

Free-Energy Barriers in MbCO Rebinding

Polina Banushkina and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

Received: April 14, 2005; In Final Form: June 20, 2005

The rebinding of CO to myoglobin (Mb) from locations around the active site is studied using a combination of molecular dynamics and stochastic simulations for native and L29F mutant Mb. The interaction between the dissociated ligand and the protein environment is described by the recently developed fluctuating three-point charge model for the CO molecule. Umbrella sampling along trajectories, previously found to sample the binding site (B) and the Xe4 pocket, is used to construct free-energy profiles for the ligand escape. On the basis of the Smoluchowski equation, the relaxation of different initial population distributions is followed in space and time. For native Mb at room temperature, the calculated rebinding times are in good agreement with experimental values and give an inner barrier of 4.3 kcal/mol between the docking site B (Mb•••CO) and the A state (bound MbCO), compared to an effective barrier, H_{eff} , of 4.5 kcal/mol and barriers into the majority conformation A₁ and the minority conformation A₃ of 2.4 and 4.3 kcal/mol, respectively. In the case of the L29F mutant, the free-energy surface is flatter and the dynamics is much more rapid. As was found in experiment, escape to the Xe4 pocket is facile for L29F whereas, for native Mb, the barriers to this site are larger. At lower temperatures, the rebinding dynamics is delayed by orders of magnitude also due to increased barriers between the docking sites.

I. Introduction

Myoglobin (Mb) plays an important role in transporting small molecules such as O₂, CO, and NO. Mb exists in muscles in high concentrations; its primary role is to store and deliver dioxygen to mitochondria in red muscle cells, but its importance for physiological processes may go way beyond that. An understanding of the reactions of ferrous Mb with these ligands is important because, as the simplest protein capable of reversible ligand binding, it is a prototype for more complex systems.^{1,2} Experimentally, the dynamics of ligand migration has been studied using picosecond and femtosecond time-resolved spectroscopy^{3,4} and Fourier transform infrared spectroscopy (FTIR) combined with temperature derivative spectroscopy (TDS).⁵ Concomitantly, computer simulations have been undertaken at different levels to study the migration of photodissociated ligands in the protein and to locate possible pathways into and out of the heme pocket.^{6–10}

An important aspect in understanding the dynamics of small ligands in the protein interior is to quantify the energetic barriers for rebinding, which can be of different origin (steric, electronic, and dynamical).^{3,8,11} Under favorable conditions, the ligand motion and the associated free-energy barriers can be described along suitably chosen (progression) coordinates. Here, we investigate the energetics and dynamics of CO migration from the heme binding site (see Figure 1) to the Xe4 pocket using molecular dynamics (MD) simulations and umbrella sampling. This region is of particular interest for the early stages following photodissociation of the ligand from the heme. The binding site (B) is of major importance because after photodissociation it is rapidly populated. From there, the ligand can either rebound directly or it can follow a largely unknown path within the protein to diffuse toward the solvent, from where it rebinds at much longer time scales.^{8,9} One possible secondary binding site in the neighborhood of site B is the Xe4 pocket.^{5,9,12,13} For native Mb, experimental evidence for population of the Xe4 pocket

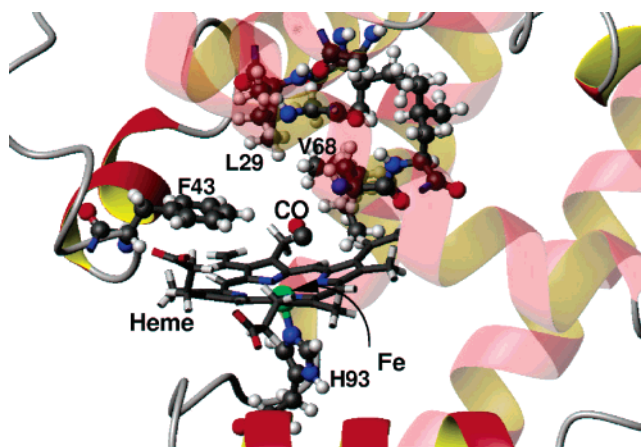


Figure 1. Active site of Mb with the heme unit (sticks), the dissociated CO molecule, and a few nearby residues shown in ball-and-stick representation.

by CO after dissociation appears to be sparse. It is only recently that FTIR-TDS experiments have found migration of photolyzed CO from the primary docking site B to the Xe4 cavity.⁵ This finding is supported by results from a 90 ns MD simulation of native MbCO which show that CO occupies the Xe4 cavity before migrating further in the protein matrix.⁹ Using time-resolved X-ray crystallography, it has been found that the Xe4 pocket in the L29F mutant is populated within about 1 ns after photodissociation.¹² It is still populated 3 ns after dissociation, whereas at 30 ns, all CO molecules have moved out of the Xe4 pocket either back to the docking site or further to cavities such as Xe1. Another mutant for which CO is found in the Xe4 pocket is L29W.¹⁴ It is interesting to note that the population of Xe4 is found in crystals at 105 K after illumination below

180 K, whereas illumination above 180 K leads to the observation of electron density in Xe1, which was attributed to the presence of CO.

Experimentally, the rebinding dynamics of CO after photodissociation has been studied extensively for native MbCO and for various mutants under different conditions^{3,14–16} by following the number of ligand molecules, $N(t)$, that have not re-bound at time t after photodissociation. Usually, rate constants (k_{XY}) are extracted by fitting a set of first-order kinetic equations which describe transitions between stable and metastable states, X and Y, visited by the ligand during the rebinding dynamics.^{2,3,5,14,17} For native and mutant MbCO, three bound substates (A_0 , A_1 , and A_3) and a number of intermediate states after photodissociation have been identified directly and indirectly.^{5,14,18} In many cases, the identity of the intermediate states (B, C, and D) is not explicitly specified,^{3,2,17} but over the past few years, evidence has been accumulated that B is the docking site and C is related to one of the internal Xe pockets.^{12,14,19,20} Recently, a sequential three-well model was convincingly fitted to rebinding data for the L29W mutant.¹⁴ In this case, structural information was also available to correlate the metastable states with docking sites within the protein. Three rate coefficients describe transitions between B and A (k_{BA}), between B and C (k_{BC}), and between C and B (k_{CB}). Here, A is the bound MbCO state. For state B, the CO is in the docking site, and for C, the CO is in the Xe4 pocket. In different Mb mutants, access to the internal Xe pockets is more or less facile or even completely blocked.⁵ Depending upon whether sequential^{14,17} or side-path schemes^{21,22} are used, the bimolecular rate constants will significantly differ. Thus, the model for an underlying reaction scheme to interpret the data is important. A graphical representation of possible ligand pathways for MbCO is sketched in Figure 2 (see, also, ref 23). After photodissociation from the bound state A (with its substates A_0 , A_1 , and A_3), the ligand accesses the docking site B and subsequently migrates via internal cavities (in an as yet unspecified sequence) to either escape into the solvent or geminately rebound from one of the internal cavities. For native MbCO, the analysis of the data shows that the rebinding dynamics is nonexponential and extends over time scales ranging from picoseconds to seconds. At longer time scales, CO is expected to diffuse to sites further away from the heme and to escape to the solvent, from where it rebinds at much slower rates.^{1,2}

