Published in final edited form as:

J Chem Theory Comput. 2016 June 14; 12(6): 2964–2972. doi:10.1021/acs.jctc.6b00071.

Kinetics of O₂ Entry and Exit in Monomeric Sarcosine Oxidase via Markovian Milestoning Molecular Dynamics

Anthony Bucci[†], Tang-Qing Yu[‡], Eric Vanden-Eijnden[‡], and Cameron F. Abrams^{*,†}

[†]Department of Chemical and Biological Engineering, Drexel University, Philadelphia, Pennsylvania 19104, United States

[‡]Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, United States

Abstract

The flavoenzyme monomeric sarcosine oxidase (MSOX) catalyzes a complex set of reactions currently lacking a consensus mechanism. A key question that arises in weighing competing mechanistic models of MSOX function is to what extent ingress of O₂ from the solvent (and its egress after an unsuccessful oxidation attempt) limits the overall catalytic rate. To address this question, we have applied to the MSOX/O₂ system the relatively new simulation method of Markovian milestoning molecular dynamics simulations, which, as we recently showed [Yu et al. *J. Am. Chem. Soc.* 2015, 137, 3041], accurately predicted the entry and exit kinetics of CO in myoglobin. We show that the mechanism of O₂ entry and exit, in terms of which possible solvent-to-active-site channels contribute to the flow of O₂, is sensitive to the presence of the substrate-mimicking competitive inhibitor 2-furoate in the substrate site. The second-order O₂ entry rate constants were computed to be 8.1×10^6 and $3.1 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ for bound and apo MSOX, respectively, both of which moderately exceed the experimentally determined second-order rate constant of $(2.83 \pm 0.07) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for flavin oxidation by O₂ in MSOX. This suggests that the rate of flavin oxidation by O₂ is likely not strongly limited by diffusion from the solvent to the active site. The first-order exit rate constants were computed to be 10^7 s^{-1} and $7.2 \times 10^6 \text{ s}^{-1}$ for the apo and bound states, respectively. The predicted faster entry and slower exit of O₂ for the bound state indicate a longer residence time within MSOX, increasing the likelihood of collisions with the flavin isoalloxazine ring, a step required for reduction of molecular O₂ and subsequent reoxidation of the flavin. This is also indirectly supported by previous experimental evidence favoring the so-called modified ping-pong mechanism, the distinguishing feature of which is an intermediate complex involving O₂, the flavin, and the oxidized substrate simultaneously in the cavity. These findings demonstrate the utility of the Markovian milestoning approach in contributing new understanding of complicated enzymatic function.

Graphical abstract

*Corresponding Author: cfa22@drexel.edu.

Notes

The authors declare no competing financial interest.

INTRODUCTION

Enzymes that reduce molecular oxygen use protein structure simultaneously to control substrate access to catalytic cofactors and to arrange active sites that greatly accelerate electron transfer to O_2 . For example, monomeric sarcosine oxidase (MSOX), a prototypical flavoprotein D-amino acid oxidase that reduces O_2 to regenerate the oxidized state of its covalently bound flavin cofactor, reduces molecular O_2 at least 1000 times faster than free flavin in aqueous solution. MSOX has been the subject of much valuable structural and kinetic characterization,^{1–8} making it a good test bed upon which to investigate questions relating the structure and function of flavoenzymes. Among such questions are (a) whether the rate of flavin oxidation is limited by the diffusion of O_2 from the (dilute) solution phase through the protein and into the active site and (b) whether O_2 reduction in that site occurs before or after release of the oxidized product.

Both of these questions bear directly on the determination of an overall kinetic model for the function of MSOX and many similar oxidases. The two simplest competing models for MSOX are the so-called ping-pong and modified ping-pong mechanisms,³ shown in Figure 1. The key distinction between the two is the location of the oxidized product P during the step in which the flavin is oxidized by O_2 : in the classical ping-pong mechanism, this occurs after release of P, while in the modified ping-pong mechanism, this occurs before the release of P. Kinetic studies of MSOX suggest that the modified ping-pong mechanism is more likely solely on the basis of the quality of the fit of temporal fluorescence data to model equations.³ There has not yet appeared any structural rationale for this possibility. Here we use all-atom classical molecular dynamics (MD) simulations to study the effect of the presence of a substrate-mimicking inhibitor on important fundamental steps in the two mechanisms involving O_2 entry and exit.

Most previous molecular simulation work in this context has focused on identifying possible routes O_2 takes between the bulk solvent and the active site in similar enzymes without specific consideration of rates. For example, using MD and implicit ligand sampling (ILS), Saam showed that O_2 diffusion through 12/15-lipoxygenase is coupled dynamically to sidechain reorientations in several channels connecting the solvent to the active site.⁹ Baron used enhanced-statistics MD to show spontaneous diffusion of O_2 to the active site in the flavoenzymes *p*-hydroxyphenylacetate hydroxylase and alditol oxidase.¹⁰ More recently, Shadrina used temperature-controlled locally enhanced sampling (TLES) to show O_2 escape routes from the heme to the solvent based on the state of HisE7 in hemoglobin.¹¹

Additionally, in the cofactor-independent oxygenase DpgC, Di Russo used MD enhanced with multiple graphics processing units to determine spontaneous O₂ diffusion routes to the active site.¹² Although these works support the idea that small gas molecules can access buried sites by multiple pathways, it remains challenging to determine which pathways contribute the most to the rate at which O₂ accesses an active site, since none provide a direct way to calculate such rates. This makes these and similar methods somewhat limited when trying to understand fundamental kinetics.

