Cooperative Role of Arg45 and His64 in the Spectroscopic A$_3$ State of Carbonmonoxy Myoglobin: Molecular Dynamics Simulations, Multivariate Analysis, and Quantum Mechanical Computations

Brita G. Schulze and Jeffrey D. Evanseck*

Contribution from the Center of Supramolecular Science, Department of Chemistry, University of Miami, 1301 Memorial Drive, Coral Gables, Florida 33146-0431

Received June 18, 1998. Revised Manuscript Received April 8, 1999

Abstract: The structural, dynamic, and electronic origin of the spectroscopically observed carbonmonoxy myoglobin (MbCO) A states has been investigated by using molecular dynamics to sample conformational space, multivariate analysis to aid in structural interpretations, and quantum mechanics to compute ligand stretch frequencies. Ten short (400 ps) and two longer time (1.2 ns) molecular dynamics simulations, starting from five different crystallographic and solution phase structures centered in a 37 Å radius sphere of water, were used to sample the native-fold of MbCO. Three discrete conformational substates resulted where the primary structural differences corresponded to a variable strength nonbond interaction between His64, Arg45, and the bound ligand. To correlate the structures from the computed substates with the experimentally observed ligand stretch frequencies, Hartree–Fock theory with the 6-31G(d) basis set was used to carry out constrained minimizations and vibrational analysis on representative model geometries from each conformational substate. The A$_0$ state (_out_ conformation) was determined to have both Arg45 and His64 removed from the heme pocket with negligible electrostatic effect on the ligand. Alternatively, His64 was determined to induce the red-shifted frequencies characteristic of the A states (A$_{1-3}$) by forming a weak hydrogen bond between its protonated N$_\delta$ and the ligand (_in/N$_\delta$_ conformation). The A$_{1,2}$ state was specifically assigned to the _in/N$_\delta$_ conformation with Arg45 removed from His64 ($\Delta\nu_{comp} = -10.0 \pm 1.8$ cm^{-1}). The second and faster translational motion engaged Arg45 in an additional and cooperative electrostatic interaction with His64 that distinguished between the A$_{1,2}$ and A$_3$ states. The strongest red-shifted ligand stretch frequency (A$_3$ state) was computed when Arg45 interacted with His64 in the _in/N$_\delta$_ conformation. The polarizing effect of the distal histidine on the CO ligand ($\Delta\nu_{comp} = -19.0 \pm 6.8$ cm^{-1}) was increased by the positive charge on Arg45. Consequently, a new A-state model, which rationalizes the A$_3$ state based upon the fluctuating electrostatic field generated by the gate-like dynamics of His64 and Arg45, is presented, which is consistent with previously reported time scales for substate interconversion.

Introduction

Background. The ability of myoglobin (Mb) to recognize and discriminate between carbon monoxide (CO) and oxygen (O$_2$) is essential to fundamental biochemical processes involving O$_2$ storage and transport.[1] Due to a wealth of structural, kinetic, and spectroscopic information, carbonmonoxy myoglobin (MbCO) has served as a classic paradigm for protein structure–function relationships.[1] In the absence of intermolecular interactions due to the surrounding protein, simple heme models bind CO up to 10^5-fold greater than O$_2$.[2] The overwhelming binding affinity favoring CO decreases by several orders of magnitude when the heme prosthetic group is inserted into the Mb protein matrix.[1] Clearly, the vast array of intermolecular interactions by the surrounding Mb polypeptide assists in the discrimination against endogenous CO inhibition. As a result of recent discoveries, the nature of intermolecular interactions that modulate the CO/O$_2$ binding ratio has become a controversial issue.

The bound CO ligand in MbCO exhibits four IR absorption bands centered at 1965, 1949, 1942, and 1932 cm^{-1}, which are identified as the A states (A$_{0-3}$) with half-bandwidths ranging up to 18 cm^{-1}.[3a–e] The A$_2$ frequency typically appears as a small shoulder of the larger A$_1$ peak, thus both are collectively referred to as A$_{1,2}$. It is widely recognized that the CO ligand stretch frequencies (A states) are used to identify the functionally distinct conformational substates (CSs) of MbCO.[3] Each CS has characteristic energetics and kinetics, and the relative population of each is a function of temperature, pressure, pH, and solvent.[3–5] There have been numerous attempts to understand the structural and electronic aspects of CSs and their connection to the A states, since the relationship is believed to explain MbCO's ability to differentiate between diatomic ligands.[4,5]

(1) (a) Springer, B. A.; Sligar, S. G.; Olson, J. S.; Phillips, G. N., Jr. _Chem. Rev._ **1994,** _94,_ 699. (b) Sage, J. T.; Champion, P. M. In _Comprehensive Supramolecular Chemistry_; Elsevier Science Inc.: New York 1996; Vol. 5, pp 171–213.

(2) (a) Collman J. P.; Herrmann, P. C.; Fu, L.; Eberspacher, T. A.; Eubanks, M.; Boitrel, B.; Hayoz, P.; Zhang, X.; Brauman, J. I.; Day, V. W. _J. Am. Chem. Soc._ **1997,** _119,_ 3481. (b) Collman, J. P. _Inorg. Chem._ **1997,** _36,_ 5145. (c) Gerothanassis, I. P.; Momenteau, M.; Hawkes, G. E.; Barrie, P. J. _J. Am. Chem. Soc._ **1993,** _115,_ 9796. (d) Harutyunyan, E. H.; Safonova, T. N.; Kuranova, I. P.; Popov, A. N.; Teplyakov, A. V.; Obmolova, G. V.; Vainshtein, B. K.; Dodson, G. G.; Wilson, J. C. _J. Mol. Biol._ **1996,** _264,_ 152. (e) Slebodnick, C.; Ibers, J. A. _J. Biol. Inorg. Chem._ **1997,** _2,_ 521.

Early neutron[6] and X-ray crystallographic[7] structures indicate that steric interference from the hydrophobic pocket (see Figure 1) was the origin of the MbCO ligand discrimination.[2] Recent work by Collman[2a,b] and others[2c] on a series of Mb active site porphyrin analogues and high-resolution crystallographic studies[2d] on carbonmonoxy leghemoglobin indirectly support the steric arguments for ligand discrimination in Mb. In addition, photoselection experiments give a correlation between each MbCO A state and unique CO ligand tilt angle.[8] This suggests that CSs can be distinguished based on different degrees of steric hindrance presented to the CO ligand. However, the conclusions of an increasing number of studies are inconsistent with a steric interpretation of ligand distortion. Instead, the polarity of the heme pocket distal residues is believed to influence the bond
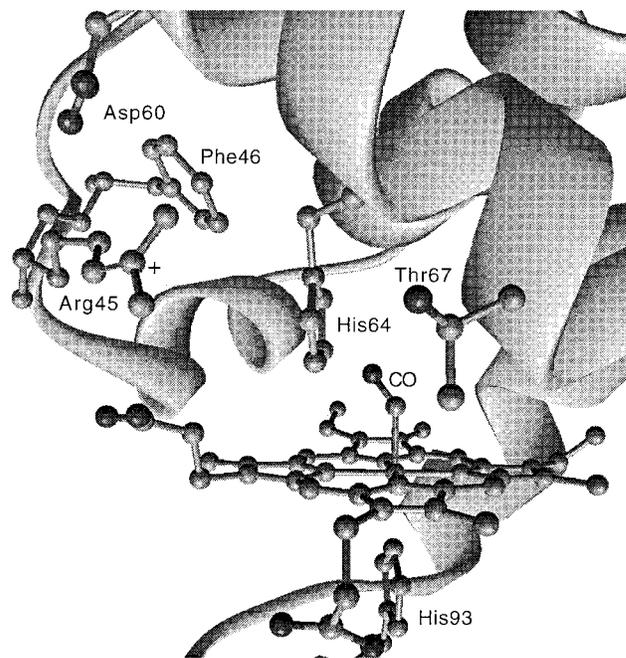
Figure 1. Schematic of MbCO heme pocket created from the X-ray crystallographic structure[7] with WebLab viewer.[46]

order of the ligand through electrostatic interactions.[1,4,5,10−12c,d] In addition, ν_{CO} and ν_{FeC} are inversely proportional to each other, suggesting that the electronic distribution in the Fe−C−O unit is responsible for frequency shifts rather than a steric bending.[5a] Interpretation of recent crystallographic,[13] spectroscopic,[5,10,11] and theoretical[9,12c,d] information shows that the originally reported large magnitude of Fe−CO ligand distortion[6,7] was overestimated. Currently, the majority of experimental and theoretical studies agree that the electronic environment of the CO ligand is controlled by changes in pH, temperature, pressure, and solvent, which in turn determines the different red-shifted IR absorption frequencies of the bound ligand (A states). Consequently, several models have been proposed to explain the A states in terms of a changing electrostatic field around the ligand.[4d,5a,b,12b,19c] However, most of these mechanisms involve a change in protonation that contradicts the time scales established for A-state interconversion.[3p]

