Supporting Information

Baron et al. 10.1073/pnas.0903809106

SI Text

Computational Details. ***Molecular model details.*** The initial configurations for the C_2 homodimer (chain A and D; 798 residues; $FMNH^-$ cofactor; PDB ID: 2jbs) (1) and the AldO monomer (418 residues; $FADH^-$ cofactor; PDB ID: 2vfr) (3) were solvated in (preequilibrated) rectangular boxes of SPC water (4), large enough to avoid any interactions between mirror images under periodic boundary conditions. Water molecules in close contact with the solute were deleted if the minimum distance between non-hydrogen atoms fell below 2.3 Å. All systems were neutralized with ions substituted to randomly chosen water molecules (15 Å minimum pair distances). 100 O_2 molecules were also substituted to randomly chosen water molecules enforcing minimum pair distances of 12 Å to reproduce a condition of oxygen saturation. The simulated systems are summarized in Table S1.

Enhanced-statistics MD simulations. A procedure for enhanced-statistics MD simulation was developed on test systems, and subsequently applied for this work. It enforces protein dynamics as close as possible to the physiological conditions, accurately describes the dynamical and structural role of protein motion, explicitly includes water dynamics effects, and O_2 spontaneous diffusion in absence of biasing forces. This simulation protocol consists of: (*i*) reduced masses (1.6 u) for the oxygen atoms in the O_2 molecules (5); (*ii*) enhanced statistics from 100 (independent and noninteracting) O_2 molecules in the system, kept at a minimum pair distance by a network of repulsive half-harmonic distance-restraining potentials during the heat-up and equilibration periods; (*iii*) a fast grid-based pairlist-construction algorithm (6); (*iv*) enhanced statistics from 5 independent multiple-copy MD runs (7–9) at 300 K and 5 corresponding independent MD runs at 350 K. For a summary of the simulations see Table S1.

Computational details. 20 MD trajectories for the C_2 and AldO systems in explicit water (5 replicas/system at 300 K, and 5 replicas/system at 350 K) were generated using the GROMOS05 biomolecular software (10). The force field parameters and charges were chosen on the basis of the GROMOS 53A6 parameter set (11) (including parameters for the reduced $FMNH^-$ and $FADH^-$ cofactors) to reproduce an apparent pH 7 (C_2) or 7.5 (AldO) (1, 3). The GROMOS-compatible SPC water model (4) and previously reported parameters for ions (12) were used. Bonded and nonbonded interactions for the reduced-mass O_2 molecule were based on standard GROMOS parameters (O–O ideal bond length of 1.23 Å, force constant of 1.66 10^7 kJ mol^{-1}·nm^{-2}). Equilibrium sampling periods of 3 (C_2) or 5 (AldO) ns were used for each run, totaling an enhanced-MD sampling time of 80 ns. For each system, a steepest-descent energy minimization was performed to relax solvent, ions, and O_2 configuration, whereas the position of protein atoms were restrained by using a harmonic potential (force constant of 2,000 kJ·mol^{-1}·nm^{-2}). Next, a steepest-descent energy minimization was performed without restraints to eliminate any residual strain. The energy minimizations ended whenever the energy change per step became <1 kJ·mol^{-1}. MD simulations were initialized from the energy-minimized configurations with atomic velocities taken from a Maxwell-Boltzmann distribution at 5 K, whereas the position of protein atoms were restrained using a harmonic potential (force constant of 5,000 kJ·mol^{-1}·nm^{-2}) and O_2 molecules were kept at a minimum pair distance of 10 Å by a network of 19,800 repulsive half-harmonic distance-restraining potentials (force constant of 5,000 kJ·mol^{-1}·nm^{-2}). Each system was then gradually brought to the desired temperature of 300 or 350 K in 11 or 19 consecutive 150 ps MD periods. During each period the reference temperature was incremented by 25 K and the force constant decreased by 250 kJ·mol^{-1}·nm^{-2}. At this stage 5 independent MD runs were initialized at each temperature by re-assigning velocities taken from Maxwell-Boltzmann distributions at 5 K. The equilibrium distance for the network of repulsive half-harmonic distance-restraining potentials was reduced to enforce minimum pair distance of 5 Å for all pairs of O_2 molecules. This ensures that each O_2 molecule behaves as in absence of the other O_2 molecules, but allows to fully capture O_2-water interactions. An initial equilibration period for each of these 20 MD runs was performed and extended for 750 ps to reach full equilibration of the energy components, and of the box volume (not shown). In all cases, only the sampling periods after equilibration of each trajectory were used for analysis. Newton's equations of motion were integrated using the leap-frog algorithm (13) with a 2 fs time step. The SHAKE algorithm (14) was applied to constrain all bond lengths with a relative geometric tolerance of 10^{-4}. The simulations were carried out in the N,p,T ensemble (at a pressure of 1 atm) by separately coupling the temperature of solute and solvent degrees of freedom to a heat bath (15) (relaxation time 0.1 ps) and by coupling the pressure (estimated based on an atomic virial) to a pressure bath (15) via isotropic coordinate scaling (relaxation time 0.5 ps; isothermal compressibility 4.574 10^{-4} [kJ·mol^{-1}·nm^{-3}]$^{-1}$). Nonbonded interactions were truncated at a distance of 1.4 nm, re-calculated every timestep in the range 0.0 - 0.8 nm and every five timesteps in the range 0.8–1.4 nm, using a twin-range cutoff scheme. A reaction-field correction (16) was applied to account for the neglected interactions beyond 1.4 nm, using a relative dielectric permittivity of 61 for the SPC water model (17). The fast grid-based pairlist-construction algorithm (6) was used (cell-mask edge of 0.4 nm; atomic-level cutoff) as implemented in the GROMOS05 MD++ module (10).

