Full Kinetics of CO Entry, Internal Diffusion, and Exit in Myoglobin from Transition-Path Theory Simulations

Tang-Qing Yu,[†] Mauro Lapelosa,[‡,§] Eric Vanden-Eijnden,[†] and Cameron F. Abrams*,[‡]

[†]Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, United States
[‡]Department of Chemical and Biological Engineering, Drexel University, Philadelphia, Pennsylvania 19104, United States

ⓢ *Supporting Information*

ABSTRACT: We use Markovian milestoning molecular dynamics (MD) simulations on a tessellation of the collective variable space for CO localization in myoglobin to estimate the kinetics of entry, exit, and internal site-hopping. The tessellation is determined by analysis of the free-energy surface in that space using transition-path theory (TPT), which provides criteria for defining optimal milestones, allowing short, independent, cell-constrained MD simulations to provide properly weighted kinetic data. We coarse grain the resulting kinetic model at two levels: first, using crystallographically relevant internal cavities and their predicted interconnections and solvent portals; and second, as a

$$q_{ij,ik} = \frac{N_{ij,ik}}{R_{ij}}$$

three-state side-path scheme inspired by similar models developed from geminate recombination experiments. We show semiquantitative agreement with experiment on entry and exit rates and in the identification of the so-called "histidine gate" at position 64 through which ≈90% of flux between solvent and the distal pocket passes. We also show with six-dimensional calculations that the minimum free-energy pathway of escape through the histidine gate is a "knock-on" mechanism in which motion of the ligand and the gate are sequential and interdependent. In total, these results suggest that such TPT simulations are indeed a promising approach to overcome the practical time-scale limitations of MD to allow reliable estimation of transition mechanisms and rates among metastable states.
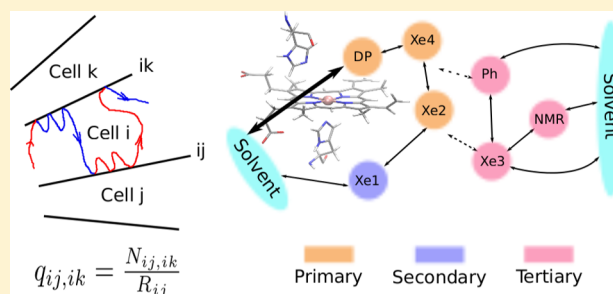
■ INTRODUCTION

A large fraction of all biochemical processes rely on small dissolved gas molecules gaining entry into buried cavities and active sites in proteins and enzymes. The ability to shield such sites from undesired molecules while permitting access by desired molecules is one of the many interesting consequences of proteins' particular 3D structures that has yet to enjoy a general and complete understanding. Myoglobin (Mb), more than perhaps any other protein, has played a central role in seeking this understanding, using a variety of experimental and all-atom simulation approaches.[1−30] Mb is a heme-containing oxygen storage and transport protein in muscle cells, and its liganded form is defined by a small gas molecule (O_2, CO, or NO) covalently bound to the heme iron cation. Much of what we know about how such molecules gain entry comes from analyzing what happens when such a state is perturbed by flash photolysis to break that bond. Models fit to kinetics of geminate recombination,[5,6,8] as well as direct observation from time-resolved X-ray crytallography,[4,9−11,13,14,16] clearly show that dissociated ligands sample internal cavities in Mb distinct from the distal pocket at the heme prior to exiting or rebinding. Careful site-directed mutatgenesis studies also show that both exit and entry occur predominantly (>70%) through a gate defined by the histidine residue at position 64.[6,7] As pointed out in a recent review, however, essentially all molecular simulation studies that model small molecules diffusing around inside Mb

agree that there are multiple internal sites and, distressingly, multiple important entry/exit portals.[31]

It is possible that this apparent disagreement between simulations and experiments is rooted in the fact that simulations have yet to provide direct estimates of entry and escape kinetics. Due to the rare-event essence of the ligand diffusion inside or out of Mb in any simulation model, regular molecular dynamics (MD) simulations are too short in duration to provide kinetic information for such long time-scale phenomenon. Though kinetics, e.g., mean-first-passage-times (MFPT's), have in the past been estimated from collections of MD trajectories of small molecules inside Mb or other globins,[23,24,32,33] questions remain regarding the ergodicity of those simulations and the potential bias this might present in the results. Free-energy calculations along the transition paths have been reported by many,[25,26,28] but rate estimations based on transition-state theory (TST) are still far from being consistent with the experimental measurements.[31]

Development of enhanced sampling techniques for rare events has accelerated in recent years. In particular, Voronoi-cell Markovian milestoning[34] circumvents the time-scale issue in straightforward MD simulation. Milestoning is generally a method in which progress along a reaction coordinate is

measured using discrete, separate MD simulations that collectively span the reaction coordinate.[34−36] Markovian milestoning is an implementation of transition-path theory (TPT), the central quantity of which is the committor function, which is a local property of state space that measures the probability of a dynamical trajectory launched from a location encountering a product state before a reactant state. The Markovian property is guaranteed by choosing milestones as isocommittors,[34,37] thus obviating the need to reinitialize the milestoning MD simtulations from a noncanonical (first hitting point) distribution on the milestones. For two metastable states with a rare transition between them, the kinetics can be recovered from the transition rates from the milestones in one state to the milestones in the other state. Because MD in a confined space is able to offer sufficient statistics within a short running time and the simulations in each cell can be parallelized easily, milestoning represents a potentially large enough gain in efficiency to make reliable estimates of ligand entry/exit kinetics possible using all-atom simulations.

Here, we applied the Markovian milestoning method to calculate rates of migration of CO among Mb cavities and solvent. We have shown that the milestoning method can treat multiple pathways forming a network, and we further developed the methodology to account for bulk ligand concentration and diffusivity into account in the calculation of ligand entry rates. We report all MFPT's between cavities and solvent and construct a full kinetic network based on the calculated rates from TPT analysis. From the kinetics, the dominant entry/exit pathway is revealed unambiguously to be the one through the "histidine gate" (HG). Thus, we also performed a detailed analysis of the gating process using string method. These studies provide a paradigm for all-atom milestoning simulations in estimating kinetics of biomolecular processes.

■ METHODS

Markovian Milestoning. Generally, milestoning is a procedure for running locally restrained, independent MD simulations that collectively sample reactive trajectories between two metastable states and combines them to estimate the rates (reciprocal MFPT's) between the two states.[35,36] Markovian milestoning uses isocommittor surfaces as milestones. These milestones are optimal in the sense that they guarantee the Markovian assumption for transition between two neighboring milestones. This avoids the need to initialize each MD simulation a noncanonical distribution (i.e., first-hitting point) on the milestones.[37] Once a dominant transition pathway (e.g., minimum free-energy pathways (MFEP's)) between any two states is found, the isocommittors can be approximated by hyperplanes locally orthogonal to it. This can be modeled as a Voronoi tessellation with centers spaced along the MFEP. The idea of Markovian milestoning using Voronoi tessellation[34] is, given a Voronoi tessellation partitioning the configuration space, running many local simulations each confined in a Voronoi cell and collecting the kinetics about the transition process. A detailed description is given in the following.

