Structural Assignment of Spectra by Characterization of Conformational Substates in Bound MbCO

Michael Devereux and Markus Meuwly*
Department of Chemistry, University of Basel, Basel, Switzerland

ABSTRACT Residue motions of the distal heme pocket and bound CO ligand of carbonmonoxy Myoglobin are studied using a combination of molecular dynamics simulations and quantum chemical methods. Using mixed quantum mechanics/molecular mechanics calculations together with sampling from molecular dynamics simulations (QM/MM(MD)), the experimentally observed spectroscopic A_0 and A_1 substates of the bound CO ligand are assigned to the open and closed conformation of His^{64} and the His^{64}_{ε} tautomer, respectively. Several previously proposed origins of the A_3 substate, including rotamers of the doubly protonated $His^{64}H^+$ side chain, $His^{64}H^+$ inside the distal pocket, and cooperative motions with Arg^{45}, are investigated with QM/MM(MD). However, the signatures of the calculated infrared spectra do not agree with the experimentally observed ones. For additional insight on this, extensive molecular dynamics simulations are used together with improved electrostatics for the bound ligand. A CO fluctuating charge model is developed to describe the ab initio dipole and quadrupole moments of the bound ligand. CO absorption spectra are then obtained directly from the dynamics simulations. Finally, the electrostatics of the heme pocket is examined in detail in an attempt to determine the structural origins of the observed spectroscopic A-states from MD simulations. However, contrary to related simulations for unbound CO in myoglobin, the shifts and splittings for carbonmonoxy Myoglobin are generally small and difficult to relate to structural change. This suggests that coupling of the CO motion to other degrees of freedom, such as the Fe-CO stretching and bending, is important to correctly describe the dynamics of bound CO in myoglobin.

INTRODUCTION

Myoglobin (Mb) is a small globular protein (153 residues) and one of the primary model systems for studying protein structure and dynamics in general, and ligand binding and dynamics in particular ((1); and see Fig. 1 for Mb protein structure). The role of Mb in storing oxygen within muscle tissue gives the system strong biological relevance. Given its fundamental role in biology and biophysics, an abundance of experimental data is available, ranging from detailed crystallographic studies (2) to infrared (IR) (3–5), Raman (6), nuclear magnetic resonance (7), and vibrational echo (8) experiments. Owing to its relatively small size and the short timescale of a number of fundamentally important processes, Mb has also become a tractable and meaningful system for computational studies (9–15).

One of the more recent successes in understanding the spectroscopy and dynamics in Mb in more detail is the assignment of the conformational substates associated with the split infrared spectrum of unbound carbon monoxide (CO) after photodissociation from the heme (16–19). In earlier work (11), vibrational relaxation of CO in the same photodissociated system was also successfully modeled, with good agreement in a range of measured dynamic properties of the excited CO molecule with experimental results. Merchant et al. (8) have successfully used MD simulations to explain the dephasing dynamics of experimental vibrational echo results for bound MbCO in terms of conformations of the adjacent His^{64}. Franzen (20) used ab initio calculations of CO Stark shifts in Mb to calibrate a model to predict CO vibrational frequency from density functional theory (DFT), bond lengths and Mulliken charges. The model was used to study the effects of nonbonded interactions in a series of Mb mutants on CO vibrational frequency, providing good agreement with experimentally measured frequency shifts.

Despite the considerable progress that has been made, interpretation of the bound states has been difficult and current knowledge is incomplete. Our aim is to apply previously successful methods used for unbound CO to bound CO. We focus on the three main spectroscopic A-states visible in IR spectra. The principal absorption bands are denoted A_0, A_1, and A_3, with vibrational frequencies $\nu(A_0) \approx 1965\ cm^{-1}$, $\nu(A_1) \approx 1949\ cm^{-1}$, and $\nu(A_3) \approx 1933\ cm^{-1}$, respectively (21). The missing A_2 band identified in earlier works is now commonly incorporated into A_1. An additional substate, denoted A_x, was recently identified by Müller et al. (22), but shows low occupation under physiological conditions. Rapid interconversion between substates is exhibited on the nanosecond timescale between A_1 and A_3, and on the microsecond timescale between A_0 and $A_{1,3}$ (4).

The A_0 substate is the best characterized of all observed A-states in the literature. Experimental data have shown that the A_0 IR absorption peak increases significantly in intensity at low pH (6), and structural data indicate that the His^{64} side chain swings out of the distal pocket and into the solvent, leading to an open conformation (2). Mutation studies with the polar His^{64} side chain replaced by an apolar substitute such as Val or Met also show absorptions clustered

Submitted October 8, 2008, and accepted for publication January 21, 2009.

*Correspondence: m.meuwly@unibas.ch

Editor: Steven D. Schwartz.

0006-3495/09/06/4363/13 $2.00

doi: 10.1016/j.bpj.2009.01.064

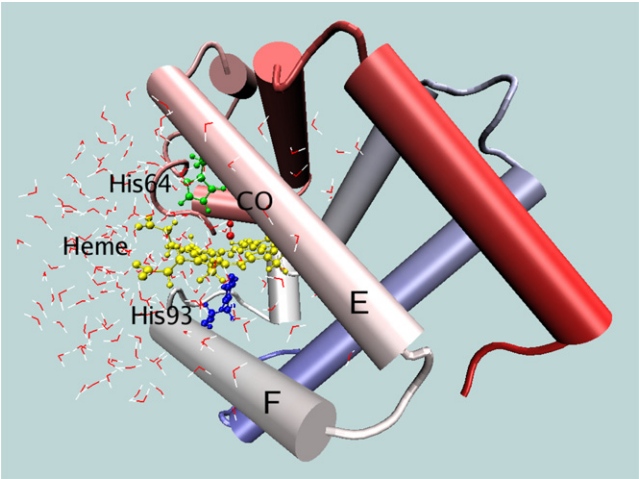

FIGURE 1 Myoglobin protein structure highlighting important features. The heme group is marked in yellow, the CO ligand in red, His^{64} in green, and His^{93} in blue. Helices are represented by cylinders; the E and F helices are labeled.

around the band at the A_0 position (6,23), indicating that the A_0 state corresponds to little or no interaction between the polar histidine imidazole group and the bound CO ligand.

An early explanation for the origin of the splitting observed between the A_1 and A_3 substates was a change in CO bond strength as a function of the Fe-C-O angle. This explanation was later discarded (24), in favor of an interpretation that the A_1 and A_3 substates result from interactions of the CO ligand with the adjacent His^{64} residue (5,21,25). A consensus is emerging toward dominance of His_{ε}^{64} as the dominant tautomerization state of the distal histidine His^{64} (5,18,21,26). Interaction between the $N_{\varepsilon}H$ proton and oxygen atom of the bound CO, perhaps via a hydrogen bond (27), then causes a red shift in the vibrational frequency of the bound ligand relative to the environment of the A_0 substate. However, the precise structural origin of the A_1 and A_3 substates is still unclear. Interconversion between the His_{ε}^{64} and His_{δ}^{64} tautomers of the His^{64} residue, in addition to producing predicted blue-shifted CO stretch frequencies, is inconsistent with the observed rapid A_1-to-A_3 interconversion rate of $\approx$1 ns (4) or less (28). This process would require the residue to first swing into the solvent, which leads to the A_0 substate and prevents direct A_1-to-A_3 interconversion. The timescale is accessible to MD simulations, and is suggestive of a conformational change between structurally related substates. Indeed, Vojtechovsky et al. (29) obtained a number of high resolution x-ray structures containing distinct protein conformations that indicate the A_1 and A_3 substates correspond to structures with His^{64} $N_{\varepsilon}\cdots$CO oxygen distances of 3.2 and 2.7 Å, respectively. Phillips et al. (5) proposed that rotation of the His_{ε}^{64} side chain toward the bound CO ligand would cause an increase in the electrostatic potential at the position of the oxygen atom that may account for the increased shift of the A_3 frequency. No origin for an energy barrier that would trap the side chain in this close-contact conformation was given. New vibrational echo results (28), however, indicate this barrier may originate from a wider conformational change taking place in the protein on a timescale of ~50 ps. Shifting of the E helix was proposed as the origin of the change in His_{ε}^{64} side-chain position, providing the observed rapid A_1, A_3 flipping. Schulze and Evanseck (30) suggested a different mechanism involving cooperative steric motion between His^{64} and Arg^{45}. The results of earlier multidimensional vibrational echo experiments (26) combined with MD simulations had also indicated that the substates represent rotamers of the His_{ε}^{64} imidazole side chain. No confirmation has yet been made by linking these rotamers to further experimental or theoretical data. The continued interest and quite divergent views on the structural origins of the spectroscopic A-states were the motivation to investigate the underlying processes in more detail.

