Distal Residue–CO Interaction in Carbonmonoxy Myoglobins: A Molecular Dynamics Study of Three Distal Mutants

Philip Jewsbury*‡ and Teizo Kitagawa*

*Institute for Molecular Science, Myodaiji, Okazaki 444, and ‡Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan

ABSTRACT Six 90-ps molecular dynamics trajectories, two for each of three distal mutants of sperm whale carbonmonoxy myoglobin, are reported; solvent waters within 16 Å of the active site have been included. In both His64Gln trajectories, the distal side chain remains part of the heme pocket, forming a "closed" conformation similar to that of the wild type $64N_{\delta}H$ tautomer. Despite a connectivity more closely resembling the $N_{\epsilon}H$ histidine tautomer, close interactions with the carbonyl ligand similar to those observed for the wild type $64N_{\epsilon}H$ tautomer are prevented in this mutant by repulsive interactions between the carbonyl O and the $64O_{\epsilon}$. The aliphatic distal side chain of the His64Leu mutant shows little interaction with the carbonyl ligand in either His64Leu trajectory. Solvent water molecules move into and out of the active site in the His64Gly mutant trajectories; during all the other carbonmonoxy myoglobin trajectories, including the wild type distal tautomers considered in an earlier work, solvent molecules rarely encroach closer than 6 Å of the active site. These results are consistent with a recent structural interpretation of the wild type infrared spectrum, and the current reinterpretation that the distal–ligand interaction in carbonmonoxy myoglobin is largely electrostatic, not steric, in nature.

INTRODUCTION

Although carbonmonoxy myoglobin (MbCO) is one of the most studied and best understood proteins, the nature of its protein–ligand interaction is still in debate: see Springer et al. (1994) for a recent review. Early interpretations, based on the x-ray structure (Kuriyan et al., 1986) and studies of model heme prosthetic groups (Collman et al, 1976), suggested that the key interaction was a steric repulsion between the CO and the distal histidine residue, the nearest residue to the ligand binding site (shown in Fig. 1). This would destabilize the binding of the poisoning CO ligand by sterically crowding its preferred perpendicular binding geometry, without affecting the physiological ligand, O_2, which prefers a bent conformation (Yu and Kerr, 1988). Neutron diffraction studies have shown that the O_2 ligand is stabilized further by a hydrogen bond to the distal residue (Phillips and Schoenborn, 1981), an interaction not present in the MbCO crystal structure (Norvell et al., 1975; Cheng and Schoenborn, 1991). This discrimination in favor of oxygen is crucial for the physiological functioning of the globin proteins, because unhindered model heme prosthetic groups show affinities for CO up to 100,000 times higher than for O_2. However, the distortions of the Fe—C—O geometry observed in the crystal structures are large, and it has been difficult to account (Case and Karplus, 1978; Ray et al., 1994) for such a large steric repulsion being delivered by the distal side chain, which, being exposed to solvent, is expected to be relatively mobile, as confirmed by the x-ray B factors (Kuriyan et al., 1986) and in molecular dynamics (MD) simulations (Kuczera et al., 1990; Jewsbury and Kitagawa, 1994).

If the distal–CO interaction is steric in nature, however, then the spectral properties of the CO ligand (ν_{CO} and ν_{FeC}) are expected to be sensitive to changes in the steric bulk of the distal side chain. Recent studies of distal mutant MbCOs with aliphatic side chains of varying steric bulk have shown little change in their infrared (IR) (Balasubramanian et al., 1993; Braunstein et al., 1993; Li et al., 1994) or resonance Raman (RR) (Sakan et al., 1993) spectra. Furthermore, the x-ray structures of a range of distal mutant MbCOs have the same Fe—C—O conformation within experimental error (Quillin et al., 1993), and only those mutants with a polar side chain have IR spectra differing significantly from that of the His64Gly mutant (Balasubramanian et al., 1993; Braunstein et al., 1993; Li et al., 1994). Thus it has been proposed that the ν_{CO} and ν_{FeC} frequencies are largely determined by electrostatic interactions with the distal pocket (Oldfield et al., 1991; Li et al., 1994; Ray et al., 1994) and that, in the absence of a significant steric repulsion between the distal side chain and the CO ligand, the true CO binding geometry in the wild type is closer to the heme normal than found in the earlier crystal structures (Ivanov et al., 1994; Ray et al., 1994).

This reinterpretation of the structure-function relationship in myoglobin is, however, incomplete. The shift to lower CO stretching frequencies observed for the wild type His64 MbCO compared with the His64Gly mutant implies, in this model, a polarization of the CO ligand by a *positive* electrostatic potential. The only candidate is the nitrogen-bound hydrogen of the distal side chain, yet the neutron diffraction structure of MbCO shows that the nearest distal nitrogen to the CO ligand, $64N_{\epsilon}$, is not protonated and that the protonated $64N_{\delta}$ nitrogen points *out of* the heme pocket (Cheng and Schoenborn, 1991) (see Fig. 1). In such a distal histidine conformation, the carbonyl ligand experiences a *negative*

Received for publication 11 October 1994 and in final form 23 December 1994.

Address reprint requests to Philip Jewsbury, ZENECA pharmaceuticals, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG, U.K. Tel. 44-1625-582828; FAX: 44-1625-583074.

0006-3495/95/04/1283/12 $2.00

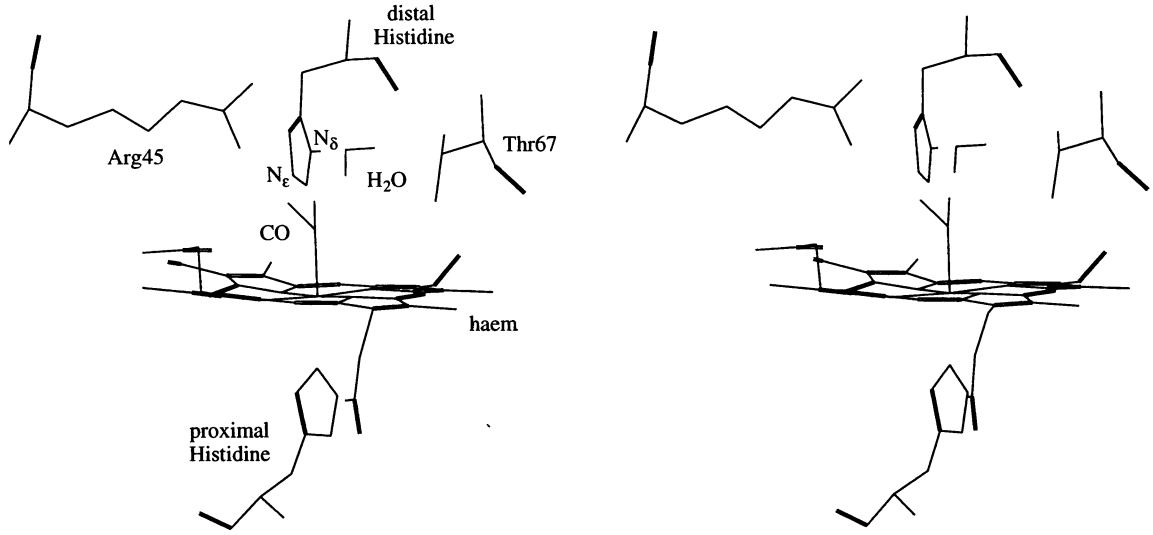

FIGURE 1 Stereo representation of certain key residues in the heme pocket of the MbCO neutron diffraction structure (PDB file 2MB5); two carbonyl positions were resolved.

electrostatic potential that is due to the lone pair of the $64N_\epsilon$ nitrogen. Such a distal conformation is inconsistent with the IR spectrum of a 10-mM wild type MbCO solution frozen to 10 K (Ormos et al., 1988; Ray et al., 1994), and a recent molecular dynamics study has suggested that, in the solution phase, the distal histidine side chain rotates so that the protonated nitrogen points into the heme pocket, interacting with the CO ligand (Jewsbury and Kitagawa, 1994). The extent of this interaction was found to depend on the tautomeric state of the distal side chain: protonation at the N_ϵ site led to a long lived interaction on the time scale of the simulation (90 ps), while protonation at the $64N_\delta$ site showed a much weaker interaction. An "open" structure, where the distal side chain lies outside the heme pocket, was also found in one $64N_\delta H$ trajectory. It was suggested that these three distal side chain conformations correspond in broad terms to the three ν_{CO} stretching frequencies observed in the IR spectra of wild type MbCO. However, a shift to lower ν_{CO} frequencies with stronger distal–CO interaction implies that the distal residue is strengthening the Fe—C bond, i.e., is *stabilizing* the CO ligand. In this model the distal–CO interaction is similar, though weaker, to the distal–O_2 interaction. Thus it is unlikely to be responsible for the relative destabilization of heme-bound CO implied by the orders of magnitude change in the relative affinities of CO and O_2 for the protein compared with those for the free prosthetic group. A recent ab initio study of a model heme prosthetic group has suggested that this destabilization may result from the distorted binding geometry of the proximal ligand (Jewsbury et al., 1994), the complementary axial ligand at the Fe binding site (see Fig. 1), implying that the large strain energies destabilizing the CO ligand are delivered mainly by the protein tertiary structure and not by the mobile distal side chain. Such a mechanism would account for the off perpendicular CO binding geometry that has been found in all MbCO X-ray structures to date, even those of the His64Gly and His64GlyVal68Ala mutants (Quillin et al., 1993).