Let $x \in \Omega \subset \mathbb{R}^d$ represent the coordinates of the system and $\theta(x) = (\theta_1(x), ..., \theta_M(x))$ are our collective variables (CV's). We define a set of points in CV space $z_i \in \mathbb{R}^M$ with $i = 1, 2, 3, ..., \Lambda$, which partition the configuration space into Λ Voronoi cells. The Voronoi cell B_i at z_i is defined as

$$B_i = \{x \in \Omega \colon \|\theta(x) - z_i\| < \|\theta(x) - z_j\| \text{ for all } j \neq i\} \quad (1)$$

where $\|\cdot\|$ is Euclidean distance. The face between any two adjacent Voronoi cells B_i and B_j is denoted S_{ij} and serves as a milestone state. Note, though these milestones are defined in CV space, they are embedded in $\mathbb{R}^d$ and thus partition the whole configuration space. If

we imagine a long MD trajectory, each milestone crossing event (well-defined according to $\theta[x(t)]$ of the trajectory) changes the coarse-grained state of the system, and the entire trajectory can then be conceptualized as a continuous-time Markovian jump process among such states. From here on, we refer to these states as "milestone states" since each corresponds to a particular milestone. Note that a direct jump only occurs between the milestones associated with the same cell; that is, a jump can occur from milestone S_{ij} to S_{ik}, which interface cell B_i to cells B_j and B_k, respectively. We use $i, j, k, ...$ to index cells, and since each milestone interfaces exactly two cells, each is uniquely identified by a pair of cell indices, ij, etc.

Full kinetic information on the level of milestones could be inferred directly from an infinitely long MD trajectory, which is coarse-grained according to the last milestone state visited S_{ij} and the elapsed time since that visit. This information is most directly captured by the quantities $N_{ij,ik}$ and R_{ij}, where $N_{ij,ik}$ records the number of times that the trajectory collides with face S_{ik} after having last hit face S_{ij}, and R_{ij} records the total time the system spends in state ij, that is the total time in which face S_{ij} was the last most recent milestone encountered. With these quantities, we can construct the rate matrix $\mathbf{Q}$ associated with the continuous-time Markov jump process interconnecting states. The nondiagonal elements of $\mathbf{Q}$ are given by

$$q_{ij,ik} = \begin{cases} \dfrac{N_{ij,ik}}{R_{ij}} & \text{if } R_{ij} \neq 0 \\[2mm] 0 & \text{if } R_{ij} = 0 \end{cases} \quad (2)$$

where $q_{ij,ik}$ is the rate from milestone state S_{ij} to S_{ik}. This reduces the original dynamics to a jump process with rate matrix $\mathbf{Q}$. Now, because we do not have an infinitely long MD trajectory, we turn to Markovian milestoning, which instead requires running short independent MD simulations confined to each cell. In practical terms, a simulation is assigned to cell B_i and constrained using reflecting boundary conditions:

$$x_i(t + \Delta t) = \begin{cases} x_i(t + \Delta t) & \text{if } (x_i(t + \Delta t) \in B_i) \\ x_i(t) & \text{otherwise} \end{cases} \quad (3)$$

and

$$v_i(t + \Delta t) = \begin{cases} v_i(t + \Delta t) & \text{if } (x_i(t + \Delta t) \in B_i) \\ -v_i(t) & \text{otherwise} \end{cases} \quad (4)$$

Note that these are applied to the positions and velocities of all atoms in the system, not just those that determine the CV. This guarantees the entire trajectory mimics fragments of an infinitely long MD trajectory. In practice, the violation check in these conditions is simply a distance check in CV space based on eq 1, which is easily implemented in any regular MD simulation and bears very little overhead.

To reconstruct estimates for the quantities $N_{ij,ik}$ and R_{ij}, the following procedure is used for the MD simulation in each cell (e.g., Voronoi cell B_i):

(1) For each adjacent cell B_j, we record the number of attempted transitions across milestone state S_{ij} from within cell B_i, $N_{i \to j}$ and calculate the quantity $k_{i \to j} = N_{i \to j}/T_i$.

(2) For all pairs of milestone states (S_{ij}, S_{ik}), we record the number of transitions from S_{ij} to S_{ik} while necessarily remaining within B_i, $N^i_{ij,ik}$ and calculate the quantity $n^i_{ij,ik} = N^i_{ij,ik}/T_i$.

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

Here T_i is the total simulation time in cell B_i. Naturally, $k_{i \to j}$, $n^i_{ij,ik}$ and r^i_{ij} should asymptote to constant values as $T_i \to \infty$. $k_{i \to j}$ 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 locate in cell B_i satisfies a balance equation (meaning the total flux in and out of each cell is zero at statistical equilibrium):
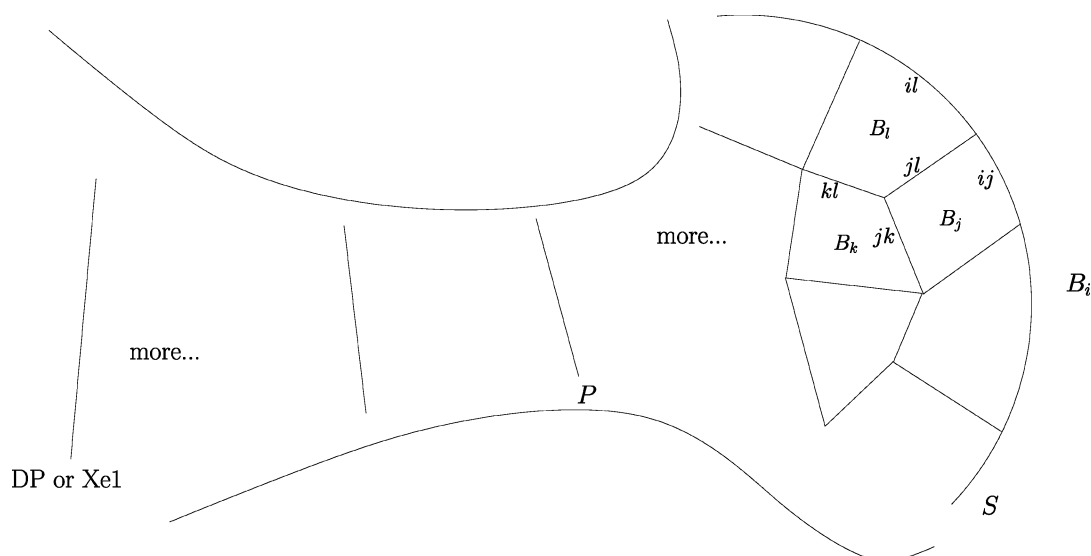
Figure 1. A sketch of cells and milestones at a portal. P is the portal milestone at HG or Xe1s. S is the spherical boundary comprised of milestones interfacing solvent. (Note that we label each milestone with double-letter indices of cells that share that milestone.)

$$\sum_{\substack{j=1 \\ j\neq i}}^{\Lambda} \pi_j k_{j\to i} = \sum_{\substack{j=1 \\ j\neq i}}^{\Lambda} \pi_i k_{i\to j}, \quad \sum_{i=1}^{\Lambda} \pi_i = 1 \tag{5}$$

The solution gives π_i, and consequently the free energy for locating the system in cell B_i, $-k_B T \ln(\pi_i)$. The cell-simulation-specific quantities $n_{ij,ik}^i$ and r_{ij}^i can then be used to evaluate

$$q_{ij,ik} = \frac{\pi_i n_{ij,ik}^i}{\pi_i r_{ij}^i + \pi_j r_{ij}^j} \tag{6}$$

Note that the $q_{ij,ik}$ and $q_{ik,ij}$ satisfy detailed balance according to

$$\pi_{ij} q_{ij,ik} = \pi_{ik} q_{ik,ij} \tag{7}$$

where $\pi_{ij} = CR_{ij}$ and $\pi_{ik} = CR_{ik}$, if $N_{ij,ik}^i = N_{ik,ij}^i$ (thus $n_{ij,ik}^i = n_{ik,ij}^i$), which is naturally fulfilled in a long enough MD trajectory in B_i or can easily be enforced.

