

Supporting online material

Crystal Structures of Fe²⁺ Dioxygenase Superoxo, Alkylperoxo, and Bound Product Intermediates

Elena G. Kovaleva and John D. Lipscomb*

Department of Biochemistry, Molecular Biology, and Biophysics,
University of Minnesota, Minneapolis, MN 55455

*To whom correspondence should be addressed. E-mail: lipsc001@umn.edu

Materials and methods:

Expression and purification

2,3-homoprotocatechuate dioxygenase (2,3-HPCD) (EC 1.13.11.15) from *Brevibacterium fuscum* was expressed in *Echerichia coli* strain BL21(DE3) and purified as described previously (S1) with the following exceptions: 1) ampicillin was included only in seed cultures, 2) Iron-Cys complex was omitted from purification buffers to eliminate auxiliary binding of iron to the enzyme. The enzyme was pure based on SDS-PAGE electrophoresis, and iron content was 75-80 % as determined by ICP-MS.

Crystallization and preparation of complexes

The crystals of full length 2,3-HPCD were grown in 18% PEG 8000, 0.2 M calcium acetate, 0.1 M sodium cacodylate, pH 6.5 (Hampton Research) by a hanging or sitting drop method at 19 °C either inside an anaerobic glove box (Belle Technology) or under air.

For complex preparation, all mother liquor solutions, crystals of 2,3-HPCD, and 4-nitrocatechol (4NC) stocks were equilibrated in the glove box atmosphere for at least 12 hours prior to mixing. Reaction was initiated by transfer of enzyme crystals to crystal growth solution (18% PEG 8000, 0.2 M calcium acetate, 0.1 M sodium cacodylate, pH 6.5) containing 2 mM 4NC at 19 °C. After soaking for 30-60 min, crystals were rapidly and sequentially transferred into mother liquor solutions containing 5, 10, 15, and then 20 % glycerol prior to cryofreezing in liquid nitrogen inside the glove box.

Complex preparation was initially carried out under air atmosphere (data not shown). However, the X-ray structure revealed that despite high overall resolution (1.5 Å), the electron density corresponding to the active site ligands was somewhat diffuse, suggesting that more than one intermediate was present, perhaps due to the high concentration of oxygen in the reaction solution. This problem was resolved by preparing the crystal complexes inside the low oxygen glove box where residual concentrations of

O₂ are still sufficient to initiate the catalytic reaction of 2,3-HPCD due to its high affinity for O₂. While it may seem surprising that the oxygenase reaction can be initiated inside an “anaerobic” glove box, ring cleaving dioxygenases are routinely used as oxygen scrubbers for systems in which true anaerobic conditions are desired (S2). Accordingly, we have observed that crystals soaked in HPCA in the glove box slowly turn yellow due to product formation beginning with the surface facing the box atmosphere.

Data collection and refinement

X-ray data were collected at the Advanced Light Source (ALS, LBL, Berkeley, CA) on beamline 4.2.2. at 1.24 Å and 100 K. All diffraction data were processed with D*Trek (Molecular Structure Corp). The structures were refined by cycles of restrained refinement with Refmac5 (S3) and model building using Coot (S4). The coordinates for truncated 2,3-HPCD (1F1X, (S5)) were used as a molecular replacement model, and the C-terminal residues (323-362) were built into the density from $F_{\text{obs}}-F_{\text{calc}}$ map. Link restraints to Fe were removed from the refinement to avoid bias in the refined Fe-ligand distances.

Ligand refinement and molecular modeling

4NC was used as the initial ligand model in the active sites. Then, based on the resultant $F_{\text{obs}}-F_{\text{calc}}$ map, appropriate ligands were modeled into the density to obtain the best fit of the data (see additional information below). Ligand dictionaries for the 4NC-semiquinone form, alkylperoxo intermediate, and open-ring products were produced with Monomer Sketcher (S6). Ligand occupancy was estimated to be ~75% for the 4NC-semiquinone and alkylperoxo intermediates and ~60% for the more flexible product based on the average temperature factors (that are similar to those of the protein at these ligand occupancies) and the lack of $F_{\text{obs}}-F_{\text{calc}}$ density contoured at 3.0 – 3.5 σ .