The solution phase structure of MbCO has recently been determined from NMR distance constraints and chemical shifts (Osapay et al., 1994), and, in contrast to the MD results, the distal side chain conformation is similar to that of the crystal, though displaced further out into the solvent. The determination of the distal histidine orientation, however, allowed for only one distal tautomer and is subject to uncertainties in the Fe—C—O geometry. Since as the solution phase IR spectrum under similar conditions shows two peaks, which have been interpreted as representing two distal side chain orientations/tautomers (the electrostatic interpretation of the distal–CO interaction) or two Fe—C—O conformations (the steric interpretation), it would be interesting to investigate what effect including these possibilities would have on the distal geometry determined from the NMR data. It is difficult to reconcile the MbCO IR spectra, which have multiple peaks suggesting multiple heme pocket–ligand interactions, with the single heme pocket geometries resolved in recent x-ray and NMR studies (Moore et al., 1988; Ormos et al., 1988; Oldfield et al., 1991; Jewsbury and Kitagawa, 1994; Li et al., 1994; Ray et al., 1994).

Therefore, in our current interpretation of the structure–function relationship in myoglobin, the destabilization of the CO ligand is caused by the proximal residue, but the final CO orientation and Fe—C—O vibrational frequencies are determined by electrostatic interactions with the distal pocket. The three ν_{CO} frequencies observed in the wild type IR spectrum (Caughey et al., 1981; Mourant et al., 1993) correspond in broad terms to a stable interaction with the $64N_\epsilon H$ distal tautomer (A_3 peak around 1932 cm^{-1}); a weaker interaction with the $64N_\delta H$ distal tautomer ($A_{1,2}$ peak around 1945 cm^{-1}); and the absence of a distal residue–CO interaction (the "open" form, A_0 peak around 1966 cm^{-1}). This paper in-

vestigates this structural interpretation of the IR spectrum further by considering three distal mutant sperm whale myoglobins for which the x-ray structures have recently become available: a polar distal mutant, His64Gln; an aliphatic distal mutant, His64Leu; and the His64Gly mutant. Two independently generated, 90-ps trajectories have been calculated for each mutant. Water molecules within 16 Å of the carbonyl C have been retained in the data generation, so that the heme pocket can be considered effectively solvated. Protein residues more than 16 Å from the carbonyl C have been weakly restrained to their starting positions to maintain the structural integrity of the solvated protein tertiary structure. The results of these trajectories are discussed with reference to the experimental IR, RR, and structural studies available in the literature. It is found that this recently developed reinterpretation of the distal–CO interaction in wild type MbCO is consistent with the simulation results of the three distal mutants reported here.

METHODS

The methodology has been described in detail elsewhere (Jewsbury and Kitagawa, 1994). Briefly, the initial coordinates for the molecular dynamics simulation were generated from the x-ray structures (including crystal waters) of each mutant (Quillin et al., 1993), Protein Databank (Bernstein et al., 1977; Abola et al., 1987) files 2MGA (His64Gly mutant), 2MGC (His64Leu mutant), and 2MGF (His64Gln mutant), following a series of heating and equilibration steps of the fully solvated protein over 20ps. The final coordinates of this process were then edited to include only those water molecules within 16 Å of the carbonyl C (~135 molecules). Four "solvated" counterions were included in their energy minimized positions close to vacuum exposed side chains. These counterions, solvent molecules more than 12 Å distant from the carbonyl C, and protein residues more than 16 Å from the carbonyl C, were all weakly restrained to their initial positions. This ensures that, while solvent molecules near the active site are free to move, there is no significant "boiling-off" of the solvent, and the solution phase tertiary structure of the protein is conserved during the data generating phase of the simulation. This coordinate set was then used to initiate a 100-ps trajectory, with data collected over the final 90 ps. The standard AMBER 4.0 united atom force field was used for the protein residues (Weiner et al., 1984); a recently developed force field was used for the heme prosthetic group (Jewsbury and Kitagawa, 1994) and TIP3P parameters for the water molecules (Jorgensen et al., 1983). A recent ab initio study has suggested that the nonperpendicular Fe—C—O geometry in MbCOs may result from the nonequilibrium orientation of the proximal residue through the coupling of the axial ligands' motions via the metal *d* orbitals (Jewsbury et al., 1994). In particular this was found to be sensitive to the Fe—N_ϵ—C_{His} angle. The failure of previous molecular dynamics simulations to find significantly distorted Fe—C—O geometries (Henry, 1993; Jewsbury and Kitagawa, 1994), while demonstrating that there is on average little steric influence over the ligand geometry, may therefore be an artifact of these MD force fields not allowing a coupling between these angles. This deficiency has not been rectified in this work, and the previous heme force field has been used unchanged to allow comparison of these trajectories with those reported earlier for the wild type MbCO (Jewsbury and Kitagawa, 1994).

The recombinant mutants whose structures are reported by Quillin et al. (1993) retain their initiator Met residues and have residue 122 as asparagine rather than the aspartate of the wild type, because of an error in their original sequence determination. These "extra" mutations are not expected to affect our MD results, since they are both more than 24 Å from the active site. The residue numbering scheme is the same as in the wild type, i.e., the initiator Met residue is numbered 0.

Two independently generated trajectories, which we label A and B, were calculated for each mutant. The calculations were run using AMBER 4.0 (Pearlman et al., 1991) at the computer centers of the Institute for Molecular Science and the Protein Engineering Research Institute.

RESULTS AND DISCUSSION

The average root mean square deviations of the protein backbone structures from the appropriate x-ray structures during each trajectory were 1.040 ± 0.032 and 0.900 ± 0.040 Å for the His64Gln simulations A and B, 0.806 ± 0.040 and 0.853 ± 0.035 Å for the His64Leu simulations A and B, and 0.849 ± 0.045 and 1.107 ± 0.056 Å for the His64Gly simulations A and B, respectively.

His64Gln mutant: polar interactions at the active site

This mutant is expected to resemble most closely the wild type protein because the glutamine side chain has similar functionality to that of a histidine residue. Fig. 2 shows the separation between the carbonyl C and the distal glutamine side chain as a function of time during the two trajectories: the glutamine side chain remains part of the heme pocket, interacting with the carbonyl ligand throughout the majority of the simulations. If a separation of 5.5 Å is taken as indicative of a "closed" pocket structure, then the side chain forms part of the heme pocket for over 73% and 66% of simulations His64Gln A and B, respectively. In "closed" pocket conformations, the His64Gln side chain is orientated with the 64N_ϵH inside the heme pocket; the 64O_ϵ oxygen points out of the pocket. A representative stereo diagram of the "closed" pocket structure in these simulations is shown in Fig. 3*A*; the x-ray conformation is shown for comparison in Fig. 3*B*.