Milestoning for Ligand Entry. Entry refers to the process by which a ligand molecule transitions from bulk solvent through a portal on the protein surface and eventually settles at a buried site. The entry rate is affected by the external concentration and diffusivity of the ligand. Here we extend the original milestoning method to take these properties into account in calculation of rate matrix. We graft a continuum-level description of the diffusion-limited flux with predefined solvent milestones into the milestoning framework. This harkens back to the original work of Smoluchowski[38] and Shoup and Szabo.[39] To begin, Figure 1 illustrates of the cells and milestones at a portal region. In Figure 1, we term cell B_i the solvent cell and B_j, B_k, B_l, ..., the inner cells. The inner cells interface directly with the solvent cell through milestones denoted S_{ij}, S_{il}, ..., which we label specially as "solvent milestones". Milestones S_{kl}, S_{jk} S_{jl}, ..., between any pair of inner cells are termed "inner milestones".

First, we need to know the cell-to-cell fluxes to calculate cell probabilities using eq 5. From constrained MD simulations in the cells, we know the fluxes between inner cells, e.g., $N_{j\to k}$. From these simulations we also have the fluxes from an inner cell to the solvent cell, e.g., $N_{j\to i}$. However, since we do not perform MD simulation in the solvent cell, the fluxes from the solvent cell to any inner cell must be obtained differently. We propose to calculate them analytically from an appropriate diffusion equation. Specifically, if we approximate the protein as a sphere with radius R, the total flux on all the portal milestones is

$$J = J_{tot} \frac{A}{4\pi R^2} \quad \text{where } J_{tot} = 4\pi RD[\text{CO}] \tag{8}$$

where $A = \sum_k A_{ik}$ is the total area of the solvent milestones, and A_{ik} is the area of solvent milestone S_{ik}, D is the bulk self-diffusion constant for CO in water at 37 °C, and $[\text{CO}]$ is the bulk concentration of CO. This gives

$$J = \frac{AD[\text{CO}]}{R} \tag{9}$$

The flux on each solvent milestone, say S_{ij}, is proportional to the ratio between the area of this milestone and the total area of the solvent milestones,

$$N_{i\to j} = \frac{JA_{ij}}{A} = \frac{D[\text{CO}]A_{ij}}{R} \tag{10}$$

where A_{ij} is the area of solvent milestone S_{ij}. We can then use these fluxes in eq 5 to get all cell probabilities, including that of the solvent cell, π_i. The rates between inner milestones, or from an inner to a portal milestone can be calculated as in eqs 2 and 7. The rate from a solvent milestone state S_{ij} to any inner milestone state, say, S_{jk} can be calculated from

$$q_{ij,jk} = \frac{\pi_j n_{ij,jk}^j}{\pi_i r_{ij}^i + \pi_j r_{ij}^j} \tag{11}$$

which is identical to eq 2. All quantities in this expression are known from the milestoning MD simulations, except for r_{ij}^i, which we calculate analytically, as follows. If we assume that the solvent is defined as a phase in which ligand can diffuse freely, then the fraction of time it is assigned to milestone S_{ij} should be equal to the ratio between the area of this milestone and the total area of all the portal milestones, that is,

$$r_{ij}^i = \frac{R_{ij}^i}{T_i} = \frac{A_{ij}}{\sum_{ik} A_{ik}} \tag{12}$$

Note that this rate satisfies detailed balance in eq 7 with respect to $q_{jk,ij}$ since $N_{jk,ij}^j = N_{ij,jk}^j$ (thus $n_{jk,ij}^j = n_{ij,jk}^j$) holds.

The rates between different solvent milestones need special treatment. We will assume that the rates are independent of the initial state, the milestone from which the system starts, i.e., $q_{ij,il} \equiv q_{ik}$. The assumption is based on the fact that when the ligand leaves a solvent milestone, it is immersed into the solvent phase quickly and "forgets" where it came from. In other words the portal it will hit next (conditional on this next milestone being another portal milestone) is

independent of the portal it hit last. Since we also require that the detailed balance should hold for the rates between solvent milestones, we must have

$$\pi_{ij} q_{ij,il} = \pi_{il} q_{il,ij} \qquad (13)$$

where $\pi_{il} \propto (\pi_i r_{il}^i + \pi_l r_{il}^l)$ and $\pi_{ij} \propto (\pi_i r_{ij}^i + \pi_j r_{ij}^j)$. Thus, we should have

$$q_{ij,il} = C(\pi_i r_{il}^i + \pi_l r_{il}^l), \qquad q_{il,ij} = C(\pi_i r_{ij}^i + \pi_j r_{ij}^j), \text{ etc} \qquad (14)$$

for some constant C with the dimension of a flux: here we will simply assume that $C = J$.

TPT Analysis of Markov Jump Processes. In addition to using the framework of TPT to justify the choice of milestone construction, we also use TPT to analyze the resulting Markov jump process spanning the milestones.[40,41] For example, we calculate the rates between predefined macrostates constructed by lumping milestones together. Here we describe briefly the version of TPT for Markov jump processes.[41]

Let $S = i$, $i = 1, ..., N$, be the set of states of the orginal jump process with rate matrix $K = (k_{ij})$, where k_{ij} is the rate from i to j. Assume we have defined the macrostates A_α with $\alpha = 1, ..., n$, each of which is certain set of the microstate i. TPT permits the calculation of the rate between any pair of macrostates. To do that, we choose certain A_α as reactant state, and all other macrostates, A_β, $\beta \neq \alpha$, as the product state. TPT gives the statistical properties of reactive trajectories for the reaction from reactant A_α to the product state. The essential quantities to compute are the forward committor function, q_i^+, the probability that the process starting at i will reach the product state before reach A_α, and the backward committor function q_i^-, defined as the probability to last come from A_α rather than the product state arriving at i. When the process is time-reversible (i.e., detailed balance is satisfied), $q_i^+ = 1 - q_i^-$. The transition rate from A_α to A_β per unit time is given by

$$F_{A_\alpha, A_\beta} = \sum_{i \in S, j \in A_\beta, i \neq j} \pi_i k_{ij} (1 - q_i^+) \qquad (15)$$

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

$$k_{A_\alpha, A_\beta} = \frac{F_{A_\alpha, A_\beta}}{\pi_{A_\alpha}} \qquad (16)$$

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

$$k_{A_\alpha, \text{product}} = \sum_{\beta \neq \alpha} k_{A_\alpha, A_\beta} \qquad (17)$$

and the fraction of transition from A_α to A_β is simply

$$f_{A_\alpha, A_\beta} = \frac{k_{A_\alpha, A_\beta}}{k_{A_\alpha, \text{product}}} \qquad (18)$$

■ IMPLEMENTATION

Simulation Details. The initial structure of Mb with CO is obtained from the PDB (2MB5).[42] The protein is solvated with TIP3P[43] water and neutralized with chlorides. All MD simulations were carried out using NAMD 2.9[44] which we modified to implement temperature-accelerated molecular dynamics (TAMD)[45,46] and cell-confined milestoning simulations. Protein translation is suppressed by restraining the center of mass of the backbone α-carbons with a spring constant of 100 kcal/mol Å^2. Protein rotation is suppressed by harmonically restrained the quaternion of the constellation of backbone α-carbons to $(1,0,0,0)$ during the MD simulation with a spring constant of 1000 kcal/mol. The effectiveness of this strategy to fix the rotation is verified by comparison of raw

RMSD from the restrained system to initial-frame-aligned RMSD from an unrestrained MD simulation. All simulations use the CHARMM22 force field.[47] Periodic boundary conditions are used and long-range electrostatics are computed using particle-mesh Ewald summation with grid spacing of 1 Å. The van der Waals interactions are cut off beyond a distance of 12 Å, and the covalent bonds in water and involving H are kept rigid via the SHAKE/RATTLE algorithm.[48,49] A Langevin thermostat[50] is used to control the temperature at 310 K with a friction coefficient of 5 ps^{-1}, and the integration time step is 2 fs. The system is equilibrated for 5 ns at constant pressure ($P = 1.01$ bar) and temperature ($T = 310$ K) by using a Langevin Nosé−Hoover piston[51] with a coupling constant of 5 ps^{-1} for the thermostat and 10 ps^{-1} for the barostat.