Analysis details. Trajectory snapshots were extracted every 2 ps along each of the 20 MD trajectories and used for analysis. Structural fitting was performed by (*i*) superimposing their centers of mass (to remove overall translation) and (*ii*) performing an atom-positional least-square fitting procedure (to remove overall rotation), using all C^α atoms (18). The stability of secondary structure elements was monitored according to the definition by Kabsch and Sander (19). Successive analyses were carried out to extract successful diffusion paths. First, those O_2 molecules entering a sphere centered on the C4a flavin carbon of 11 Å radius were selected. Next, these trajectories were further processed to extract only those O_2 paths entering the active site cavity. The cavity was defined as a sphere of 5.5 Å centered on the C^α atom of the key residues Phe-266 (C_2) or Ala-105 (AldO). These choices are motivated by the fact that no O_2 path could be observed on the *si*-side of the cofactor molecules. However, because of residue flexibility and thermal fluctuations, the final complete successful paths were selected only after further visual inspection of all of the screened paths. O_2 molecules entering the protein matrix only, but not displaying successful complete paths were extracted whenever falling beyond a cutoff distance of 3 Å from any protein atom and by analysis of their distance time series from the C4a carbon cofactor. These distance time series were also used to characterize reversible or permanently resident paths. In all cases a molecular radius was used, i.e., a molecule is considered to be inside the sphere whenever its center of mass fell within the cutoff distance. Ensemble-averaged iso-surfaces

to identify favorable O_2 locations were calculated by using the VMD software (20).

Experimental Data. ***Product analysis of the single-turnover reaction of Phe266Gly:FMNH$^-$:HPA with oxygen.*** The amount of DHPA formed during a single-turnover reaction of the mutant Phe266Gly was measured to understand the hydroxylation ratio (also called coupling ratio, i.e., the hydroxylated product formed/amount of reduced flavin used) in the reaction of the Phe266Gly mutant. Solutions of Phe266Gly or C_2 wild-type (80 μM) and FMN (50 μM) in 50 mM sodium phosphate buffer pH 7.0 were reduced with dithionite in an anaerobic environment (21). HPA (2 mM) from a sidearm of the cuvette was mixed with the main solution and then, the cuvette was opened to air. The solution was ultrafiltrated using a Centricon YM-10 concentrator to remove the protein. These samples were analyzed for DHPA by using HPLC as described previously (22). In the wild-type reaction, DHPA formed at 89 ± 3% under the condition described whereas it formed at 62 ± 5% for the reaction of Phe266Gly. The data indicate that Phe266Gly is less efficient in hydroxylating the substrate and more prone to H_2O_2 elimination. It also suggest that the active site of Phe266Gly is likely to be more solvent accessible than that of the wild-type enzyme.