No long-term stable interaction between the glutamine polar hydrogens and the carbonyl O is formed, as can be seen in Fig. 4 where their separations are shown as a function of time (their average separations are 4.40 and 4.31 Å for His64Gln simulations A and B, respectively); also shown for comparison are the separations found for the "closed" pocket trajectories of each of the distal histidine tautomers from an earlier study of wild type MbCO (Jewsbury and Kitagawa, 1994). In that study, the 64N_ϵH tautomer was found to remain in a "closed" pocket conformation for over 96% of both of the 90-ps simulations and formed a stable electrostatic interaction between the 64N_ϵH hydrogen and the carbonyl O (average separation 3.22 and 2.86 Å in the two simulations reported); the 64N_δH tautomer, however, was found to move more freely between "open" and "closed" conformations (remaining part of the heme pocket for ~70% of the "closed" pocket simulation) and formed only weak interactions with the carbonyl ligand (average separation 5.02 Å). The interaction between the His64Gln side chain and the carbonyl ligand is therefore similar to that found for the "closed" pocket structure of the 64N_δH tautomer of wild type MbCO. This is surprising, because the connectivity of the glutamine side chain more closely resembles the histidine N_ϵH tautomer. Two factors will largely determine how closely the

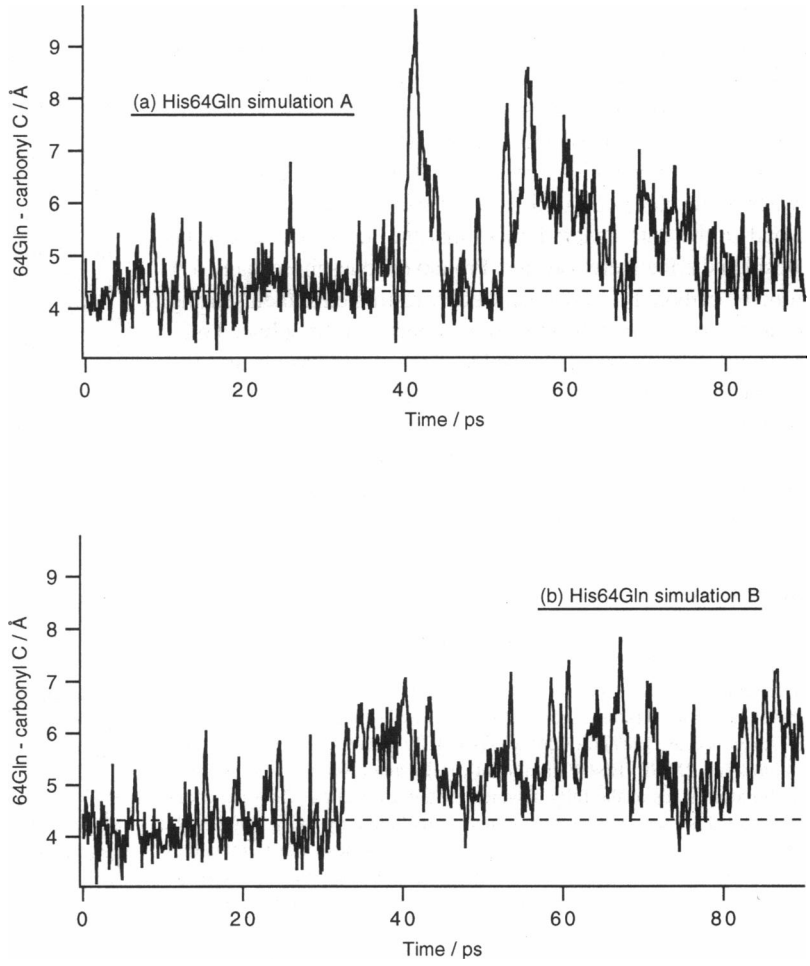

FIGURE 2 Separation between the 64Gln side chain (defined as the average of the C_γ, C_δ, N_ϵ, and O_ϵ positions at a given time t) and the carbonyl C for the His64Gln trajectories. The dotted line shows the separation, 4.32 Å, in the x-ray structure (PDB file 2MGF).

His64Gln distal side chain can mimic the histidine wild type in these molecular dynamics simulations: the similarity of their charge distributions and the extent to which the glutamine side chain can adopt a "histidine-like" conformation under the influence of a weak force field. Table 1 shows the atomic charges assigned to these residues in the AMBER 4.0 force field; it is clear that, while the connectivities of their side chains are similar, their charge distributions are quite different, with the glutamine side chain being significantly more polar. Fig. 5, *A* and *C* shows the spatial overlap of the energy minimized glutamine side chain with that of each of the histidine tautomers: the backbone atoms have been overlaid to give a minimum root mean square (rms) separation. Fig. 5, *B* and *D* shows the best overlap geometries when the conformational energy of the glutamine residue has been minimized subject to weak harmonic restraints (2 kcal/molÅ^2) of its atoms to their equivalents in each histidine tautomer; the equivalent atoms have been overlaid to minimize their rms separation. During the majority of the His64Gln trajectories it is the 64N_ϵH hydrogen *trans* to the 64O_ϵ oxygen that is closest to the carbonyl O (see Figs. 3 and 4). This orientation is preferred because it keeps the negatively charged 64O_ϵ oxygen pointing away from the heme pocket; however, it also prevents a close N_ϵH–carbonyl O interaction because only the *cis* N_ϵH hydrogen in the glutamine side chain can closely overlap the N_ϵH hydrogen of the histidine side chain (see Fig. 5, *A* and *B*). Consequently, although its connectivity more closely resembles that of the histidine N_ϵH tautomer, the dynamics of the glutamine side chain are similar to those of the N_δH tautomer.

These MD results are consistent with those of the IR spectra reported for these MbCOs (Balasubramanian et al., 1993; Li et al. 1994). Wild type MbCO shows two peaks in its room temperature IR spectrum, $A_{1,2}$ (1945 cm^{-1}, 70% intensity) and A_3 (1932 cm^{-1}, 30% intensity); these were interpreted in terms of differing degrees of electrostatic polarization of the carbonyl ligand by the distal side chain (Li et al., 1994). We have recently attributed each peak to one of the two distal histidine tautomers observed under physiological conditions, the $A_{1,2}$ peak to the 64N_δH tautomer and the A_3 peak to the 64N_ϵH tautomer, on the basis of MD simulations (Jewsbury and Kitagawa, 1994). The distal glutamine mutant, on the other hand, shows dynamics similar to those of the wild type 64N_δH tautomer and has only one broad band in its IR spectrum, peaking at 1945 cm^{-1} (Balasubramanian et al., 1993; Li et al., 1994), supporting our interpretation of the wild type spectrum. The IR spectrum of the His64Gln mutant has a higher intensity in the 1950–1965 cm^{-1} range than the wild type, probably as a result of differences in the charge distributions of the two side chains (see Table 1). Attempts to