Complementary to the existence of secondary binding sites within Mb, it was suggested that heme relaxation may be involved in the observation of the nonexponential rebinding times. This hypothesis was captured in a theoretical model proposed by Agmon and Hopfield.²⁴ They considered two coordinates, a protein configuration coordinate x and the Fe–ligand separation r coordinate, to describe a qualitative two-dimensional potential energy surface. Thus, the protein is treated with a single coordinate. Following this, Srajer and co-workers have discussed an explicit dependence of the interaction potential from an additional iron–porphyrin coordinate.²⁵ Using experimental constraints, model parameters were determined to best reproduce the measured kinetic data. These studies assumed a simplified, low-dimensional potential energy surface to describe the rebinding process. As studies on MbNO have shown, both effects (existence of secondary binding sites and distributed barriers due to internal relaxation) are likely to contribute to the observed rebinding dynamics.¹¹

Here, we use molecular dynamics simulations for bound and dissociated CO interacting with myoglobin using validated interaction potentials to derive free-energy profiles at different

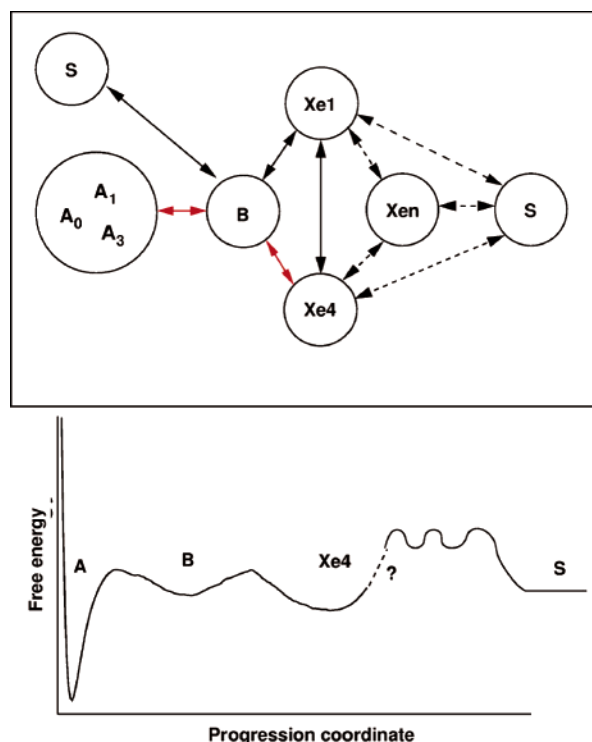


Figure 2. Graphical representation of different CO rebinding paths among the bound state (A), the docking site (B), and various internal cavities (Xe1, Xe4, and Xen (i.e., Xe2 and Xe3)). Solid lines indicate migration pathways that have been observed directly by experiment, the red arrows indicate the path followed in the present work, and dashed lines are further possible pathways; although, the Xe2 and Xe3 pockets are traditionally not considered to be involved in ligand migration. They include the $A \leftrightarrow B \leftrightarrow Xe4$ (for L29W),¹⁴ the $A \leftrightarrow B \leftrightarrow Xe4 \leftrightarrow Xe1 \leftrightarrow B \leftrightarrow A$ (for native Mb),^{5,20} the calculated $A \leftrightarrow B \leftrightarrow Xe4 \leftrightarrow Xe3$ (for native Mb via ghost states) $\leftrightarrow Xe1 \leftrightarrow Xe2 \leftrightarrow Xe1$,⁹ and the $A \leftrightarrow B \leftrightarrow Xe1$ (for native Mb).¹⁶ Scott et al. have proposed a side-path scheme for MbO₂ for which $A \leftrightarrow B \leftrightarrow C$ competes with $A \leftrightarrow B \leftrightarrow S$, and C is not specified.^{21,22} The one-dimensional projection of the free-energy landscape is shown in the lower panel. For smaller values of the progression coordinate, it is the free-energy profile calculated for native MbCO between A, B, and Xe4 (see Figure 4), and for larger values of the progression coordinate, it is a cartoon. The schematic (above) is a two-dimensional projection of the three-dimensional situation in the real protein, and the FEP (below) is a further reduction in complexity.

temperatures. In particular, the recently developed three-point fluctuating charge model for CO which correctly describes the splitting of the CO vibrational band is used to accurately model the electrostatic interaction between unbound CO and the protein environment.²⁶ This allows for a detailed investigation of the interactions and dynamics between CO and the cavity-forming residues around the primary binding site. In a next step, the free-energy profiles for CO migration around the heme pocket in native and mutant (L29F) Mb are used to follow the decay of different initial populations as a function of time. From this, rebinding times for the various free-energy profiles (FEPs) are estimated.

The work is organized as follows: First, the simulation and theoretical methods employed are presented. Next, results for native and mutant (L29F) myoglobin are discussed. Finally, the results are compared with experimental data and conclusions are drawn.

II. Theoretical Methods

A. Setup of the Systems. All molecular dynamics (MD) simulations were carried out with the CHARMM program²⁷ and

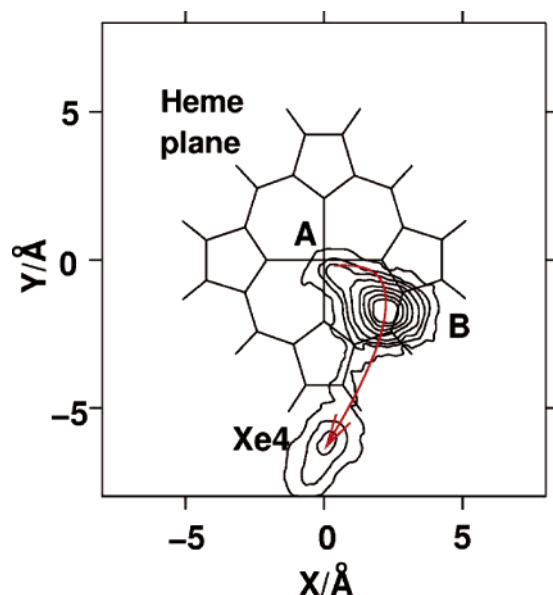


Figure 3. Sample escape path along which FEPs were calculated. The path follows the center of mass of CO from the heme binding site (A) via the docking site (B) to the Xe4 pocket. The contours give the probability (with maximum probability in the B site) from a 1 ns MD simulation²⁶ for the dissociated state *without* application of an umbrella potential.