In this work, atomistic MD simulations with refined electrostatic models or combined with DFT calculations are used to explore the structural origins of the A_0, A_1, and A_3 states. Previously it was found that both (18) quantum mechanics/molecular mechanics (molecular dynamics) (QM/MM(MD)) and MD simulations employing higher-order multipole models (19) are suitable to better characterize the dynamics and spectroscopy of photodissociated (unbound) CO in the active site of Mb. In the QM/MM(MD) simulations, the trajectories from atomistic simulations are used to identify distinct conformational substates of MbCO. The substates are then characterized in more detail using DFT. Because such a procedure is computationally demanding, there is insufficient sampling of conformational space to provide statistically representative vibrational frequency distributions. Improved electrostatic models allow greater conformational sampling, providing additional insight, so refined models are developed here for the bound CO ligand. First, the computational methods used are described. Then, results from QM/MM(MD) simulations and from refined charge models are discussed. Finally, the results are put into perspective relative to experimental data and other computational work.

THEORETICAL METHODS

MD simulations

MD simulations were carried out using the CHARMM program (31), and the CHARMM22 force field (32). The MbCO structure was taken from the x-ray study of Kuriyan et al. (33) (PDB No. 1mbc), with hydrogen atoms added within CHARMM. For this work, the protonation state of the His^{64} residue is essential, and simulations with all three possibilities (protonation at the N_{δ} position (His_{δ}^{64}), N_{ε} position (His_{ε}^{64}), and double protonation (denoted $His^{64}H^{+}$)) were carried out. Since the simulation is focused on the region surrounding the heme group, the stochastic boundary method was used to increase computational efficiency (34). 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 using Langevin dynamics. This setup resulted in a total of 2534 protein and ligand atoms and

TABLE 1 Charge parameters used in CO electrostatic models I, II, and III

	q_C	q_O	q_{COM}
Model I	0.02	− 0.02	0.0
Model II	$0.444R - 0.367$	$-0.49R - 0.367$	0.0
Model III	$0.82R - 1.067$	$0.468R - 0.995$	$q_C + q_O - 0.08$

Carbon, oxygen, and center-of-mass charges (COM) are expressed as a function of CO bond length R. Distances are in Angstrom, charges are in atomic units.

~178 water molecules, although the exact number of water molecules varied slightly depending on the protein conformation at the time of solvent overlay. Water molecules were treated with a TIP3P potential (35). The nonbonded interactions were truncated at a distance of 9 Å 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 water-molecule oxygen atoms and remaining nonhydrogen atoms in the remainder of the system, respectively. The entire system was equilibrated at 300 K for 90 ps. Trajectories were typically 1 ns in length, and generally at least six such trajectories were averaged to adequately sample a particular conformational substate.

The CO bond potential used in the simulations is a Morse potential of the form

$$V(r) = D_e(1 - \exp(-\beta(r - r_e)))^2, \quad (1)$$

where D_e represents the well depth (260.6 kcal/mol), r_e is the equilibrium bond length (1.128 Å), and the parameter β was set to 2.046 $Å^{-1}$. CO vibrational frequencies obtained from MD simulations using these parameters were generally between 1950 and 1975 cm^{-1}, depending on the charge model applied to the ligand atoms, lying near the experimental A_0 and A_1 values of 1949 and 1965 cm^{-1}, respectively. Three different electrostatic models for the CO ligand were explored. Model I used CHARMM22 CO charges (Table 1). Model II used fluctuating charges on the C and O atoms that correctly capture the fragment dipole moment of bound CO at different bond lengths. The dipole moment was defined as the dipole moment of the atoms-in-molecules (AIM) (36,37) C and O atoms as partitioned from the geometry-optimized molecular charge density of a heme-His^{93}-CO model system at the B3LYP/6-311+G(2d,p) level of theory. All AIM analysis was performed with a modified, in-house version of the MORPHY98 program (38). Atomic integration errors of $\sim L(\Omega) = 0.001$ atomic units or less were obtained, as measured by the integrated sum of the Laplacian of the electron density over the quadrature of each atomic basin, Ω. (39) Model III used an additional charge site on CO at the ligand center of mass, and again used a fluctuating charge, this time to reproduce the total fragment charge, dipole, and quadrupole moments along the CO axis.

IR spectra $C(\omega)$ from the MD simulations are calculated from the Fourier Transform of the real-time dipole-dipole autocorrelation function, $C(t)$ (40). To construct $C(t)$, the CO dipole moment is recorded at 2^n consecutive time steps, typically corresponding to the first or last 524 ps (2^{19} 1 fs steps) of a 1-ns trajectory. Where fluctuating charge models are used for the CO ligand, the CO dipole moment was used directly to construct $C(t)$. The resulting data are then transformed via Fourier transform with a Blackman filter (41). The IR absorption spectrum is calculated using

$$A(\omega) = \omega(1 - \exp(-\hbar\omega/(kT)))C(\omega), \quad (2)$$

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

Quantum chemical calculations

All electronic structure calculations were performed using the Gaussian 03 suite of programs (42) and DFT with the hybrid B3LYP functional. The heme group, bound CO, distal His^{64}, and proximal His^{93} residue were included explicitly in each DFT calculation. His^{64} was represented by its imidazole side chain, capped at the connection point with a methyl group. A double-ζ VDZ basis set was used throughout for the central CO ligand, Fe, and N atoms, whereas a smaller 3-21G basis set was applied to the remainder of the QM system. Remaining protein and solvent atoms were represented as point charges from the CHARMM22 force field placed at the respective atomic coordinates. A procedure used in a recent study (18) on photodissociated CO in Mb was then followed to obtain the fundamental vibration frequency from the one-dimensional Schrödinger equation: The CO bond energy was sampled at eleven 0.025 Å intervals of r, centered at the equilibrium distance r_e. The corresponding V_r stretching energies were fitted to a Morse function where β and D_e were allowed to vary, and the LEVEL (43) code was used to calculate the fundamental vibration. Vibrational frequencies calculated in this way allow quantitative estimation of vibrational frequency shifts from a DFT bond energy profile.

RESULTS AND DISCUSSION

The results are discussed in three separate subsections. First, results from mixed QM/MM(MD) simulations are described. Second, the electrostatic model for the heme-unit is analyzed in detail. Thirdly, the electrostatics of the CO ligand is examined and extensive MD simulations are used to assess the influence of more detailed heme and CO models on the calculated IR frequency using conformational sampling.

Mixed QM/MM(MD) simulations

The A_0 state

The open conformation associated with the A_0 vibrational substate at low pH was first observed in a 1-ns MD trajectory with $His^{64}H^+$ starting inside the distal pocket. QM calculations were used to obtain CO vibrational frequencies for 40 snapshots taken from the trajectory as described in Theoretical Methods. During the simulation, the $His^{64}H^+$ protonated imidazole side chain was observed to swing away from the ligand and into the solvent, as also observed in the low-pH x-ray structure (2). The change takes place after just 250 ps in the MD trajectory (Fig. 2), and the side chain remains in the open conformation. Similar conformational

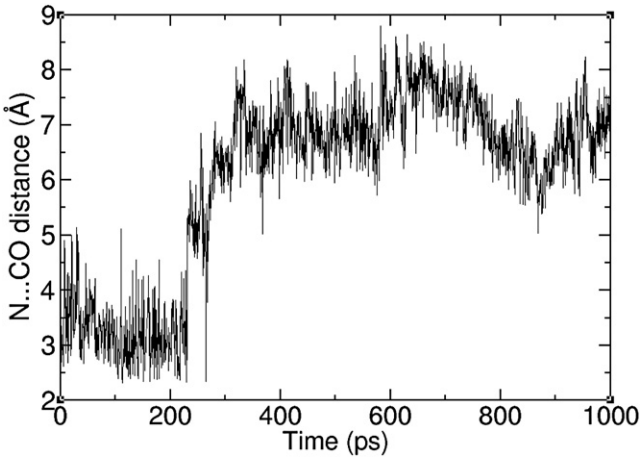

FIGURE 2 $N_\varepsilon\cdots CO$ distance (Å) as a function of time for the $His^{64}H^+$ residue. The side chain swings out of the distal pocket and into the solvent after ~250 ps.

changes on the nanosecond timescale were obtained from additional MD trajectories (see below). To ensure that the opening of the His^{64} gate was induced by protonation of the side chain, the open conformation was used as the starting point for a further 1-ns trajectory with neutral His^{64}_{ε}. The open conformation did not persist in the new simulation, and the His^{64}_{ε} side chain rapidly returns to a stable position inside the active site, resembling the original crystal structure. A histogram of 20 CO stretch frequencies from snapshots with $His^{64}H^{+}$ in the open conformation was used to locate the position of the QM/MM(MD) A_0 absorption peak, which is found at ~2103 cm^{-1}, and serves as a reference for all subsequent calculations. Vibrational coupling is not fully incorporated when calculating stretch frequencies in our DFT calculations, so reliable relative frequency shifts are obtained but absolute values are systematically overestimated. We therefore report only frequency shifts relative to the assigned A_0 frequency throughout the remainder of the text.

