Structure-based Insights into Mechanism of Endoperoxidase FtmOx1 Catalyzed Reactions

Fei Wang[a,b], Yanqing Gao[c], Chunxi Wang[a], Wenxian Lan[a], JianHua Gan[c*], Chunyang Cao[a*]

Table of contents

Supplemental materials and methods

Supplemental figures

residues Y68, T134 and Y224, upon the structures of ftmOx1·Fe^{2+} complex (4Y5T, cyan) and ftmOx1·Fe^{2+}·α-KG binary complex (4Y5S, green) being overlapped. The residues Y68, T134 and Y224, and co-factor α-KG were displayed in stick mode, the proteins were displayed in 80% transparent cartoon mode. The balls stand for iron ions.

(3) Figure S3 Structural analysis of apo ftmOx1 Y224F variant. (a) Four ftmox1 Y224F variant monomers are in each asymmetric unit. (b) All monomers in (a) are identical to one other. (c) The "face-to-face contact at C-terminus" like dimer structures (monomers were displayed in blue and green, respectively). (d) The reported dimer structures of ftmOx1 or its variants (for example, observed in structures with pdb codes 4Y5S, 4Y5T, 7ETK and 6OXJ and 6OXH, monomers were displayed in pink and green, respectively). (e) The "point-to-point contact at N-terminus" like dimer structure (monomers were displayed in blue and grey, respectively). (f) The monomers in (a) are identical to those in (d). (g) The orientation of the side-chain of Y224 in wild-type ftmOx1. (h) The orientation of the side-chain of Phe224 in ftmOx1 Y224F variant. In (b-f), the N- and C- termini were labeled in each structure.

(4) Figure S4 Y224 forms hydrogen-bond with with α-KG in structures of (a) ftmOx1 Y68F •Co^{2+}•α-KG ternary complex and (b) ftmOx1 Y140F•Fe^{2+}•α-KG. In (a and b), the structures of the proteins were displayed in 80% transparent cartoon mode. All residues were displayed in stick mode. The iron ions were presented as balls. The distances were shown in dotted lines, indicating hydrogen-bonds formation.

(5) Figure S5 The conformation changes of Y68 and Y224 in structures of 4Y5S (in cyan stick) and 7ETK (in green stick) upon their being overlapped. The substrate was displayed in orange line mode. The distances between oxygen of Y68 -OH group and C-13, C-21 and C-23 of the substrate were in dotted-lines. The atom numbers of C-13, C-21 and C-23 were labeled in red. The orange balls stand for iron ions, α-KG was displayed in green, or cyan stick modes. The protein structures were in 80%

transparency cartoon mode.

(6) **Figure S6** T134 have hydrophobic interactions with the substrate or its analogue in structures (a) 7ETK and (b) 7WSB. All residues, the substrate and its analogue were displayed in stick mode. The iron ions were presented as balls. The distances were shown in dotted lines. The orientations of the –OH group and methyl group in T134 side-chains in 7ETK and 7WSB are different.

(7) **Figure S7** Structural determination of compound 9 by NMR (a and b) and MALDI-TOF (c) spectra. In 1D ^{1}H-NMR (a) and ^{13}C-NMR (b) spectra, signal assignments were marked based on the structure (d).

(8) **Figure S8** ESI-HR-MS indicates compound 10 with chemical formula $C_{27}H_{33}N_3O_8$ and molecular weight 528Da are produced in the catalytic reaction. (a) High-resolution chromatography of reaction compound 10. (b) The compound 10 with retention time 11.7 mins showed low MFG Diff (ppm) value, which is consistent with our expected geminal diol chemical intermediate; (c) Compound 10 with a retention time 13.5 mins also accords with our expected geminal diol chemical intermediate. (d) The chemical structure of geminal diol intermediate.

Supplemental tables

(1) **Table S1** Summary of data collection and refinement statistics for the x-ray crystal structure of apo ftmOx1 Y224F.