the CHARMM22 force field.²⁸ Other parameters required to describe the heme–CO interactions were taken from the literature.^{11,29} The systems contained a total of 2532 heme protein atoms (native Mb) and 2533 heme protein atoms (L29F), respectively, the CO ligand, and 181 water molecules, which were represented by a modified TIP3P potential.³⁰ The heme pocket was solvated by a 16-Å sphere of equilibrated water molecules centered around the heme, and a solvent boundary potential was applied to constrain the water molecules. The “reaction region” of radius 12 Å, inside which the system was propagated with Newtonian dynamics, was centered on the heme, and the dynamics of the buffer region between 12 and 16 Å was described using Langevin dynamics. Friction coefficients of 62 and 250 ps^{−1} were applied to the water oxygens and the remaining non-hydrogen atoms, respectively. The FEPs were calculated using the umbrella-sampling method, as implemented in CHARMM³¹ along escape paths (see Figure 3) from 1 ns molecular dynamics simulations previously calculated²⁶ with the same setup as described above.

B. Construction of the Free-Energy Profiles. *Interaction Potentials.* For the bound state (native and mutant (L29F) MbCO), standard molecular mechanics parameters^{28,29} for MbCO were used, with the only difference that Fe and C interact via a Morse potential of the form

$$U(R) = D_e[1 - \exp(-\beta(R - R_{eq}))]^2 \quad (1)$$

Here, $R_{eq} = 1.9$ Å is the equilibrium Fe–C separation,^{18,32} $\beta = 3.2$ Å^{−1} is the steepness of the repulsive wall, and $D_e = 21.4$ kcal/mol is the binding energy.³³ For the equilibrium separation R_{eq} , an average of two different experimentally determined structures (1.93 Å from EXAFS³² and 1.82 Å from X-ray¹⁸) was used, and for the steepness parameter β , the value from a recent study on MbNO was taken.¹¹ EXAFS experiments can give precise bond lengths, although it is known that they may have systematic errors because they do not include anharmonicity. Neutron diffraction experiments give a significantly longer Fe–C distance of 2.12 Å.³⁴ It is noted that the precise

values of these parameters are not critical for the present investigation because we only require that the bound state is sufficiently strongly bound to provide a clearly defined minimum. An energy-minimized structure of bound MbCO was heated to 300 K and equilibrated at that temperature for 1 ns of Langevin dynamics.

Next, umbrella sampling³¹ was carried out in the range $R \in (1.6, 7.0)$ Å, divided into 42 intervals, (δ_i, δ_{i+1}) . MD simulations allow collection of statistics (n_i) for given intervals ($\Delta\delta_i$) along a particular reaction coordinate (δ) by applying a biasing potential of the form $V(\delta) = k_i(\delta - \delta_{0i})^2$. Here, k_i is the force constant and δ_{0i} is the value around which δ should be constrained. Associated free energies, $G(\delta)$, are extracted from n_i via

$$f(i) = -kT \ln n_i + c$$

where k is the Boltzmann constant, T is the temperature, and c is an arbitrary constant. Because each interval of the FEP is only defined up to the constant c , the segments are joined to obtain a continuous FEP over the entire range of the coordinate δ . Here, the constant c is calculated as the value that minimizes the functional $\min[\sum_i (f_2(i) - f_1'(i))^2]$ in the overlap region i of two adjoining windows. $f_1(j_1)$ and $f_2(j_2)$ are sets of points in the first and second interval, respectively, and $f_1'(j_1) = f_1(j_1) - c$. To find suitable parameters (k_i and δ_{0i}) for the biasing potential (see above), a short 5 ps simulation for every interval was run. From this, an approximate form of the FEP was extracted and, if necessary, the parameters k_i and δ_{0i} were modified for better sampling. Next, statistics were collected during 50 ps for each window, (δ_i, δ_{i+1}) , from which the final FEP was derived.

For the unbound (dissociated) state, the connectivity between Fe and C was deleted and replaced by nonbonded interactions including Lennard-Jones parameters and by r -dependent charges on the C and O atoms (the fluctuating three-point charge model) and on another site at the center of mass of CO,^{26,35} where r is the CO bond length. The interatomic CO potential is an anharmonic rotational Rydberg–Klein–Rees (RRKR) model.³⁶ Because we are interested in the FEP from A via B to the Xe4 pocket, umbrella sampling was carried out along CO escape paths found in MD simulations (see above). The reaction coordinate R , which is the Fe–C distance, was varied between 2.4 and 10.0 Å. Once appropriate parameters, k_i and δ_{0i} , were found, the system was equilibrated for 10 ps, after which statistics were collected during 60 ps for each of the 49 windows. Further calculations at 300 K with extended sampling of 250 ps for each window yielded virtually indistinguishable FEPs for both native and L29F MbCO. The FEPs thus constructed are based on more than 10 ns of MD simulation.

Diabatization of the Free-Energy Profiles. Because, for the propagation of an initial population distribution, one continuous free-energy profile is required, the two FEPs have to be diabaticized to eliminate the crossing. Diabatization of two crossing potential energy curves is a physically well-defined procedure.³⁷ In the present case, it requires the two FEPs, their asymptotic separation (which shifts the zero of the FEPs with respect to each other), and a coupling matrix element. The asymptotic separation $\Delta(R \rightarrow \infty)$ between the two FEPs (bound and unbound) is estimated to be around 5 kcal/mol.³⁸ Because this value is approximate and is based on calculations for a model system (heme–CO in the gas phase), Δ is a parameter and changing its value can be used to explore the sensitivity of the results (see below). To diabaticize the surfaces, $G_{\text{bound}}(R)$ and $G_{\text{unbound}}(R)$ are coupled around the crossing point R_c by a Gaussian coupling potential,³⁷ $V_c(R) = c \exp(-(R - R_c)^2/2\sigma^2)$,

