SUPPORTING INFORMATION

Structural Basis for Substrate and Oxygen Activation in Homoprotocatechuate 2,3-Dioxygenase: Roles of Conserved Active Site Histidine-200

Elena G. Kovaleva,[†]* Melanie S. Rogers,[‡] and John D. Lipscomb[‡]*

[†]*Institute of Molecular and Cellular Biology, University of Leeds, Leeds, LS2 9JT, UK and* [‡]*Department of Biochemistry, Molecular Biology and Biophysics, and Center for Metals in Biocatalysis, University of Minnesota, Minneapolis, MN, USA 55455*

Experimental Procedures:

Reagents. All chemicals for cell growth and enzyme purification were obtained from commercial vendors and used without further purification. All reagent solutions and media were prepared using water treated with a Millipore MilliQ or SuperQ water system to minimize trace metal-ion contamination. Crystallization solutions were prepared using LC-MS grade water (Sigma-Aldrich). Substrates homoprotocatechuate (HPCA) and 4-nitrocatechol (4NC) were recrystallized from water at 4 °C to remove minor contaminants. 4-Sulfonylcatechol was the generous gift of James W. Whittaker.

Site-Directed Mutagenesis. Recombinant homoprotocatechuate 2,3-dioxygenase (FeHPCD) from *Brevibacterium fuscum* (EC 1.13.11.15) was expressed in *E. coli* BL21(DE3) cells carrying plasmid pYZW204.[1] Single mutations in plasmid pYZW204 to generate H200E, H200Q, and H200N variants were prepared previously.[2] A similar procedure was carried out to generate single mutations in pYZW204 plasmid for D272N and Y269F variants. Oligonucleotide primers were purchased from Sigma, and presence of the desired mutation was verified by gene sequencing (GATC Biotech) using FeHPCD primer 5'-CTGAGGACGTCGACAAGGC-3'.

Mutation	Oligonucleotide primer
D272N+	5'-GAGATCTACACCCAGAACTACTACACCGGCG-3'
D272N–	5'-CGCCGGTGTAGTAGTTCTGGGTGTAGATCTC-3'
Y269F+	5'-CACCGCATCGAGATCTTCACCCAGGACTACTAC-3'
Y269F–	5'-GTAGTAGTCCTGGGTGAAGATCTCGATGCGGTG-3'

Protein Expression and Purification. The WT and mutant enzymes of FeHPCD were expressed and purified using previously described procedures.[1-3] Protein concentration was determined using absorbance at 280 nm (1.2 mg/ml/cm^{-1}) as previously described.[4] Metal quantification was determined by Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES) using a Thermo Scientific iCAP 6500 dual view ICP-OES or using a Thermo Scientific XSERIES 2 ICP-MS ICP-MS. Iron site occupancy for each of the H200X variants was >90%.

Preparation of Fully-Reduced FeHPCD Variants. Residual active site ferric ion present following protein purification was reduced using 0.5 – 1.0 equivalents of sodium dithionite under anaerobic conditions. Excess sodium dithionite was removed from the enzyme solution using a 10 ml Sephadex G-25 column (PD-10, GE Healthcare Life Sciences) equilibrated with 200 mM MOPS, pH 7.5.

X-ray Crystallography. Crystals of FeHPCD variants were grown by the hanging-drop method at 20 °C in either 13 – 15 % PEG6000, 0.1 M calcium chloride or acetate, 0.1 M Tris-HCl or MOPS buffer pH 6.5 – 7.0. For X-ray data collection at 100K, crystals were cryo-cooled in liquid nitrogen following a brief transfer into cryoprotectant solution containing either 20–25% PEG400 or 20% glycerol in the mother liquor solution. Crystal morphology, cell dimensions, crystal packing and structure are not affected by the choice of buffer or calcium salt used in our earlier or present studies of the FeHPCD and its variants.[5-7] Specific crystallization conditions used for each structure are stated in the Methods section of each PDB Database entry.

For anaerobic complex preparation, all mother liquor solutions, crystals of FeHPCD variants and ligand stocks were equilibrated in the anaerobic glove box atmosphere (Belle Technology) for at least 18 h prior to mixing. The soaking reactions were initiated by anaerobic addition of 5 mM aromatic substrates to crystals of FeHPCD. After 10 – 30 min incubation, crystal complexes were rapidly transferred into mother liquor solution containing either 25% PEG400 or 20% glycerol prior to cryo-cooling in liquid nitrogen inside the anaerobic glove box.

Diffraction data were processed using the XDS package.[8] The coordinates of the full length FeHPCD (PDB 3OJT) were used as an initial model in rigid body refinement followed by cycles of restrained refinement with Refmac5[9, 10] as part of the CCP4 program suite[11] and model building using Coot.[12] TLS (translation, libration, small movements) was used in the final round of restrained refinement, with a single subunit defined as a TLS group. Link restraints to the iron were removed from the refinement to avoid bias in the refined metal-ligand distances. Non-crystallographic symmetry restraints were not used during refinement, and the 4 subunits of the single enzyme molecule present in the asymmetric unit were refined independently. Ligand refinement protocols were essentially the same as that described previously.[5, 7] X-ray data processing and refinement statistics are summarized in Tables S1, S2, S3 and S5. All structure figures were produced using Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC).

Table S1. X-ray data collection and refinement statistics for the H200E FeHPCD variant in free form and anaerobic complexes with HPCA, 4SC and 4NC [a].

