CHIMIA **2005**, *59*, No. 7/8

Chimia 59 (2005) 517–521

ISSN 0009–4293

Computational Chemistry for Elucidating Protein Function: Energetics and Dynamics of Myoglobin–Ligand Systems

David R. Nutt, Polina Banushkina, and Markus Meuwly*

Abstract: State-of-the-art computational methodologies are used to investigate the energetics and dynamics of photodissociated CO and NO in myoglobin (Mb···CO and Mb···NO). This includes the combination of molecular dynamics, *ab initio* MD, free energy sampling, and effective dynamics methods to compare the results with studies using X-ray crystallography and ultrafast spectroscopy methods. It is shown that modern simulation techniques along with careful description of the intermolecular interactions can give quantitative agreement with experiments on complex molecular systems. Based on this agreement predictions for as yet uncharacterized species can be made.

Keywords: *Ab initio* MD · Free energy surface · Molecular dynamics · Myoglobin · QM/MM simulations · Smoluchowski equation

1. Introduction

The function of proteins is intimately linked to their three-dimensional structure and fluctuations around this structure [1]. Over the last two decades impressive progress has been made, both experimentally and theoretically, to unveil and characterize functional aspects of protein dynamics. Myoglobin interacting with small ligands such as CO or NO (MbXO) has been extensively studied because of its biological importance. In most cases, photodissociation of the iron–ligand bond is used to form a reactive, unbound state. Cleavage of the Fe–X bond is followed by a series of spectroscopic and structural changes which may ultimately lead to ligand rebinding and reformation of MbXO. The dynamics of the free ligand (XO) within the protein matrix after photodissociation is characteristic of the ligand–protein interactions. This can be detected by infrared (IR) spectroscopy of the XO stretching vibration. For example, CO bound to the heme-Fe has three strong IR absorption bands (A-bands) around 1950 cm^{-1} while the bands associated with photodissociated CO are weaker and appear around 2130 cm^{-1} (B-bands) [2][3]. It is suggested that the B-bands originate from CO vibrating in the docking site (Fig.1) although still some controversy exists as to the orientation of the CO molecule [4]. Two well-resolved bands have been associated with two distinct CO orientations in the docking site, separated by a low free energy barrier. Spectroscopic measurements also provide information about rates for relaxation and migration processes since spectroscopic properties are quite sensitive to structural changes [5]. In most cases, however, the structural changes cannot be directly observed, but are inferred indirectly from the spectral changes. The development of ultrafast time-resolved X-ray crystallography has provided supplementary experimental information in such cases.

Computational investigations of the energetics, dynamics and spectroscopy of complex systems have become feasible in the past few years. This is mainly due to the increased power of computers but also because new and improved numerical algorithms have become available. Myoglobin is a small globular protein (153 residues) which makes it an ideal subject for computational studies. Despite abundant information from experimental and theoretical studies, a microscopic description of ligand dynamics within the protein remains elusive and fundamental questions remain unanswered. They include for example the origin of the splitting of the IR bands in MbCO, the pathways and energetic barriers for ligand migration between the cavities of Mb, and the understanding of the time scales for ligand (CO, NO) rebinding after dissociation from the heme. In the following, computational methods to address such questions and to propose microscopic pictures for elucidating ligand dynamics in proteins are discussed. They are then applied to CO and NO interacting with myoglobin.

2. Theoretical Methods

Molecular dynamics (MD) simulations described here are carried out with the CHARMM program [6] and the CHARMM22 force field [7]. One particular concern is the way in which the electrostatic interactions between the free ligand (CO, NO) and the environment are treated [8–10]. The fluctuating charge model developed for this purpose correctly reproduces the experimentally observed existence of the docking

*Correspondence: Prof. M. Meuwly
Department of Chemistry
University of Basel
Klingelbergstrasse 80
CH-4056 Basel
Tel.: +41 61 267 3821
E-Mail: m.meuwly@unibas.ch