Model bias associated with ligand refinement was evaluated by examining ligand-omit difference maps. Active site ligands were removed from the final refined protein model, and ligand-free model was allowed to refine with 5 cycles of restrained refinement. The resultant positive density in the difference (ligand omit) $F_{\text{obs}}-F_{\text{calc}}$ maps was found to agree with the refined positions and structures of the active site ligands in the full models (e.g. Fig 3).

In the free enzyme, strong spherical density is observed in the pocket created by His248/Arg243/Arg293 at the bottom of the substrate-binding cleft, and it is assigned to Cl⁻ anion based on the following observations: 1) refinement of a water molecule results in unreasonably low temperature factor and an excess of positive $F_{\text{obs}}-F_{\text{calc}}$ density remains at the site, 2) Cl⁻ is the major anion present in the last purification step (~ 0.3 M NaCl) and can account for all of the observed electron density. The position of the bound anion (Cl⁻) in the free enzyme essentially overlaps with the position of the oxygen atom from the nitro substituent of the 4NC-derived ligands. Since ligands in the reactive complex structures appear to be present at incomplete occupancy, it is likely that the anion would occupy the His248/Arg243/Arg293 pocket in their absence. Inclusion of Cl⁻ anion at partial occupancy during refinement normalized the temperature factor for the oxygen atoms of nitro moiety and eliminated slight excess of positive $F_{\text{obs}}-F_{\text{calc}}$ density at the site.

The $2F_{\text{obs}}-F_{\text{calc}}$ density for C3-C4 bond of 4NC and its intact ring derivatives bound to 2,3-HPCD was often found to be weaker than that for other C-C bonds of the ring. X-ray structures of several inactive enzyme complexes with 4NC, where presence of intact 4NC ring is confirmed spectrophotometrically (data to be described elsewhere), also exhibit lower electron density for C3-C4 bond. The 4NC ring is also planar in these complexes in contrast to that in the superoxo- Fe^{2+} -4NC semiquinone structure described here. The weaker density for C3-C4 bond of 4NC may be due to ring puckering or partial single bond character during species interconversion. Alternatively, it may be due to dynamic fluctuations resulting from non-optimal binding due to shorter ring substituent (nitro group vs carboxymethyl).

Reproducibility

The distribution of reactive states in the tetrameric enzyme reported here (1 substrate-semiquinone, 2 alkylperoxy intermediates and 1 product) was also observed in several additional independently refined structural data sets obtained from crystals prepared using the same procedure described (data not shown). This fact affirms a physical basis for the observed differential subunit reactivity in crystalline form.

Table S1. X-Ray data collection and refinement statistics^a.

Dataset	WT 2,3-HPCD	WT reacted with 4NC/O ₂
PDB code	2IG9	2IGA
Space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2
Monomer/AU	4	4
Cell dimensions (Å)	152.8, 110.4, 99.7	153.0, 110.7, 96.4
Cell angles (deg)	90, 90, 90	90, 90, 90
Resolution range ^a (Å)	53.2-1.90 (1.97)	52.7-1.95 (2.02)
Reflections (observed/unique)	422505/132886	219014/105058 ^f
R_{merge} ^{a, b} (%)	9.9 (53.2)	9.5 (37.9)
Mean $\langle I \rangle / \sigma \langle I \rangle$ ^a	7.1 (2.0)	5.4 (1.9)
Completeness (%) ^a	99.8 (99.9)	87.8 (91.6)
R , R_{free} , test (%) ^c	18.7, 23.4, 5.0	18.5, 24.5, 5.0
RMSD ^d bond length (Å)	0.014	0.018
RMSD ^d angles (deg)	1.455	1.676
ESU ^e (Å)	0.118	0.123
Ramachandran Plot		
Allowed regions (%)	99.7	99.7
Additional regions (%)	0.3	0.3