TAMD/Single-Sweep. The collective variables for TAMD/ single-sweep are the Cartesian coordinates of the center of mass of the CO molecule. Thirty independent TAMD trajectories of 5 ns each, starting with CO in the DP site, are used to explore the ligand-accessible regions inside Mb and all escape pathways. TAMD was run with coupling constant $\kappa = 200$ kcal/mol/Å^2, fictitious friction $\bar{\gamma} = 250$ ps^{-1}, and fictitious temperature $k_B\bar{T} = 5.0$ kcal/mol. From the aggregate TAMD sweeps, we harvest a total of 707 samples ("centers") that constitute an irregular mesh of CO locations with a discretization of 2 Å. This is a finer mesh than was used previously,[28] required to adequately sample the portal regions for later milestoning MD.

The mean force is calculated at each mesh point from 4 ns of MD simulation in which the CV's are restrained at that point using a coupling constant of 200 kcal/mol/Å^2. With the mean forces computed at each mesh point, we reconstruct the free energy surface (FES) as an expansion in Gaussian radial basis functions (RBF) following a least-squares fitting procedure.[52] The optimal σ minimizing the fitting error is 3.2 Å. The MFEP's connecting local minima are computed from the zero-temperature string method[53,54] on the resulting 3D FES.

String Method in 6D Collective-Variable Space for DP-HG Escape. To study the role of His64 gating in the mechanism of ligand escape through the HG portal, we include the center of mass (COM) coordinates of the His64 imidazole ring in the CV set, resulting in a 6D CV space. We do not perform single-sweep on this space to ultimately arrive at an MFEP; instead, we perform a string method calculation directly.[55] The initial string is composed of 16 images which are taken from TAMD, one piece of trajectory in which CO successfully escapes from DP to solvent through HG. For each iteration, the mean force on CVs are calculated from 200 ps restrained simulation with coupling constant 200 kcal/mol/Å^2. The string converged within 400 iterations, as shown in Figure S3 in the SI. We get a final smoother string from the averaging of the last 120 iterated strings, and for each component of the final 6D string, we perform a polynomial fit to obtain a continuous MFEP. Figure S4 shows the fitting results.

Milestoning. Since optimal milestones are the isocommitor surfaces, which can be approximated by hyperplanes locally orthogonal to an MFEP, we choose the Voronoi centers z_i as discrete points along the MFEP's provided by string method on the FES. We require any two neighboring centers to be 0.6−1.3 Å away from each other. We also verify that the Markovian assumption is valid by calculating the velocity autocorrelation function and the first passage time distribution (see Supporting Information). The initial configuration for the MD simulation in each cell is taken from TAMD trajectories as the snapshot that has the COM of CO closest to that cell's center. The

system is first equilibrated for 20 ps with the COM of CO fixed at the center and after that, confined MD is run for 4 ns to collect the statistics in that cell. For the pathway from DP-HG, we also performed milestoning using as CVs both the COM of CO and the COM of imidazole ring on His64. The Voronoi centers are chosen along the resulting string (MFEP) from the 6D string method (see above), with a distance from 0.36 to 0.57 Å. The initial configuration for simulation in each cell is inherited from the string method simulations.

We extend Markovian milestoning simulation in the portal regions (between the portal milestone and a sphere boundary) to obtain escape rates. Because the FES is very broad with only shallow gradients in the portal regions and there is no a single narrow transition pathway carrying most of probability flux associated with ligand transit from solvent to the Mb interior, we cannot do optimal milestoning that puts Voronoi centers along a MFEP. Rather, we use Voronoi centers uniformly distributed outside the portal. Since TAMD/single-sweep deposits enough centers even in the portal regions, they are used to generate the Voronoi tessellation. Note, these centers are 2 Å away from each other. This large distance results in large Voronoi cells, but it does not hurt the milestoning efficiency because the ligand moves more freely in these cells, and there are still sufficient transition attempts from milestone to milestone. A sketch of the portal region with Voronoi tessellation is shown in Figure 1. The simulation time in these cells is 4 ns. Thus, the total number of Voronoi cells used is 251, including those at portal region, and the total simulation time for milestoning MD is around 1 μs. All simulations are done using a modified version of NAMD 2.9 containing a collective-variables module that defines a broadcast that "rewinds" positions and "rewinds-reverts" velocities by one time-step upon detection of a Voronoi cell violation by the CV.

■ RESULTS AND DISCUSSION

CO Sites and Migration/Escape Pathways Revisited. Following Maragliano et al.,[28] we first used the method of single-sweep free-energy reconstruction[52] to compute the FES of the CO COM in a protein-centered coordinate system. The present work expands upon the findings of that previous work in that our reconstruction used a larger and more finely resolved irregular mesh (see Implementation section). Representative isosurfaces of FES are shown in Figure 2. Local minima in this FES correspond clearly to previously known localization sites for CO in Mb, including the canonical Xe sites (Xe1 to Xe4),[2] the phantom site,[19] and a final site we termed "NMR", to correspond to the same location found previously using NMR scanning.[1] Relative to the xenon sites, which are observed more often in experiments, the phantom and NMR sites are relatively less important due to their lower stabilities. Relatively large transition rates out of these sites (as will be shown in the next section) support this finding.

We identify MFEP's interconnecting these sites with each other and with the solvent phase using direct zero-temperature string method calculations on this FES.[53,54] We postulate that the MFEP's represent the major CO diffusion channels in Mb, and we indicate these pathways in Figure 2. The pathway network mapped out in our study is consistent with the previous single-sweep results[28] and remarkably similar in structure to the pathways observed in brute-force MD.[24]

Kinetics of One CO Molecule in Mb. To illustrate the raw results one obtains from milestoning simulations, we visualize the milestones by rendering the hitting points on each
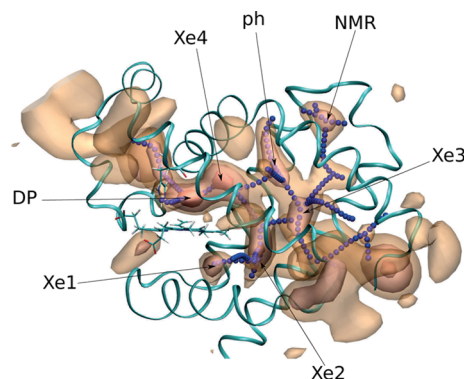

Figure 2. Overlay of three selected isosurfaces (−12.0, −8.0, and −5.4 kcal/mol) of the 3D FES of the CO COM inside Mb. Discretized MFEP's computed by direct string method on the full FES are represented by blue balls. Free energies for the cavities are (in kcal/mol): DP, −15.3; Xe4, −14.2; Xe2, −13.1; Xe1, −11.6; ph, −10.9; Xe3, −10.8; NMR, −9.8.

milestone for each cell simulation in Figure 3. The resulting Markovian jump process contains 759 milestone states with a
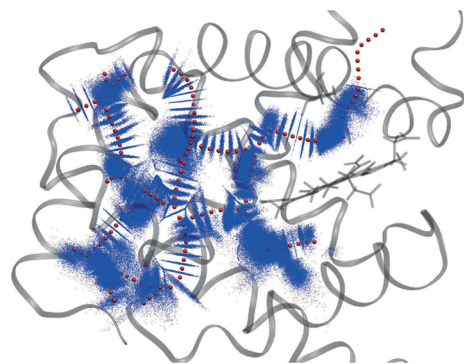

Figure 3. Hitting points (blue) on cell boundaries from all milestoning MD simulations. Discretized MFEP's computed by direct string method on the full FES are presented by red balls. A full 360° 3D view of the hitting points is available in Supporting Information.

rate matrix **Q** of 759 × 759 (a text file containing **Q** is attached in the Supporting Information). This rate matrix includes all the relevant kinetic information about CO migration in Mb.