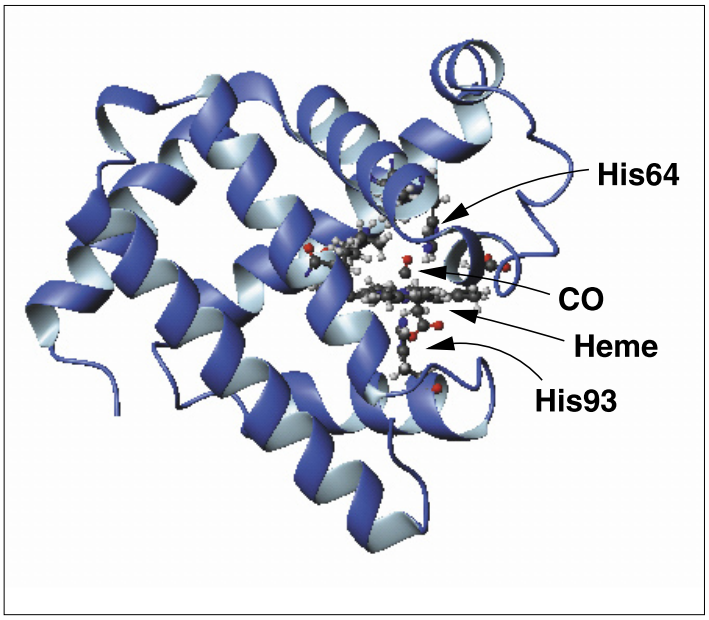

Fig. 1. The active site region of MbCO. The residues involved in forming the docking site at the edge of the distal heme pocket and the xenon 4 pocket are shown in a ball-and-stick representation. The heme unit is drawn in a line representation. Helix A is in the foreground and all helices are shown transparent for clarity. The solid arrow points towards the entrance of the xenon-4 pocket.

site within the distal heme pocket for MbCO (Fig. 2) as well as the migration of CO through the protein to the xenon-4 pocket (Xe4), which has been observed for various mutants [11][12]. The time-scales and energy barriers for motion from the binding site to the docking site [8] and from the docking site to the Xe4 pocket [13] (see below) were also found to be in almost quantitative agreement with experiment.

The IR spectrum is calculated from the Fourier transform $C(\omega)$ of real-time dipole–dipole autocorrelation function, $C(t)$. From $C(\omega)$ the adsorption spectrum is then calculated by evaluating

$$A(\omega) = \omega\{1-\exp[-\omega/(kT)]\}C(\omega) \qquad (1)$$

where k is the Boltzmann constant and T is the temperature in Kelvin.

With MD simulations, free energy profiles (FEPs) for ligand migration within protein active sites can be calculated along selected progression coordinates using umbrella sampling. In umbrella sampling, statistics n_i for given intervals Δq along a particular coordinate q are collected [14]. Associated free energies, $G(q)$, 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. For ligand migration (Fig. 2(i)) from the binding site (A) *via* the docking site (B) to the Xe4 pocket, the reaction coordinate R is the Fe–C distance. FEPs were calculated at 300 K for the bound (^1A and the unbound (^3A) states. The two adiabatic curves are diabatized since the rebinding dynamics take place on the lower diabat $G(R)$ (Fig. 3).