^a Values for the highest resolution shell are given in parenthesis. ^b $R_{\text{merge}} = \sum |I_i - \langle I \rangle| / \sum |I_i|$, where I_i is the integrated intensity of a given reflection, and $\langle I \rangle$ is the mean value for that reflection. ^c $R = (\sum |F_{\text{obs}} - kF_{\text{calc}}|) / \sum |F_{\text{obs}}|$, where k is a scale factor. The R_{free} value was calculated with the indicated percentage of reflections not used in the refinement. ^d Root-mean-square deviation (RMSD) from ideal geometry in the final models. ^e Estimated overall coordinate error (ESU) based on maximum likelihood. ^f Higher redundancy data (e.g. 6-12) does not improve or affect observed ligand density as inferred from X-ray data collected for various enzyme-substrate complexes prepared similarly (data to be described elsewhere).

Additional Supporting Information:

Ligand modeling

As stated earlier, active site ligands were not included in the initial refinement of the protein models and were built into the density from $F_{\text{obs}}-F_{\text{calc}}$ map. When 4NC was used as the initial ligand model in the active sites of all subunits, it was clear from the resultant $F_{\text{obs}}-F_{\text{calc}}$ map that: 1) none of the sites contain 4NC in an unaltered form, 2) three different species derived from 4NC are represented in the four subunits, and 3) a small molecule larger than water was also present in iron ligation in three of the four subunits. Therefore, rather than treating all subunits (including ligands) as identical, model building and refinement of the active site ligands for each subunit was approached individually. This was accomplished by not using NCS restraints and by sequentially fitting the density features present in each subunit with candidate ligands drawn from the mechanistic hypotheses that we and others have put forward until the best fit of the data was obtained. Ligands considered included 4NC (Fig. 1, I), 4NC-semiquinone (Fig. 1, II), alkylperoxy (Fig.1, III) and lactone (Fig.1, IV) intermediates, open-ring products (Fig. S3), as well as epoxide and dioxetane intermediates.

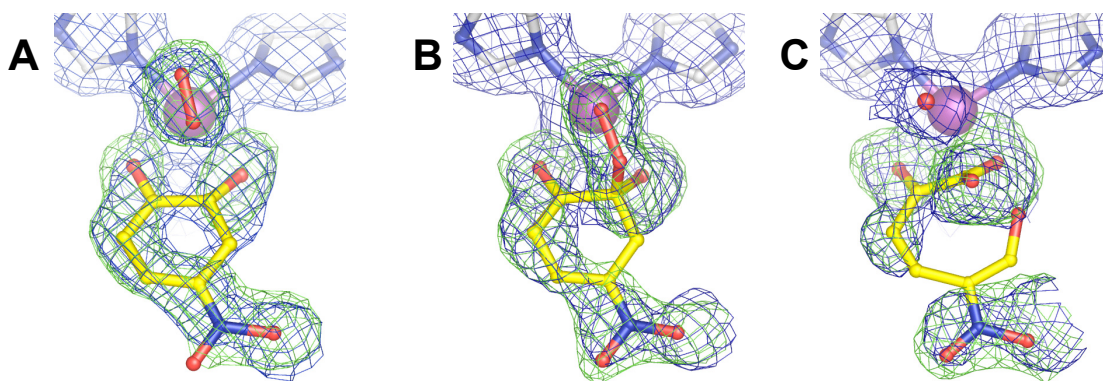


Fig. S1. Additional views for the structures of reactive species shown in Fig. 3. **A)** Ternary complex (subunit C) with 4NC semiquinone and side-on bound dioxygen species. **B)** Alkylperoxy intermediate complex (subunit D). **C)** Complex with open-ring product (subunit A). The blue $2F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at 1.0σ (A, B) and 0.75σ (C). The green $F_{\text{obs}}-F_{\text{calc}}$ maps were computed without ligands present in the model and are contoured at $+4 \sigma$ (A), $+3.75 \sigma$ (B) and $+3 \sigma$ (C). Atom color code: gray, carbon (enzyme residues); yellow, carbon (ligands); blue, nitrogen; red, oxygen; purple, iron.