(3) (a) Caughey, W. S.; Alben, J. O.; McCoy, S.; Boyer, S. H.; Carache, S.; Hathaway, P. *Biochemistry* **1969**, *8*, 59. (b) Austin, R. H.; Beeson, K. W.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. C.; Marshall, V. P. *Science* **1973**, *181*, 541. (c) Austin, R. H.; Beeson, K. W.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. C. *Biochemistry* **1975**, *14*, 5355. (d) Alberding, N.; Austin, R. H.; Chan, S. S.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. C.; Nordlund, T. M. *J. Chem. Phys.* **1976**, *65*, 4701. (e) Caughey, W. S.; Shimada, H.; Choc, M. G.; Tucker, M. P. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2903. (f) Doster, W.; Beece, D.; Bowne, S. F.; DiIorio, E. E.; Eisenstein, L.; Frauenfelder, H.; Reinisch, L.; Shyamsunder, E.; Winterhalter, K. H.; Yue, K. T. *Biochemistry* **1982**, *21*, 4831. (g) Agmon, N. Hopfield, J. J. *J. Chem. Phys.* **1983**, *79*, 2042. (h) Young, R. D.; Bowne, S. F. *J. Chem. Phys.* **1984**, *81*, 3730. (i) Ansari, A.; Berendzen, J.; Braunstein, D.; Cowen, B. R.; Frauenfelder, H.; Hong, M. K.; Iben, E. T.; Johnson, J. B.; Ormos, P.; Sauke, T. B.; Scholl, R.; Schulte, A.; Steinbach, P. J.; Vittitow, J.; Young, R. D. *Biophys. Chem.* **1987**, *26*, 337. (j) Steinbach, P. J.; Ansari, A.; Berendzen, J.; Braunstein, D.; Chu, K.; Cowen, B. R.; Ehrenstein, D.; Frauenfelder, H.; Johnson, J. B.; Lamb, D. C.; Luck, S.; Mourant, J. R.; Nienhaus, G. U.; Ormos, P.; Philipp, R.; Xie, A.; Young, R. D. *Biochemistry* **1991**, *30*, 3988. (k) Mourant, J. R.; Braunstein, D.; Chu, K.; Frauenfelder, H.; Nienhaus, G. U.; Ormos, P.; Young, R. D. *Biophys. J.* **1993**, *65*, 1496. (l) Post, F.; Doster, W.; Karvounis, G.; Settles, M. *Biophys. J.* **1993**, *64*, 1833. (m) Mayer, E. *Biophys. J.* **1994**, *67*, 862. (n) Gilch, H.; Dreybrodt, W.; Schweitzer-Stenner, R. *Biophys. J.* **1995**, *69*, 214. (o) Zharikov, A. A.; Fischer, S. F. *Chem. Phys. Lett.* **1996**, *263*, 749. (p) Johnson, J. B.; Lamb, D. C.; Frauenfelder, H.; Müller, J. D.; McMahon, B.; Nienhaus, G. U.; Young, R. D. *Biophys. J.* **1996**, *71*, 1563. (q) Galkin, O.; Buchter, S.; Tabirian, A.; Schulte, A. *Biophys. J.* **1997**, *73*, 2752. (r) Franzen, S.; Boxer, S. G. *J. Biol. Chem.* **1997**, *272*, 9655. (s) Tian, W. D.; Sage, J. T.; Champion, P. M.; Chien, E.; Silgar, S. G. *Biochemistry* **1996**, *35*, 3487.

(4) (a) Makinen, M.; Houtchens, R. A.; Caughey, W. S. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 6042. (b) Li, X.-Y.; Spiro, T. G. *J. Am. Chem. Soc.* **1988**, *110*, 6024. (c) Braunstein, D.; Ansari, A.; Berendzen, J.; Cowen, B. R.; Egeberg, K. D.; Frauenfelder, H.; Hong, M. K.; Ormos, P.; Sauke, T. B.; Scholl, R.; Schulte, A.; Sligar, S. G.; Springer, B. A.; Steinbach, P. J.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8497. (d) Oldfield, E.; Guo, K.; Augspurger, J. D.; Dykstra, C. E. *J. Am. Chem. Soc.* **1991**, *113*, 7537. (e) Park, K. D.; Guo, K.; Adebodun, F.; Chiu, M. L.; Sligar, S. G.; Oldfield, E. *Biochemistry* **1991**, *30*, 2333. (f) Braunstein, D. P.; Chu, K.; Egeberg, K. D.; Frauenfelder, H.; Mourant, J. R.; Nienhaus, G. U.; Ormos, P.; Sligar, S. G.; Springer, B. A.; Young, R. D. *Biophys. J.* **1993**, *65*, 2447.

(5) (a) Ray, G. B.; Li, X.-Y.; Ibers, J. A.; Sessler, J. L.; Spiro, T. G. *J. Am. Chem. Soc.* **1994**, *116*, 162. (b) Li, T.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *Biochemistry* **1994**, *33*, 1433. (c) Lim, M.; Jackson, T. A.; Anfinrud, P. A. *Science* **1995**, *269*, 962. (d) Navarro, A. M.; Maldonado, M.; Gonzalez-Lagoa, J.; Lopez-Mejia, R.; Lopez-Garriga, J.; Colon, J. L. *Inorg. Chim. Acta* **1996**, *243*, 161. (e) Unno, M.; Christian, J. F.; Olson, J. S.; Sage, J. T.; Champion, P. M. *J. Am. Chem. Soc.* **1998**, *120*, 2670. (f) McMahon, M. T.; deDios, A. C.; Godbout, N.; Salzmann, R.; Laws, D. D.; Le, H.; Havlin, R. H.; Oldfield, E. *J. Am. Chem. Soc.* **1998**, *120*, 4784. (g) Christina, J. F.; Unno, M.; Sage, J. T.; Champion, P. M. *Biochemistry* **1997**, *36*, 11198. (h) Olson, J. S.; Phillips, G. N., Jr. *J. Biol. Inorg. Chem.* **1997**, *2*, 544. (i) Spiro, T. G.; Kozlowski, P. M. *J. Biol. Inorg. Chem.* **1997**, *2*, 516.

(6) (a) Cheng, X.; Schoenborn, B. P. *Acta Crystallogr.* **1990**, *B46*, 195. (b) Cheng, X.; Schoenborn, B. P. *J. Mol. Biol.* **1991**, *220*, 381.

(7) Kuriyan, J.; Wilz, S.; Karplus, M.; Petsko, G. A. *J. Mol. Biol.* **1986**, *192*, 133.

(8) (a) Ormos, P.; Braunstein, D.; Frauenfelder, H.; Hong, M. K.; Lin, S.-L.; Sauke, T. B.; Young, R. D. *Proc. Natl. Acad. Sci.* **1988**, *85*, 8492. (b) Hong, M. K.; Braunstein, D.; Cowen, B. R.; Frauenfelder, H.; Iben, I. E. T.; Mourant, J. R.; Ormos, P.; Scholl, R.; Schulte, A.; Steinbach, P. J.; Xie, A.-H.; Young, R. D. *Biophys. J.* **1990**, *58*, 429. (c) Moore, J. N.; Hansen, P. A.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci.* **1988**, *85*, 5062.

(9) (a) Jewsbury, P.; Kitagawa, T. *Biophys. J.* **1994**, *67*, 2236. (b) Spiro, T.; Kozlowski, P. M. *J. Am. Chem. Soc.* **1998**, *120*, 4524.

(10) (a) Ivanov, D.; Sage, J. T.; Keim, M.; Powell, J. R.; Asher, S. A.; Champion, P. M. *J. Am. Chem. Soc.* **1994**, *116*, 4139. (b) Sage, J. T.; Jee, W. *J. Mol. Biol.* **1997**, *274*, 21.

(11) (a) Rector, K. D.; Rella, C. W.; Hill, J. R.; Kwok, A. S.; Sligar, S. G.; Chien, E. Y. T.; Dlott, D. D.; Fayer, M. D. *J. Phys. Chem. B* **1997**, *101*, 1468. (b) Morikis, D.; Champion, P. M.; Springer, B. A.; Sligar, S. G. *Biochemistry* **1989**, *28*, 4791. (c) Tian, W. D.; Sage, J. T.; Champion, P. M. *J. Mol. Biol.* **1993**, *233*, 155.

(12) (a) Kushkuley, B.; Stavrov, S. S. *Biophys. J.* **1996**, *70*, 1214. (b) Kushkuley, B.; Stavrov, S. S. *Biophys. J.* **1997**, *72*, 899. (c) Jewsbury, P.; Yamamoto, S.; Minato, T.; Saito, M.; Kitagawa, T. *J. Phys. Chem.* **1995**, *99*, 12677. (d) Ghosh, A.; Bocian, D. F. *J. Phys. Chem.* **1996**, *100*, 6363. (e) Straub, J.; Karplus, M. *Chem. Phys.* **1991**, *158*, 221. (f) Augspurger, J. D.; Dykstra, C. E.; Oldfield, E. *J. Am. Chem. Soc.* **1991**, *113*, 2447. (g) Ma, J.; Huo, S.; Straub, J. E. *J. Am. Chem. Soc.* **1997**, *119*, 2541. (h) Rovira, C.; Kunc, K.; Hutter, J.; Ballone, P.; Parrinello, M. *J. Phys. Chem. A* **1997**, *101*, 8914.

(13) Quillin, M. L.; Arduini, R. M.; Olson, J. S.; Phillips, G. N., Jr. *J. Mol. Biol.* **1993**, *234*, 140.

(14) Ösapay, K.; Theriault, Y.; Wright, P. E.; Case, D. A. *J. Mol. Biol.* **1994**, *244*, 183.

(15) (a) Bashford, D.; Case, D. A.; Dalvit, C.; Tennant, L.; Wright, P. E. *Biochemistry* **1993**, *32*, 8045. (b) Bhattacharya, S.; Sukits, S. F.; MacLaughlin, K. L.; Lecomte, J. T. J. *Biophys. J.* **1997**, *73*, 3230. (c) Bhattacharya, S.; Lecomte, J. T. J. *Biophys. J.* **1997**, *73*, 3241.

The most probable candidate to change its conformation in the immediate ligand environment is His64[1,9] (distal histidine), which is relatively flexible as confirmed by its X-ray B-factors.[7] His64 moves significantly and changes the electronic environment of the CO ligand. As suggested initially by single-crystal Raman spectroscopy,[16e] and later confirmed by an X-ray structure of MbCO at pH 4,[16a] His64 was found to swing out of the pocket and open a channel to the binding site. Furthermore, His64 is thought to control ligand entry, since the heme iron is buried deep within the protein and some type of gating mechanism must be assumed to allow ligand access.[16b] Later, Case and Karplus showed that the energy barrier for ligand escape decreases when His64 moves out of the heme pocket.[16c,d] In addition, the mobility of His64 increases with decreasing size of residue 46, allowing it to rotate about its C_β–C_γ bond or to swing out and away from the ligand.[17] In the X-ray structure of the Phe46Val mutant, His64 is refined in two positions, inside and outside of the heme pocket.[17] Furthermore, mutation of His64 leads to the disappearance of the A_{1-3} bands, and to an increase of A_0 intensity.[4f,5b,11b] This is interpreted such that the outward position of His64 corresponds to the A_0 state. Mutating the residues from the CD corner (Phe43, Arg45, and Phe46) leads to more complex responses, but always correlates with a net increase of the A_0 band, which is thought to arise from the outward movement of His64.[4f,5b] No clear correlation between mutation of individual residues and the relative intensities of the A_{1-3} states has been reported. The importance of His64 is clear, but the influence of other heme pocket residues on the structure and function of MbCO is less straightforward.