(2) **Table S2** NMR signal assignment of ^{1}H and ^{13}C atoms in compound 9

Supplemental materials and methods

Expression and purification of ftmOx1

The DNA coding sequences of wild type ftmOx1 or its variants was optimized for expression in *Escherichia coli* system, and then sub-cloned into the *Nde*I and *Xho*I restriction sites of pET-22b vector with a 6xHis Tag at C-terminus by GenScript Company (Nanjing, China). Plasmids were then transformed into *Escherichia coli* BL21(DE3) cells. A single colony was used to incubate 100 ml of LB medium with 10 μg/ml ampicillin overnight at 37°C. Next 10 ml production was transferred into 1 L LB medium with 10 μg/ml Ampicillin for another 3 hours until OD_{600} value is equal to 0.8, and then induced by IPTG with a final concentration 0.3 mM for 20 hours at 18°C before harvesting. The following steps were all performed at 4°C. Harvested cell was resuspended in 50 ml buffer containing 20 mM Tris pH 7.5 and 300 mM NaCl with protease inhibitor (PMSF) and lysed at 20 kpsi using a hydraulic cell disruption system (Constant System JINBO Benchtop) (Guangzhou Juneng Biology and Technology Co., LTD, Guangzhou, China). The lysate was centrifuged at 12000 rpm and 4°C for 1 hour to remove cellular debris. The 0.1% (w/v) Streptomycin sulfate was then added into the supernatant and stirred for 30 min at 4°C, and centrifuged for 1 hour again to remove nucleic acid. The supernatant was loaded into a Ni-NTA resin (GE health) followed by a gradient elution with Ni^{2+}-binding buffer (20 mM Tris, pH 7.5, 300 mM NaCl and different concentrations of imidazole, 20 mM, 50 mM, 100 mM and 500 mM). Fractions containing ftmOx1 protein were dialyzed against 20 mM Tris, pH 7.5 and 150 mM NaCl at 4°C. After that, the protein was further purified by cation exchange column (HiTrap™ 5ml Q FF, GE health) with high salt concentration buffer (20 mM Tris, pH 7.5, 1 M NaCl) and low salt buffer (20 mM Tris, pH 7.5). The elution protein was concentrated and purified on a Superdex75 16/300 GL column (GE Health), which was previously equilibrated with buffer containing 20 mM Tris, pH 7.5 and 50 mM NaCl.

Expression and purification of 3-^{19}F-L-Tyrosine labeled ftmOx1

Plasmids were transformed into *Escherichia coli* BL21(DE3) cells for protein expression. A single colony was used to incubate 100 mL of TY medium with 10 μg/ml ampicillin overnight at 37°C. The harvested cell was resuspended in 1L M9 medium and incubated at 37°C for about 3h until OD_{600} value equals to 0.4, then 0.5 g N-(phosphonomethyl)-glycine, 70 mg 3-fluoro-L-tyrosine, 60 mg L-tryptophan and 60 mg L-phenylalanine were added into system. The culture was incubated at 37°C until OD_{600} value is equal to 0.8, followed by adding 0.3 mM IPTG to induce protein expression for 20 hours at 18°C. The ^{19}F-labeled protein was purified in the same way to its unlabeled wild-type protein.

Enzymatic reaction assay of wild-type ftmOx1 and its variants

The reaction mixture was fixed as 100 μl, containing 0.5 mM fumitremorgin B (compound 1), 0.11 mM ftmOx1 (or its variant), 1mM α-KG, 1 mM ammonium iron (II) sulfate (i.e., $Fe(NH_4)_2(SO_4)_2$), 1 mM L-ascorbic acid and 50 mM Tris pH 7.5. After incubation at 37°C for 30 min, the reaction was quenched by adding 100 μL acetonitrile to precipitate protein. Then the solution was centrifuged at 12000 rpm for 1 minute. After that, 200 μL chloroform was added to extract organic phase. The organic phase was drying by N_2 and dissolved by DMSO for HPLC analysis.

Enzymatic reaction products were analyzed by HPLC using an Agilent ZORBAX SB-C18 reversed phase column (4.6 x 250 mm, 5 μm; USA) on the Agilent 1200 series HPLC system (Agilent Technologies Inc., USA). A linear gradient of 30%-100% (v/v) acetonitrile in water was run for 30 min with a flow rate of 1 ml/min, followed by 100% (v/v) acetonitrile for another 5 min. Before the next run, the column was equilibrated with 30%(v/v) acetonitrile for 3 min. The components were detected with a Photo Diode Array detector at 300 nm.

After reaction components were separated by HPLC system, ESI-high resolution MS (ESI-HR-MS) analysis was carried out on the 6230B Accurate-Mass TOF LC/MS system or 6530 Accurate-Mass Q-TOF LC/MS system (Agilent Technologies Inc., USA) to

determine chemical formula. Meanwhile, the data was analyzed by Agilent MassHunter Qualitative Analysis software. NMR experiments were performed on 600 MHz Varian spectrometers at 25°C and data was processed by MestReNova software.

Structural determination of compound 9 by NMR and Mass spectra

To determine the structure of reaction product compound 9, it was first purified by using HPLC system, and then MALDI-TOF was run to confirm its molecular weight. After that, one dimensional 1H-NMR and 13C-NMR spectra, and two-dimensional ^{1}H-^{1}H COSY and TOCSY, ^{1}H-^{13}C HSQC and HMBC were acquired on this compound dissolved in d^{6}-DMSO on 600MHz Varian spectrometers at 25^{o}C. NMR data was processed by NMRPipe and assigned by Sparky. The assignment of NMR signals belonging to the ^{1}H and ^{13}C atoms were summarized in **table S2**.

