# Crystal Structures of Myoglobin-Ligand Complexes at Near-Atomic Resolution

Jaroslav Vojtěchovský,* Kelvin Chu,# Joel Berendzen,# Robert M. Sweet,§ and Ilme Schlichting*
*Max Planck Institut für Molekulare Physiologie, Abteilung Physikalische Biochemie, 44227 Dortmund, Germany; #Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA; and §Biology Department, Brookhaven National Laboratory, Upton, New York 11973 USA

ABSTRACT We have used x-ray crystallography to determine the structures of sperm whale myoglobin (Mb) in four different ligation states (unligated, ferric *aquomet*, oxygenated, and carbonmonoxygenated) to a resolution of better than 1.2 Å. Data collection and analysis were performed in as much the same way as possible to reduce model bias in differences between structures. The structural differences among the ligation states are much smaller than previously estimated, with differences of $<$0.25 Å root-mean-square deviation among all atoms. One structural parameter previously thought to vary among the ligation states, the proximal histidine (His-93) azimuthal angle, is nearly identical in all the ferrous complexes, although the tilt of the proximal histidine is different in the unligated form. There are significant differences, however, in the heme geometry, in the position of the heme in the pocket, and in the distal histidine (His-64) conformations. In the CO complex the majority conformation of ligand is at an angle of 18 $\pm$ 3° with respect to the heme plane, with a geometry similar to that seen in encumbered model compounds; this angle is significantly smaller than reported previously by crystallographic studies on monoclinic Mb crystals, but still significantly larger than observed by photoselection. The distal histidine in unligated Mb and in the dioxygenated complex is best described as having two conformations. Two similar conformations are observed in MbCO, in addition to another conformation that has been seen previously in low-pH structures where His-64 is doubly protonated. We suggest that these conformations of the distal histidine correspond to the different conformational substates of MbCO and $MbO_2$ seen in vibrational spectra. Full-matrix refinement provides uncertainty estimates of important structural parameters. Anisotropic refinement yields information about correlated disorder of atoms; we find that the proximal (F) helix and heme move approximately as rigid bodies, but that the distal (E) helix does not.

## INTRODUCTION

Myoglobin (Mb) is a globular protein of 153 residues that binds molecular oxygen ($O_2$) and other small ligands at a ferrous ($Fe^{II}$) heme iron. Mb is involved in $O_2$ storage and transport in muscle tissues (Antonini and Brunori, 1971; Dickerson and Geis, 1983) and is an important model system for studying the physics and dynamics of reactions in proteins. We address four open questions about the structure, function, and dynamics of Mb with improved structural data on the molecule in different ligation states.

### What are the geometries of bound CO ligands?

Although the physiological function of Mb appears to be $O_2$ binding, carbon monoxide (CO) is also a biologically significant ligand for Mb and other heme proteins because it is an endogenous poison. Most of the hemes in the human body would be poisoned if the specific affinity for CO over $O_2$ ($K_{CO}/K_{O2}$) were as high in heme proteins as it is in model compounds (e.g., protoheme). Differences between CO and $O_2$ binding to heme proteins are thus physiologically relevant and constitute a key to understanding the relationship between structure and function.

The textbook explanation of the lowered specific affinity of heme proteins focuses on residue His-64, the distal histidine, which forms the side of the ligand binding pocket closest to the bound $O_2$. The distal histidine moderates the specific affinity for CO over $O_2$ by providing a hydrogen bond to $O_2$ ligands, and it is said to sterically hinder binding of CO ligands (CO prefers to bind perpendicularly to the heme plane, whereas $O_2$ prefers to bind at a slight angle to the heme normal in unencumbered model compounds). However, the latter part of this explanation cannot be correct because it predicts that the effect of the protein is mostly in the "on" rates for CO binding, when in fact the largest contribution to the lowered specific affinity is the lowered "off" rate for $O_2$ (Springer et al., 1994). Moreover, structural studies of mutants of the distal histidine with smaller side chains at that position show CO bound similarly off-axis as to wild-type MbCO (Quillin et al., 1993). Understanding the origins of the ligand-binding geometry is crucial to understanding function, but at present there is some dispute about what the geometry is.