Dataset	H200E	H200E-[HPCA]	H200E-[4SC]	H200E-[4NC]
(PDB Code)	(PDB 4Z6L)	(PDB 4Z6O)	(PDB 4Z6R)	(PDB 4Z6U)
Wavelength	0.9801 Å	0.9801 Å	0.9795 Å	0.9801 Å
Synchrotron (beamline)	Synchrotron Soleil (Proxima I)	Synchrotron Soleil (Proxima I)	DLS (I-02)	Synchrotron Soleil (Proxima I)
Spacegroup	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell dimensions (Å)	110.3, 150.6, 96.2	110.3, 150.5, 96.0	110.5, 150.5, 96.0	110.3, 150.9, 95.9
Cell angles (deg)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range[a] (Å)	48.1 – 1.65 (1.74)	48.0 – 1.63 (1.72)	29.6 – 1.70 (1.79)	45.8 – 1.48 (1.56)
Reflections (observed/unique)	968653/191307	852108/197371	704029/145565	1177706/262655
R_{rim} [a, b] (%)	8.4 (79.5)	7.4 (68.0)	8.5 (71.1)	6.1 (83.5)
R_{pim} [a, c] (%)	4.1 (38.8)	3.9 (39.3)	4.1 (36.3)	3.1 (42.5)
Mean $\langle I\rangle/\sigma\langle I\rangle$ [a]	10.5 (1.9)	12.5 (1.9)	12.2 (2.0)	14.6 (2.0)
Completeness (%) [a]	99.5 (97.8)	99.3 (97.9)	83.5 (87.9)	99.2 (99.4)
R, R_{free}, test (%) [d]	14.8, 17.8, 5.0	13.8, 16.7, 5.0	14.1, 17.4, 5.0	11.7, 16.2, 5.0
RMSD [e] bond length (Å)	0.013	0.013	0.014	0.012
RMSD [e] angles (deg)	1.502	1.511	1.580	1.469
ESU [f] (Å)	0.064	0.057	0.068	0.044
Ramachandran Plot				
Allowed regions (%)	99.7	99.7	99.7	99.7
Additional regions (%)	0.3	0.3	0.3	0.3

[a] Values for the highest resolution shell are given in parentheses. [b] Redundancy-independent merging R factor: $R_{r.i.m} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [c] Precision-indicating merging R factor: $R_{p.i.m} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [d] $R = (\sum||F_{obs}|-k|F_{calc}||)/\sum|F_{obs}|$, where k is a scale factor. The R_{free} value was calculated with the indicated percentage of reflections not used in the refinement. [e] Root-mean-square deviation (RMSD) from ideal geometry in the final models. [f] Estimated overall coordinate error (ESU) based on maximum likelihood.

Table S2. X-ray data collection and refinement statistics for the H200Q FeHPCD variant in free form and anaerobic complexes with HPCA, 4SC and 4NC [a].

Dataset (PDB Code)	H200Q (PDB 4Z6M)	H200Q-[HPCA] (PDB 4Z6P)	H200Q-[4SC] (PDB 4Z6S)	H200Q-[4NC] (PDB 4Z6V)
Wavelength	1.0000 Å	0.9801 Å	1.0000 Å	0.9334 Å
Synchrotron (beamline)	SLS (X-06DA)	Synchrotron Soleil (Proxima I)	SLS (X-06DA)	ESRF (ID-14-1)
Spacegroup	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell dimensions (Å)	110.5, 152.1, 96.2	110.7, 151.2, 96.4	110.6, 152.0, 96.4	110.2, 150.6, 96.1
Cell angles (deg)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range[a] (Å)	48.1 – 1.35 (1.42)	48.0 – 1.75 (1.84)	47.9 – 1.42 (1.50)	47.8 – 1.37 (1.44)
Reflections (observed/unique)	1766318/338753	722159/160940	1339975/299060	1340516/325847
R_{rim} [a, b] (%)	9.4 (76.9)	7.8 (77.0)	7.3 (76.7)	6.4 (78.6)
R_{pim} [a, c] (%)	4.5 (36.6)	3.9 (38.8)	3.8 (41.0)	3.5 (43.2)
Mean $\langle I\rangle/\sigma\langle I\rangle$ [a]	11.3 (2.4)	12.6 (2.1)	11.8 (2.0)	14.2 (1.9)
Completeness (%) [a]	96.1 (93.0)	98.8 (99.5)	98.4 (96.6)	97.9 (95.9)
R, R_{free}, test (%) [d]	11.2, 14.0, 5.0	14.6, 17.7, 5.0	12.4, 16.0, 5.0	11.9, 15.7, 5.0
RMSD [e] bond length (Å)	0.009	0.011	0.010	0.009
RMSD [e] angles (deg)	1.369	1.365	1.379	1.357
ESU [f] (Å)	0.027	0.067	0.038	0.034
Ramachandran Plot				
Allowed regions (%)	99.7	99.5	99.7	99.7
Additional regions (%)	0.3	0.5	0.3	0.3

[a] Values for the highest resolution shell are given in parentheses. [b] Redundancy-independent merging R factor: $R_{r.i.m} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [c] Precision-indicating merging R factor: $R_{p.i.m} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [d] $R = (\sum||F_{obs}|-k|F_{calc}||)/\sum|F_{obs}|$, where k is a scale factor. The R_{free} value was calculated with the indicated percentage of reflections not used in the refinement. [e] Root-mean-square deviation (RMSD) from ideal geometry in the final models. [f] Estimated overall coordinate error (ESU) based on maximum likelihood.