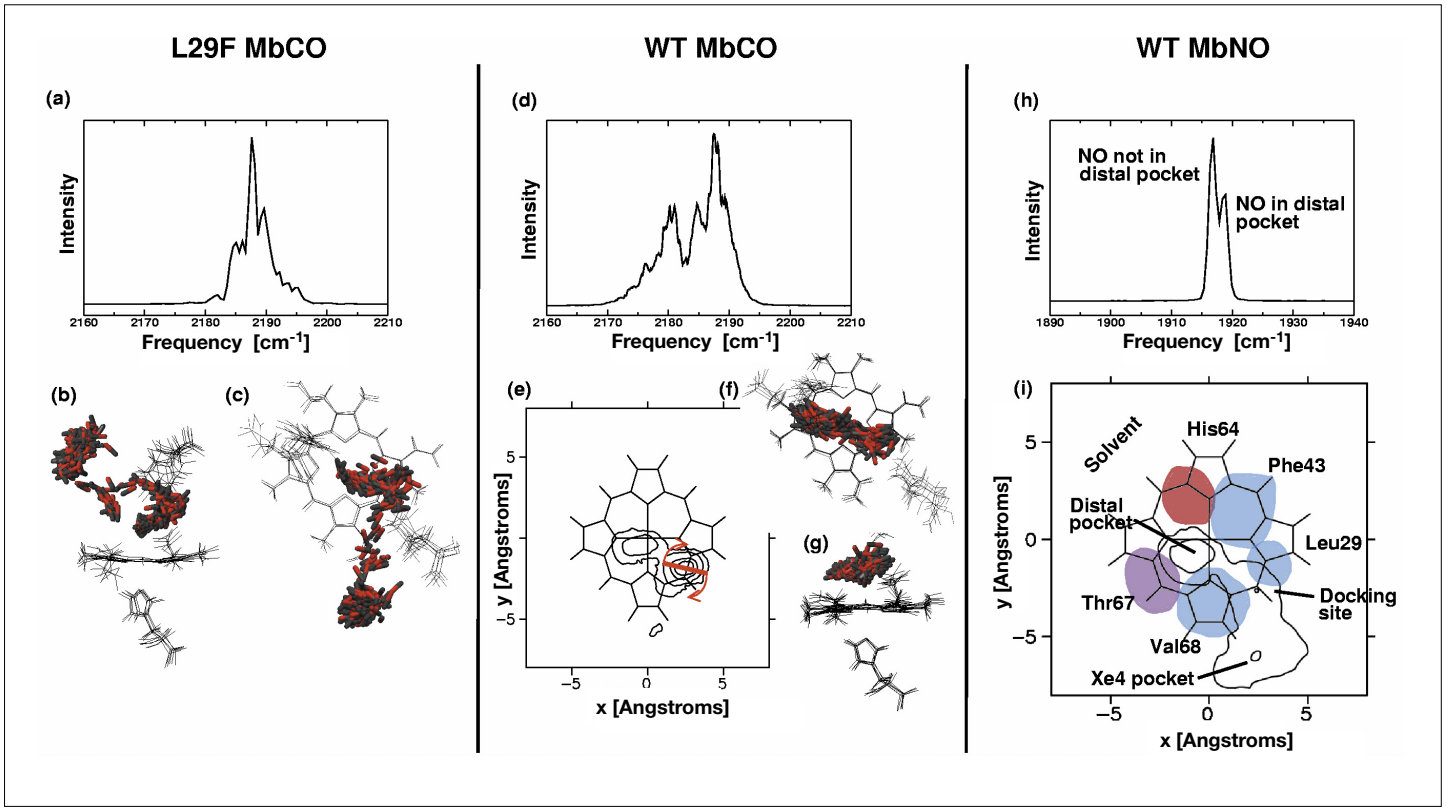

Fig. 2. (a), (d), and (h) show the calculated IR spectra for CO in L29F MbCO, wild-type MbCO and for NO in wild-type MbNO. The frequency scale is the same in all cases, illustrating the different widths of the peaks. In (h), the two peaks can be assigned to NO in the distal heme pocket and NO away from the distal pocket by correlating over the relevant parts of the trajectories. (b) and (c) illustrate side and top views of the CO motion in L29F MbCO as calculated from QM/MM simulations. Motion from the distal heme pocket to the xenon-4 pocket can be observed. (e), (f) and (g) show the configuration space explored by CO in wild-type MbCO. Part (e) shows the projection of the CO centre of mass onto the heme plane from a 1 ns molecular dynamics simulation. The rotation of the ligand in the docking site is indicated schematically. Parts (f) and (g) show top and side views of equivalent QM/MM calculations. (i) shows a schematic drawing of the configuration space available to NO (and also to CO) in wild-type Mb, together with a calculated probability distribution for NO. A selection of residues forming the distal heme pocket are shown and colored according to their type; red = charged, purple = polar and blue = non-polar. The docking site and Xe4 pocket are also indicated.