1. Alfieri A, et al. (2007) Structure of the monooxygenase component of a two-component flavoprotein monooxygenase. *Proc Natl Acad Sci USA* 104:1177–1182.
2. Sucharitakul J, Chaiyen P, Entsch B, Ballou DP (2006) Kinetic mechanisms of the oxygenase from a two-component enzyme, p-hydroxyphenylacetate 3-hydroxylase from Acinetobacter baumannii. *J Biol Chem* 281:17044–17053.
3. Forneris F, et al. (2008) Structural analysis of the catalytic mechanism and stereoselectivity in Streptomyces coelicolor alditol oxidase. *Biochemistry* 47:978–985.
4. Berendsen HJC (1981) *Interaction Models for Water in Relation to Protein Hydration* (Pullman, B. E., Dordrecht).
5. Hofacker I, Schulten K (1998) Oxygen and proton pathways in cytochrome c oxidase. *Proteins* 30:100–107.
6. Heinz TN, Hünenberger PH (2004) A fast pairlist-construction algorithm for molecular simulations under periodic boundary conditions. *J Comput Chem* 25:1474–1486.
7. Huber T (1998) SWARM-MD: Searching conformational space by cooperative molecular dynamics. *J Phys Chem A* 102:5937–5943.
8. Shirts MR, Pande VS (2001) Mathematical analysis of coupled parallel simulations. *Phys Rev Lett* 86:4983–4987.
9. Adcock SA, McCammon JA (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev* 106:1589–1615.
10. Christen M, et al. (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J Comput Chem* 26:1719–1751.
11. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676.
12. Åqvist J (1990) Ion Water Interaction Potentials Derived from Free-Energy Perturbation Simulations. *J Phys Chem* 94:8021–8024.
13. Hockney RW (1970) The potential calculation and some applications. *Methods Comput Phys* 9:136–211.
14. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys* 23:327–341.
15. Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR (1984) Molecular-Dynamics with Coupling to an External Bath. *J Chem Phys* 81:3684–3690.
16. Tironi IG, Sperb R, Smith PE, van Gunsteren WF (1995) A Generalized Reaction Field Method for Molecular-Dynamics Simulations. *J Chem Phys* 102:5451–5459.
17. Heinz TN, van Gunsteren WF, Hünenberger PH (2001) Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J Chem Phys* 115:1125–1136.
18. McLachlan AD (1979) Gene Duplications in the Structural Evolution of Chymotrypsin. *J Mol Biol* 128:49–79.
19. Kabsch W, Sander C (1983) Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* 22:2577–2637.
20. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graphics 14:33-* 38:27–38.
21. Sucharitakul J, et al. (2007) Kinetics of a Two-Component p-Hydroxyphenylacetate Hydroxylase Explain How Reduced Flavin Is Transferred from the Reductase to the Oxygenase. *Biochemistry* 46:8611–8623.
22. Chaiyen P, Suadee C, Wilairat P (2001) A novel two-protein component flavoprotein hydroxylase. *Eur J Biochem* 268:5550–5561.

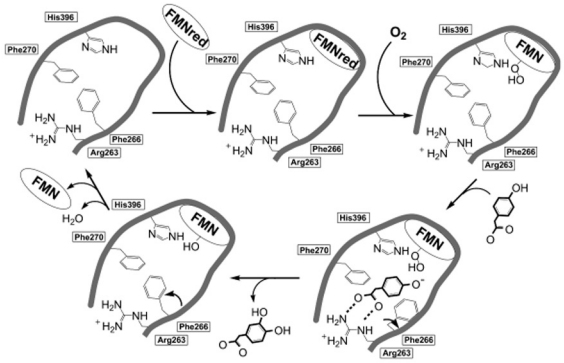


Fig. S1. The catalytic cycle of C_2 monooxygenase. Catalysis starts with the binding of $FMNH^-$ to the apoenzyme. Next, O_2 reacts with the flavin forming a stable C4a-hydroperoxyflavin intermediate. The subsequent binding of the HPA substrate involves a small movement of the Phe-266 side chain. The HPA binding site is located on the *re*-side of the flavin in proximity of the cavity that hosts the hydroperoxide moiety of the C4a-hydroperoxyflavin. This geometry promotes HPA hydroxylation, which results in the formation of a C4a-hydroxyflavin molecule. The reaction is completed by the release of a water molecule that generates the final oxidized flavin product. The order of these catalytic steps is not strictly obligate because HPA can bind also before C4a-hydroperoxyflavin formation. However, prior binding of HPA causes a 100-fold reduction in the O_2 reaction and intermediate formation (Table 1). Adapted from Ref (1).