Crystallization and structural determination of ftmOx1 Y224F variant

The ftmOx1 Y224F variant was concentrated with the final concentration 10 mg/ml in buffer, containing 20 mM Tris pH 7.5 and 50 mM NaCl. Crystallization was set up using the sitting-drop vapour diffusion method by mixing protein and crystallization buffer (100 mM Bicine (pH 6.5), 300 mM $MgCl_2$, 30% (w/v) PEG2000 and 1.2% (w/v) myo-Inositol) at a ratio of 1:2 at 18°C. After 7 days, ship-like crystal formed. Before being placed in liquid nitrogen, the crystals were cryo-protected by the mother liquor containing 30% glycerol. The diffraction data were collected at BL19U1 beamline of the Shanghai Synchrotron Radiation Facility (SSRF, China) using a DECTRIS PILATUS3 6M detector at 98K. The wavelength of the radiation was 0.9793 Å, the distance between detector and crystal was 450 nm and the exposure time was set as 1s. All 360 frames were collected with a 1°oscillation. Then the data were indexed, integrated and scaled by using HKL3000 program suite. The structure of ftmOx1 Y224F variant was finally solved by molecular replacement (MR) with the reported structure of wild-type ftmOx1 as initial mode (pdb code 4y5t). Then, using PHENIX 1.10.1_2155, protein structure was obtained, and followed by doing several corrections in COOT. Structural figures were prepared using Pymol (http://www.pymol.org). Statistics of the structure

refinement and the quality of the structure model were also summarized in **table S1**.

^{19}F-NMR spectra acquired on 3-^{19}F-L-Tyrosine labeled ftmOx1 variants

All 3-^{19}F-L-tyrosine labelled protein samples were concentrated to 0.5 mM in 10% D_2O buffer, containing 150mM NaCl and 25mM PBS pH 7.5, the final concentration of the substrate is kept at 0.5 mM. ^{19}F spectra were acquired on Bruker 600MHz spectrometers equipped with 5-mm H/F (C, N) triple resonance cryoprobes at 293K, and the spectral width was set as 30 ppm with a duty cycle delay of 2.0s.

Supplemental figures and tables

Figure S1 Two previously suggested mechanisms for ftmOx1 catalyzed endo-peroxidation. (a) COX-like mechanism. In this mechanism, oxo-ferryl species directly abstracts hydrogen from Tyr224, forming Tyr224-O• radical. Then, this radical abstract hydrogen from C-21 of fumitremorgin B, followed by dioxygen installation. The second hydrogen radical rebound produces major product verruculogen (*i.e.*, compound 2). (b) CarC-like mechanism reported by Dunham et al. In this mechanism, once the oxo-ferryl species is formed, a coordination rearrangement of iron centre happens to adapt substrate binding. Oxo-ferryl species directly abstracts hydrogen from C-21 of fumitremorgin B. After dioxygen addition, the radical is transferred to C-26 atom (C26•). This radical is quenched by residue Tyr68 to form verruculogen. The Tyr68 radical (Tyr68-O•) and Fe (III)-OH species, will be reduced by ascorbate to their initial states for another catalytic cycle.