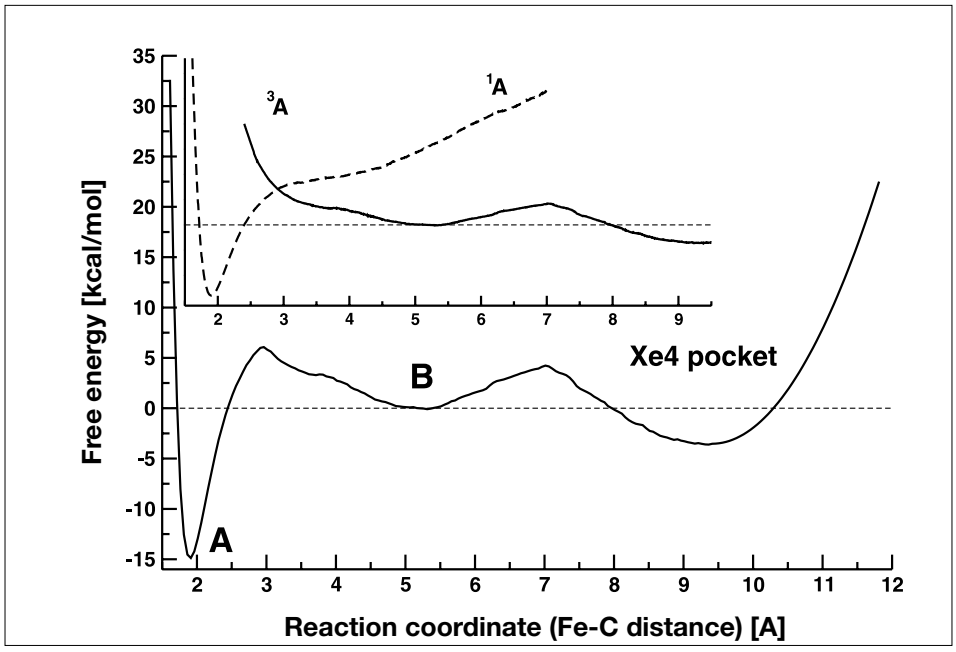

Fig. 3. Adiabatic (main figure) and diabatic (inset) FEPs. In the inset, the dashed line corresponds to the bound state FEP while the solid line is the unbound FEP. For the ^1A state the FEP becomes repulsive at long range due to the presence of protein residues.

Once $G(R)$ has been calculated, the relaxation of an initial population of CO along R can be followed by solving the Smoluchowski equation.

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

Eqn. (2) describes the decay of a non-equilibrium distribution $p(R,t=0)$ to the equilibrium $p_{eq}(R)$ governed by $G(R)$. To solve Eqn. (2) the discrete approximation (DA) is used [15]. Depending on the roughness of the FEP more efficient algorithms than the simple DA are available [16].

3. Vibrational Spectroscopy and Ligand Motion

3.1. *Vibrational Spectroscopy for MbCO*

The IR spectrum of the photodissociated CO molecule is calculated from the dipole moment time series from the MD trajectories. Spectra from different MD simulations are slightly different, reflecting the different trajectories of the CO molecules and the continually changing protein environment. In experiments, the spectra recorded correspond to many CO molecules in different environments and therefore give an *average* picture of the CO environment. For wild type MbCO, the averaged spectrum over several 1 ns trajectories (Fig. 2(d)) shows two principal bands separated by ≈ 8 cm^{-1} with intensities approximately in the ratio 2:1, in almost quantitative agreement with experiment [2][3]. Since there are several processes occurring following the photodissociation of CO from Mb (docking, diffusion through the protein, loss of the CO to the solvent) the IR spectra of more extended trajectories might show additional features due to ligand diffusion through the protein environment. In agreement with previous experimental and theoretical results (Fig. 2(e)), the CO spends significant time in the region at the edge of the distal heme pocket (a region surrounded by residues *Leu*29, *Phe*43, *Val*68 and *Ile*107, Fig. 2(h)) [2][3][17][18].