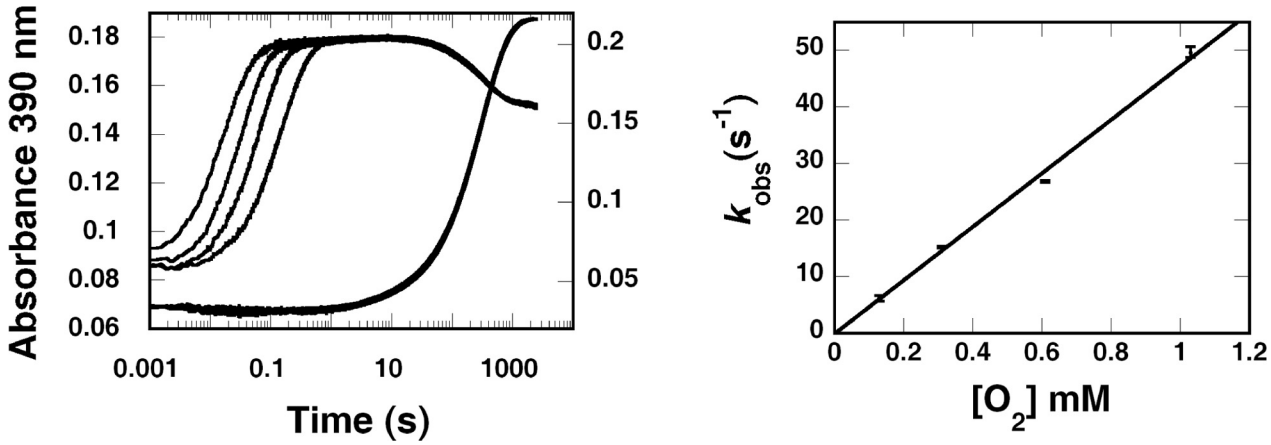


Fig. S2. The O_2 reactivity of the C_2(Phe266Trp) mutant in the absence of HPA. A solution of C_2(Phe266Trp) (80 μM) plus $FMNH^-$ (32 μM) was mixed with buffers containing various O_2 concentrations (0.13, 0.31, 0.61 and 1.03 mM) in the stopped-flow spectrophotometer. The reaction was performed in 50 mM sodium phosphate buffer pH 7.0 at 277 K and monitored at 390 (*Left y* axis) and 446 (*Right y* axis) nm. Kinetic traces show two phases. The first (up to 0.8 s) depends on O_2 concentration. During the first phase C4a-hydroperoxy-FMN formation occurs; traces from right to left are from low to high O_2 concentrations. The second phase (0.8 s–end) is independent of O_2 concentration, shows a large increase of absorbance at 446 nm. This phase involves flavin re-oxidation and can be fitted with a rate constant of 0.0032 s^{-1}. (*Right*) Plot of k_{obs} of the first phase vs. O_2 concentrations. A second-order rate constant of 4.85×10^4 $M^{-1} \cdot s^{-1}$ for the first phase was calculated from the slope of this plot. Error bars were determined from standard deviations at each concentration.

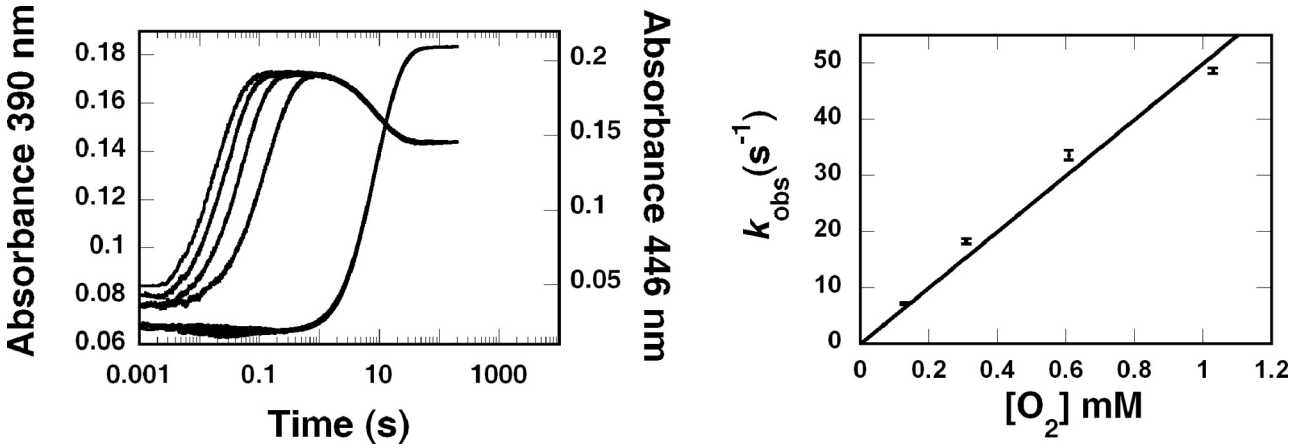


Fig. S3. The O_2 reactivity of the C_2(Phe266Trp) mutant in the presence of HPA. A solution of C_2(Phe266Trp) (80 μM), $FMNH^-$ (32 μM) and HPA (2 mM) was mixed in the stopped-flow spectrophotometer with buffers containing HPA (2 mM) and various O_2 concentrations. The reactions were monitored at 390 and 446 nm. Kinetic traces show three phases. The first phase depends on O_2 concentration (0.13, 0.31, 0.61 and 1.03 mM, traces from right to left) and involves C4a-hydroperoxyflavin formation. The second phase is a lag of absorbance increase at 390 nm. The third phase is the flavin re-oxidation. Both second and third phases are independent of O_2 concentration. (*Right*) Plot of k_{obs} of the first phase versus O_2 concentrations. A second-order rate constant of 4.58×10^4 $M^{-1}{\cdot}s^{-1}$ for the first phase was calculated from the slope of this plot. Error bars were determined from standard deviations at each concentration.

