CO migration in native and mutant myoglobin: Atomistic simulations for the understanding of protein function

David R. Nutt and Markus Meuwly[†]
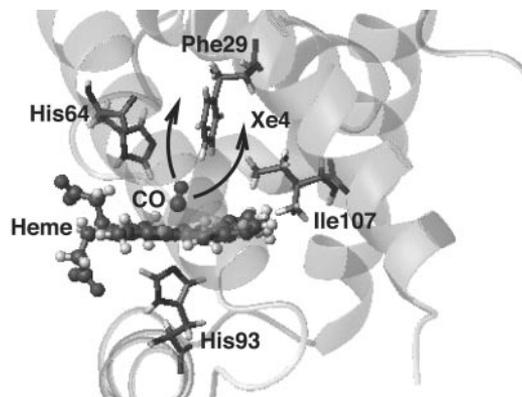
Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

Molecular dynamics simulations of the events after the photodissociation of CO in the myoglobin mutant L29F in which leucine is replaced by phenylalanine are reported. Using both classical and mixed quantum-classical molecular dynamics calculations, we observed the rapid motion of CO away from the distal heme pocket to other regions of the protein, in agreement with recent experimental results. The experimentally observed and calculated infrared spectra of CO after dissociation are also in good agreement. We compared the results with data from simulations of WT myoglobin. As the time resolution of experimental techniques is increased, theoretical methods and models can be validated at the atomic scale by direct comparison with experiment.

Fig. 1. Overview of the structure of the L29F mutant. The important residues surrounding the heme unit are drawn and labeled. The arrows indicate the movement of the CO ligand observed by Schotte *et al.* (1), either to a region between residues 64 and 29 or to a region behind Phe-29, which is also known as the xenon-4 pocket. The xenon-4 pocket actually extends in front of the plane of the figure. The figure was produced by using MOLMOL (26).

Recently, the time-resolved infrared spectrum and x-ray diffraction of photodissociated ^{13}CO in myoglobin (Mb) have been reported for a particular mutant, L29F (1). In this mutant the leucine (L) residue at position 29 is replaced by a phenylalanine (F) (see Fig. 1). This mutation introduces a bulky phenyl side chain into the distal heme pocket that can be expected to have an impact on the dynamics after photodissociation. It was previously found that the equilibrium constants for O_2 binding are larger by a factor of 15 in L29F compared to native Mb (2). This difference was interpreted as due to stabilizing interactions between the bound dioxygen and atoms of the phenyl ring. The infrared experiments, carried out at 283 K, found that the characteristic infrared absorption band for CO in the native Mb changes profoundly in the case of the L29F mutant (1). The well known bimodal distribution, indicative of the population of different binding sites [conformational substates B1 and B2, which differ in the orientation of the CO (3, 4), and possibly B3, which can only be populated by extended illumination in native MbCO (5)] is fundamentally altered. For L29F, the bimodal structure of the infrared spectrum disappears after ≈100 ps, whereas for native Mb it persists for 100 ns. This effect was explained by the escape of CO from the primary docking site within the distal heme pocket (Fig. 1) to neighboring sites within the protein on a time scale three orders of magnitude faster than in the case of native Mb (1).

Previous theoretical investigations of the ligand dynamics of CO in myoglobin focused mostly on the short time scale and have considered effects after dissociation of the ligand from the iron on the picosecond time scale (6, 7). Some of the investigations used a three-site model for the CO ligand (8). This model correctly describes the static dipole and quadrupole moment of CO. Recently, we developed a refined model that uses fluctuating charges on the C and O atom sites and at the center of mass (9). This model reproduces high-level *ab initio* calculations of the dipole and quadrupole moment over a wide range of intermolecular separations (10, 11) and allows the calculation of the infrared spectrum associated with the CO vibration. The model correctly describes the experimentally observed existence of the docking site within the distal heme pocket as well as the migration of the ligand through the protein to the xenon-4 pocket, which has been observed for various mutants (12, 13). The time scales and energy barriers for motion from the binding site to the docking site were also found to be in almost quantitative agreement with experiment. These calculations were based on multiple 1-ns trajectories that extensively sample the configurational space available to the dissociated CO molecule within the binding site (9).