Further evidence for the assigned A_0 position was then obtained by creating a simple mutation model. The atoms of the His^{64} imidazole side chain were deleted, and the fundamental vibration frequency calculated as before. The snapshot chosen initially had His^{64}_{ε} inside the distal pocket, so the background electric field originating from the protein resembles that of the closed conformation. After the side chain has been removed, the system is analogous to an MbCO mutant with His^{64} replaced by a small, apolar residue such as alanine. The CO stretch frequency without His^{64} was calculated to be (red) shifted by just -5 cm^{-1} from the peak of the open conformation. With the side chain added back into the calculation, the vibrational frequency is red-shifted significantly further to -15 cm^{-1}. The experimentally observed increase in intensity at the frequency of the A_0 absorption peak relative to the $A_{1,3}$ peaks both after protonation at low pH and after His^{64} mutation to an apolar substitute therefore agrees with QM/MM(MD) results. A QM/MM(MD) value for the absorption band of the A_0 substate has been obtained, and can be used as a reference to identify and characterize the remaining A_1 and A_3 substates from further MD trajectories.

His^{64} tautomerization state in A_1 and A_3 substates

The next issue addressed was the tautomerization state of the His^{64} side chain. Although consensus points increasingly toward His^{64}_{ε} (5,20,26), some recent studies (25,44) continue to use the His^{64}_{δ} tautomer. To compare both tautomers, two separate 1-ns trajectories were run starting from similar MbCO structures with either His^{64}_{ε} or His^{64}_{δ}. Forty snapshots were taken at regular time intervals from each trajectory, and DFT stretch frequencies were calculated. A histogram of all calculated CO frequencies is shown in Fig. 3. The resulting peaks are red-shifted by -12 cm^{-1} from the A_0 position for the His^{64}_{ε} tautomer, and $+16$ cm^{-1} blue-shifted for the His^{64}_{δ} tautomer. The experimental IR spectrum shows that

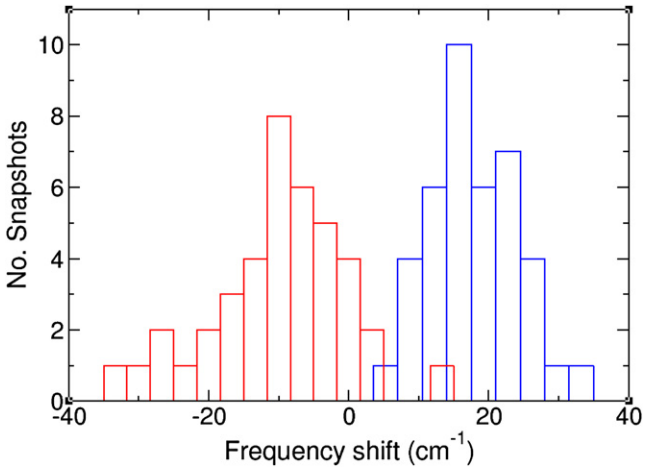

FIGURE 3 (*Red histogram*) DFT frequency shifts relative to the A_0 position for 40 snapshots of a 1-ns His^{64}_{ε} trajectory. (*Blue histogram*) Frequency shifts for 40 snapshots of His^{64}_{δ} trajectory. The His^{64}_{ε} trajectory gives red-shifted (negative) frequencies, whereas the His^{64}_{δ} trajectory gives blue shifts.

both the A_1 and A_3 substates should be red-shifted from the A_0 substate, indicating that the His^{64}_{δ} form does not significantly contribute. Furthermore, the -12 cm^{-1} red shift for the His^{64}_{ε} protonation state is in fair agreement with the experimentally observed red shift of -16 cm^{-1} between the A_0 and A_1 substates.

Rotamers of the His^{64}_{ε} side chain as a possible origin for the A_3 substate

After the identification of candidates for both the A_0 and A_1 substates, attempts were made to characterize the remaining A_3 substate. In vibrational echo experiments (8), the A_1 and A_3 substates were attributed to two rotamers of the imidazole side chain of His^{64}_{ε}. The first rotamer, R_{ε}, corresponds to the closed conformation observed in the His^{64}_{ε} trajectories (see above). The second rotamer, B_{ε}, with the imidazole side chain rotated ~180° about the C_{β}-C_{γ} bond, had not been observed in any of the previous trajectories. Energy optimization of the B_{ε} conformation gave a new stable rotamer of the His^{64}_{ε} side chain with the CO oxygen atom in contact with the atoms of the face of the imidazole ring (Fig. 4). This new conformation was found to persist for at least 1 ns of MD simulation. Twenty snapshots were taken from this trajectory, and corresponding CO stretch frequencies were calculated. The resulting calculated vibrational spectrum exhibited a single, broad peak, with frequencies ranging from -6 cm^{-1} red-shifted from the A_0 position to $+24$ cm^{-1} blue-shifted, with the maximum of the distribution at around the A_0 frequency. Because experiment predicts a 32 cm^{-1} red shift of the A_3 peak from the A_0 position, but a broad peak containing predominantly blue-shifted values is visible created from the DFT snapshots, it is unlikely that this rotated conformational state represents the missing A_3 substate. The conformer may, however, contribute to the visible A_0 absorption peak.

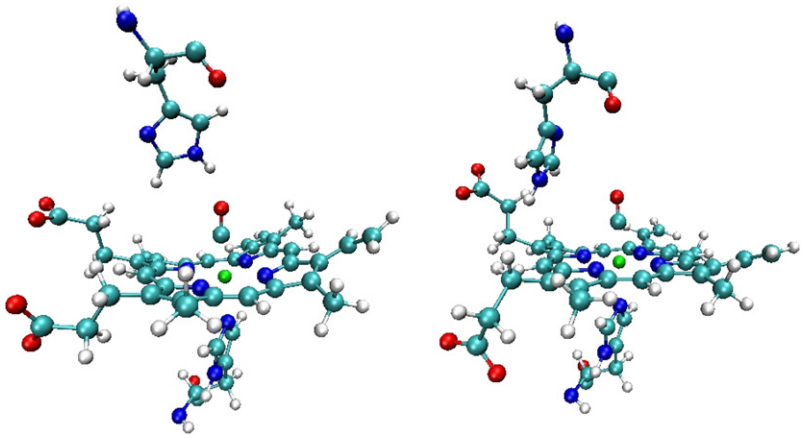

FIGURE 4 Snapshot from trajectory using standard (x-ray) His_{ε}^{64} conformer (*left*) and rotated conformer (*right*).

Proton transfer as a possible origin for the A_3 substate

Larger red shifts of the CO stretch frequency can originate from stronger (positive) local electric fields around the CO ligand. Although the preference for the $His^{64}H^{+}$ side chain is to swing away from the bound ligand and out into the solvent, it seemed feasible that the A_3 substate could originate from interaction with the protonated side chain while it was still in the closed conformation. The doubly-protonated $His^{64}H^{+}\cdots CO$ interaction was therefore examined more closely. To reduce the propensity of $His^{64}H^{+}$ to swing out of the distal pocket, 1 ns of MD simulation was carried out at 273 K, and 40 snapshots were taken from the first 800 ps of the trajectory. For each of them the corresponding CO stretch frequency was calculated. Visual inspection of the trajectories reveals that the $His^{64}H^{+}$ side chain inside the distal pocket is pulled toward the negative charge of the propionates, which differs from the conformations of the charge neutral His_{ε}^{64} trajectories. The corresponding histogram of CO stretch frequencies exhibits a broader range of values, but the main peak is red-shifted by only -12 cm^{-1} from the A_0 frequency, which does not correspond to the A_3 position. Close $N_{\varepsilon}H\cdots CO$ contacts result in more strongly red-shifted values at ~ -32 cm^{-1}, but configurations corresponding to the average red-shifted frequencies at ~ -12 cm^{-1} are still the most occupied. Overall, the results are similar to the charge neutral His_{ε}^{64} results in Fig. 3. It is therefore unlikely that $His^{64}H^{+}$ inside the pocket is responsible for the missing A_3 IR absorption band.

It would also be difficult to reconcile $His^{64}H^{+}$ interaction with CO as a source of the A_3 peak with experimental evidence. Firstly, the considerable increase in intensity of the A_0 absorption peak relative to the A_1 peak at low pH was not seen for the A_3 substate, which increases only slightly in intensity relative to A_1 before essentially disappearing at low pH (3,6). A larger relative increase might be expected for a mechanism involving His^{64} proton acquisition from the solvent. Secondly, simulations suggest that the open A_0 conformation is favored by $His^{64}H^{+}$, but the IR A_3 peak is significantly more intense than the A_0 peak. Finally, regular acquisition and subsequent loss of a proton from the N_{δ} atom of His_{ε}^{64} under experimental conditions would be necessary to provide the observed rapid A_1, A_3 interconversion rate. Intramolecular proton transfer from an adjacent residue seems the most likely mechanism, but the only obvious source is the adjacent Arg^{45} side chain. The acidic protons of Arg^{45}, however, interact strongly with the negatively charged propionate groups of the heme and are generally unavailable to interact with His^{64}. Calculations carried out on a simple model system also suggest proton transfer from Arg^{45} to His_{ε}^{64} is energetically unfavorable. In these calculations, a model system with $His_{\varepsilon}64$ side chain cleaved at the C_{β} position and Arg^{45} side chain cleaved at the C_{δ} position was oriented so that a proton of the terminal $-NH_2^{+}$ group of Arg^{45} pointed toward the His^{64} N_{δ} position with a $N_{\delta}\cdots H$ separation of 2 Å. The N-H bond was then optimized with the proton joined to the Arg^{45} side chain at the B3LYP 6-311+G(2d,p) level of theory, and again with the proton joined to His^{64} N_{δ}. The total B3LYP 6-311+G(2d,p) energy difference with the proton attached to $Arg^{45}$and His_{ε}^{64} is $\approx$19 kcal/mol, making proton transfer from Arg to His unlikely.