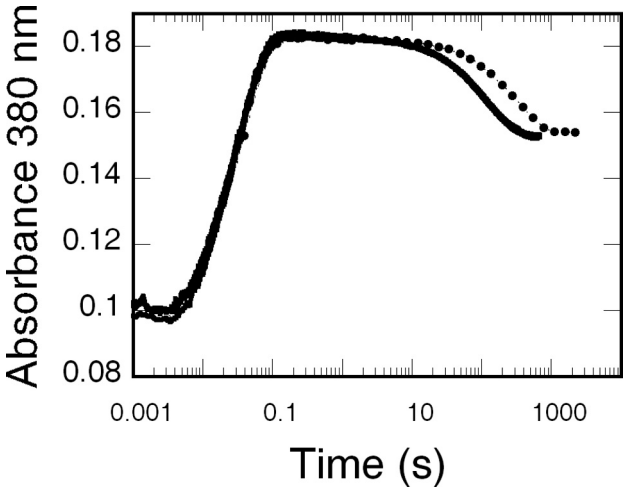


Fig. S4. HPA can only bind to C_2(Phe266Trp) after C4a-hydroperoxy-FMN formation. Unlike the wild-type enzyme (Table 1), the reduced C_2(Phe266Trp) enzyme reacts with O_2 with identical rates regardless of the presence of HPA (Fig. S3). This prevents employing the difference in the rate of C4a-hydroperoxyflavin formation as a means to investigate HPA binding before the oxygen reaction, as previously described in the case of C_2 wild type (2). Therefore, HPA binding to the reduced mutants was investigated by monitoring the flavin oxidation at the final phase using double mixing stopped-flow spectrophotometry. During the first mix, HPA was added, under anaerobic conditions, to C_2(Phe266Trp)-$FMNH^-$. The second mix combined the previous mixture with O_2 at various age times (0.01, 10, and 120 s). The final concentrations after double mixing were 16 μM C_2(Phe266Trp)-$FMNH^-$ (40 μM C_2(Phe266Trp) + 16 μM $FMNH^-$), 100 μM HPA and 0.61 mM O_2. Reactions were monitored with absorbance at 380 nm. *Solid lines* are the superimposed reaction traces at 380 nm for C_2(Phe266Trp)-$FMNH^-$ + HPA at various age times (0.01, 10 and 120 s). For comparison, a dotted line represents the reaction of C_2(Phe266Trp)-$FMNH^-$ with O_2 under identical conditions, but in absence of HPA. It can be seen that the reactions at different age times result in the same rate of flavin oxidation (0.0093 s^{-1}). Therefore, upon adding HPA in the first mix, it is unlikely that the C_2(Phe266trp)-$FMNH^-$-HPA complex was formed because all of the reactions of various age times resulted in the same rate of flavin oxidation. The result suggests that HPA binds to the enzyme only after the C4a-hydroperoxyflavin is formed and before the hydroxylation takes place.

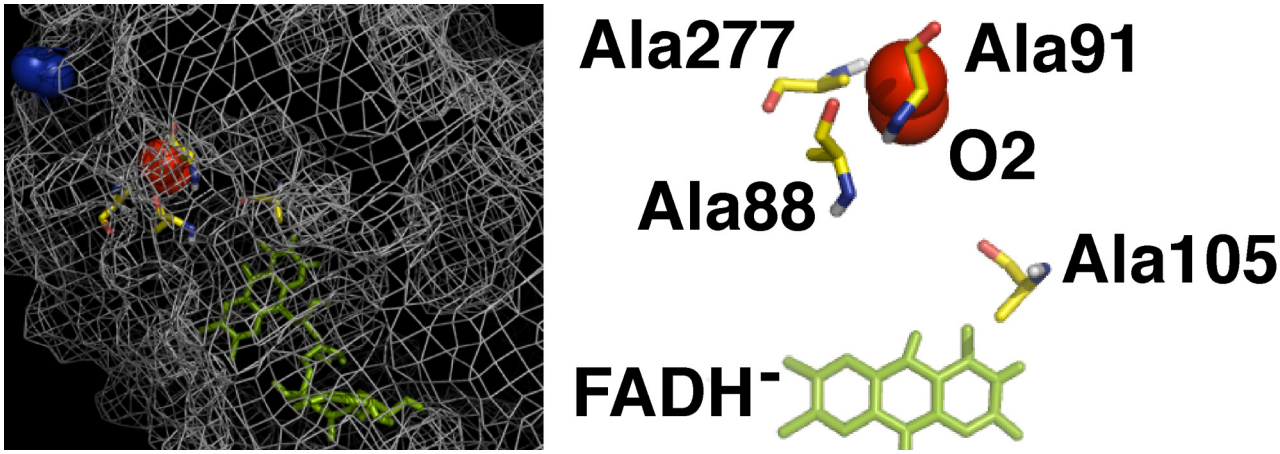


Fig. S5. The all-Ala cavity formed by AldO residues Ala-88, Ala-91, and Ala-277 is an example of O_2 surface pocket. O_2 molecules reside in this pocket before diffusion to the C4a flavin cavity defined by Ala-105. (*Left*) O_2 molecule when entering AldO surface (blue spheres) and trapped into the pocket (red spheres). (*Right*) Positions of residues and FAD cofactor (sticks representation) when O_2 molecule in the pocket (red spheres). The snapshot corresponds to stage A in Fig. 4*B*. All 10 simulations display O_2 molecules transiently occupying this cavity.