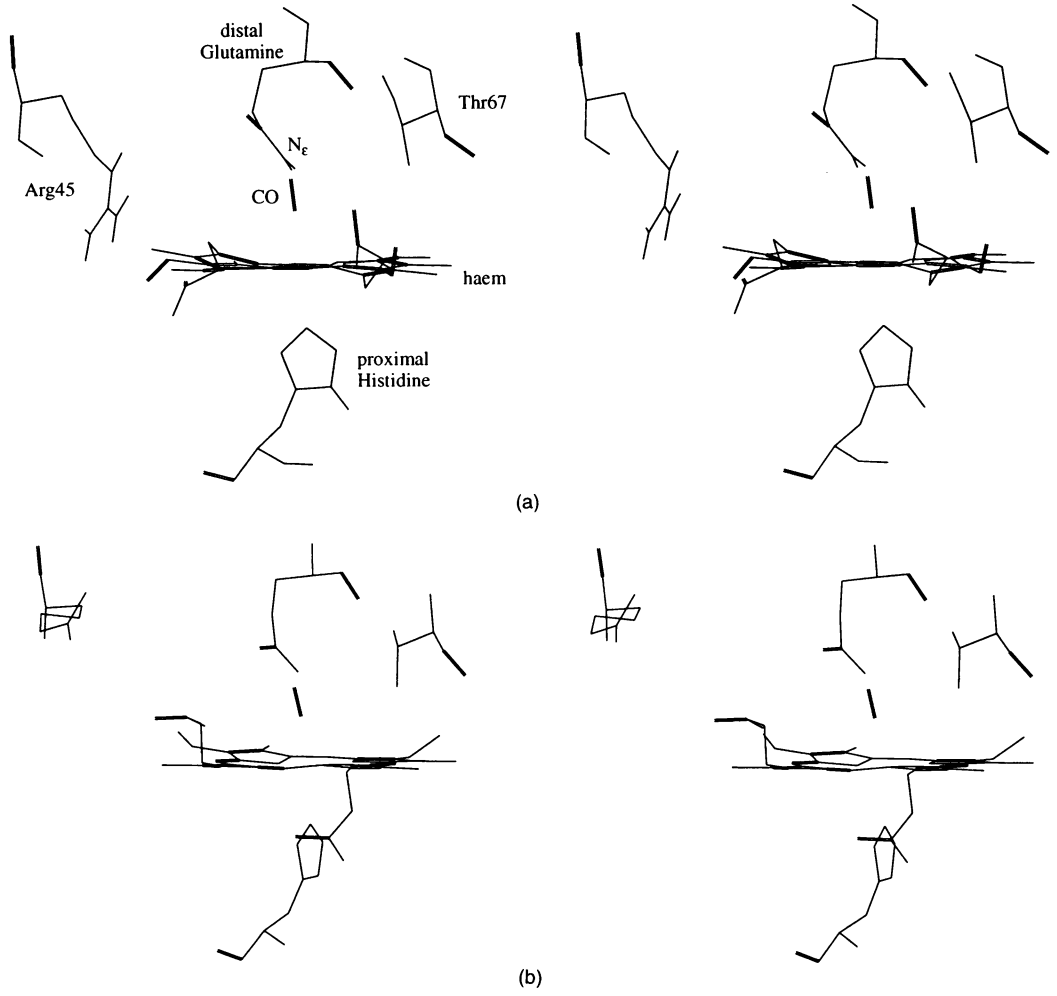

FIGURE 3 Stereo representations of certain key residues in the heme pocket of His64Gln MbCO. (A) The "closed" heme pocket of His64Gln simulation A. (B) The "closed" heme pocket of the x-ray structure (PDB file 2MGF).

investigate the changes in the electric field at the ligand due to the distal mutation by solving the Poisson–Boltzmann equation for the protein structures were inconclusive because of the large field gradient across the heme pocket and the proximity of the protein/solvent boundary.

One simulation of the wild type $64N_\delta H$ tautomer had the distal histidine outside the heme pocket for over 74% of the 90-ps trajectory in an "open" pocket conformation stabilized by a hydrogen bond to the Asp60 main chain O (Jewsbury and Kitagawa, 1994). Such an "open" distal conformation has no distal–CO interaction and has been identified with the A_0 state in the wild type IR spectrum (Morikis et al., 1989). Although this peak accounts for only $\approx$5% of the spectral density at netural pH, the "open" conformation may be functionally important by contributing significantly to the ligand binding process (Tian et al., 1993). The corresponding separations during the His64Gln simulations reported here are shown in Fig. 6; although hydrogen bonds stabilizing the "open" conformation are observed, these are quickly broken, and, unlike for the wild type trajectory, no long-term stable "open" structure develops. The observation of such an "open" structure in a wild type trajectory originating from a "closed" pocket x-ray structure may seem surprising because the interconversion between A states in MbCO is known to occur on the 10-μs time scale. However, the kinetics of processes in proteins are frequently determined by energy barriers, that is, the macroscopic rate constant is small because the system only rarely has sufficient energy in the appropriate degrees of freedom to surmount the barrier. Once the energy is available, the *dynamics* of the reaction/transition takes place on a time scale of picoseconds (Karplus, 1993). Since the distal residue lies on the surface of the protein, the motion of its side chain between the heme pocket and the solvent is largely unhindered; thus the *dynamics* of the "open" $\leftrightarrow$ "close" transition may be on the picosecond time scale observed in the MD simulation.

It should be noted that the "closed" conformation of the x-ray structure 1MBC, used to initiate the wild type trajectories in Jewsbury and Kitagawa (1994), differs significantly from the "closed" conformations found to be stable in the simulations. This will enhance the probability of observing a transition to the "open" structure during the trajectory.

The slow $A_0 \leftrightarrow A_{1,2,3}$ interconversion rate observed experimentally for the wild type, therefore, probably reflects

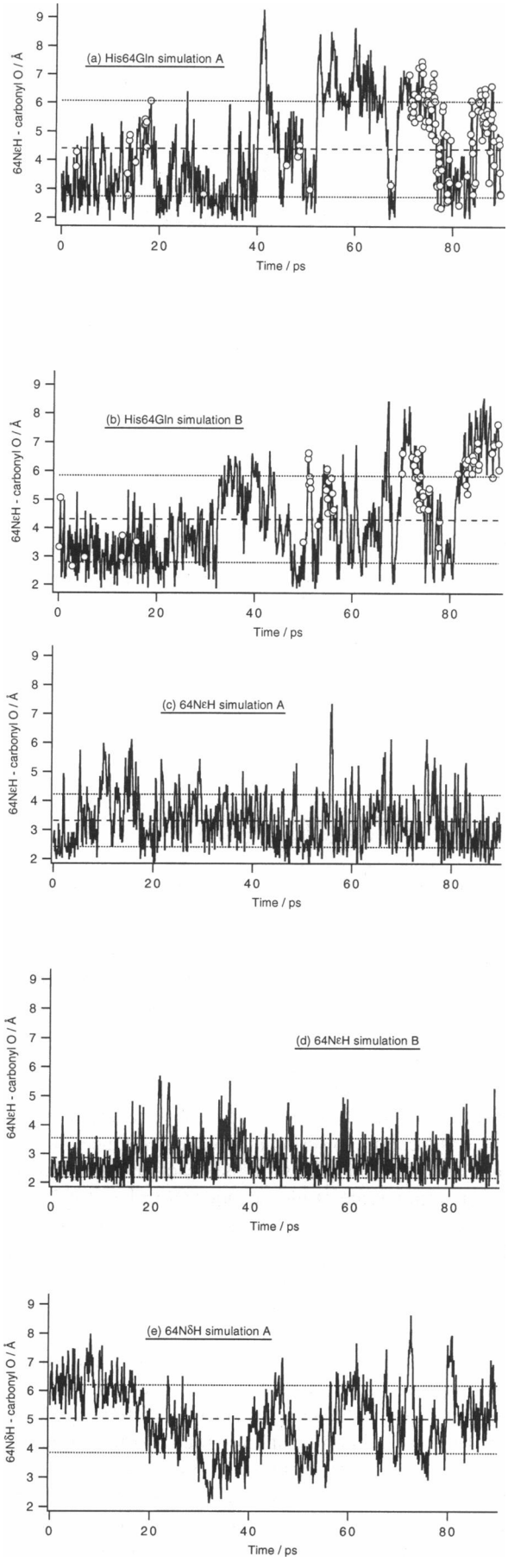

TABLE 1 Atomic charges (in a.u.) for the histidine ($N_\epsilon H$ tautomer), glutamine, leucine and glycine side chains in the Amber 4.0 force field; equivalent atoms listed first.

Atom	His	Gln	Leu	Gly
C_α	0.219	0.210	0.204	0.246
C_β	0.060	0.053	0.016	—
C_γ	0.112	−0.043	0.054	—
C_δ	0.122	0.675	−0.014	—
N_ϵ	−0.444	−0.867	—	—
$H_{N\epsilon}$	0.320	0.344	—	—
C_ϵ	0.384	—	—	—
N_δ	−0.527	—	—	—
O_ϵ	—	—	—	—

the stability of the A state protein conformations. Thus it is significant that, once the "open" pocket has formed in the wild type trajectory, it remains stable on a 100-ps timescale. The $N_\epsilon H$ hydrogen *trans* to the $64O_\epsilon$ oxygen in the His64Gln side chain is the electron acceptor at all hydrogen bonding geometries in the "open" conformer of the mutant because a hydrogen bond to the *cis* hydrogen would bring the distal and Asp60 main chain oxygens close together. However, in the mutant, such a hydrogen bond is weakened, relative to that adopted by the histidine $64N_\delta H$ tautomer, by a less favorable interaction geometry (the glutamine side chain only imperfectly overlaps the histindine $N_\delta H$ tautomer; see Figs. 4, *C* and *D*). In one simulation, His64Gln A around t = 60 ps, the Gln side chain forms a hydrogen bond with the Asp60 side chain; this is, however, short lived because of the inherent mobility of surface side chains. The tendency of the wild type "open" conformation to be stabilized by hydrogen bonding to the *main chain* oxygen offers a possible explanation for its apparent stability and for why the mutation of Asp60 has only a limited effect on ligand binding kinetics (Balasubramanian et al., 1994).