We recently adapted the MD-based method of Markovian milestoning¹³ to compute rates of diffusion, entry, and exit of small molecules in proteins and applied it to the study of CO entry and exit from myoglobin (Mb).¹⁴ Markovian milestoning uses milestoning to develop a Markov state model describing the rates of transitions between important states. It is similar in spirit to using Brownian dynamics and a reduced system description to develop a state model describing transitions. In fact, milestoning has been shown to yield results comparable to those from both brute-force MD and Brownian dynamics to develop a Markov state model.^{15,16} The purpose of this paper is to report the application of Markovian milestoning to the MSOX/O₂ system with the aim of addressing the two major questions posed above. We compare full O₂ transport kinetics in apo (free) and holo (inhibitor-bound) structures of MSOX. We first present a brief outline of the computational methods, followed by detailed protocols. We then discuss the results in terms of the O₂ entry and exit kinetic model for each of the two systems and the overall entry and exit rate constants. Differences in the portal-by-portal rate contributions are discussed in terms of structure, and differences in overall entry and exit rates are used to argue in favor of the modified ping-pong mechanism.

METHODS

Summary

Milestoning generally refers to a method in which the rate of some rare event is computed using many independent MD simulations initiated at discrete locations along a reaction coordinate. These simulations evolve freely until transitioning to another location. The transitions provide input for the statistical analysis and subsequent rate calculations.^{17,18}

Here we apply Markovian milestoning in Voronoi tessellations (MMVT), an application of transition-path theory (TPT)¹³ to milestoning that increases the accuracy and efficiency. MMVT has been presented in detail in the past, and the variant we employ was applied in detail in our CO/Mb study.¹⁴ The major steps in the method are as follows:

1. Compute $F(z)$, the free energy associated with z , the Cartesian position of the center of mass of an O₂ molecule in a coordinate system defined by the protein.
2. Analyze the function F to find local minima and minimum free energy pathways (MFEPs) interconnecting them and solvent portals.
3. Define a set of Voronoi cells with centers discretely chosen along all MFEPs and, for each cell, run confined MD.

4. Analyze the cell-by-cell MD results to determine the mean first passage times (MFPTs) for an O₂ molecule to transit along each pathway in either direction.
5. Incorporate a bulk diffusion model to compute second-order entry rate constants based on the solvent-to-site MFPTs and the bulk O₂ concentration.

We reported previously on steps 1 and 2.¹⁹ There we used the method of single-sweep reconstruction²⁰ to compute F for the apo and holo MSOX states. To apply single sweep, we first sampled the interior of MSOX using the collective variable (CV)-based enhanced sampling technique temperature-accelerated molecular dynamics (TAMD).²¹ Since we were interested in mapping the free energy of O₂ as a function of position, our CV of interest became the center of mass of the O₂ present in either the apo or holo MSOX state. From the long composite TAMD trajectory we selected frames uniformly distributed throughout the interior, where the mean force acting on O₂ was determined. Since the mean force is the derivative of the free energy $F(z)$, integration using a series expansion of radial basis functions (RBFs) was employed, completing the single sweep method. Using free energy reconstruction, we identified several deep minima residing on the *re* face of the flavin isoalloxazine ring, consistent with the established active-site cavity. We then analyzed four or five pathways connecting the minimum closest to the inhibitor with solvent portals, beginning at step 3 above, in the apo and holo states, respectively. The methodological details behind each of the steps above are essentially identical to those presented in their entirety in our previous study of the CO/Mb system,¹⁴ so we present only a brief outline of them here.

Markovian Milestoning

The Voronoi cell B_i centered on CV-space point z_i is defined as all of the configurations whose mappings into CV space are closer to z_i than to any other center. The face between any two adjacent Voronoi cells B_i and B_j is denoted as S_{ij} and is termed a “milestone state”. In our previous work with Mb, we showed that a spacing of 0.667 Å between cells resulted in velocity decorrelation in under 100 fs for CO, a ligand of similar size.¹⁴ Here we increased the initial distance between cell centers to 1 Å. A transition generally can occur from milestone S_{ij} to milestone S_{ik} , which interface cell B_i to cells B_j and B_k , respectively. The MD simulation assigned to any cell B_i is constrained using boundary conditions that are applied to *all* atoms in the system such that the simulation remains in cell B_i at any time step at which the system is detected to be outside B_i ; all of the atom positions are rewound by one time step and all of the atom velocities are rewound by one time step and negated. The state of the simulation refers to the latest milestone state encountered. The MD simulation in each cell B_i with total simulation time T_i is subject to the following analysis:

1. For each adjacent cell B_j , we record $N_{i \rightarrow j}$, the number of attempted transitions across milestone state S_{ij} from within cell B_i , and calculate the quantity $k_{i \rightarrow j} = N_{i \rightarrow j} / T_i$.
2. For all pairs of milestone states (S_{ij}, S_{ik}) , we record $N_{ij,ik}^i$, the number of transitions from S_{ij} to S_{ik} while necessarily remaining within B_i , and calculate the quantity $n_{ij,ik}^i = N_{ij,ik}^i / T_i$.