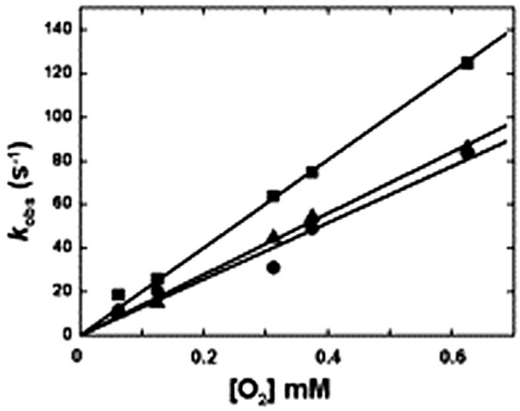


Fig. S6. Dithionite-reduced AldO variants (wt-AldO: filled circles; A105S-AldO: filled triangles; A105G-AldO: filled squares) were mixed with buffers containing various O_2 concentrations (0.0625, 0.125, 0.3125, 0.375 and 0.625 mM) in the stopped-flow spectrophotometer. The observed rates of flavin reoxidation are plotted against the O_2 concentrations. Second-order rate constants (k_{ox}) were calculated from the slope of the plots.

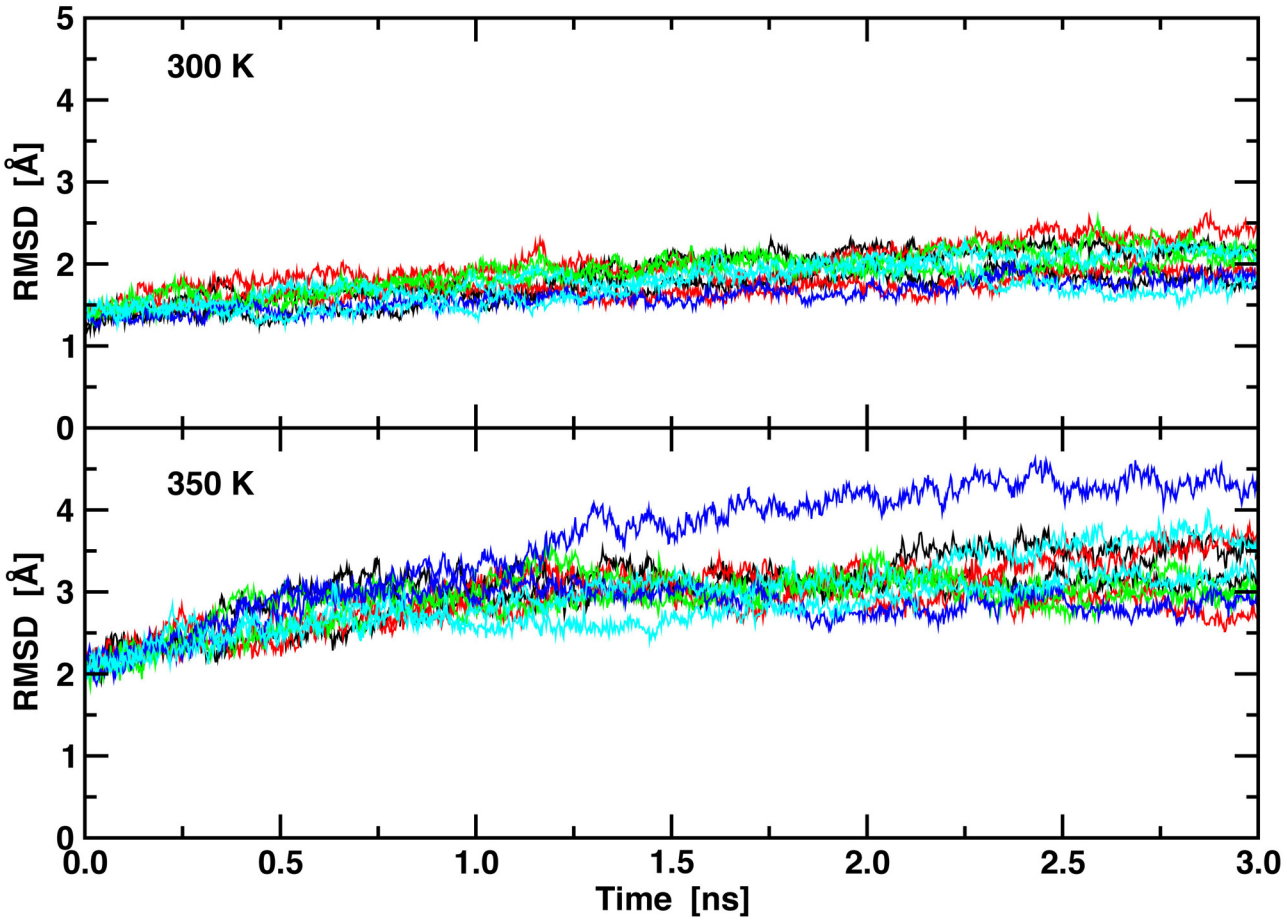


Fig. S7. Root-mean-square deviation, RMSD, of C_2 backbone C^{α}-atoms from the x-ray structure. Different colors represent independent MD runs, same color is used for C_2 chain A and D.