Since experimental methods do not describe the A-state phenomenon unambiguously, several theoretical studies have been employed to better understand the A-state origin. Augspurger et al. used high-level ab initio calculations to establish a correlation between external electric perturbations and ν_{CO}.[12f] Similarly, Straub and Karplus recognized the importance of the distal histidine and its protonation state upon ν_{CO} of photodissociated CO.[12e] They used Hartree–Fock calculations to deduce that the B-state frequency shifts originate from a His64 hydrogen bond with the ligand. The CO distortion, energetics, and vibrational characteristics of the Fe–C–O unit have been studied with density functional theory on a model consisting of a porphine, imidazole, and CO.[12d] Proximal His93 tilting by 15° from the minimum did not affect the CO orientation, which is in contrast to that reported by Jewsbury et al.[12c] Kushkuley and Stavrov have reported a systematic investigation of ν_{CO} dependency on the porphyrin ring, CO distortion, Fe atom position in the porphyrin, and external point charges.[12a,b] Similar to Augspurger et al., the most significant effect on ν_{CO} and NMR parameters came from electrostatic interactions between the ligand and a positive point charge. Consequently, the authors

concluded that the A_0 state involved His64 displaced from the hydrophobic pocket, and that the A_{1-3} states involved direct interactions between the ligand and His64 tautomers.[12a,b] Interestingly, the A_3 state was thought to arise from His64 with partial positive charge buildup due to interaction with an unidentified amino acid.

Given the physiological importance and controversial understanding of how myoglobin recognizes and discriminates between diatomic ligands, it is both appropriate and timely to develop a simple and unified understanding of A-state interconversion. Multivariate analysis is used to interpret the complex structural variation from multiple MD trajectories. Quantum mechanical calculations are used to estimate the vibrational frequency changes in the CO ligand from representative CS structures. As a result, a new structure–function model based upon the distal face consistent with previously reported structural and kinetic data is proposed to explain the spectroscopic A states of MbCO. For the first time, Arg45 is invoked to explain the origin of the A_3 state.

Methodology

Atomic Coordinates. Five high-resolution structures of sperm-whale MbCO were selected from the Brookhaven Protein Data Bank:[33] both conformations of an X-ray structure at 1.5 Å resolution where 7 residues

(16) (a) Yang, F.; Phillips, G. N., Jr. *J. Mol. Biol.* **1996**, *256*, 762. (b) Perutz, M. F.; Mathews, F. S. *J. Mol. Biol.* **1966**, *21*, 199. (c) Case, D. A.; Karplus, M. *J. Mol. Biol.* **1979**, *132*, 343. (d) Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 7690. (e) Zhu, L.; Sage, J. T.; Rigos, A. A.; Morikis, D.; Champion, P. M. *J. Mol. Biol.* **1992**, *224*, 207. (f) Ramsden, J.; Spiro, T. G. *Biochemistry* **1989**, *28*, 3125.

(17) Lai, H. H.; Li, T.; Lyons, D. S.; Phillips, G. N., Jr.; Olson, J. S.; Gibson, Q. H. *Proteins: Struct. Funct. Gen.* **1995**, *22*, 322.

(18) (a) Brooks, C. L., III; Karplus, M.; Pettit, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. Advances in Chemical Physics*; J. Wiley & Sons: New York, 1988; Vol. LXXI. (b) McCammon, J. A.; Harvey, S. C. *Dynamics of proteins and nucleic acids*; Cambridge University Press: Cambridge, 1987.

(19) (a) Elber, R.; Karplus, M. *Science* **1987**, *235*, 318. (b) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598. (c) Nienhaus, G. U.; Müller, J. D.; McMahon, B. H.; Frauenfelder, H. *Physica D* **1997**, *107*, 297. (d) Leeson, D. T.; Wiersman, D. A.; Fritsch, K.; Friedrich, J. *J. Phys. Chem. B* **1997**, *101*, 6331.

(20) (a) Clarage, J. B.; Romo. T.; Andrews, B. K.; Pettitt, B. M.; Phillips, G. N., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3288. (b) Hodel, A.; Simonson, T.; Fox, R. O.; Brunger, A. T. *J. Phys. Chem.* **1993**, *97*, 3409. (c) Caspar, D. *Structure* **1995**, *3*, 327. (d) Hunenbeger, P. H.; Mark, A. E.; van Gunsteren, W. F. *J. Mol. Biol.* **1995**, *252*, 492.

(21) (a) Dagett, V.; Levitt, M. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 353. (b) Troyer, J. M.; Cohen, F. E. *Proteins: Struct. Funct. Gen.* **1995**, *23*, 97. (c) Ciocchetti, A.; Bizzarri, A. R.; Cannistraro, S. *Biophys. Chem.* **1997**, *69*, 185. (d) van Gunsteren, W. F.; Hunenbeger, P. H.; Mark, A. E.; Smith, P. E.; Tironi, I. G. *Comput. Phys. Commun.* **1995**, *91*, 305. (e) Cheatham, T. E.; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434.

(22) (a) Elofsson, A.; Nilsson, L. *J. Mol. Biol.* **1993**, *233*, 766. (b) Straub, J. E.; Rashkin, A. B.; Thirumalai, D. *J. Am. Chem. Soc.* **1994**, *116*, 2049. (c) Auffinger, P.; Westhof, E. *Biophys. J.* **1996**, *71*, 940. (d) Caves, L. S. D.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649. (e) Carlson, M. L.; Regan, R. M.; Gibson, Q. H. *Biochemistry* **1996**, *35*, 1125.

(23) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671.

(24) Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893.

(25) Elber, R.; Karplus, M. *J. Am. Chem. Soc.* **1990**, *112*, 9161.

(26) Straub, J. E.; Thirumalai, D. *Proteins: Struct. Funct. Gen.* **1993**, *15*, 360.

(27) (a) Manly, B. F. J. *Multivariate Statistical Methods*; Chapman & Hall: London, 1995. (b) Joliffe, I. T. *Principal Component Analysis*; Springer-Verlag: New York, 1986. (c) Krzanowski, W. J. *Principles of multivariate analysis: A user's perspective*; Oxford University Press: Oxford, 1988.

(28) Koradi, R.; Billeter, M.; Wüthrich, K. J. *J. Mol. Graphics* **1996** 14, 51.

(29) Edwards, S. L.; Poulos, T. L. *J. Biol. Chem.* **1990** 265, 2588.

(30) (a) Diamond, R. *J. Mol. Biol.* **1974**, *82*, 371. (b) Levitt, M. *J. Mol. Biol.* **1983**, *168*, 621. (c) Murray-Rust, P.; Raftery, J. *J. Mol. Graphics* **1985**, *3*, 50. (d) Caves, L. S. D.; Nguyen, D. T.; Hubbard, R. E. In *Molecular Dynamics: Applications in Molecular Biology*; Goodfellow, J., Ed.; Macmillan Press: London, 1991; pp 27–68. (e) Romo, T. D.; Clarage, J. B.; Sorensen, D. C.; Phillips, G. N. *Proteins: Struct. Funct. Gen.* **1995**, *22*, 311. (f) Becker, O. *Proteins: Struct. Funct. Gen.* **1997**, *27*, 213. (g) Abagyan, R.; Argos, P. *J. Mol. Biol.* **1992**, *225*, 519. (h) Murray-Rust, P.; Motherwell, S. *Acta Crystallogr.* **1978**, *B34*, 2534. (i) Domenicano, A.; Murray-Rust, P.; Vaciago, A. *Acta Crystallogr.* **1983**, *B39*, 457. (j) Susnow, R.; Schutt, C.; Rabitz, H. *J. Comput. Chem.* **1994**, *15*, 963. (k) Hayward, S.; Kitao, A.; Go, N. *Proteins: Struct. Funct. Gen.* **1995**, *23*, 177. (l) Garcia, A. E. *Phys. Rev. Lett.* **1992**, *68*, 2696.

(31) Brünger, A. T. X-PLOR, Version 2.1, Yale University, 1992.

(32) Brünger, A. T.; Karplus, M. *Proteins: Struct. Func. Gen* **1988**, *4*, 148.

(33) (a) Abola, E. E.; Bernstein, F. C.; Byant, S. H.; Koetzle, T. F.; Weng, J. In *Protein Data Bank*; Allen, F. H., Bergerhoff, G., Sievers, R., Eds.; Data Commission of the International Union of Crystallography: Bonn/Cambridge/Chester, 1987; p 107. (b) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.
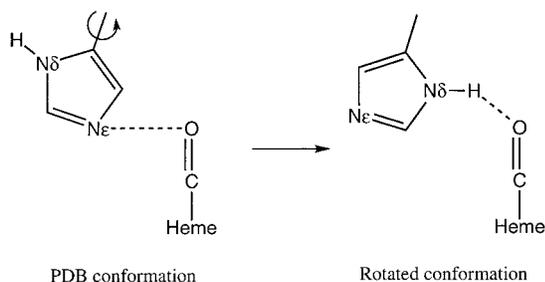
Figure 2. Schematic of the two His64 isomers used as starting points.

and the CO ligand were modeled in 2 conformations (entry 1MBC),[7] a neutron diffraction structure at 1.8 Å resolution (entry 2MB5),[6] and the first two NMR structures from a set of 12 conformations obtained at pH 5.7 (entry 1MYF).[14] Missing hydrogen atom positions were generated with the HBUILD algorithm.[36] Of the 12 His residues, 8 were modeled in their protonated form (i.e., charge +1) consistent with the NMR structures (residues 12, 36, 48, 81, 97, 113, 116 ,and 119).[14] The other histidines residues, 24, 82, proximal (93), and distal (64), were assumed to have the proton at the N_δ nitrogen (HSD).[14] The structures were used as starting points for five trajectories which are called Xray-A$_s$, Xray-B$_s$, Neutron$_s$, NMR1$_s$, and NMR2$_s$. In addition, the distal His64 was rotated in each of the PDB structures about its C_β−C_γ bond (see Figure 2) and five additional trajectories were started from these conformations, called Xray-A$_{s-r}$, Xray-B$_{s-r}$, Neutron$_{s-r}$, NMR1$_{s-r}$, and NMR2$_{s-r}$. A detailed explanation of this modification will be given in the Results Section.