3. For all S_{ij} we record R_{ij}^i , the time the confined trajectory accumulates after having hit S_{ij} before next hitting any other milestone state, and calculate the quantity $r_{ij}^i = R_{ij}^i / T_i$.

We require that $k_{i \rightarrow j}$, $n_{ij,ik}^i$ and r_{ij}^i saturate to constant values as $T_i \rightarrow \infty$; in practice, this requires between 5 and 50 ns for this system. $k_{i \rightarrow j}$ is the rate estimate for the system to escape from cell B_i to B_j . We require that the equilibrium probability π_i for the system to be located in cell B_i satisfies the following balance equation:

$$\sum_{j=1/j \neq i}^{\Lambda} \pi_j k_{j \rightarrow i} = \sum_{j=1/j \neq i}^{\Lambda} \pi_i k_{i \rightarrow j}, \quad \sum_{i=1}^{\Lambda} \pi_i = 1 \quad (1)$$

The solution gives π_i and consequently the free energy for the system to be located in cell B_i , which is equal to $-k_B T \ln(\pi_i)$. The cell-simulation-specific quantities $n_{ij,ik}^i$ and r_{ij}^i can then be used to evaluate elements of the overall rate matrix $\mathbf{Q}$:

$$q_{ij,ik} = \frac{\pi_i n_{ij,ik}^i}{\pi_i r_{ij}^i + \pi_j r_{ij}^j} \quad (2)$$

From TPT, we know that because the Voronoi faces are locally orthogonal to the MFEPs that connect local minima in F , they are good approximations to *isocommittor surfaces*. The advantage of using Voronoi cell-based milestoneing is that we do not need to generate initial data on the milestones. This task is difficult, as the data must be generated from a nonequilibrium distribution of hitting points and then reinitialized after each iteration.^{17,18} Since the probability that a trajectory launched from any point on an isocommittor surface will reach the product state before returning to the reactant state is invariant, the initial location of O_2 in a particular cell is arbitrary. Avoiding the use of specific milestones is founded not only in the theory of milestoneing²² but also in further work on trajectory parallelization²³ and nonequilibrium umbrella sampling (NEUS).²⁴ These methods are exact in all cases.

Ligand Entry

In the CO/Mb work, our group detailed a new method for handshaking bulk transport of the dissolved gas molecule with the milestoneing description.¹⁴ The basic idea of this method is that we graft a continuum-level description of the diffusion-limited flux into the milestoneing framework with predefined “solvent milestones”. First, for each portal we identify the outermost cell along the MFEP that “leaks” to solvent, that is, a cell in which O_2 accesses the solvent without hitting a milestone. That cell is disregarded, and the space defined by a sphere centered at the center of the next cell inward from that cell and tessellated into a set of “portal cells”. The size of the sphere is chosen such that it represents a boundary between bulk solvent and solvent within interaction distance of the protein atoms that define the portal. Milestoneing MD is run in these, and their data are included in the network of cells.

The outermost set of these cells have milestones that interface with the bulk solvent, which we label specially as “solvent milestones”, S_{sj} .

The key feature of the method is an estimate of the total flux on all solvent milestones:

$$J = \frac{AD[\text{O}_2]}{R} \quad (3)$$

where A_{sj} is the area of solvent milestone S_{sj} , $A = \sum_j A_{sj}$ is the total area of the solvent milestones, D is the bulk self-diffusion constant for O_2 in water at 37 °C, $[\text{O}_2]$ is the bulk concentration of O_2 , and R is the radius of the solvent sphere. The flux $N_{s \rightarrow j}$ on each solvent milestone S_{sj} is proportional to the ratio of the area of this milestone to the total area of the solvent milestones:

$$N_{s \rightarrow j} = \frac{JA_{sj}}{A} = \frac{D[\text{O}_2]A_{sj}}{R} \quad (4)$$

We can then use these fluxes to get all of the cell probabilities, including that of the fictitious “cell” representing bulk solvent, π_s . The rate from a solvent milestone state S_{sj} to any inner (or nonsolvent) milestone state, say, S_{jk} , can be calculated as

$$q_{sj,jk} = \frac{\pi_j n_{sj,jk}^j}{\pi_s r_{sj}^s + \pi_j r_{sj}^j} \quad (5)$$

All of the quantities in this expression are known from the milestone MD simulations except for the fraction of time it is assigned to milestone S_{sj} , r_{sj}^s . Since O_2 is assumed to freely diffuse in the solvent, we can further assume that

$$r_{sj}^s = \frac{A_{sj}}{\sum_{sk} A_{sk}} \quad (6)$$

For pairs of solvent milestone states, we use detailed balance to calculate π_j :

$$\pi_{sj} q_{sj,sl} = \pi_{sl} q_{sl,sj} \quad (7)$$

where $\pi_{sl} \propto (\pi_s r_{sl}^s + \pi_l r_{sl}^l)$ and $\pi_{sj} \propto (\pi_s r_{sj}^s + \pi_j r_{sj}^j)$. Thus, we should have

$$q_{sj,sl} = C(\pi_s r_{sl}^s + \pi_l r_{sl}^l), \quad q_{sl,sj} = C(\pi_s r_{sj}^s + \pi_j r_{sj}^j), \quad \text{etc.} \quad (8)$$

for some constant C with dimensions of a flux; here we simply assume that $C = J$.