The L29F mutant is an ideal system for which to investigate the infrared spectrum of a photodissociated ligand in more detail. Most importantly, the change in the infrared spectrum takes place on the subnanosecond time scale, which is accessible to meaningful computational investigations. Because the vibrational spectrum is experimentally known to change on a time scale of 10 ps, even mixed quantum mechanical/molecular mechanics simulations for photodissociated Mb + CO are possible. This situation is similar to ligand rebinding in MbNO (Mb + NO → MbNO), where the rebinding process occurs in the subnanosecond time regime, and thus is ideal for studying by computer simulations (14). In the latter case, part of the nonexponential behavior in the rebinding process could be attributed to protein relaxation (15).

Methods: Details of the Molecular Dynamics (MD) Simulations

All MD simulations were carried out with the CHARMM program (16) and the CHARMM 22 force field (17). Other parameters required to describe the heme–ligand interactions were taken from Kuczera *et al.* (18) and Meuwly and coworkers (15).

The computational setup follows a similar procedure to previous studies of the photodissociation of MbNO and MbCO (8, 9, 15). The native protein structure was taken from the x-ray structure of Kuriyan and coworkers (19), to which hydrogen atoms were added. Because the simulation is focused on the region surrounding the heme group, the stochastic boundary method was used to increase computational efficiency (20). The heme pocket was solvated by three sequential overlays of a 16-Å sphere of equilibrated water molecules, centered on the heme, and a solvent boundary potential with radius 16 Å was applied to constrain the water molecules. A "reaction region" of radius 12 Å centered on the heme was defined, inside which the system was propagated with Newtonian dynamics. The dynamics of the buffer region between 12 and 16 Å from the center was described by using Langevin dynamics. The system contained a total of 2,533 heme protein atoms, the CO ligand, and 179 water molecules, which were represented by a modified TIP3P potential (where the van der Waals interactions are shared between the H and O sites) (21). The nonbonded interactions were truncated at a distance of 9 Å by using a shift function for the electrostatic terms and a switch algorithm for the van der Waals terms. Friction coefficients of 62 ps^{-1} and 250 ps^{-1} were applied to the oxygen site of water and the remaining nonhydrogen atoms, respectively. After equilibration, the L29F mutation was made by deleting the side-chain atoms of Leu-29 and replacing them with a Phe residue. The structure was then further equilibrated to remove any strain or bad contacts present.

The photodissociation event was modeled by the "sudden" approximation, as used previously (8, 9, 15). The Fe–C bond is deleted and the potential parameters for the bound state are replaced by those for the dissociated state. A repulsive term of the form r-12 was added, and all nonbonded interactions between the CO and the heme "gate" were switched off. After 0.2 ps of dynamics, by which time the Fe–C bond is fully dissociated, the repulsive term was switched off and the nonbonded interactions were reintroduced. At the same time, the three-point fluctuating charge model for CO described above was activated.

Twenty-four 100-ps trajectories and five 1-ns trajectories of photodissociated MbCO were calculated, and the results are described below. In addition, one simulation for native Mb + CO and the L29F mutant was carried out by using a mixed quantum mechanics/molecular mechanics (QM/MM) scheme to compare to the results of the pure MM calculations. For this, the CO molecule was treated at the B3LYP/6–31G** level (QM part), whereas the Mb and the L29F mutant are described with the CHARMM 22 force field (MM part). These calculations are very time consuming, and therefore they were only carried out for 75 ps.

