Electronic Supplementary Information for:

Oxygen diffusion pathways in a cofactor-independent dioxygenase

Natali V. Di Russo, Heather L. Condurso, Kunhua Li, Steven D. Bruner*, Adrian E. Roitberg*

Department of Chemistry, University of Florida, Gainesville, FL 32611-7200, USA.
*Corresponding Authors: roitberg@ufl.edu (AER), bruner@chem.ufl.edu (SDB)

Structure of DPA-CoA and proposed reaction mechanism

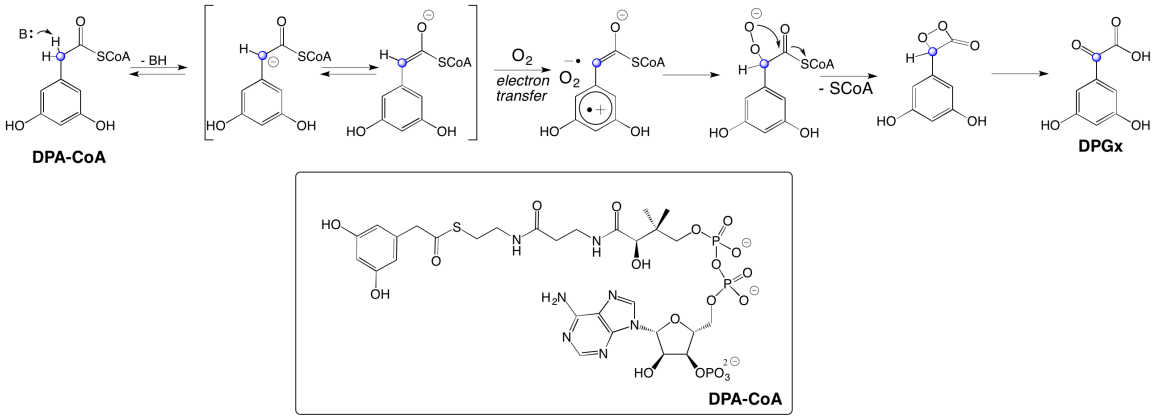

Figure S1. Proposed mechanism for the conversion of DPA-CoA to DPGx by DpgC. The carbon atom that reacts with O_2 is highlighted in blue.

Statistics for O_2 complete entries and escapes in WT DpgC

Complete entries

Complete entries are defined as events where O_2 starts completely outside the protein surface and ends in a position less than 6 Å away from the DPA-CoA carbon atom where it reacts.

Table S1. Complete entries observed in each independent simulation, classified according to the pathway used.

Pathway	Sim1	Sim2	Sim3	Sim4	Sim5	All sim	% of total
L1	4	13	6	5	11	39	**70.91**
L2	0	4	2	2	0	8	**14.55**
U1	0	1	2	1	1	5	**9.09**
S1	0	1	0	0	0	1	**1.82**
Other	0	1	1	0	0	2	**3.64**
Total	**4**	**20**	**11**	**8**	**12**	**55**	**100.00**

Complete escapes

Complete escapes are defined as events where O_2 starts from a position less than 6 Å away from the DPA-CoA carbon atom where it reacts, and ends completely outside the protein surface.

Table S2. Complete escapes observed in each independent simulation, classified according to the pathway used.

Pathway	Sim1	Sim2	Sim3	Sim4	Sim5	All sim	% of total
L1	5	13	7	6	15	46	**63.89**
L2	1	2	2	2	1	8	**11.11**
U1	2	3	1	3	2	11	**15.28**
S1	1	0	0	0	0	1	**1.39**
Other	1	1	3	1	0	6	**8.33**
Total	**10**	**19**	**13**	**12**	**18**	**72**	**100.00**

Occupation of O_2 pockets along complete entries or escapes

For the following analysis, only complete trajectories are considered. Time spent by O_2 in the pockets that does not contribute to a complete trajectory is not taken into account.

Table S3. Average time in ns that O_2 spends in each pocket, for each independent simulation. Values are averages for the 6 O_2 present in each simulation.

Pocket	Sim1	Sim2	Sim3	Sim4	Sim5	Average	Standard dev.
Lower	3.48	8.97	7.39	20.33	9.87	10	6
Upper	3.61	3.60	1.44	2.33	4.15	3	1
SI	11.26	0.56	17.12	0.00	0.00	6	8
SII-SIII	6.91	2.67	0.00	0.00	0.00	2	3
SIV	0.68	0.07	0.00	0.00	0.00	0.2	0.3
Total time (ns)	25.93	15.87	25.94	22.66	14.01	21	6
% simulations*	24.60	15.16	24.87	22.17	14.32	20	5

* Simulations are slightly different in length: 105.43, 104.65, 104.32, 102.19 and 97.84 ns respectively.

Table S4. Total time for each simulation that O_2 spends at a distance less than 6 Å away from the DPA-CoA atom where it reacts ("reactive time"), calculated adding the time each O_2 fulfills this requirement with respect to the DPA-CoA in any of the protein subunits. On average, each O_2 spends 3% of a trajectory in this reactive position.

	Sim1	Sim2	Sim3	Sim4	Sim5	Total
Reactive time (ns)	9.74	9.57	10.83	56.14	11.19	97.47

The time O_2 spends in this "reactive" position can be used to estimate the concentration of O_2 in this region. Since O_2 is only found on one face of DPA-CoA, the volume of this region can be roughly approximated as a half sphere of 6 Å radius (452.39 $Å^3$). Considering that DpgC is a hexamer, this adds up to a total volume of 2714 $Å^3$. This volume is occupied 1/10 to 1/2 of the time depending on the simulation, resulting in local concentrations ranging from 0.06 to 0.31 M. A uniform distribution of O_2 (as expected from passive diffusion) would result in a 5mM concentration, making the probability of finding O_2 in a reactive position 10 to 60-fold higher.

In a similar way, the relative exclusion of water from the active site can be assessed by counting the number of waters found in a 5 Å radius around the same DPA-CoA carbon atom at every saved trajectory time step. This information can be used to calculate the probability of finding different numbers of water molecules in this volume (Figure S2). The expectation value derived from this probability distribution, 1.58 water molecules, is in good agreement with the crystal structure where 2 water molecules are observed in two of the sites and none in the third one. Assuming that the complete sphere volume is available (523.6 $Å^3$), the local water concentration is 5.01 M. This is ~10 times lower that the water concentration in the complete simulated volume (assuming an uniform distribution, ~40 M).

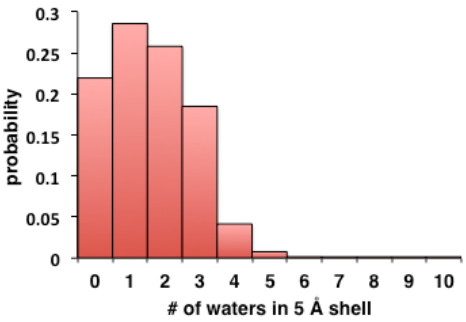

Figure S2. Average probability of finding a given number of water molecules in a 5 Å radius from the DPA-CoA carbon atom that becomes deprotonated.

Dynamic nature of the O_2 pockets

MDpocket[1] is a fast open source geometry-based cavity detection method for the analysis of MD-generated ensembles. Using default settings, we applied the algorithm to a set of 1055 equally spaced snapshots from one of the five independent simulations after stripping water and O_2 molecules, and counterions. The normalized frequency map obtained represents the fraction of time that any given grid point is accessible during the trajectory. Figure S3 shows the density maps obtained for the O_2 pockets, at 50% and 75% frequency isovalues. The overall shape of the density is similar for all the subunits composing the hexamer. The maps reveal that the lower pocket is available 50% of the time, but only in subunit 3 there is density in this region with 75% frequency. In contrast, the upper pocket cannot be observed at these frequency levels and is only available 40 to 10% of the time, depending on the subunit. The O_2 pockets corresponding to the S1 pathway display more variability. SIII is available at least 50% of the time in all cases, and pockets SI and SII are also open in some of the subunits.