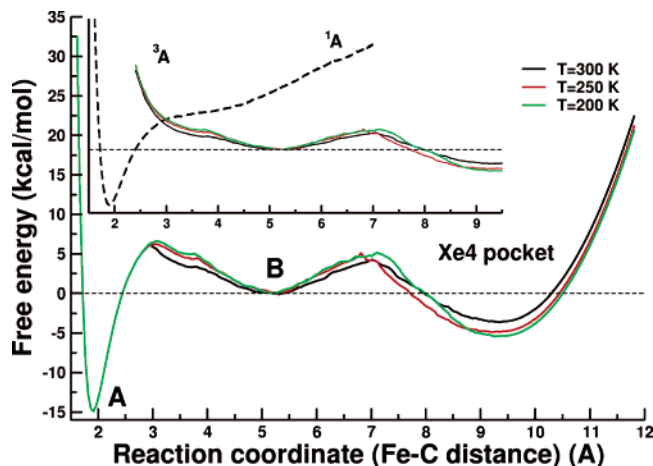


Figure 4. Diabatic (main figure) and adiabatic (inset) FEPs for different temperatures. In the inset, the dashed line corresponds to the bound state FEP and the solid lines are the unbound FEPs for different T . For the 1A state, the FEP becomes repulsive at long ranges because of interactions with nearby protein residues (see Figure 1). The energies of the FEPs for the three different temperatures have been shifted vertically to coincide in site B to allow direct comparison of the barriers.

with strength c and width σ . $G(R)$ is the lower eigenvalue of the matrix

$$M = \begin{pmatrix} G_{\text{bound}}(R) & V_c(R) \\ V_c(R) & G_{\text{unbound}}(R) \end{pmatrix}$$

This gives a lower $G_{\text{low}}(R)$ and an upper $G_{\text{up}}(R)$ diabat. $G_{\text{low}}(R)$ is an approximation to the $G(R)$ on which the rebinding reaction ($\text{Mb}\cdots\text{CO} \rightarrow \text{MbCO}$) takes place. The parameters $R_c = 3.0$ Å and $\sigma = 0.2$ Å characterize the crossing seam,¹¹ and for the coupling strength, $c = 0.5$ is used. This choice is not critical for the resulting FEP or for the propagation of the probability distributions.

C. Solution of the Smoluchowski Equation. With $G(R)$ calculated at a particular temperature T , the relaxation of an initial CO population, $p(R, t)$, along R can be followed by solving the Smoluchowski equation:

$$\frac{\partial p(R, t)}{\partial t} = \frac{\partial}{\partial R} D(R) e^{-\beta G(R)} \frac{\partial}{\partial R} [e^{\beta G(R)} p(R, t)] \quad (2)$$

Equation 2 describes the decay of a nonequilibrium distribution, $p(R, t = 0)$, to the equilibrium $p_{\text{eq}}(R)$ governed by $G(R)$. To solve eq 2, the discrete approximation is used.^{39,40} For large R , reflecting boundary conditions were imposed (Figure 4). $D(R)$ is the R -dependent diffusion constant, $G(R)$ is the FEP, and $\beta = 1/kT$ is the Boltzmann factor. D was calculated in the distal pocket from mean-square displacements at $T = 200, 250$, and 300 K.⁴¹ For this, we ran 300 short-time trajectories of dissociated CO in B at all three temperatures. The resulting diffusion coefficients were $D = 2.2$ Å²/ps at $T = 300$ K, $D = 1.3$ Å²/ps at $T = 250$ K, and $D = 1.0$ Å²/ps at $T = 200$ K. In the following, D is position independent. Three cases with different initial populations were considered: $p_1(R, 0) = \delta(R_0 = 5.3$ Å) (corresponding to the location of the CO immediately after photodissociation; see Figures 3 and 4), $p_2(R, 0) = \delta(R_0 = 9.3$ Å) (the minimum of the FEP in Xe4), and $p_3(R, 0) = \frac{1}{2}\delta(R_0 = 5.3) + \frac{1}{2}\delta(R_0 = 9.3)$ (a mixture of the two which describes a possible situation shortly after dissociation, although quantitative populations of the different sites are not known from experimental data). For lower T , rebinding times can become

TABLE 1: Barrier Heights (in kcal/mol) for Transitions between Bound State (A), Distal Pocket (B), and Xe4 Pocket for Native MbCO and the L29F Mutant as a Function of Temperature^a

	native MbCO			L29F mutant 300 K
	200 K	250 K	300 K	
A \rightarrow B	21.6	21.2	21.1	21.4
A \leftarrow B	5.7	5.5	4.3	1.6
B \rightarrow Xe4	5.2	5.2	4.3	1.4
B \leftarrow Xe4	10.5	10.0	7.8	1.5

^a For $T = 200$ and 250 K, $\Delta = 5$ kcal/mol, and for $T = 300$ K, the value of $\Delta = 4$ kcal/mol is used, which reproduces the experimentally observed rebinding time. The barriers affected by the value of Δ are $H_{A \rightarrow B}$ and $H_{A \leftarrow B}$. Barriers for other values of Δ are mentioned in the text.

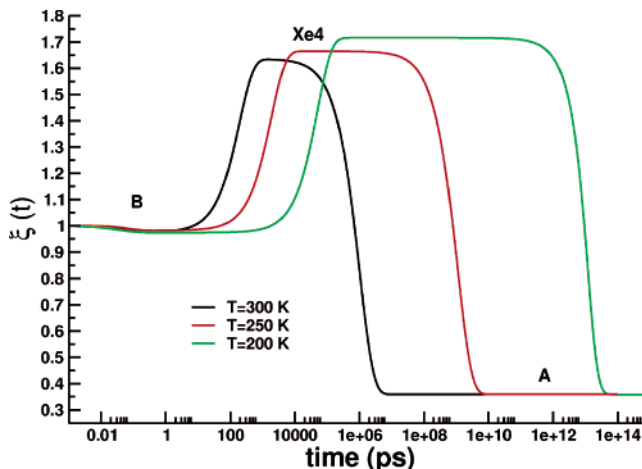


Figure 5. Time evolution of $\xi(t) = R_{\text{cm}}(t)/R_0$ of the distribution function for CO diffusing in native myoglobin on the diabatic FEPs at $T = 300, 250$, and 200 K (time on a logarithmic scale). At $\xi(t) = 1$, the distribution is in the B site, at $\xi(t) > 1$, it migrates toward the Xe4 pocket, and for $\xi(t) < 1$, it approaches the bound A state.

very long. To ensure stability of the propagation, the propagation matrix was explicitly symmetrized.