Results

We investigated the infrared spectrum for unbound ^{12}CO from the MD trajectories. To this end, the CO dipole moment was correlated over 2^{15} time origins, the Fourier transform of the dipole–dipole autocorrelation function was taken, and the resulting spectrum was weighted with the appropriate Boltzmann factor (22). Fig. 2 shows the infrared spectrum of several independent trajectories for dissociated CO for both forms of Mb. The spectrum extends essentially over the range of 2,170–2,200 cm^{-1}, except for two outlying peaks at lower frequency that could be attributed to unusual conformations of the protein in two (of twenty-four) 100-ps trajectories of the L29F mutant (off scale in Fig. 2). The overall spectra are shifted by ≈40 cm^{-1} from the experimental ^{12}CO spectrum. This shift is mostly due to the use of the time correlation function to evaluate the spectrum and to remaining uncertainties in the CO stretching potential that is influenced by the protein matrix (9). Use of the time-correlation function to calculate the infrared spectrum of ^{12}CO gives a fundamental frequency of 2,183 cm^{-1} compared to 2,144 cm^{-1}
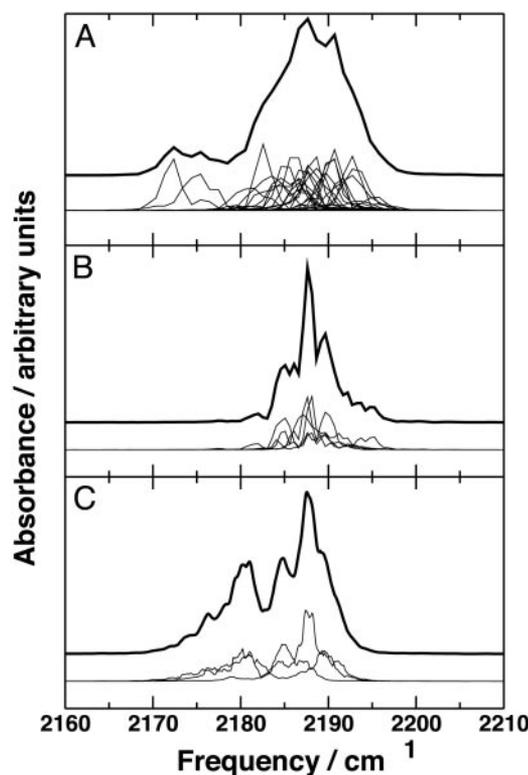

Fig. 2. Infrared spectra calculated from the simulations of photodissociated mutant (L29F) and WT MbCO. (*A*) Spectra from twenty-four 100-ps trajectories of L29F MbCO with a resolution of ≈1 cm^{-1}. (*B*) Spectra from five 1-ns trajectories of L29F with a resolution of ≈0.5 cm^{-1}. (*C*) Spectra from three 1-ns trajectories of WT MbCO with a resolution of ≈0.5 cm^{-1}. The thick line shows the average spectrum (enlarged and shifted vertically for clarity), and the thin lines show the individual spectra.

from the exact solution of the Schrödinger equation for the CO rotational Rydberg–Klein–Rees potential developed by Huffaker (23, 24) that is in almost perfect agreement with experiment (2,143.3 cm^{-1}) (25). Most importantly, the infrared spectrum appears for both, native and mutant Mb, in exactly the same frequency range. This finding is in agreement with an experiment where ^{13}CO was used (1).

In WT MbCO, the photodissociated CO molecules remain in the distal heme pocket (formed by residues 29, 43, 64, 68, and 107 and the heme) for >1 ns (3, 9). This behavior gives rise to the average spectrum shown in Fig. 2*C*. It consists of two main peaks (2,180 cm^{-1} and 2,188 cm^{-1}), separated by 8 cm^{-1}, which can be attributed to CO in the docking site at the side of the distal heme pocket (3, 9). The effect of temperature on the infrared absorption of CO is of the order of 1 cm^{-1}, whereas the mobility of the CO within the protein matrix is likely to be temperature-dependent and may also influence the absorption spectrum (9).