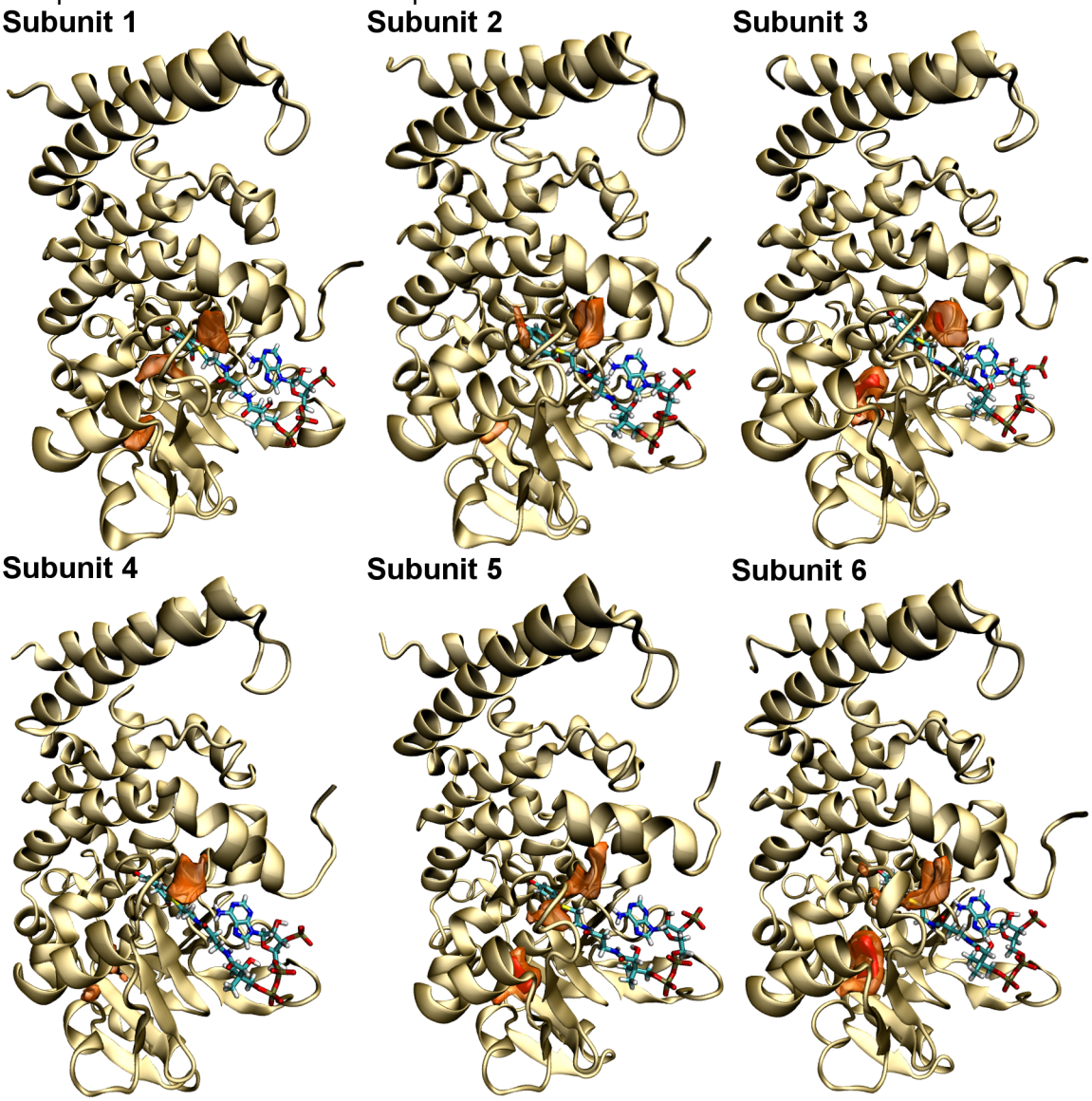

Figure S3. Frequency map for the O_2 pockets. Orange density corresponds to 50% frequency and red density to 75% frequency. Other cavities observed at this isovalues are ommited for clarity.

Figure S3 only shows the density corresponding to the O_2 pockets occupied along complete entries and escapes, but other cavities were also detected, especially between protein subunits (Figure 4B). These are regularly occupied by O_2 along the trajectories and play an important role in allowing O_2 to reach the access point to the S1 pathway, highlighting the importance of simulating the system as a hexamer instead of using only one subunit.

Assessment of convergence

In order to assess the convergence of the results for the WT simulations, we performed an additional set of simulations with roughly six times the concentration of O_2 (a total of 60 O_2 molecules). In this case no O_2 molecules started in the crystallographically observed position: all O_2 molecules were randomly positioned in the solvent. Five independent simulations of this type were carried out, each of them approximately 60 ns long. Table S5 and Table S6 summarize the complete entries and escapes observed in these simulations, which are in reasonable agreement with those observed at lower O_2 concentrations. In this case no complete entries or escapes were observed along the S1 pathway, but O_2 was seen along this path without completing an entry or escape during the simulation time. An important difference between this set of simulations and those presented above is that the larger concentration of O_2 significantly increases the likelihood of having two O_2 molecules in the active site at the same time. In the simulations with only 6 O_2 molecules this scenario was only observed once, while in those with 60 O_2 molecules this happened in 4 of the 5 independent simulations. Holding two O_2 molecules in the lower and upper pockets partially distorts them, increasing the frequency of the use of pathways other than L1 and L2.

Table S5. Complete entries observed in each independent simulation with 60 O_2 molecules, classified according to the pathway used.

Pathway	Sim1	Sim2	Sim3	Sim4	Sim5	All sim	% of total
L1	9	7	8	9	17	50	**59.52**
L2	2	8	0	4	1	15	**17.86**
U1	2	2	2	0	3	9	**10.71**
S1	0	0	0	0	0	0	**0.00**
Other	1	1	0	5	3	10	**11.90**
Total	**14**	**18**	**10**	**18**	**24**	**84**	**100.00**

Table S6. Complete escapes observed in each independent simulation with 60 O_2 molecules, classified according to the pathway used.

Pathway	Sim1	Sim2	Sim3	Sim4	Sim5	All sim	% of total
L1	5	8	10	10	16	49	**62.82**
L2	2	7	0	4	3	16	**20.51**
U1	0	0	0	1	2	3	**3.85**
S1	0	0	0	0	0	0	**0.00**
Other	6	1	0	3	0	10	**12.82**
Total	**13**	**16**	**10**	**18**	**21**	**78**	**100.00**

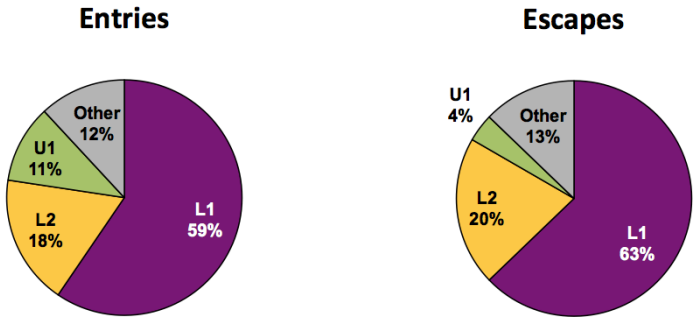

Figure S4. Distribution of O_2 pathway uses in simulations with 60 O_2 molecules.

Comparison with literature data

In order to compare the number of events observed to those reported in the literature, we calculated the normalized events according to the following equation, where the terms are defined in Table S7:

$$normalized\ events = \frac{E}{AS \times [O_2] \times S \times L}$$

Table S7. Comparison of the number of entries and escapes observed in DpgC with those reported in the literature for other systems.