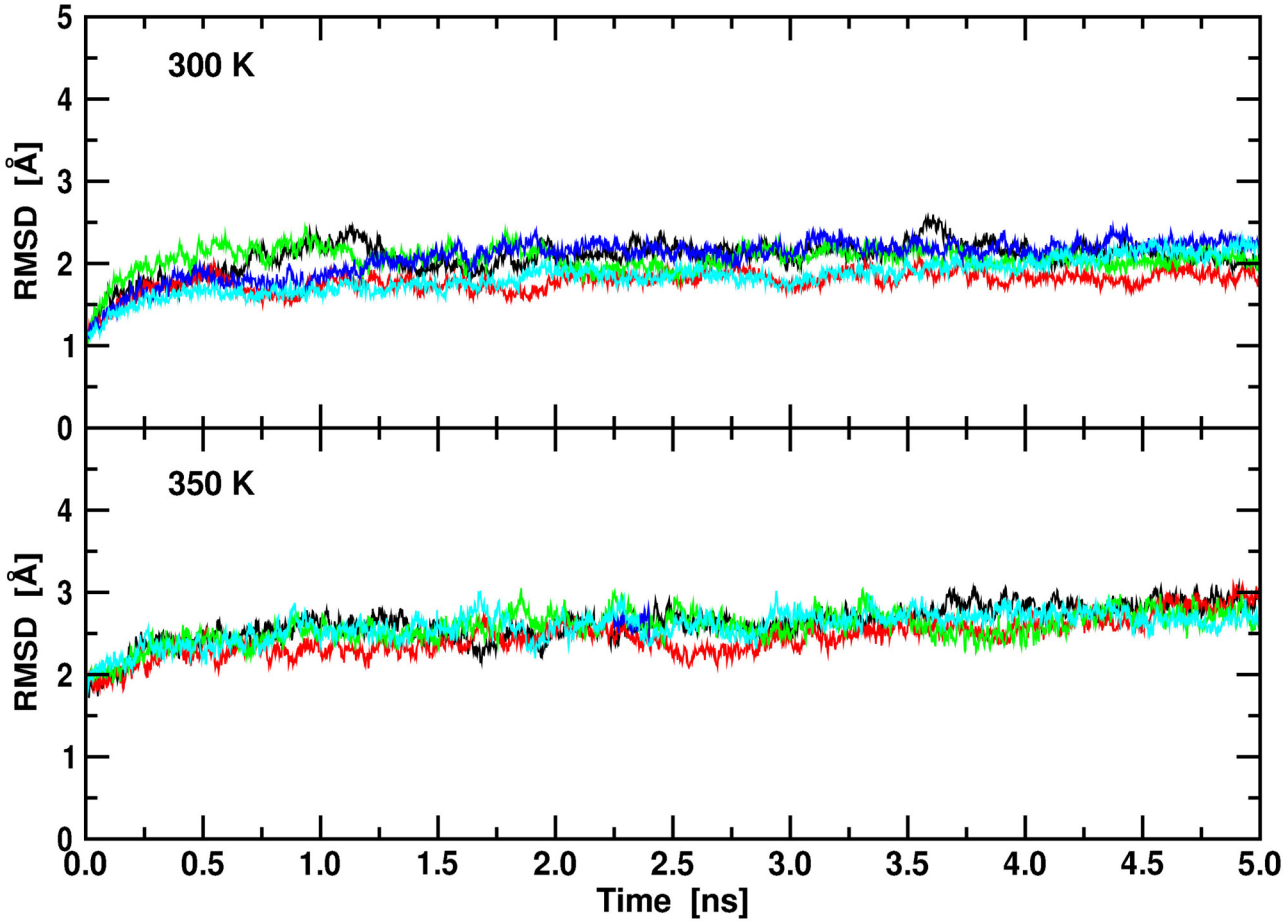
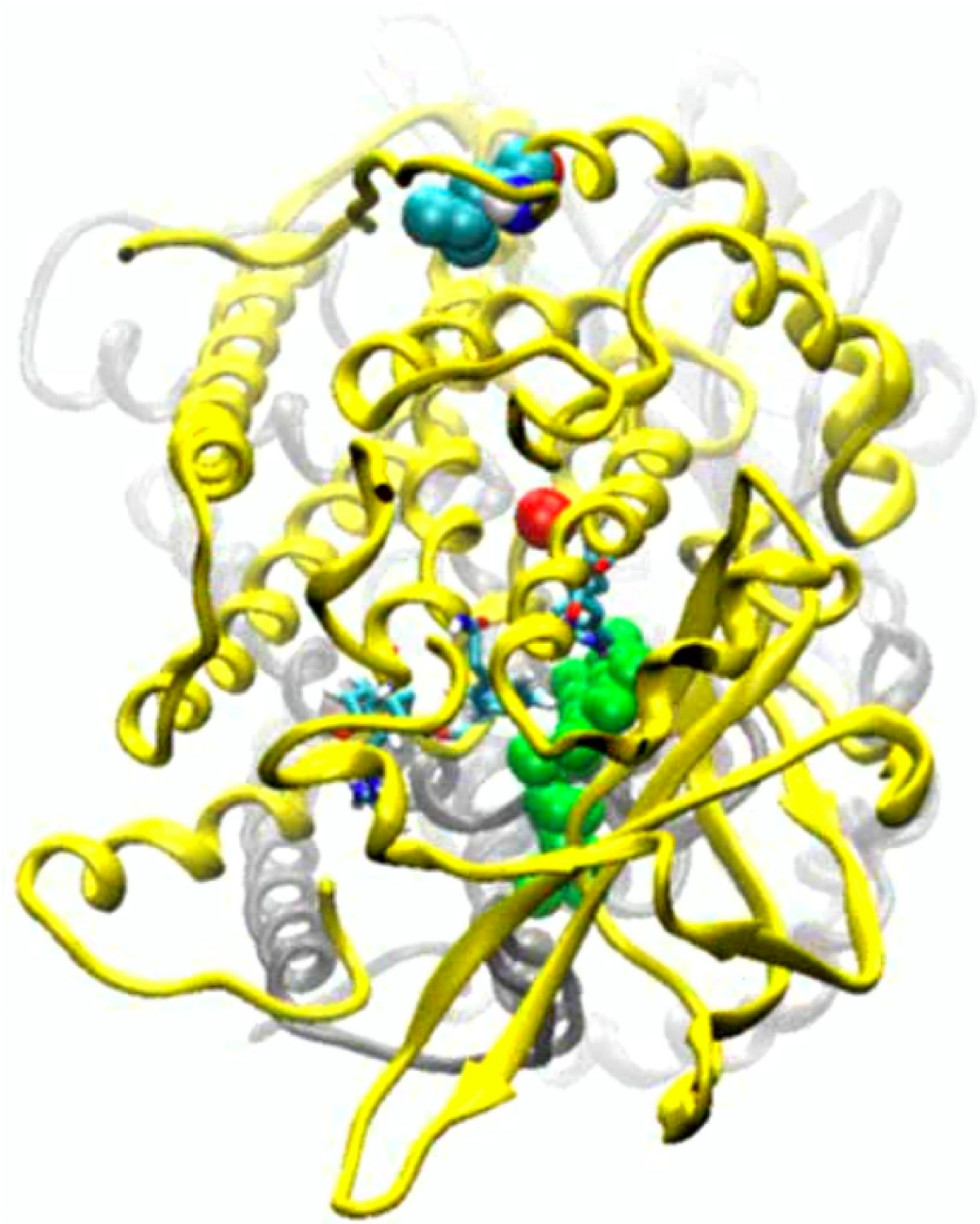