Product complex

The discontinuous ligand density in the active site of subunit A (Fig. 3C) is consistent with the presence of a more flexible ligand than the 4NC-derived ligands with an intact ring. This assignment is supported by similarity of density distribution observed in the structure of the enzyme-product complex prepared by direct soaking of the enzyme

with the open-ring product (Fig. S2, data to be described elsewhere). The common density features for the enzyme-product complex include: 1) strong density for the Fe-bonded atoms and for the carboxylate moiety, which is rotated slightly to distance its non-bonded oxygen atom from the Fe-bound water in the adjacent site, 2) strong, but globular density near the anion-binding site corresponding to the aldehyde/nitro moieties, 3) weak density around C5 consistent with a more flexible open ring ligand, 4) *cis* conformation rather than the thermodynamically favored extended conformation is observed due to constraints in the active site.

As discussed in the main text, the orientation of the aldehyde and nitro substituents for the enzyme-product complex could not be resolved and multiple conformations could account for the observed density (Fig. S3).

In subunit A of the reactive complex, lack of density between C2 and C3 also rules out an epoxide as a possible ligand.

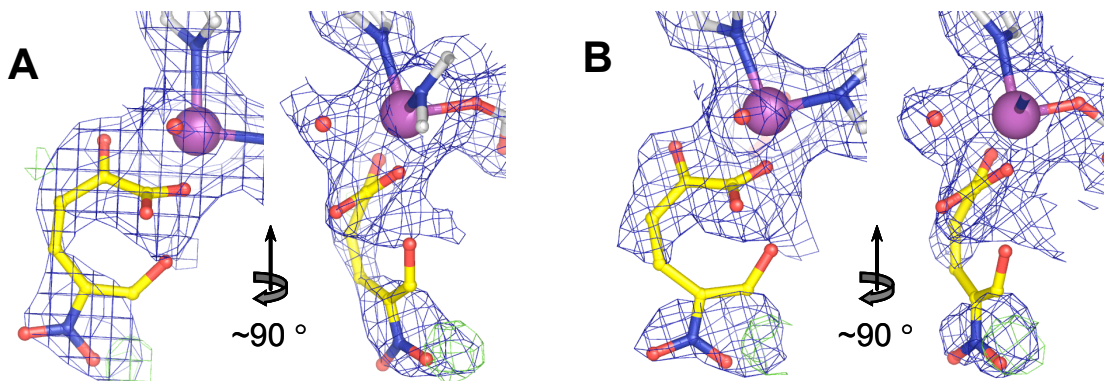


Fig. S2. Examples of electron density maps observed for enzyme-product complex prepared by direct soaking of the open-ring product rather than by carrying out the reaction *in crystallo* (c.f. Fig 3C). A) subunit C, B) subunit D. Structures (to be described elsewhere) were refined to 2.40 Å with R/R_{free} values of 18.5% and 25.3%, respectively. The blue $2F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at 0.75 σ (same as structures shown in Fig. 3C), and the $F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at -3.25 σ (red) and 3.25 σ (green). Alternative conformations of the open-ring product could account for the observed density; the same conformation as in Fig 3C is shown.

