supplementary Material (ESI) for Catalysis Science & Technology.

Supporting Information

Engineering of D-Fructose-6-Phosphate Aldolase A for Improved Activity towards Cinnamaldehyde

*Xiaohong Yang, Lidan Ye, Aipeng Li, Chengcheng Yang, Huilei Yu, Jiali Gu, Fei Guo, Ling Jiang, Fan Wang and Hongwei Yu**

* Corresponding author:

Institute of Bioengineering, College of Chemical and Biological Engineering, Zhejiang University, Hangzhou 310027, PR China.

Tel.: +86 571 8795 1873; Fax: +86 571 8795 1873.

E-mail: yuhongwei@zju.edu.cn

Table of Content

Materials and methods

Chemicals and enzymes

Cinnamaldehyde (**1a**), α-bromocinnamal (**1b**), 2-pyridinecarboxaldehyde (**1d**) and D-Fructose 6-phosphate disodium salt hydrate were purchased from Aladdin Ltd. (Shanghai, China); 4-nitrocinnamaldehyde (**1c**) and hydroxyacetone(HA, **2**) were purchased from *Alfa Aesar* Ltd. (Shanghai, China), and methanol was purchased from Adamas-beta Ltd. (Shanghai, China). Other chemicals were purchased from Sangon Biotech Ltd. (Shanghai, China). Taq DNA polymerase, restriction endonucleases, T4 DNA ligase and PrimerSTAR™ HS DNA polymerase were purchased from TAKARA Ltd. (Dalian, China). α-Glycerophosphate Dehydrogenase-Triosephosphate Isomerase from rabbit muscle was purchased from Sigma-Aldrich Ltd. (Shanghai, China). DNA and protein markers were purchased from Sangon Biotech Ltd. (Shanghai, China).

Gene amplification, plasmid construction and site-directed mutagenesis

Gene encoding FSAA (mipB, Gene Bank Accesion number BAA13552.1) was amplified from genomic DNA of *E. coli* K12 by PCR using PrimerSTAR™ HS DNA polymerase with the following primers: FSAAF-*Bgl*Ⅱ(5'- CGCGCGAGATCTGATGGAACTGTATCTGGATACTT-3') and FSAAR-*Hind*Ⅲ(5'-CGCGCGAAGCTTTTAAATCGACGTTCTGCCAAACG-3'). The amplified sequence was inserted into the *Bgl*Ⅱ and *Hind*Ⅲ sites of pET30a (+), creating pET-30a-*fsa*A, which was then transformed into chemically competent *E. coli* BL21 (DE3) following standard protocols. After sequencing, the recombinant plasmid pET-30a-*fsa*A was extracted and transformed into chemically competent *E. coli* BL21-pGro7[1] (DE3). *E. coli* BL21-pGro7 (DE3) contains a vector pGro7 harboring groEL-groES, which is a class of common molecular chaperones that can help proteins to fold correctly. The expression of groEL-groES was induced by L-arabinose. PCR-based site-directed mutagenesis was performed following the QuikChange™ method (Stratagene, La Jolla, CA) using the recombinant plasmid pET-30a-*fsa*A as the template. Then the entire mutant gene was sequenced to exclude the introduction of unwanted mutations into the nucleotide sequence. Though heat treatment is often used for direct purification of wild-type FSAA, potential instability of mutants may occur during this process.[2] Therefore, a His_6-tagged FSAA construct was used in mutagenesis, which meanwhile facilitated the purification of FSAA mutants by HisTrap affinity chromatography.

Protein expression and purification

Wild-type FSAA and mutants were cultured in 100 ml Luria–Bertani (LB) media supplemented with 50 μg ml^{-1} of kanamycin and 20 μg ml^{-1} of chloramphenicol at 37 °C with shaking. When the culture reached an OD_{600} of 0.6, isopropyl-β-thiogalactopyranoside (IPTG, final concentration 100 μM) and L-arabinose (final concentration 1 mg mL^{-1}) were added to induce protein expression. Then the temperature was reduced to 25 °C and the cultures were incubated for another 7 h. Cells were

collected by centrifugation at 4000 × g, 4 °C and washed twice with phosphate buffer (50 mM, pH 7.4). The cell precipitation was resuspended in phosphate buffer (50 mM, pH 8.0), and then lysed by sonication. Cell debris was removed by centrifugation at 12,000 × g and 4 °C for 20 min, and the enzyme was purified from the supernatant by HisTrap affinity column (GE Healthcare) following the manufacturer's instructions. Eluted fractions containing pure protein were pooled and loaded onto HiTrap™ desalting columns to remove imidazole and then redissolved in citric acid-sodium citrate buffer (0.1 M, pH 6.5). The expression of the wild-type FSAA as well as the mutants was examined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and the protein concentration was determined by the Bradford method.[3]

The complete protein sequence of His_6-tagged wild-type FSAA was:

MHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDL*MELYLDTSDVVAVKALSRIFPLAGVTTNPSIIAAGK KPLDVVLPQLHEAMGGQGRLFAQVMATTAEGMVNDALKLRSIIADIVVKVPVTAEGLAAIKMLKAEGIPTLGTAVY GAAQGLLSALAGAEYVAPYVNRIDAQGGSGIQTVTDLHQLLKMHAPQAKVLAASFKTPRQALDCLLAGCESITLPLD VAQQMISYPAVDAAVAKFEQDWQGAFGRTSI*.

His_6-tag is underlined, and the sequence of wild-type FSAA is in italic.

Enzyme assay

Before testing the enzyme activity, reaction conditions for the aldol reactions (buffer, pH, temperature) were optimized. First of all, the spontaneous aldol addition of cinnamaldehyde to hydroxyacetone (HA) in the absence of enzyme was examined. Different buffering substances, glycylglycine, HEPES, phosphate, MOPS, citrate or borate were used in the pH range of 6.0-8.5. In order to control the yield of spontaneous reaction below 1% (after 12 h reaction) and meanwhile to achieve the optimal FSAA activity, citric acid-sodium citrate buffer (0.1 M, pH 6.5) was selected as the most suitable buffer. In addition, 30 °C was found to be the optimal temperature for the reaction of interest.

An initial screening was conducted to test the obtained mutants towards cinnamaldehyde (**1a**, table S4). Reactions (0.3 mL total volume) were performed using 1.5 mL test tubes at 30 °C in citric acid-sodium citrate buffer (0.1 M, pH6.5) containing 1 mM dithiothreitol (DTT) with agitation (220 rpm). Concentrations of acceptor **1a** and HA were 10 mM and 50 mM, respectively. Amount of enzymes and reaction time were adjusted to control the conversion below 15% in order to calculate the initial reaction rate (v_0, nmol min^{-1} mg^{-1} protein). After reaction, samples (300 μL) were dissolved in methanol (900 μL) and incubated in ice bath for 3 h, and then analyzed by high performance liquid chromatography (HPLC). The mutants with enhanced activity were chosen for calculation of the initial reaction rate and maximum conversion rate using **1a-1d** as acceptor substrates. Concentrations of acceptors (**1a-1d**) and HA were 100 mM and 500 mM, respectively. Reaction conditions were the same as described above. Conversion rates were calculated after agitating for 24 h when the product concentration became constant.

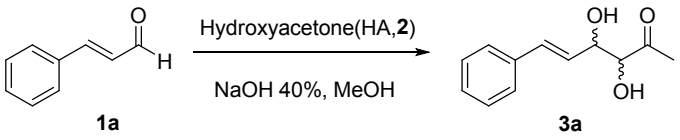
The reaction mixtures containing **1a** and HA were extracted using ethyl acetate after agitating for 24 h. The samples were dried and redissolved in isopropanol for diastereomeric ratio (dr) and enantiomeric excess (ee) analysis.

The specific activity of wild-type FSAA and the selected mutants towards D-fructose-6-phosphate was determined by measuring the formation of D-glyceraldehyde-3-phosphate (D-G3P) and dihydroxyacetone from D-fructose-6-phosphate (D-F6P) by spectrophotometric monitoring in

glycylglycine buffer (1 mL, 50 mM, pH 8.5) at room temperature. Standard spectrophotometric assay for D-G3P detection was performed as described previously using a coupled assay containing NADH 0.5 mM, glycerol phosphate dehydrogenase (1.3 U) and triose phosphate isomerase (12.6 U).[4, 5] The oxidation of NADH was monitored at 340 nm. 1 µmol of NADH oxidized was considered equivalent to 1 µmol of D-F6P cleaved.

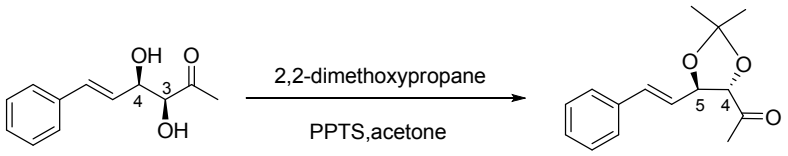
Synthesis of product 3a as standard reference and derivative 4 for determination of the absolute configuration of enzymatically catalyzed product 3a

(3*RS*,4*RS*,*E*)- and (3*RS*,4*SR*,*E*) -3,4-dihydroxy-6-phenylhex-5-en-2-one (**3a**) were prepared as described by Acetti and Gatti,[6] and the spectral data are in agreement with the reported values (see **Fig. S5**).



Scheme S1. Chemocatalyzed aldol addition of cinnamaldehyde (**1a**) to hydroxyacetone (HA, **2**).

The reaction containing cinnamaldehyde (5 mmol) and hydroxyacetone (25 mmol) was conducted in 25 mL capped conical-bottom flask with 10 mL total volume, then incubated with recombinant whole cells overexpressing FSAA Q59T (80 g/L, wet cell weight) in citric acid-sodium citrate buffer (0.1 M, pH 6.5) with agitation (220 rpm) at 30 ºC for 48 h. Then the reaction mixture was dissolved in excess methanol and incubated in ice bath for 3 h. Finally, the mixture was extracted with a 2:1 mixture of ethyl acetate-hexane and the product was separated by column chromatography with increasing amounts of ethyl acetate in petroleum ether affording enzymatically catalyzed diol **3a** (639 mg, 3.1mmol, 62% isolated yield)**.** Then the enzymatically catalyzed product **3a** (0.5 g, 2.43 mmol) was dissolved in methanol and treated as described by Acetti and Gatti with dimethoxypropane (0.38 ml) in acetone (5 ml) to get compound **4**, in the presence of a catalytic amount of pyridinium *p*-toluensulfonte at 50 ºC for 4 h.[6] Then compound **4** was separated by column chromatography with increasing amounts of ethyl acetate in petroleum ether and was used to determine the absolute configuration of enzymatically catalyzed **3a** through electronic circular dichroism (ECD) simulation.



(3*S*, 4*R*)-**3a** (4*S*, 5*R*)-**4**

Scheme S2. Derivatization of enzymatically catalyzed product **3a** to afford compound **4** (PPTS: pyridinium *p*-toluensulfonate).