III. Results

The FEPs contain, by construction, all coordinates of MbCO. Thus, all couplings between the coordinates present in the force field are included in $G(R)$. This is different from constructing a model potential along selected coordinates.^{24,25} Because the molecular dynamics simulations are carried out for one particular (average) temperature, FEPs derived from them should only be used in following the evolution of an initial population distribution at the same temperature (see eq 2). In the following, FEPs and population dynamics for native MbCO and the L29F mutant are discussed separately.

A. Native MbCO. Figure 4 (inset) shows FEPs for bound and unbound CO calculated at different T along R . For direct comparison, the curves for unbound CO were shifted so that the minimum in the distal pocket (B) for each curve was at $G = 0$. The main Figure 4 shows the diabatic FEPs for the three temperatures. The “exit” barriers ($B \rightarrow \text{Xe4}$) $H_{B \rightarrow \text{Xe4}}$ do not vary appreciably with temperature, whereas the “entry” barriers ($B \leftarrow \text{Xe4}$) $H_{B \leftarrow \text{Xe4}}$ differ somewhat more. This is shown in Table 1 for both native MbCO, at the three simulation temperatures, and the L29F mutant, at 300 K.

In Figure 5, the time evolution of the center of mass of the probability distribution $\xi(t) = R(t)/R_0$ for CO at $T = 300, 250$, and 200 K starting from $p_1(R, 0)$ (the center of mass of the distribution is in site B) is shown. After 3 ps (at $T = 300$ K),

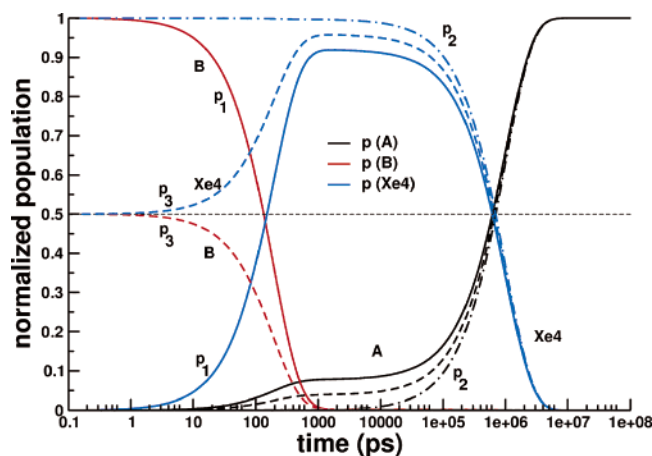


Figure 6. Time evolution for $p(R,t)$ in the A site (black lines), the B site (red line), and the Xe4 pocket (blue line) for CO diffusing on the FEPs at $T = 300$ K for the three different initial conditions (see text). For $p_1(R,t)$, the population and depopulation of the Xe4 pocket can be easily followed. Also, for $p_1(R,t)$, it is not possible to move the entire population into the Xe4 pocket because roughly 10% rebound to the A state without migration to Xe4. For initial condition $p_2(R,0)$, (dashed/dotted line) the population of the B site (red) is very small and not visible.

20 ps ($T = 250$ K), and about 600 ps ($T = 200$ K), $\xi(t)$ starts to move to the Xe4 pocket. Depending on the temperature, $\xi(t)$ reaches values between 1.6 and 1.7, indicating that $p_1(R,t)$ is primarily localized in the Xe4 pocket. After 3 ns ($T = 300$ K), 1.8×10^3 ns ($T = 250$ K), and 2.3×10^7 ns ($T = 200$ K), the distribution overcomes the barriers H_{B-Xe4} and H_{A-B} (see Figure 4) and $\xi(t)$ moves to its equilibrium value, which is 0.36. This corresponds to the A state.

The integrated probability distribution of CO in A, B, and Xe4 as a function of time gives further information on the rebinding dynamics. In Figure 6, the temporal evolution of $p(R,t)$ is reported for the three different initial conditions at 300 K. Up to 0.5 ps, the system remains in B, and after 2 ps, 0.01% of the particles make the transition to the Xe4 pocket. The population in Xe4 increases rapidly after 20 ps, whereas only 0.01% of the distribution rebounds after 40 ps (solid black line). Between 0.8 ns and 25 ns, 90% of the population remains in Xe4, and by $t = 600$ ns, the distribution is the same in Xe4 as in the A site. By $t = 2.3 \times 10^3$ ns, 90% have re-bound. A similar behavior is found for the initial conditions $p_2(R,t)$ and $p_3(R,t)$ (dot-dashed and dashed lines, respectively, in Figure 6), although the details of the rebinding process differ somewhat.

The rebinding time τ is calculated from $\tau = \int_0^\infty dt \Sigma(t)$ where $\Sigma(t) = \int p(R,t) dR$ and the second integration extends over all possible values of R .⁴² At 300 K, this yields a τ between 280 ns and $1.77 \mu\text{s}$ depending on the barrier height H_{A-B} . The barrier is directly related to the asymptotic separation Δ between the 1A and the 3A surface (see Figure 4) because changing Δ modifies the value of H_{A-B} . To explore the effect of different values of Δ on H_{A-B} and the rebinding times, population dynamics was run for a range of Δ values. A value of $\Delta = 5$ kcal/mol gives $H_{A-B} = 5.17$ kcal/mol and $\tau = 280$ ns. Further calculations were carried out up to $\Delta = 7.5$ kcal/mol in increments of 0.25 kcal/mol. For $\Delta = 7.5$ kcal/mol, $H_{A-B} = 6.48$ kcal/mol and $\tau = 1.77 \mu\text{s}$.

Experimental investigations of the $A \leftarrow B$ rebinding have provided Arrhenius constants and an effective rebinding barrier height.¹⁵ Using these values, one finds a room-temperature rebinding time of ≈ 100 ns, which is also consistent with nanosecond time-resolved spectroscopy.³ From the results

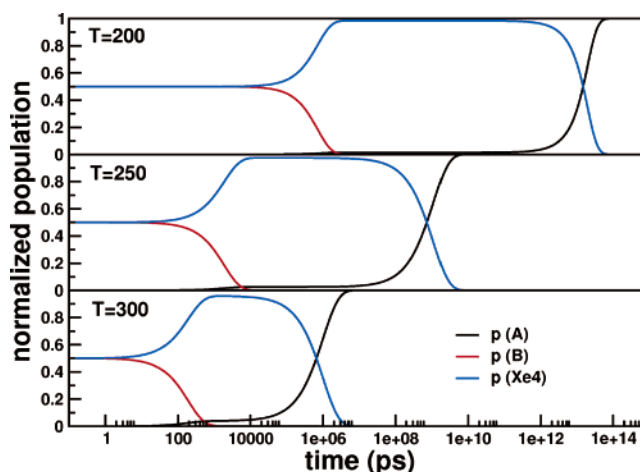


Figure 7. Time evolution for $p(R,t)$ in the A site (black line), the B site (red line), and the Xe4 pocket (blue line) for CO diffusing on the FEPs at $T = 300$, 250, and 200 K (from bottom to top), starting from $p_3(R,0)$ (initial populations are 50% in B and 50% in Xe4).