TPT Analysis of a Markov Jump Process

The final piece of the method is the extraction of meaningful rate constants and MFPTs from the rate matrix $\mathbf{Q}$. This amounts to using TPT directly on $\mathbf{Q}$ to compute rates between “macrostates”, i.e., chosen subsets of states. To prevent notational proliferation, let us consider in this section only a *general* jump process with rate matrix $\mathbf{K}$ with elements k_{ij} estimating the escape rate (probability per unit time) from cell B_i to cell B_j , linking pairs of states indexed as $i=1, 2, \dots, N$. For any macrostate α that is the union of some set of microstates, TPT permits the calculation of the rate between any pair of macrostates. Let α be the reactant state, with all other macrostates β ($\beta \neq \alpha$) as the aggregate product state. TPT gives the statistical properties of reactive trajectories for the reaction going from the reactant state α to the product state β . The essential quantities to compute are the forward committor function, q_i^+ , defined as the probability that the process starting at i will reach the product state before reaching α , and the backward committor function q_i^- , defined as the probability to last come from α rather than the product state arriving at i . When the process is time-reversible (i.e., the detailed balance is satisfied), $q_i^+ = 1 - q_i^-$. The discrete rate matrix $\mathbf{Q}$ is used to calculate the forward committor function q^+ by solving the following constrained linear problem:²⁵

$$\mathbf{Q}\mathbf{q}^+ = \mathbf{0}, \text{ where } \begin{cases} q_i^+ = 0 & \forall i \in \alpha \\ q_i^+ = 1 & \forall i \in \beta \end{cases} \quad (9)$$

in which α and β represent the reactant and product states, respectively. The transition rate from α to β per unit time is then given by

$$F_{\alpha,\beta} = \sum_{i \in S, j \in \beta, i \neq j} \pi_i k_{ij} (1 - q_i^+) \quad (10)$$

where π_i is the equilibrium probability of state i . Thus, the normalized reaction rate is

$$k_{\alpha,\beta} = \frac{F_{\alpha,\beta}}{\pi_\alpha} \quad (11)$$

where $\pi_\alpha = \sum_{i \in S} \pi_i (1 - q_i^+)$. The total reaction rate from α to the product state is just

$$k_{\alpha,\text{product}} = \sum_{\beta \neq \alpha} k_{\alpha,\beta} \quad (12)$$

Finally, by dividing the rates of interest (i.e., the product and reactant) by all of the rates contributing to the reactive trajectory, we arrive at the fraction of transitions from α to β :

$$f_{\alpha,\beta} = \frac{k_{\alpha,\beta}}{k_{\alpha,\text{product}}} \quad (13)$$

Using this formalism, we can compute pseudo-first-order rate constants for entry from any solvent portal or for all portals at once. We perform this calculation for a series of O₂ concentrations and verify that the second-order rate constant is a constant with respect to O₂ concentration.

Simulation System Setup

Heavy-atom coordinates for MSOX were taken from PDB entry 2gf3.^{1,5} This structure contains several crystal waters and the inhibitor 2-furoic acid (FOA). The holo (FOA-bound) MD system was generated by adding hydrogens to the 2gf3 coordinates where needed, solvating in TIP3P water,²⁶ and neutralizing with Na⁺ ions. This initial system was subjected to 1000 steps of conjugate gradient minimization followed by 130 ns of *NPT* MD equilibration at 310 K and 1 bar. The apo system was created by deleting the FOA from this minimized system and subjecting the resulting system to 130 ns of *NPT* equilibration. After these equilibrations, a molecule of O₂ was added to each system by mutating one of the crystal waters coordinating Lys265. Each system comprised approximately 33 000 atoms. We are unaware of any published structure of MSOX with its natural substrate, sarcosine. Rather than building an experimentally unverifiable model of sarcosine docked in the MSOX binding site, we opted to use the FOA-bound structure directly, with the interpretation that FOA is a “substrate-mimicking” inhibitor.

All of the milestone simulations employed the same MD protocol as in our previous work on single-sweep free energy reconstruction for O₂ in MSOX.¹⁹ Briefly, all of the milestone MD simulations employed the CHARMM22 force field,^{27,28} periodic boundary conditions, a nonbonded cutoff of 10 Å, a particle-mesh Ewald spacing of 2 Å, a time step of 1 fs, and rigid bonds set to “all”, meaning all O–H and H–H distances in water molecules are fixed, and other bonds to H atoms have fixed lengths. Additionally, each hydrogen and the atom it is bonded to are similarly constrained. The temperature was held at 310 K using a standard Langevin thermostat and the pressure at 1 bar using a Langevin–Nosé–Hoover barostat²⁹ with a damping coefficient of 5 ps^{−1} and Langevin hydrogens on. Finally, because the milestone analysis requires that the protein position and orientation define the lab-frame coordinate system, weak positional/rotational restraints were applied. These consisted of Cartesian harmonic restraints on the C_α atoms of residues 25, 100, and 370 with a common spring constant of 1 kcal mol^{−1} Å^{−2}. All of the simulations were conducted with NAMD version 2.9³⁰ customized to support velocity reversal and position rewind upon detection of a Voronoi violation.