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were obtained using a Jasco J-815 circular dichroism spectrometer (Japan Spectroscopic, Tokyo) at room temperature. CD spectrum for compound **4** was recorded in

MeOH (c= 0.4167 mg mL^{-1}) in a quartz cuvette with a 0.1 cm light path over a wavelength range of 190-425 nm.

Electronic circular dichroism (ECD) simulation:

The absolute configuration of compound **4** was determined by quantum chemical time-dependent density functional theory (TDDFT) calculations of its theoretical ECD spectrum. Firstly, conformational analysis was carried out via Monte Carlo searching using molecular mechanism with MMFF94 force field in the Spartan 08 program (Wavefunction, Inc., Irvine, CA). The results showed four conformers with relative energy below 2.0 kcal/mol. The conformers were reoptimized using density functional theory (DFT) at the B3LYP/6-311++G (2d, 2p) level in vacuum by the Gaussian 09 program.[7] The B3LYP/6-311++G (2d, 2p) harmonic vibrational frequencies were further calculated to confirm their stability. The energies, oscillator strengths, and rotational strengths of the first 10 electronic excitations were calculated using the TDDFT methodology at the B3LYP/6-311++G (2d, 2p) level in vacuum. The ECD spectra for each conformer were simulated by the overlapping Gaussian function (σ=0.40 eV),[8] in which velocity rotatory strengths of the first 8 exited states were adopted. To get the overall ECD spectrum, the simulated spectra of the lowest energy conformers were averaged according to the Boltzmann distribution theory and their relative Gibbs free energy (ΔG).

ECD spectrum of each conformation was simulated according to the overlapping Gaussian functions expressed as:

$$\Delta\varepsilon(E)=\frac{1}{2.296\times10^{-39}\sqrt{\pi}\sigma}\sum_{i}^{A}\Delta E_i R_i e^{[-(E-\Delta Ei)^2/\sigma^2]}$$

Where σ is half the bandwidth at 1/e peak height and expressed in energy units. The parameters ΔE_i and R_i are the excitation energies and rotational strengths for the transition *i*, respectively.

The above function is converted to Δε, λ (wavelength) correlation as:

$$\Delta\varepsilon(\lambda)=\frac{1}{2.296\times10^{-39}\sqrt{\pi}\sigma}\sum_{i}^{A}\Delta E_i R_i e^{[-(1240/\lambda-\Delta Ei)^2/\sigma^2]}$$

and then simulation was accomplished using Excel 2003 and Origin 7.0.

To get the final spectra, all the simulated spectra of conformations of each compound were averaged according to their energy and the Boltzmann distribution theory expressed as:

$$\frac{N_i^*}{N}=\frac{g_i e^{-\varepsilon_i/k_B T}}{\sum g_i e^{-\varepsilon_i/k_B T}}$$

Synthesis of products 3b-3d as standard reference

The reaction containing acceptor (**1b** or **1c**, 200 mM) and hydroxyacetone (1 M) was conducted in 25 mL capped conical-bottom flask with 10 mL total volume, and then incubated with recombinant whole cells overexpressing the FSAA Q59T (80 g/L, wet cell weight) in citric acid-sodium citrate buffer (0.1 M, pH 6.5) with agitation (220 rpm) at 30 °C for 48 h. Then the reaction mixture was dissolved in excess methanol and incubated in ice bath for 3 h. Finally, the mixture was extracted with a 2:1 mixture of ethyl acetate-hexane and the products were separated by column chromatography with

increasing amounts of ethyl acetate in petroleum ether affording corresponding diols **3b** (20 mM, 10% isolated yield) or **3c** (46.8 mM, 23.4% isolated yield).

2-pyridinecarboxaldehyde (**1d**, 300 mM) and hydroxyacetone (1.5 M) were mixed in 25 mL capped conical-bottom flask with 10 mL total volume, and catalyzed by purified enzyme FSAA Q59T (3 mg/mL) in citric acid-sodium citrate buffer (0.1 M, pH 6.5) with agitation (220 rpm) at 30 ℃ for 24 h. Then the reaction mixture was extracted with ethyl acetate and the product was separated by column chromatography with increasing amounts of ethyl acetate in petroleum ether affording diol **3d** (125 mM, 42% isolated yield).

Determination of apparent kinetic parameters

Apparent kinetic parameters for the nucleophile HA and acceptor **1a** including Michaelis constant (K_m^{app}) and maximum velocity (V_{max}^{app}) were determined via a Lineweaver-Burk double reciprocal plot. Analytical scale reactions (total volume was 300 μL) were conducted in Eppendorf tubes (1.5 mL), stirred with a vortex mixer (MS-100, AOSHENG) at 1000 rpm and 30 °C.

Kinetic parameters for cinnamaldehyde. HA (200 mM) was dissolved in citric acid-sodium citrate buffer (0.1 M, pH 6.5) containing 1 mM DTT. Different concentrations of cinnamaldehyde were added into the reaction mixture (5, 10, 20, 30, 45, 60, 100, 150 and 200 mM for wild-type FSAA; 5, 10, 20, 30, 60, 100, 150 and 200 mM for FSAA Q59N; 5, 10, 20, 30, 60, 100, 150 and 200 mM for FSAA Q59T; 10, 20, 30, 45, 60, 100 and 200 mM for FSAA Q59L). To this mixture, the wild-type FSAA, FSAA Q59N, FSAA Q59T, and FSAA Q59L (containing 0.032 to 0.45 mg of protein) were added to start the reaction (total final volume, 300 μL). Conversion was controlled below 5%. Reaction was stopped by adding 900 μL methanol at different time (in a range between 3-5 min), centrifuged, and then analyzed by HPLC.

Kinetic parameters for HA. Cinnamaldehyde (80 mM) was dissolved in citric acid-sodium citrate buffer (0.1 M, pH 6.5) containing 1 mM DTT. Different concentrations of HA were added into the reaction mixture (20, 30, 80, 100, 150, 200 and 250 mM for FSAA wild-type; 20, 30, 40, 60, 80, 100, 150, 250 and 300 mM for FSAA Q59N; 30, 40, 60, 80, 100 and 150 mM for FSAA Q59T; 20, 40, 60, 80, 150 and 250 mM for FSAA Q59L). To this mixture, wild-type FSAA, FSAA Q59N, FSAA Q59T or FSAA Q59L (containing 0.032 to 0.45 mg of protein) was added to start the reaction (total final volume, 300 μL). Conversion was controlled below 5%. Reaction was stopped by adding 900 μL methanol at different time (in a range between 3-5 min) and centrifuged, and then subsequently analyzed by HPLC.

Analytical methods

The enzyme activity for **1a-1c** was calculated from the results of HPLC using a C18 column and a UV detector at 254 nm. The mobile phase was composed of methanol and water (56/44, v/v) at a flow rate of 1 mL min^{-1}.

For detection of activity towards **1d**, HPLC analysis was performed on a CHIRALPAK AD-RH, 5μm, 4.6×150 mm column at 254 nm. The mobile phase was composed of acetonitrile and water (20/80, v/v) at a flow rate of 0.6 mL min^{-1}.

The diastereomeric ratio (dr) and enantiomeric excess (ee) of the product **3a** were analyzed via HPLC using an AD-H chiral column and a UV detector at 254 nm. The mobile phase was composed of *n*-hexane and *iso*-propanol (90/10, v/v) at a flow rate of 1 mL min^{-1}.

High-performance liquid chromatography-tandem mass spectrometry

The high-performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS) of **3a-3d** was conducted in ZHEJIANG NHU CO.,LTD (Zhejiang, China) using a Q Exactive mass spectrometer (Thermo Fisher Scientific) with electrospray ionization (ESI) as ionization source. In positive ion mode, the electrospray ionization voltage was 3700 V with sheath gas setting at 35 psi and aux gas flow at 15 arb. In negative ion mode, the electrospray ionization voltage was -3200 V with sheath gas setting at -40 psi and aux gas flow at -20 arb. The sweep gas pressure was set at 0 psi, the capillary temperature was set at 320 ^{o}C, and the heater temperature was set at 200 ^{o}C . Scan Mode was set as full MS/dd-MS^2, and the scan range was 80-1200 m/z. The isolation window was set at 1.0 m/z. The normalized collision energy (NCE) was set at 30 eV with the stepped NCE setting at 50 %.

HPLC was performed on U3000 (Thermo Fisher Scientific) using an Amethyst C18-H column (5μm, 4.6×150 mm) at 254 nm for **3a-3c** separation and a CHIRALPAK AD-RH column (5μm, 4.6×150 mm) at 254 nm for **3d** separation. The mobile phase of **3a-3c** separation was composed of methanol and water (56/44, v/v) at a flow rate of 1 mL min^{-1}. The mobile phase of **3d** separation was composed of acetonitrile and water (20/80, v/v) at a flow rate of 0.6 mL min^{-1}.

Computational methods

Molecular docking towards D-fructose-6-phosphate

Based on the structural analysis and the proposed mechanism of FSAA, automatic molecular docking was carried out to predict the possible position and orientation of D-fructose-6-phosphate in protein binding site. In order to assure productive docking geometry, several modifications on substrate and enzyme parameters were carried out.[9] The δ values (atom diameter) of NZ within Lys 85 and substrate carbonyl carbon atom (CT) were reduced (to 1.4 and 1.5 Å, respectively) to assure a reasonable intermolecular distance. In practice, increased vdw well depth (ε) was applied to NZ (ε=70 kcal/mol) and CT (ε=30 kcal/mol) to enhance the attractive interaction between NZ and CT. Additionally, the catalytic water was treated as a part of the protein. Finally, we used Autodock 4.0 (AD4) program to dock the ligand into FSAA with the yielded docking energy.

Complex for molecular dynamic simulation

The initial coordinates of the wild-type *E.coli* FSAA were obtained from the protein data bank (PDB ID: 1L6W). In the crystal, the enzyme forms a decameric structure by two packed pentamers.[10] An important interaction between subunits in each pentamer involves the C-terminal helix (residues 199-215), which runs across the active site of the neighboring subunit. So a large fragment from chain A with residues 1-199 and a small fragment from chain B with residues 199-220 were used to build the minimal functional unit for calculation.[11] Models of the mutants were constructed and energy-minimized in silico using Swiss-Pdb Viewer.[12] To obtain productive binding models, the hemiaminal intermediate of HA and acceptor **1a** was covalently bound to Lys85. In order to function as a nucleophile, Lys85 was treated as uncharged amine Lyn85.[13] Nonpolar hydrogen atoms were added to the enzyme structure using the AMBER 11 simulation package.[14] Counterions of Na^{+} were used to

neutralize the system. The whole system was immersed in a rectangle box of TIP3P water molecules. The box was extended 12 Å in all three dimensions from the edges of the protein structure.