discussed in the previous paragraph, it is now possible to determine the asymptotic separation (the only free parameter in the present model) which reproduces a rebinding time of ≈ 100 ns. With $\Delta = 4$ kcal/mol for the adiabatic separation between the 1A and 3A surfaces, the inner barrier is $H_{A-B} = 4.3$ kcal/mol and the rebinding time becomes $\tau = 106$ ns. This value for H_{A-B} compares favorably with the effective barrier $H_{\text{eff}} = 4.5$ kcal/mol previously reported;¹⁵ H_{eff} is the effective barrier in the fully relaxed and fluctuonally averaged protein near 300 K, as described in ref 15, and is thus most relevant for comparison with the present simulations. However, it should be noted that rebinding barriers H_{A-B} into the majority conformation A_1 and the minority conformation A_3 are 2.4 and 4.3 kcal/mol, respectively.¹⁵ The value of $\Delta = 4$ kcal/mol agrees quite favorably with the previous estimate of 5 kcal/mol.³⁸ However, it should be noted that the meaning of these two numbers is quite different. While the separation of 5 kcal/mol between the 1A and the 3A surfaces in MbCO is calculated for heme-CO in vacuo, the value of $\Delta = 4$ kcal/mol corresponds to a conformationally averaged value. Explicit calculation of this number by ab initio calculations would be very time consuming because it requires the calculation of total energies of Mb-CO in its 1A and 3A states for many different conformations ideally sampled from MD simulations.

Rebinding studies at lower temperatures exhibit much delayed dynamics for the rebinding from Xe4 via B to A, in accord with an activated process governed by an Arrhenius relationship. In Figure 7, the temporal evolution of $p_3(R,t)$ in the three sites is reported for the three different temperatures. The depopulation of the B state occurs on time scales of 100 ps to 1 ns, and population of the A state takes from 100 μs to seconds at 200 K.^{3,15} As for 300 K, calculations for a range of Δ values between 5.0 and 7.5 kcal/mol (corresponding to H_{A-B} between 5.5 and 7.1 kcal/mol at 250 K and between 5.7 and 7.2 kcal/mol at 200 K, respectively) were carried out. Rebinding times, τ , varied between 300 μs and 3.3 ms at 250 K and from 2 to 10 s at 200 K.

B. L29F Mutant MbCO. For the L29F mutant, rapid migration of CO after photodissociation from B to the Xe4 pocket was observed.¹² The process occurred on a nanosecond time scale and was also found in MD simulations.^{19,26} The rapid time scale suggests that the barriers for migration between B and Xe4 should be small. The simulations were carried out at 300 K. The FEP along the reaction coordinate is flat with only

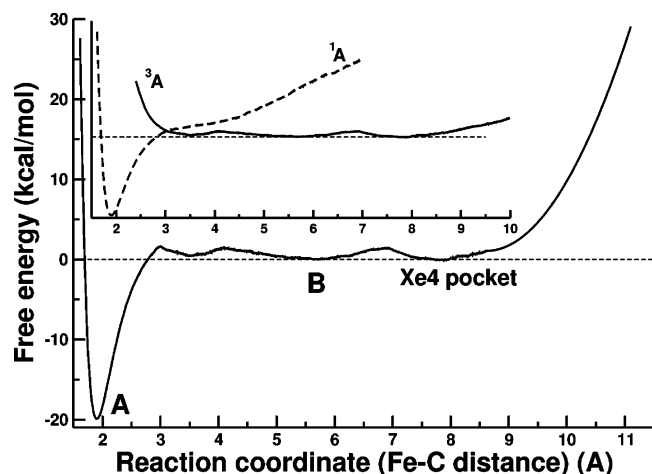


Figure 8. Adiabatic (main figure) and diabatic (inset) FEPs for L29F MbCO at 300 K. In the inset, the dashed line corresponds to the bound state FEP and the solid line is the unbound FEP. The barriers between the B and Xe4 binding sites are much smaller than those for native MbCO (see Figure 4). The energy scale is shifted vertically to set the free energy in the B state to 0. This allows comparison with FEPs in Figure 4.

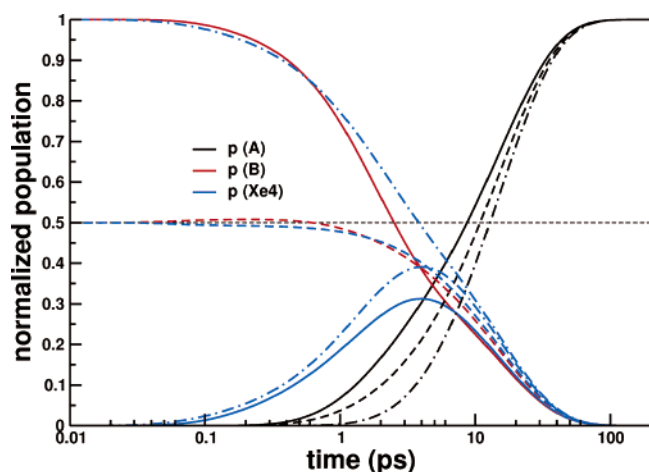


Figure 9. Time evolution of $p_i(R,t)$ for the L29F mutant; the index i labels the three different initial conditions mentioned in the text. Because the barriers involved are small, all rebinding times are short and similar for the three different initial conditions. No plateaus appear for the population of the bound A state, as is the case for native MbCO (see Figure 6). For initial condition $p_2(R,0)$ (dashed/dotted line), the population of the B site (red) is very small and not visible.

minor barriers $H_{A \leftarrow B} = 1.6$ kcal/mol and $H_{B \leftarrow \text{Xe4}} = 1.4$ kcal/mol (see Figure 8). This indicates that facile migration of CO in L29F after photodissociation is possible, in agreement with the rapid change in the infrared spectrum of photodissociated CO observed experimentally.¹² To investigate the rebinding dynamics and assess rebinding times from Xe4 via B to A, absorbing boundary conditions for $R > 10$ Å were applied to the FEP and the Smoluchowski equation was solved for different initial populations. The resulting populations of $p_i(R,t)$ in the different wells, A, B, and Xe4, are shown in Figure 9, where i stands for the different initial conditions.