Previous studies of MbCO have disagreed about the geometry of bound CO, as seen in Table 1. Early diffraction studies refined a single CO conformation. Later, evidence

*Received for publication 30 December 1997 and in final form 15 June 1999.*

Address reprint requests to Dr. Ilme Schlichting, Max Planck Institut für Molekulare Physiologie, Abteilung Physikalische Biochemie, Otto-Hahn-Str. 11, 44227 Dortmund, Germany. Tel.: 49-231-133-2738; Fax: 49-231-133-2699; E-mail: ilme.schlichting@mpi-dortmund.mpg.de.

Kelvin Chu's present address is Dept. of Physics, Cook Building, University of Vermont, Burlington, VT 05405-0125.

Coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank (access codes 1A6G, 1A6M, 1A6K, 1A6N).



0006-3495/99/10/2153/22 $2.00

**TABLE 1 Determinations of the angle between the C—O bond and the heme plane**

| Technique | 1st author | Date | Angle(s) (°) |
|---|---|---|---|
| X-ray diffraction | Kuriyan | 1986 | 40 |
| | | | 61 |
| IR photoselection | Ormos | 1988 | 75 ± 4 |
| | | | 62 ± 2 |
| | | | 57 ± 4 |
| IR photoselection | Moore | 1988 | 20 ± 3.5 |
| | | | 35 ± 3.5 |
| Neutron diffraction, pD 5.7 | Cheng | 1991 | 47 |
| | | | 47 |
| X-ray diffraction,* pH 7 | Quillin | 1993 | 19 |
| IR optical crystallography | Ivanov | 1994 | <10 |
| X-ray diffraction,* pH 9 | Schlichting | 1994 | 32 |
| IR photoselection | Lim | 1995 | 0 ± 7 |
| X-ray diffraction, pH 6.0 | Yang | 1996 | 42 |
| X-ray diffraction, pH 4.0 | Yang | 1996 | 30 |
| IR optical crystallography | Sage | 1997 | 6.7 ± 0.9 |
| X-ray diffraction, pH 6.0 | This work | 1998 | 18 ± 3 |

*This determination was on a mutant Mb in a hexagonal form; all others were carried out in monoclinic crystals. Multiple entries indicate a claim for multiple conformations of the ligand.

was found for multiple conformations that differed substantially in orientation with respect to the heme plane (Kuriyan et al., 1986; Cheng and Schoenborn, 1991). More recently, only a single orientation has been seen again, with a wide range of conformations that suggest considerable experimental uncertainty (Quillin et al., 1993; Schlichting et al., 1994; Yang and Phillips, 1996).

Infrared linear dichroism studies have obtained differing results about the bound CO geometry. MbCO has three strong absorbance bands in the mid-infrared due to the carbonyl stretch that have provided a wealth of information about MbCO (Alben et al., 1982). Solution measurements of linear dichroism after photoselection can determine the angle between the C–O stretch transition dipole moment (thought to lie along the C–O axis) and the heme transition dipole moment in the visible (which lies in the heme plane) (Hofrichter and Eaton, 1976). Early experiments of this type indicated multiple conformations of the CO, with the angle between the CO and the heme plane similar to that seen in the diffraction studies of the time (Ormos et al., 1988; Moore et al., 1988). However, more recent work on photoselection indicates that the CO transition dipole moment has a single orientation in all three bands that is close to normal to the heme plane (Lim et al., 1995). Measurements of static linear dichroism in crystals gave similar results, with the most recent estimate of the angle between the CO transition dipole moment and heme plane at 6.7 ± 0.9° (Ivanov et al., 1994; Sage and Jee, 1997; Sage, 1997). These findings appear to disagree with all previous determinations of the CO geometry made by diffraction methods, but that is not clear because previous determinations do not include uncertainty estimates. The spectroscopic camp has typically pointed to the large scatter in the diffraction results (Ray et al., 1994) while the diffraction camp lays claim to having a much more direct measurement of geometry. New, better data on CO binding geometry, with uncertainty estimates and preferably free from model bias and restraints, are needed to resolve this question.