Molecular dynamics simulation

Protein-transition state complexes of the enzymes and the model hemiaminal intermediates were constructed for wild-type FSAA and mutants (FSAA Q59L and FSAA Q59T). The distances between the oxygen of the catalytic water and Tyr131 as well as Thr109 were restrained to 3.0 Å to form hydrogen bonds. The MD simulations were conducted as described below: 2500 cycles of steepest descent and 2500 cycles of conjugate gradient minimization were first applied to remove potential collision contacts with the added water molecules. The system was then gradually heated to 303 K over the course of 50 ps followed by a 50 ps of density equilibration at a constant temperature. During these two steps, only water molecules were allowed to move freely, and the rest parts of the system were restrained. The whole system was allowed to move after density equilibration. Then, a 1 ns of constant-pressure and constant-temperature (NPT) ensemble simulation was conducted to equilibrate the system. Finally, the simulation was continued at 303 K for another 5ns (or longer for sufficient equilibrium) production run in the NPT ensemble. The temperature was maintained by coupling to the Langevin thermostat. The SHAKE algorithm was used to constrain bond distances of the hydrogen atoms. The nonbonding cutoff distance was set to 8.0 Å and simulation snapshots were saved every 2 ps for analysis. The ptraj program in the AMBER package was used to analyze the conformation of each simulation. The calculation of the binding free energy of each complex and the decomposition of the binding free energy into individual residues were evaluated using the *mmpbsa* program in AMBER.

Aldol addition of cinnamaldehyde (1a) to hydroxyacetone (HA, 2) catalyzed by whole cells of recombinant *E. coli* overexpressing variants of FSAA.

The activity of whole-cell catalysts (total volume 300 µL) were determined at 30 °C in citric acid-sodium citrate buffer (0.1 M, pH 6.5) with agitation (220 rpm). Concentrations of acceptor substrate **1a** were 250 mM and 500 mM, and the corresponding concentrations of donor substrate **2** were 1.25 M and 2.5 M, respectively. The catalyst loading was 80 g/L (wet cell weight) resting cells. Reaction was stopped by adding 900 µL methanol at different time, centrifuged, and then analyzed by HPLC as described above.

Supplementary Tables

Table S1 Energy analysis for the ECD of compound **4**.

Conformation	Gibbs free energy (298.15 K)		
	G (Hartree)	ΔG (Kcal/mol)	Boltzmann Distribution
C1	-808.090882	0.00062751	0.388754
C2	-808.090289	0.37274094	0.207384
C3	-808.090883	0	0.389166
C4	-808.087791	1.94026092	0.014696

Table S2 ECD Data for compound **4**.

State	C1		C2		C3	
	Excitation energies(eV)	Rotatory Strengths*	Excitation energies(eV)	Rotatory Strengths*	Excitation energies(eV)	Rotatory Strengths*
1	4.2815	41.1277	4.1207	3.7981	4.2816	41.1045
2	4.578	-48.3188	4.6028	25.5158	4.5778	-48.2563
3	4.6921	-2.8409	4.7327	-9.2718	4.6919	-2.8459
4	4.8492	-9.9667	4.7979	3.6747	4.8491	-9.9352
5	5.353	-0.7691	5.0386	-3.2971	5.3528	-0.766
6	5.3706	0.9464	5.45	2.4301	5.3706	0.9403
7	5.5177	3.7203	5.5466	1.7473	5.5176	3.7089
8	5.5302	3.0246	5.5742	2.4711	5.5302	3.0325
9	5.5953	-4.5905	5.6164	-16.0053	5.5953	-4.5798
10	5.6821	11.2163	5.6485	-0.595	5.682	11.1875

* R(velocity) 10**-40 erg-esu-cm

State	C4				
	Excitation energies(eV)	Rotatory Strengths*	State	Excitation energies(eV)	Rotatory Strengths*
1	4.3471	37.3742	6	5.3933	-0.6445
2	4.5937	-24.632	7	5.4577	-2.6554
3	4.7171	14.6871	8	5.5768	0.0907
4	4.828	7.1706	9	5.6435	0.9899
5	5.2148	2.9226	10	5.6982	12.0671

* R(velocity) 10**-40 erg-esu-cm

Table S3 PCR primers used for gene amplification and mutant construction.

Oligonucleotides	Primers[a]
FSAA T109S	TTCCGACGCTGGGATCGGCGGTATATGGCGCAGCACAAG
	GCGCCATATACCGCCGATCCCAGCGTCGGAATCCCTTCC
FSAA Y131F	AATATGTTGCGCCTTTTGTTAATCGTATTGATGCTCAGG
	TCAATACGATTAACAAAAGGCGCAACATATTCCGCACCT
FSAA Q59N	GGCGTCTGTTTGCCAACGTAATGGCTACCACTGCCGAAG
	GTGGTAGCCATTACGTTGGCAAACAGACGCCCCTGACCG
FSAA Q59E	GGCGTCTGTTTGCCGAAGTAATGGCTACCACTGCCGAAG
	GTGGTAGCCATTACTTCGGCAAACAGACGCCCCTGACC
FSAA Q59T	GGCGTCTGTTTGCCACCGTAATGGCTACCACTGCCGAAG
	GTGGTAGCCATTACGGTGGCAAACAGACGCCCCTGACCG
FSAA Q59L	GGCGTCTGTTTGCCCTGGTAATGGCTACCACTGCCGAAG
	GTGGTAGCCATTACCAGGGCAAACAGACGCCCCTGACCG
FSAA Q59H	GGCGTCTGTTTGCCCATGTAATGGCTACCACTGCCGAAG
	GTGGTAGCCATTACATGGGCAAACAGACGCCCCTGACCG
FSAA Q59Y	GGCGTCTGTTTGCCTATGTAATGGCTACCACTGCCGAAG
	GTGGTAGCCATTACATAGGCAAACAGACGCCCCTGACCG

[a] The mutated sites are underlined.

Table S4 Wild-type FSAA and mutants catalyzed aldol reaction of cinnamaldehyde (**1a**) with HA (**2**) for initial screening.

Catalyst	v_0/nmol min^{-1} mg^{-1}	dr[a] (syn: anti)	ee[b] (syn %)
FSAA WT	23.3 ± 0.27	97 : 3	> 99%
FSAA T109S	19.1 ± 0.12	_[c]	_[c]
FSAA Y131F	10.9 ± 0.03	_[c]	_[c]
FSAA Q59E	18.5 ± 0.06	90 : 10	> 99%
FSAA Q59N	76.4 ± 0.8	94 : 6	> 99%
FSAA Q59L	238 ± 3.9	95 : 5	> 99%
FSAA Q59T	236 ± 1.2	93 : 7	> 99%
FSAA Q59Y	27.7 ± 0.06	89 : 11	> 99%
FSAA Q59H	20.2 ± 0.02	97 : 3	> 99%

[a] dr, diastereoselectivity ratio;

[b] ee, enantiomeric excess;

[c] Not checked.

Table S5 Specific activity of wild-type FSAA and selected FSAA variants towards D-fructose-6-phosphate.

Protein	Specific activity(U mg^{-1})[a]
Wild-type FSAA	3.5 ± 0.02
FSAA Y131F	n.d.[b]
FSAA Q59N	1.1 ± 0.1
FSAA Q59L	0.26 ± 0.03
FSAA Q59T	0.5 ± 0.01

[a]One Unit (U) of FSAA cleaves 1 μmol of D-fructose-6-phosphate per min to afford D-glyceraldehyde-3-phosphate and dihydroxyacetone in glycylglycine buffer (1 mL, 50 mM, pH 8.5) at room temperature.

[b] n.d., not detected (specific activity ≤ 0.01 U mg^{-1}).

The specific activity of the wild-type FSAA and selected mutants towards D-fructose-6-phosphate (D-F6P) was determined using the assay described above. Unlike the increased activity towards acceptors **1a-1d**, all of these mutants showed decreased activities towards D-F6P. Mutant Y131F was completely inactive against D-F6P, which is in accordance with the previous study.[2] Docking was performed to give insights into the difference between the wild type and the variants (**Fig. S7**).

As indicated by the model of the wild type, hydrogen bonds were formed between D-F6P and key residues (Asp6, Asn28 and Tyr131), respectively, for substrate stabilization. Between the catalytic water and D-F6P, a hydrogen bond was also found. Residues Arg134 and Ser166 appeared to accommodate D-F6P, as ionic interactions were revealed between these two residues and the phosphate group of D-F6P, respectively.

Compared with the aldehydes **1** containing a hydrophobic benzene ring in this study, D-F6P is much more hydrophilic. When Gln was mutated into hydrophobic Leu in Q59L, the orientation of the hydroxyl groups in D-F6P became quite different from that in the wild type. As a result, these hydrogen bonds failed to form. This may contribute to the decreased activity of Q59L towards D-F6P.

In Q59T, a more outstretched conformation was indicated by molecular dynamic simulation towards acceptor **1a**. On one hand, the hydrogen bond between Asn28 and D-F6P failed to form, as the side chain of Asn28 was closer to the polar Thr59. On the other hand, D-F6P was more inclined to get close to the entrance as it could be stabilized by the polar residues located at the enlarged entrance, which was not beneficial for the formation of the effective covalent bond between D-F6P and the catalytic Lys85.

Table S6 Binding free energies of the enzyme-substrate complexes for the wild-type FSAA and selected mutants.

Entry	Enzyme	$\triangle G_{binding}$(kcal mol^{-1})[a]
1	FSAA-WT	-11.7±3.7
2	FSAA-Q59L	-16.9±2.0
3	FSAA-Q59T	-12.6±3.3

[a] The binding free energy is calculated by the *mmpbsa* program of AMBER software package.

Table S7 Decomposition of binding free energy for individual residues in the wild-type FSAA and selected mutants

Key residue number	Residue	Total energy (kcal mol^{-1})		
		FSAA WT	FSAA Q59L Gln/Leu[59]	FSAA Q59T Gln/Thr[59]
28	Asn	-1.55 ± 6.38	-2.04 ± 5.23	-0.18 ± 8.62
59	Gln	0.08 ± 4.58	-0.31 ± 4.48	-0.15 ± 6.56
85[a]	Lys	2.12 ± 6.68	1.01 ± 7.62	0.92 ± 7.95
109	Thr	2.73 ± 5.07	1.46 ± 5.65	1.22 ± 5.88
131	Tyr	2.91 ± 5.71	0.28 ± 4.04	-0.19 ± 5.24
207	Phe	-0.03 ± 4.73	-0.28 ± 3.39	-0.04 ± 5.04

[a] The catalytic lysine

Table S8 Calculation of distance between the αC of residues involved in the interface interaction of the wild-type FSAA and mutant FSAA-Q59T.