System	DpgC		DpgC (higher $[O_2]$)		Flavin-dependent hydroxylase[2]	Flavin-dependent oxidase[2]	Myoglobin[3]	
	Entries	Escapes	Entries	Escapes	Entries	Entries	Entries	Escapes
E: # of events	55	72	84	78	4	5	16	11
S: # of independent simulations	5	5	5	5	10	10	6	6
L: Length of each simulation (ns)	100	100	60	60	30	50	90	90
AS: # of active sites per simulation	6	6	6	6	2	1	1	1
$[O_2]$ (M)$^{\neq}$	0.005	0.005	0.03	0.03	0.14	0.3	~0.1*	~0.1*
Normalized events ($M^{-1}.ns^{-1}$)	**3.7**	**4.8**	**1.56**	**1.44**	**0.283**	**0.033**	**0.3**	**0.2**

*Estimated concentration based on the size of the solvent box.
$^{\neq}$[CO] in the case of myoglobin.

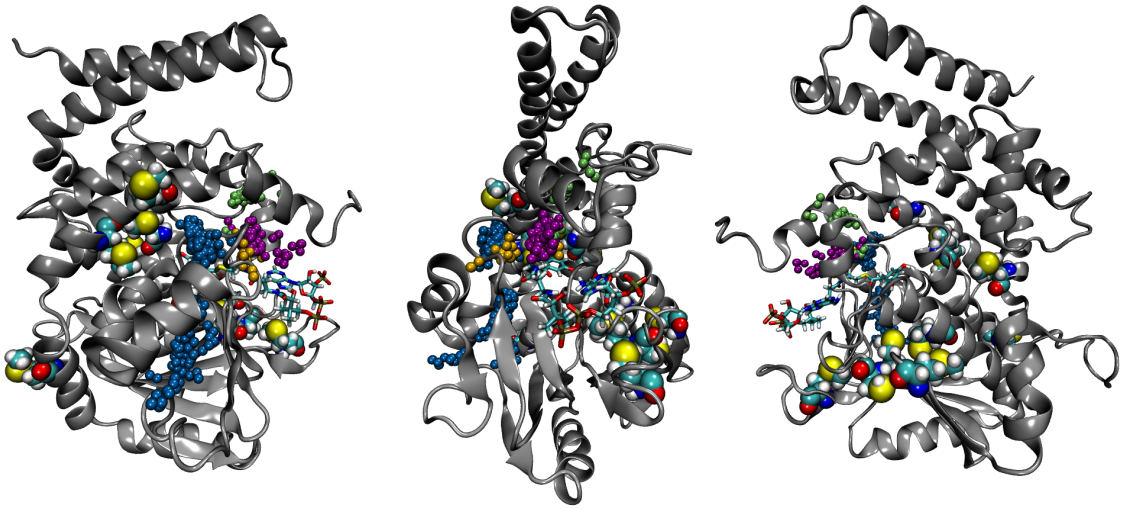

Figure S5. Location of sulfur-containing residues (space-filling representation) in a DpgC monomer, shown in three different views. The L1 (purple), L2 (yellow), S1 (blue), and U1 (green) pathways are also shown.

Supplementary data for molecular dynamics simulations of mutants

L1 mutants

Using MDpocket, we analyzed the 9 ns long simulations of F432W and F432Y and compared them to the first 9 ns of one of the WT trajectories. In order to assess the effect of the mutations on the percentage of time that the L1 gate is open, we employed the MDpocket parameters recommended to locate small channels.[1] The resulting frequency grid files were visually inspected to determine the percentage of time the gate is open in each protein subunit as the maximum isovalue that still shows a connection between the lower pocket and the protein surface between residues K428 and 432. The results are summarized in Table S8.

Table S8. Percentage of time the L1 gate is open in each protein subunit for WT DpgC and the F432W and F432Y mutants.

Mutant	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Average
WT	45	75	75	67	50	67	**63**
F432W	20	67	60	57	62	70	**56**
F432Y	43	50	47	53	37	50	**47**

U1 and L2 mutants

We performed 9 ns long simulations of V429L, V429A and I426A (U1 mutants). The V429L mutation was ruled out because it is likely it will also affect L1 due to a change in the position of F432 the opens up the L1 access point (Figure S6, A). The V429A mutation also affects the mobility of F432 and in some of the protein subunits it causes the collapse of the U1 "hole" (Figure S6, B). The I426A mutation also causes the U1 "hole" collapse in some of the subunits and, in addition, the smaller residue allows O_2 access through additional access points that were never observed in the WT simulations (Figure S6, C). These mutants were ruled out from further consideration because of the possibility of affecting both L1 and U1 at the same time, and because mutations to smaller residues could both ease or hinder O_2 diffusion, depending on whether the U1 "hole" effectively collapsed, and on whether other U-type pathways became possible.

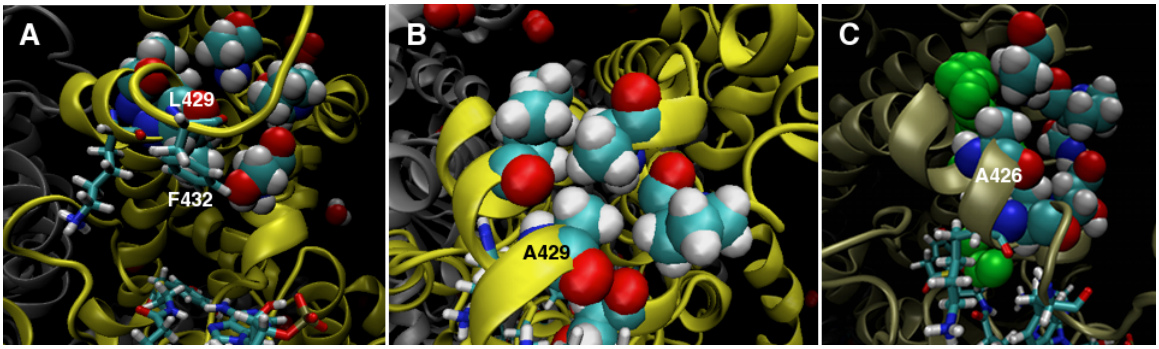

Figure S6. A) Change in the position of F432 caused by the V429L mutation. B) Collapse of the U1 "hole", shown in this case for V429A. C) The I426A mutation allows O_2 access through a pathway not observed in the WT protein. O_2 positions along the trajectory shown in green.

We also ran molecular dynamics simulations of A319L (L2 mutant). This mutation perturbs the interactions of K428 with neighboring acid residues, affecting the L1 pathway as well as the L2 pathway, and was ruled out from further consideration.

Implicit Ligand Sampling

The free energy profile for O_2 migration in DpgC was computed using implicit ligand sampling (ILS). Table S9Table S12 detail the ILS results obtained for each subunit, and Figure S7 summarizes the results graphically.

Table S9. ILS free energy profile along the L1 pathway. All values in kcal/mol.

L1 pathway	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Average
Lower pocket	-4.4	-4.8	-4.3	-4.1	-4.2	-4.2	**-4.3**
Barrier 1	-2.6	-2.3	-2.6	-2.2	-1.7	-2.5	**-2.3**
L1 sink	-3.2	-3.4	-3.7	-3.1	-3.6	-3.7	**-3.5**

Table S10. ILS free energy profile along the L2 pathway. All values in kcal/mol.

L2 pathway	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Average
Lower pocket	-4.4	-4.8	-4.3	-4.1	-4.2	-4.2	**-4.3**
Barrier 1	-2.6	-1.6	-2.2	-3.2	-1.6	-1.2	**-2.1**
L2 sink	-2.8	-3.8	-3.5	-4.3	-3.7	-3.5	**-3.6**

Table S11. ILS free energy profile along the U1 pathway. All values in kcal/mol.

U1 pathway	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Average
Lower pocket	-4.4	-4.8	-4.3	-4.1	-4.2	-4.2	**-4.3**
Barrier 1	-2.8	-2.8	-1.5	-0.9	-2.9	-1.3	**-2.0**
Upper pocket	-3.9	-3.7	-3.1	-1.8	-3.7	-2.7	**-3.2**
Barrier 2	-1.7	-0.8	-2.1	-1.5	-0.6	-1.8	**-1.4**
U1 sink	-3.8	-3.7	-4	-3.5	-2.1	-3.4	**-3.4**

Table S12. ILS free energy profile along the S1 pathway. All values in kcal/mol.