## What are the structural differences among the infrared A substates of MbCO?

The infrared absorption spectrum of CO in MbCO shows three distinct bands that are conventionally labeled $A_0$, $A_1$, and $A_3$ (Alben et al., 1982). Each band shows distinct kinetics of CO rebinding after flash photolysis, thus implying different functional properties (Alben et al., 1982). The populations of the bands at cryogenic temperatures are sensitive to a variety of external conditions, including pH (Fuchsman and Appleby, 1979; Müller, 1997), cooling rate (Chu et al., 1993) and crystal form (Makinen et al., 1979; Mourant et al., 1993). $A_0$ is favored at pH lower than 4.6, but the ratio of $A_3$ to the dominant $A_1$ in solution is nearly independent of pH above 6 (Fuchsman and Appleby, 1979; Müller, 1997), with $A_1$ dominant by almost a factor of 10. The A substates interconvert reversibly at room temperature on the microsecond time scale, but exchange is frozen out below the glass transition temperature of the protein and solvent near 175 K (Young et al., 1991).

When different CO conformations were identified in MbCO by diffraction studies (Kuriyan et al., 1986; Cheng and Schoenborn, 1991) and IR photoselection/linear dichroism experiments (Ormos et al., 1988; Moore et al., 1988), the A substates were thought to be associated with different CO orientations. Recent spectroscopic data (Lim et al., 1995) make this explanation appear unlikely, and a calculation of the energies required to bend the Fe–CO bond also cast doubt on this interpretation (Ghosh and Bocian, 1996). The $A_0$ substate has been identified by diffraction at low pH with a doubly protonated conformation of the distal histidine, which is swung out of the pocket toward the solvent (Yang and Phillips, 1996). Presently, it is believed that $A_1$ and $A_3$ are associated with different conformers or protonation states of the distal histidine (Park et al., 1991; Ray et al., 1994; Jewsbury and Kitagawa, 1994; Jewsbury et al., 1994), which would result in different electric fields at the bound CO. Due to the low occupancy of $A_3$ and the presumably small differences between the $A_1$ and $A_3$ conformations, well-modeled atomic-resolution data are a prerequisite for making structural distinctions between $A_1$ and $A_3$.

## What are the origins of the non-photolyzable fraction of $O_2$?

Carbon monoxide binding in Mb is easier to study than $O_2$ binding for two reasons. First, the infrared C–O stretch band is a convenient spectroscopic marker. Second, the apparent quantum yield for photolysis of $MbO_2$ on the time scale of 10 ps or longer at low temperatures is in the range 30–70% at neutral pH (Austin et al., 1975; Chance et al., 1990;

Miller et al., 1996) as opposed to 100% for CO. Investigations of the weak IR O–O stretch of $MbO_2$ show two distinct bands, one is photolyzable at low temperatures (as measured on the time scale of microseconds) and the other is not (Potter et al., 1987; Miller and Chance, 1995). It is thought that the low apparent quantum yield of $MbO_2$ is due to extremely rapid barrierless rebinding on the subpicosecond time scale of a distinct population in the sample (Miller et al., 1996). These populations do not exchange at cryogenic temperatures. Although there are differences in the mechanism of thermal dissociation and photodissociation, identification of the barrierless substates would have implications for understanding the relatively high affinity of the myoglobin heme for $O_2$. Are there structural differences between the photolyzable and non-photolyzable fractions?

## What are the dynamics of the molecule?

Examination of a space-filling model of the structure of Mb quickly demonstrates that there is no open channel for ligands such as $O_2$ and CO to enter and exit the heme pocket. Fluctuations in the conformation of the protein must take place in order for ligand binding and escape to occur. These transient openings and closings are too fast and have too low a population to appear in the NMR structure of the molecule (Ösapay et al., 1994). However, some indication of the dynamics is retained in the conformational disorder of myoglobin, even at low temperature. For data of moderate resolution, this disorder has typically been modeled as an isotropic Debye-Waller (temperature) factor for each non-hydrogen atom (Frauenfelder, 1989). High-resolution data permit a more sophisticated analysis of conformational flexibility, including refinement of anisotropic Debye-Waller factors. These can give insight into the character of concerted thermal motions of the protein at equilibrium.