Enzyme	Distance (Å)		
	:29@CA-:207@CA	:30@CA-:210@CA	:33@CA-:214@CA
FSAA WT	6.28 ± 0.40	5.98 ± 0.50	5.42 ± 0.26
FSAA Q59T	8.93 ± 1.03	7.24 ± 0.97	6.39 ± 0.99

Supplementary Figures

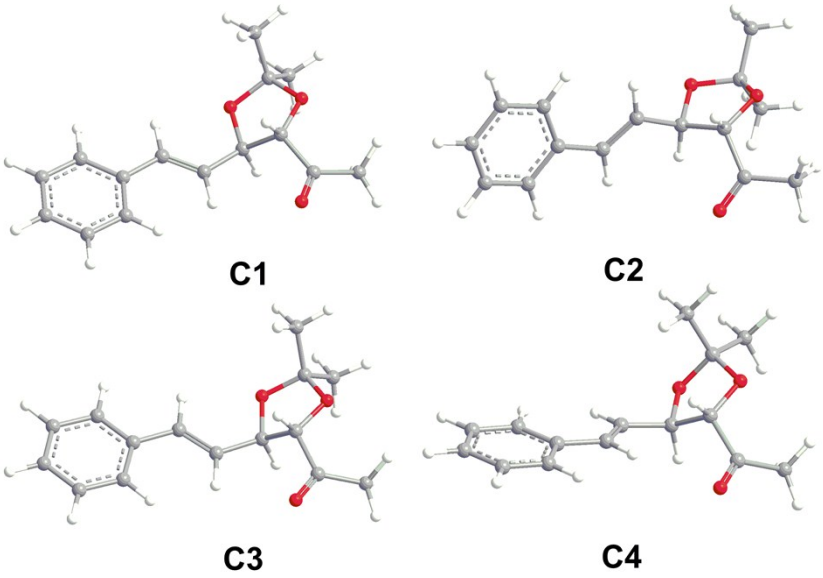


Fig. S1 Optimized lowest energy 3D conformers of compound **4** by Gaussian 09 program at the B3LYP/6-311++G (2d, 2p) level in vacuum.

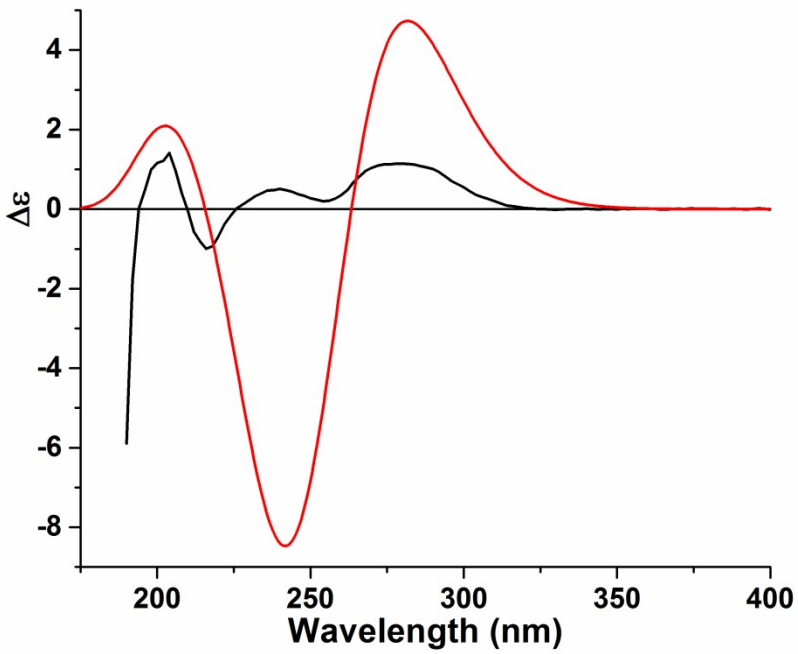


Fig. S2 Experimental (black line) and B3LYP/6-311++G (2d, 2p)//B3LYP/6-311++G (2d, 2p) calculated (red line) ECD spectra of **4**.

In the 175-425 nm range (Fig. **S2**), both the experimental and the theoretical ECD spectra showed first positive and last positive cotton effects at around 205 nm and 280 nm, respectively. Qualitative analysis of the calculated and experimental ECD spectra allowed the assignment of the absolute configurations of **4** as 4*S*5*R*. So the absolute configuration of enzymatically catalyzed product **3a** was confirmed as 3*S*4*R*.

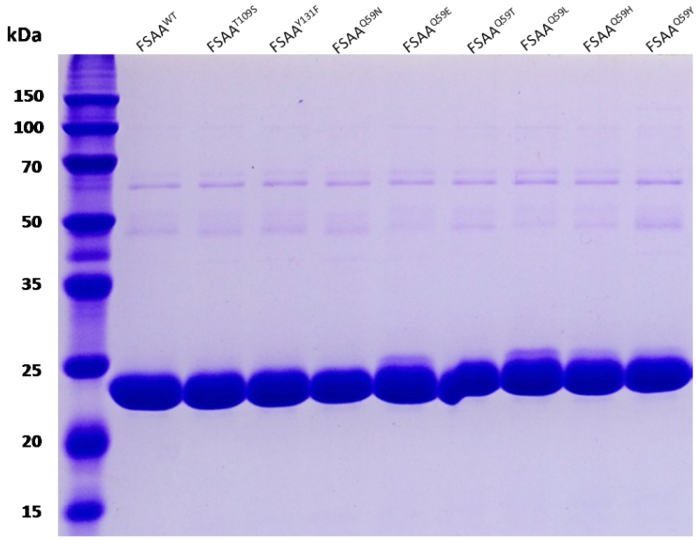
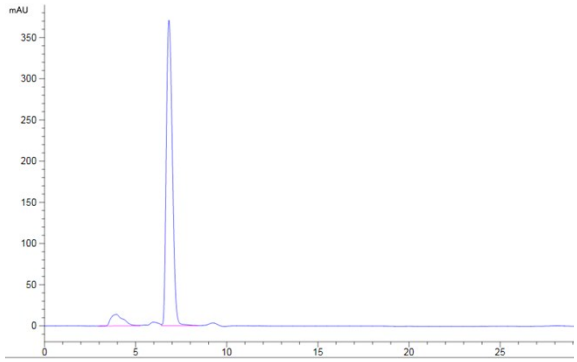


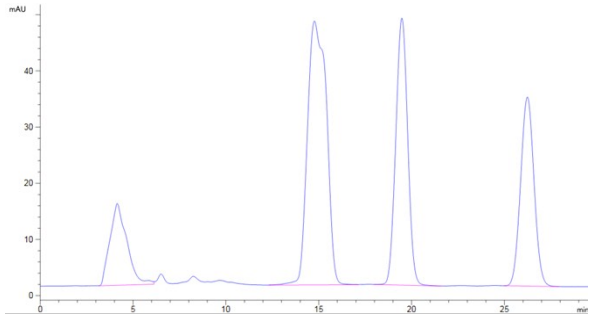
Fig. S3 SDS-PAGE analysis of the purified wild-type FSAA and selected mutants.

The full-length sequence of gene *fsaA* was 663 bp, encoding a 220-amino acid polypeptide with a calculated molecular mass of 23 kDa. The purified enzymes with an N-terminal His_6-tag migrated as a single band with a molecular weight of approximately 24 kDa on SDS-PAGE (**Fig. S3**), consistent with the theoretical molecular mass of the His_6-tagged protein.

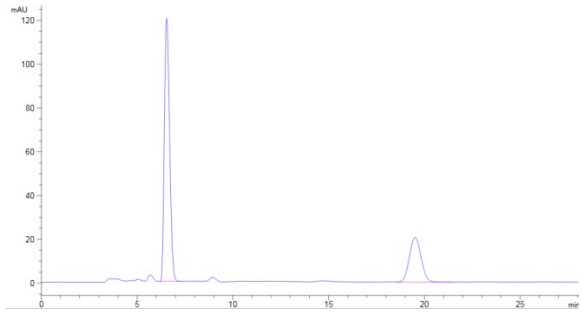
Cinnamaldehyde (substrate **1a)**:



Products (aldol addition of **1a** to **2** catalyzed by NaOH):



Wild-type FSAA:



FSAA Q59L:

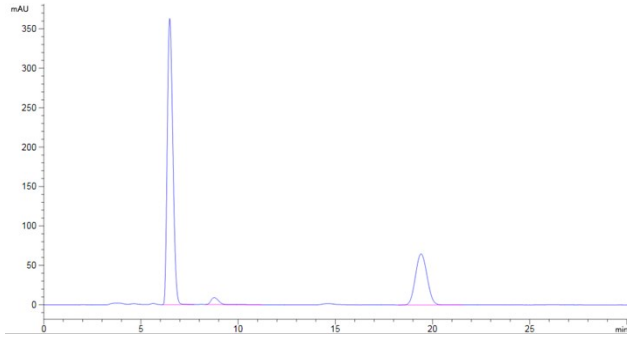
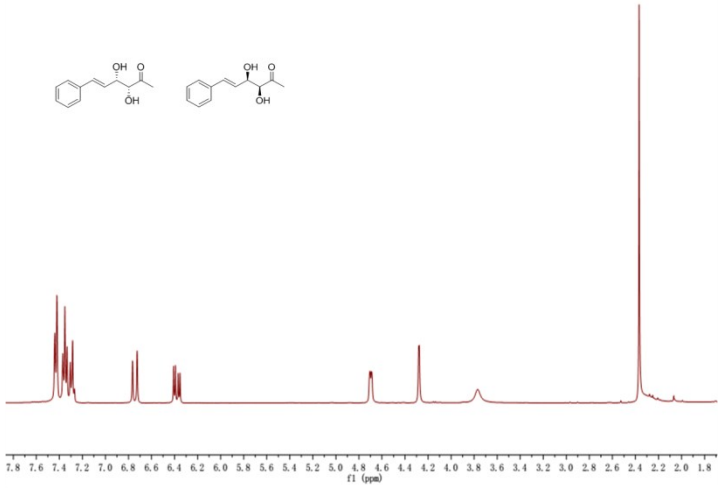


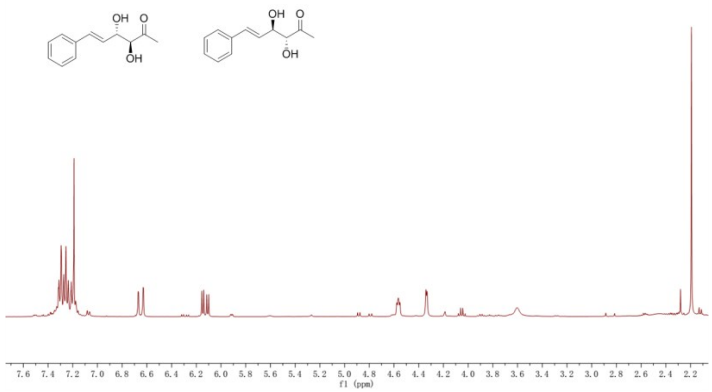
Fig. S4 HPLC analysis (chiral AD-H column) of the product catalyzed by the wild-type FSAA and mutant FSAA Q59L (concentrations of cinnamaldehyde and hydroxyacetone were 10 mM and 50 mM, respectively). Retention time: cinnamaldehyde, 6.8 min; anti-products, 15 min (inseparable); syn-products, 19.5 min and 26.4 min.

Fig. S5 Spectra of products **3a**-**3d** and compound **4**.