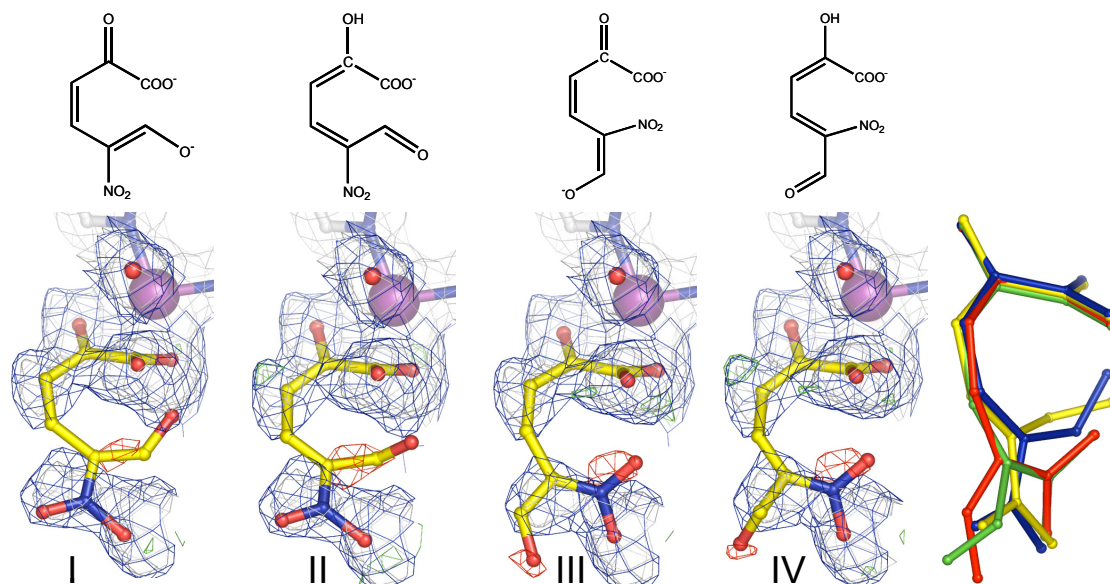


Fig. S3. Alternative conformations and resonance forms of open-ring product modeled in subunit A of the reactive complex structure. The $2F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at 1σ (grey) and 0.5σ (blue). The $F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at -3.0σ (red) and 3.0σ (green). Product overlay color code: blue (I), yellow (II), red (III) and green (IV). Structure I corresponds to that shown in Fig. 3C.

Side-on oxygen binding

The following observations support the presence of a diatomic molecule bound side-on to the Fe in subunit C: 1) a single water molecule modeled in place of dioxygen results in excess of positive $F_{\text{obs}}-F_{\text{calc}}$ density (Fig. S4), 2) two solvent molecules at that location refine to an unreasonably close value of 1.3 \AA apart, 3) partial tetrahedral character of C2 of the bound 4NC-semiquinone as indicated by the out of plane position of the C2-hydroxyl and puckering of the ring (Fig. 3A) is possible only in the presence of bound oxygen (as an electron acceptor) and not solvent, 4) there is no density to support the end-on binding mode of dioxygen. These results do not rule out the possibility that O_2 initially binds end-on before assuming the side-on orientation observed.

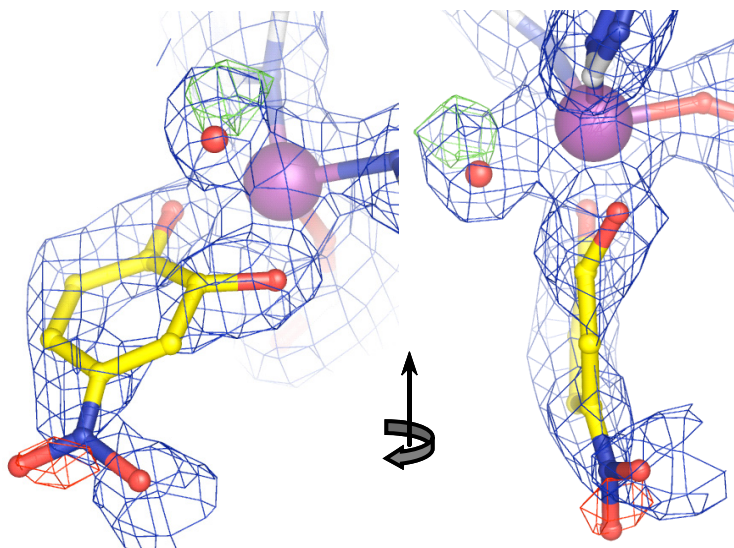


Fig. S4. Electron density maps for 4NC-semiquinone and a solvent (instead of dioxygen) in the active site of subunit C of the reacted complex structure. The $2F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at 1σ (blue), and the $F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at -3.5σ (red) and 3.5σ (green).