To have a more accessible view of the kinetics, we construct a "site-specific" kinetic network using as nodes those metastable states identified on the FES. Each state in this network is a conglomeration of several states from the set of 759 milestone states: e.g., the Xe4 state is composed of those milestones states at the Xe4 site; the dissociated state (CO in solvent) is composed of solvent milestone states. The rates between node states on the network are calculated from TPT analysis (see Methods). In Figure 4, we have assigned these forward and backward MFPT's for the important site-to-site diffusion channel and constructed the kinetic network for modeling escape of one CO from the protein interior. Generally, our calculated rates for the main site-to-site transitions range from several to hundreds of ns. They are at the same level as rough estimates of transit rates based on unrestricted MD simulations of Ruscio et al.,[24] though ours have a higher degree of statistical certainty.

Based on the rates, the site-specific network can be partitioned into two parts connected by two channels: one
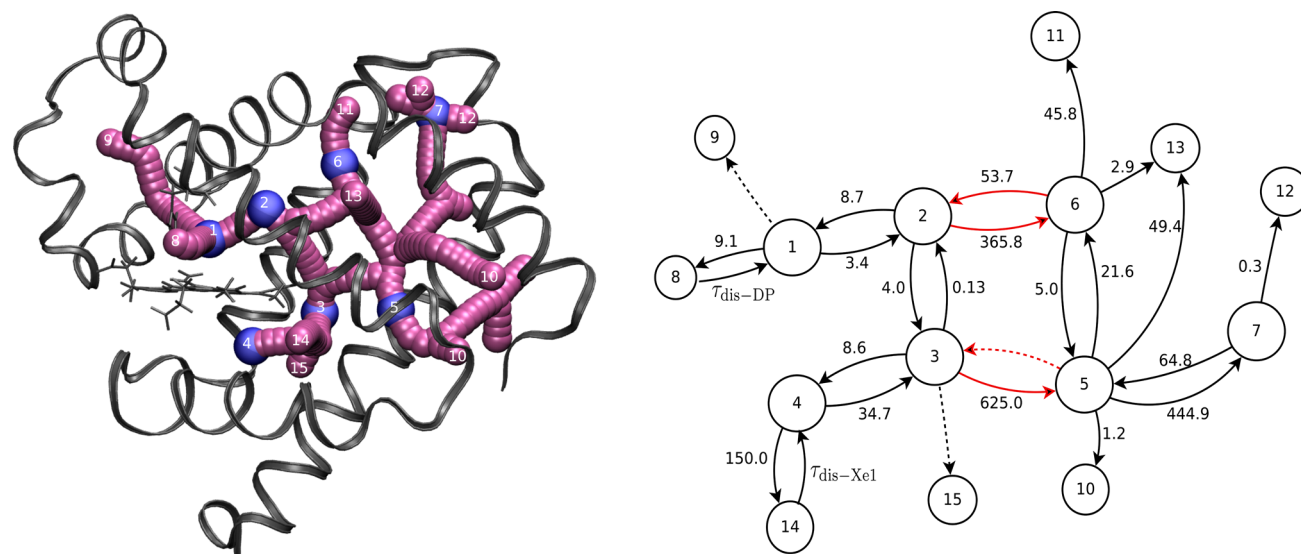
Figure 4. Left panel: Pathways of ligand diffusion in Mb (Magenta). Blue spheres indicate metastable ligand localization sites. Cavity and portal sites are indicated by integer labels which correspond to states of the kinetic networks in the right panel. The sites, PHE33s (9), Xe3s (10), Phs (11), NMRs (12), MIDs (13), Xe2s (15), HG (8), and Xe1s (14) are portal sites, which may be buried in or on the surface of protein, but has channels to solvent. The sites DP (1), Xe4 (2), Xe2 (3), Xe1 (4), Xe3 (5), ph (6), and NMR (7) are cavities. Right panel: the site-specific kinetic network modeling escape of one CO from Mb. Nodes are numbered as "states" with ligand at the site corresponding to the same number in the left-hand figure. The MFPT's are shown on each edge in ns. Only the MFPT's on edge 1−8 or 4−14 are between the state (with CO in DP or Xe1) and the dissociated state (with CO in solvent). Dashed arrows indicate that the MFPT is above 2000 ns. The MFPT's $\tau_{\text{dis-DP}}$ and $\tau_{\text{dis-Xe1}}$ depend on the concentration of CO in solution (see Table 1 for the values at several different concentrations).

between Xe4 and Ph and the other between Xe2 and Xe3; the two channels are indicated in red in Figure 4. Having low rates, these two channels are bottlenecks in the ligand migration process between the two halves of this network. These bottlenecks were already qualitatively apparent in the previous MD work.[24] We report here quantitatively how small the rates are through these bottlenecks. The two bottlenecks partially explain why one observes from recent time-resolved X-ray crystallography that the electron density of CO flows only to the DP, Xe4, and Xe1 cavities after photolysis.[10,11,13]

The main discrepancy between simulation and experiment is about how many escape pathways exist for ligand diffusion in Mb.[7,31] In all the simulation studies, multiple escape pathways have been found; however, most experiments support the HG hypothesis: the main escape pathway for ligand is through histidine gate (E7). The geminate recombinations on mutated MbO₂ predict HG is the dominant pathway carrying over >70% of the escaping flux.[5−8] The time-resolved crystallographic studies[10,15] and IR spectroscopy analysis[12] on wild and mutated MbCO also show that the escape from Xe1 to solvent should be very slow.

The question of what pathways ligands use to escape Mb after photolysis contains a two-fold meaning. First, how many escape pathways exist? Second, if the ligand starts from the DP site after photolysis, which pathway is it most likely to take? For the first question, as in the other simulation studies, our TAMD enhanced-sampling runs identify three main pathways: DP-HG-solvent; Xe1-Xe1s-solvent; and ph/Xe3-solvent. Exits via a few other pathways were observed very rarely in TAMD. Now, because our milestoning procedure provides all the rates, we can answer the second question. Using TPT analysis (see Methods), we obtained the rate from DP to solvent through HG and the rate from DP to the portals other than HG: the former is $(34.9 \text{ ns})^{-1}$ and the latter is $(394.4 \text{ ns})^{-1}$. That is, 92% of all CO's leaving DP and entering the solvent phase do so

through the HG. Our simulation results therefore seem to strongly support the HG hypothesis. Note, the direct rate from DP to solvent through HG is only $(9.1 \text{ ns})^{-1}$, and this is not an accurate estimate of the ligand escape rate from DP. The true rate requires accounting for the rates associated with CO rapidly equilibrating between DP and Xe4. This is a point that might be missed by the simulation community.