Table S3. X-ray data collection and refinement statistics for the H200N FeHPCD variant in free form and anaerobic complexes with HPCA, 4SC and 4NC. [a]

Dataset (PDB Code)	H200N (PDB 4Z6N)	H200N-[HPCA] (PDB 4Z6Q)	H200N-[4SC] (PDB 4Z6T)	H200N-[4NC] (PDB 4Z6W)
Wavelength	0.9801 Å	0.9801 Å	0.9173 Å	0.9801 Å
Synchrotron (beamline)	Synchrotron Soleil (Proxima I)	Synchrotron Soleil (Proxima I)	DLS (I04-1)	Synchrotron Soleil (Proxima I)
Spacegroup	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell dimensions (Å)	110.2, 150.7, 96.1	110.5, 150.1, 96.0	110.2, 150.0, 95.8	110.5, 150.6, 96.0
Cell angles (deg)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution range[a] (Å)	48.0 – 1.52 (1.60)	47.9 – 1.57 (1.65)	29.5 – 1.50 (1.58)	48.0 – 1.57 (1.65)
Reflections (observed/unique)	891886/240127	985392/217448	1212267/248661	1452866/218224
R_{rim} [a, b] (%)	6.1 (69.4)	7.5 (79.9)	9.0 (79.6)	9.0 (81.3)
R_{pim} [a, c] (%)	3.5 (41.2)	3.8 (40.8)	4.4 (39.1)	3.7 (36.7)
Mean $\langle I\rangle/\sigma\langle I\rangle$ [a]	12.5 (1.9)	13.1 (1.9)	11.6 (2.0)	12.6 (2.2)
Completeness (%) [a]	98.3 (95.3)	98.2 (97.6)	98.6 (95.7)	98.1 (95.3)
R, R_{free}, test (%) [d]	11.7, 16.7, 5.0	14.1, 16.5, 5.0	11.6, 16.2, 5.0	13.7, 16.2, 5.0
RMSD [e] bond length (Å)	0.013	0.012	0.011	0.013
RMSD [e] angles (deg)	1.516	1.469	1.407	1.530
ESU[f] (Å)	0.048	0.051	0.045	0.050
Ramachandran Plot				
Allowed regions (%)	99.6	99.7	99.7	99.7
Additional regions (%)	0.4	0.3	0.3	0.3

[a] Values for the highest resolution shell are given in parentheses. [b] Redundancy-independent merging R factor: $R_{r.i.m} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [c] Precision-indicating merging R factor: $R_{p.i.m} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [d] $R = (\sum||F_{obs}|-k|F_{calc}||)/\sum|F_{obs}|$, where k is a scale factor. The R_{free} value was calculated with the indicated percentage of reflections not used in the refinement. [e] Root-mean-square deviation (RMSD) from ideal geometry in the final models. [f] Estimated overall coordinate error (ESU) based on maximum likelihood.

Table S4. Comparison of selected distances for FeHPCD and H200X variants in the resting state.

	FeHPCD[a] (PDB 3OJT)	H200E[a] (PDB 4Z6L)	H200Q[a] (PDB 4Z6M)	H200N[a] (PDB 4Z6N)
Fe – H155 (Å)	2.20 ± 0.03	2.26 ± 0.03	2.21 ± 0.02	2.20 ± 0.03
Fe – H214 (Å)	2.22 ± 0.03	2.25 ± 0.02	2.21 ± 0.02	2.21 ± 0.02
Fe – E267 (Å)	2.04 ± 0.03	2.11 ± 0.03	2.03 ± 0.02	2.05 ± 0.04
Fe – Wat1[b] (Å)	2.23 ± 0.05	2.23 ± 0.04	2.17 ± 0.01	2.21 ± 0.03
Fe – Wat2[b] (Å)	2.19 ± 0.06	2.11 ± 0.04	2.21 ± 0.00	2.21 ± 0.09
Fe – Wat3[b] (Å)	2.13 ± 0.02	2.17 ± 0.03	2.16 ± 0.01	2.17 ± 0.01
Y257 – Wat2 (Å)	2.43 ± 0.06	2.41 ± 0.07	2.54 ± 0.03	2.53 ± 0.05
N157 – Wat3 (Å)	2.87 ± 0.10	2.90 ± 0.05	2.92 ± 0.01	2.90 ± 0.04
Res200 – Wat3 (Å)	2.87 ± 0.10	2.6 ± 0.1	2.62 ± 0.01	3.36 ± 0.04
Res200 – Wat1 (Å)	3.64 ± 0.03	2.69 ± 0.08	3.02 ± 0.04	4.80 ± 0.04
Res200 – Y269 (Å)	4.77 ± 0.06	3.13 ± 0.08	3.18 ± 0.03	4.0 ± 0.1
D272 – Y269 (Å)	2.82 ± 0.05	3.00 ± 0.03	2.73 ± 0.02	2.74 ± 0.02
WatN[c] – N200 (Å)				2.64 ± 0.04
WatN[c] – Wat1 (Å)				2.90 ± 0.05
WatN[c] – Y269 (Å)				2.9 ± 0.1

[a] average values and standard deviations for distances calculated for 4 monomeric subunits of FeHPCD and H200X variants. [b] Wat1, Wat2 and Wat3 designate solvent-derived ligands coordinated to the iron in the resting state of the enzyme in the positions trans to the residues H214, H155 and E267, respectively. In the high-resolution structures of all FeHPCD enzymes, slightly asymmetric rather than spherical electron density is apparent at Wat2 coordination sites. Since no small ligand accounting for the observed density is present in purification or crystallization solutions, a single solvent was modeled at these sites. [c] WatN designates new crystallographically observed solvent in the active site of the H200N variant of FeHPCD, located in the pocket formed by Y269/Wat1/N200.