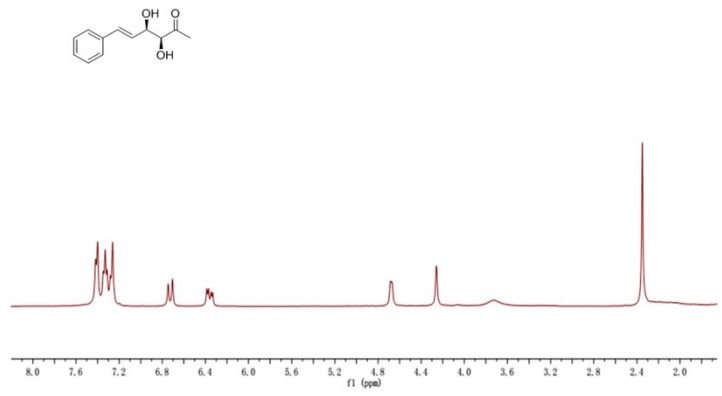
Syn-**3a**, ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.43 – 7.24 (m, 5H, Ar*H*), 6.72 (d, J = 15.9 Hz, 1H, ArC*H*=CH), 6.35 (dd, 1H, ArCH=C*H*), 4.67 (d, J = 6.1 Hz, 1H, =CHC*H*OH), 4.25 (d, J =1.6 Hz, 1H, COC*H*OH), 2.34 (s, 3H, COCH_3).



Anti-**3a**, ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.32 – 7.14 (m, 5H, Ar*H*), 6.66 (d, J = 16.0 Hz, 1H, ArC*H*=CH), 6.13 (dd, 1H, ArCH=C*H*), 4.61 (m, 1H, =CHC*H*OH), 4.34 (d, J = 3.5 Hz, 1H, COC*H*OH), 2.19 (s, 3H, COCH_3).

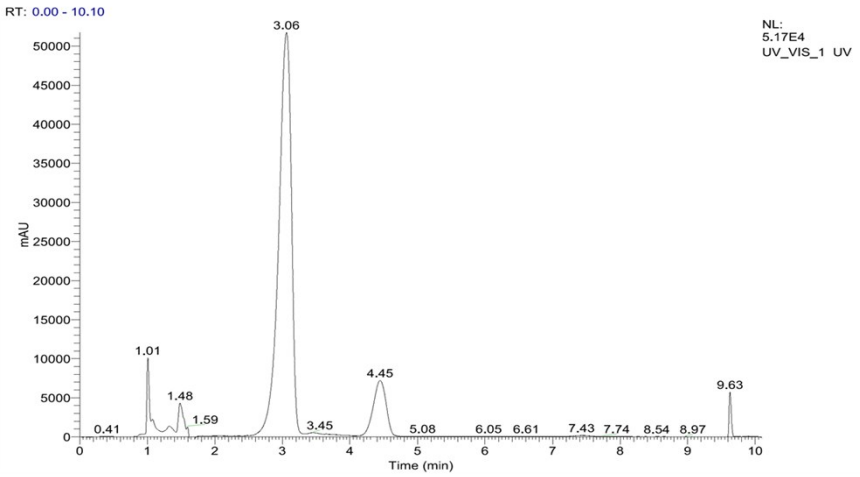


(3*S*, 4*R*)-**3a**, ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.44 – 7.24 (m, 5H, Ar*H*), 6.73 (d, *J* = 15.9 Hz, 1H, ArC*H*=CH), 6.36 (dd, ArCH=C*H*), 4.68 (d, J = 6.4 Hz, 1H, =CHC*H*OH), 4.26 (s, 1H, COC*H*OH), 2.35 (s, 3H, COCH_3).

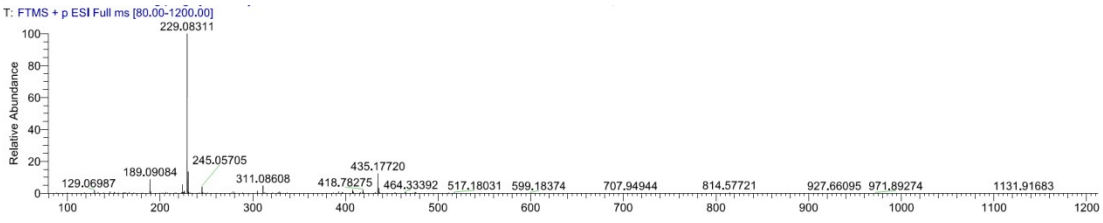


HPLC-MS/MS spectrometry of (3*S*, 4*R*)-**3a**

a) Measured HPLC spectrum of (3*S*, 4*R*)-**3a**. Retention time: 3.06 min.



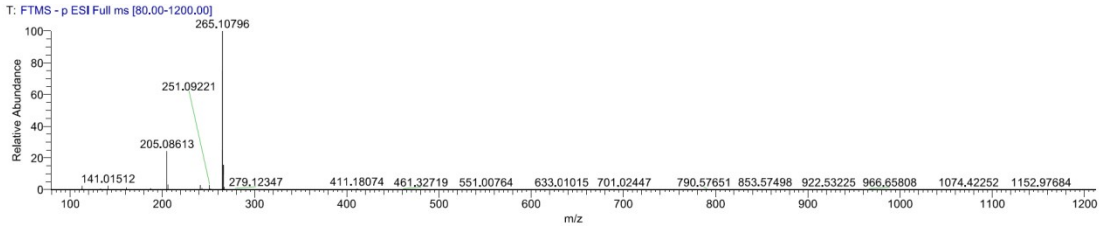
b) Measured HPLC-MS/MS spectrum of (3*S*, 4*R*)-**3a**.



Elemental composition search on mass 229.08311

m/z= 224.08311-234.08311

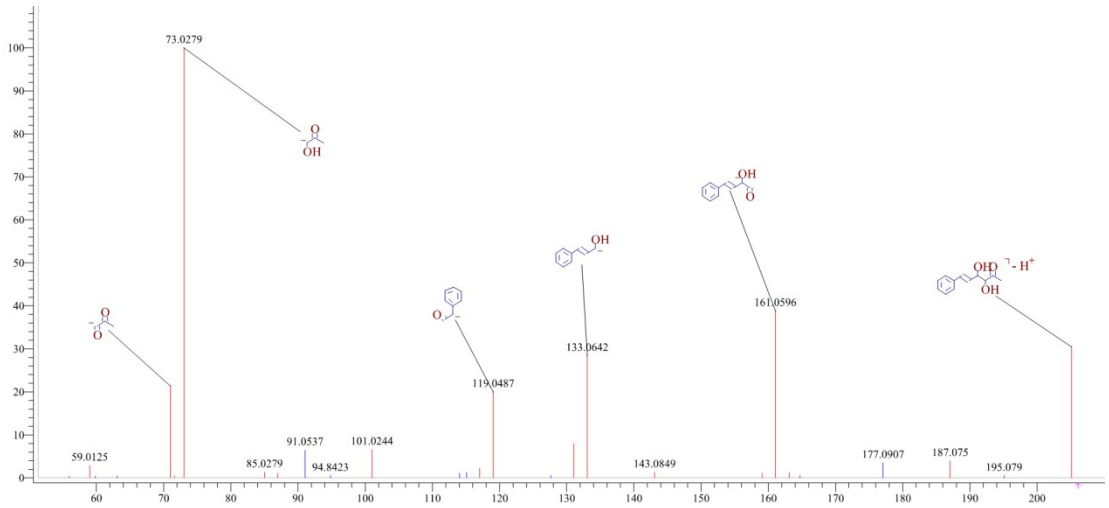
m/z	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
229.08311	229.08352	-1.77	5.5	$C_{12}H_{14}O_3Na$



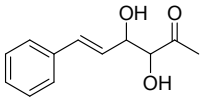
Elemental composition search on mass 205.08613

m/z= 200.08613-210.08613

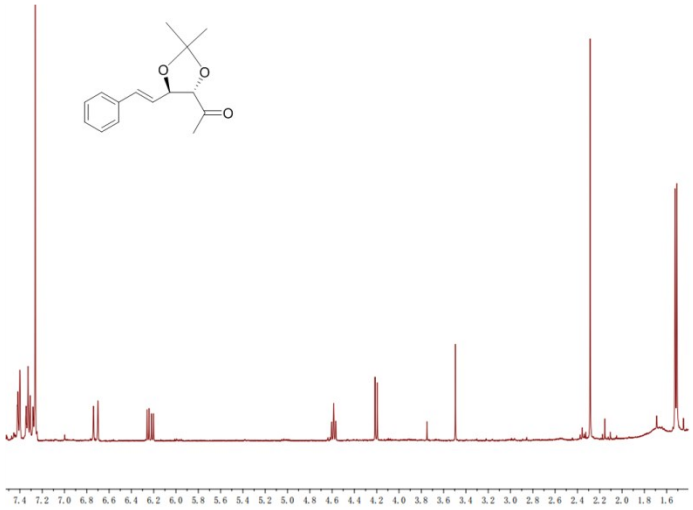
m/z	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
205.08613	205.08702	-4.33	6.5	$C_{12}H_{13}O_3$



The structure of the peak at 3.06 min was indicated to be:

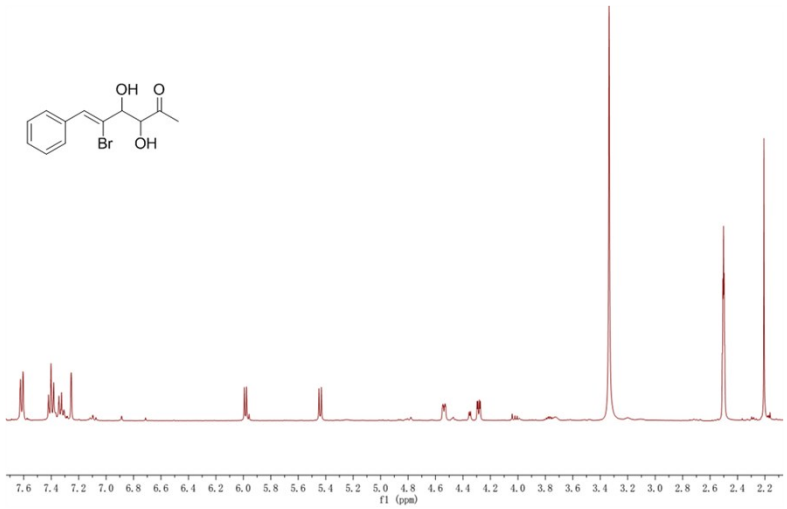


(4*S*, 5*R*)-**4**, ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.43 – 7.29 (m, 5H, Ar*H*), 6.72 (d, *J* = 15.8 Hz, 1H, ArC*H*=CH), 6.23 (dd, 1H, ArCH=C*H*), 4.59 (t, 1H, =CH-C*H*-O), 4.20 (d, *J* = 7.8 Hz, 1H, CO-C*H*-O), 2.29 (s, 3H, COCH_3), 1.52 (s, 3H, CCH_3), 1.51 (s, 3H, CCH_3).