Protonation of the flavin was made consistent with its reduced form, FADH[−]. CHARMM-style parameters for the adenine and sugar portions of FADH[−] were adapted from existing parameters for NADH. The flavin ring was parametrized using the AMBER Antechamber procedure with parameters (including charges) determined independently using geometry optimization at the B3LYP/6-311G* level using Gaussian and translated into CHARMM-

style units. FOA was similarly parametrized. The parameter sets for FADH⁻ and FOA were not further optimized. Charmm-style topology and parameter files for FADH⁻ and FOA are available in the supporting information of our previous paper.¹⁹

Infrequent, spontaneous opening of the flavin cleft was observed in the apo equilibration, as evidenced by the breakage of a contact between Asn41 and Arg282. Initial milestone results in the region of the open cleft showed O₂ exiting into the solvent from deep within the interior along both MFEPs traversing the cleft. These large deviations from the MFEP indicate a breakdown in the assumption that our ligand should travel along a tubelike path before reaching the solvent at the end of the MFEP. We were unable to confine O₂ to the MFEP without solvent transitions outweighing those along the MFEP, and we concluded that with a completely open cleft, MSOX is porous and not suitable for analysis with milestone. Additionally, our computed entry rates were predicted to be near the diffusion limit for the closed-cleft system already, and this was unlikely to be different for a state where O₂ entry/exit is easier by inspection. We therefore removed the open cleft configuration from further analysis.

Milestone Protocols and Pathway Refinement

On the basis of the free energy reconstructions, we previously identified four MFEPs leading from the reactive re face of the flavin ring to the solvent for each system, which we term the “Lys”, “Gate”, “Lower”, and “Sarc(osine)” pathways.¹⁹ For each path, depending on its length, between 50 and 80 centers were chosen along it to form the initial Voronoi centers. MD simulations were run with the O₂ center of mass initialized in the center of each cell, and on the basis of these results, some pathway refinements and further milestone MD simulations were performed. The major objective of the first set of milestone MD simulations was to identify the solvent milestones for each pathway. Once these were identified, the portal cells were constructed and subjected to milestone MD to complete the data set. Figure 2 illustrates this procedure with hitting-point data on solvent milestones for the Apo-Sarc gateway.

A second issue that required refinement of the set of Voronoi centers arose from an infrequent problem we termed “pinching” in certain interior cells. Because the milestones are planes in 3-space that are defined by local orthogonality to an MFEP, they can intersect in regions where the MFEP is not perfectly straight. The hope in using such milestones is that the MD simulations inside each cell will be naturally confined to sampling near the MFEP and therefore far from the regions of milestone intersections. In such regions, of course, the milestones are quite far from being isocommitters, which by definition foliate space and therefore *cannot* intersect. However, if the free energy is not steeply graded down to the MFEP in some cells, MD can result in sampling of the region of the cell near a pinch. This is problematic because the transition rates in such cases become unreasonably large as the milestone-to-milestone transition events become artificially too frequent (e.g., when the distance between the milestones is about the same as the average displacement of an atom in one time step of numerical integration). After some trial and error, we determined that a suitable way to mitigate this behavior is to merge cells by deleting centers of cells in which pinching occurs. This led to several relatively large cells (centers 5 Å or more away from

nearest neighbors), which required 20 ns or more to saturate the transition statistics. After construction of all of the portal cells and deletion of the centers of initially pinching cells, the final set of Voronoi cells consisted of 178 centers for apo MSOX and 177 for holo MSOX.

There are some important details regarding how we established the portal cells for the apo and holo systems, which we illustrate in Figure 3A,B, respectively. First, the Gate and Lys pathways in the apo system have portals that are relatively close together on the protein surface. We therefore used a single portal sphere to encompass both (Figure 3A), and thus, their contribution to the entry rate is reported as combined. Also in the apo system, during the initial set of milestone MD simulations we discovered a cell that, although not on the protein surface, could escape into the solvent. This “leaky” cell occurred along the Lys pathway, and its corresponding portal was distant from all of the others, therefore requiring its own portal sphere. We refer to this new portal as “Mid-Lys” (Figure 3A), and its contributions to entry and exit are delineated from all others.

Finally, in the holo system, one MFEP exits the active site through a channel that was observed to be blocked in MD simulations by the side chain of Phe256. When milestone MD was run along this pathway, the rigidity of the side-chain torsion of Phe256 prevented transit in one cell. We hypothesized that the MFEP detected here was an artifact resulting from the finite irregular mesh of points in the singlesweep reconstruction, with one point placed on either side of Phe256. We therefore ran an independent adaptive biasing force (ABF) simulation³¹ to assess the free energy associated with flipping the side chain of Phe256 into a state that would permit transit. As can be seen in Figure 4A, the free energy barrier required to open the channel is about 12 kcal/mol, and the closed state is 10 kcal/mol more stable than the open state. We therefore chose to ignore the possibility that Phe256 spontaneously opens and remains open, and we did not perform milestone along the Lys pathway for the holo system (Figure 3B).