Fig. S8. Root-mean-square deviation, RMSD, of Aldobackbone C^{α}-atoms from the x-ray structure. Different colors represent independent MD runs.



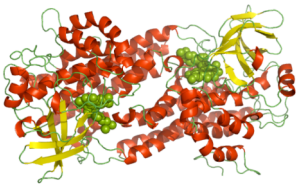
Movie S1. An example diffusion path for C_2 monooxygenase. Enhanced-statistics MD simulations capture spontaneous unbiased diffusion of O_2 molecules toward the reactive side of the flavin cofactor. Corresponding key snapshots are summarized in Fig. 1*b*, black line.

Movie S1 (PG)

Other Supporting Information Files

Table S1

Table S1. Simulated systems

Enzyme	C_2		AldO	
				
PDB ID initial coordinates	2jbs	2jbs	2vfr	2vfr
T [K]	300	350	300	350
nr. O_2 molecules	100	100	100	100
cofactor	$FMNH^-$	$FMNH^-$	$FADH^-$	$FADH^-$
nr. solute atoms	8267	8267	4009	4009
nr. solvent molecules	35028	35028	15342	15342
nr. atoms in the system	113360	113360	50245	50245
nr. of ions	12 Na^+	12 Na^+	1 Cl^-, 3 Mg^{2+}, 6 K^+	1 Cl^-, 3 Mg^{2+}, 6 K^+
total system charge	0	0	0	0
equilibration period [ns]	1.8	2.7	2.7	2.7
nr. MD replicas	5	5	5	5
equilibrium period [ns]	3.0	3.0	5.0	5.0