Infrared spectra of CO in L29F calculated from the twenty-four 100-ps trajectories are presented in Fig. 2*A*. The average spectrum exhibits two peaks, one broad and dominant, separated by ≈15 cm^{-1}. This finding corresponds favorably with the time-resolved spectra presented by Schotte *et al.* (1) taken between 1 and 100 ps that also show a dominant feature at higher wave numbers and a smaller peak ≈15 cm^{-1} below. The smaller peak disappears at longer observation times. Because Schotte *et al.* (1) observed a peak that decays with a time constant of 140 ps, it is conceivable that the corresponding peak in our calculated spectrum should be smaller, as is found to be the case. The spectra calculated from five 1-ns trajectories are shown in Fig. 2*B*. One single peak is obtained, in agreement with the exper-
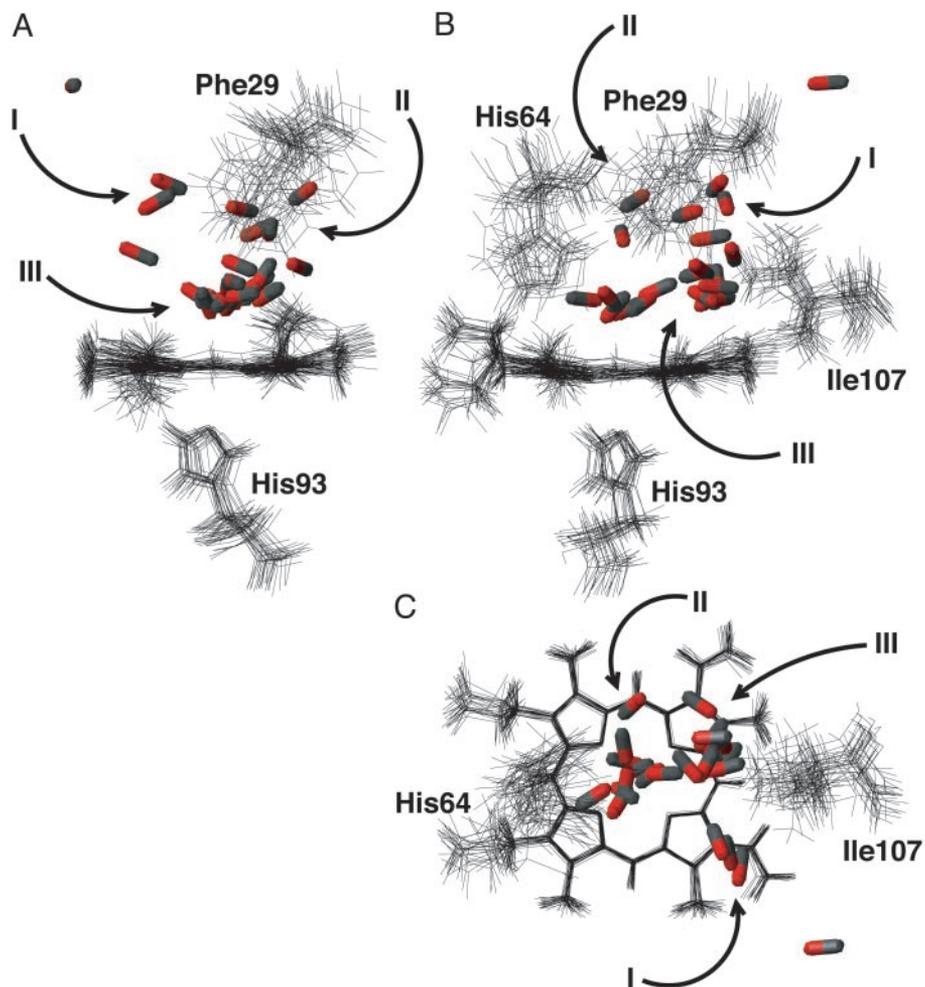
Fig. 3. Side (*A*), front (*B*), and top (*C*) views of the final structures obtained from the twenty-four 100-ps trajectories. Residues 29, 64, 93, and 107 are shown when they do not obscure the features of interest. The frames were fitted to the heme unit of the first structure and the rms deviation of the fit was minimized. It can be seen that a wide range of final configurations are obtained. Specific regions in which the CO ligands are found are labeled with roman numerals. They correspond to the xenon-4 pocket (I), the region between residues 29 and 64 (II), and the distal heme pocket (III). There is one CO ligand that is distant from the others, having left the xenon-4 pocket, probably through the gap between residues 28 and 69. The figure was produced by using MOLMOL (26).