Detailed inspection of the trajectories reveals that the motion of CO in the B site is essentially restricted to a plane parallel to the heme plane (Fig. 2(e)). From these trajectories the barrier for rotation Fe···CO→ Fe···OC is calculated from the relationship $G(q) = -RT\ln(P(q)) + G_0$, where R is the ideal gas constant, T is the temperature, G_0 is a constant, and q is the angle describing the rotation. The free energy barrier to rotation of CO parallel to the heme plane was found to be approximately 0.9 kcal mol^{-1} for the transition 0°→155° and 1.2 kcal mol^{-1} for the reverse transition. The difference in free energy between the two orientations is approximately 0.3 kcal mol^{-1}, with the orientation ϕ = 155° lower in free energy. This is in quantitative agreement with the experimental results of Lim *et al.* [3] who found a free energy difference of 0.29 kcal mol^{-1} between the two orientations, with the orientation with the O atom closer to the centre of the heme being lower in free energy.

For wild type MbCO the distance between the CO and the heme plane increases significantly in rare cases, and the CO molecule leaves the docking site at the edge of the distal heme pocket (Fig. 2(e)). This corresponds to the CO entering the 'xenon-4 cavity' as has been observed *e.g.* by Ostermann and coworkers [19]. The 'xenon 4' cavity is a hydrophobic pocket adjacent to the heme binding pocket, lined by the residues *Gly*25, *Ile*28, *Leu*29, *Gly*65, *Val*68, *Leu*69, *Leu*72, *Ile*107 and *Ile*111, which was found to bind Xe atoms. The CO is then observed to return to the docking site. Elber and Karplus [20] also found that in wild-type MbCO transitions of CO between cavities are rare events that are rapid and involve the crossing of barriers. The kinetic data of Ostermann and coworkers suggests a rate constant of the order of milliseconds and barriers of between 2.7 and 5.4 kcal mol^{-1} for motion to and from this pocket in the mutant L29W [19]. This also suggests that even longer simulations may reveal behavior which is not observable on the nanosecond time scale. Bossa *et al.* recently published results from a 90 ns trajectory, in which diffusion pathways between the xenon pockets (four in total) were determined [21]. However, such extensive simulations are not yet routinely possible.

Experimentally, the IR spectrum of dissociated CO was recently observed for the L29F mutant of myoglobin [22]. It was found that the IR spectrum consists of a single line for all but the shortest times after dissociation and that CO escapes from the main binding site on a sub-ns time scale. MD simulations using both the fluctuating point-charge model and QM/MM simulations (which treat the CO molecule with B3LYP/6-31G**) show that in the L29F mutant the CO escapes on a timescale of several 10 ps after dissociation (Fig. 2(b)(c)). The corresponding IR spectra for native and mutant MbCO are characteristically different: for CO in the native protein, the spectrum is split into two or possibly more bands that can be attributed to different binding sites within the distal heme pocket, while for the mutant, largely unstructured spectra are found from the simulations, as shown in Fig. 2(a).

One continuing question in the literature concerns the origin of the splitting of the B band and its relation to structural biology. Agreement seems to have reached in that the two bands are attributed to two different orientations of the CO molecule with respect to the heme-Fe. The original assignment of B_1 to Fe···OC and B_2 to Fe···CO has recently been challenged based on results from *ab initio* calculations [4]. Since these calculations assume a static protein environment (neglecting the dynamics of the protein environment), MD simulations at 100 K were run. The lower temperature prevents the CO from rotating in the docking site and allows the calculation of IR

spectra for both orientations separately. The spectra for the Fe···OC and Fe···CO conformation are split by about 2.5 cm^{-1}, a factor of 4 less than observed experimentally. This may also be due to the reduced temperature. Simulations at higher temperature still allow to extract IR spectra over the time for CO in one of the two possible orientations. At temperatures of 200 and 300 K, the spectra exhibit a splitting of between 2 and 3 cm^{-1}, however the correlation between the two peaks and the two orientations is unclear. It is possible that the splitting is inherent to the IR spectrum of the ligand in the docking site and that no direct one-to-one assignment to a particular structure can be made.