S1 pathway	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Average
Lower pocket	-4.4	-4.8	-4.3	-4.1	-4.2	-4.2	**-4.3**
Barrier 1	-0.3	0.2	-3.3	-0.1	-2	-3.5	**-1.5**
SI	-4.6	-4.3	-4.1	-3.7	-4.4	-4.3	**-4.2**
Barrier 2	4	6.4	10	5.4	3.3	4.9	**5.7**
SII	-5.6	-5	-4.5	-4.1	-4.9	-5	**-4.9**
Barrier 3	-2	-3.3	-3.1	-2.6	-3.6	-3.1	**-3.0**
SIII	-2.4	-4.8	-3.8	-3.8	-5.2	-4.1	**-4.0**
Barrier 4	-0.8	-1.5	-0.9	-1.2	0.5	-0.9	**-0.8**
SIV	-3.2	-3.8	-3.1	-4.1	-1	-2.7	**-3.0**
Barrier 5	-1.8	-0.7	0	-0.5	5	1.7	**0.6**
S1 sink	-2.2	-3	-2.8	-2.1	-3	-2.9	**-2.7**

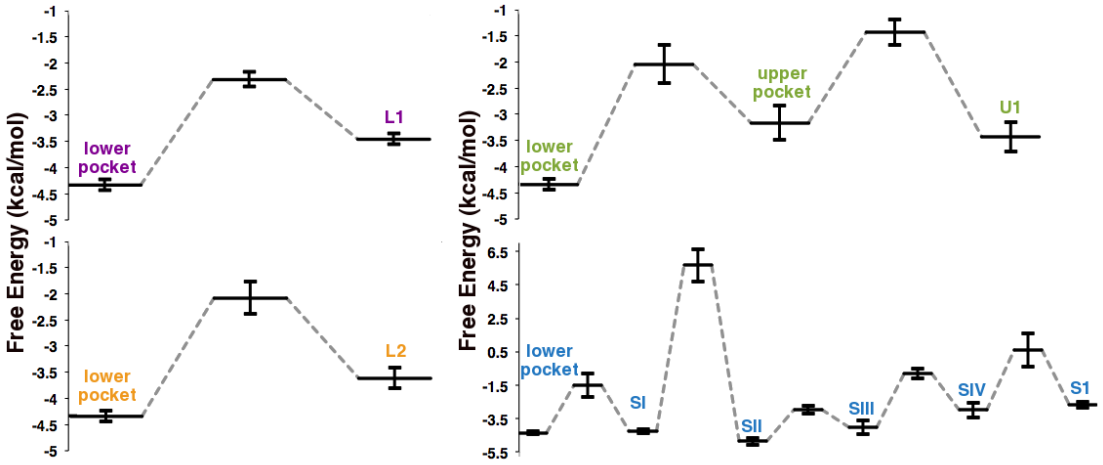

Figure S7. Free energy profile for O_2 migration along pathways L1, L2, U1, and S1. Free energies are relative to O_2 in the solvent. Error bars correspond to the standard error for the six different subunits.

ILS results can also provide insight into why Xe binds in the Xe1 and Xe2 cavities. First, Xe binding to the O_2 site in the active site is unlikely considering that the site is solvent-accessible, while Xe binds preferentially to very hydrophobic sites. Taking a look at the ILS profiles obtained for all subunits, there are only two regions that consistently display very favorable free energies, the Xe-binding region and a region next to an alpha-helix (h-right) near V263, R264, E282, R393, N396, L397, and D400. However, the latter is polar and thus not a good candidate for Xe binding.

.

Table S13. ILS free energies for the Xe-binding site and the h-right site. All values in kcal/mol.

Site	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Sub 6	Average
Xe1	-6	-4.3	-4.9	-4.3	-3.8	-4.4	**-4.6**
Xe2	-4.1	-5.1	-3.9	-4.1	-5.1	-4.2	**-4.4**
h-right	-4.7	-5.1	-4.4	-4.6	-5.1	-5	**-4.8**

Detailed Materials and Methods

Molecular Dynamics simulations

All simulations were performed using the AMBER12 suite.[4] The ff99SB force field parameters were used for all residues[5] except DPA-CoA. Parameters for this molecule were generated combining published parameters for adenosine triphosphate[6] and HF/6-31G* Gaussian09[7] calculations. The five missing C-terminal residues in the crystal structure of DpgC (PDB: 2NP9, 2.45 Å resolution) were added manually and the substrate-analogue was substituted for DPA-CoA. O_2 was kept in the crystallographically observed position. All simulations were performed in an octahedral box of TIP3P water using periodic boundary conditions, with a size consistent with an O_2 concentration in the order of an O_2 saturated solution (5 mM; for comparison, an O_2 saturated buffer is ~1.2 mM). The net negative charge of the system was neutralized adding Na^+ atoms.

The system was minimized in two stages: an initial minimization of the solvent with the protein and O_2 restrained using a 20 kcal.mol^{-1}Å^{-2} weight, and subsequent minimization of the side chains, keeping the backbone carbon atoms and O_2 molecules restrained with the same weight. Then, the system was linearly heated from 50 K to 300 K in 10 ps and maintained at 300 K for 90 ps using 20 kcal.mol^{-1}Å^{-2} restraints in the backbone carbon atoms and O_2 molecules. These restraints were lowered in three 250 ps stages using 10, 5 and 1 kcal.mol^{-1}Å^{-2} weights. The final 250 ps equilibration stage kept 1 kcal.mol^{-1}Å^{-2} restrains only on the position of the O_2 molecules, which were removed during the production stage.

All production simulations were performed at 300 K, maintained using a Langevin thermostat with a 2 ps^{-1} constant for temperature coupling. The SHAKE algorithm was used to keep bonds involving hydrogen atoms at their equilibrium length. Newton's equations were integrated using a 2 fs time step. Frames were collected every 5000 steps. 5 independent production simulations were performed, each approximately 100 ns long.

In the case of simulations with 60 O_2 molecules (30 mM), these were placed at random initial positions within the solvent, and equilibration and production were performed as described above, except that the duration of each of the five independent simulations was approximately 60 ns long. Mutant simulations were also performed starting with 60 O_2 molecules outside the protein.

General Experimental Considerations

Unless otherwise stated, all chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Fischer Scientific (Pittsburgh, PA, USA). Protein concentrations were determined at 280 nm using the Nanodrop 2000 spectrophotometer (Nano-drop Technologies, Wilmington, DE). 3,5-Dihydroxyphenylacetyl-CoA (DPA-CoA) and inhibitor DPA-NH-CoA were synthesized by coupling of 3,5-dihydroxyphenylacetic acid with coenzyme A and amino-desulfocoenzyme A, respectively, and purified by preparative HPLC following published procedures.[8,9] O_2 saturated solutions were prepared by continuously bubbling O_2 into 20 mM Tris·HCl, 50 mM NaCl, pH 7.5 buffer for 30 minutes, and kept in a septum capped flask while in use.

Site-Directed Mutagenesis

Point mutations of DpgC were constructed using DpgC from the *Streptomyces toyocaenisis* A47934 gene cluster cloned into the NcoI and HindIII sites of the pET30A vector as a template. PCR amplification was performed using Q5 hot start high-fidelity DNA polymerase, Q5 reaction buffer, Q5 high GC enhancer (New England Biolabs), dNTP mix and the following primers and their reverse complements (modified sequences underlined):

F432W 5'- G ATC GAC AAG GTC GGC CGA TGG GGA GGG CGG -3'
F432Y 5'- G ATC GAC AAG GTC GGC CGA TAC GGA GGG CGG -3'
F432L 5'- G ATC GAC AAG GTC GGC CGA CTG GGA GGG CGG -3'
L361A 5'- GAA CCC GAG GCG CGC GCG CTG GTC GAC GAG GTC GTG -3
A329L 5'-C ATC ATC CCG GGA GCC CTG AAC CTG CGG CTC -3'

The PCR program consisted of an initial hold of 98°C for 30 s, followed by 30 cycles of 98°C for 1 30 s, 55°C for 30 s, and 72°C for 4.5 min, and a final hold of 72°C for 2 min. The template DNA was digested with 10U DpnI for 1 or 2 hours at 37°C before transformation into XL10-Gold Ultracompetent cells. The mutagenesis products were confirmed by DNA sequencing.