Table S5. X-ray data collection and refinement statistics for FeHPCD anaerobic complex with 4SC [a].

Dataset	FeHPCD-[4SC]
(PDB Code)	(PDB 4Z6Z)
Wavelength	0.9762 Å
Synchrotron (beamline)	DLS (I-03)
Spacegroup	$P2_12_12$
Cell dimensions (Å)	110.4, 151.2, 96.2
Cell angles (deg)	90, 90, 90
Resolution range[a] (Å)	29.7 – 1.52 (1.60)
Reflections (observed/unique)	1094337/239601
R_{rim} [a, b] (%)	7.7 (76.8)
R_{pim} [a, c] (%)	3.4 (35.9)
Mean $\langle I\rangle/\sigma\langle I\rangle$ [a]	12.0 (2.1)
Completeness (%) [a]	97.5 (93.9)
R, R_{free}, test (%) [d]	12.1, 16.1, 5.0
RMSD [e] bond length (Å)	0.009
RMSD [e] angles (deg)	1.289
ESU [f] (Å)	0.047
Ramachandran Plot	
Allowed regions (%)	99.7
Additional regions (%)	0.3

[a] Values for the highest resolution shell are given in parentheses. [b] Redundancy-independent merging R factor: $R_{r.i.m} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [c] Precision-indicating merging R factor: $R_{p.i.m} = \sum_{hkl}[1/(N-1)]^{1/2} \sum_i|I_i(hkl)-\langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean value of $I(hkl)$.[13] [d] $R = (\sum||F_{obs}|-k|F_{calc}||)/\sum|F_{obs}|$, where k is a scale factor. The R_{free} value was calculated with the indicated percentage of reflections not used in the refinement. [e] Root-mean-square deviation (RMSD) from ideal geometry in the final models. [f] Estimated overall coordinate error (ESU) based on maximum likelihood.

Table S6. Comparison of selected distances for FeHPCD and H200X variants in complex with optimal substrate HPCA.

	FeHPCD-HPCA[a] (PDB 4GHG)	H200E-HPCA[b] (PDB 4Z6O)	H200Q-HPCA[a] (PDB 4Z6P)	H200N-HPCA[b] (PDB 4Z6Q)
Fe – Wat3[c] (Å)	2.43	2.31 ± 0.06	2.40	2.6 ± 0.1
N157 – Wat3[c] (Å)	2.98	2.84 ± 0.05	2.81	2.87 ± 0.02
Res200 – Wat3[c] (Å)	3.00	2.67 ± 0.01	2.72	3.2 ± 0.2
$O_{C1\text{-}HPCA}$ – Wat3[c] (Å)	3.19	2.94 ± 0.04	3.09	2.98 ± 0.01
$O_{C2\text{-}HPCA}$ – Wat3[c] (Å)	3.12	2.88 ± 0.04	3.04	3.16 ± 0.08
Fe-$O_{C1\text{-}HPCA}$ (Å)	2.19	2.13 ± 0.09	2.09	2.12 ± 0.06
Fe-$O_{C2\text{-}HPCA}$ (Å)	2.09	2.05 ± 0.03	2.14	2.06 ± 0.01
$C1_{HPCA}$-$O_{C1\text{-}HPCA}$ (Å)	1.38	1.38 ± 0.05	1.39	1.37 ± 0.02
$C2_{HPCA}$-$O_{C2\text{-}HPCA}$ (Å)	1.36	1.35 ± 0.01	1.36	1.31 ± 0.01
$O_{C2\text{-}HPCA}$ – Y257 (Å)	2.57	2.62 ± 0.03	2.61	2.61 ± 0.04
$O_{C1\text{-}HPCA}$ – Res200 (Å)	3.05	2.51 ± 0.09	2.82	4.44 ± 0.03
Y269 – Res200 (Å)	4.44	3.03 ± 0.00	3.11	3.96 ± 0.04
Y269 – D272 (Å)	2.70	3.14 ± 0.01	2.76	2.80 ± 0.00
WatN[d] – Res200 (Å)				2.50 ± 0.04
WatN[d] – $O_{C1\text{-}HPCA}$ (Å)				2.64 ± 0.03
WatN[d] – Y269 (Å)				2.79 ± 0.01

[a] values for distances calculated for subunit C. [b] average values and standard deviations for distances calculated for subunits C and D. [c] Wat3 designate solvent-derived ligands coordinated to the iron in the position trans to the residue E267. [d] WatN designates new crystallographically observed solvent in the active site of the H200N variant of FeHPCD, located in the pocket formed by Y269/Wat1/N200.

Table S7. Comparison of selected distances for FeHPCD and H200X variants in complex with alternative substrate 4SC.