The Role of Arg^{45}

Schulze and Evanseck (30) suggested a cooperative motion between His_{δ}^{64} and nearby Arg^{45} as the origin of the A_3 substate. Use of the His_{δ}^{64} tautomer in place of His_{ε}^{64} was justified by their observation that a red-shifted CO stretch frequency would be obtained if the imidazole group were rotated to provide contact between $N_{\delta}H$ and CO. In this model, the protonated side chain of Arg^{45} would interact with the basic Nitrogen atoms on His^{64}, inducing a change in the $His^{64}\cdots CO$ interaction. The proposed induced motion was investigated within our own His_{ε}^{64} QM/MM(MD) framework. Firstly, close H-bonding interactions between N_{δ} and either water or the nearby protonated Arg^{45} were added to the DFT model system. It was found, however, that simply adding H-bonding of Arg^{45} or water to His_{ε}^{64} induced negligible shifts in the CO stretch frequency via

through-space interactions or polarization of the imidazole side chain. A second study then examined the possibility of cooperative motions, such as collisions between Arg^{45} and His_{ε}^{64} that push His^{64} closer to the bound ligand. To search for such an effect, one of the heme propionate groups was rotated away from Arg^{45} in the starting structure of the simulation. Without this rotation, a stable $Arg^{45}\cdots COO^{-}$ interaction prevents contact with His^{64}. A 1-ns MD trajectory was run on the rotated conformation, with the positively charged Arg^{45} side chain now free to interact. In the new trajectory, the propionate ultimately rotates back to its favored position next to Arg^{45}, whereas the necessary cooperative motion with His^{64} was not observed. In fact, the His^{64} side chain is pulled away from the CO ligand toward the now available positive charge of the Arginine. As a result, frequencies do not show red shifts.

In summary, the QM/MM(MD) results presented so far provide strong candidates for conformational substates that are likely the A_0 and A_1 substates found in experimental spectra. Data support the open conformation of His^{64} as the origin of the A_0 substate, and a closed conformation of the His_{ε}^{64} tautomer as the origin of A_1. No strong candidate has yet emerged to explain the origin of the A_3 substate. Given the various previous proposals tested here, and the considerable effort that has been made to characterize this substate in terms of local conformational changes, it seems increasingly unlikely that a simple static explanation involving only local binding site moieties exists. Focus must then shift to examining the wider dynamics of the protein, requiring greater sampling of conformational space on longer timescales to observe further possible candidates such as the helix sliding motion suggested by Ishikawa et al. (28)

Analysis of electrostatic interactions in the active site

As a first step toward developing improved intermolecular interactions around the active site, the quality of the force field was considered in more detail. In particular, the electrostatics around the heme, His^{64}, and CO moieties, was analyzed and compared with results from ab initio calculations.

First, the accuracy of the electrostatics around the His_{ε}^{64} imidazole side chain was assessed. Direct comparisons between CHARMM and ab initio results were made by use of the AIM partitioning method (36). Within this framework, the molecular charge density of a methyl-capped imidazole group was partitioned to isolate only the charge density belonging to the atoms of the imidazole ring. Then, the electrostatic potential on the 0.001 au isodensity surface (chosen as it lies near the water-accessible surface, relevant to H-bonding) generated by the charge density of the ring atoms was evaluated. This allows direct comparison with the electrostatic potential from CHARMM atomic charges at the same points on the isodensity surface. The ab initio charge density was generated from DFT calculations at the B3LYP level of theory. A modified version of the MORPHY98 (38) program was used to partition the molecular isodensity surface and generate the electrostatic potential across a fine grid of $\approx$7400 evenly spaced points across the molecular surface, following a procedure described previously (45).

It is important to note that every partitioning method introduces an artificial boundary to define discrete regions of space associated with atoms from a continuous molecular wave function. However, the electrostatic potential generated by one entire functional group using either a continuous charge distribution or discrete point charges should be similar, if the point charges are intended to recreate the electrostatics of the same molecular charge density. Some differences may arise if charges are fitted primarily to interaction energies. For more details on how to interpret different partitionings, the reader is referred to the literature (45). Comparing the CHARMM and DFT results (Fig. 5) reveals that the imidazole side chain of His_{ε}^{64} is quite well parameterized, and there is good qualitative agreement. Quantitative agreement is measured via the mean unsigned error (MUE),

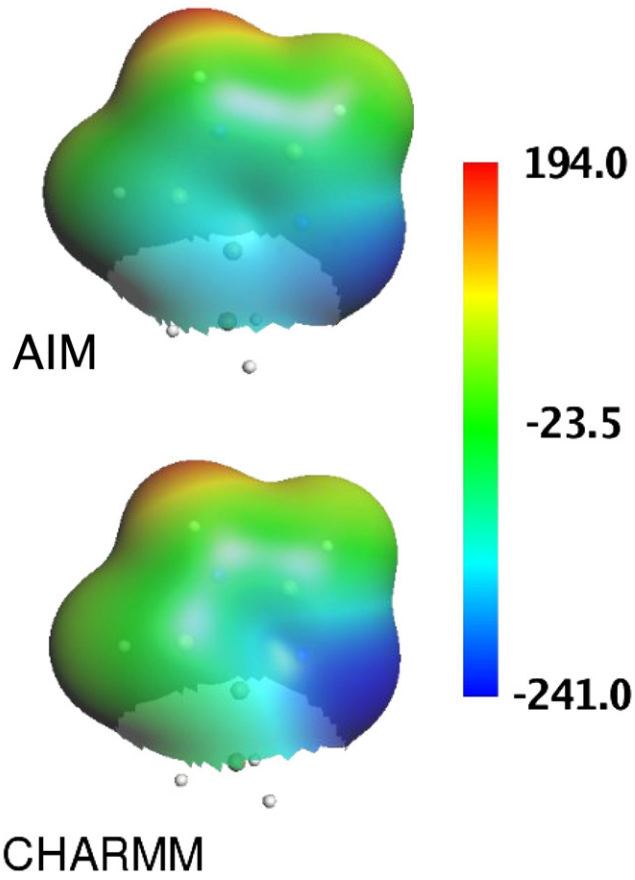

FIGURE 5 Comparison of ab initio (AIM) and CHARMM electrostatic potential distribution on 0.001 au isodensity surface of an imidazole ring fragment. Good qualitative and quantitative agreement is present.

$$MUE = \frac{1}{N}\sum_{i=1}^{N} |V_{\mathrm{DFT}}(\mathbf{r}_i) - V_{\mathrm{CHARMM}}(\mathbf{r}_i)|, \quad (3)$$

where the sum runs over all N grid points. The MUE across the imidazole surface is only 4.6 kcal/mol, although the largest error is 16.5 kcal/mol. Negative charge is overestimated to some degree at the N_δ position. Similar results were obtained for the His_δ^{64} tautomer.

Next, the electrostatics around the heme group was studied. As the heme represents a relatively large system for ab initio calculations, it was decided initially to focus on the electrostatic potential in the region around the CO ligand. This is the region most relevant to interactions with the His^{64} side chain. A detailed charge model for the ligand multipole moments would later be required to more realistically represent the electric field in the binding site (see below), and needed to be fitted separately. It was therefore desirable to develop a heme model that was not dependent on CO ligand parameters, meaning the ligand contribution to the electrostatic potential had to be removed. To this end, a quantity $V_{\mathrm{AIM}}(\mathbf{r})$, resulting from only the heme and His^{93} contribution was defined as

$$V_{\mathrm{AIM}}(\mathbf{r}) = \int d\tau' \frac{\boldsymbol{\rho}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \int_{\Omega_C} d\tau' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \int_{\Omega_O} d\tau' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \quad (4)$$

The first volume integral of the electron density $\rho(\mathbf{r}')$ extends over all space, the second over the AIM carbon atomic basin Ω_C, and the third over the AIM oxygen atomic basin Ω_O. $|\mathbf{r}-\mathbf{r}'|$ is the scalar distance between the current point $\mathbf{r}'$ of the volume integral and the position at which we are evaluating the potential.