3.2. *Vibrational Spectroscopy for MbNO*

The related wild-type MbNO system has been less widely studied than MbCO. This is probably because rebinding of photodissociated NO to the heme unit occurs rapidly, on a timescale of picoseconds. Infrared studies are complicated by the weak IR signal of the ligand and spectral congestion due to the protein amide bands in the region corresponding to bound NO (1615 cm^{-1}). A three-point fluctuating charge model similar to the one developed for CO allows the calculation of IR spectra of photodissociated NO [10]. Using MD simulations and analysis comparable to that described for Mb···CO, the spectra for MbNO are generally found to be narrow, with a line width of approximately 5 cm^{-1}, and observed in the range 1'915–1'922 cm^{-1} (Fig. 2(h)). This is in contrast to wild-type MbCO where the experimental and calculated spectra span up to 30 cm^{-1}. Calculating vibrational frequencies from the dipole–dipole correlation function introduces a slight overall shift (≈10 cm^{-1}) of the spectrum which has to be included. Thus, the experimental peak position is expected to occur in the range 1905–1912 cm^{-1}. This slight blue shift for NO (1–8 cm^{-1}) is smaller than the corresponding red shifts observed for CO (8–26 cm^{-1}) [2]. It is known that the stretching frequencies of ligands such as CO and NO are sensitive to the local electrostatic field [23]. One possible interpretation is that a red shift of 5 to 26 cm^{-1} for CO corresponds to a weakening of the CO bond, suggesting that the effective CO potential along a MD trajectory is slightly flatter than the corresponding gas phase potential. This flattening originates from collisions and nonbonded interactions with the protein. Taking CO as the reference, the blue shift of 1 to 8 cm^{-1} for NO indicates that the weakening of the bond is essentially compensated through stronger electrostatic interaction. From this perspective, the electrostatic interactions seem to have a stronger influence on the NO vibration.

In the calculated spectra for NO, a splitting is observed due to two opposite orientations of the ligand within the docking site at the side of the distal pocket. The calculated splitting is 0.5 cm^{-1}, much smaller than the splitting observed for CO, which is more than an order of magnitude larger. A larger splitting (2 cm^{-1}) was found between the signals from NO in the distal heme pocket and NO in the xenon-4 pocket. Decreasing the temperature below 300 K has a small effect on the position of the peak. The line shapes vary depending on the specific ligand position and the presence of barriers between conformations. These barriers are small and disappear once the system is warmed above 200 K.

In the L29F mutant, no significant changes in the NO dynamics or IR spectrum are observed. This is in stark contrast to the behavior of the MbCO system. Here, the CO rapidly leaves the distal heme pocket, passing either side of the bulky Phe29 residue. Schotte *et al.* describe the phenyl residue as "*rapidly sweeping CO away from its primary docking site*" [22]. The corresponding IR spectra also show large changes over the first nanosecond following dissociation, a time scale directly accessible to molecular dynamics simulations [22]. Analysis of the dynamics suggests that rapid rebinding of NO to the heme is influenced by the fact that the dissociated ligand remains largely in a region above the centre of the heme, within the distal heme pocket. This is in contrast to MbCO, where the free ligand rapidly moves to a docking site at the edge of the distal heme pocket. The different behavior appears to be due to favorable electrostatic interactions between NO and the distal histidine which are lacking in the case of CO.

The observation of NO within the docking site at the edge of the distal heme pocket and in the xenon-4 pocket, as observed for CO [8], suggests that these docking sites are universal features of myoglobin and do not depend on the ligand. However, the stability of the ligand within these docking sites, and hence their relative importance in the dynamics following photodissociation, will depend on the identity of the ligand. In addition, the pockets of myoglobin are not simply a static environment in which ligands can move. The dynamics of the NO and CO ligands in the L29F mutant shows that different ligands can behave very differently in the same protein environment. Determining which features are due to the protein and which to the ligand will be an important step towards understanding ligand discrimination and protein function.