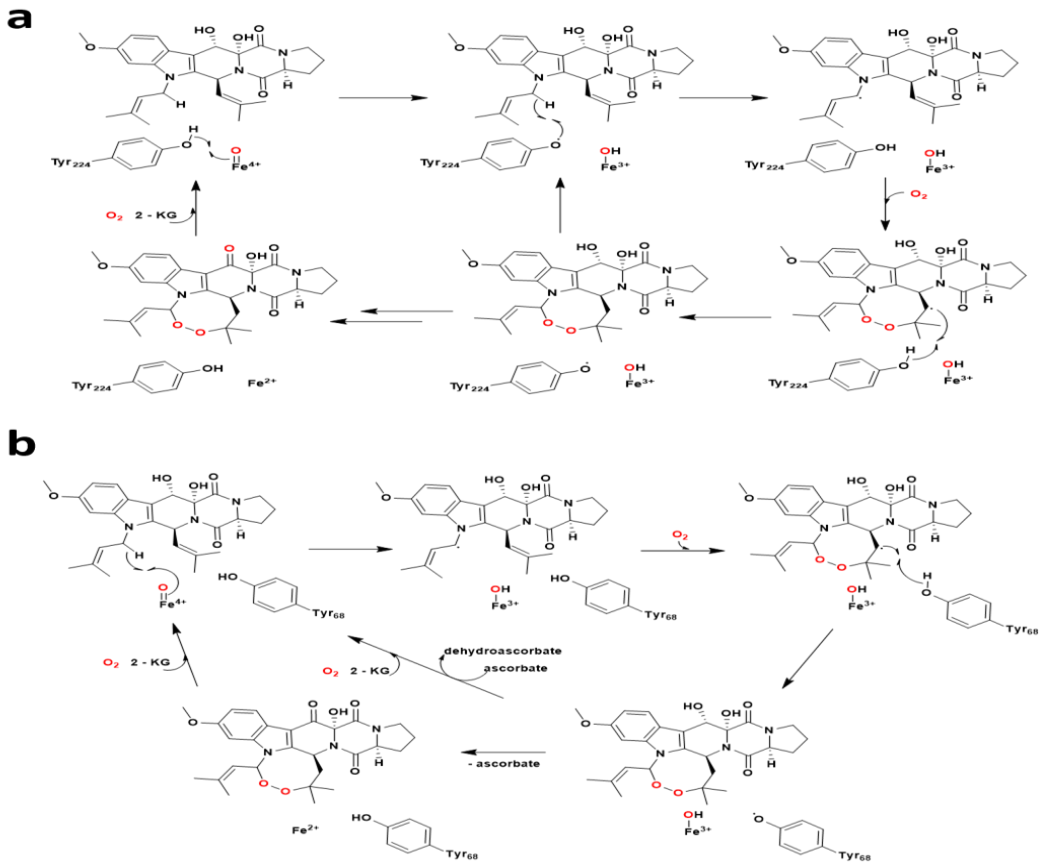


Figure S2 α-KG binding does not result in obvious changes in the conformations of residues Y68, T134 and Y224, upon the structures of ftmOx1·Fe^{2+} complex (4Y5T, cyan) and ftmOx1·Fe^{2+}·α-KG binary complex (4Y5S, green) being overlapped. The residues Y68, T134 and Y224, and co-factor α-KG were displayed in stick mode, the proteins were displayed in 80% transparent cartoon mode. The balls stand for iron ions.

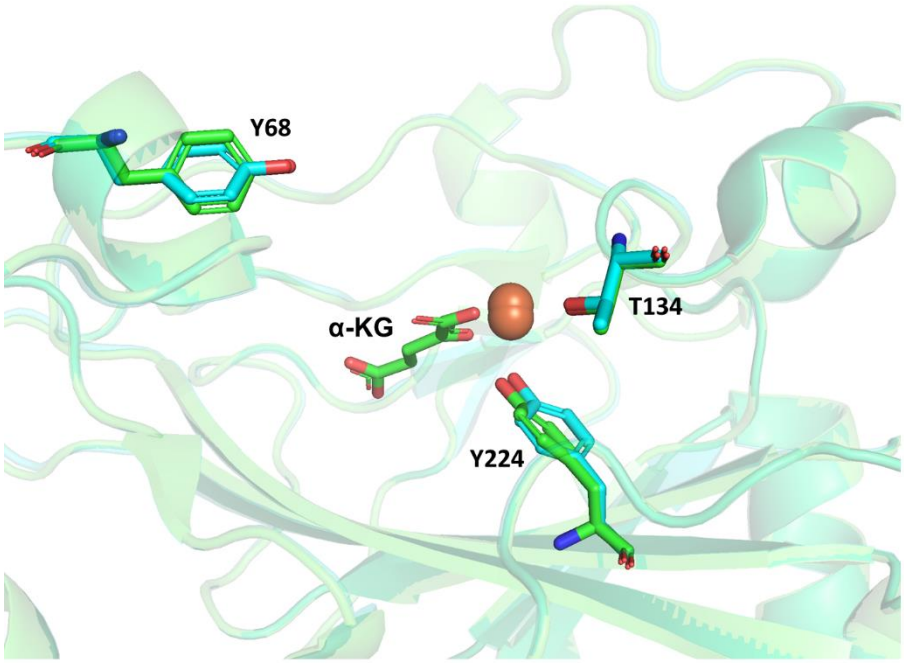


Figure S3 Structural analysis of apo ftmOx1 Y224F variant. (a) Four ftmox1 Y224F variant monomers are in each asymmetric unit. (b) All monomers in (a) are identical to one other. (c) The "face-to-face contact at C-terminus" like dimer structures (monomers were displayed in blue and green, respectively). (d) The reported dimer structures of ftmOx1 or its variants (for example, observed in structures with pdb codes 4Y5S, 4Y5T, 7ETK and 6OXJ and 6OXH, monomers were displayed in pink and green, respectively). (e) The "point-to-point contact at N-terminus" like dimer structure (the monomers were displayed in blue and grey, respectively). (f) The monomers in (a) are identical to those in (d). (g) The orientation of the side-chain of Y224 in wild-type ftmOx1. (h) The orientation of the side-chain of F224 in ftmOx1 Y224F variant. In (b-f), the N- and C- termini were labeled in each structure.