Within the protein interior, each cell required 10–50 ns to attain sufficient sampling, as determined by saturation of the running average for all within a given cell. With the interior completed, the solvent portals were examined. Since the free energy gradients are rather shallow over large distances in this region, the MFEP, although calculable, does not really correspond to the center of a compact tubelike transport channel, and we could not do optimal milestone with Voronoi centers along such an MFEP. Instead, we included TAMD frames uniformly distributed within a sphere of 10 or 12.5 Å centered on the last string image center with an approximate spacing of 2.5 Å. Only frames in the direction of the solvent were kept. Those falling on the MFEP were excluded. These cells were prepared in the same way as the MFEP and run to saturation. The solvent cells had a much larger spacing (2.5 Å) compared with the majority of the interior (~0.7–1 Å) yet typically required between 10 and 25 ns to saturate.

RESULTS

Raw Milestoning Results

To give an impression of the amount of simulation data that is processed and to provide a view of the 3D structure of the pathways, we show milestoning MD hitting points on all of the milestones along the pathways in Figure 5 for the holo system. (The impression given by the same view of the apo system is identical, though the data of course differ in detail.) The milestone planes are for the most part easily observable. It should be noted that each plane has hitting points that arise from collisions occurring in two MD simulations, those in cells on either side of the milestone.

Rates and Mechanisms of O₂ Entry

In Figure 6, we summarize the entire set of milestoning results in the form of two kinetic networks, one for the apo system and the other for the holo system, that represent individual pathways of O₂ entry into and exit from the MSOX hydrophobic cavity on the re face of the flavin ring. For these individual representations, the Markov state model defines the active-site macrostate as the node from which all pathways extend. Each portal macrostate is the set of solvent milestones that interface with the solvent at each respective portal. Each arrow is labeled by the MFPT in microseconds. We first consider the entry pathways.

In the apo system (Figure 6 left), the dominant pathways carrying O₂ from the solvent to the active site are Gate/Lys, Mid-Lys, and Sarc. The MFPTs from the solvent portal to the active site are about 1190, 1886, and 842 μ s for the Sarc, Gate/Lys, and Mid-Lys pathways, respectively, at an O₂ concentration of 209 μ M, corresponding to saturation at 37 °C and 1 bar in equilibrium with air. Hence, in the apo system, we see that the protein is essentially porous, with three entry pathways that are all significant contributors to the overall rate of entry.

In the holo system (Figure 6 right), the situation changes markedly. Here the MFPT along the Sarc pathway is greatly reduced while that along the Gate pathway is increased so that it becomes effectively infinite. Also, as noted earlier, the holo system does not have a Lys or Mid-Lys pathway thanks to Phe256. Therefore, in the holo system, O₂ predominantly enters by the Sarc pathway. In both systems, the MFPT along the lower pathway is so large that it is effectively infinite, meaning that although there is a path there, it does not contribute at all to the O₂ entry rate in either system.

The overall second-order rate constant for entry can be computed at any O₂ concentration, and it is observed to be essentially invariant across O₂ concentrations for both the apo and holo systems. The calculation of second-order rate constants is based on the overall MFPTs presented in Table 1 since it encompasses the entire interconnected network rather than a lone pathway. Interestingly, although the weights of the contributions of the pathways to the overall entry rate shift between the apo and holo systems, the overall rate itself is not strongly affected by substrate-mimic binding. We predict that $k_{\text{entry}} = (8.1 \pm 0.1) \times 10^6$ and $(3.12 \pm 0.1) \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ for the holo and apo states, respectively. These values are about an order of magnitude higher than the second-order rate constant for oxidation of reduced wild-type MSOX, $(2.83 \pm 0.07) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, determined during mutagenesis studies⁵ (i.e.,

k_3 in Figure 1) but about the same as the second-order rate constant for O_2 consumption by the most active flavoenzymes.³²

Because $k_{\text{entry}} > k_3$ in both the apo and holo cases considered here, it is unlikely that entry of O_2 limits the rate of flavin oxidation. It should be borne in mind that we define entry as an event in which an O_2 transits from the solvent to the local free energy minimum associated with the active site, as computed using single-sweep reconstruction, and that the standard molecular mechanics force field used (CHARMM) cannot model specific charge-transfer interactions. The oxidation steps in the standard mechanistic picture can therefore be subdivided into more fundamental steps of entry to the active-site cavity, acquisition of a specific O_2 /flavin complex (likely involving polarization), and then electron transfer to complete the reaction. Our second-order rate constant corresponds to the first of these fundamental steps, and given that the overall rate of the oxidation step is lower than our computed entry rate, we conclude that the fundamental step of entry is not rate-limiting.

Rates and Mechanisms of O_2 Exit

We look now at the individual exit rates for apo and holo MSOX. Oxygen exit is faster in the apo state, as the exit times are on the order of nanoseconds rather than microseconds. We also note that no one path clearly contains the majority of the flux. As with entry, the Sarc, Gate/Lys, and Mid-Lys pathways exhibit comparable exit MFPTs of 22, 33, and 14 ns, respectively. Once again we see the apo state exhibiting porous behavior.

In the holo system, we see one dominant exit route. The Sarc pathway displays an MFPT of 140 ns, which is nearly 2 orders of magnitude shorter than Gate at 10 μ s. As with entry, exit MFPTs on the Lower pathway are so large as to be considered infinite, effectively not contributing to the overall exit of O_2 .