Alkylperoxo intermediate

The formation of a bond between oxygen and the ring of the substrate is evident from the ligand omit maps that show continuous positive density from C2 of the substrate to the diatomic ligand. The difference maps were calculated in the *absence* of either the entire ligand (e.g. Fig. 3B) or in the absence of peroxo moiety (i.e. semiquinone used instead of alkylperoxo intermediate in subunits B and D) (Fig. S5).

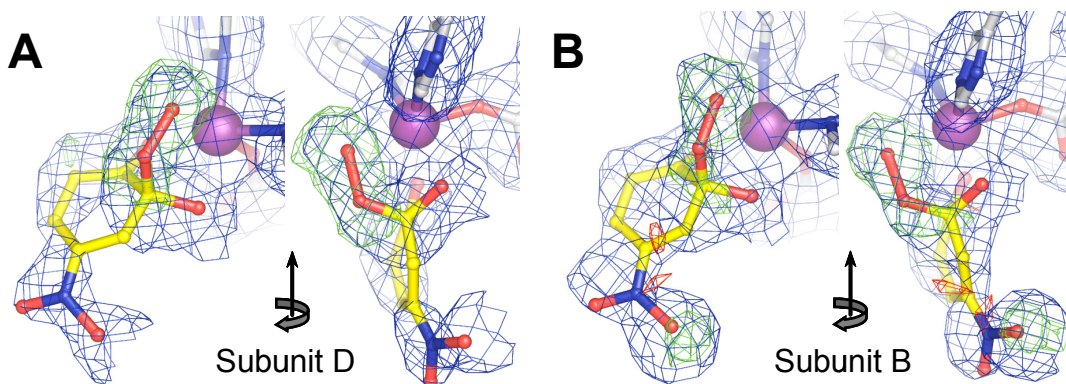


Fig. S5. Omit maps for dioxygen-derived atoms calculated using semiquinone instead of alkylperoxo intermediate in subunits B and D. The $2F_{\text{obs}}-F_{\text{calc}}$ maps are contoured at 0.75σ (A) and 0.5σ (B), and the $F_{\text{obs}}-F_{\text{calc}}$ omit maps calculated in the absence of peroxo moiety (i.e. semiquinone) are contoured at 3.5σ (A) and 3.25σ (B).

Short hydrogen bonds

The structures shown in Fig. 3 contain several short hydrogen bonds indicating particular strength. The short bonds occur in three places. One is from His200 to O1 of

the bound O₂ in the structure of the superoxo intermediate (Fig. 3A). Transient kinetic and mutagenesis studies suggest that His200 is the critical active site acid that donates a proton to allow the O-O bond to break, thus a short bond here reflects the chemistry of the reaction. The second short bond is from the same His200 to the solvent bound to iron in the product complex (Fig. 3C). Again, this His is ideally placed for the role it must serve, and the solvent is marking the spot where O₂ will bind in the next round of the cycle. It may also play a role in the regulation of the reaction by stabilizing solvent in this iron ligation site, and thus inhibiting O₂ association prior to binding of the anionic substrate, as experimentally observed. Finally, the third short bond is between Tyr257 and the substrate C2 diol oxygen in the alkylperoxo complex (Fig. 3B). In this position, the hydrogen bond stabilizes the tetrahedral sp³ hybridized carbon over the original sp² aromatic carbon. Thus, it promotes the key transition state in the dioxygenase reaction. The short hydrogen bonds appear to all support the proposal for the mechanism of extradiol dioxygenases that have emerged from chemical and computational studies.