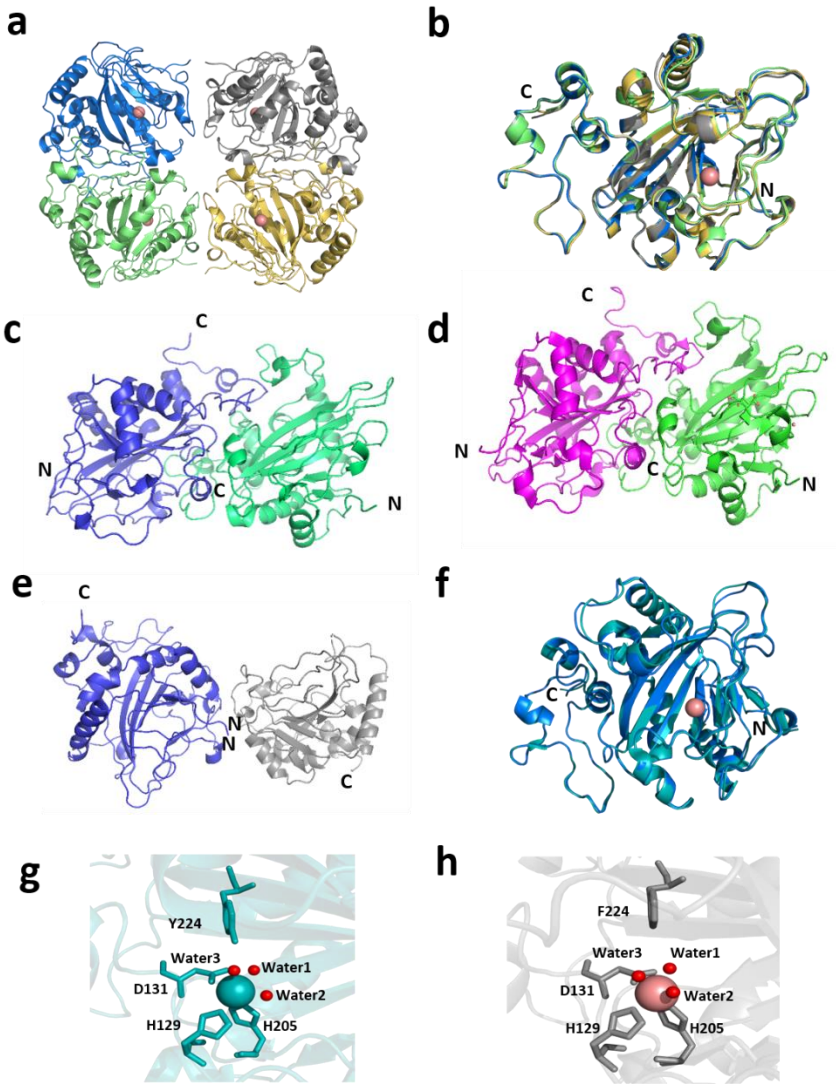


Figure S4 Y224 forms hydrogen-bond with α-KG in structures of (a) ftmOx1 Y68F •Co^{2+}•α-KG ternary complex and (b) ftmOx1 Y140F•Fe^{2+}•α-KG. In (a and b), the structures of the proteins were displayed in 80% transparent cartoon mode. All residues were displayed in stick mode. The iron ions were presented as balls. The distances were shown in dotted lines, indicating hydrogen-bonds formation.