	FeHPCD-4SC[a] (PDB 4Z6Z)	H200E-4SC[b] (PDB 4Z6R)	H200Q-4SC[a] (PDB 4Z6S)	H200N-4SC[b] (PDB 4Z6T)
Fe – Wat3[c] (Å)	2.35 ± 0.03	2.27	2.38 ± 0.08	2.59
N157 – Wat3[c] (Å)	2.91 ± 0.02	2.76	2.85 ± 0.09	2.86
Res200 – Wat3[c] (Å)	2.83 ± 0.05	2.75	2.70 ± 0.03	3.06
$O_{C1\text{-}4SC}$ – Wat3[c] (Å)	2.97 ± 0.09	2.88	2.84 ± 0.04	2.77
$O_{C2\text{-}4SC}$ – Wat3[c] (Å)	2.98 ± 0.08	2.89	3.00 ± 0.00	3.27
Fe – $O_{C1\text{-}4SC}$ (Å)	2.11 ± 0.03	2.10	2.02 ± 0.04	2.00
Fe – $O_{C2\text{-}4SC}$ (Å)	2.10 ± 0.02	2.15	2.11 ± 0.05	2.11
$C1_{4SC}$ – $O_{C1\text{-}4SC}$ (Å)	1.39 ± 0.00	1.39	1.37 ± 0.01	1.36
$C2_{4SC}$ – $O_{C2\text{-}4SC}$ (Å)	1.36 ± 0.01	1.34	1.36 ± 0.01	1.35
$O_{C2\text{-}4SC}$ – Y257 (Å)	2.53 ± 0.02	2.42	2.53 ± 0.06	2.52
$O_{C1\text{-}4SC}$ – Res200 (Å)	3.1 ± 0.1	2.47	2.76 ± 0.04	4.22
Y269 – Res200 (Å)	3.78 ± 0.01	2.99	3.15 ± 0.03	3.81
Y269 – D272 (Å)	2.77 ± 0.01	3.05	2.78 ± 0.02	2.79

[a] average values and standard deviations for distances calculated for subunits C and D. [b] values for distances calculated for subunit C. [c] Wat3 designates solvent-derived ligands coordinated to the iron in the position trans to the residue E267.

Table S8. Comparison of selected distances for FeHPCD and H200X variants in complex with alternative substrate 4NC.

	FeHPCD-4NC[a] (PDB 4GHH)	H200E-4NC[b] (PDB 4Z6U)	H200Q-4NC[b] (PDB 4Z6V)	H200N-4NC[c] (PDB 4Z6W)
Fe – Wat3[d] (Å)	2.39 ± 0.06	2.16	2.24	2.43 ± 0.08
N157 – Wat3[d] (Å)	2.91 ± 0.01	2.82	2.78	2.81 ± 0.04
Res200 – Wat3[d] (Å)	2.93 ± 0.06	2.72	2.83	3.15 ± 0.04[f]
$O_{C1\text{-}4NC}$ – Wat3[d] (Å)	2.95 ± 0.20	2.81	2.82	2.96 ± 0.02
$O_{C2\text{-}4NC}$ – Wat3[d] (Å)	3.0 ± 0.1	2.88	2.93	3.2 ± 0.1
Fe – $O_{C1\text{-}4NC}$ (Å)	2.13 ± 0.01	2.10	2.08	2.10 ± 0.02
Fe – $O_{C2\text{-}4NC}$ (Å)	2.17 ± 0.01	2.10	2.07	2.10 ± 0.04
$C1_{4NC}$ – $O_{C1\text{-}4NC}$ (Å)	1.37 ± 0.00	1.35	1.34	1.34 ± 0.03
$C2_{4NC}$ – $O_{C2\text{-}4NC}$ (Å)	1.35 ± 0.00	1.34	1.36	1.33 ± 0.01
$O_{C2\text{-}4NC}$ – Y257 (Å)	2.54 ± 0.08	2.46	2.62	2.52 ± 0.05
$O_{C1\text{-}4NC}$ – Res200 (Å)	2.98 ± 0.07	2.52	2.77	4.49 ± 0.05[f]
Y269 – Res200 (Å)	4.51 ± 0.08	3.16	3.18	3.85 ± 0.07[f]
Y269 – D272 (Å)	2.75 ± 0.02	2.99	2.72	2.81 ± 0.03
WatN[e] – Res200 (Å)				2.3 ± 0.1[f] 3.5 ± 0.2[g]
WatN[e] – $O_{C1\text{-}4NC}$ (Å)				2.72 ± 0.08
WatN[e] – Y269 (Å)				2.91 ± 0.09

[a] average values and standard deviations for distances calculated for subunits C and D. [b] values for distances calculated for subunit C. [c] average values and standard deviations for distances calculated for subunits A – D. [d] Wat3 designate solvent-derived ligands coordinated to the iron in the position trans to the residue E267. [e] WatN designates new crystallographically observed solvent in the active site of the H200N variant of FeHPCD, located in the pocket formed by Y269/Wat1/N200. [f] N200 conformer-1 ("in"). [g] N200 conformer-2 ("out").