On the basis of the overall exit MFPTs, we find that $k_{\text{exit}} = 10^7$ and $7.2 \times 10^6 \text{ s}^{-1}$ for the apo and holo states, respectively. Differing by an order of magnitude, it is more likely that exit is limiting in the case of O_2 oxidation of the flavin. A longer residence time within MSOX increases the likelihood of collisions with the flavin isoalloxazine ring, which are required for reduction of molecular O_2 and subsequent reoxidation of the flavin. This is further supported by the fact that holo MSOX shuts down all but one exit route, the so-called Sarc pathway. Evolution appears to have designed MSOX in such a way as to be porous without a substrate in the active site, while ensuring the highest probability of O_2 being in contact with the isoalloxazine ring when a substrate is present in the active site.

Presence of Inhibitor and Closure of the Flavin Cleft Correlate to Slower O_2 Exit

In comparing the holo and apo states, there are no major structural differences within the vicinity of the flavin isoalloxazine ring (save the presence of the inhibitor FOA itself) that lend themselves to an easy explanation for the shutdown of the Gate pathway in the holo state. Our long MD trajectory does not show any opening or closing of residues that would facilitate or hinder the progress of O_2 . However, further out from the ring, we see that the cofactor-admitting cleft exhibits two unique configurations when holo and apo MSOX are compared. It remains tightly closed in the holo state yet is less tightly packed in the apo state, as shown in Figure 7. There is no pathway analogous to the apo state's Gate path in the

holo state that does not overlap residues. This likely explains why the only MFEP identified through this portal in the holo state does not permit transit during milestone MD, as described previously in regard to Phe256. We therefore see a direct effect of inhibitor binding on a variety of transport pathways.

Binding Site Desolvation Correlates to Faster O₂ Entry

On the basis of the portal-specific MFPTs, inhibitor binding effectively closes off all of the pathways available in the apo state except one: the Sarc pathway. However, along this pathway the rate of entry almost doubles relative to that in the apo state. This may be explained by the fact that the presence of the inhibitor desolvates the active site, excluding approximately three molecules of water that easily access the *re* face, as shown in Figure 8. Addition of O₂ to the *re* face removes another water molecule, resulting in a net loss of three to four water molecules from the active site. Desolvation at the active site has also been proposed as important for reoxidation of the flavin in other flavoenzymes.³² We observe that the more hydrophobic environment in the binding site created by the inhibitor both speeds the entry of O₂ along the Sarc pathway and slows its exit.

CONCLUSION

We have used Markovian milestone MD simulations to compute the entry and exit kinetics of O₂ in MSOX. The overall second-order entry rate constant increases by a factor of about 2.5 upon inhibitor binding, while the exit rate decreases by about an order of magnitude. The net effect is the prediction of increased O₂ residence time in the active site when the inhibitor is bound relative to the apo case, which provides indirect support for the modified ping-pong mechanism. We observe that binding of the inhibitor shuts down multiple routes of entry and exit. Combined with substrate-mediated desolvation, this evidence suggests that it is kinetically favorable that flavin reoxidation in MSOX operates via the modified pingpong mechanism.

Acknowledgments

Funding

This work was supported by the National Institutes of Health under Grant R01GM100472. The calculations reported in this work were performed at the TACC Supercomputing Center using an XSEDE allocation (TG-MCB070073N). Some of the work reported here was run on hardware supported by Drexel's University Research Computing Facility.

References

1. Trickey P, Wagner MA, Jorns MS, Mathews FS. Structure. 1999; 7:331–345. [PubMed: 10368302]
2. Wagner MA, Trickey P, Chen Z-W, Mathews FS, Jorns MS. Biochemistry. 2000; 39:8813–8824. [PubMed: 10913292]
3. Wagner MA, Jorns MS. Biochemistry. 2000; 39:8825–8829. [PubMed: 10913293]
4. Zhao G, Jorns MS. Biochemistry. 2002; 41:9747–9750. [PubMed: 12146940]
5. Zhao G, Bruckner RC, Jorns MS. Biochemistry. 2008; 47:9124–9135. [PubMed: 18693755]
6. Hassan-Abdallah A, Zhao G, Chen Z-W, Mathews FS, Schuman Jorns M. Biochemistry. 2008; 47:2913–2922. [PubMed: 18251505]
7. Jorns MS, Chen Z-W, Mathews FS. Biochemistry. 2010; 49:3631–3639. [PubMed: 20353187]