By excluding the CO contribution, the heme parameters can be modified to more accurately represent the heme electric field without introducing a dependence on the CO charge parameters. To obtain $V_{\mathrm{AIM}}(\mathbf{r})$, an 8.25 × 9.05 × 8.35 Å grid with grid points spaced 0.05 Å apart was placed over the CO ligand. The molecular electrostatic potential and electron density values were evaluated at each grid point using the Gaussian03 program, at the B3LYP/6-311+G(2d,p) level of theory. A larger basis set was used to better represent the negative charge on the heme anion, as well as the metal center. MORPHY98 (38) was again used to partition the molecular charge density and obtain the density corresponding to the bound CO ligand. $V_{\mathrm{AIM}}(\mathbf{r})$ was calculated across the 0.001 au CO isodensity surface to give the electrostatic envelope arising from just the heme group and the proximal His^{93}, which can be directly compared with the potential generated using CHARMM point charges. The standard CHARMM heme point charge model is termed model ''HA'' throughout the remainder of the text, and the associated heme electrostatic potential is defined as

$$V_{\mathrm{HA}}(\mathbf{r}) = \sum_{n=1}^{\mathrm{nheme}} \frac{q_n}{|\mathbf{r}-\mathbf{r}_n|}. \quad (5)$$

The sum runs over all ''nheme'' heme and His^{93} atoms with atomic charge q_n. The difference $V_{\mathrm{AIM}}(\mathbf{r})$ - $V_{\mathrm{HA}}(\mathbf{r})$ was larger than the corresponding difference between DFT and CHARMM for the imidazole group. The MUE on the 0.001 au CO isodensity surface was 15.5 kcal/mol, with a maximum of 29.2 kcal/mol (Fig. 6). Across the more distant 10^{-4} and 10^{-5} au isodensity surfaces, the error is smaller but still at 10.4 kcal/mol on the 10^{-5} au isodensity surface. The error is asymmetric, with negative charge overestimated on the side of the isodensity surface closest to the

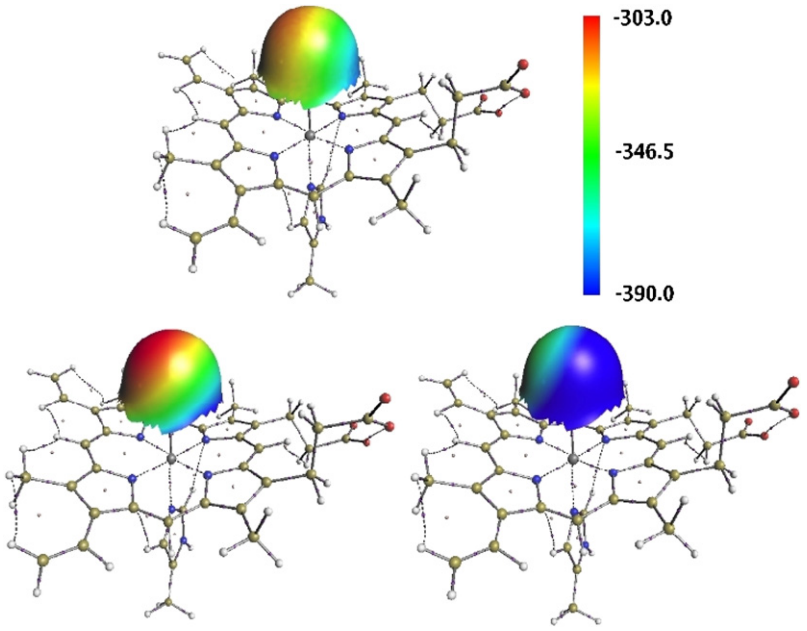

FIGURE 6 Comparison of ab initio (*top*), Autenrieth (*bottom-left*), and CHARMM22 (*bottom-right*) electrostatic potential distribution from the atoms of the heme group and proximal His^{93} on the 0.001 au isodensity surface of the CO ligand. The contribution to the electrostatic potential from the charge density of the CO ligand has been removed in each case. The Autenrieth model clearly agrees better with ab initio results. AIM analysis was used to partition the molecular isodensity surface. Topological features including atomic interaction lines (*solid* and *dotted paths* running between nuclei), bond critical points (*small purple spheres* lying along *interaction lines*), and ring critical points (*small pink spheres* enclosed by *atomic interaction lines*) are also shown.

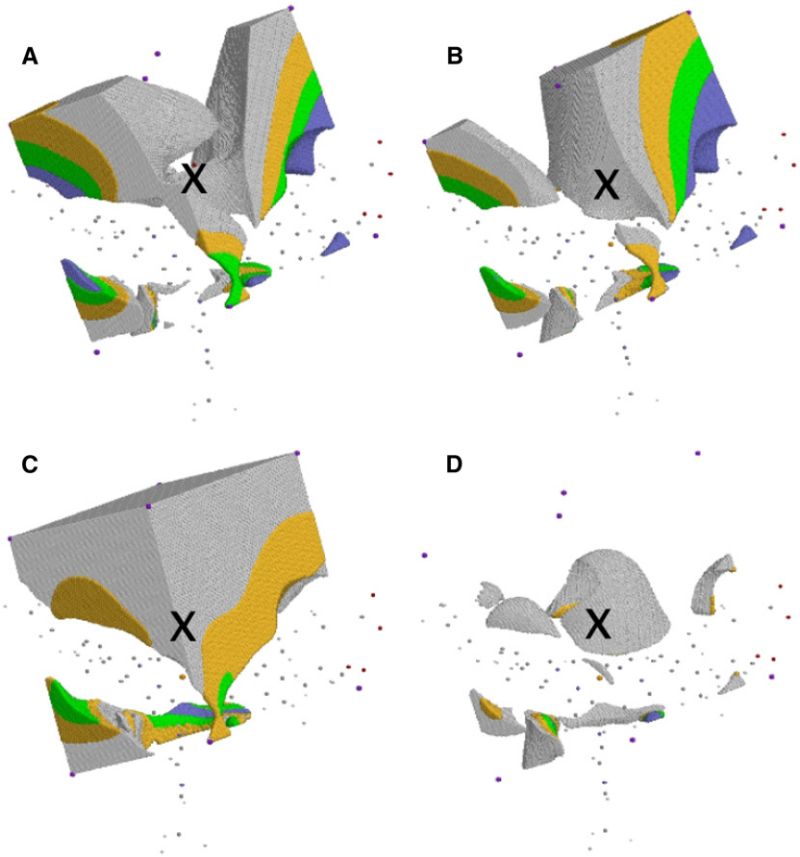

FIGURE 7 Comparison of model using gas phase CO parameters and CHARMM22 heme charges (*A*), fluctuating quadrupole CO model and original CHARMM22 heme charges (*B*), fluctuating quadrupole CO model and original Autenrieth heme charges (*C*), and fluctuating quadrupole CO model with modified Autenrieth heme charges (*D*). Grids are placed across the bound CO ligand (position of CO is marked with a *cross*), and are viewed from the side of the heme plane with propionate groups on the right-hand side. Gray areas of each grid represent an MUE from the DFT electrostatic potential of between 4.8 and 9.6 kcal/mol, yellow regions differ from DFT by between 9.6 and 14.3 kcal/mol, green regions differ by between 14.3 and 19.1 kcal/mol, and blue areas by >19.1 kcal/mol.

heme propionates, whereas positive charge is overestimated at the opposite end of the CO isodensity surface. The asymmetric error cannot be corrected using the missing CO point charges lying perpendicular to the heme plane, so it seems important to adjust some of the charges on the heme group.

An alternative heme model (model HB) has been developed by Autenrieth et al. (46) for use in cytochrome *c*. A comparison of $V_{HB}(\mathbf{r})$ with $V_{AIM}(\mathbf{r})$ and $V_{HA}(\mathbf{r})$ is also included in Fig. 6. $V_{HB}(\mathbf{r})$ is in much better agreement with $V_{AIM}(\mathbf{r})$ than $V_{HA}(\mathbf{r})$. The MUE across the 0.001 au CO surface is 3.1 kcal/mol compared with 15.5 kcal/mol from model HA. Across the more distant 10^{-4} and 10^{-5} au isodensity surfaces, the electrostatic potential is, however, systematically too positive. It should be noted that the total net charge of the heme in model HB is −1.62 au compared with the expected −2.00 au.

Next, the contribution of the CO ligand charge density to the electrostatic potential generated by models HA and HB was added. Only the results using charge model III for CO (see Theoretical Methods), the most detailed CO charge model, are presented here. The combinations of heme and CO charge model III are then denoted HA(III) and HB(III). With the addition of CO charge model III, the complete molecular electrostatic potential at point **r**, generated by the charge density of the entire His^{93}-heme-CO system, was evaluated. The molecular electrostatic potential obtained across a large number of grid points at the B3LYP level of theory was compared with the electrostatic potential at the same points generated using the charge models HA(III) and HB(III). The first three-dimensional grid had dimensions of 21.0 × 18.5 × 23.0 Å with coarse grid points spaced 0.5 Å apart to cover the entire heme and His^{93} groups. The electrostatic potential was compared at regions outside the 0.001 au isodensity surface of the molecule. Model HA(III) gave an MUE across the entire grid of 7.6 kcal/mol. The strong asymmetry in the error across the heme plane was now more clearly visible. The negative charge around the carboxyl groups of the heme propionates was considerably overestimated, and excessive positive charge was located at the opposite end of the heme system. Interestingly, the CO ligand is located directly at the point of crossover, where the error partially cancels. The effect can be seen after zooming in on a small grid across the CO position and plotting envelopes of increasing error as shown in Fig. 7. Fig. 7, *a* and *b* (based on model HA), show the gap of low error above the ring plane around the CO position. Model HB(III) performed slightly better across the large grid, but still had an MUE of 6.6 kcal/mol. In contrast to model HA(III), the electrostatic potential was systematically too positive across the entire grid, and the error was quite evenly distributed.