imental study of Schotte *et al.* (1), who, at longer time scales (>100 ps), observed a single signal, which they assign to CO inside the protein matrix but not in the docking site at the edge of the distal heme pocket. The detailed structure of the individual spectra can be understood by carefully analyzing the trajectories as was done previously (9).

In the case of the L29F mutant, the CO molecule escapes from the distal pocket on a time scale of a few tens of picoseconds to several hundred picoseconds. This behavior is reflected by the wide range of configurations found at the end of the twenty-four 100-ps trajectories (Fig. 3). The mutation at residue 29 sufficiently alters the nature of the docking site at the edge of the distal heme pocket to induce this change. It can be seen that CO molecules access the same regions as found in the time-resolved x-ray diffraction study of Schotte *et al.* (1). In particular, we observed CO molecules moving to the xenon-4 pocket and to a region between residues 29 and 64. Some CO molecules remain within the distal heme pocket, indicating that in the simulations the time constant for escape from the distal pocket is of the order of 100 ps. This time scale is comparable to 140 ps from the analysis of the experiments and considerably shorter than 190 ns found for native Mb (1). We also find one CO molecule, which travels beyond the xenon-4 pocket to a region surrounded by residues 24, 28, 114, and 115 (between the B and G helices). The

distribution of the CO positions 100 ps after dissociation is significantly altered if the three-point fluctuating charge model is replaced with the fixed charge model of Straub and Karplus (8). The CO ligand is then found to explore regions of space not observed by Schotte *et al.* (1), demonstrating the sensitivity of the results to modeling the electrostatic interactions within the distal heme pocket.