Simulation Protocol. MD simulations were carried out with the EGO_VIII software,[47] which employs a combined structure-adapted multipole method and multiple time step algorithm to efficiently calculate electrostatic forces so that no cutoff has to be employed. The CHARMM all-atom parameter set[35] was used. The crystal waters (137 in the X-ray structures, 89 in the neutron diffraction structure) were

(34) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(35) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(36) Brünger, A. T.; Karplus, M. *Proteins: Struct. Funct. Gen.* **1988**, *4*, 148.

(37) (a) Wiberg, K. B. *J. Am. Chem. Soc.* **1965**, *87*, 1070. (b) Powell, M. J. D. *Math. Prog.* **1977**, *12*, 241. (c) Levitt, M. *J. Mol. Biol.* **1983**, *168*, 595.

(38) Quanta 3.3. Molecular Simulations, Inc.

(39) ERG Program (Eigenanalysis Routines for General Biophysical Analysis) Version 1.0, Schulze, B. G.; Chou, T.; Evanseck, J. D.; University of Miami, 1998.

(40) Press, W. H.; Flannery, S. A.; Teukolsky, W. T.; Vetterling, W. T. *Numerical Recipes*; Cambridge University Press: Cambridge, 1986

(41) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.

(42) (a) van Gunsteren, W. F.; Berendsen, H. J. C *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992. (b) Steinbach, P. J.; Brooks, B. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9135. (c) Lounnas, V.; Pettitt, B. M. *Proteins: Struct. Funct. Gen.* **1994**, *18*, 133. (d) Steinbach, P. J.; Loncharich, R. J.; Brooks, B. R. *Chem. Phys.* **1991**, *158*, 383.

(43) Gaussian 94, Revision B.3, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A., Gaussian, Inc.: Pittsburgh, PA, 1995.

(44) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley and Sons: New York, 1986.

(45) (a) Adachi, S.; Sunohara, N.; Ishimori, K.; Morishima, I. *J. Biol. Chem.* **1992**, *267*, 12614. (b) Balasubramanian, S.; Lambright, D. G.; Marden, M. C.; Boxer, S. G. *Biochemistry* **1993**, *32*, 2202.

(46) WebLab Viewer, Version 2.0, Molecular Simulations, Inc.

(47) Eichinger, M.; Grubmüller, H.; Heller, H.; Tavan, P. *J. Comput. Chem.* **1997**, *18*, 1729.

retained and each conformation was enveloped in a 37 Å sphere of equilibrated TIP3P water molecules.[48] The surface region of the water was subjected to an SBOUND potential[32] to confine water molecules within the sphere. The system was equilibrated in several steps. First, all water molecules were relaxed with a gradient minimizer (400 steps) and then equilibrated for 10 ps at 300 K (protein constrained). Next, the whole system was minimized (400 steps) and subsequently equilibrated for 30 ps at 300 K. During equilibration, the system was coupled to a heat bath with a time constant of 10^{-13} s to achieve the desired temperature of 300 K. For the subsequent MD simulation (370 ps), a weaker coupling (time constant 10^{-11} s) was employed for infrequent and slight adjustment of the temperature. Structural characterization is based only on the 370 ps time period. For visualization in conformational space, structures from the whole 400 ps time were included. Two trajectories (Xray-B$_s$ and Neutron$_s$) were extended to 1200 ps to compare the sampling of conformational space between the shorter and longer time trajectories. Thus, a total simulation time of 5.6 ns was computed. Analysis of the trajectories was carried out by using standard tools of CHARMM[34] and XPLOR.[31]

Collection of Structures and Analysis. Since the extraction of functionally relevant data from such a large set of MD trajectories is not trivial, PCA as a form of multivariate analysis was applied. One structure every 4 ps was collected for each of the 10 short trajectories (7 from equilibration, 93 from simulation) to produce 1000 structures. In addition, 400 snapshots from the extended Xray-B$_s$ and Neutron$_s$ trajectories (400−1200 ps) and the 5 PDB structures were added to yield a total of 1405 conformations. This ensemble was created to visualize the evolution of the MD trajectories, compare the extent of conformational space covered by the short and long trajectories, evaluate the effect of the His64 orientation on global and local protein structure, and determine any significant global or local protein motion. No differences were found when structures were collected more frequently (every 1 ps) or averaged over the 4 ps time period. Average structures of each simulation were computed from MD transient conformations over 30 to 400 ps. An overall average structure was computed from the transients of all trajectories. A PDB average from the 5 PDB structures served as a reference for all structural comparisons. Before analysis, structures were superimposed onto the PDB average to eliminate differences due to rotation or translation. Time series were computed from the trajectories by first reorienting all structures with respect to the PDB average, and then computing the desired property for one frame every 200 fs. This frequency was selected since it suppressed nonspecific high-frequency fluctuations due to bond stretches (time scale 10 to 100 fs), and allowed observation of all other motions.[18]

Proper sampling of a biomacromolecular energy landscape is not trivial.[20] Several methods attempt to probe more of the molecule's conformational space, such as the extension of the simulation to longer time periods,[21] multiple short-time trajectories,[22] simulated annealing,[23] conformational flooding,[24] and locally enhanced sampling.[25] In this study, we report the simulation of multiple trajectories, since recent studies indicate that this approach improves the sampling of conformational space.[22a,d,e] Steinbach and Brooks have demonstrated that although solvation increases the mobility of surface atoms of the protein, it slows global protein motions by introducing friction.[42b,d] Thus, in MD simulations of solvated proteins, the sampling problem becomes even more critical, as shown by the RMSD fluctuations of the 1.2 ns simulations in Figure 3.

Principal Component Analysis (PCA). The multivariate analysis technique PCA can be used to simplify multidimensional data and has been described in detail.[27] It is implemented in the ERG software program.[39] The general approach of PCA is to express a set of given data vectors (correlated to each other) in terms of new vectors (uncorrelated to each other) which might allow a reduction of the data. For example, p Cartesian coordinates $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_p$ of the protein (e.g., the coordinates of the 153 α-carbon atoms of MbCO) in n structures produced during the MD simulations can be arranged in a matrix $\mathbf{X}$ with p columns and n rows. The element X_k^l in the k-th column and l-th row is defined as coordinate k of structure l. The goal is to express

(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
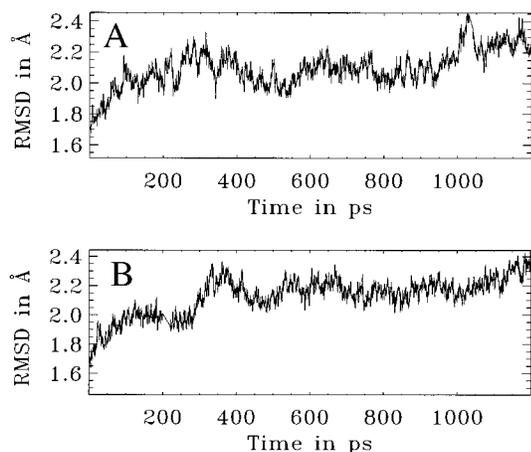
Figure 3. The time series for RMSD based on all heavy atoms for the two long trajectories illustrates that in both cases up to 300 ps are required for the protein to fully relax from its crystal structure.

the original data as a multivariate Gaussian distribution. For this, a covariance matrix **C** is computed from **X**:

$$C_{ml} = \frac{1}{n}\sum_{j=1}^{n}(X_m^j - \langle X_m \rangle)(X_l^j - \langle X_l \rangle) \qquad (1)$$

where j is the index over the n structures, X_m^j is the m-th coordinate in structure j, and $\langle X_m \rangle$ is the average m-th coordinate. By diagonalizing **C**, one obtains its eigenvectors Ψ and the diagonal eigenvalue matrix Λ:

$$\mathbf{C} = \Psi\Lambda\Psi^T \qquad (2)$$

The eigenvalues indicate the magnitude of the variances in the direction of their corresponding eigenvectors. These eigenvectors are sorted by decreasing magnitude of their respective eigenvalue. The original n coordinate vectors collected during MD are projected onto the new basis given by the p eigenvectors resulting in the so-called principal components (PCs). The m-th PC for the first structure can then be expressed as follows:

$$PC_m^1 = \begin{pmatrix} X_1^1 \\ X_2^1 \\ \vdots \\ X_p^1 \end{pmatrix} \cdot \Psi_m \qquad (3)$$

Whereas the original coordinate vectors can be correlated to each other (e.g., coordinates of atoms from the same residue), the PCs are uncorrelated. The structural information in the PCs of the first few eigenvectors $\Psi_1, \Psi_2, \cdots, \Psi_p$ (typically, $p < 5$) is assumed to be most relevant. A plot of the scalar products from eq 3 of all coordinate vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ and the most significant eigenvectors allows the visualization of the data in reduced space. The relative inertia of an eigenvector is defined as the contribution of its eigenvalue to the sum of all eigenvalues. It measures how much of the structural variation in **D** is expressed by the corresponding PC. It has been shown recently for crambin[22d] or for the tetrapeptide isobutyryl-(Ala)$_3$−NH-methyl[30f] that the first three PCs contain already 57% of the total structural variation. The next step involves the visual inspection of the PCA plot to find groups of similar conformation to determine the origin of structural deviation.

   Quantum Chemical Computations. Calculations were performed with Gaussian 94.[43] Polar residues of the heme pocket were examined to better understand their influence on the ligand. The model included CO, His64, Arg45, and the proprionate chain of the heme pointing into the distal pocket. The effect of solvent molecules around the heme pocket was evaluated by including the water molecule most closely interacting with the distal histidine. As shown in Figure 4, His64 was approximated by an imidazole ring, Arg45 by a guanidinium ion (charge
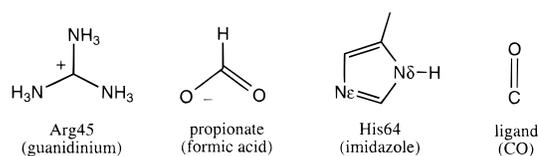
Figure 4. Simplified model illustrating which heme pocket residues were included in the quantum chemical calculations and how they were approximated.