Because the site-specific kinetic network in Figure 4 shows that Xe2 and DP relax quickly to Xe4, all milestones at these sites and on the diffusion channels interconnecting them can be merged into one state, which we call the "primary" state. Considering the longer relax time from Xe1 to Xe4, we assign all milestones at Xe1 as "secondary". By the same reasoning, ph and Xe3 can be combined into a "tertiary" state. Using TPT analysis on the full Markov-state model (containing 759 milestone states), we calculated the rates among the primary, secondary, and tertiary states and obtained a more simplified network shown in Figure 5. Due to the slow rates interconnecting the primary, secondary, tertiary states, the migration of CO after photolysis probably only involves the liganded state, the primary state, and the dissociated state (CO in solvent). This finding not only supports the three state model used in geminate recombination experiments[3,5,6] but further gives an interpretation of what sites each of the three states could correspond to. Fitted to the three state model (side-path mechanism), the geminate recombination experiments predict the escape rate from primary to solvent for CO is about 189 ns⁻¹ [3] and for O₂ is about 159 ns⁻¹.[5,6] However, our escape rate from primary to the dissociated state is only 28.0 ns⁻¹, 6-fold faster. This corresponds to 1.14 kcal/mol too low a free-energy barrier. Considering the error bar of free energy calculation in simulation is around 1 kcal/mol typically, this is a reasonable degree of quantitative agreement.

We also see that the escape rate from the secondary to dissociated state is $(139.1 \text{ ns})^{-1}$, which is much slower than the
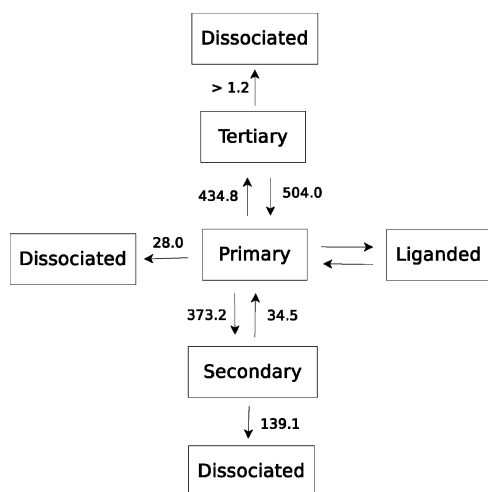
Figure 5. A coarse-grained kinetic network with MFPT's in ns indicated on the arrows.

rate from the primary to dissociated state. A simple investigation of the Mb structure gives an stereochemical explanation for the slow rate. First, residues Leu89 and Ile75 form a narrow pathway for ligand to escape to the portal site Xe1s. But Phe138 could play a more important role since it partially blocks the entrance of the pathway.

The special role played by Xe4 site found in our study has been noted by Olson et al.:[8] based on germinate recombination on mutated Mb, they suggested that Xe4 and DP should be in a single well (what we here identify as the primary state). Recent crystallographic studies[10,13] provide another endorsement of the role played by Xe4. They revealed that the transition between DP and Xe4 is fast after photolysis in several types of mutated Mb (this phenomenon is not observed in the wild-type Mb). Even though the dynamical behavior of Mb-CO system may be different under crystallographic vs physiological conditions, there is no doubt that the Xe4 site plays a special role indeed.

In 2010, Cho et al. studied MbCO dynamics in solution using time-resolved X-ray scattering.[56] They fit their raw data to a five-state sequential model, and predicted that the MbCO after photolysis goes through several intermediate states A′, A, and B (named by us for convenience), and then from state B, CO escapes into solvent or rebinds to the heme. The escape rate is predicted to be 220 ns^{-1} consistent with early geminate recombination studies using three states model. The transition rate from A′ to A is <100 ps^{-1}, from A to B is 2.7 ns^{-1}. These states are not identified by any other information. Based on our calculations, we suggest that A is the state with CO in DP, while B is the primary state in Figure 5, that is, with CO mainly in Xe4 and only a small fraction of the time in DP. Interestingly, our theoretical rate from DP to Xe4, 3.4 ns^{-1}, is very close to the experimental A−B transition rate (2.7 ns)$^{-1}$. More than that, our results suggest that CO should go through HG from B to solvent. However, in 2014, Oang et al. used the same experimental data to fit 18 candidate models including the five state model utilized by Cho.[57] They found the optimal one (with smallest fitting error) is one in which CO migrates from DP to Xe4, then to Xe1 and finally escapes into solvent from Xe1. Our results, like essentially all of the previous experimental work, do not support this mechanism.

Role of Gating in DP-HG Escape. Movement of the His64 side-chain at the gate is correlated to the ligand entry to and

escape from Mb.[29,58−63] From our TAMD trajectories, we noticed that His64 is fully open when CO takes the DP-HG pathway into solvent, which does not allow us to conclude whether or not the rotation of His64 side chain is potentially the rate-determining step of the escaping process. With this observation in mind, we peformed a more detailed, 6D string method calculation for movement of CO along the DP-HG pathway with motion of the His64 side-chain. We included the COM Cartesian coordinates of the imidazole ring of His64 with the COM Cartesian coordinates of the ligand in the CV set. The MFEP in this 6D space is shown in Figure 6, projected
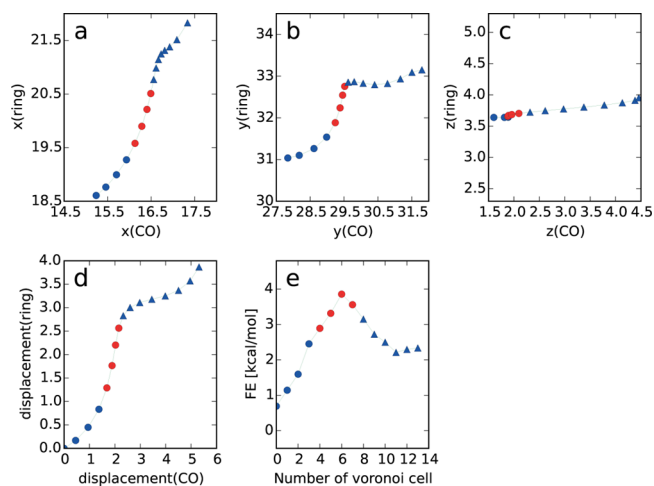

Figure 6. 6D MFEP for the DP-HG migration projected into 2D subspaces of CO COM and COM of the imidazole ring: (a) x of ring COM vs x of CO COM; (b) y of ring COM vs y of CO COM; (c) z of ring COM vs z of CO COM. (d) Projection of the 6D MFEP into the space of displacement $((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2)^{1/2}$ of CO COM and of ring COM. (e) Free-energy profile along the MFEP. Circles denote stage 1 and triangles stage 2 (see text). The first image (index 0) corresponds to DP milestone and the last one (index 13) corresponds to the HG milestone.

onto various 2D subspaces. Interestingly, the mechanism of ligand escape through HG is composed of two stages. In the first stage, CO approaches the gate, and during its approach to the gate, the gate swings outward a very small amount. This is immediately followed by a swift opening of the gate as the imidiazole side chain swings out into solution while CO remains stationary (red circles in Figure 6). This process can clearly be seen from the displacements of COM's of the CO and the imidazole ring (Figure 6 e): an initial simultaneous displacement of CO COM and ring COM (blue circles), followed displacement of the ring during a stand-still of the CO (red circles). In the second stage, CO moves all the way outside of the HG portal while the gate adjusts itself a little and remains open. This can be seen from the displacement plot as well: the displacement of CO COM increases from 2.5 to 5.5 Å, while the displacement of ring COM only increases by <1 Å (up-triangles).

Note that the bottleneck in the process (where the free-energy is maximal) occurs when CO approaches His64 and seemingly pushes it open in the first stage. The picture we have from this 6D MFEP is a knock-on mechanism where the gate-opening is made more facile by the presence of CO, rather than a lucky coincidence that CO attempts to transit when the gate has fluctuated open spontaneously.

We performed Markovian milestoning in this 6D space as well. We find that the DP-solvent MFPT calculated in this way is 9.1 ns, which is only a little larger than the MFPT computed in the 3D space (5.3 ns). Note that in the kinetic network of Figure 4, we used 9.1 ns as the DP-to-HG MFPT.