His64Leu mutant: steric interactions at the active site

All aliphatic distal mutants, with the exception of His64Ala (see below), have very similar IR (Balasubramanian et al., 1993; Li et al., 1994) and RR (Sakan et al., 1993) spectra, implying that their distal side chains have little steric influence over the CO binding site. Therefore, although the Fe—C—O unit was resolved as nonperpendicular in the His64Leu X-ray structure (Quillin et al., 1993), there is no steric crowding of the CO binding site by the distal side chain, as can be seen in the space filled model of the active site shown

FIGURE 4 Separation between the distal $N_\epsilon H$ hydrogens and the carbonyl O during the two His64Gln trajectories; only the nearest of the two hydrogens is shown for any given time t (open circles correspond to the hydrogen *cis* to the $64O_\epsilon$ oxygen; otherwise the *trans* hydrogen is closest). The separations found for the closed conformations of each of the wild type distal tautomers are also shown for comparison (Jewsbury and Kitagawa, 1994): the $64N_\epsilon H$ tautomer in (C) and (D) and the $64N_\delta H$ tautomer in (E). Dashed lines show the average separation over the whole trajectory: (A) 4.40 Å, (B) 4.31 Å, (C) 3.32 Å, (D) 2.85 Å, and (E) 5.02 Å; dotted lines show one standard deviation.

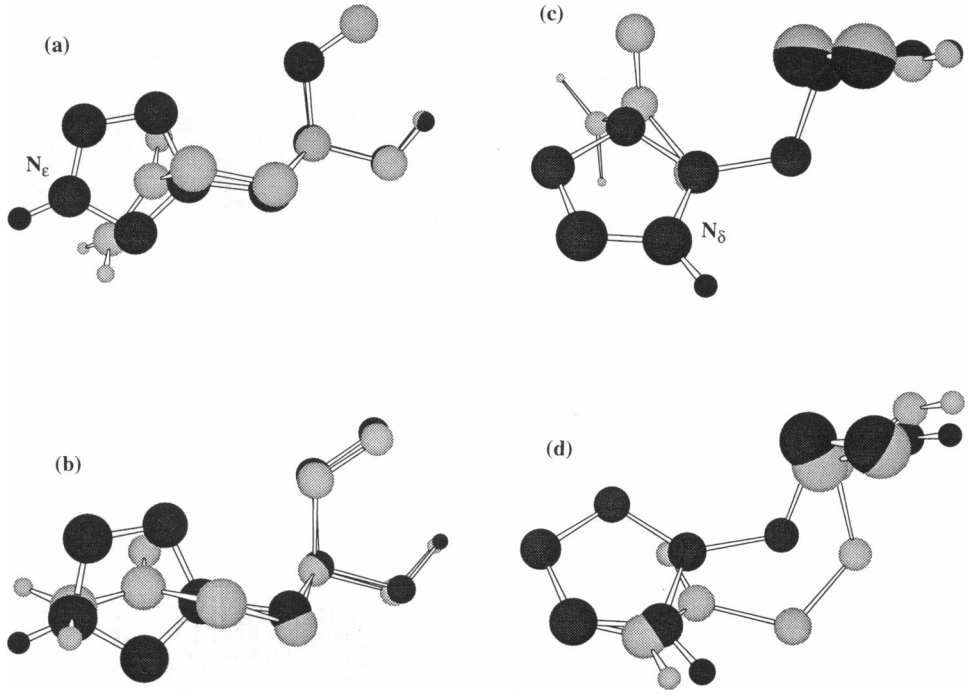

FIGURE 5 Conformational overlap of the histidine (dark gray) and glutamine (light gray) side chains. (A) Overlay of the minimum energy conformations of the glutamine and histidine $N_\epsilon H$ residues (rms deviation of backbone atoms: 0.060 Å). (B) Overlay of the glutamine and histidine $N_\epsilon H$ residues after the glutamine atoms have been restrained during conformational optimization by a 2 kcal/mol $Å^2$ harmonic potential to their geometrical equivalents in an energy minimized histidine template (rms deviation of equivalent atoms: 0.696 Å). (C) As in (A) but in comparison with the histidine NδH tautomer (rms deviation of backbone atoms: 0.007 Å). (D) As in (B) but in comparison with the histidine $N_\delta H$ tautomer (rms deviation of equivalent atoms: 0.407 Å). All conformational energies were calculated using the AMBER 4.0 force field.

in Fig. 7. Since the average position of the ligand atoms is close to the heme normal in both His64Leu simulations (the angles between the CO bond and the heme normal are 5.1° and 3.5° in the average structures of trajectories A and B, respectively), no significant steric interaction between the heme pocket and the CO ligand is found in these trajectories. Thus, these MD simulations are consistent with such a structural interpretation of the IR and RR spectra of these aliphatic distal mutants. In fact, all the simulations reported here, and those of wild type MbCO reported in earlier studies (Henry, 1993; Jewsbury and Kitagawa, 1994), have the average CO position within a few degrees of the heme perpendicular. That is, steric interactions are not primarily determining the distal–CO interaction in any of these MbCO trajectories. In the absence of interactions with residues on the distal side of the heme pocket, or of strain at the proximal side chain, the force fields used in MD simulations have a minimum energy for a perpendicular Fe—C—O conformation. It is possible that the nonperpendicular Fe—C—O geometry observed in all MbCO x-ray structures to date, including those of His64Gly and His64GlyVal68Ala MbCOs, where steric interactions with the carbonyl ligand should be minimized, arises from the orientation of the proximal side chain, an influence over the Fe—C—O conformation that is not allowed for, and therefore not observed in, the MD studies.

The experimental and theoretical evidence therefore suggests that the peak in the IR spectra of the aliphatic distal mutants, which is similar to the A_0 frequency in the wild type spectrum, represents an absence of significant distal–CO interaction. The His64Ala mutant, however, shows broader IR (Balasubramanian et al., 1993; Li et al., 1994) and RR (Sakan et al., 1993) spectra than the other aliphatic distal mutants. This has been attributed to the influence of solvent waters, which may enter the heme pocket on account of the smaller alanine side chain (Li et al., 1994). A similar effect is found in the His64Gly mutant, where solvent waters are resolved as hydrogen bound to the carbonyl ligand in the x-ray structure (Quillin et al., 1993) and are found inside the heme pocket in the MD simulations (see below).

It should be noted that, although the spectral frequencies of these aliphatic distal mutants are similar to those attributed to an "open" distal conformation in the wild type, they do not necessarily represent an "open" structure as such but rather the absence of distal–CO interactions. The His64Leu side chain remains inside the heme pocket for the majority of both the simulations reported here.

His64Gly mutant: solvent waters in the active site

The His64Gly mutant MbCO has a broad IR spectrum, peaking at 1966 cm^{-1} but covering the range 1935–1980 cm^{-1} (Braunstein et al., 1988; Li et al., 1994); the RR spectrum is also broad. Morikis et al. (1989) have assigned the RR spectrum as a convolution of two peaks at 492 and 506 cm^{-1}, which they assigned as the A_0 and $A_{1,2}$ states based on the similarity to their assignment of the wild type frequencies; Sakan et al. (1993), however, did not resolve two peaks but assigned their broad band as peaking at 505 cm^{-1}. The broad nature of the His64Gly spectra has been attributed to the influence of solvent waters in the heme pocket (Li et al., 1994), two of which are resolved in the x-ray structure within hydrogen bonding distance of the carbonyl ligand (Quillin et al., 1993). Fig. 8 shows the distance from the carbonyl ligand to the nearest water molecule during each of the tra-