Differential subunit reactivity

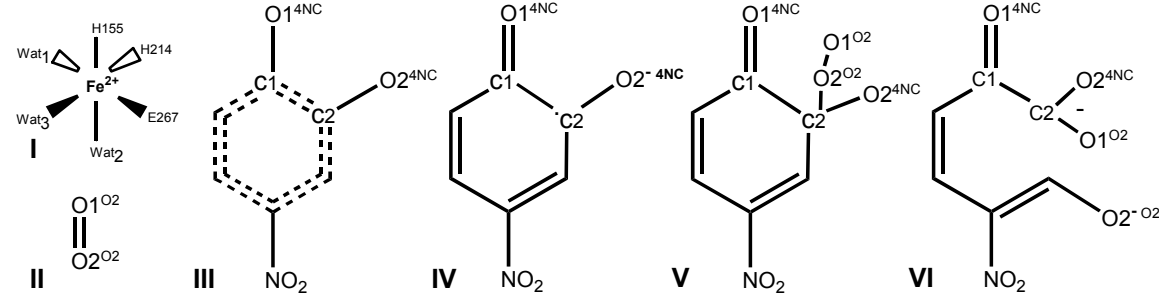
The superposition analysis of the individual subunits in both free and reacted complex structures shows that subunits have similar overall structures as indicated by rmsd values of 0.2 to 0.7 Å. Analysis of the surface area involved in intermolecular (crystal) and intersubunit contacts also shows no indication consistent with the observed differences in subunit reactivity (Table S2, column Δ ASA). In fact, subunits A and B have comparable contact surface areas as do subunits C and D, while only subunits B and D contain the same intermediate. Therefore, it is unlikely that the total amount of contact surface directs the reactivity of the subunits; rather, the specific location and type of crystal packing interactions may be the determining factor.

It is interesting to note that comparison of the surface area involved in intermolecular (crystal) contacts for each subunit in the presence and absence of a bound ligand (i.e. Δ ASA_(free) – Δ ASA_(enzyme-ligand)) does show a correlation with the type of the complex formed. As shown in Table S2 (see column $\Delta\Delta$ ASA), bound semiquinone and product result in the largest and the least increase in crystal contact areas, respectively, relative to the ligand-free state. The $\Delta\Delta$ ASA value is similar for the two subunits that contain the alkylperoxo intermediate. Thus, there is a correlation between the *changes* that occur in the accessible surface area and the specific intermediate in the site. These ligand specific changes may contribute to the dynamics of intermediate interconversion in the *P2₁2₁2* crystalline space group.

Table S2. Analysis of accessible surface area for subunits of 2,3-HPCD due to crystal packing and due to complex formation^a.

	ASA ^b (Å ²)	Isolated ASA ^c (Å ²)	ΔASA ^d (Å ²)	ΔΔASA ^e (Å ²)
Free WT 2,3-HPCD enzyme (PDB code 2IG9)				
Subunit A	10028 (3015)	10844 (3257)	-816 (-242)	
Subunit B	9908 (2974)	10717 (3216)	-809 (-242)	
Subunit C	10363 (3112)	10698 (3213)	-335 (-101)	
Subunit D	10220 (3069)	10686 (3207)	-466 (-138)	
WT 2,3-HPCD enzyme reacted with 4NC/O ₂ (PDB code 2IGA)				
Subunit A	10164 (3052)	11021 (3306)	-857 (-254)	41 (12)
Subunit B	9634 (2893)	10611 (3184)	-977 (-291)	168 (49)
Subunit C	10108 (3032)	10810 (3243)	-702 (-211)	367 (110)
Subunit D	10267 (3083)	10925 (3278)	-658 (-195)	192 (57)