Rebinding times for the three different initial conditions are 14, 17, and 16 ps, respectively. They are considerably faster than those for native Mb because of the lower barriers between the binding sites. However, rebinding in possible experiments on the L29F mutant will also have longer time scales due to CO rebinding from sites further away than Xe4 (e.g., the solvent).

IV. Discussion and Conclusions

Ligand rebinding from sites near the heme in native and mutant (L29F) Mb was investigated using a combination of molecular dynamics and stochastic simulations to follow an initial population on a realistic free-energy curve. The FEP was calculated along an appropriately chosen progression coordinate using state-of-the-art interaction potentials between CO and the surrounding protein matrix. To our knowledge, this study is the first that quantitatively characterizes the FEP governing CO migration within Mb on the basis of detailed simulations rather than model potentials. All relevant coordinates, in particular, the Fe out-of-plane motion and all heme-related coordinates, are explicitly included in the FEP. One limitation of the present approach is to restrict the progression coordinate to describe ligand migration from state A via B to Xe4, which does not include effects of rebinding from sites further away, e.g., from Xe1. Focusing on the immediate surroundings of the heme group is meaningful because it allows us to treat and understand the early stages of the rebinding process and the local dynamics in more detail. Also, the Fe–C coordinate loses its meaning as a valid progression coordinate beyond the Xe4 pocket because of the complex internal structure of Mb.

Umbrella sampling yields a rebinding barrier of $H_{B \leftarrow \text{Xe4}} = 7.8$ kcal/mol. This value is larger than an estimate from the 90 ns trajectory by Bossa et al., which is 2.4 kcal/mol.⁹ In this work, the height of the barrier comes from a limited number of crossing events (11) between B and Xe4. Such a small number of crossings can lead to systematic errors in estimating free-energy barriers because the barrier region is not sufficiently sampled by the MD simulations. Furthermore, electrostatic interactions between CO and the protein environment are described by the validated fluctuating three-point charge model in the present work and by a conventional two-point model in the work by Bossa et al. However, to ensure that the larger barrier calculated here is not an artifact, a second set of umbrella sampling was run at 300 K for native MbCO, starting from the structure at the transition state ($R = 7.0$ Å; see Figure 4). The rebinding barrier is 6.8 kcal/mol, which is 1 kcal/mol smaller than that from the first simulation but still considerably larger than the value of 2.4 kcal/mol. Support for substantial rebinding barriers $H_{B \leftarrow \text{Xe4}}$ comes from observations on the L29W mutant.¹⁴ For this system, the $B \leftarrow C$ barrier is 5.5 kcal/mol, considerably larger than the $A \leftarrow B$ barrier which is 1.5 kcal/mol. Thus, a value of 6.8–7.8 kcal/mol for the $B \leftarrow \text{Xe4}$ barrier, as calculated in the present work, is realistic. Umbrella sampling may somewhat overestimate the rebinding barrier because the ligand and the protein environment can adjust to each other and optimize their interaction. However, to quantify such an effect is difficult. To demonstrate how the ligand and the protein influence each other, the final structures after the umbrella sampling for native and mutant L29F MbCO at $R = 8.05$ and 9.70 Å are shown in Figure 10. Although the shapes of the Xe4 pocket for native and mutant Mb are still quite similar for $R = 8.05$ Å, several pocket-forming residues have considerably shifted for the simulation with $R = 9.70$ Å. This is indicated by the green arrow. Such an optimization of the interaction between the ligand and the protein environment could contribute to some overestimation of the stabilization in a local minimum on the potential energy surface.

The combination of umbrella sampling and stochastic simulations provides new and fundamental insight into the ligand rebinding $\text{Mb} \cdot \text{CO} \rightarrow \text{MbCO}$ for native and mutant (L29F) myoglobin. Direct simulations for converged barrier-crossing statistics are very difficult because the time scales involved in

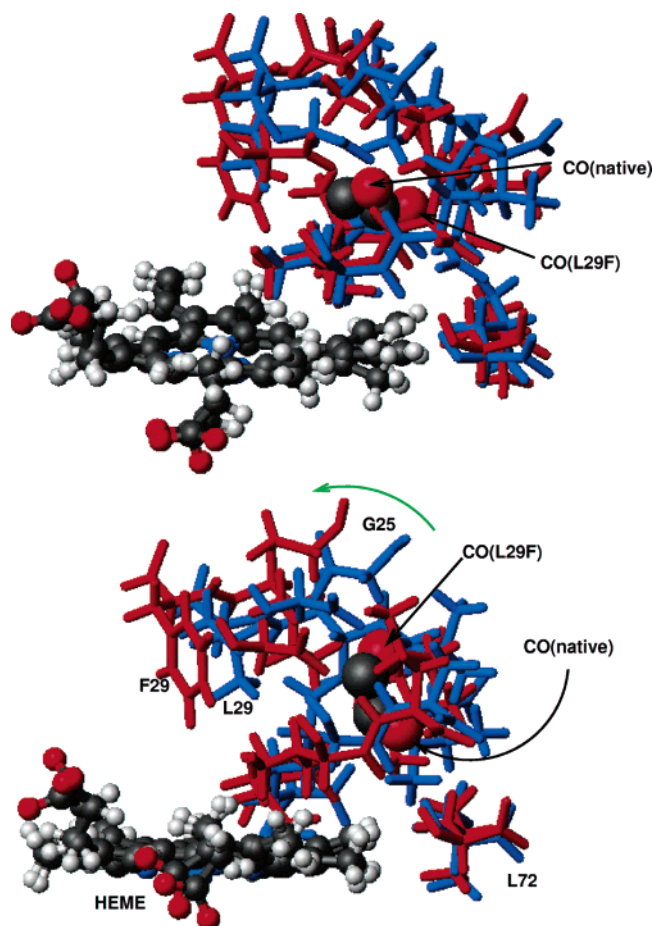


Figure 10. Snapshots for $R = 8.05 \text{ \AA}$ (above) and $R = 9.70 \text{ \AA}$ (below) of the structure of the Xe4 pocket with the CO ligand (red/black) after umbrella sampling at 300 K with native (blue wires) and mutant (red wires) MbCO. The two structures are superimposed to minimize the root-mean-square deviation among the heme groups of the two structures. Important residues are labeled. The green arrow shows the direction in which the residues of L29F are displaced with respect to the positions of the residues of native Mb. Note that L72 remains essentially at the same position.

the rebinding process can be long, and sufficient sampling of crossing events (e.g., $B \leftarrow \text{Xe4}$) by standard MD simulations become difficult. With the present approach, we were able to characterize the population dynamics of CO molecules in binding sites (A, B, and Xe4) around the heme. The present study does not allow us to take into account direct rebinding $A \leftarrow \text{Xe4}$ because the rebinding path includes migration via the distal pocket (B) explicitly. However, to the best of our knowledge, there is neither experimental evidence nor indications from simulations of such a direct process.⁹ In conclusion, FEPs were calculated for CO migration between the bound and the B and Xe4 sites, respectively, for native and mutant (L29F) Mb. The population dynamics among the three sites considered supports the existence of multiple time scales for native MbCO, whereas for the L29F mutant, the rebinding dynamics is much faster and only one time scale appears. Where available, rebinding barriers are in qualitative (L29F not measured but inferred from the rapid population of Xe4) to good agreement with experimental data.