8. Kommoju P-R, Chen Z-w, Bruckner RC, Mathews FS, Jorns MS. *Biochemistry*. 2011; 50:5521–5534. [PubMed: 21568312]
9. Saam J, Ivanov I, Walther M, Holzhütter H-G, Kuhn H. *Proc Natl Acad Sci U S A*. 2007; 104:13319–13324. [PubMed: 17675410]
10. Baron R, Riley C, Chenprakhon P, Thotsaporn K, Winter RT, Alfieri A, Forneris F, van Berkel WJH, Chaiyen P, Fraaije MW, Mattevi A, McCammon JA. *Proc Natl Acad Sci U S A*. 2009; 106:10603–10608. [PubMed: 19541622]
11. Shadrina M, Peslherbe G, English A. *Biochemistry*. 2015; 54:5279–5289. [PubMed: 26226401]
12. DiRusso N, Conductor H, Li K, Bruner S, Roitberg A. *Chem Sci*. 2015; 6:6341–6348. [PubMed: 26508997]
13. Vanden-Eijnden E, Venturoli M. *J Chem Phys*. 2009; 130:194101. [PubMed: 19466815]
14. Yu T-Q, Lapelosa M, Vanden-Eijnden E, Abrams CF. *J Am Chem Soc*. 2015; 137:3041–3050. [PubMed: 25664858]
15. Luty BA, El Amrani S, McCammon JA. *J Am Chem Soc*. 1993; 115:11874–11877.
16. Votapka LW, Amaro RE. *PLoS Comput Biol*. 2015; 11:e1004381. [PubMed: 26505480]
17. Faradjian AK, Elber R. *J Chem Phys*. 2004; 120:10880–10889. [PubMed: 15268118]
18. Májek P, Elber R. *J Chem Theory Comput*. 2010; 6:1805–1817. [PubMed: 20596240]
19. Bucci A, Abrams C. *J Chem Theory Comput*. 2014; 10:2668–2676. [PubMed: 25061440]
20. Maragliano L, Vanden-Eijnden E. *J Chem Phys*. 2008; 128:184110. [PubMed: 18532802]
21. Maragliano L, Vanden-Eijnden E. *Chem Phys Lett*. 2006; 426:168–175.
22. Vanden-Eijnden E, Venturoli M, Ciccotti G, Elber R. *J Chem Phys*. 2008; 129:174102. [PubMed: 19045328]
23. Vanden-Eijnden E, Venturoli M. *J Chem Phys*. 2009; 131:044120. [PubMed: 19655850]
24. Dickson A, Warmflash A, Dinner AR. *J Chem Phys*. 2009; 130:074104. [PubMed: 19239281]
25. Held M, Metzner P, Prinz J-H, Noe F. *Biophys J*. 2011; 100:701–710. [PubMed: 21281585]
26. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. *J Chem Phys*. 1983; 79:926.
27. MacKerell AD Jr, et al. *J Phys Chem B*. 1998; 102:3586–3616. [PubMed: 24889800]
28. MacKerell AD Jr, Banavali NK. *J Comput Chem*. 2000; 21:105–120.
29. Melchionna S, Ciccotti G, Holian BL. *Mol Phys*. 1993; 78:533–544.
30. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. *J Comput Chem*. 2005; 26:1781–1802. [PubMed: 16222654]
31. Darve E, Rodriguez-Gomez D, Pohorille A. *J Chem Phys*. 2008; 128:144120. [PubMed: 18412436]
32. Mattevi A. *Trends Biochem Sci*. 2006; 31:276–283. [PubMed: 16600599]

Figure 1.

Elementary mechanisms of the MSOX catalytic cycle.³ The lower branch represents the classical ping-pong mechanism, while the upper branch represents the modified ping-pong mechanism. Reduced and oxidized states of the flavin are shown as E_{red} and E_{ox} , respectively. Substrates and products are denoted as S and P, respectively.

Figure 2.

Sample portal setup. The Apo-Sarc MFEP is discretized into its interior Voronoi tessellation centers (cyan), with the first cell exiting to the solvent shown as a larger blue sphere. Cell centers defining the tessellation in a 12.5 Å sphere (cyan) are shown along with the first 5 ns of data for any transition out of the sphere (red). Nearby protein structure is shown in cartoon representation (black) and the cofactor in space-filling representation (orange).

Figure 3. Schematic representations of pathways and portals in the (A) apo and (B) holo systems, established by refinement of MFEPs from the string method using milestone MD as described in the text. The large black dot designates the active-site cavity in which the free energy minimum for O_2 is located for both systems. Arcs at each portal signify the spherical portal boundaries.

Figure 4.

(A) Free energy profile from a 10 ns adaptive biasing force (ABF) MD simulation sampling the C-C α -C β -C γ torsion of Phe256, χ_1 . Open (*) and closed (**) configurations are indicated. (B) Snapshots from the ABF trajectory illustrating the (left) closed and (right) open configurations of the Phe256 side chain. Only atoms in Phe256 (colored on the basis of atom name), the flavin (gray), and Lys265 (white) are shown. (C) Trace of Phe256 χ_1 vs time in the ABF calculation, illustrating that χ_1 has repeatedly sampled its domain.

Figure 5. Representative selection of milestone hitting points (red) for the holo system, with the protein represented in black. Major pathways (cyan) and their portals are also labeled. The first 5 ns of data is shown in each cell.

Figure 6. Kinetic networks for the (left) apo and (right) holo systems. Labels indicate MFPTs in microseconds. Entry times correspond to entry from aqueous solution at an O_2 concentration of $209 \mu M$.

Figure 7.

Aligned holo (yellow, space-filling) and apo (blue, spacefilling) structures showing the relative states of the cofactor-admitting cleft during equilibrium MD. The cofactor (pink, space-filling) is shown as well as the apo-Lys MFEP (red, spheres). The holo structure is tightly packed around the cofactor, while apo is more relaxed.

Figure 8. Water occupancy distributions in the active site from long equilibrium MD (defined as a sphere of radius 6 Å centered at atom C₁₀ of the flavin ring) for apo and holo MSOX.

Table 1

Overall Entry and Exit MFPTs for Apo and Holo MSOX; Entry Times Correspond to Entry from Aqueous Solution at an O₂ Concentration of 209 μ M

	apo	holo
entry	1553 μ s	590 μ s
exit	10 ns	138 ns