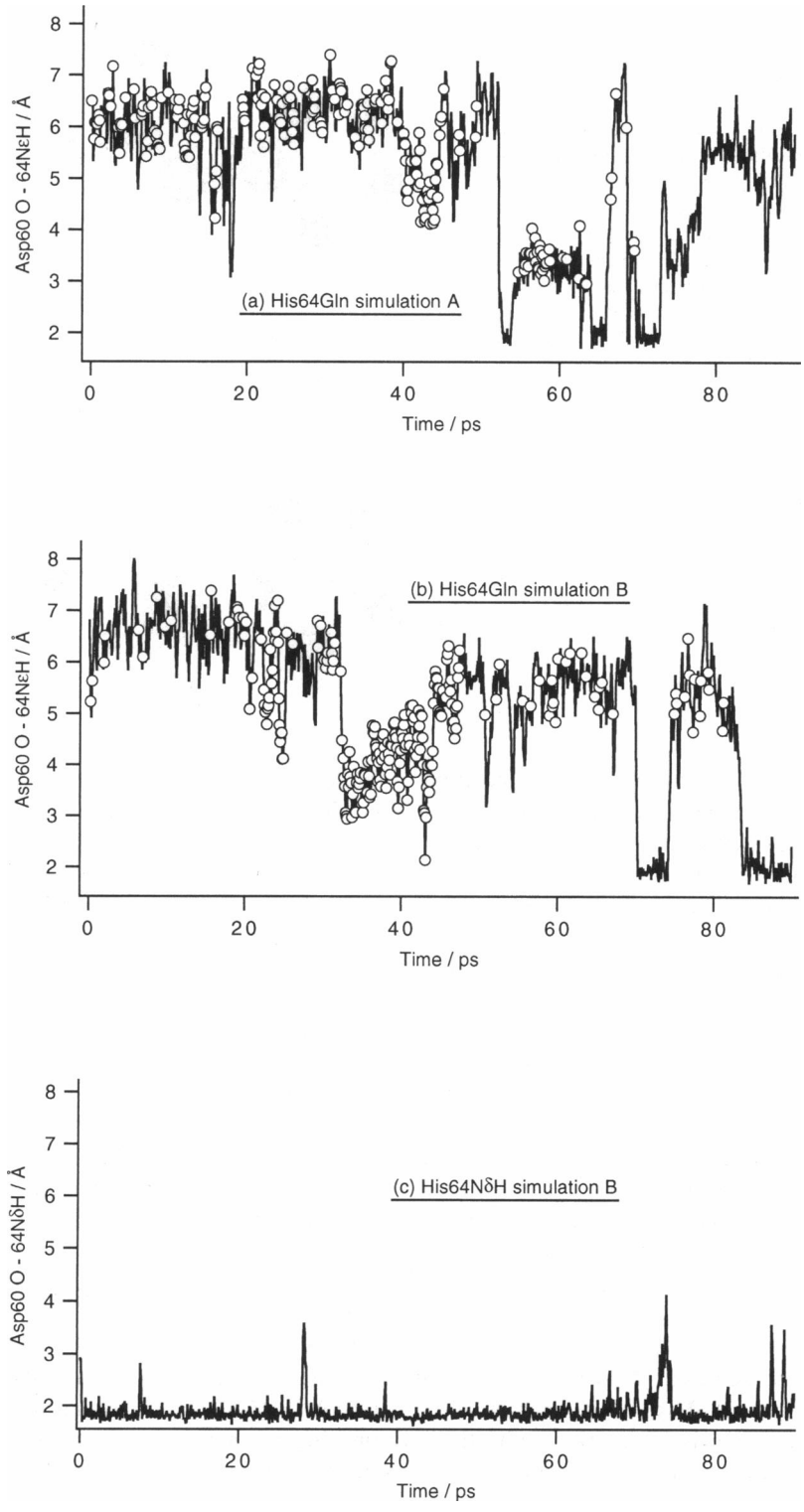

FIGURE 6 Separation between the distal $N_\epsilon H$ hydrogens and the Asp60 main chain O during the two His64Gln trajectories; only the nearest of the two hydrogens is shown for any given time t (open circles correspond to the hydrogen *cis* to the $64O_\epsilon$ oxygen; otherwise the *trans* hydrogen is closest). The separation found in the "open" conformation of the wild type $64N_\delta H$ tautomer is shown for comparison (Jewsbury and Kitagawa, 1994).

jectories reported here and the four wild type trajectories reported in an earlier work (Jewsbury and Kitagawa, 1994). Only the His64Gly mutant trajectories, in particular His64Gly simulation B (Fig. 8*B*), have significant periods with solvent waters close to the carbonyl ligand, consistent with the interpretations of the IR and RR spectra made above. The "closed" pocket simulation of the distal $64N_\delta H$ tautomer of wild type MbCO ($64N_oH$ trajectory A shown in Fig. 8*G*) also has a 20-ps period with a solvent molecule close to the carbonyl ligand; once this molecule leaves the heme pocket, however, no solvent molecules encroach within 6 Å of the carbonyl O for the remainder of the trajectory. Surprisingly, the "open" $64N_\delta H$ tautomer does not have significant periods with solvent molecules inside the heme pocket. It should be

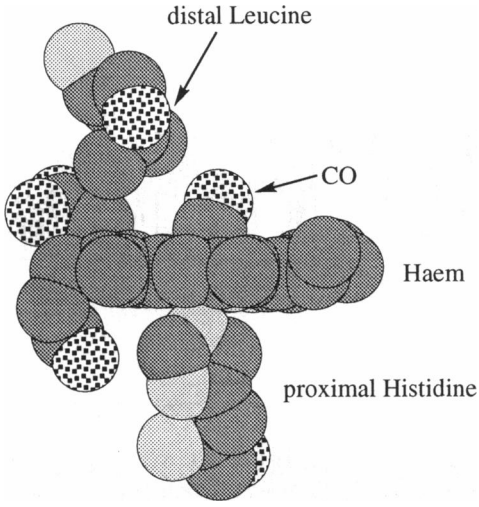

FIGURE 7 Space filled model of the His64Leu active site (PDB file 2MGC).

noted however, that although the protein conformations included the x-ray crystal waters and were fully solvated during the heating and equilibration phases, the motion of the solvent molecules during the data collection phase will not fully mimic that of the true solvent: the solvent restraints used in the methodology, and the approximate nature of the TIP3P force field, in particular with respect to solvent-solute interactions, will lead to artifacts in the solvent motion. However, it appears that solvent molecules on average lie ~6 Å from the carbonyl ligand; only in the case of the His64Gly mutant is significant motion of solvent molecules into and out of the heme pocket observed.

Other distal mutants

The results of these MD studies are consistent with the structural interpretations of IR and RR spectra made in recent experimental studies. This success is partly because the distal residue, since it lies on the surface of the protein, can accommodate a range of side chain mutations with little effect on the protein structure. Thus it is likely that conservative structural interpretations of the spectra of other distal mutants, whose x-ray structures are not available, can be made. The IR spectra of 11 distal mutants are currently available; interpretations of some of these spectra have been made above and elsewhere (Morikis et al., 1989; Oldfield et al., 1991; Balasubramanian et al, 1993; Braunstein et al.; 1993; Jewsbury and Kitagawa, 1994; Li et al., 1994; Ray et al., 1994). It is likely that mutants having a single A_0 peak in their IR spectra have either an "open" pocket structure or a predominantly "closed" pocket structure but with no distal-CO interaction. The two peaks observed in the His64Trp IR spectrum (Li et al., 1994) probably arise one (1969 cm^{-1}) from an "open" structure where the bulky side chain is held outside the heme pocket and the other (1942 cm^{-1}) from a conformation with the $N_\epsilon H$ group inside the heme pocket, though the bulk of the side chain prevents the close interaction with the carbonyl ligand seen for the wild type $64N_\epsilon H$ tautomer. The only other distal mutant reported to have two peaks in its IR spectrum is His64Met MbCO. Braunstein et al. (1993) report only one peak (1967 cm^{-1}) for this mutant at room temperature and pH 7.3, but at cryogenic temperatures, or acidic pH, a second band appears at 1960 cm^{-1}. Li et al (1994) report two bands (1962 and 1953 cm^{-1}) in their neutral pH spectrum. It is possible that the higher-frequency peak results from a predominantly "closed" pocket structure with the terminal methyl group oriented inside the heme pocket. The exposed sulfur lone pairs could be stabilized by interactions with solvent molecules or the Arg45 side chain. Braunstein et al. suggested that the lower-frequency peak could arise from a pH dependent structural change, possibly as a result of protonation of His97 on the proximal side of the heme pocket. Alternatively it could arise from a solvent molecule forming a hydrogen bonded bridge between the distal sulfur lone pairs and the carbonyl oxygen; this may account for its enhancement at low temperatures when the bridging water would be less mobile. Low temperature studies of other distal mutants show that the IR spectra of aliphatic mutants are largely unaltered, though the His64Ala spectrum shows the largest change, presumably due to the reduced mobility of waters in the heme pocket. The His64Gln mutant also shows strong temperature dependence, with a peak shift to 1960 cm^{-1} at 100 K from 1945 cm^{-1} at 300 K (Balasubramanian et al., 1993). It is possible that the same His97 driven structural change in the heme pocket causes this shift to higher frequency at lower temperature in the His64Gln spectrum. This structural change may also be present in the wild type, though its effect on the IR spectrum would be masked by temperature-, or pH-, dependent changes at the distal histidine. It is not clear what this structural change may be; determining the IR spectra of appropriate double mutants would test Braunstein et al.'s proposal. Accounting for this temperature dependence of the IR spectra is a challenge for future studies.