3.3. *Ligand Dynamics and Rebinding in MbCO*

Since the rebinding times in Mb···CO are of the order of several hundred nanoseconds [24], direct MD simulations with sufficient statistics are not possible. An alternative is to calculate the free energy surface (FES) $G(R)$ for the rebinding along a well-defined reaction coordinate R (see 2. Theoretical Methods) [13][16]. Through umbrella sampling, $G(R)$ contains information about the entire protein motion and thus can be used to solve the Smoluchowski equation from which rebinding times are calculated. Starting from $p(R,0) = 1/2\delta(R_0=5.3)+1/2\delta(R_0=9.3)$ (population of B and Xe4, which is a possible situation shortly after dissociation) the time evolution of $p(R,t)$ can be followed (Fig. 4). The actual rebinding time τ is then calculated from $\tau = \int_0^\infty dt\Sigma(t)$ where $\Sigma(t) = \int p(R,t)dR$

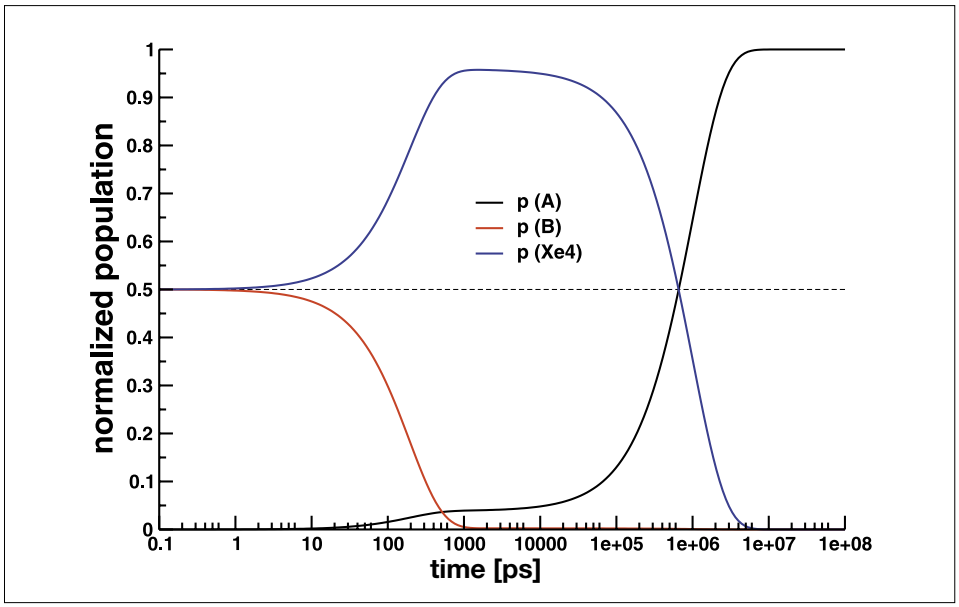

Fig. 4. Time evolution for $p(R,t)$ in the A site (black line), the B site (red line) and the Xe4 pocket (blue line) for CO diffusing on the FEP at T = 300.

[25]. Since the asymptotic separation Δ of the ^{1}A (MbCO) and the ^{3}A (Mb···CO) surface is not known exactly, this is used as a parameter. *Ab initio* studies on a model heme–CO system estimated Δ ≈ 5 kcal/mol [26]. Calculating τ for a range of values of Δ allows the determination of the asymptotic separation that best reproduces the experimentally determined rebinding time τ≈100 ns for Mb···CO. This yields Δ = 4.0 kcal/mol and leads to an inner barrier ($H_{A\leftarrow B}$) of 4.3 kcal/mol which compares favorably with the effective barrier of 4.5 kcal/mol reported by Steinbach *et al.* [5].

As from experiment, the rebinding dynamics described by $p(R,t)$ gives two separate time scales: a geminate phase on a 100 ps time scale and rebinding from the more distant Xe4 pocket on a 100 ns time scale [24]. This can be seen in Fig. 4. It is also found that the docking site is depopulated on a ns time scale and rebinding from the Xe4 pocket occurs in one step without stabilizing in the intermediate B state.