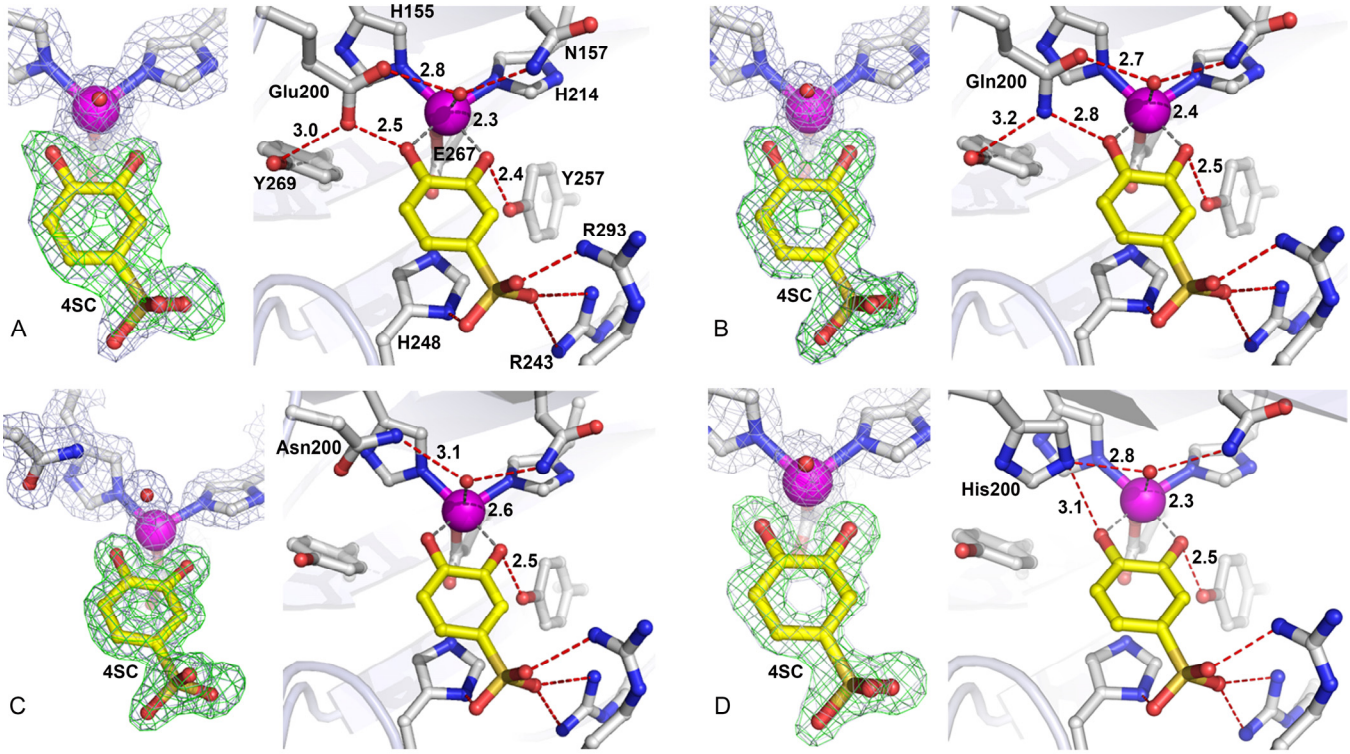

Figure S1. Substrate coordination environments in the active sites of FeHPCD and H200X variants in complex with 4SC. (A) H200E-[4SC] complex (PDB 4Z6R), (B) H200Q-[4SC] complex (PDB 4Z6S), (C) H200N-[4SC] (PDB 4Z6T), (D) FeHPCD-[4SC] (PDB 4Z6Z). The blue $2F_{obs}-F_{calc}$ maps are contoured at 1.0 – 1.4 σ. The green ligand omit $F_{obs}-F_{calc}$ maps are contoured at 6 – 8 σ. Atom color code: gray, carbon (enzyme); yellow, carbon (substrate); blue, nitrogen; red, oxygen; purple, iron. Red dashed lines show hydrogen-bonds (Å). Gray dashed lines indicate bonds or potential bonds to iron (Å). Cartoons depict secondary structure elements. Representative complexes of H200X variants with HPCA and 4NC are shown in Figure 3.

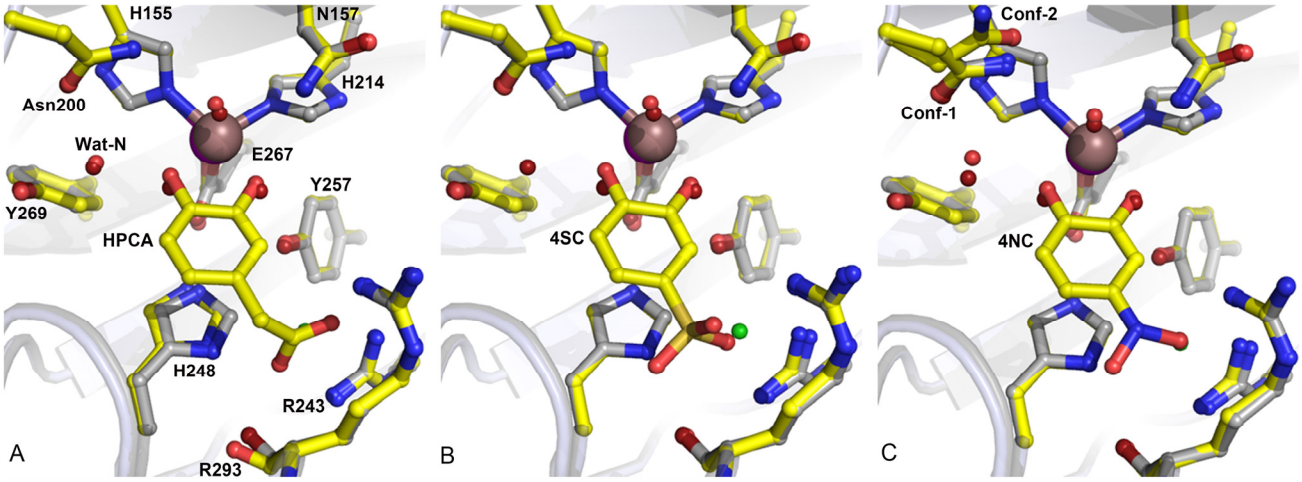

Figure S2. Active site comparison of H200N in the resting state and in complex with aromatic substrates (subunits C shown). H200N variant (PDB 4Z6N) and complex with HPCA (A, PDB 4Z6Q), 4SC (B, PDB 4Z6T) and 4NC (C, PDB 4Z6W). Atom color code: gray, carbon (resting state); yellow, carbon (complex); blue, nitrogen; dark red, oxygen (resting state); red, oxygen (complex); bronze, iron (resting state); purple, iron (complex); green, chlorine (resting state).