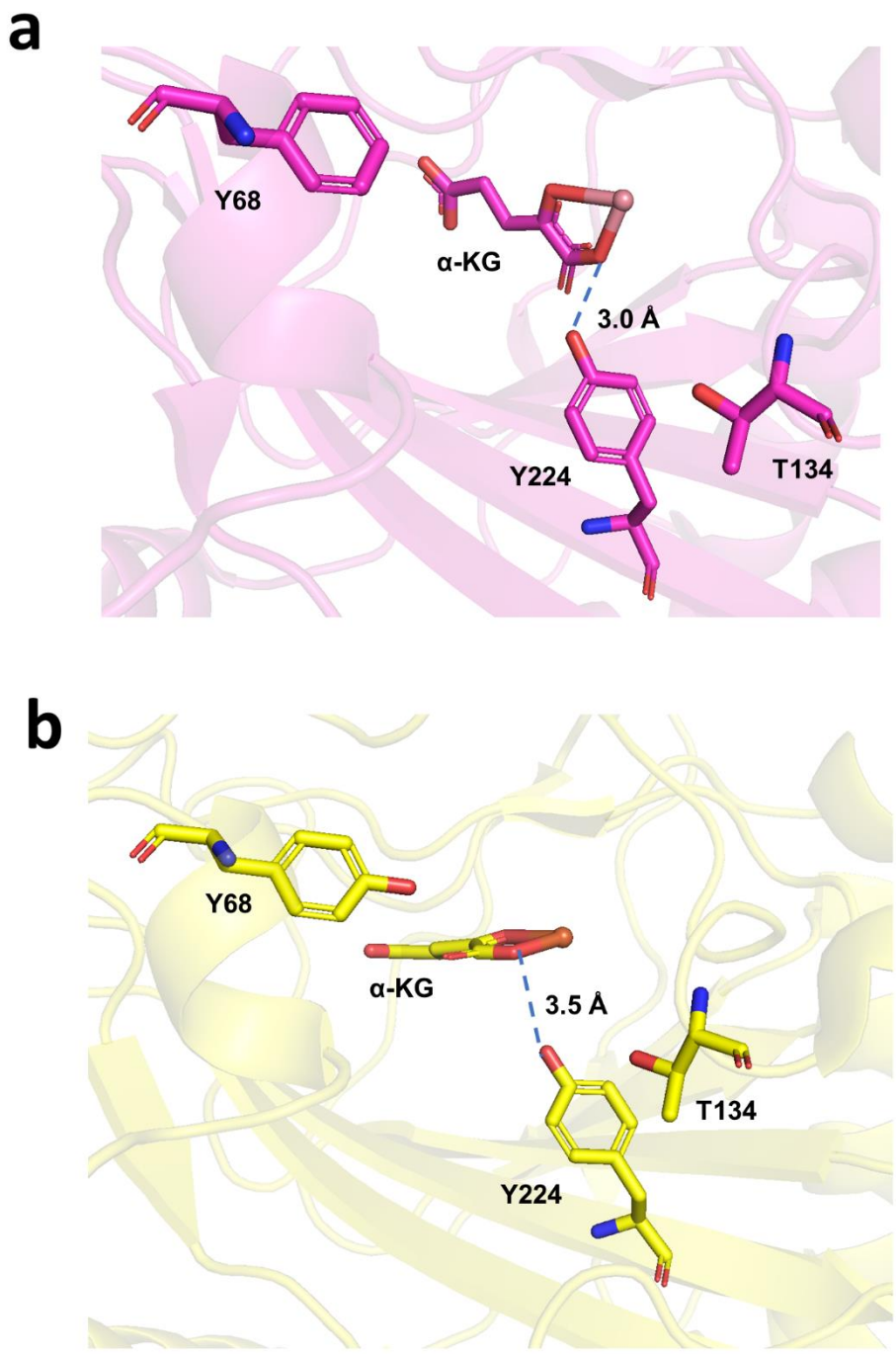


Figure S5 The conformation changes of Y68 and Y224 in structures of 4Y5S (in cyan stick) and 7ETK (in green stick) upon their being overlapped. The substrate was displayed in orange line mode. The distances between oxygen of Y68 -OH group and C-13, C-21 and C-23 of the substrate were in dotted-lines. The atom numbers of C-13, C-21 and C-23 were labeled in red. The orange balls stand for iron ions, α-KG was displayed in green, or cyan stick modes. The protein structures were in 80% transparency cartoon mode.

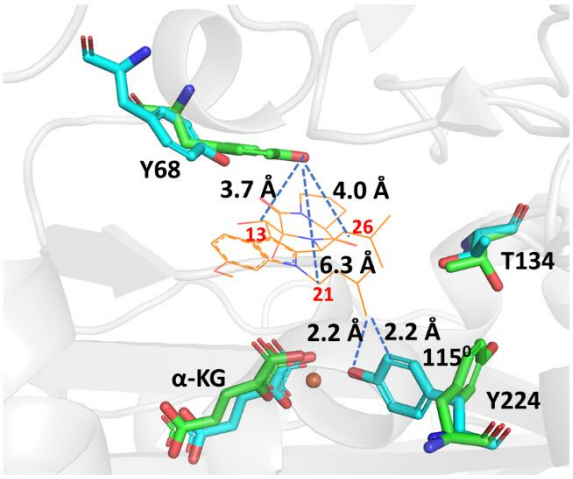


Figure S6 T134 should have hydrophobic interactions with the substrate or its analogue in structures (a) 7ETK and (b) 7WSB. All residues, the substrate and its analogue were displayed in stick mode. The iron ions were presented as balls. The distances were shown in dotted lines. The orientations of the –OH group and methyl group in T134 side-chains in 7ETK and 7WSB are different.