Table 1. RMSDs (in Å) from the Reference Structure (PDB Average)

		trajectory averages			
trajectory	PDB (heavy atoms)	heavy atoms	sphere[c] $r = 10$ Å	sphere[c] $r = 5$ Å	His64
Xray-A$_s$	1.1	1.3	1.2	1.2	2.4
Xray-B$_s$	1.1	1.4	1.2	0.9	1.3
Neutron$_s$	1.1	1.5	1.2	1.3	1.8
NMR1$_s$	1.9	2.1	1.4	0.9	1.1
NMR2$_s$	1.2	1.4	1.1	0.8	1.1
averages[a]		**1.5**	**1.2**	**1.0**	**1.5**
Xray-A$_{s-r}$		1.6	1.2	0.9	1.4
Xray-B$_{s-r}$		1.3	1.1	0.9	1.2
Neutron$_{s-r}$		1.4	1.1	1.0	1.3
NMR1$_{s-r}$		2.2	1.5	1.6	2.6
NMR2$_{s-r}$		1.5	1.0	1.0	1.5
averages$_{s-r}$[a]		**1.6**	**1.2**	**1.1**	**1.8**
av (all trajectories)		**1.6**	**1.2**	**1.0**	**1.6**
total av structure[b]		**0.9**	**0.9**	**0.8**	**1.5**
Neutron$_s$ (1.2 ns)		1.4	1.2	1.3	2.0
Xray-B$_s$ (1.2 ns)		1.5	1.1	0.8	1.1

   [a] Averages of the five trajectories. [b] Computed from *one* structure obtained by averaging over all MD structures. [c] Centered at the oxygen atom of CO.

+1), and the heme propionate by formic acid (charge −1). First, the structural units CO, imidazole, guanidinium ion, formic acid, and water were energy minimized individually by using Hartree−Fock theory and the 6-31G(d) basis set.[44] The MD geometries were rebuilt with use of the energy-minimized components, while retaining the essential geometric features from the representative MD structures (a distance, angle, and dihedral angle for each individual component). With the geometry of the model fixed, the CO bond length was allowed to relax and energy minimize. The CO stretch frequency in each representative conformation was then determined from a vibrational analysis.

Results and Discussion

   MD Structural Comparison. To evaluate how the trajectories evolve and move away from their respective starting experimental structures, each of the five individual PDB structures and twelve trajectory averages were compared to the computed PDB average. The root-mean-square deviations (RMSDs)[30b] for several atom selections are summarized in Table 1. The RMSDs pertaining to the NMR1 structure and corresponding trajectories deviate strongly from those of the other PDB structures and trajectories, which is mainly due to variations in the D helix, as well as in the EF and FG loops. In addition, the table indicates that the orientation of the distal histidine does not influence the evolution of the global or local structure, since the respective heavy atom and His64 RMSDs from the *s* and *s*−*r* trajectories do not deviate significantly. The total average structure, formed from all MD transients, shows best agreement with the PDB reference (0.9 Å). A possible explanation is that each individual trajectory becomes caught in one substate (even if simulated for 1.2 ns) and averaging over an ensemble of independent trajectories yields better agreement with experiment. Thus, the average structure from
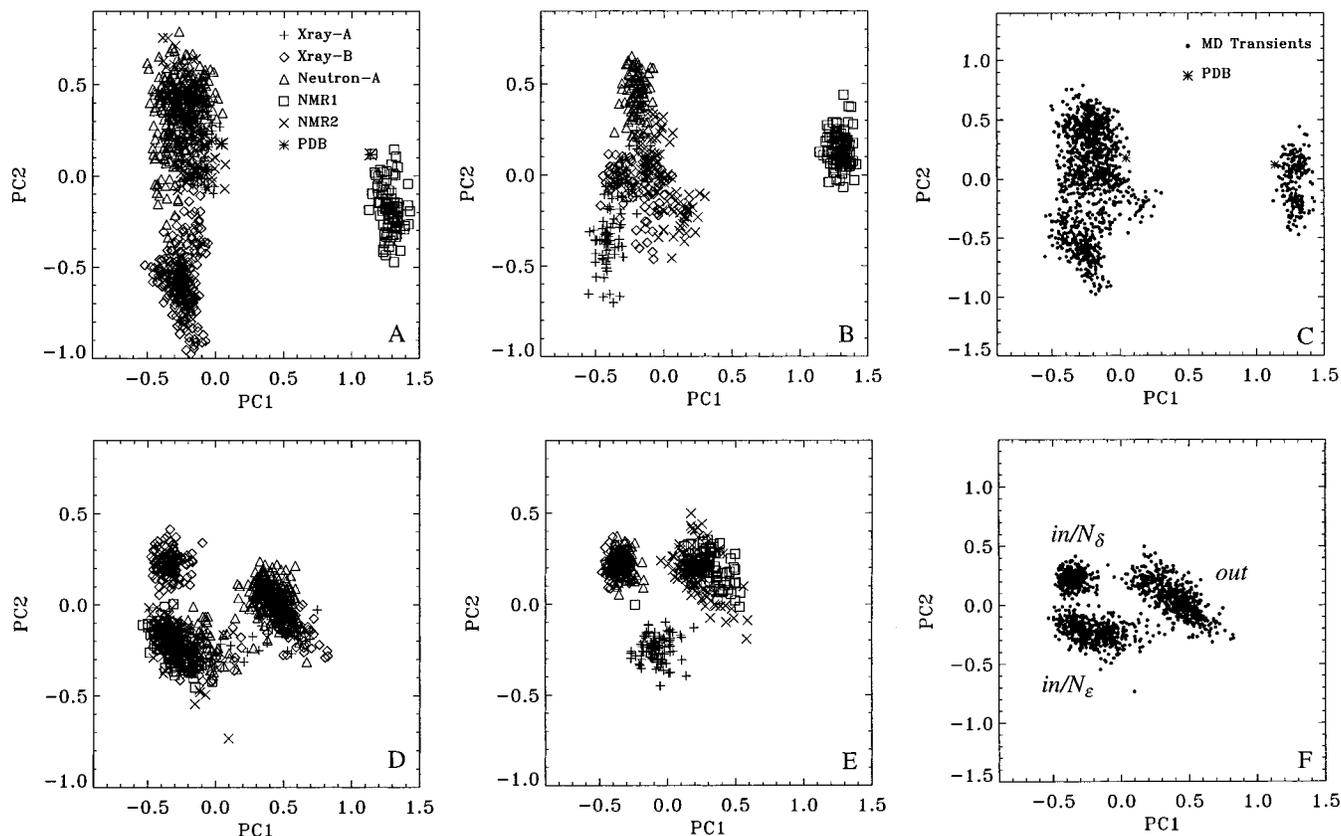
Figure 5. PCA plots of MD transients based on the atomic coordinates of all α-carbon atoms (A−C) and His64 (D−F). In parts A and D only the structures from the *s* trajectories are displayed, in parts B and E only those form the *s−r* trajectories are shown. Parts C and F show all structures from the two ensembles.

the MD simulations is interpreted to better represent the MbCO native-fold, as compared to a single long-time trajectory.

PCA Plots. To visualize global and local motions, PCA was carried out on the α-carbon and His64 atomic coordinates, respectively. The projections of the original 1405 structures onto the first two eigenvectors are shown in Figures 5A−C for the α-carbon atoms and in Figures 5D−F for His64. In Figures 5A−C, the first two PCs display 33.6% of all structural variation, in Figures 5D−F they contain 61.8%. The left-hand plots (5A and 5D) display only the projections of the *s* trajectories, the center plots (5B and 5E) represent the *s−r* trajectories, and the combination of both is displayed in the right-hand plots (5C and 5F).

The most prominent feature in Figures 5A−C is the clear division of structures into two clusters. All structures in the right group derive from NMR1 trajectories, whereas the remaining trajectories fall into the other cluster. Apparently, the structural deviation of the NMR1 PDB conformation in the D region (affecting distal heme pocket) and in the EF and FG regions (affecting proximal side) is so strong that it is maintained during the 400 ps simulation. Thus, the NMR1 PDB structure is also found in the right cluster (coordinates 1.12, 0.12), whereas the remaining PDB conformations are in the other (coordinates around −0.3, 0.45). Extending the simulation for 1.2 ns (Xray-Bₛ, Neutronₛ) did not enhance the sampling of conformational space significantly, as compared to the structural variation introduced into the ensemble by multiple trajectories. Thus, each trajectory stays in a defined region of Figures 5A−C during the entire simulation and no crossing between clusters is found. The orientation of His64 does not have an effect on the structure, since the *s* and *s−r* trajectories populate similar areas.

In Figures 5D−F, three clusters are identified. Visual

inspection of representative structures from each substate shows that the position of His64 differs significantly in these groups. In the lower left cluster, the distal histidine is oriented such that its N$_\epsilon$ atom points toward the oxygen of the CO ligand (refer to Figure 1). This is called the *in/N$_\epsilon$* conformation and is found in all five PDB structures. In the upper left group of conformations, the His64 has rotated around its C$_\beta$−C$_\gamma$ bond and now the hydrogen atom of the N$_\delta$ atom points toward O(CO). This geometry is called the *in/N$_\delta$* conformation. In the right cluster, His64 has undergone a translational movement away from the CO ligand (both N$_\delta$ and N$_\epsilon$ are more than 5 Å from CO) and is referred to as the *out* conformation.

Both in the *s* and *s−r* trajectories, all three clusters are sampled which indicates that the His64 rotation in the *s−r* trajectories was not an irreversible modification and that all three conformations can convert into each other. Whereas in Figures 5A−C each trajectory samples only one specific area, in Figures 5D−F a few trajectories move between the clusters. In Figure 5D, Xray-Aₛ and NMR2ₛ move between the *out* and *in/N$_\epsilon$* clusters, and Xray-Bₛ samples all three groups. In Figure 5E, the NMR1$_{s−r}$ and NMR2$_{s−r}$ trajectories move between the *out* and *in/N$_\delta$* clusters. From a comparison of Figures 5A−C with 5D−F, we conclude that in each case PCA displayed the most prominent structural differences between the ensembles. Global conformations of the PDB structures were maintained (Figures 5A−C) on the time scales of our simulations, and local structural features such as the orientation of His64 (Figures 5D−F) had the ability to change.