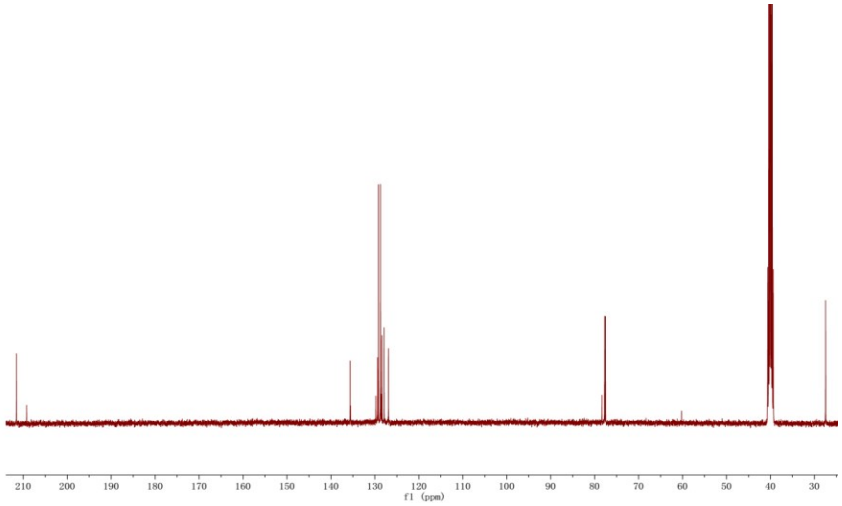
Biosynthetic **3b**. a) ^{1}HNMR, b) ^{13}CNMR and c) HPLC-MS/MS spectrometry

a)



^{1}H NMR (400 MHz, DMSO): δ 7.65-7.3 (m, 5H, Ar*H*), 7.25 (s, 1H, ArC*H*=CBr), 5.98 (d, J = 6.3 Hz, 1H, =CBrCHO*H*), 5.44 (d, J = 6.9 Hz, 1H, COCHO*H*), 4.54 (dd, 1H, =CBrC*H*OH), 4.29 (dd, 2.5 Hz, 1H, COC*H*OH), 2.21 (s, 3H, COCH_3).

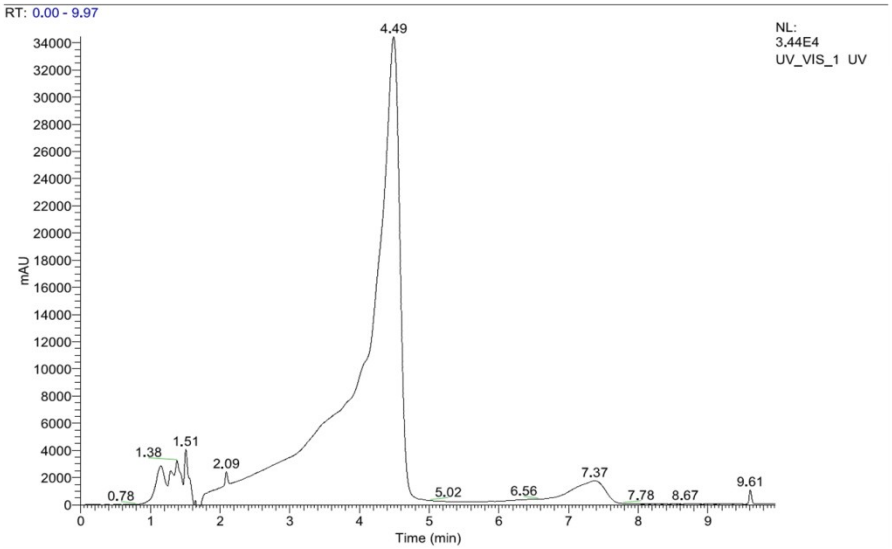
b)



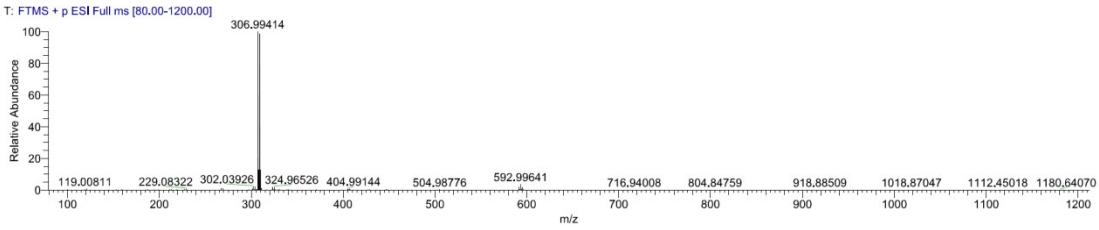
^{13}C NMR (101 MHz, DMSO): δ 27.47, 77.63, 78.35, 126.40, 126.9, 127.89, 128.23, 128.38, 128.71, 129.29, 135.58, 211.54.

c) HPLC-MS/MS spectrometry of **3b**

i) Measured HPLC spectrum of **3b**. Retention time: 4.49 min.



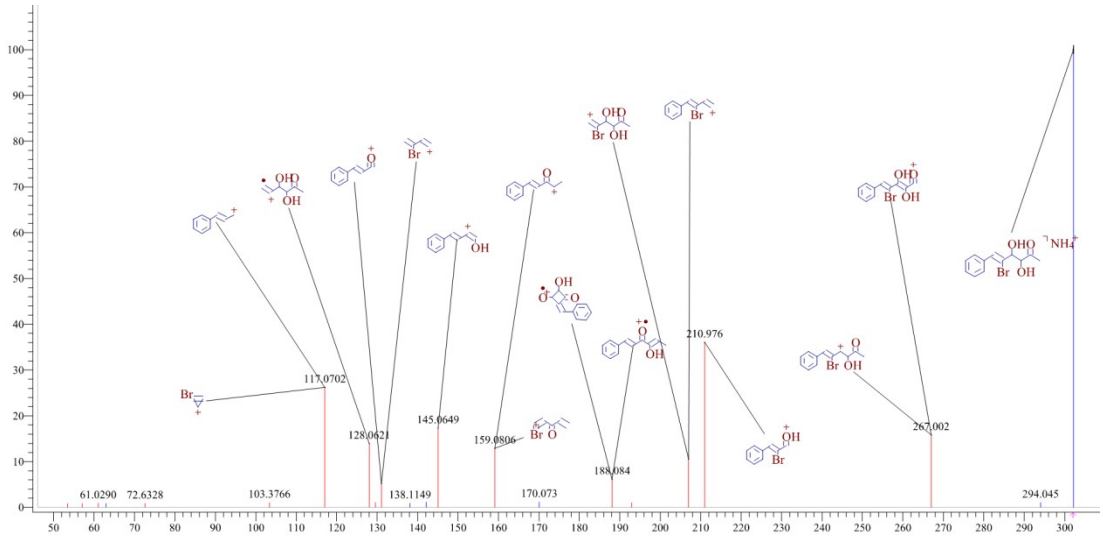
ii) Measured HPLC-MS/MS spectrum of **3b**.



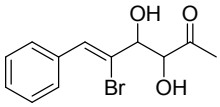
Elemental composition search on mass 306.99414

m/z= 301.99414-311.99414

m/z	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
306.99414	306.99403	0.36	5.5	$C_{12}H_{13}O_3BrNa$

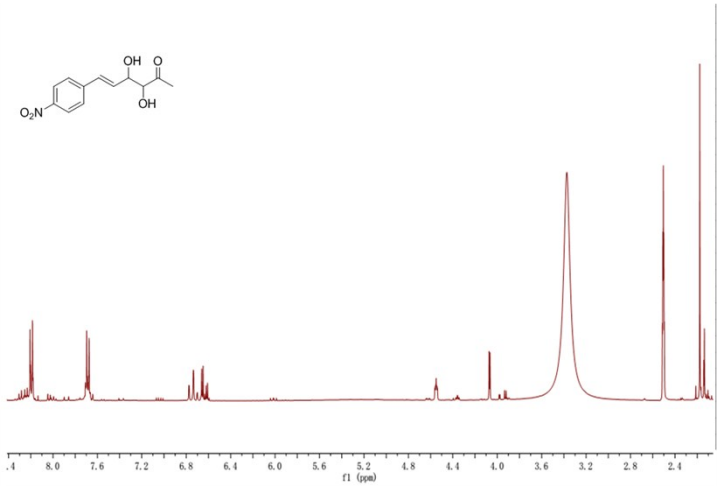


The structure of the peak at 4.49 min was indicated to be:



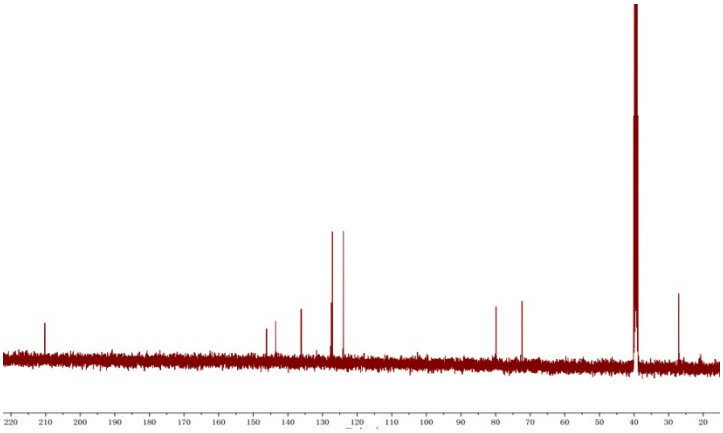
Biosynthetic **3c**. a) ^{1}HNMR, b) ^{13}CNMR and c) HPLC-MS/MS spectrometry

a)



^{1}H NMR (400 MHz, DMSO): δ 8.23 – 8.16 (m, 2H, Ar*H*), 7.73 – 7.64 (m, 2H, Ar*H*), 6.76 (d, J = 16.1 Hz, 1H, ArCH=C*H*), 6.65 (d, J = 5.1 Hz, 1H, ArC*H*=CH), 4.55 (m, 1H, =CHC*H*OH), 4.07 (d, J = 3.3 Hz, 1H, COC*H*OH), 2.17 (s, 3H, COCH_3).