The MD(QM/MM) calculations for dissociated CO from native and mutant Mb, respectively, corroborate the present findings. In Fig. 4, two representative trajectories of 75 ps for both the native MbCO (*a–c*) and the L29F mutant MbCO (*d–f*) are shown. Six frames for the heme and residues 29, 64, 93, and 107 separated by 12.5 ps and fitted to the heme unit are overlaid, whereas for the CO molecule 300 configurations separated by 0.25 ps are shown. As in Fig. 3, residues obstructing the view are omitted. For native MbCO, the CO molecule does not escape from the distal heme pocket on the time scale of the simulation (Figs. 4 *a* and *b*). Trajectories starting from a different dissociation run confirm this result. The dissociated CO molecule extensively samples the binding pocket, in accord with previous MD simulations (9) and with the experimental work (3, 4). In the case of the L29F mutant (Fig. 4 *d* and *e*), the CO molecule leaves the primary binding site after 30 ps. This result is confirmed by multiple simulations starting from different dissociation trajectories. It is also of interest to calculate the infrared
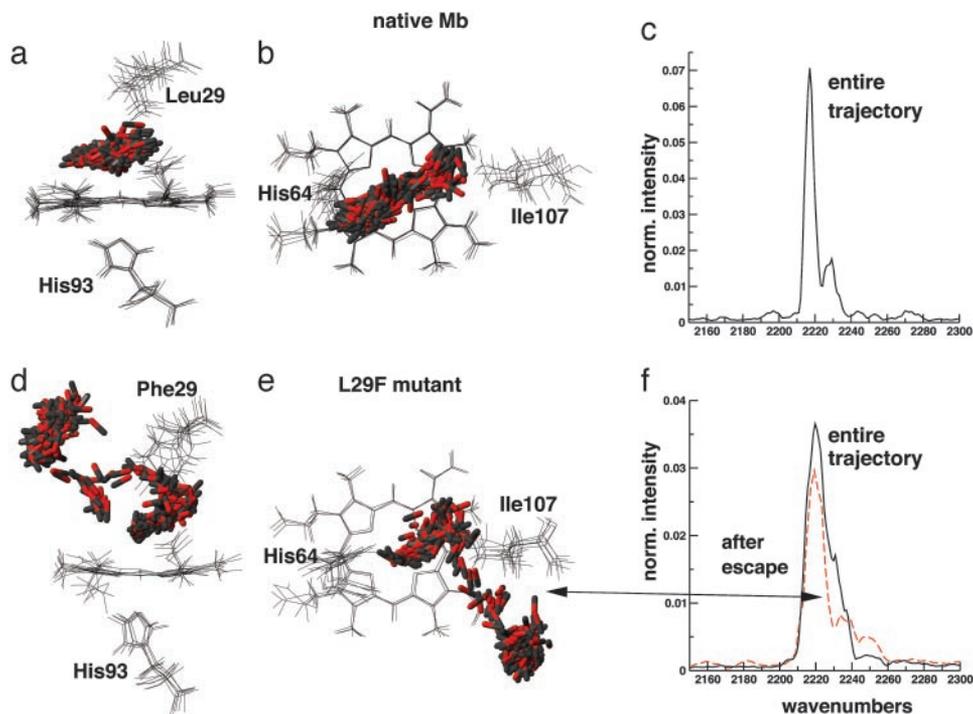
Fig. 4. Side (*a*) and top (*b*) views of six structures along a 75-ps MD(QM/MM) trajectory of CO in native Mb. The CO molecule is described by B3LYP/6–31G**. The frames are fitted to the heme unit and the six frames shown are separated by 12.5 ps. In addition, the 300 CO positions (separation of 0.25 ps) along the same trajectory are shown. In *d* and *e*, equivalent views are shown for the L29F mutant. (*c* and *f*) Corresponding infrared spectrum calculated from the dipole moment autocorrelation function. All spectra are normalized to unity. The peaks in *c* are separated by 11 cm^{-1}. The figure was produced by using MOLMOL (26).

spectrum from these MD(QM/MM) trajectories (Fig. 4 *c* and *f*). The main peaks in both cases are centered at around 2,215 cm^{-1}. The overall shift of 80 cm^{-1} is mainly due to the level of theory used in the present QM calculations (B3LYP/6–31G**) and the use of the dipole–dipole autocorrelation function to calculate the infrared spectrum. As was found for the pure MD simulations with the fluctuating three-point charge model, the peaks for native MbCO are split by 11 cm^{-1}, which is in turn in quite good agreement with experiment. For the L29F mutant, the infrared spectrum has a different shape. Because this trajectory samples two largely different regions within the protein matrix, the spectrum can be evaluated over the entire trajectory (0–75 ps) or only over the time interval during which the CO molecule is outside of the distal heme pocket (30–75 ps) (Fig. 4*f*). The spectrum over the entire trajectory is essentially featureless and broad and extends over ≈30 cm^{-1}. After leaving the primary binding site, the spectrum broadens further (i.e., the feature ≈2,250 cm^{-1} gains intensity relative to the other peaks). Both spectra in Fig. 4*f* have little in common with the spectrum in Fig. 4*c*. However, it is worthwhile to emphasize that the spectra have their maxima at essentially the same wave number, in agreement with experiment. The MD(QM/MM) calculations thus confirm the findings from the much longer MD/MM simulations and are also in agreement with experiment.