His64 Dynamics. The conformational changes of His64 were studied in more detail, since the position and orientation of this residue relative to the CO ligand is thought to be the most decisive factor for the occurrence of the spectroscopic A states

Table 2. Classification of His64 Geometries Based on Distances (Å) between His64 and CO

| | distance H(N$_\delta$)H64−O(CO)[a] | | distance (N$_\epsilon$)H64−O(CO) | | |
trajectory	PDB struct	MD av	PDB struct	MD av	position of His64[b]
Xray-A$_s$	5.6	5.9 ± 0.5	2.7	3.9 ± 0.3	*in/N$_\epsilon$*
		8.0 ± 0.7		7.1 ± 0.9	*out* (72−400 ps)
Xray-B$_s$	6.6	6.6 ± 1.0	3.9	4.8 ± 1.3	*in/N$_\epsilon$*
		3.6 ± 0.6		5.1 ± 0.5	*in/N$_\delta$* (105−668 ps)
		7.6 ± 1.4		6.2 ± 2.2	*out* (669−1200 ps)
Neutron$_s$	6.5	6.6 ± 0.8	3.5	6.2 ± 1.0	*out*
NMR1$_s$	5.0	6.3 ± 0.7	3.1	4.2 ± 0.5	*in/N$_\epsilon$*
NMR2$_s$	5.0	6.5 ± 0.9	3.1	4.6 ± 0.8	*in/N$_\epsilon$*
Xray-A$_{s-r}$	3.2	6.8 ± 0.5	3.8	4.3 ± 0.5	*in/N$_\epsilon$*
Xray-B$_{s-r}$	4.4	3.6 ± 0.4	4.9	4.9 ± 0.4	*in/N$_\delta$*
Neutron$_{s-r}$	4.0	3.5 ± 0.4	4.6	4.8 ± 0.4	*in/N$_\delta$*
NMR1$_{s-r}$	3.3	6.6 ± 0.5	4.3	5.8 ± 0.5	*out*
		4.0 ± 0.5		6.9 ± 0.5	*in/N$_\delta$* (120−400 ps)
NMR2$_{s-r}$	3.3	5.9 ± 1.2	4.3	8.8 ± 1.0	*out*
		4.6 ± 0.6		7.2 ± 0.5	*in/N$_\delta$* (250−400 ps)

[a] Distance measured from the hydrogen to the ligand oxygen. [b] Classification based upon the average structure over the time interval indicated.

in MbCO.[1,4,5,9] In addition, PCA in Figures 5D−F indicates that distinct substates were sampled in the simulations. This high mobility of the distal histidine is also known from crystallography[16a] and from previous MD simulations.[9a,12g]

The distances of the His64 N$_\delta$ and N$_\epsilon$ atoms to the ligand's oxygen atom are summarized in Table 2. For the *s* trajectories, the X-ray (PDB) is *in/N$_\epsilon$* with a N$_\epsilon$(His64)−O(CO) distance between 2.7 and 3.9 Å. Only two trajectories remained in this conformation for the duration of the simulations, while three simulations ended up in the *out* conformation. In one trajectory (Xray-B$_s$), the His64 ring rotated by approximately 120° about its C$_\beta$−C$_\gamma$ bond and formed an *in/N$_\delta$* geometry for approximately 500 ps. Although a complete ring flip was not achieved, the simulation time was sufficient to bring the protonated N$_\delta$ closer to O(CO) than the N$_\epsilon$. We hypothesize that if all the simulations were extended to longer times, such ring flips could occur more often and result in a hydrogen bond between the proton of the N$_\delta$ and the ligand. This was the rationale for us to perform the second set of simulations starting from the PDB structures with a 180° ring flip of His64 (*s−r* trajectories). In the *s−r* trajectories, the X-ray is *in/N$_\delta$* with an HN$_\delta$(His64)−O(CO) distance between 3.2 and 4.4 Å (see Table 2). Here the distal histidine remained in the *in/N$_\delta$* conformation in two trajectories, in two it moved outward and then back to *in/N$_\delta$*, and only in one trajectory it reverted to *in/N$_\epsilon$*. Thus, His64 is observed to be very flexible, which is in agreement with previous work.[16a,9a,12g]

Experimentally, the transition rates between the A$_0$ and A$_{1-3}$ states have been reported to occur between 1.4 × 10^6 (aqueous solution, 273 K) and 2.3 × 10^4 s^{-1} (75% glycerol, 273 K).[3s] Therefore, the transition time between the His64 "open" and "closed" states ranges from 0.71 to 43 μs at 273 K. His64 was computed to switch between the *in* and *out* conformations four times over the 5.6 ns of total simulation time at 300 K (Table 2). Three of the transitions originated from starting structures which could be accelerated due to a strained conformation. Only one true transition (Xray-B$_s$ from 105 to 668 ps) was computed, and its time constant was 563 ps (Table 2). The ca. 10^3 discrepancy between the kinetically determined transition times (0.71 μs at 273 K) and the simulated values reported here (563 ps at 300 K) and by other groups[12c,g] demands further examination. To bridge an appropriate comparison between theory and experiment, it is imperative to realize that the experimental interconversion times between the "open" and "closed" states of MbCO are known to be dependent upon solvent composition and temperature.[3s]

The simulations were carried out at 300 K, as compared to the experiments which were measured at 273 K. The strong dependency of the interconversion times between the "open" and "closed" states on temperature has been previously shown by a study in 75% glycerol solution.[3s] Transition rates of 1.2 × 10^5, 2.3 × 10^4, and 8.1 × 10^3 s^{-1} at 293, 273, and 264 K were measured, respectively.[3s] Simple logarithmic extrapolation of the data to 300 K gives a transition rate of 2.4 × 10^5 s^{-1}, or a transition time of 4.2 μs in 75% glycerol. Thus, the transition time reduces by a factor of 10 (2.3 × 10^4 s^{-1}/2.4 × 10^5 s^{-1}) when considering an increase of temperature from 273 to 300 K in the 75% glycerol solution. By analogy, if one assumes that the same reduction applies in aqueous solution, then the 0.71 μs transition time (1.4 × 10^6 s^{-1}, aqueous solution, 273 K) adjusted to 300 K is estimated to be 69 ns. As a result, the computed transition of 563 ps is ca. 10^2 times faster than the 69 ns value provided by experiment at 300 K. Many factors, such as the ability of the solvent model to represent frictional forces and viscosity effects, incomplete sampling of full conformational change, and the short time scale of the simulations, can rationalize the computed and observed time difference. Due to the limited number of observed transitions in this study, and absence of direct experimental measurement in aqueous solution at 300 K, explanation of the transition time difference between theory and experiment is not attempted in this work. The theoretical assessment of solvent effects on protein dynamics is a continuing topic of research in this laboratory, and will be reported at a later date.

Quantum Chemical Computations. Since the rotation and outward translation of His64 is considered to be a key factor in influencing the vibrational frequency of the ligand, quantum mechanical calculations were carried out. The polar residues of the heme pocket (His64, Thr67, Asp60, Arg45, and propionate chains of the heme) and the water molecule most closely associated with His64 were examined to better understand their influence on the ligand. The residue closest to the ligand is His64. Thr67 was typically 7 Å away from the heme and was therefore excluded from further analysis. Asp60 forms a hydrogen bond with His64 only when the histidine is in the *out* conformation, i.e., relatively far away from CO. Therefore, Asp60 can also be excluded as an influence to the immediate electric field around CO. The proprionate, Arg45, and water interact with CO via His64. Heme pocket conformations were selected randomly from each of the substates as defined by PCA in Figure 5F. The quantum chemical treatment of this model is described in the Methodology Section. Table 3 summarizes the

Table 3. Vibrational Frequency Shifts (cm^{-1}) in CO and Selected Distances (Å)^a

structure	geometry class	N_δ(H64)−O(CO)	N_ϵ(H64)−O(CO)	N_ϵ(H64)−N_η(R45)	N(H64)−O(water)^b	N(R45)−proprionate	freq shift, cm^{-1}
Neutron$_s$, 400 ps	*out*	7.6	7.0	8.0	2.9 (N_δ)	7.8	−2.3
NMR2$_{s-r}$, 80 ps	*out*	7.2	8.8	3.4	2.7 (N_ϵ)	2.7	−2.4
NMR2$_{s-r}$, 120 ps	*out*	7.4	9.5	3.1	3.2 (N_ϵ)	2.8	−1.4
average							**−2.0 ± 0.6**
Xray-A$_s$, 60 ps	*in/N_ϵ*	4.8	3.1	6.3	4.0 (N_δ)	3.0	+10.2
Xray-A$_{s-r}$, 200 ps	*in/N_ϵ*	6.0	4.0	4.9	2.9 (N_δ)	2.9	+0.5
Xray-B$_s$, 60 ps	*in/N_ϵ*	5.3	3.3	6.4	2.8 (N_δ)	2.8	+11.7
NMR1$_s$, 80 ps	*in/N_ϵ*	6.1	4.4	3.6	2.8 (N_δ)	3.0	−6.0
NMR1$_s$, 100 ps	*in/N_ϵ*	5.2	4.1	7.2	3.0 (N_δ)	3.0	+1.1
average							**+3.5 ± 7.3**
Xray-B$_s$, 521 ps*	*in/N_δ*	3.6	4.4	4.8	2.9 (N_ϵ)	2.9	−10.4
Xray-B$_s$, 600 ps	*in/N_δ*	3.5	4.4	7.4	2.9 (N_ϵ)	5.2	−12.6
Xray-B$_{s-r}$, 120 ps*	*in/N_δ*	3.4	4.0	4.9	2.9 (N_ϵ)	2.9	−8.6
Neutron$_{s-r}$, 320 ps	*in/N_δ*	4.3	5.1	5.9	4.5 (N_ϵ)	4.6	−10.4
NMR1$_{s-r}$, 240 ps*	*in/N_δ*	3.3	4.7	4.5	2.8 (N_ϵ)	3.0	−8.2
average (n = 5)							**−10.0 ± 1.8**
Xray-B$_s$, 128 ps	*in/N_δ*	3.6	4.6	3.1	3.0 (N_ϵ)	4.5	−23.9
Neutron$_{s-r}$, 280 ps*	*in/N_δ*	3.2	4.1	4.3	3.1 (N_ϵ)	2.8	−21.9
NMR1$_{s-r}$, 320 ps	*in/N_δ*	4.3	6.4	3.0	2.9 (N_ϵ)	2.7	−11.2
average (n = 3)							**−19.0 ± 6.8**
average (n = 8)							**−13.4 ± 6.0**

^{a} The model includes CO, His64 (imidazole), Arg45 (guanidinium ion), heme proprionate (formic acid), and water. ^{b} Shown in parentheses is the nitrogen atom to which hydrogen bond was formed.

conformations and their characteristic distances. The table lists the difference between the CO frequency computed for each geometry and that of the free CO.