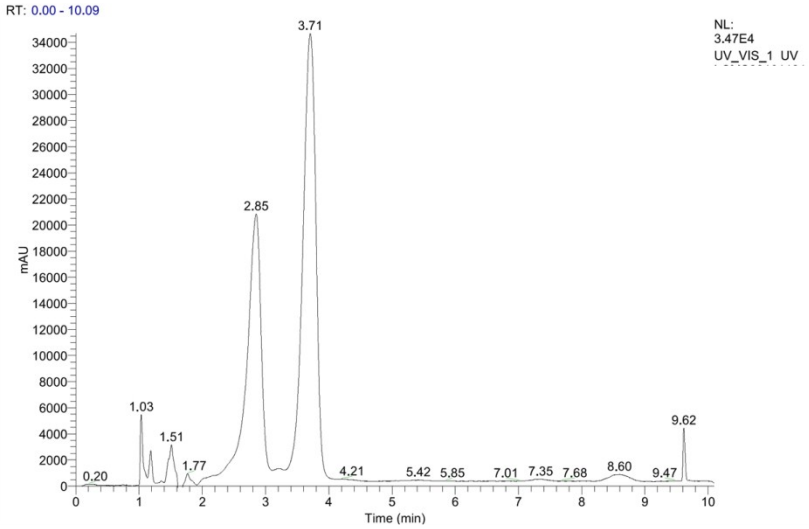
b)



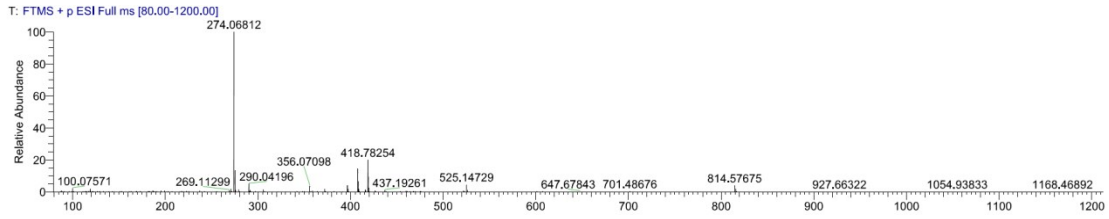
^{13}C NMR (101 MHz, DMSO): δ 27.11, 72.36, 79.89, 123.96, 127.14, 127.46, 136.15, 143.56, 146.19, 210.23.

c) HPLC-MS/MS spectrometry of **3c**

i) Measured HPLC spectrum of **3c**. Retention time: 2.85 min.



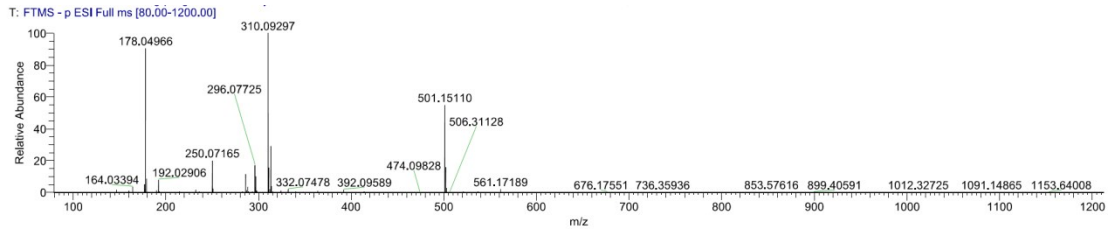
ii) Measured HPLC-MS/MS spectrum of **3c**.



Elemental composition search on mass 274.06812

m/z= 269.06812-279.06812

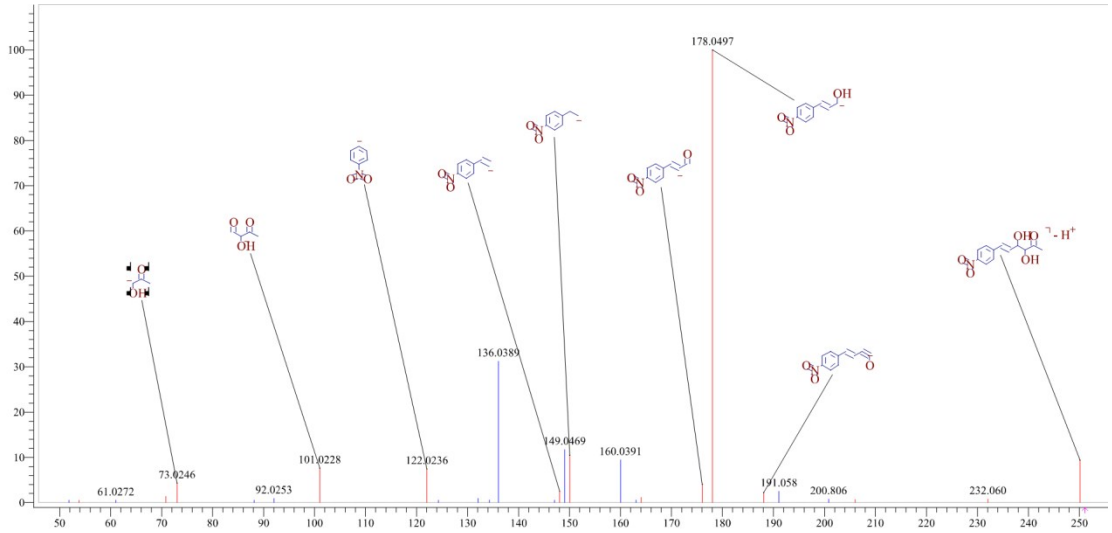
m/z	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
274.06812	274.06859	-1.73	6.5	$C_{12}H_{13}O_5NNa$



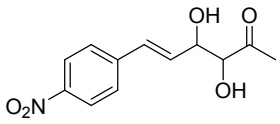
Elemental composition search on mass 250.07165

m/z= 245.07165-255.07165

m/z	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
250.07165	250.07210	-1.78	7.5	$C_{12}H_{12}O_5N$

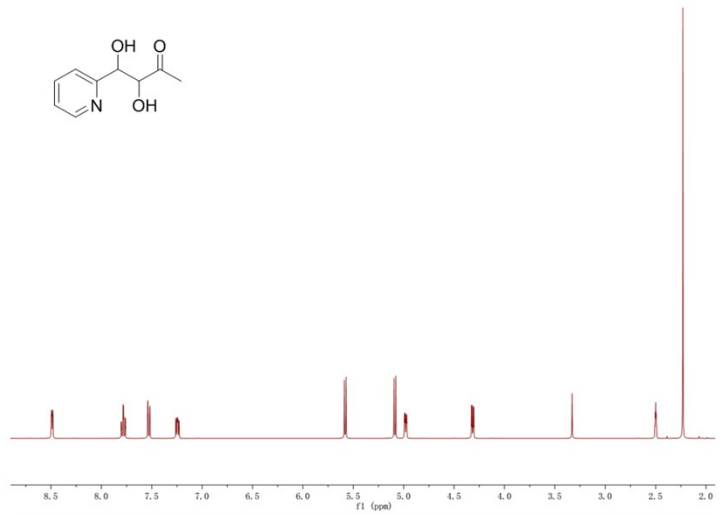


The structure of the peak at 2.85 min was indicated to be:



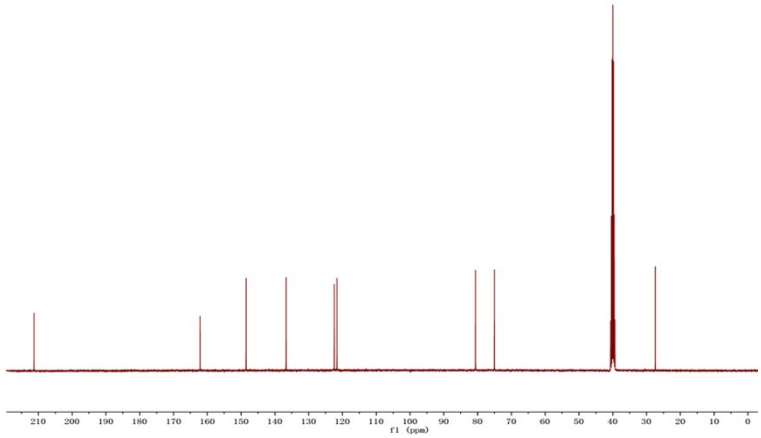
Biosynthetic **3d**. a) ^{1}HNMR, b) ^{13}CNMR and c) HPLC-MS/MS spectrometry

a)



^{1}H NMR (400 MHz, DMSO): δ 8.49 (m, 1H, Ar*H*), 7.78 (m, 1H, Ar*H*), 7.53 (d, J = 7.9 Hz, 1H, Ar*H*), 7.24 (m, 1H, Ar*H*), 5.58 (d, J = 6.7 Hz, 1H, ArCHO*H*), 5.09 (d, J = 7.0 Hz, 1H, COCHO*H*), 4.98 (dd, 1H, ArC*H*OH), 4.32 (dd, 1H, COC*H*OH), 2.23 (s, 3H, COCH_3).

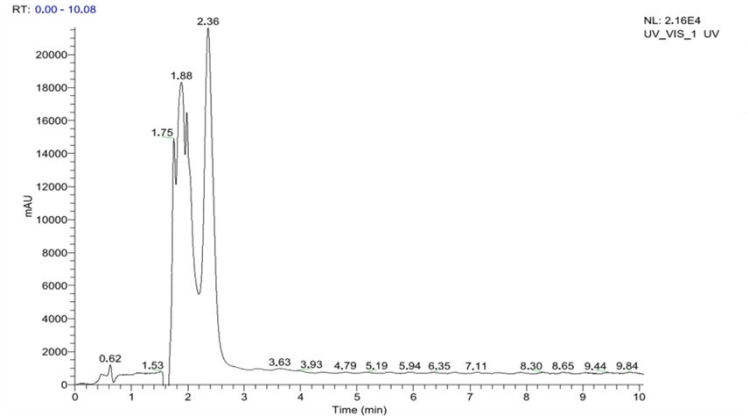
b)



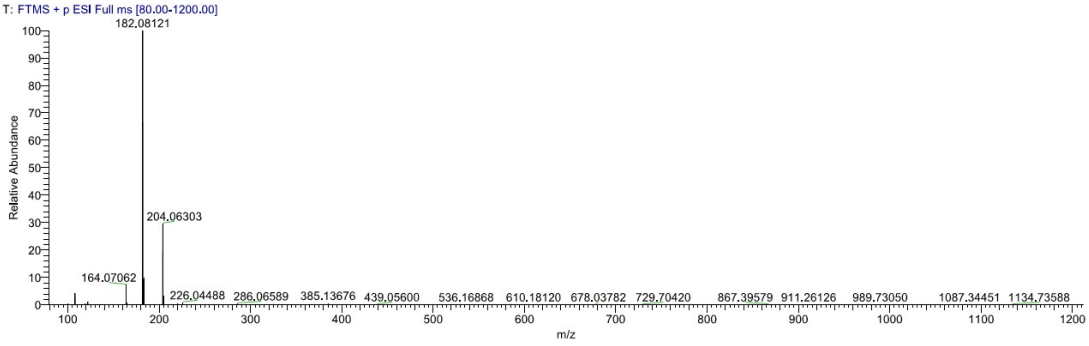
^{13}C NMR (101 MHz, DMSO): δ 27.41, 75.03, 80.64, 121.64, 122.43, 136.68, 148.53, 162.14, 211.3.

c) HPLC-MS/MS spectrometry of **3c**

i) Measured HPLC spectrum of **3d**. Retention time : 2.36 min.