WT and mutant DpgC purification.

The vectors containing the target mutation were then transformed into BL21(DE3) cells and grown in LB media at 37°C until cell density reached O.D. = 0.6. Overexpression was induced by adding IPTG (50 μM), followed by overnight incubation at 18°C. Cells were pelleted by centrifugation and lysed using a microfluidizer (MicroFluidics Corp.). The enzyme was then purified using a Ni-NTA affinity resin (Qiagen), followed by fast protein liquid chromatography using a HiTrap-Q (GE Bioscience) ion exchange column.

Kinetic parameters for DPA-CoA

UV-Vis kinetic assays were performed using the DTNB [5,5'-dithiobis(2-nitrobenzoic acid)] reporter assay as described elsewhere[8] with one modification: the concentration of enzyme used was 0.8 μM. A freshly prepared solution of DTNB (1 mM), Tris·HCl (250 mM, pH=7.5), and increasing concentrations of DPA-CoA were mixed with the enzyme (0.8 μM) at 24 °C (1 mL reaction volume). The reactions were monitored in a UV-Vis spectrophotometer (Agilent) at 412 nm for 5 min. The observed initial rate of absorbance increase was converted to initial reaction velocity (V_0) using ε_{412nm} = 13600 $M^{-1}.cm^{-1}$. K_M^S and k_{cat}^S were determined by performing a non-linear regression to a plot of V_0 vs [DPA-CoA] using gnuplot.

Kinetic parameters for O_2

The kinetic parameters for O_2 were determined as described elsewhere[10] using a fluorescence quenching based probe (Ocean Optics) in an airtight cuvette. All experiments were performed at 25 °C with continuous 1000 rpm stirring, using 83 μM DPA-CoA and varying concentrations of O_2. The reaction buffers were prepared by injecting different volumes of O_2-saturated buffer into the cuvette containing varying volumes of air-equilibrated buffer and DPA-CoA. The assay was initiated by injection of DpgC (1.67 μM) into the reaction vessel. Data was collected every 0.1 s. Initial velocities were determined from the initial change in oxygen concentration observed after accounting for non-enzymatic decrease in oxygen levels. Since the $K_M^{O_2}$ for DpgC is

larger than the concentration of O_2 in a saturated solution,[9] it is not possible to obtain data above $K_M^{O_2}$ and it is more convenient to work under the assumption that $K_M^{O_2} >> [O_2]$ and fit the data to a linear equation with slope $V_{max}^{O_2}/K_M^{O_2}$.

Isothermal titration calorimetry

Isothermal titration calorimetry experiments were performed using the MicroCal ITC200 system (GE Healthcare). In all cases, the temperature was maintained at 30°C and the contents of the reaction cell were stirred at 1000 rpm. The enzyme solution (12.4 μM) was placed in the calorimetric cell and titrated with 100 μM DPA-NH-CoA using sixteen 2.49 μL injections performed every 4.98 s. Using the software supplied by MicroCal, the experimental data was fit to a standard one-site model after correcting for the heat change caused by dilution of the inhibitor in the buffer using ΔH (enthalpy of binding), K (binding constant), and n (number of binding sites) as variables. From these values, ΔS (change in entropy upon binding) and ΔG (binding free energy) can be derived using simple thermodynamic relationships.

Xe-pressurized crystallography

DpgC-Xenon data was collected on the X12B beamline of NSLS at Brookhaven National Laboratory at a wavelength 1.500 Å. The data set was initially indexed in a primitive orthorhombic space group with unit cell parameters a = 139.9 Å, b = 156.6 Å, c = 171.0 Å. Analysis of systematic absences showed h00, 0k0 absences for odd reflections, suggesting that the space group could be $P2_12_12$. However, as we reported in the previous structures, only one of the two trimers was interpretable in primitive orthorhombic and R-factors are characteristically high.[9,11] We reasoned that the presence of pseudo-translational symmetry provided us the inaccurate space group. The data set was reprocessed to space group P1 in same cell parameters for further interpretation. The native monotrimer structure (PDB 2NP9) was employed as a search model using molrep[12,13] and the rotation function was conducted within a resolution range of 40 - 3.0 Å. A model hexamer was generated from the best solution applying pseudo-translation vectors generated from Patterson function analysis. Eight hexamers (twenty-four monomers) were identified in space group P1, and the final space group $P2_1$ (a = 139.1 Å, b = 171.0 Å, c = 156.0 Å, β = 90.02 °) was determined through Zanuda[14] based on P1 molecular replacement map and the model.

Structural refinement of the Xe soaked DpgC structure was performed in refmac[15] with NCS restraints and twin law -h, -k, l. Alternative conformations of F315 were added manually in Coot.[16] Based on the refined model, an anomalous difference electron density map was calculated with phenix.maps.[17] Visible Xe sites with an anomalous map level above 3.0 σ were added into the model and all Xe atoms were refined with occupancy and anisotropic B-factors sequentially. Water molecules were added with Coot and evaluated based on inspection of the maps. Sigma-weighted simulated composite omit maps were used to judge and verify structures throughout refinement.

Table S14. X-ray data collection parameters for Xe soaked DpgC crystals.

Data Collection	DpgC/Xe/DPA-NHCoA complex
Wavelength (Å)	1.500
Resolution range (Å)	39.52 - 2.58 (2.672 - 2.58)
Space group	P 1 21 1
Unit cell (Å)	139.135,170.941, 156.003
	90.00, 90.02, 90.00
Total reflections	758372 (73084)*
Unique reflections	228105 (21905)
Multiplicity	3.4 (3.3)
Completeness (%)	99.70 (97.54)
anomalous	96.5 (89.4)
Mean I/sigma(I)	12.36 (3.56)
Wilson B-factor	38.02
R_{merge}	0.07636 (0.3614)
R_{meas}	0.0907
$CC1/2^{18}$	0.999 (0.973)
Refinement	
R_{work}/ R_{free}	0.1675/0.2047
Number of atoms	40818
protein	39080
ligand	723
water	1015
Protein residues	5062
RMS deviations	
bond lengths (Å)	0.020
bond angles (°)	1.89
Ramachandran favored	
favored (%)	98
outliers (%)	0.24
Clashscore	4.79
Average B-factors	38.70
protein	38.90
ligand/ion	40.00
water	29.00
PDB entry	4YLH

*Highest resolution shell is shown in parenthesis.

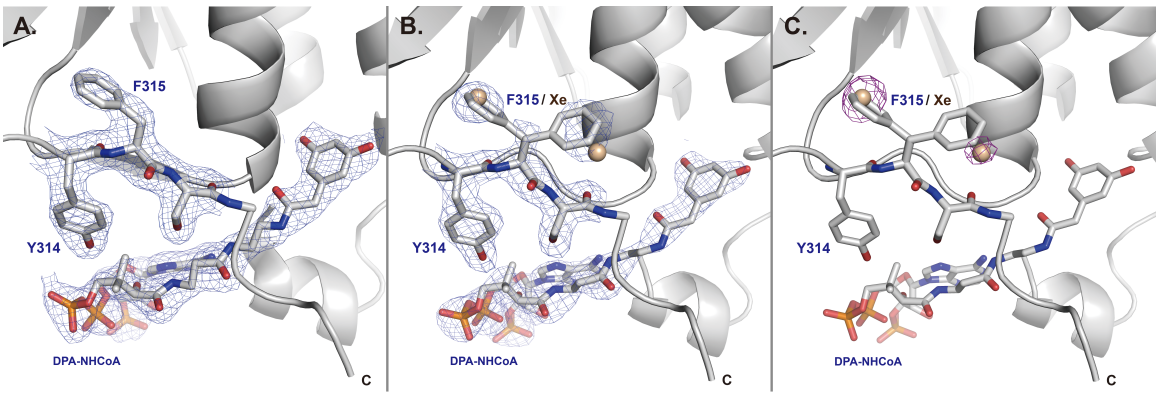

Figure S8. Electron density map of the DpgC/DPA-NHCoA complex near Phe315 and substrate mimic DPA-NHCoA. 2mFo-DFc map at 2.0σ from native (A) and Xenon (B) data sets are illustrated in blue. The anomalous difference map (3.0 σ contour level) from Xenon data sets is shown in magenta (C). The presence of Xe atom forces the alternative conformation Phe315.