In this approximation, the influence of the heme and its complete environment is neglected. However, previous calculations have shown that variations in the Fe−C bond length or in the proximal histidine do not affect the CO vibrational frequencies to the degree that is observed in experiments.[12b,d] The heme itself does influence ν_{CO} (back-bonding via Fe),[5a,b] but it is not expected to cause the distinct absorption frequencies that have been correlated to the A states. Instead, the heme and especially the attached proximal histidine are thought to play a role in ligand differentiation by affecting the electronic distribution in the iron which prepares it for the approaching ligand.[12i] On the other hand, there is evidence through mutation studies that polar fluctuations in the distal cavity largely determine the CO vibrational shifts.[5b] Based on results of these mutation studies, it appears more critical to evaluate residues in the distal pocket which exhibit large positional fluctuations.

For the *in/N_ϵ* conformations, the CO vibrational frequencies were computed to be blue-shifted by 3.5 ± 7.3 cm^{-1}. The *out* conformation provided an average shift of −2.0 ± 0.6 cm^{-1}. Only the eight *in/N_δ* conformations resulted in consistently large negative frequency shifts (−13.4 ± 6.0 cm^{-1}). Therefore, the red-shifted frequencies of the A states are interpreted to be caused by *in/N_δ* and not *in/N_ϵ* conformations. These results agree qualitatively with frequency calculations on CO−imidazole models by Straub and Karplus, who tested different orientations between CO and His64 and computed a red-shift only when the ligand's oxygen is oriented toward the protonated imidazole nitrogen.[12e]

Further analysis of the data for the *in/N_δ* structures indicated that the frequencies depend on the N_δ(His64)−O(CO) distances (correlation coefficient 0.25). However, better correlation was achieved between the frequency shifts and the sum of the His64−CO and the N_ϵ(His64)−N_η(Arg45) distances (correlation coefficient 0.44), suggesting that the combined motion of both His64 and Arg45 is responsible for the frequency shifts. Therefore, the eight entries in Table 3 for the *in/N_δ* conforma-
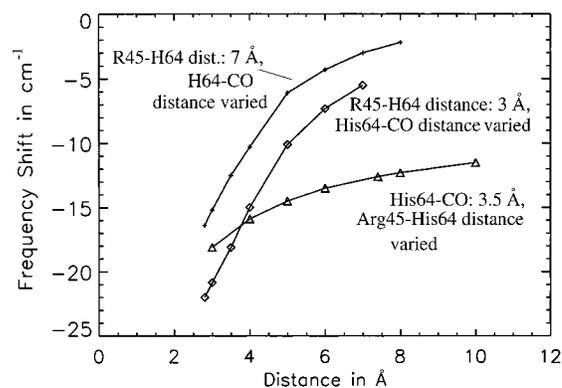
Figure 6. Induced CO frequency shifts dependent upon the proximity of CO, His64, and Arg45.

tions are divided into two groups where Arg45 is either greater than 4.5 Å away from His64 ($\Delta\nu_{CO}$= −10.0 ± 1.8 cm^{-1}) or less than 4.5 Å with a correspondingly higher frequency shift ($\Delta\nu_{CO}$= −19.0 ± 6.8 cm^{-1}).

We picked one *in/N_δ* conformation (Xray-B$_s$, 600 ps) and varied systematically both the N_δ(His64)−O(CO) and N_ϵ-(His64)−N_η(Arg45) distances. The CO frequencies computed for each case are plotted in Figure 6. The data show that in our model system, His64 caused frequency shifts by up to −16 cm^{-1}. With Arg45 close to His64, the frequency shift in CO increased to −22 cm^{-1}. Thus, despite the neglect of the heme or of the remaining protein matrix, which probably led to the systematic underestimation of frequency shifts by 10 to 15 cm^{-1}, the critical role of His64 and Arg45 has been identified. The data show that His64 has a primary effect of red-shifting the CO vibrational frequencies, and that Arg45 has a lesser yet significant role in modulating ν_{CO}. These results are the basis of our A state model, as described below.

Effect of Individual Heme Pocket Components on ν_{CO}. For selected *in/N_δ* conformations listed in Table 3 (indicated with an asterisk), we determined the individual influences of His64, water, Arg45, and the heme propionate upon the CO frequency red-shift (see Methodology). His64 was identified as the primary amino acid that interacted with the ligand, since it

Table 4. Individual Contributions to the Vibrational Frequency Shifts (cm^{-1})[a]

structure	His64	Arg45	water	proprionate	CO freq shift
Xray-B$_s$, 521 ps	+	−	−	−	−8.6
	+	+	−	−	−9.4
	+	+	+	−	−10.7
	+	+	−	+	−9.6
	+	−	+	+	−7.6
	+	+	+	+	−10.4
Xray-B$_{s-r}$, 120 ps	+	−	−	−	−4.3
	+	+	−	−	−10.0
	+	+	+	−	−11.8
	+	+	−	+	−6.8
	+	−	+	+	−3.5
	+	+	+	+	−8.6
Neutron-A$_{s-r}$, 280 ps	+	−	−	−	−15.9
	+	+	−	−	−24.4
	+	+	+	−	−27.0
	+	+	−	+	−18.8
	+	−	+	+	−14.4
	+	+	+	+	−21.9
NMR1$_{s-r}$, 240 ps	+	−	−	−	−4.5
	+	+	−	−	−9.2
	+	+	+	−	−11.5
	+	+	−	+	−6.0
	+	−	+	+	−3.5
	+	+	+	+	−8.2

[a] Distances between the components of each structure are listed in Table 3. Plus (+) indicates its presence in the quantum mechanical calculations, whereas minus (−) indicates its absence.

accounted for $65 \pm 15\%$ of the total CO frequency shift (Table 4). To account for the remaining 35% of the total frequency shift, the heme propionate, water, and Arg45 models were individually removed from the complete system to determine how they modulated the effect provided by His64. Elimination of the water molecule resulted in *a decrease of the ν_{CO} red-shift* by $17 \pm 8\%$ (i.e., red-shifting power of water). Since the simulations resulted with at least one water molecule near His64, the water's contribution to ν_{CO} was considered to be constant. Leaving out the heme propionate yielded *an increase in the ν_{CO} red-shift* by $26 \pm 17\%$ (i.e., blue-shifting effect of the propionate). From the simulation data, the heme propionate was 33% of the time in solution and 67% of the time involved in a salt bridge with the Arg45 residue. Arg45 had the strongest effect, since leaving out the residue resulted in a *decrease of the ν_{CO} red-shift* by $44 \pm 16\%$ (i.e., red-shifting ability of Arg45). The breaking and formation of the Arg45 and heme propionate salt bridge indirectly affects the large polarizing ability of Arg45 on His64 and its ultimate influence on the CO ligand. Thus, the individual contributions of the two red-shifting components, Arg45 and water, and the blue-shifting effect of the heme propionate account for approximately 35% of the total frequency shift, with the remaining 65% attributed to His64.

A$_0$ Substate. From mutation studies, it was found that when His64 is replaced by a smaller residue, only the A$_0$ band remains.[4f,5b,11b] Consequently, the interaction between His64 and CO is minimal in the A$_0$ substate. X-ray crystallography at low pH values (4.5) confirmed this hypothesis, and showed that His64 actually swings out of the heme pocket.[16a] Both our results and previous calculations[9] indicate that the neutral His64 is also very mobile and can move out of the heme pocket. The outward movement of His64 is thought to be biologically relevant because it might create a ligand entry pathway to the heme.[16] Thus, it seems plausible that His64 actually swings out before becoming protonated, making its protonation easier and leading to an increased intensity of the A$_0$ band. From our quantum chemical calculations the His64(N$_\delta$)−CO(O) distance is esti-
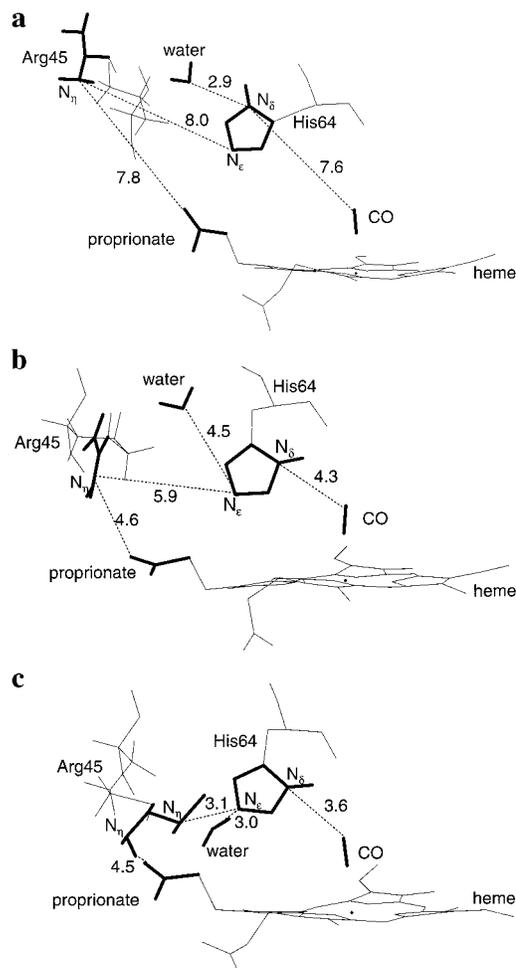
Figure 7. (a) Transient MD structure from the Neutron$_s$ trajectory after 400 ps. This configuration is classified as the *out* conformation with both Arg45 and His64 far removed from the ligand (A$_0$ state). Distances are indicated in angstroms. (b) Transient MD structure from the Neutron$_{s-r}$ trajectory after 320 ps. This configuration is classified as the *in/N$_\delta$* conformation with Arg45 far removed from His64 (A$_{1,2}$ state). Distances are indicated in angstroms. (c) Transient MD structure from the X-ray B$_s$ trajectory after 521 ps. This configuration is classified as the *in/N$_\delta$* conformation with His64 close to the CO ligand and Arg45 close to His64 (A$_3$ state). Distances are indicated in angstroms. All plots were made with MOLMOL.[28]

mated to be greater than 4.5 Å for the A$_0$ state, as exemplified with a snapshot from the Neutron$_s$ trajectory in Figure 7a.