Entry Kinetics. As recognized by Elber,[31] most all-atom simulations agree that there exist multiple solvent-accessible pathways to the DP, and experiments agree that the HG is the dominant entry and escape portal (>70%).[5−8] This is not a direct contradiction; in fact, determination of the dominant pathway in simulations should rest on the same observables on which the experimental findings rest, namely, entry kinetics.

As detailed in the Methods section, we grafted a simple diffusion-equation-based model onto the milestoning framework to account for concentration-dependent entry kinetics. We consider rates of entry through direct pathways from solvent portals to both DP (through HG) and Xe1 (through Xe1s). The reason we focus on these two is that CO most likely enters the DP site through these two pathways, solvent-DP and solvent-Xe1-Xe2-Xe4-DP, since no substantial barrier exists along these pathways compared to other pathways, e.g., solvent-Xe3-Xe4-DP. Furthermore, DP and Xe1 are the most relevant sites for the ligand escape/entry as indicated in several previous studies.[8,24,64,65] The rates from the dissociated state to DP and Xe1 are calculated from the TPT analysis on the Markovian jump process (see Methods section). Values of pseudo-first-order rate constants k_{dis-DP} and $k_{dis-Xe1}$ are presented in Table 1

Table 1. Entry Rates at Different CO Concentrations (in μM)[a]

[CO]	solvent-DP pathway		solvent-Xe1 pathway	
	τ_{dis-DP} (= k_{dis-DP}^{-1})	k_{entry}	$\tau_{dis-Xe1}$ (= $k_{dis-Xe1}^{-1}$)	k_{entry}
985.4 × 1	56.2	17.8	513.8	2.0
× 5	11.3	17.7	103.0	2.0
× 10	5.6	18.1	51.4	2.0
× 20	2.8	17.9	25.9	2.0
× 50	1.1	18.2	10.3	1.9

[a]The diffusion constant (CO in water) $D = 2.03 \times 10^{-5}$ cm^2/s = 203 Å^2/ns at 1 atm and 298 K.[66] 985.4 μM = 27.4 mg/L is the solubility of CO in water at a partial pressure of 1 atm and 298 K. The pseudo-first-order rate constant $k_{dis-DP(Xe1)}$ is in unit of us^{-1} and the MFPT $\tau_{dis-DP(Xe1)}$ is in unit of us. The second-order rate constant k_{entry} has units of μM^{-1} s^{-1} ($k_{dis-DP(Xe1)} = k_{entry}[CO]$).

as functions of bulk [CO]. Note, since we neglect the ligand−ligand interaction, the bulk concentrations considered in must be low enough so that the solution is dilute; as CO is only sparingly soluble, this is reasonable.

We note that the ratio of the two rate constants in Table 1, k_{dis-DP}:$k_{dis-Xe1}$, is 9:1, for all ligand concentrations. This means that, compared to entry through Xe1s, 90% of CO entry events occur through the HG portal, independent of concentration, which supports experimental observations.[5,6,31] Under the same conditions (1 atm and 298 K, [CO] = 985.4 μM), the theoretical prediction of the second order rate constant is k_{entry} = 18.0 μM^{-1} s^{-1}, which falls well inside the range reported experimental entry rates of ligands: 81 μM^{-1} s^{-1} or 17 μM^{-1} s^{-1} for CO,[3,67] and 34 μM^{-1} s^{-1} for O$_2$.[5]

■ CONCLUSION

We have presented a new framework for using all-atom MD simulations to compute a full kinetic picture of CO entry into, exit from, and diffusion in Mb. The framework overcomes the time-scale limitations of practical MD through an interpretation of the free-energy surface for ligand localization in the protein using transition-path theory, which provides a basis for efficient parallelization through Markovian milestoning MD. In addition, we employ a simple model to handshake entry kinetics to the internal and exit kinetics. The result is a full kinetic network accounting for the dissociated state and all noncovalent associated states of CO and Mb. This network predicts exit rates that are in near-quantitative agreement with kinetics extracted from geminate recombination experiments while providing a new picture of the roles played by the various crystallographically determined interal cavities. Importantly, because the method allows independent calculation of exit rates through every plausible portal, we are able to conclude that the gate at His64 is likely the portal through which more than 90% of all exit events occur. This result broadly supports experimental findings and directly addresses the apparent disagreement between the experimental and simulation communities regarding dominant diffusion pathways into and out of Mb. We further show with detailed string method calculations that gating in a minimal free-energy sense appears to be a "knock-on" process in which the ligand position and gate state are interdependent. Our kinetic treatment also predicts that of the two major plausible entry pathways that can deliver ligand from solvent to the distal pocket, again the histidine gate is rate dominant by 90%.

We expect this methodology to be useful for other theoretical kinetic studies of ligand diffusion for which MD alone is not suitable due to sampling limitations. It relies on simple parallelization into multiple independent MD runs, each assigned to a cell in a tessellation such that the cell faces are optimal milestones. Importantly, we have shown that one way to approximate these milestones is to choose them locally orthogonal to minimum free-energy pathways, which are computable using, for instance, the string method in collective variables, which is even not restricted to 3D or even 6D collective-variable spaces. A particularly enticing example for a potential future milestoning study is the mini-hemoglobin CerHb, whose dominant ligand exit pathway was recently shown by careful mutagenesis to be 180° in the opposite direction from the histidine gate, through an apolar tunnel.[7] The surprising difference in exit pathways between CerHb and Mb could potentially be explained in atomic detail using our method.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information
Figure illustrating string converged within 400 iterations. An animated view of milestone hitting points, the full rate matrix in a human-readable, ZIP'ed data file, and four supplementary data figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*cfa22@drexel.edu