The typical F135 conformation (shown in Figure 5 pointing down) overlaps with the Xe1 site. In this conformation, the N-CA-CB-CG dihedral angle ranges from -60 to 100 degrees (Figure S9). Integrating the histogram data in the 130 to180 and -180 to -130 degree range provides an estimation of the percentage of time that F135 spends in the alternative conformation. On average, F135 spends 9% of the simulation time in the alternative conformation, which overlaps with the Xe2 site.

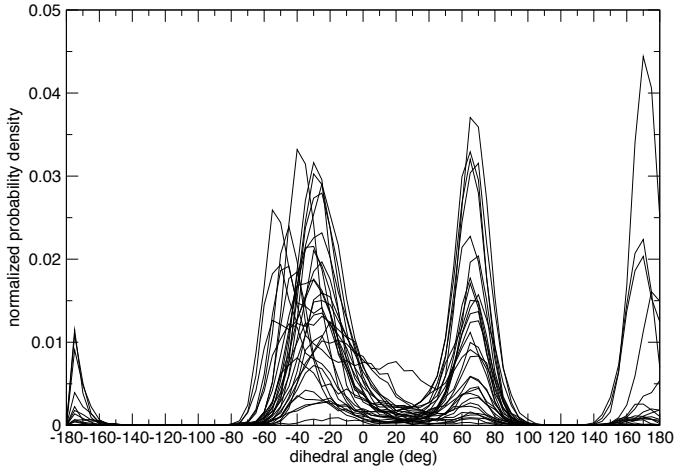

Figure S9. Histogram of the N-CA-CB-CG dihedral angle of F135 for all subunits of all simulations.

Supplementary Experimental Information

Kinetic parameters for DPA-CoA in L1 mutants

Figure S10 depicts the dependence of the reaction rate with the concentration of DPA-CoA for WT DpgC and mutants F432W, F432Y, and F432L.

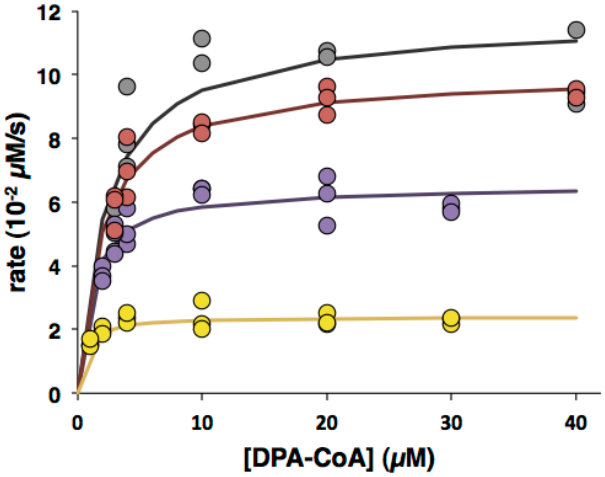

Figure S10. Reaction rate as a function of [DPA-CoA] for WT DpgC (gray), F432W (violet), F432Y (red) and F432L (yellow), fit to the Michaelis-Menten equation with parameters given in Table 1.

These results can be analyzed using the following kinetic model, which has been previously applied to lipoxygenases[19]:

$$E \underset{k_{-1}}{\overset{k_{+1}(S)}{\rightleftharpoons}} E.S \xrightarrow{k_s} E.S^* \underset{k_{-0}}{\overset{k_{+0}(O_2)}{\rightleftharpoons}} E.S^*O_2 \xrightarrow{k_P} \text{Products}$$

This model can be reasonably applied to DpgC because:

- Previous kinetic studies on a different cofactor-independent oxygenase (Hod) are consistent with a compulsory-order ternary-complex mechanism in which the organic substrate binds before O_2.[20]
- UV-Visible and EPR data suggest that in Hod, deprotonation (substrate activation to S* in the model) takes place before O_2 binding.[20]
- Mass spectrometry data from isotopic labeling experiments suggests that in DpgC, the C_2-carbanion (S*) forms reversibly under anaerobic conditions with a rate of 14.2 min^{-1}, suggesting that it is an early intermediate.[21]

From the model, we can derive the following kinetic equations:

(1) $\frac{d[ES]}{dt} = k_{+1}[E][S] - (k_{-1} + k_s)[ES]$

(2) $\frac{d[ES^*]}{dt} = k_s[ES] - k_{+0}[ES^*][O_2] + k_{-0}[ES^*O_2]$

(3) $\frac{d[ES^*O_2]}{dt} = k_{+0}[ES^*][O_2] - k_{-0}[ES^*O_2] - k_P[ES^*O_2]$

(4) $\frac{d[P]}{dt} = k_P[ES^*O_2] = v$

Assuming quasi-stationarity and rearranging (1):

(5) $[ES] = \frac{k_1[E][S]}{k_{-1}+k_s}$

From (2),

(6) $[ES^*] = \frac{k_s[ES]+k_{-0}[ES^*O_2]}{k_{+0}[O_2]}$

And from (3),

(7) $[ES^*O_2] = \frac{k_{+0}[ES^*][O_2]}{k_{+0}+k_P}$

Replacing (7) into (6):

$[ES^*] = \frac{k_s}{k_{+0}}\frac{[ES]}{[O_2]} + \frac{k_{-0}[ES^*]}{k_{-0}+k_P}$, which can be rearranged into

(8) $[ES^*] = \frac{k_s}{k_P}\frac{k_{-0}+k_P}{k_{+0}}\frac{[ES]}{[O_2]}$

Replacing (8) into (7):

(9) $[ES^*O_2] = \frac{k_s}{k_P}[ES]$

If we assume that the total concentration of enzyme is conserved,

(10) $[E_{tot}] = [E] + [ES] + [ES^*] + [ES^*O_2]$

Replacing (5), (8) and (9) into (10) we obtain:

$[E_{tot}] = [E] + \frac{k_{+1}[E][S]}{k_{-1}+k_s} + \frac{k_s}{k_P}\frac{k_{-0}+k_P}{k_{+0}}\frac{[ES]}{[O_2]} + \frac{k_s}{k_P}[ES]$, which can be rearranged to give

(11) $[E] = \frac{[E_{tot}]}{1+\frac{k_1}{k_{-1}+k_s}[S]\left[1+\frac{k_s}{k_P}\left(1+\frac{1}{[O_2]}\frac{k_{-0}+k_P}{k_{+0}}\right)\right]}$

Going back to (4) for the definition of the reaction rate, replacing (9) into (4), we obtain:

$v = k_P[ES^*O_2] = k_s[ES]$, and now using (5), we obtain

(12) $v = k_s\frac{k_{+1}[E][S]}{k_{-1}+k_s}$

Combining (12) with (11), and using the following definitions:

$K_M^* = \frac{k_{-1}+k_s}{k_{+1}}$ and $K^{\#} = \frac{k_{-0}+k_P}{k_{+0}}$

The resulting reaction rate is:

(13) $v = k_s\frac{[S]}{K_M^*}\frac{[E_{tot}]}{1+\frac{[S]}{K_M^*}\left[1+\frac{k_s}{k_P}\left(1+\frac{K^{\#}}{[O_2]}\right)\right]}$

To obtain the equivalent to the Michaelis-Menten K_M^S and V_{max}^S parameters derived from kinetic experiments where [S] is variable and $[O_2]$ is constant, we can compare the reciprocal of (13) to the Lineweaver-Burk linearization of the Michaelis-Menten reaction rate:

$\frac{1}{v} = \frac{K_M^S}{V_{max}^S}\frac{1}{[S]} + \frac{1}{V_{max}^S}$

Taking the reciprocal of (13):

$\frac{1}{v} = \frac{K_M^*}{k_s[S][E_{tot}]} + \frac{1}{k_s[E_{tot}]}\left(1+\frac{k_s}{k_P}\left(1+\frac{K^{\#}}{[O_2]}\right)\right)$

Making the first terms equal we find that

(14) $V_{max}^S = \frac{k_s[E_{tot}]}{1+\frac{k_s}{k_P}\left(1+\frac{K^{\#}}{[O_2]}\right)}$

Doing the same for the second terms, and using (14) we find

(15) $K_M^S = \frac{K_M^*}{1+\frac{k_s}{k_P}\left(1+\frac{K^{\#}}{[O_2]}\right)}$

Based on (14) and (15), if a mutation decreases the rate of O_2 diffusion by decreasing k_{+0} with everything else being equal, $K^{\#}$ increases and as a result K_M^S and V_{max}^S should decrease by the same factor.