As models HA(III) and HB(III) showed quite large MUEs, an improved model (HC) was developed starting from

HB(III). Negative charge on the heme propionate oxygen atoms was enhanced from −0.64 au to −0.685 au, based on visual inspection of the difference maps (Fig. 7) and monitoring of improvement in the MUE. Two slightly enhanced positive carbon counter charges were added around the center of the heme ring. The resulting model HC has a total heme charge of −1.83 au in place of −1.62 au, giving improved long-range electrostatics. Model HC(III) has an MUE of 0.9 kcal/mol across the large molecular grid, a significant improvement on the earlier models. Finally, the performance of model HC was evaluated in the important region around the CO ligand. The same 8.25 × 9.05 × 8.35 Å grid was used as before, and similar trends were observed as for the larger molecular grid (Fig. 7). Model HC(III) is again much closer to the DFT results than models HA(III) and HB(III), with an MUE of 1.8 kcal/mol compared with 8.9 kcal/mol and 7.7 kcal/mol for HA and HB, respectively.

To summarize, the original CHARMM22 heme parameters, model HA, when treated on their own or combined with a CO charge model lead to substantial errors in the electrostatic potential around the bound CO ligand when compared with DFT results. Significant improvement can be achieved by using heme model HB, developed for use in cytochrome *c*, although the total charge in this model is not sufficiently negative when applied to the MbCO system. Small modifications to the HB model lead to model HC, which leads to much better agreement in the electrostatics. The improved electrostatic model with corrected net negative charge can be directly applied in MD simulations. With the electrostatic potential generated by His^{64} quantitatively examined, and the heme electrostatic model refined, the local electric field at the ligand position should be sufficiently accurate to model CO vibrational frequency shifts. Further improvement could, however, still be made by examining evolution of the electric field as a function of dynamic motions and conformational change.

A CO fluctuating charge model for vibrational spectra from MD simulations

The final ingredient to calculate CO- IR spectra from MD simulations (see Eq. 2) is a suitable charge model for the ligand (models I–III, see Theoretical Methods). Within this methodology, the ligand vibrational response is caused by electrostatic interactions between the force-field charges on the ligand atoms and the electric field they encounter. From previous work on unbound CO in the protein cavities (19,14,15) a meaningful way forward was to consider increasingly accurate models of the charge distribution. However, improvement is not guaranteed, as accurate description of the fragment electrostatics from atomic contributions only ensures accurate net forces on the total charge density of the ligand, and not accurate distribution of forces between nuclei.

Spectra from point charge model HA(I)

Twenty-six 1-ns MD trajectories were run using the original CHARMM model HA(I). Ten trajectories were run for His^{64}_{ε}, ten for $His^{64}H^{+}$ starting in the closed conformation, and six for $His^{64}H^{+}$ starting in the open conformation. Averages over the spectra calculated for each trajectory are shown in Fig. 8 *a*. Several of the trajectories using $His^{64}H^{+}$ that started in the closed conformation led to opening of the His^{64} gate, supporting the earlier QM/MM(MD) results. The $His^{64}H^{+}$ side chain was also observed to adopt a number of partially open conformations, leading to a relatively broad spectrum with all peaks centered at ~1972 cm^{-1} (Fig. 8 *a*). There is no significant blue shift of the vibrational frequency of the open conformation with respect to the closed conformation, however, in contrast to the experimental and earlier QM/MM(MD) results. The CHARMM charge model HA(I) with anharmonic-bonded potential therefore does not show sufficient vibrational sensitivity to local electrostatic interactions with His^{64} between the open and closed conformations.

Spectra from fluctuating dipole model HC(II)

Model HC(II), reproducing the AIM dipole moment of the CO ligand, was then tested in a similar fashion. Table 2 shows the results of the AIM multipole analysis of the CO ligand. The global axis system for this analysis was rotated so that the CO bond defined the new *z* axis for the multipole moments. The Q[10] and Q[20] (*Z* and *ZZ*) multipole components therefore dominate; however, polarization of the ligand caused largely by the heme propionate groups causes deviation from cylindrical symmetry and significant off-axis multipole components to emerge. This asymmetry cannot be captured within the current fluctuating point charge models, where all charge sites lie along the *z* axis. A linear function was fitted to obtain values for fluctuating atomic charges as a function of bond length to describe the charge and dipole moments correctly. The Q[10] component is described quite well by the linear expression, as shown in the table, whereas the off-axis components are not.

Model II was combined with newly fitted heme model, HC, meaning we make two changes from the simulations involving model HA(I), with the earlier simulations acting as a reference to compare with the performance of the more detailed models. Spectra generated using model HC(II) are shown in Fig. 8 *b*. The first (black) spectrum is calculated from six 1-ns trajectories of the His^{64}_{ε} tautomer in the closed conformation. The second (red) spectrum used six trajectories of $His^{64}H^{+}$ in the open conformation. The position of the main peak in each spectrum is roughly the same, around 1949 cm^{-1}, demonstrating that we still do not obtain the experimentally observed CO frequency shift between the open and closed conformations. The model does, however, lead to broadened spectra. Noteworthy are the red-shifted peaks in the spectrum of the $His^{64}H^{+}$ (red) open conformation. These peaks seem to be caused by

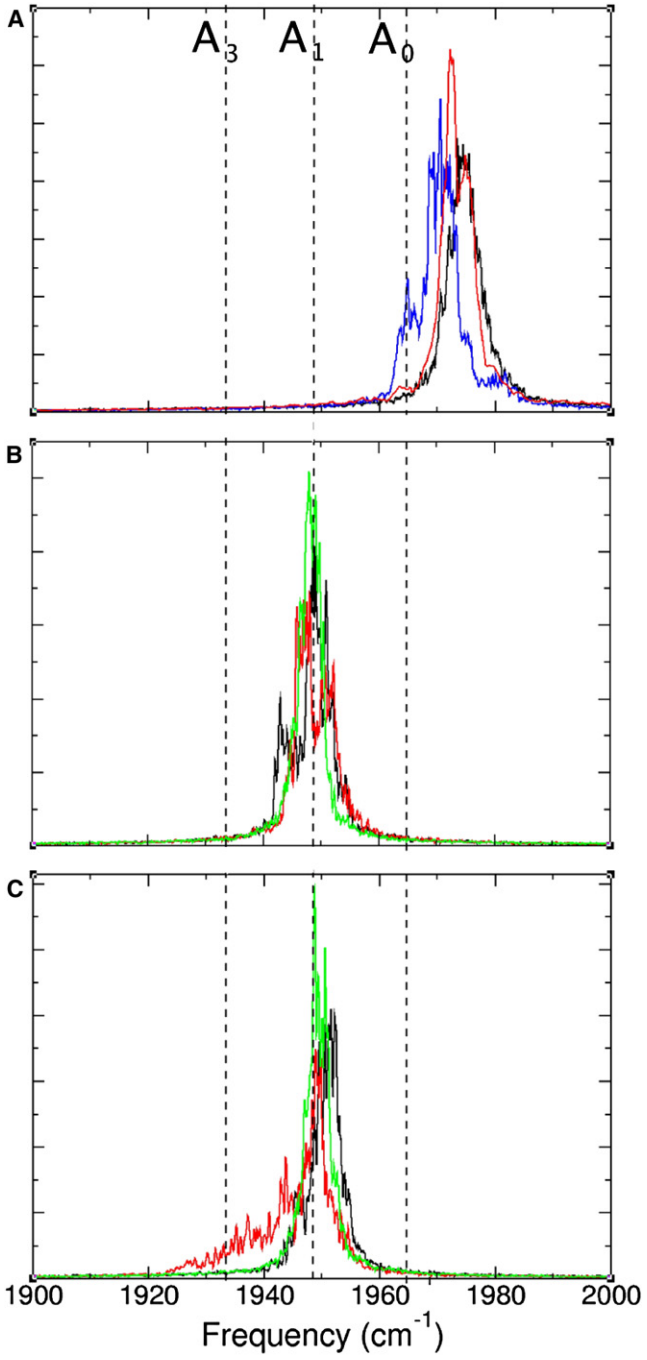

FIGURE 8 CO stretch frequencies generated from CO dipole moment and the dipole-dipole autocorrelation function. (*a*) Averaged spectra from 10 trajectories using Model HA(I) with His_{ε}^{64} in the closed conformation (*black*), 10 using $His^{64}H^{+}$ starting in the closed conformation (*blue*), and six with $His^{64}H^{+}$ starting in the open conformation (*red*). (*b*) Averaged spectra from six trajectories using model HC(II) with His_{ε}^{64} in the closed conformation (*black*), six using $His^{64}H^{+}$ in the open conformation (*red*), and six with His_{δ}^{64} (*green*). (*c*) Averaged spectra from six trajectories using model HC(III) with His_{ε}^{64} in the closed conformation (*black*), six with $His^{64}H^{+}$ in the open conformation (*red*), and six with His_{δ}^{64} (*green*). Experimental frequency shifts between open and closed conformations are still not present.

internal heating of the CO bond. The heating seems to occur after interaction with water molecules in contact with CO in the active site, able to reach the ligand as the $His^{64}H^{+}$ side chain swings into the solvent. Finally, six trajectories were run using the His_{δ}^{64} tautomer, with results shown as the green spectrum in Fig. 8 *b*. Again no significant frequency shift is observed, suggesting model HC(II) is not sufficiently sensitive to changes in electric field to study the vibrational splitting in question.