Acknowledgment. The authors acknowledge financial support from the Swiss National Science Foundation through a Förderungsscholarship to M.M. We thank Dr. D. R. Nutt for comments on the manuscript.

References and Notes

- (1) Austin, R. H.; Beeson, K. W.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. C. *Biochemistry* **1975**, *14*, 5355.
- (2) Olson, J.; Philipps, G. *J. Biol. Chem.* **1996**, *271*, 17593.
- (3) Ansari, A.; Jones, C. M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Biochemistry* **1994**, *33*, 5128.
- (4) Lim, M.; Jackson, T. A.; Anfinsen, P. A. *Proc. Natl. Acad. Sci.* **1993**, *90*, 5801.
- (5) Nienhaus, K.; Deng, P.; Kriegl, J. M.; Nienhaus, G. U. *Biochemistry* **2003**, *42*, 9647.
- (6) Case, D. A.; Karplus, M. *J. Mol. Biol.* **1979**, *132*, 343.
- (7) Tilton, R.; Singh, U.; Weiner, S.; Connolly, M.; Kuntz, I.; Kollman, P.; Max, N.; Case, D. A. *J. Mol. Biol.* **1986**, *192*, 443.
- (8) Elber, R.; Karplus, M. *J. Am. Chem. Soc.* **1990**, *112*, 9161.
- (9) Bossa, C.; Massimiliano, A.; Roccatano, D.; Amadei, A.; Vallone, B.; Brunori, M.; Nola, A. D. *Biophys. J.* **2004**, *86*, 3855.
- (10) Agmon, N. *Biophys. J.* **2004**, *87*, 1537.
- (11) Meuwly, M.; Becker, O. M.; Stote, R.; Karplus, M. *Biophys. Chem.* **2002**, *98*, 183.
- (12) Schotte, F.; Lim, M.; Jackson, T. A.; Smirnov, A. V.; Soman, J.; Olson, J. S.; Phillips, G. N., Jr.; Wulff, M.; Anfinsen, P. A. *Science* **2003**, *300*, 1944–1947.
- (13) Tilton, R. F., Jr.; Kuntz, I. D., Jr.; Petsko, G. A. *Biochemistry* **1984**, *23*, 2849–2857.
- (14) Ostermann, A.; Waschipky, R.; Parak, F. G.; Nienhaus, G. U. *Nature* **2000**, *404*, 205.
- (15) Steinbach, P. J.; Ansari, A.; Berendzen, J.; Braundstein, D.; Chu, K.; Cowen, B. R.; Ehrenstein, D.; Frauenfelder, H.; Johnson, J. B.; Lamb, D. C.; Luck, S.; Mourant, J. R.; Nienhaus, G. U.; Ormos, P.; Philipp, R.; Xie, A.; Young, R. D. *Biochemistry* **1991**, *30*, 3988.
- (16) 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.
- (17) Carver, T. E.; Rohlfis, R. J.; Olson, J. S.; Gibson, Q. H.; Blackmore, R. S.; Springer, B. A.; Sligar, S. G. *J. Biol. Chem.* **1990**, *265*, 20007.
- (18) Vojtechovsky, J.; Chu, K.; Berendzen, J.; Sweet, R. B.; Schlichting, I. *Biophys. J.* **1999**, *77*, 2153.
- (19) Nutt, D. R.; Meuwly, M. *Proc. Natl. Acad. Sci.* **2004**, *101*, 5998.
- (20) Frauenfelder, H.; McMahon, B. J.; Austin, R. H.; Chu, K.; Groves, J. T. *Proc. Natl. Acad. Sci.* **2001**, *98*, 2370.
- (21) Scott, E. E.; Gibson, Q. H.; Olson, J. S. *J. Biol. Chem.* **2001**, *276*, 5177.
- (22) Scott, E. E.; Gibson, Q. H. *Biochemistry* **1997**, *36*, 11909.
- (23) Brunori, M.; Bourgeois, D.; Vallone, B. *J. Struct. Biol.* **2004**, *147*, 223.
- (24) Agmon, N.; Hopfield, J. J. *J. Chem. Phys.* **1983**, *79*, 2042.
- (25) Srajer, V.; Reinisch, L.; Champion, P. M. *J. Am. Chem. Soc.* **1988**, *110*, 6656.
- (26) Nutt, D.; Meuwly, M. *Biophys. J.* **2003**, *85*, 3612.
- (27) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
- (28) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
- (29) Kuczera, K.; Kuryian, J.; Karplus, M. *J. Mol. Biol.* **1990**, *213*, 351.
- (30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (31) Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 7690.
- (32) Powers, L.; Sessler, J. L.; Woolery, G. L.; Chance, B. *Biochemistry* **1984**, *23*, 5519.
- (33) Keyes, M. H.; Falley, M.; Lumry, R. *J. Am. Chem. Soc.* **1971**, *93*, 2035.
- (34) Cheng, X.; Schoenborn, B. P. *J. Mol. Biol.* **1991**, *220*, 381.
- (35) Straub, J. E.; Karplus, M. *Chem. Phys.* **1991**, *158*, 221.
- (36) Huffaker, J. N. *J. Chem. Phys.* **1976**, *64*, 3175.
- (37) Shin, C.; Shin, S. *J. Chem. Phys.* **2000**, *113*, 6528.
- (38) Harvey, J. N. *J. Am. Chem. Soc.* **2000**, *122*, 12401.
- (39) Bicout, D.; Szabo, A. *J. Chem. Phys.* **1998**, *109*, 2325.
- (40) Banushkina, P.; Meuwly, M. *J. Chem. Theo. Comput.* **2005**, *1*, 208.
- (41) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Adv. Chem. Phys.* **1988**, *LXXI*, 1.
- (42) Szabo, A.; Schulten, K.; Schulten, Z. *J. Chem. Phys.* **1980**, *72*, 4350.