Kinetic parameters for O_2 in L1 mutants

Figure S11 depicts the dependence of the reaction rate with the concentration of O_2 for WT DpgC and two mutants expected to hinder diffusion through L1: F432W and F432Y. Since the $K_M^{O_2}$ for DpgC is larger than the concentration of O_2 in a saturated solution, it is not possible to obtain data above $K_M^{O_2}$ and it is more convenient to work under the assumption that $K_M^{O_2} >> [O_2]$ and fit the data to a linear equation with slope $V_{max}^{O_2}/K_M^{O_2}$.

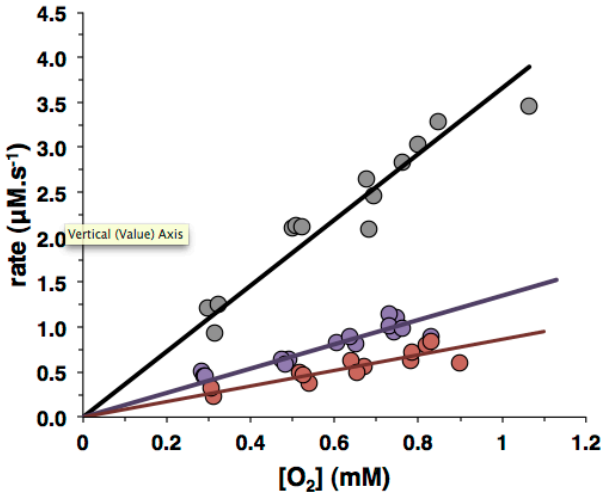

Figure S11. Reaction rate as a function of $[O_2]$ for WT DpgC (gray), F432W (violet) and F432Y (red), fit to a linear equation with slopes 3.65, 1.35, 0.87 and R^2 0.99, 0.99, 0.98, respectively.

Following the same model as in the previous section, we can obtain the equivalent to the Michaelis-Menten $K_M^{O_2}$ and $V_{max}^{O_2}$ parameters derived from kinetic experiments where $[O_2]$ is variable and [S] is constant, we can compare the reciprocal of (13) to the Lineweaver-Burk linearization of the Michaelis-Menten reaction rate:

$$\frac{1}{v} = \frac{K_M^{O_2}}{V_{max}^{O_2}} \frac{1}{[O_2]} + \frac{1}{V_{max}^{O_2}}$$

from Lineweaver-Burk and ,taking the reciprocal of (13),

$$\frac{1}{v} = \frac{K^{\#}}{k_P[E_{tot}][O_2]} + \frac{K_M^*}{k_s[S][E_{tot}]} + \frac{1}{k_s[E_{tot}]} + \frac{1}{k_P[E_{tot}]}$$

Making the first terms equal we find that:

(16) $$V_{max}^{O_2} = \frac{k_s[S][E_{tot}]}{K_M^* + [S] + \frac{k_s}{k_P}[S]}$$

Now doing the same for the second terms, and using (16) we find

(17) $$K_M^{O_2} = \frac{K^{\#} k_s[S]}{K_M^* + [S] + \frac{k_s}{k_P}[S]}$$

Based on these equations, a decrease in k_{+0}, with everything else constant should have no effect on $V_{max}^{O_2}$ and should produce a decrease on the relative $K_M^{O_2} = \frac{K_M(mut)}{K_M(WT)} = \frac{K^{\#}(mut)}{K^{\#}(WT)}$.

Kinetic parameters for DPA-CoA in S1 mutants

Figure S12 depicts the dependence of the reaction rate with the concentration of DPA-CoA for WT DpgC and mutants F432W, F432Y, and F432L.

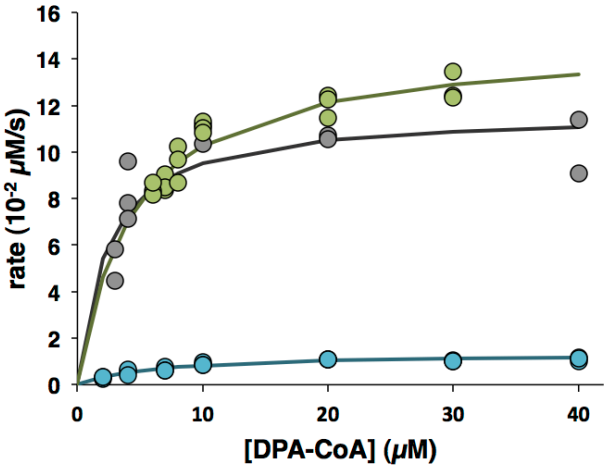

Figure S12. Reaction rate as a function of [DPA-CoA] for WT DpgC (gray), L361A (green), and A329L (turquoise), fit to the Michaelis-Menten equation with parameters given in Table 1.

Also for these mutations, based on equations (14) and (15), if the mutation decreases the rate of O_2 diffusion by decreasing k_{+0} with everything else being equal, $K^{\#}$ increases and as a result K_M^S and V_{max}^S should decrease by the same factor. Instead, if a mutation increases the rate of O_2 diffusion by increasing k_{+0} with everything else being equal, $K^{\#}$ decreases and as a result K_M^S and V_{max}^S should increase by the same factor.

Kinetic parameters for O_2 in S1 mutants

Figure S13 depicts the dependence of the reaction rate with the concentration of O_2 for WT DpgC, a mutant expected to hinder diffusion through S1 (A329L) and one expected to ease it (L316A). Since the $K_M^{O_2}$ for DpgC is larger than the concentration of O_2 in a saturated solution, it is not possible to obtain data above $K_M^{O_2}$ and it is more convenient to work under the assumption that $K_M^{O_2} >> [O_2]$ and fit the data to a linear equation with slope $V_{max}^{O_2}/K_M^{O_2}$.

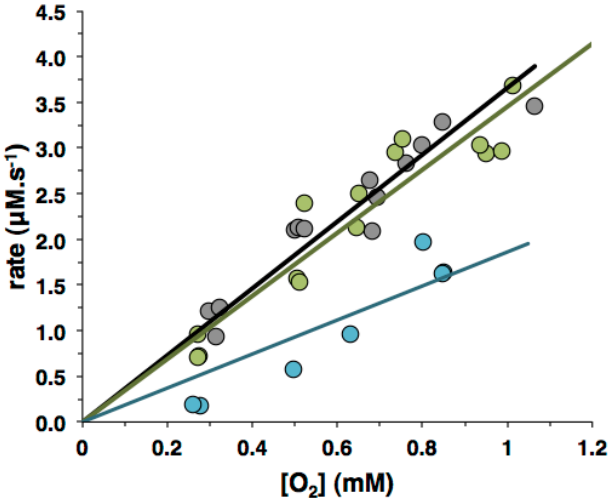

Figure S13. Reaction rate as a function of $[O_2]$ for WT DpgC (gray), L361A (green) and A329L (turquoise), fit to a linear equation with slopes 3.65, 3.45, 1.86 and R^2 0.99, 0.98, 0.94, respectively.

Based on equations (16) and (17), a decrease in k_{+0}, with everything else constant should have no effect on $V_{max}^{O_2}$ and should produce a decrease on the relative $K_M^{O_2} = \frac{K_M(mut)}{K_M(WT)} = \frac{K^{\#}(mut)}{K^{\#}(WT)}$ while an increase in k_{+0}, with everything else constant should have no effect on $V_{max}^{O_2}$ and should produce an increase on the relative $K_M^{O_2}$.