^a Total accessible surface areas (ASA) for monomeric subunits were calculated using program AREAIMOL (S6) Surface areas due to intersubunit contacts are shown in parenthesis. ^b Observed ASA for each monomeric subunit calculated with symmetry-related atoms generated using symmetry operations and lattice translations. ^c Isolated ASA for a each monomeric subunit calculated without symmetry-related atoms (i.e. without accounting for surface area involved in crystal contacts) ^d Differences in surface area due to crystal (intermolecular) contacts for each subunit. ^e Comparison of surface area involved in crystal (intermolecular) contacts for each subunit in the presence and absence of a bound ligand, i.e. $\Delta\text{ASA}_{(\text{free})} - \Delta\text{ASA}_{(\text{enzyme-ligand})}$. In the complex structure, subunit A contains an open-ring product (Fig. 3C), subunits B and D contain alkylperoxo intermediate (e.g. Fig. 3B) and subunit C contains substrate-semiquinone (Fig. 3A).

Table S3. Summary of distances and angles for ligand coordination to Fe^a


Chemical structures I-VI showing ligand coordination to Fe²⁺. Structure I shows Fe²⁺ coordinated to three water molecules (Wat1, Wat2, Wat3) and residues H155, H214, and E267. Structure II shows Fe²⁺ coordinated to two oxygen atoms of a diatomic oxygen molecule (O1^{O2}, O2^{O2}). Structure III shows Fe²⁺ coordinated to O1^{4NC} and O2^{4NC} of a 4NC ligand. Structure IV shows Fe²⁺ coordinated to O1^{4NC} and O2^{4NC} of a 4NC-semiquinone ligand. Structure V shows Fe²⁺ coordinated to O1^{4NC}, O2^{4NC}, and O1^{O2}, O2^{O2} of an alkylperoxy intermediate. Structure VI shows Fe²⁺ coordinated to O1^{4NC}, O2^{4NC}, and O1^{O2} of an open-ring product.

Distances/Angles	Ligands coordinated to Fe in resting form and complexes			
	Solvent ^b	Semiquinone/ Dioxygen ^c	Alkylperoxy intermediate ^d	Product/ Solvent ^e
Fe – Wat1	2.1 Å			2.2 Å
Fe – Wat2	2.2 Å			
Fe – Wat3	2.2 Å			
Fe – O1 ^{4NC}		2.2 Å	2.2 Å	2.2 Å
Fe – O2 ^{4NC}		2.2 Å	2.3 Å	2.0 Å
Fe – O1 ^{O2}		2.5 Å	2.1 Å	
Fe – O2 ^{O2}		2.4 Å	2.8 Å	
C2 – O2 ^{O2}		2.4 Å	1.4 Å	
∠ Fe – O1 ^{O2} – O2 ^{O2}		71 °	101 °	
∠ O1 ^{O2} – O2 ^{O2} – C2		151 °	118 °	

^a For simplicity, nomenclature based on the molecular origin of the atom is used, since the chemical nomenclature differs for substrate, intermediate and open ring ligands. Structures shown represent solvent bound to Fe in the free enzyme (I), dioxygen (II), 4NC (III), 4NC-semiquinone (IV), alkylperoxy intermediate (V) and open-ring product (VI). ^b Average distances calculated for 4 monomeric subunits of free enzyme (PDB code 2IG9). Errors for distances were 3% or less. ^c Values corresponding to structure shown in Fig. 3A. ^d Values corresponding to structure shown in Fig. 3B. ^e Values corresponding to structure shown in Fig. 3C.

References:

- S1. Y. Z. Wang, J. D. Lipscomb, *Protein Exp. Purif.* **10**, 1-9 (1997).
- S2. P. V. Patil, D. P. Ballou, *Anal. Biochem.* **286**, 187-192 (2000).
- S3. G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr.* **D53**, 240-255 (1997).
- S4. P. Emsley, K. Cowtan, *Acta Crystallogr.* **E60**, 2126-2132 (2004).
- S5. M. W. Vetting, L. P. Wackett, L. Que, Jr., J. D. Lipscomb, D. H. Ohlendorf, *J. Bacteriol.* **186**, 1945-1958 (2004).
- S6. CCP4, *Acta Crystallogr.* **D50**, 760-763 (1994).