Spectra from fluctuating quadrupole model HC(III)

The final model (model HC(III)) includes a third charge site to reproduce the AIM charge, dipole, and quadrupole moments. The first six 1-ns trajectories obtained using model HC(III) again examine the closed conformation of the His_{ε}^{64} tautomer, and are compared with a further six trajectories of the open conformation of $His^{64}H^{+}$ in Fig. 8 *c*. The figure shows that the frequency is shifted by 5 cm^{-1} from the open conformation, in comparison with the roughly 16 cm^{-1} red shift visible in the experimental spectra. Larger shifts had also been observed in the original ab initio results. It was noticed, however, that heating of the CO bond was again taking place in several of the open conformation trajectories.

Six further 1-ns trajectories were then run using the His_{δ}^{64} tautomer. The resulting averaged vibrational spectrum is overlaid on the open and closed conformations as the green peak in Fig. 8 *c*. The peak is only slightly shifted from the A_0 position, and this time slightly red-shifted rather than blue-shifted, as was the case in previous studies (20) and our earlier DFT results. Model III, then, shows vibrational splitting in the $His^{64}H^{+}$ trajectory, but does not give the relative frequency shifts observed in experimental spectra. Structural comparison of trajectories with the earlier QM/MM(MD) results did, however, reveal one significant change resulting from the new charge model and extended simulation times. The His_{δ}^{64} imidazole side chain is now observed to intermittently rotate, so that the $N_{\delta}H$ group points toward the CO ligand and reduces the negative electric field at the CO position. This motion was previously energetically unfavorable with model HA because of repulsion from the stronger negative charge of the propionate groups. Increased QM/MM(MD) sampling would reveal the effect the new conformation has on the spectra. The continued dominance of the $N_{\varepsilon}\cdots CO$ interaction, however, makes it unlikely that the blue-shifted components of the His_{δ}^{64} spectra will disappear.

The different charge models tested have confirmed the observed behavior of the opening of the $His^{64}H^{+}$ gate, generally within 1 ns of protonation of the imidazole group, although a number of partially open transitional conformations were also observed to exist. The experimentally and DFT-observed vibrational splitting between the open and closed conformations (A_0 and A_1) could not be calculated quantitatively from extensive MD simulations using increasingly detailed and carefully optimized electrostatic models. This is in contrast with the success of a related approach

TABLE 2 CO fragment AIM dipole and quadrupole moments from B3LYP/6-311+G(2d,p) charge density

	r(CO) (Ångstrom)	Q[10]	Q[11c]	Q[11s]	Magnitude	Q[20]	Q[21c]	Q[21s]	Q[22c]	Q[22s]	Magnitude
AIM	1.015	−0.172	−0.046	−0.049	0.185	−0.635	0.170	0.234	−0.031	−0.023	0.699
3-point	1.015	−0.170	0.000	0.000	0.170	−0.634	0.000	0.000	0.000	0.000	0.634
AIM	1.140	−0.269	−0.047	−0.049	0.278	−0.567	0.158	0.255	−0.035	−0.028	0.643
3-point	1.140	−0.263	0.000	0.000	0.263	−0.594	0.000	0.000	0.000	0.000	0.594
AIM	1.265	−0.369	−0.049	−0.048	0.375	−0.482	0.144	0.277	−0.041	−0.036	0.577
3-point	1.265	−0.372	0.000	0.000	0.372	−0.479	0.000	0.000	0.000	0.000	0.479

Results are compared with those of the fitted three-point fluctuating charge model, labeled "3-point".

previously applied to the vibrational spectra of photodissociated CO. The difficulties encountered may be related to additional factors beyond capturing the electrostatic interactions for bound MbCO compared with unbound CO in the protein environment. They include the direct coupling of the CO bond to the Fe-C, Fe-N_{His93}, and Fe-N_{heme} bonds and the Fe-C-O valence angle. In other words: For free CO in Mb, the accurate representation of the electrostatics, together with a faithful representation of the intermolecular CO bond potential, was found sufficient to quantitatively explain structural, dynamical, and spectroscopic properties of the ligand (15,19,14). However, for bound CO, the explicit coupling of the CO-coordinate *r* to other degrees of freedom (see above) is also likely to be involved. This is in line with experimental work on CO stretch relaxation measurements in MbCO, which suggested that the CO stretching motion could be coupled to the Fe-CO stretching and bending modes and the heme-ring modes (47). It is, however, reassuring that improving the electrostatic model for CO and the heme leads to observable effects in the spectroscopy (Fig. 8), which is in line with previous work on unbound CO. In other words, the calculated IR spectra are sensitive to the details of the electrostatic models used. Previously it has been proposed that charge polarization in the FeC bond should affect back-bonding between the Fe and C atoms, causing a strengthening or weakening of the CO bond and which explicitly couples the Fe-C and the C-O coordinates (48). This results in increased sensitivity of the CO bond in bound MbCO to external electric fields (49), requiring a force-field model that is significantly more sensitive than for the unbound case. Furthermore, new experimental evidence shows that the A_1 and A_3 substates may be linked by a large-scale conformational change in the protein, exhibiting a low energy barrier and taking place on short timescales. Investigating the large-scale structural changes will benefit from a combination of different sampling approaches and expanded simulations of the protein rather than improved charge models.

CONCLUSIONS

The structural origins of the observed spectroscopic A-states of bound MbCO have been investigated using a combination of DFT and MD techniques. The DFT calculations, examining the calculated vibrational frequencies of conformational snapshots taken from individual MD trajectories, support the hypothesis that the His^{64}_{ε} tautomer is the major form contributing to the experimental IR spectra. There is further evidence to support that the open conformation, with the $His^{64}H^{+}$ side chain swung into the solvent, is the origin of the A_0 substate. The frequency shift associated with the A_1 substate is correctly described by a $N_{\varepsilon}H\cdots CO$ interaction similar to that of the x-ray crystal structure. None of the remaining conformational or structural changes tested provide a convincing assignment of the missing A_3 substate, although conformational sampling is clearly an issue with the relatively large DFT model system used.

MD simulations were used to draw conclusions about the way the system responds to different changes, such as protonation of the His^{64} side chain at low pH. In the case of His^{64} protonation, the closed conformation was observed to convert to the open conformation on the 1-ns timescale in most of the simulations. The reliability of the MD simulations was addressed by refining the electrostatics of the underlying heme and CO models to agree with DFT results. The point charges of the heme model were refined using AIM partitioning to examine the contributions of the charge density of different groups to the total molecular charge density. The CO electrostatics were reproduced in more detail, using increasingly advanced fluctuating charge models to represent the CO fragment charge, dipole, and quadrupole moments.

Finally, the modified CO charge model was used to calculate the CO-vibrational spectrum directly from MD trajectories. This improves the sampling and it is possible to examine the response of the vibrational frequency to changing protein conformation. However, although charge models were developed that exhibited spectral peak broadening in response to their local environment, the shifts and splittings were generally small and difficult to relate to structural change.

In conclusion, this work shows that by using state-of-the art QM/MM(MD) simulations, the A_0 and A_1 states can be assigned to the open conformation of $His^{64}H^{+}$ and the closed conformation of the His^{64}_{ε} tautomer, respectively. For A_3, all previously suggested structural interpretations can be ruled out. Together with results from experiments this leads us to conclude that A_3 is related to A_1 through a larger structural change. Furthermore, it is found that using carefully balanced electrostatic models together with pure MD simulations is not sufficient to explain the structural origins of A_0,

A_1, or A_3. Therefore, including additional factors, such as coupling between the ligand and the protein degrees of freedom, is necessary.

The authors gratefully acknowledge financial support from the Swiss National Science Foundation.