The present work used two different approaches to describe a CO molecule within the heme pocket of native and mutated (L29F) Mb. With the fluctuating point-charge model, long MD simulations can be run and analyzed, whereas the MD(QM/MM) calculations are computationally more demanding and more thorough in treating the inter- and intramolecular interactions. Both approaches convey a similar picture: In native Mb the CO molecule never/very rarely escapes the heme pocket on the time scale of the simulations [75 ps for MD(QM/MM) and 1–3 ns for MD/MM], whereas in the L29F mutant the CO escapes on a timescale of several 10 ps after dissociation. The trajectories obtained depend on the model or level of theory used, but their agreement with experiment gives confidence in their value. The corresponding infrared spectra 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, whereas for the mutant, largely unstructured spectra are found from the simulations. These results are in agreement with recent time-resolved infrared experiments and previous investigations of the CO motion within the heme pocket and confirm the utility of the fluctuating three-point model to describe the dynamics of dissociated CO within the Mb protein matrix. This study shows how experimental and theoretical advances can stimulate each other in the quest to understand the functioning of proteins at an atomistic level.

1. Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N., Jr., Wulff, M. & Anfinrud, P. A. (2003) *Science* **300,** 1944–1947.
2. Carver, T. E., Brantley, R. E., Jr., Singleton, E. W., Arduini, R. M., Quillin, M. L., Phillips, G. N., Jr., & Olson, J. S. (1992) *J. Biol. Chem.* **267,** 14443–14450.
3. Lim, M., Jackson, T. A. & Anfinrud, P. A. (1995) *J. Chem. Phys.* **102,** 4355–4366.
4. Lim, M., Jackson, T. A. & Anfinrud, P. A. (1997) *Nat. Struct. Biol.* **4,** 209–214.

5. Nienhaus, G. U., Mourant, J. R., Chu, K. & Frauenfelder, H. (1994) *Biochemistry* **33,** 13413–13430.
6. Meller, J. & Elber, R. (1998) *Biophys. J.* **74,** 789–802.
7. Vitkup, D., Petsko, G. A. & Karplus, M. (1997) *Nat. Struct. Biol.* **4,** 202–208.
8. Straub, J. E. & Karplus, M. (1991) *Chem. Phys.* **158,** 221–248.
9. Nutt, D. R. & Meuwly, M. (2003) *Biophys. J.* **85,** 3612–3623.
10. Maroulis, G. (1996) *J. Phys. Chem.* **100,** 13466–13473.
11. Maroulis, G. (2001) *Chem. Phys. Lett.* **334,** 214–219.

12. Lamb, D. C., Nienhaus, K., Arcovito, A., Draghi, F., Miele, A. E., Brunori, M. & Nienhaus, G. U. (2002) *J. Biol. Chem.* **277,** 11636–11644.
13. Nienhaus, G. U. & Nienhaus, K. (2002) *J. Biol. Phys.* **28,** 163–172.
14. Petrich, J. W., Lambry, J.-C., Kuczera, K., Karplus, M., Poyart, C. & Martin, J.-L. (1991) *Biochemistry* **30,** 3975–3987.
15. Meuwly, M., Becker, O., Stote, R. & Karplus, M. (2002) *Biophys. Chem.* **98,** 183–207.
16. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4,** 187–217.
17. 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., *et al.* (1998) *J. Phys. Chem. B* **102,** 3586–3616.
18. Kuczera, K., Kuriyan, J. & Karplus, M. (1990) *J. Mol. Biol.* **213,** 351–373.
19. Kuriyan, J., Wilz, S., Karplus, M. & Petsko, G. (1986) *J. Mol. Biol.* **192,** 133–154.
20. Brooks, C. L., III, & Karplus, M. (1983) *J. Chem. Phys.* **79,** 6312–6325.
21. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983) *J. Chem. Phys.* **79,** 926–935.
22. McQuarrie, D. A. (1976) *Statistical Mechanics* (Harper & Row, New York).
23. Huffaker, J. N. (1976) *J. Chem. Phys.* **64,** 3175–3181.
24. Huffaker, J. N. (1976) *J. Chem. Phys.* **64,** 4564–4570.
25. Ewing, G. E. (1962) *J. Chem. Phys.* **37,** 2250–2256.
26. Koradi, R., Billeter, M. & Wüthrich, K. (1996) *J. Mol. Graphics* **14,** 51–55.