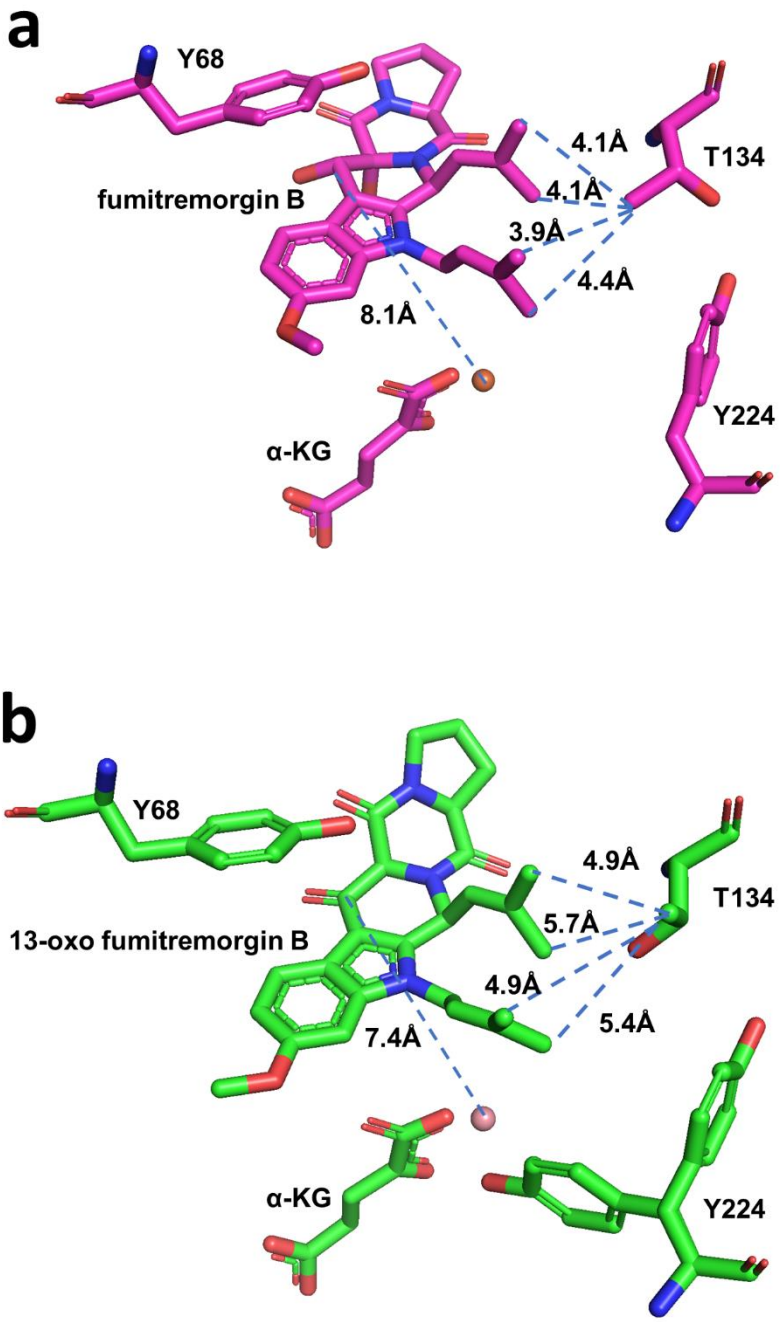
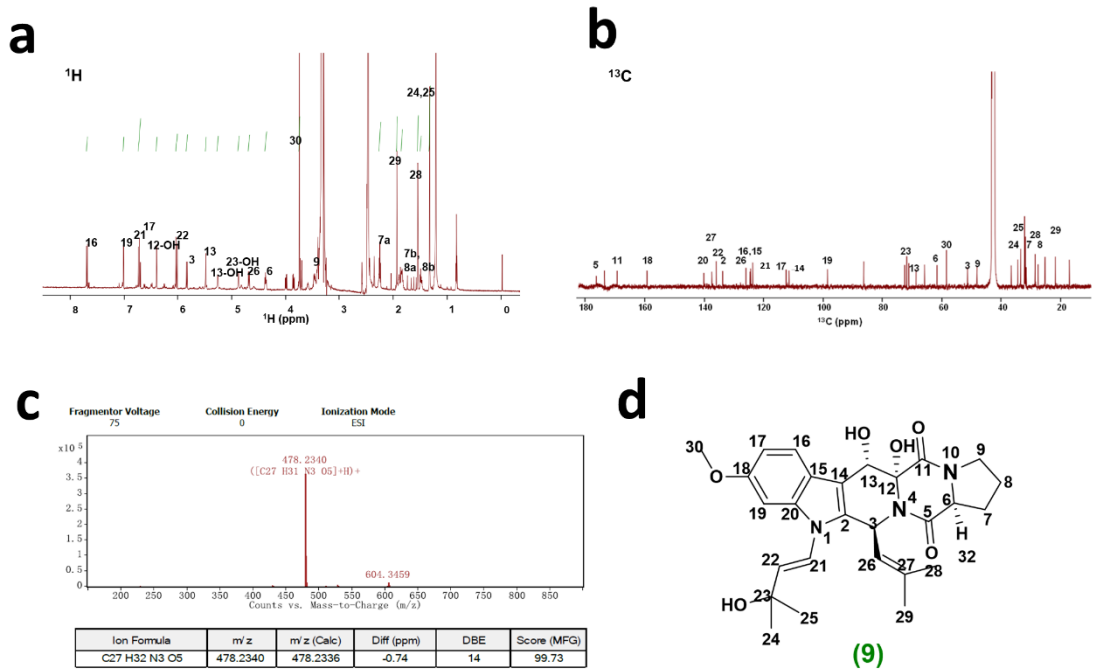


Figure S7 Structural determination of compound 9 by NMR (a and b) and MALDI-TOF (c) spectra. In 1D ^{1}H-NMR (a) and ^{13}C-NMR (b) spectra, signal assignments were marked based on the structure (d).



Ion Formula	m/z	m/z (Calc)	Diff (ppm)	DBE	Score (MFG)
C27 H32 N3 O5	478.2340	478.2336	-0.74	14	99.73

Figure S8 ESI-HR-MS indicates compound 10 with chemical formula $C_{27}H_{33}N_3O_8$ and molecular weight 528Da are produced in the catalytic reaction. (a) High-resolution chromatography of reaction compound 10. (b) The compound 10 with retention time 11.7 mins showed low MFG Diff (ppm) value, which is consistent with our expected geminal diol chemical intermediate; (c) Compound 10 with a retention time 13.5 mins also accords with our expected geminal diol chemical intermediate. (d) The chemical structure of geminal diol intermediate.