REFERENCES

1. Frauenfelder, H., and B. McMahon. 2002. Hydration, slaving and protein function. *Biophys. Chem.* 98:35–48.
2. Yang, F., and G. N. Phillips, Jr. 1996. Crystal structures of CO-, deoxy- and Met-myoglobins at various pH values. *Biophys. J.* 256:762–774.
3. Ansari, A., J. Berendzen, D. Braunstein, B. R. Cowen, H. Frauenfelder, et al. 1987. Rebinding and relaxation in the myoglobin pocket. *Biophys. Chem.* 26:337–355.
4. Johnson, J. B., D. C. Lamb, H. Frauenfelder, J. D. Mueller, B. McMahon, et al. 1996. Ligand binding to heme proteins. VI. Interconversion of taxonomic substates in carbonmonoxymyoglobin. *Biophys. J.* 71:1563–1573.
5. Phillips, G., M. Teodoro, T. Li, B. Smith, and J. Olson. 1999. Bound CO is a molecular probe of electrostatic potential in the distal pocket of myoglobin. *J. Phys. Chem. B.* 103:8817–8829.
6. Morikis, D., P. M. Champion, B. A. Springer, and S. G. Sligar. 1989. Resonance Raman investigations of site-directed mutants of myoglobin: effects of distal histidine replacement. *Biochemistry.* 28:4791–4800.
7. Osapay, K., Y. Theriault, P. E. Wright, and D. A. Case. 1994. Solution structure of carbonmonoxy myoglobin determined from nuclear magnetic resonance distance and chemical shift constraints. *J. Mol. Biol.* 244:183–197.
8. Merchant, K. A., W. G. Noid, D. E. Thompson, R. Akiyama, R. F. Loring, et al. 2003. Structural assignments and dynamics of the A substates of MbCO: spectrally resolved vibrational echo experiments and molecular dynamics simulations. *J. Phys. Chem. B.* 107:4–7.
9. Elber, R., and M. Karplus. 1987. Multiple conformational states of proteins—a molecular-dynamics analysis of myoglobin. *Science.* 235:318–321.
10. Kottalam, J., and D. A. Case. 1988. Dynamics of ligand escape from the heme pocket of myoglobin. *J. Am. Chem. Soc.* 110:7690–7697.
11. Straub, J. E., and M. Karplus. 1991. Molecular dynamics study of the photodissociation of carbon-monoxide from myoglobin: ligand dynamics in the first 10 ps. *Chem. Phys.* 158:221–248.
12. Sagnella, D. E., and J. E. Straub. 1999. A study of vibrational relaxation of B-state carbon monoxide in the heme pocket of photolyzed carboxymyoglobin. *Biophys. J.* 77:70–84.
13. Meuwly, M., O. M. Becker, R. Stote, and M. Karplus. 2002. NO rebinding to myoglobin: a reactive molecular dynamics study. *Biophys. Chem.* 98:183–207.
14. Nutt, D. R., and M. Meuwly. 2003. Theoretical investigation of infrared spectra and pocket dynamics of photodissociated carbonmonoxy myoglobin. *Biophys. J.* 85:3612–3623.
15. Nutt, D. R., and M. Meuwly. 2004. CO migration in native and mutant myoglobin: atomistic simulations for the understanding of protein function. *Proc. Natl. Acad. Sci. USA.* 101:5998–6002.
16. Lim, M., T. A. Jackson, and P. A. Anfinrud. 1995. Mid-infrared vibrational spectrum of CO after photodissociation from heme: evidence for a ligand docking site in the heme pocket of hemoglobin and myoglobin. *J. Chem. Phys.* 102:4355–4366.
17. Nienhaus, K., J. S. Olson, S. Franzen, and G. U. Nienhaus. 2005. The origin of stark splitting in the initial photoproduct state of MbCO. *J. Am. Chem. Soc.* 127:40–41.
18. Meuwly, M. 2006. On the influence of the local environment on the CO stretching frequencies in native myoglobin: assignment of the B-states in MbCO. *ChemPhysChem.* 10:2061–2063.
19. Plattner, N., and M. Meuwly. 2008. The role of higher CO-multipole moments in understanding the dynamics of photodissociated carbonmonoxide in myoglobin. *Biophys. J.* 94:2505–2515.
20. Franzen, S. 2001. An electrostatic model for the frequency shifts in the carbonmonoxy stretching band of myoglobin: correlation of hydrogen bonding and the Stark tuning rate. *J. Am. Chem. Soc.* 124:13271–13280.
21. Rovira, C., B. Schulze, M. Eichinger, J. D. Evanseck, and M. Parrinello. 2001. Influence of the heme pocket conformation on the structure and vibrations of the Fe-CO bond in myoglobin: a QM/MM density functional study. *Biophys. J.* 81:435–445.
22. Müller, J. D., B. J. McMahon, E. Y. T. Chien, S. G. Sligar, and G. U. Nienhaus. 1999. Connection between the taxonomic substates and protonation of histidines 64 and 97 in carbonmonoxy myoglobin. *Biophys. J.* 77:1036–1051.
23. Braunstein, D. P., K. Chu, K. D. Egeberg, H. Frauenfelder, J. R. Mourant, et al. 1993. Ligand binding to heme proteins. III. FTIR studies of His-E7 and Val-E11 mutants of carbonmonoxymyoglobin. *Biophys. J.* 65:2447–2454.
24. Ivanov, D., J. T. Sage, M. Keim, J. R. Powell, S. A. Asher, et al. 1994. Determination of CO orientation in myoglobin by single-crystal infrared linear dichroism. *J. Am. Chem. Soc.* 116:4139–4140.
25. Loccisano, A., O. Acevedo, J. DeChancie, B. Schulze, and J. Evanseck. 2004. Enhanced sampling by multiple molecular dynamics trajectories: carbonmonoxy myoglobin 10 μs A_0 to A_{1-3} transition from ten 400 picosecond simulations. *J. Mol. Graph. Model.* 22:369–376.
26. Merchant, K. A., W. G. Noid, R. Akiyama, I. J. Finkelstein, A. Goun, et al. 2003. Myoglobin-CO substate structures and dynamics: multidimensional vibrational echoes and molecular dynamics simulations. *J. Am. Chem. Soc.* 125:13804–13818.
27. Unno, M., J. F. Christian, J. S. Olson, J. T. Sage, and P. M. Champion. 1998. Evidence for hydrogen bonding effects in the iron ligand vibrations of carbonmonoxy myoglobin. *J. Am. Chem. Soc.* 120:2670–2671.
28. Ishikawa, H., K. Kwak, J. K. Chung, S. Kim, and M. D. Fayer. 2008. Direct observation of fast protein conformational switching. *Proc. Natl. Acad. Sci. USA.* 105:8619–8624.
29. Vojtechovsky, J., K. Chu, J. Berendzen, R. M. Sweet, and I. Schlichting, 1999. Crystal structures of myoglobin-ligand complexes at near-atomic resolution. *Biophys. J.* 77:2153–2174.
30. Schulze, B., and J. D. Evanseck. 1999. Cooperative role of Arg^{45} and His^{64} in the spectroscopic A_3 state of carbonmonoxy myoglobin: molecular dynamics simulations, multivariate analysis, and quantum mechanical computations. *J. Am. Chem. Soc.* 121:6444–6454.
31. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, et al. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* 4:187–217.
32. MacKerell, Jr., A. D., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
33. Kuriyan, J., S. Wilz, M. Karplus, and G. A. Petsko. 1986. X-ray structure and refinement of carbon-monoxy (Fe II)-myoglobin at 1.5 Angstrom resolution. *J. Mol. Biol.* 192:133–154.
34. Brooks, C. L., and M. Karplus. 1983. Deformable stochastic boundaries in molecular-dynamics. *J. Chem. Phys.* 79:6312–6325.
35. Jorgensen, W. L., J. D. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
36. Bader, R. F. W. 1990. Atoms in Molecules—A Quantum Theory. Oxford University Press, Oxford, UK.
37. Popelier, P. L. A. 2000. Atoms in Molecules. An Introduction. Pearson Education, London, UK.
38. Popelier, P. L. A. 1998. MORPHY98. UMIST, Manchester, England.
39. Aicken, F. M., and P. L. A. Popelier. 2000. Atomic properties of selected biomolecules. Part 1. The interpretation of atomic integration errors. *Can. J. Chem.* 78:415–426.

40. McQuarrie, D. A. 1976. Statistical Mechanics. Harper's Chemistry Series, New York.

41. Allen, M. P., and D. J. Tildesley. 1987. Computer Simulation of Liquids. Clarendon Press, Oxford, UK.

42. Frisch, M.J., G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, et al. 2004. Gaussian 03, Rev. C.01. Gaussian, Wallingford, CT.

43. Le Roy, R.J. 1996. LEVEL 6.1. *Chem. Phys. Res. Reports* CP-555R.

44. Bossa, C., A. Amadei, I. Daidone, M. Anselmi, B. Vallone, et al. 2005. Molecular dynamics simulation of sperm whale myoglobin: effects of mutations and trapped CO on the structure and dynamics of cavities. *Biophys. J.* 89:465–474.

45. Popelier, P. L. A., M. Devereux, and M. Rafat. 2004. The quantum topological electrostatic potential as a probe for functional group transferability. *Acta Crystallogr.* 60:427–433.

46. Autenrieth, F., E. Tajkhorshid, E. Baudry, and Z. Luthey-Schulten. 2004. Classical force field parameters for the heme prosthetic group of cytochrome *c*. *J. Comput. Chem.* 25:1613–1622.

47. Owrutsky, J. C., M. Li, B. Locke, and R. M. Hochstrasser. 1995. Vibrational relaxation of the CO stretch vibration in hemoglobin-CO, myoglobin-CO, and protoheme-CO. *J. Phys. Chem.* 99:4842–4846.

48. Li, X.-Y., and T. G. Spiro. 1988. Is bound carbonyl linear or bent in heme proteins? Evidence from resonance Raman and infrared spectroscopic data. *J. Am. Chem. Soc.* 110:6024–6033.

49. Park, E. S., S. S. Andrews, R. B. Hu, and S. G. Boxer. 1999. Vibrational Stark spectroscopy in proteins: a probe and calibration of electrostatic fields. *J. Phys. Chem. B.* 103:9813–9817.