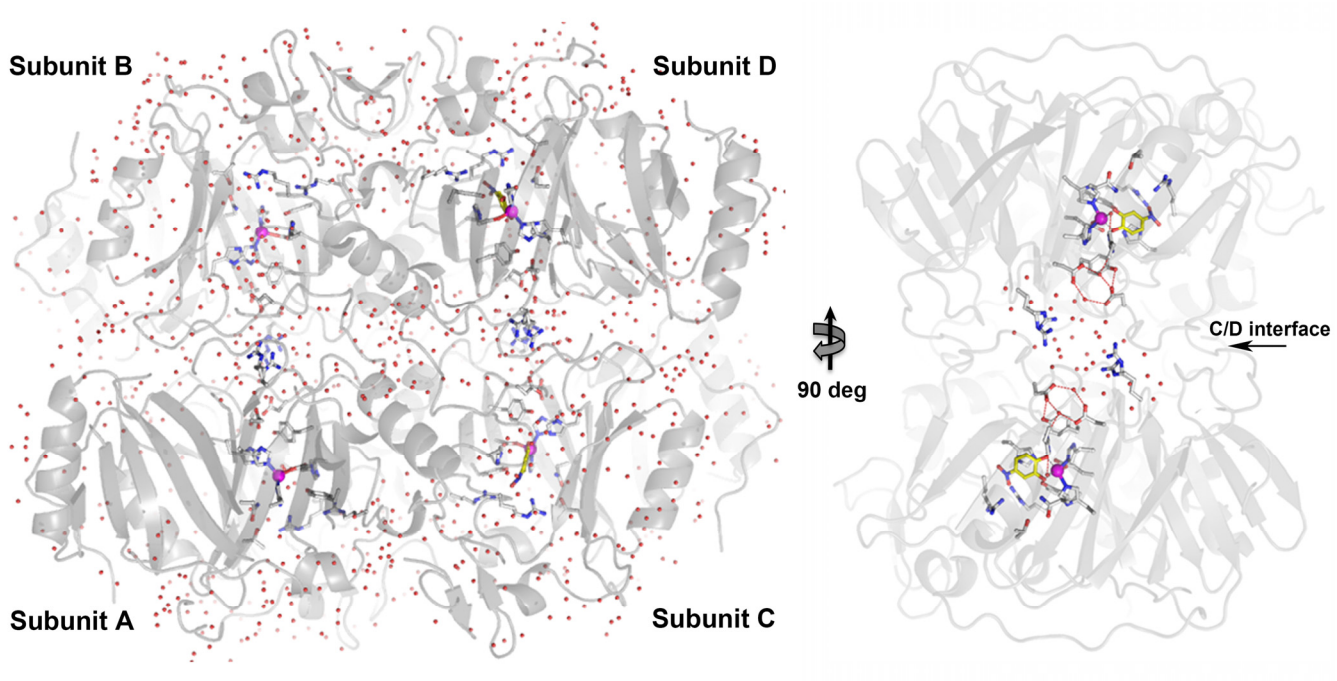

Figure S3. Solvent accessible channels within FeHPCD tetramer and dimer interface. Crystallographically observed solvent is shown as red spheres. For clarity, only selected solvent molecules in the vicinity of R152 are shown in the dimer interface panel. Atom color code: gray, carbon (enzyme); yellow, carbon (substrate); blue, nitrogen; red, oxygen; purple, iron. Red dashed lines show hydrogen-bonds defining proposed proton relay pathway (e.g. Figure 5, PDB 4GHH[7]) that connects 4NC bound in the active site to solvent in the intersubunit space (Å). Cartoons depict secondary structure elements.