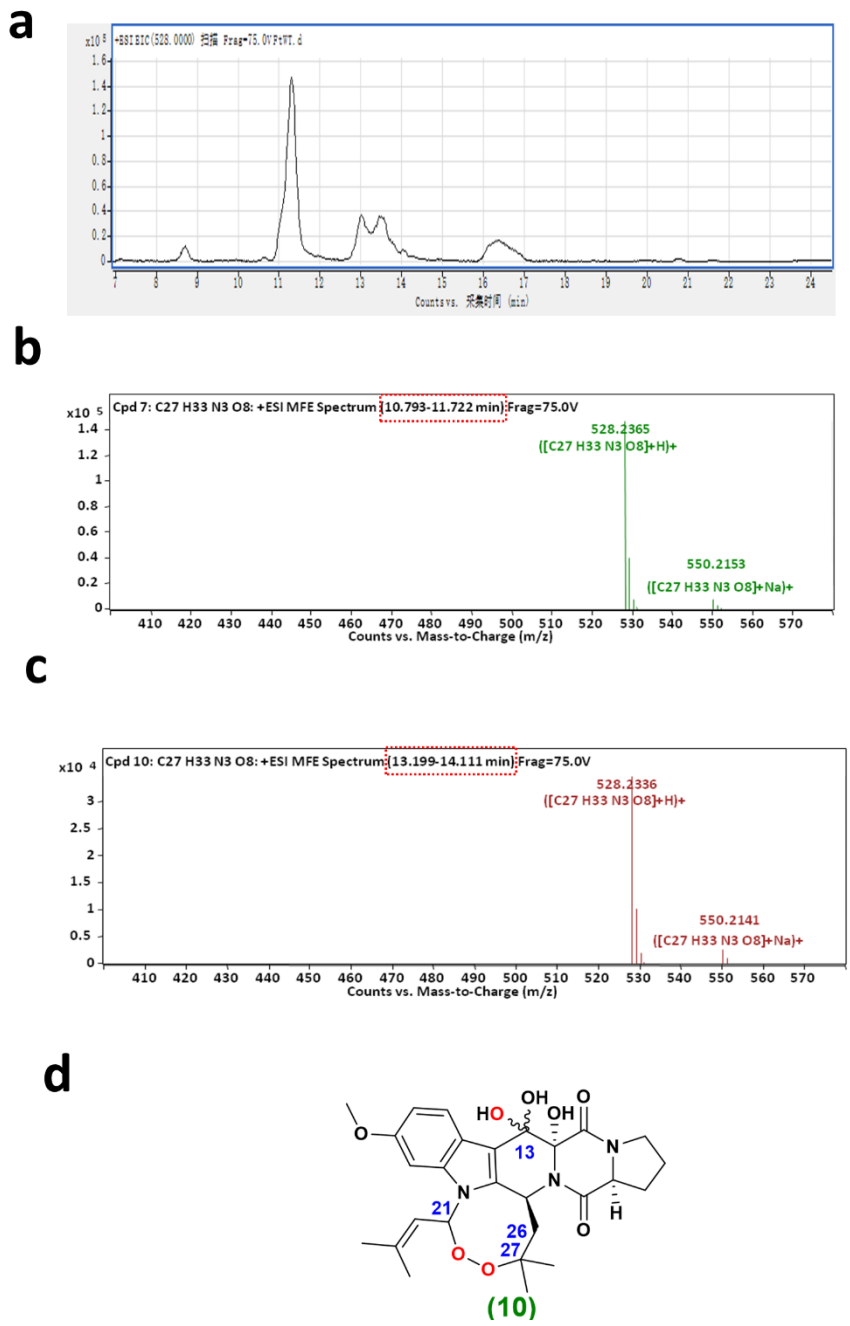


Table S1 Summary of data collection and refinement statistics for the X-ray crystal structure of apo ftmOx1 Y224F.

	ftmOx1-Y224F
Data collection	
Wavelength	0.97890 Å
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	45.436 ,136.740 ,180.628
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	30.00-2.40(2.49-2.40)
Rmerge	0.123 (0.429)
< I/σ>	14.1 (2.3)
CC1/2	0.989
Completeness (%)	98.9 (94.6)
Redundancy	7.9 (4.3)
Refinement	
Resolution (Å)	30.00-2.40
No. reflections	39205
Rwork / Rfree	0.2155/0.2631
No. atoms	
Protein	8935
Ion/Ligand	4
Water	47
B-Factors	
Protein	30.1
Ion/Ligand	36.8
Water	21.7
r.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.866
Ramachandran statistics (%)	
favored	95.0
outliers	0

Table S2 NMR signal assignment of ^{1}H and ^{13}C atoms in compound 9

Atom number	^{1}H (ppm)	^{13}C (ppm)
1-N	-	-
2	-	130.50
3	5.86(d,1H)	48.17
4-N	-	-
5	-	173.01
6	4.59(m,1H)	58.42
7	1.80-1.925 (m,0.5H),2.225-2.275(m,0.5H)	28.48
8	1.450-1.525(m,0.5H),1.80-1.925(m,0.5H)	24.48
9	3.35-3.425(m,2H)	45.00
10-N	-	-
11	-	165.96
12	-	Not assigned
12-OH	6.42(s,1H)	-
13	5.50(s,1H)	67.93
13-OH	5.28(s,1H)	-
14	-	108.17
15	-	121.07
16	7.72(d,1H)	121.34
17	6.75(d,1H)	109.18
18	-	155.95
19	7.04(d,1H)	95.24
20	-	136.85
21	6.75(d,1H)	120.40
22	6.05(d,1H)	132.65
23	-	68.63
23-OH	4.89(s,1H)	-
24-CH_3	1.33(s,3H)	31.31
25-CH_3	1.33(s,3H)	30.39
26	4.70(d,1H)	122.67
27	-	134.18
28-CH_3	1.55(s,3H)	25.43
29-CH_3	1.94(s,3H)	18.68
30-CH_3	3.76(s,3H)	55.31