Isothermal Titration Calorimetry on L1 mutants

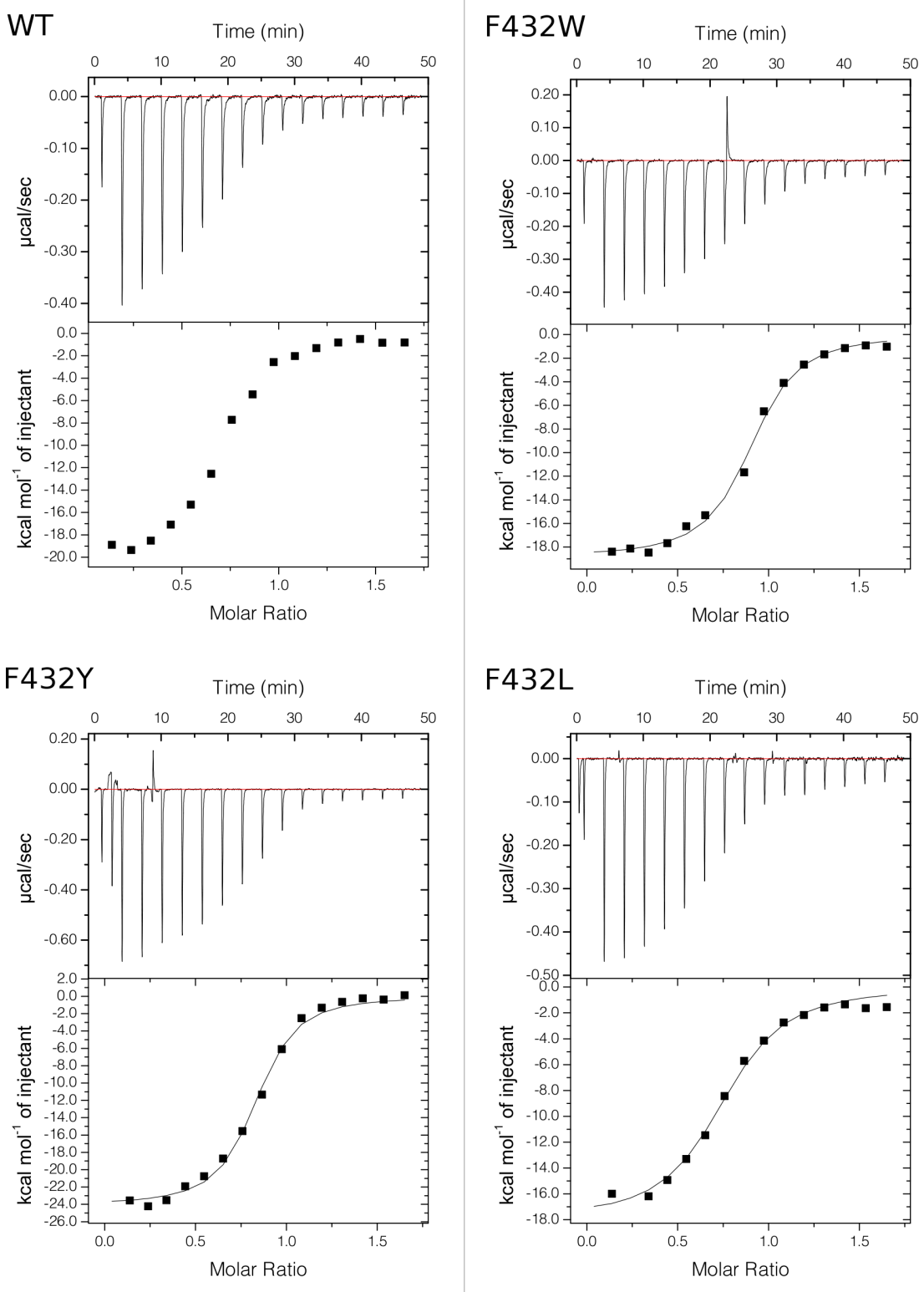

Figure S14. Isothermal titration calorimetry results for WT DpgC and the F432W, F432Y and F432L mutants, fitted using the parameters in **Table S15**.

Table S15. Thermodynamic parameters derived from fitting the isothermal titration calorimetry data of WT DpgC and the F432W, F432Y and F432L mutants.

mutant	n	K (10^6 M^{-1})	∆G (kcal/mol)	∆H (kcal/mol)	∆S (cal/K.mol)
WT	0.685	3.43	-9.05	-20.23	-36.8
F432W	0.877	4.23	-9.19	-18.83	-31.8
F432Y	0.800	5.35	-9.34	-24.11	-48.7
F432L	0.744	1.93	-8.71	-17.97	-30.5

References

1 P. Schmidtke, A. Bidon-Chanal, F. J. Luque and X. Barril, *Bioinformatics*, 2011, btr550.

2 R. Baron, C. Riley, P. Chenprakhon, K. Thotsaporn, R. T. Winter, A. Alfieri, F. Forneris, W. J. H. van Berkel, P. Chaiyen, M. W. Fraaije, A. Mattevi and J. A. McCammon, *Proc. Natl. Acad. Sci.*, 2009, **106**, 10603–10608.

3 J. Z. Ruscio, D. Kumar, M. Shukla, M. G. Prisant, T. M. Murali and A. V. Onufriev, *Proc. Natl. Acad. Sci.*, 2008, **105**, 9204–9209.

4 D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G., Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. and Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman, *AMBER 12*, University of California, San Francisco, 2012.

5 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins Struct. Funct. Bioinforma.*, 2006, **65**, 712–725.

6 K. L. Meagher, L. T. Redman and H. A. Carlson, *J. Comput. Chem.*, 2003, **24**, 1016–1025.

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. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zkrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision A.1*, 2009.

8 H. Chen, C. C. Tseng, B. K. Hubbard and C. T. Walsh, *Proc. Natl. Acad. Sci.*, 2001, **98**, 14901–14906.

9 P. F. Widboom, E. N. Fielding, Y. Liu and S. D. Bruner, *Nature*, 2007, **447**, 342–345.

10 N. V. Di Russo, S. D. Bruner and A. E. Roitberg, *Anal. Biochem.*, 2015, **475**, 53–55.

11 E. N. Fielding, P. F. Widboom and S. D. Bruner, *Biochemistry*, 2007, **46**, 13994–14000.

12 A. Vagin and A. Teplyakov, *Acta Crystallogr. D Biol. Crystallogr.*, 2010, **66**, 22–25.

13 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. a. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.*, 2011, **67**, 235–242.

14 A. a. Lebedev and M. N. Isupov, *Acta Crystallogr. D Biol. Crystallogr.*, 2014, **70**, 2430–2443.

15 G. N. Murshudov, P. Skubák, A. a. Lebedev, N. S. Pannu, R. a. Steiner, R. a. Nicholls, M. D. Winn, F. Long and A. a. Vagin, *Acta Crystallogr. D Biol. Crystallogr.*, 2011, **67**, 355–367.

16 P. Emsley and K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.*, 2004, **60**, 2126–2132.

17 P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr. D Biol. Crystallogr.*, 2010, **66**, 213–221.
18 Q. Liu, T. Dahmane, Z. Zhang, Z. Assur, J. Brasch, L. Shapiro, F. Mancia and W. Hendrickson, *Science*, 2012, **336**, 1030–1033.
19 J. Saam, I. Ivanov, M. Walther, H.-G. Holzhütter and H. Kuhn, *Proc. Natl. Acad. Sci.*, 2007, **104**, 13319–13324.
20 U. Frerichs-Deeken, K. Ranguelova, R. Kappl, J. Hüttermann and S. Fetzner, *Biochemistry*, 2004, **43**, 14485–14499.
21 C. C. Tseng, F. H. Vaillancourt, S. D. Bruner and C. T. Walsh, *Chem. Biol.*, 2004, **11**, 1195 – 1203.