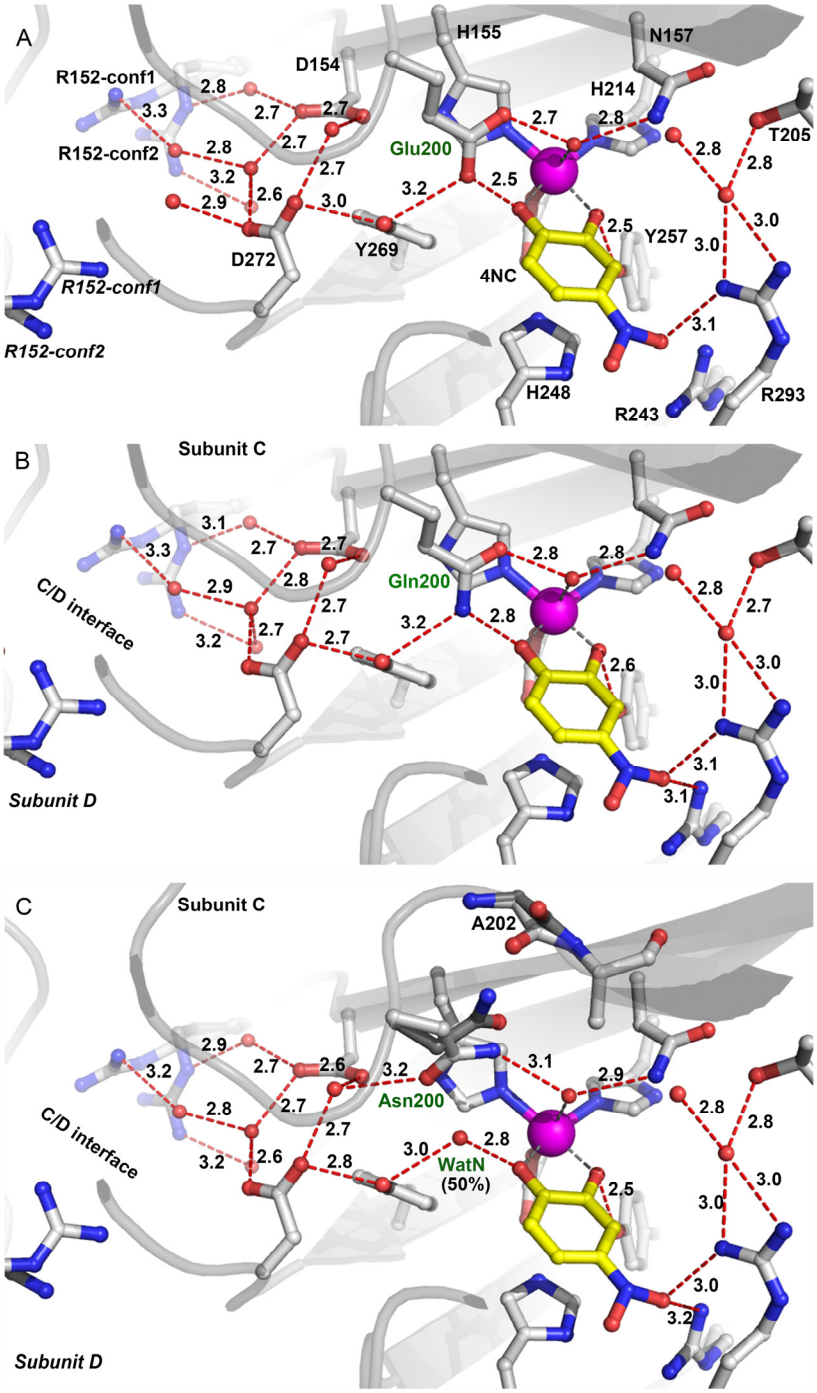

Figure S4. Proton relay pathways in the active sites of the H200E, H200Q and H200N variants in complex with 4NC. (A) H200E-[4NC] complex (PDB 4Z6U, subunit C). (B) H200Q-[4NC] complex (PDB 4Z6V, subunit C). (C) H200N-[4NC] complex (PDB 4Z6W, subunit C). Two conformers of N200 and A202 are shown. Atom color code is the same as that described for Figure S1. Red dashed lines show hydrogen-bonds (Å). Gray dashed lines indicate bonds or potential bonds to iron. Pathway for the corresponding complex in FeHPCD is depicted in Figure 5.

SUPPLEMENTAL REFERENCES:

(1) Wang, Y. Z., and Lipscomb, J. D. (1997) Cloning, overexpression, and mutagenesis of the gene for homoprotocatechuate 2,3-dioxygenase from *Brevibacterium fuscum*, *Protein Exp. Purif. 10*, 1-9.

(2) Groce, S. L., and Lipscomb, J. D. (2005) Aromatic ring cleavage by homoprotocatechuate 2,3-dioxygenase: Role of His200 in the kinetics of interconversion of reaction cycle intermediates, *Biochemistry 44*, 7175-7188.

(3) Mbughuni, M. M., Meier, K. K., Münck, E., and Lipscomb, J. D. (2012) Substrate-mediated oxygen activation by homoprotocatechuate 2,3-dioxygenase: Intermediates formed by a tyrosine 257 variant, *Biochemistry 51*, 8743-8754.

(4) Miller, M. A., and Lipscomb, J. D. (1996) Homoprotocatechuate 2,3-dioxygenase from *Brevibacterium fuscum* - A dioxygenase with catalase activity, *J. Biol. Chem. 271*, 5524-5535.

(5) Kovaleva, E. G., and Lipscomb, J. D. (2007) Crystal structures of Fe^{2+} dioxygenase superoxo, alkylperoxo, and bound product intermediates, *Science 316*, 453-457.

(6) Kovaleva, E. G., and Lipscomb, J. D. (2008) Intermediate in the O-O bond cleavage reaction of an extradiol dioxygenase, *Biochemistry 47*, 11168-11170.

(7) Kovaleva, E. G., and Lipscomb, J. D. (2012) Structural basis for the role of tyrosine 257 of homoprotocatechuate 2,3-dioxygenase in substrate and oxygen activation, *Biochemistry 51*, 8755-8763.

(8) Kabsch, W. (2010) XDS, *Acta Crystallogr D Biol Crystallogr. 66*, 125-132.

(9) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D53*, 240-255.

(10) Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Longa, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Cryst. D67* 355-367.

(11) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., and Evans, P. R. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D 67*, 235-242.

(12) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. E60*, 2126-2132.

(13) Weiss, M. S. (2001) Global indicators of X-ray data quality, *J. Appl. Cryst. 34*, 130-135.