In conclusion, the present results give detailed information about the motion of small ligands in the neighborhood of the active site of Mb. Central to all investigations was the development of an accurate electrostatic model to describe higher moments of the ligand. This allows to capture energetic and dynamical effects and to compare results from simulations with experiment.

Acknowledgment

The authors thank the Swiss National Science Foundation for financial support. M.M. is a Förderungsprofessor of the SNSF.

Received: June 10, 2005

[1] M. Brunori, *Biochem.* **2000**, *30*, 221.
[2] M. Lim, T.A. Jackson, P.A. Anfinrud, J. *Chem. Phys.* **1995**, *102*, 4355.
[3] M. Lim, T.A. Jackson, P.A. Anfinrud, *Nat. Struct. Biol.* **1997**, *4*, 209.
[4] K. Nienhaus, J.S. Olson, S. Franzen, G.U. Nienhaus, *J. Am. Chem. Soc.* **2005**, *127*, 41.
[5] P.J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B.R. Cowen, D. Ehrenstein, H. Frauenfelder, J.B. Johnson, D.C. Lamb, S. Luck, J.R. Mourant, G.U. Nienhaus, P. Ormos, R. Philipp, A. Xie, R.D. Young, *Biochem.* **1991**, *30*, 3988.
[6] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, *J. Comp. Chem*. **1983**, *4*, 187.
[7] A.D. MacKerell, Jr., D. Bashford, M. Bellott, R.L. Dunbrack, Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586.
[8] D.R. Nutt, M. Meuwly, *Biophys. J.* **2003**, *85*, 3612.
[9] D.R. Nutt, M. Meuwly, *Proc. Natl. Acad. Sci.* **2004**, *101*, 5998.
[10] D.R. Nutt, M. Meuwly, *ChemPhysChem* **2004**, *5*, 1710.
[11] D.C. Lamb, K. Nienhaus, A. Arcovito, F. Draghi, A.E. Miele, M. Brunori, G.U. Nienhaus, *J. Biol. Chem.* **2002**, *277*, 11636.
[12] G.U. Nienhaus, K. Nienhaus, *J. Biol. Phys.* **2002**, *28*, 163.
[13] P. Banushkina, M. Meuwly, *J. Phys. Chem. B*, in print **2005**.
[14] J. Kottalam, D.A. Case, *J. Am. Chem. Soc.* **1988**, *110*, 7690.
[15] D. Bicout, A. Szabo, *J. Chem. Phys.* **1998**, *109*, 2325.
[16] P. Banushkina, M. Meuwly, *J. Chem. Theo. Comp.* **2005**, *2*, 208.
[17] D. Vitkup, G.A. Petsko, M. Karplus, *Nat. Struct. Biol.* **1997**, *4*, 202.
[18] J. Meller, R. Elber, *Biophys. J.* **1998**, *74*, 789.
[19] A. Ostermann, R. Waschipky, F.G. Parak, G.U. Nienhaus, *Nature* **2000**, *404*, 205.
[20] R. Elber, M. Karplus, *J. Am. Chem. Soc.* **1990**, *112*, 9161.
[21] C. Bossa, M. Anselmi, D. Roccatano, A. Amadei, B. Vallone, M. Brunori, A. Di Nola, *Biophys. J.* **2004**, *86*, 3855.
[22] F. Schotte, M. Lim, T.A. Jackson, A.V. Smirnov, J. Soman, J.S. Olson, G.N. Phillips Jr., M. Wulff, P.A. Anfinrud, *Science* **2003**, *300*, 1944.
[23] E.S. Park, S.S. Andrews, R.B. Hu, S.G. Boxer, *J. Phys. Chem. B* **1999**, *103*, 9813.
[24] A. Ansari, C.M. Jones, E.R. Henry, J. Hofrichter, W.A. Eaton, *Biochem.* **1994**, *33*, 5128.
[25] A. Szabo, K. Schulten, Z. Schulten, *J. Chem. Phys.* **1980**, *72*, 4350.
[26] J.N. Harvey, *J. Am. Chem. Soc.* **2000**, *122*, 12401.