SUMMARY

The distal side chain of His64Gln mutant MbCO shows dynamics relative to the carbonyl ligand that are similar to those found earlier for the "closed" pocket conformation of the wild type $64N_\delta H$ tautomer; in spite of its having a connectivity more closely resembling that of a wild type $64N_\epsilon H$ tautomer, only a weak interaction with the carbonyl O atom was found in both simulations. This is consistent with the IR spectrum of the His64Gln mutant, which peaks at 1945 cm^{-1} (Balasubramanian et al., 1993; Li et al., 1994), similar to the $A_{1,2}$ state of the wild type spectrum that we have previously attributed to the $64N_\delta H$ wild type tautomer (Jewsbury and Kitagawa, 1994). Close interactions between the His64Gln side chain and the carbonyl ligand, similar to those found in the wild type $64N_\epsilon H$ tautomer, are prevented by a repulsive interaction between the $64N_\epsilon$ oxygen and the carbonyl oxy-

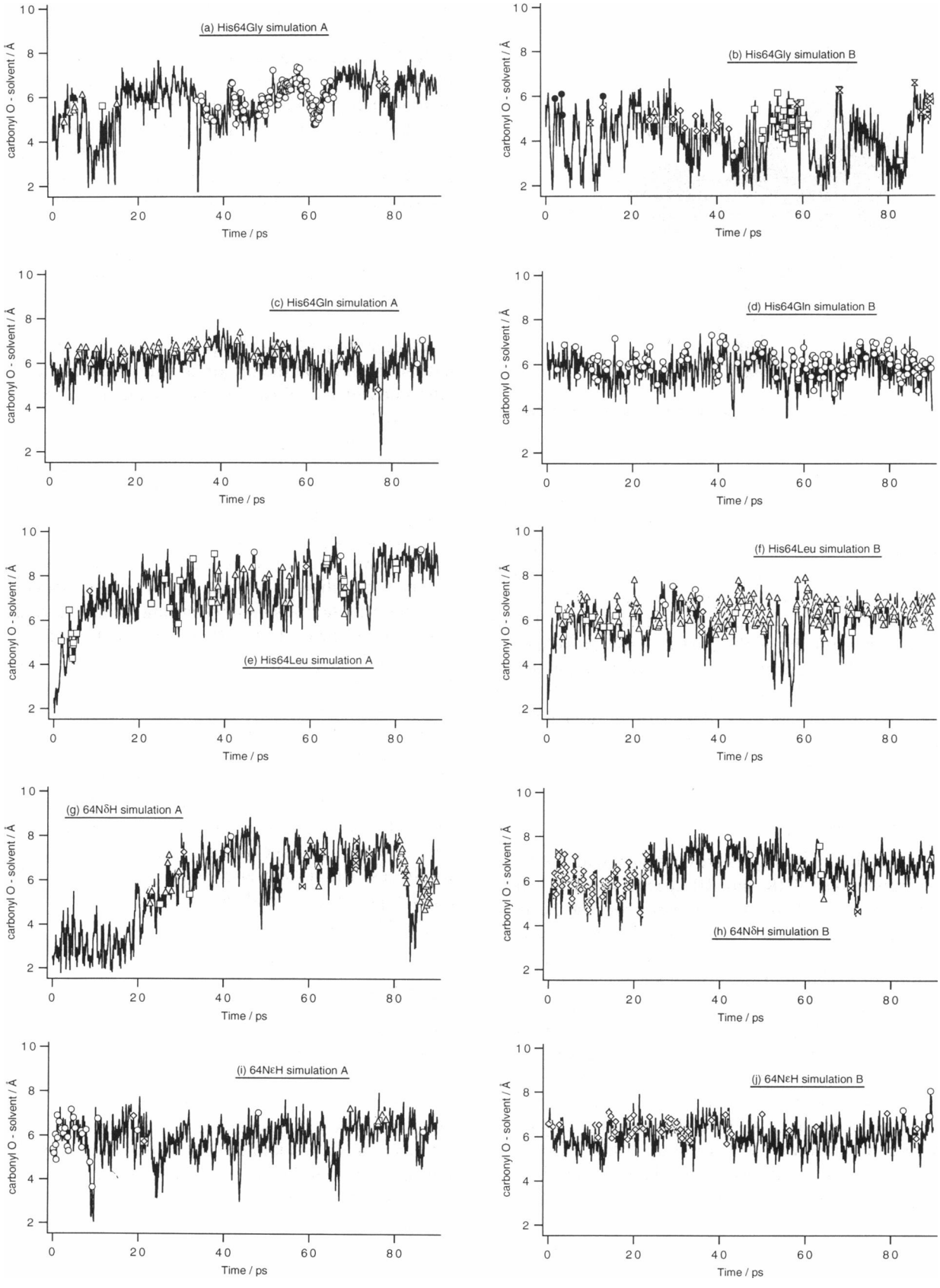

FIGURE 8 Separation between the carbonyl O atom and the nearest solvent atom at a given time for the (A), (B) His64Gly, (C), (D) His64Gln, and (E), (F) His64Leu mutant MbCO trajectories; also shown for comparison are those of the two wild type distal tautomers reported in an earlier study (Jewsbury and Kitagawa, 1994): (G) "closed" pocket trajectory of the $64N_\delta H$ tautomer, (H) "open" pocket trajectory of the $64N_\delta H$ tautomer, and (I), (J) the $64N_\epsilon H$ tautomer trajectories. Open markers indicate where a solvent oxygen atom is closest to the carbonyl O; different markers represent different solvent molecules.

gen. The His64Gln side chain is not found to adopt an "open" structure for significant periods during either simulation, unlike the wild type $64N_8H$ tautomer, because of a poor interaction geometry with the Asp60 main chain O.

The aliphatic side chain of the His64Leu mutant shows little influence over the carbonyl ligand: the average CO position is close to the heme normal, consistent with little steric control over the CO binding geometry. The IR (Balasubramanian et al., 1993; Li et al., 1994) and RR (Sakan et al., 1993) spectra of the distal aliphatic mutants, with the exception of His64Ala, where water is thought to enter the heme pocket (Li et al., 1994), are all very similar to the A_0 state of the wild type spectra. The A_0 state has been interpreted as representing little distal–CO interaction in the wild type (Morikis et al., 1989), a structural interpretation supported by the MD results reported here. Unlike the wild type A_0 state, however, the A_0 peaks in the aliphatic mutant MbCO spectra do not necessarily represent an "open" distal structure; rather they represent an absence of any significant distal–CO interaction.

Both the IR (Li et al., 1994) and RR (Morikis et al., 1989; Sakan et al., 1993) spectra of the His64Gly mutant are significantly broader than those of the aliphatic distal mutants. This has been attributed to the influence of solvent waters (Li et al., 1994), two of which are found in hydrogen bonding positions in the His64Gly X-ray structure (Quillin et al., 1993). Of the three distal mutants reported here, and the two distal histidine tautomers of wild type MbCO reported in an earlier study (Jewsbury and Kitagawa, 1994), only the His64Gly trajectories show significant periods with solvent molecules moving into and out of the active site. In the other cases, solvent water molecules only rarely encroach within 6Å of the carbonyl O.

The results of these molecular dynamics simulations of distal mutant MbCOs are consistent with a recent structural interpretation of the wild type IR spectrum (Jewsbury and Kitagawa, 1994; Li et al., 1994), and the current reinterpretation of the distal-ligand interaction in this important family of proteins as largely electrostatic, rather than steric, in nature (Oldfield et al., 1991; Jewsbury & Kitagawa, 1994; Li et al., 1994; Ray et al., 1994).

The authors thank Dr. M. Saito of PERI, Japan, for stimulating discussions, and Dr G.N. Phillips Jr., of Rice University, Houston, for communicating the mutant MbCO x-ray coordinates prior to their release by the Protein Databank. P.J. thanks the Japan Society for the Promotion of Science for a postdoctoral fellowship under grant 04NP0301 to T.K. for the new program "Intelligent molecular systems with controlled functionality" from the Ministry of Education, Science and Culture of Japan.