Present Address
§(M.L.) Genentech Inc. (member of Roche), 1 DNA way, South San Francisco, CA 94080.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Tilton, R. F.; Kuntz, I. D. *Biochemistry* **1982**, *21*, 6850−6857.
(2) Tilton, R. F.; Kuntz, I. D.; Petsko, G. A. *Biochemistry* **1984**, *23*, 2849−2857.
(3) Henry, E. R.; Sommer, J. H.; Hofrichter, J.; Eaton, W. A. *J. Mol. Biol.* **1983**, *166*, 443−451.
(4) Srajer, V.; Teng, T. Y.; Ursby, T.; Pradervand, C.; Ren, Z.; Adachi, S.; Schildkamp, W.; Bourgeois, D.; Wulff, M.; Moffat, K. *Science* **1996**, *274*, 1726−1729.
(5) Scott, E. E.; Gibson, Q. H. *Biochemistry* **1997**, *36*, 11909−11917.
(6) Scott, E. E.; Gibson, Q. H.; Olson, J. S. *J. Biol. Chem.* **2001**, *276*, 5177−5188.
(7) Salter, M. D.; Blouin, G. C.; Soman, J.; Singleton, E. W.; Dewilde, S.; Moens, L.; Pesce, A.; Nardini, M.; Bolognesi, M.; Olson, J. S. *J. Biol. Chem.* **2012**, *287*, 33163−33178.
(8) Olson, J. S.; Soman, J.; Phillips, G. N. *IUBMB Life* **2007**, *59*, 552−562.
(9) Ostermann, A.; Waschipky, R.; Parak, F. G.; Nienhaus, G. U. *Nature* **2000**, *404*, 205−208.
(10) Srajer, V.; Ren, Z.; Teng, T. Y.; Schmidt, M.; Ursby, T.; Bourgeois, D.; Pradervand, C.; Schildkamp, W.; Wulff, M.; Moffat, K. *Biochemistry* **2001**, *40*, 13802−13815.
(11) Schotte, F.; Lim, M. H.; Jackson, T. A.; Smirnov, A. V.; Soman, J.; Olson, J. S.; Phillips, G. N.; Wulff, M.; Anfinrud, P. A. *Science* **2003**, *300*, 1944−1947.
(12) Nienhaus, K.; Deng, P.; Kriegl, J. M.; Nienhaus, G. U. *Biochemistry* **2003**, *42*, 9633−9646.
(13) Schotte, F.; Soman, J.; Olson, J. S.; Wulff, M.; Anfinrud, P. A. *J. Struct. Biol.* **2004**, *147*, 235−246.
(14) Hummer, G.; Schotte, F.; Anfinrud, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15330−15334.
(15) Schmidt, M.; Nienhaus, K.; Pahl, R.; Krasselt, A.; Anderson, S.; Parak, F.; Nienhaus, G. U.; Srajer, V. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11704−11709.
(16) Bourgeois, D.; Vallone, B.; Arcovito, A.; Sciara, G.; Schotte, F.; Anfinrud, P. A.; Brunori, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4924−4929.
(17) Tomita, A.; Sato, T.; Ichiyanagi, K.; Nozawa, S.; Ichikawa, H.; Chollet, M.; Kawai, F.; Park, S.-Y.; Tsuduki, T.; Yamato, T.; Koshihara, S.-Y.; Adachi, S.-I. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 2612−2616.
(18) Elber, R.; Karplus, M. *J. Am. Chem. Soc.* **1990**, *112*, 9161−9175.
(19) Bossa, C.; Anselmi, M.; Roccatano, D.; Amadei, A.; Vallone, B.; Brunori, M.; Di Nola, A. *Biophys. J.* **2004**, *86*, 3855−3862.
(20) Bossa, C.; Amadei, A.; Daidone, I.; Anselmi, M.; Vallone, B.; Brunori, M.; Di Nola, A. *Biophys. J.* **2005**, *89*, 465−474.
(21) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. *Biophys. J.* **2006**, *91*, 1844−1857.
(22) Elber, R.; Gibson, Q. H. *J. Phys. Chem. B* **2008**, *112*, 6147−6154.
(23) Anselmi, M.; Nola, A. D.; Amadei, A. *Biophys. J.* **2008**, *94*, 4277−4281.
(24) Ruscio, J. Z.; Kumar, D.; Shukla, M.; Prisant, M. G.; Murali, T. M.; Onufriev, A. V. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9204−9209.
(25) Ceccarelli, M.; Anedda, R.; Casu, M.; Ruggerone, P. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 1231−1236.
(26) Nishihara, Y.; Hayashi, S.; Kato, S. *Chem. Phys. Lett.* **2008**, *464*, 220−225.
(27) Scorciapino, M. A.; Robertazzi, A.; Casu, M.; Ruggerone, P.; Ceccarelli, M. *J. Am. Chem. Soc.* **2009**, *131*, 11825−11832.
(28) Maragliano, L.; Cottone, G.; Ciccotti, G.; Vanden-Eijnden, E. *J. Am. Chem. Soc.* **2010**, *132*, 1010−1017.
(29) Plattner, N.; Meuwly, M. *Biophys. J.* **2012**, *102*, 333−341.
(30) Lapelosa, M.; Abrams, C. F. *J. Chem. Theory Comput.* **2013**, *9*, 1265−1271.
(31) Elber, R. *Curr. Opin. Struct. Biol.* **2010**, *20*, 162−167.
(32) Mishra, S.; Meuwly, M. *Biophys. J.* **2010**, *99*, 3969−3978.
(33) Cazade, P.-a.; Berezovska, G.; Meuwly, M. *Biochim. Biophys. Acta, Gen. Subj.* **2014**, 2−11.
(34) Vanden-Eijnden, E.; Venturoli, M. *J. Chem. Phys.* **2009**, *130*, 194101.
(35) Faradjian, A. K.; Elber, R. *J. Chem. Phys.* **2004**, *120*, 10880−10889.
(36) Elber, R. *Biophys. J.* **2007**, *92*, L85−L87.
(37) Vanden-Eijnden, E.; Venturoli, M.; Ciccotti, G.; Elber, R. *J. Chem. Phys.* **2008**, *129*, 174102.
(38) Smoluchowski, M. V. *Z. Phys. Chem.* **1917**, *92*, 129−168.
(39) Shoup, D.; Szabo, A. *Biophys. J.* **1982**, *40*, 33−39.
(40) E, W.; Vanden-Eijnden, E. *J. Stat. Phys.* **2006**, *123*, 503−523.
(41) Metzner, P.; Schuette, C.; Vanden-Eijnden, E. *J. Chem. Phys.* **2006**, *125*, 84110.
(42) Cheng, X.; Schoenborn, B. *Acta Crystallogr., Sect. B* **1990**, *46*, 195−208.
(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(44) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(45) Maragliano, L.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2006**, *426*, 168−175.
(46) Abrams, C. F.; Vanden-Eijnden, E. *Proc. Nat. Acad. Sci. U.S.A.* **2010**, *107*, 4961−4966.
(47) Foloppe, N.; MacKerell, J. *J. Comput. Chem.* **2000**, *21*, 86−104.
(48) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327−341.
(49) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24−34.
(50) Brünger, A.; B, C. L., III; Karplus, M. *Chem. Phys. Lett.* **1984**, *105*, 495−500.
(51) Quigley, D.; Probert, M. I. J. *J. Chem. Phys.* **2004**, *120*, 11432−11441.
(52) Maragliano, L.; Vanden-Eijnden, E. *J. Chem. Phys.* **2008**, *128*, 184110.
(53) N, E. W.; Ren, W. Q.; Vanden-Eijnden, E. *Phys. Rev. B* **2002**, *66*, 52301.
(54) N, E. W.; Ren, W. Q.; Vanden-Eijnden, E. *J. Chem. Phys.* **2007**, *126*, 164103.
(55) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 24106.
(56) Cho, H. S.; Dashdorj, N.; Schotte, F.; Graber, T.; Henning, R.; Anfinrud, P. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 7281−7286.
(57) Oang, K. Y. K.; Kim, J. J. J. G.; Yang, C.; Kim, T. W. T.; Kim, Y.; Kim, K. H.; Ihee, H. *J. Phys. Chem. Lett.* **2014**, *5*, 804−808.
(58) Ringe, D.; Petsko, G. A.; Kerr, D. E.; Demontellano, P. R. O. *Biochemistry* **1984**, *23*, 2−4.
(59) Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Rohlfs, R. J.; Mathews, A. J.; Olson, J. S. *J. Biol. Chem.* **1989**, *264*, 3057−3060.
(60) Johnson, K. A.; Olson, J. S.; Phillips, G. N. *J. Mol. Biol.* **1989**, *207*, 459−463.
(61) Bellelli, A.; Blackmore, R. S.; Gibson, Q. H. *J. Biol. Chem.* **1990**, *265*, 13595−13600.
(62) Olson, J. S.; Mathews, A. J.; Rohlfs, R. J.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G.; Tame, J.; Renaud, J. P.; Nagai, K. *Nature* **1988**, *336*, 265−266.
(63) Steinbach, P. J.; et al. *Biochemistry* **1991**, *30*, 3988−4001.
(64) Tetreau, C.; Blouquit, Y.; Novikov, E.; Quiniou, E.; Lavalette, D. *Biophys. J.* **2004**, *86*, 435−447.
(65) Nishihara, Y.; Sakakura, M.; Kimura, Y.; Terazima, M. *J. Am. Chem. Soc.* **2004**, *126*, 11877−11888.
(66) Cussler, E. L. *Diffusion: Mass Transfer in Fluid Systems*, 2nd ed.; Cambridge University Press: New York, 1997; pp 10013−2473.

(67) Goldbeck, R. A.; Bhaskaran, S.; Ortega, C.; Mendoza, J. L.; Olson, J. S.; Soman, J.; Kliger, D. S.; Esquerra, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1254−1259.