A$_{1,2}$ and A$_3$ Substates. When His64 is inside the hydrophobic heme pocket, vibrational spectroscopy indicates the presence of two conformations,[5] whereas crystallography or NMR cannot resolve the different conformations.[6,7,14] On the basis of observations from vibrational spectroscopy of wild-type and mutant MbCO, several slightly different models have emerged. Common to all of them is the assumption that the different IR frequencies arise from a variation in the C−O bond order, which is inversely proportional to the Fe−C bond order due to back-bonding with Fe.[4b,d,e,5a,b,e] The Fe−C−O unit is polarizable and can be influenced by a close electric field. If no back-bonding is present, then the bonds in the FeCO unit can be formally written as Fe−C≡O$^+$, which is thought to be the situation in the A$_0$ substate. If the oxygen of CO is polarized by positively charged residues or if it forms a hydrogen bond, then it can assume a more negative partial charge. Thus, the bond order of CO decreases and the approximate bonding scheme Fe=C=O emerges. Structures of this type are assumed for the A$_{1,2}$
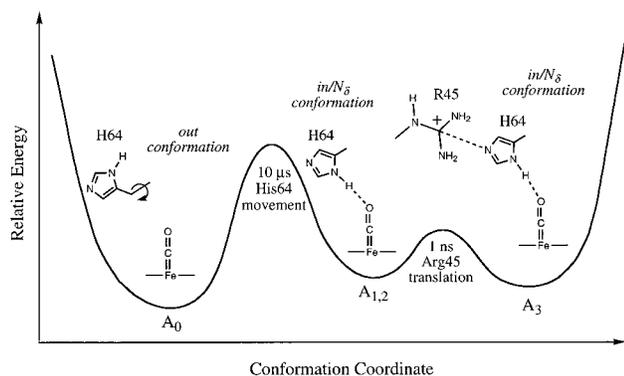
Figure 8. Schematic summary of the three conformations leading to the A_{0-3} states. Barrier heights are drawn based on time scales for A-state interconversion given by Johnson et al.[3p] and are not to scale.

states.[4d,5a,b,12b] Controversy still surrounds the structure of the A_3 state. Some groups have invoked polarization of the ligand's carbon atom by a negative charge (such as a lone pair of electrons at the N_ϵ of His64 or at the oxygen atom of Thr67).[5a,b] Others attribute the frequency shifts to different orientations of the dipole associated with His64.[4d] Recently, it has been suggested that the four A states result from the stepwise protonation of His64 and His97 resulting in a concurrent increase in positive charge around CO.[19c]

On the basis of the current simulation and quantum mechanical data, we propose that the $A_{1,2}$ states in MbCO can be induced by the proximity of His64 (<4.5 Å), with Arg45 relatively far away, as shown in Figure 7b. The A_3 state results when Arg45 comes simultaneously close to His64 (Figure 7c). The data in Table 3 and Figure 6 suggest that the minimum distance between His64 and Arg45 should be below 4.5 Å for the A_3 state. The observation that cytochrome c peroxidase has a CO frequency of 1905 cm^{-1} through a water-assisted hydrogen bond between CO and a distal arginine residue underscores that a positively charged amino acid can have a significant and large effect on ν_{CO} as compared to a neutral residue with polarizing capability.[29] Our assignment is in agreement with the $A_{1,2}$ models established by other groups and with the recent spectroscopic[5e] and computed[49] evidence for a hydrogen bond formed between His64 and CO. However, our A_3 model is novel and contradicts previous assumptions. MbCO can interconvert between the A_0 and $A_{1,2}$ state by the swinging-out motion of His64, and between the $A_{1,2}$ and A_3 states by the appropriate translational motion of Arg45, as schematically shown in Figure 8. The model is in qualitative agreement with the time scales reported, where the $A_0 \rightarrow A_{1-3}$ interconversion times have been estimated to range between 10^{-6} and 10^{-4} s depending upon solvent composition and temperature[3p,s] and the $A_{1,2} \rightarrow A_3$ time at 10^{-9} s.[3p] Thus, the movement of His64 would take ca. 3–5 orders of magnitude longer than the movement of Arg45, which is plausible because His64 is more buried than Arg45 and attached to helix E which is a more rigid unit than the CD loop where Arg45 is located. Also, the distances to be traveled by His64 are larger. The simulations also indicate that the Arg45 interaction with His64 is fast and frequent, since its energy of interaction is weakened due to Arg45 solvation.

It is well-known that solvent can alter the time evolution of atomic motions and molecular properties.[18] Molecular dynamics simulations have been used to show that solvent slows protein motions by random collisions between the solvent and protein, and by introducing friction through a Stoke's-like dissipative force dependent upon the solvent's viscosity, velocity, and size. Consequently, as the residues participating in the dynamical process become more solvent exposed and travel longer distances, the more influential the solvent composition upon their motion. The observed $A_{1-3} \rightarrow A_0$ transition suggests that solvent plays a significant role in the process, since the rates measured for water and water/glycerol (75/25) mixtures differ by 2 orders of magnitude.[3s] The observed solvation effect for the $A_{1-3} \rightarrow A_0$ transition is supported qualitatively by our data, where His64 makes contact with solvent molecules in all trajectories, and the average displacement (*out–in* conformation) of His64 (N_ϵ(H64)−O(CO)) is ca. 4 Å (Table 3). For the $A_{1,2} \rightarrow A_3$ transition, Arg45 spends 33% of its time in solvent and its remaining time in the hydrophobic pocket. The (N_ϵ(H64)−N_η(R45)) distance traveled during this transition is ca. 2 Å (Table 3). Since His64 moves ca. 2 Å more than Arg45, the $A_{1-3} \rightarrow A_0$ interconversion should depend more strongly on the nature of the solvent than the $A_{1,2} \rightarrow A_3$ transition, consistent with experimental results.[3s]

It has been found that the relative population of the A_3 state in the crystal is larger than in solution.[4a] A possible explanation could be that Arg45 in solution is more mobile and interacts with water molecules, whereas in the crystal it is closer to His64 and induces a stronger A_3 band. Furthermore, a triple MbCO mutant with increased aromatic character near the ligand resulted in an enhanced A_3 state population.[5d] The strong cation−π intermolecular interaction[50] between Arg45 and the aromatic residues may anchor the positive charge in place near the ligand, and subsequently red-shift the stretch frequency. The relative population of the A states is also different in human Mb where the A_3 state amounts to only 2−3%.[45] Interestingly, human Mb has a lysine residue in position 45. Since lysine is smaller and structurally different than arginine, the type of interaction with His64 is expected to be less intense and lead to a dramatic decrease in the A_3 intensity. In the Arg45Glu mutant, Li et al. found a decrease in the intensity of the $A_{1,2}$ and A_3 states, and an increase in the A_0 state at pH 7.[5b] The authors attribute this shift in the relative populations to the outward movement of His64 which is no longer confined to the heme pocket. This is further support that the two residues interact with each other. The fact that two distances (His64−CO and Arg45−His64) play a role in the frequency modulation could also explain the broad peaks and large half-bandwidths observed spectroscopically. Obviously, besides the distances, the relative orientation of the groups affects ν_{CO} and leads to large variations in the frequencies even for similar distances (refer to Table 3).

The proposed model avoids the need to invoke protonation changes between N_ϵ and N_δ which have been used repeatedly to account for the frequency shifts.[4d,5a] In the trajectories of this study, solvent approaches only one side of His64 (pointing away from the heme pocket), even after 1.2 ns simulation time. However, to aid in the transfer of a proton leading to a change in protonation from N_ϵ to N_δ, His64 would have to swing out of the heme pocket. Other groups have already pointed out that there is no plausible mechanism for a change in protonation on the time scale of a nanosecond, which is the estimated interconversion time between $A_{1,2}$ and A_3.[3p] We based our choice of protonation at the N_δ atom of His64 on observations by neutron diffraction studies and NMR spectroscopy.[6,14] Recently, it was reported that His64 is primarily protonated at N_ϵ.[15b] The current MD simulations suggest that if His64 were protonated at N_ϵ then the proton would still be close enough to O(CO) to induce the frequency changes observed by IR

(49) Sigfridsson, E.; Ryde, U. *J. Biol. Inorg. Chem.* **1999**, *4*, 99.

(50) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303.

spectroscopy. The important point is that as long as a proton from His64 is close to the ligand it can cause the red-shifted frequencies characteristic of the A states, and at the same time its electrostatic effect on the CO ligand can be enhanced by interaction between its unprotonated nitrogen and Arg45. The in/N_ϵ conformation, as defined in this study, is not involved in the formation of the A_{1-3} states.

Conclusion

On the basis of the simulation and quantum mechanical data, a new model based upon a combination of two side chain movements is established to explain the spectroscopic A states of MbCO. In agreement with previous reports, the geometric positioning of His64 with respect to the hydrophobic pocket was found to define the open (A_0) and closed (A_{1-3}) states of MbCO. It is proposed that the MbCO $A_{1,2}$ states can be induced by a hydrogen bond between His64 and the ligand, and the A_3 state results from an additional electrostatic interaction between Arg45 and His64 while in the in/N_δ conformation. This is in agreement with the recent spectroscopic evidence for a hydrogen bond formed between His64 and the ligand, and a previously reported time scale determination of A-state interconversions. Therefore, the model supports the idea that the origin of the MbCO spectroscopic A states is determined by the fluctuating electrostatic field generated by the dynamics of His64 and Arg45 residues.