REFERENCES

Abola, E. E., F. C. Bernstein, S. H. Bryant, T. F. Koetzle, and J. Weng. 1987. Protein Data Bank. *In* Crystallographic Databases—Information Content, Software Systems, Scientific Applications. F. H. Allen, G. Bergerhoff, and R. Sievers, editors. Data Commission of the International Union of Crystallography, Bonn. 107–132.

Balasubramanian, S., D. G. Lambright, and S. G. Boxer. 1993. Perturbations of the distal heme pocket in human myoglobin mutants probed by infrared spectroscopy of bound CO: correlation with ligand binding kinetics. *Proc. Natl. Acad. Sci. USA.* 90:4718–4722.

Balasubramanian, S., D. G. Lambright, J. H. Simmons, S. J. Gill, and S. G. Boxer. 1994. Determination of the carbon monoxide binding constants of myoglobin mutants: comparison of kinetic and equilibrium methods. *Biochemistry.* 33:8355–8360.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The Protein Data Bank: a computer based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Braunstein, D., A. Ansari, J. Berendzen, B. R. Cowen, K. D. Egeberg, H. Frauenfelder, M. K. Hong, P. Ormos, T. B. Sauke, R. Scholl, A. Schulte, S. G. Silgar, B. A. Springer, P. J. Steinbach, and R. D. Young. 1988. Ligand binding to synthetic mutant myoglobin (His-E7 Æ Gly): role of the distal histidine. *Proc. Natl. Acad. Sci. USA.* 85:8497–8501.

Braunstein, D. P., K. Chu, K. D. Egeberg, H. Frauenfelder, J. R. Mourant, G. U. Nienhaus, P. Ormos, S. G. Silgar, B. A. Springer, and R. D. Young. 1993. Ligand binding to heme proteins: III. FTIR studies of His-E7 and Val-E11 mutants of carbonmonoxymyoglobin. *Biophys. J.* 65: 2447–2454.

Case, D. A., and M. Karplus. 1978. Stereochemistry of carbon monoxide binding to myoglobin and hemoglobin. *J. Mol. Biol.* 123:697–701.

Caughey, W. S., H. Shimada, G. C. Miles, and M. P. Tucker. 1981. Dynamic protein structures: infrared evidence of four rapidly interconverting conformers at the carbonmonoxide binding site of bovine heart myoglobin. *Proc. Natl. Acad. Sci. USA.* 78:2903–2907.

Cheng, X., and B. P. Schoenborn. 1991. Neutron diffraction study of carbonmonoxy myoglobin. *J. Mol. Biol.* 220:381–399.

Collman, J. P., J. I. Brauman, T. R. Halbert, and K. S. Suslick. 1976. Nature of O_2 and CO binding to metalloporphyrins and heme proteins. *Proc. Natl. Acad. Sci. USA.* 73:3333–3337.

Henry, E. R., 1993. Molecular dynamics simulations of heme reorientational motions in myoglobin. *Biophys. J.* 64:869–885.

Ivanov, D., J. T. Sage, M. Keim, J. R. Powell, S. A. Asher, and P. M. Champion. 1994. Determination of CO orientation in myoglobin by single-crystal infrared linear dichroism. *J. Am. Chem. Soc.* 116: 4139–4140.

Jewsbury, P., and T. Kitagawa. 1994. The distal residue–CO interaction in carbonmonoxy myoglobins: a molecular dynamics study of two distal histidine tautomers. *Biophys. J.* 67:2236–2250.

Jewsbury, P., S. Yamamoto, T. Minato, M. Saito, and T. Kitagawa. 1994. The proximal residue largely determines the CO distortion in carbonmonoxy globin proteins. An *ab initio* study of a haem prosthetic unit. *J. Am. Chem. Soc.* 116:11586–11588.

Jorgensen, W. L., J. 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.

Karplus, M. 1993. Preface. *In* Computer simulation of biomolecular systems. Theoretical and experimental applications. Vol. 2. W. F. van Gunsteren, P. K. Weiner, and A. J. Wilkinson, editors. ESCOM, Leiden. xi.

Kuczera, K., J. Kuriyan, and M. Karplus. 1990. Temperature dependence of the structure and dynamics of myoglobin. *J. Mol. Biol.* 213: 351–373.

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 Å resolution. *J. Mol. Biol.* 192:133–154.

Li, T., M. L. Quillin, G. N. Phillips, Jr., and J. S. Olson. 1994. Structural determinants of the stretching frequency of CO bound to myoglobin. *Biochemistry.* 33:1433–1446.

Moore, J. H., P. A. Hansen, and R. M. Hochstrasser. 1988. Iron–carbonyl bond geometries of carboxymyoglobin and carboxyhemoglobin in solution determined by picosecond time-resolved infrared spectroscopy. *Proc. Natl. Acad. Sci. USA.* 85:5062–5066.

Morikis, D., P. M. Champion, B. A. Springer, and S. G. Silgar. 1989. Resonance Raman investigations of site-directed mutants of myoglobin: effects of distal histidine replacement. *Biochemistry.* 28:4791–4800.

Mourant, J. R., D. P. Braunstein, K. Chu, H. Frauenfelder, G. U. Nienhaus, P. Ormos, and R. D. Young. 1993. Ligand binding to heme proteins: II transitions in the heme pocket of myoglobin. *Biophys. J.* 65:1496–1507.

Norvell, J. C., A. C. Nunes, and B. P. Schoenborn. 1975. Neutron diffraction analysis of myoglobin: structure of the carbon monoxide derivative. *Science.* 190:568–570.

Oldfield, E., K. Guo, J. D. Augspurger, and C. E. Dykstra. 1991. A molecular model for the major conformational substates in heme proteins. *J. Am. Chem. Soc.* 113:7357–7541.

Ormos, P., D. Braunstein, H. Frauenfelder, M. K. Hong, S. Lin, T. B. Sauke, and R. D. Young. 1988. Orientation of carbon monoxide and structure-function relationship in carbonmonoxy myoglobin. *Proc. Natl. Acad. Sci. USA.* 85:8492–8496.

Osapay, K., Y. Theriault, P. E. Wright, and D. A. Case. 1994. Solution structure of carbonmonoxy myoglobin determined from NMR distance and chemical shift constraints. *J. Mol. Biol.* 244:183–197.

Pearlman, D. A., D. A. Case, J. C. Caldwell, G. L. Siebel, U. C. Singh, P. Weiner, and P. A. Kollman. 1991. *AMBER 4.0.* University of California, San Francisco.

Phillips, S. E. V., and B. P. Schoenborn. 1981. Neutron diffraction reveals oxygen-histidine hydrogen bond in oxymyoglobin. *Nature.* 292:81–82.

Quillin, M. L., R. M. Arduni, J. S. Olson, and G. N. Phillips, Jr. 1993. High resolution crystal structures of distal histidine mutants of sperm whale myoglobin. *J. Mol. Biol.* 234:140–155.

Ray, G. B., X.-Y. Li, J. A. Ibers, J. L. Sessler, and T. G. Spiro. 1994. How far can proteins bend the FeCO unit? Distal polar effects in heme proteins and models. *J. Am. Chem. Soc.* 116:162–176.

Sakan, Y., T. Ogura, T. Kitagawa, F. A. Fraunfelter, R. Mattera, and M. Ikeda-Saito. 1993. Time-resolved resonance Raman study of the binding of carbon monoxide to recombinant human myoglobin and its distal mutants. *Biochemistry.* 32:5815–5824.

Springer, B. A., S. G. Silgar, J. S. Olson, and G. N Phillips, Jr. 1994. Mechanisms of ligand recognition in myoglobin. *Chem. Rev.* 94:699–714.

Tian, W. D., J. T. Sage, and P. M. Champion. 1993. Investigations of ligand association and dissociation rates in the "open" and "closed" states of myoglobin. *J. Mol. Biol.* 233:155–166.

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, Jr., and P. Weiner. 1984. A new force field for molecular dynamics simulations of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.

Yu, N.-T., and E. A. Kerr. 1988. Vibrational modes of coordinated CO, CN^-, O_2 and NO. *In* Biological Applications of Raman Spectroscopy. Vol. 3: Resonance Raman Spectra of Heme and Metalloproteins. T. G. Spiro, editor. Wiley-Interscience, New York. 39–95.