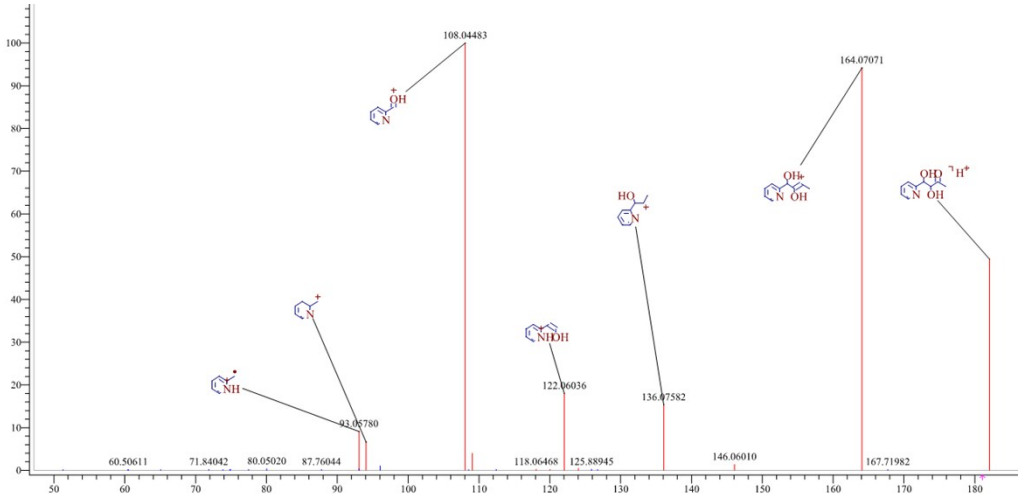
ii) Measured HPLC-MS/MS spectrum of **3d**.



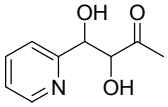
Elemental composition search on mass 182.08121

m/z= 177.08121-187.08121

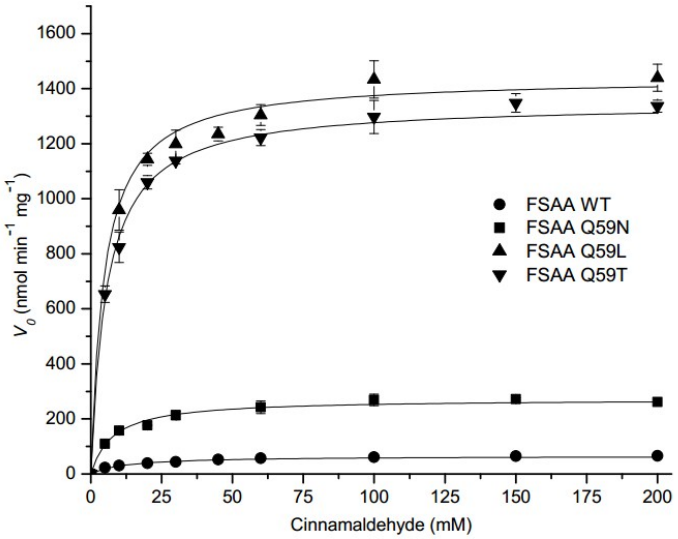
m/z	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
182.08121	182.08117	0.22	4.5	$C_9H_{12}O_3N$



The structure of the peak at 2.36 min was indicated to be:



(a)



(b)

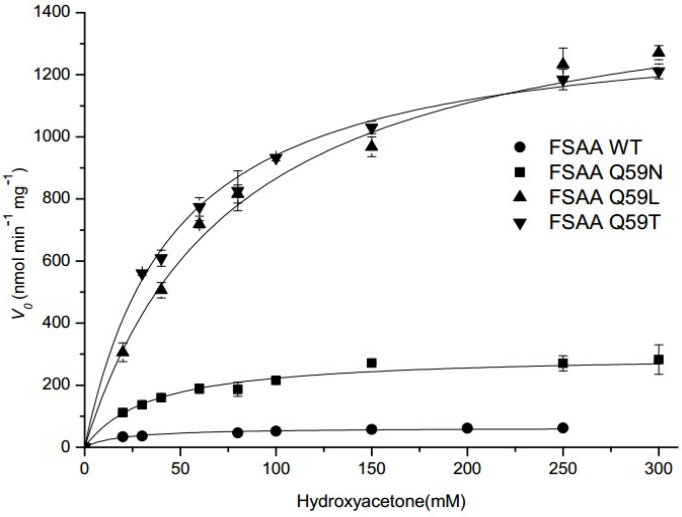
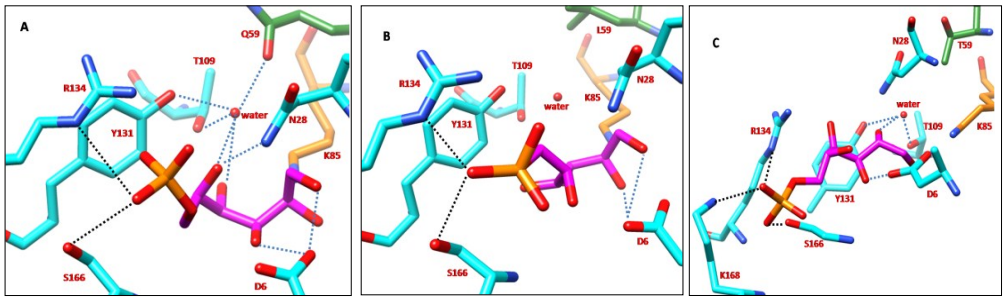


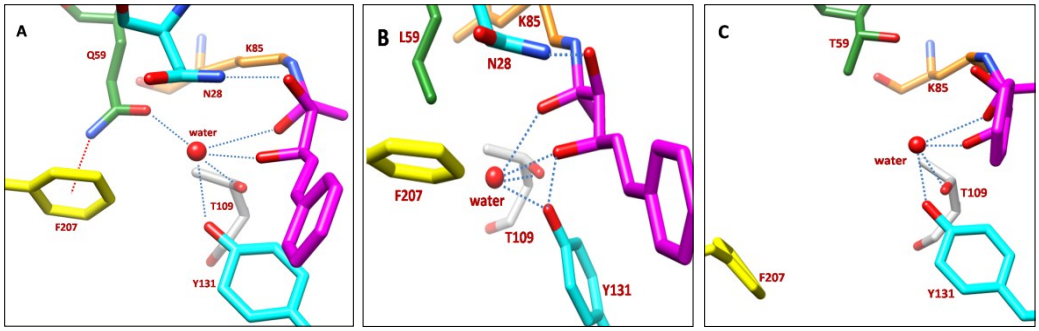
Fig. S6 The kinetic Michaelis-Menten equation fitted by Origin 8 software. (a) Curves for cinnamaldehyde; (b) Curves for hydroxyacetone.

Fig. S7 Models derived from automatic molecular docking (1) and molecular dynamics (MD) simulations (2).

(1) Protein and D-fructose-6-phosphate complexes of wild-type FSAA (A), FSAA Q59L (B), and FSAA Q59T (C) derived from automatic molecular docking using Autodock. Ligand is shown in magenta, the catalytic Lys85 in orange and site 59 in forest green. The catalytic water molecule is shown as a red ball. Hydrogen bonds are shown as dashed lines in wathet. Ionic interactions are shown as dashed lines in black.



(2) Protein and substrate complexes of wild-type FSAA (A), FSAA Q59L (B), and FSAA Q59T (C) derived from the molecular dynamics (MD) simulations. Ligand is shown in magenta, the catalytic Lys85 in orange and site 59 in forest green. The catalytic water molecule is shown as a red ball. Hydrogen bonds are shown as dashed lines in wathet. The proposed NH/π interaction is shown as a dashed line in red.



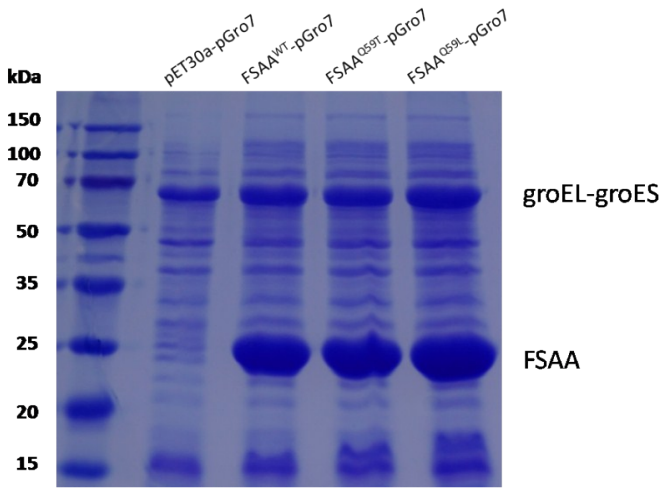


Fig. S8 SDS-PAGE analysis of cell-free extract of the whole-cell catalysts (wild-type FSAA and mutants).

E. coli BL21 (DE3) containing plasmids pET30a and pGro7 was used as control (pET30a-pGro7). Vector pGro7 harbors groEL-groES, which is a class of common molecular chaperones that can help proteins to fold correctly.

References

1. S. K. Kim, S. W. Nam, *Korean Journal of Life Science* 2002, **12** 688-693.
2. L. Stellmacher, T. Sandalova, S. Leptihn, G. Schneider, G. A. Sprenger and A. K. Samland, *ChemCatChem* 2015, **7**, 3140 – 3151.
3. M. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248–254.
4. M. Schürmann and G. A. Sprenger, *J. Biol. Chem.*, 2001, **276**, 11055-11061.
5. I. Sánchez-Moreno, L. Nauton, V. Théry, A. Pinet, J.-L. Petit, V. de Berardinis, A. K. Samland, C. Guérard-Hélaine and M. Lemaire, *J. Mol. Catal. B: Enzym.*, 2012, **84**, 9-14.
6. D. Acetti, E. Brenna and F. G. Gatti, *Eur. J. Org. Chem.*, 2010, **23**, 4468-4475.
7. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. H. Li, H. P., A. F. Izmaylov, J. Bloino, G. Zheng and J. L. Sonnenberg, *Gaussian, Inc.*, Wallingford, 2009.
8. P. J. Stephens and N. Harada, *Chirality*, 2010, **22**, 229-233.
9. J. Liu, X. L. Tang and H. W. Yu, *Biotechnol. Bioeng.*, 2010, **105**, 687-696.
10. S. Thorell, M. Schürmann, G. A. Sprenger and G. Schneider, *J. Mol. Biol.*, 2002, **319**, 161-171.
11. M. Gutierrez, T. Parella, J. Joglar, J. Bujons and P. Clapés, *Chem. Commun.*, 2011, **47**, 5762-5764.
12. N. Guex and M. C. Peitsch, *Electrophoresis*, 1997, **18**, 2714-2723.
13. A. Szekrenyi, A. Soler, X. Garrabou, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, J. Bujons and P. Clapés, *Chemistry*, 2014, **20**, 12572-12583.
14. D. A. Case, X. Wu, T. Darden, S. R. Brozell, T. E. C. III, T. Steinbrecher, C. Simmerling, H. Gohlke, J. Wang, Q. Cai, R. E. Duke, X. Ye, R. Luo, J. Wang, R. C. Walker, M.-J. Hsieh and W. Zhang, *AMBER 11*, University of Calfornia, San Francisco, 2010..