## Technical improvements

Over the past 30 years, many structures of Mb have been determined by different laboratories, often setting the standards for their time. Examination of the differences among these structures led us to conclude that the true differences between structures of different ligation states were being swamped by variations in the way the data were collected and handled. This outlook has shaped our approach, which is to re-determine the structures of Mb in four different ligation states as a single set. In addition to collecting and analyzing the data in as much the same way as possible, we have also sought improvements in the quality of each structure through implementing recent advances in crystallographic practice. The most important of these improvements are as follows.

### *Rapid ligation*

Diffusion of small molecules into crystals often requires many minutes, and binding rates for CO and $O_2$ can be dramatically lower in crystals. In oxygenated crystals, autooxidation of the heme iron to form *aquomet*-Mb can occur on a time scale of hours to days. Myoglobin crystals of appreciable size are optically thick, so it is difficult to determine the ligation state without first dissolving the crystal. We used $O_2$ and CO gas at 50–100 bars of pressure to rapidly and completely ligate the crystals and minimize the opportunity for autooxidation.

### *Cryocrystallography*

Maintaining the crystal at cryogenic temperatures permits collection of a complete data set from a single crystal with minimal opportunity for chemical and physical changes (e.g., autooxidation and radiation damage) during data collection (Garman and Schneider, 1997). This is particularly important in a ligation study: crystal-to-crystal variation in ligation across a data set could lead to pronounced differences in a particular region of reciprocal space corresponding to data collected from a bad crystal. In addition, the electron density corresponding to water molecules is more localized at cryogenic temperatures due to the absence of liquid-like motions.

### *Better instrumentation*

Synchrotron sources and image-plate detectors offer great advantages over rotating-anode sources and previous generations of detectors in collecting high-resolution, high-precision data sets. The high collimation and brilliance of synchrotron radiation, coupled with the low noise, high dynamic range, and large working area of modern detectors enable data to be collected with a higher signal-to-noise ratio than was previously possible. Better data that extend to higher resolution enable the experimenter to reduce the weight given to prior constraints on refinement (such as the heme geometry). In addition, lower noise levels allow one to detect minority conformations at lower levels of occupancy in the electron density. Higher resolution provides better localization of individual atoms, more accurate delineation of the effects of partial occupancy and conformational disorder, and improved analysis of effects related to dynamics of the molecule such as anisotropic refinement of Debye-Waller factors.

### *Statistical methods*

In recent years, advanced statistical techniques for data reduction, modeling, and refinement of macromolecular diffraction data have been developed. These include better algorithms for integration and scaling of diffraction data, appreciation of the $R_{free}$ value as an unbiased indicator of model quality (Brünger, 1992), and Bayesian methods for weighting macromolecular data in refinement (Terwilliger and Berendzen, 1996a, b). These techniques can improve both the absolute quality of a structure given a set of data

and also improve estimates of differences among closely related structures.

By using the strategies described above, we re-determined the structures of unligated (deoxy) Mb, and of complexes with water (*aquomet*-Mb), carbon monoxide (MbCO), and oxygen ($MbO_2$) to a higher precision than was previously possible. The presented data sets extend to better than 1.2 Å in all cases, which is a noticeable improvement over many of the existing Mb structures in the Protein Data Bank (Bernstein et al., 1977). We use the differences in structure seen in the different ligation states to address questions about the structure, function, and dynamics of myoglobin.

## MATERIALS AND METHODS

### Crystal preparation and data collection

*aquomet*-Mb crystals were grown at room temperature using the batch method. Solid ammonium sulfate (AS) was added to a solution of 50 mg/ml sperm whale myoglobin in 50 mM potassium phosphate ($KP_i$) buffer pH 7.0 until the protein started to precipitate. Then, water was added until the solution started to clarify. Monoclinic crystals formed within a week.

To obtain ferrous unligated myoglobin crystals, *aquomet*-Mb crystals were reduced by soaking in a nitrogenated solution containing 50 mM sodium dithionite, 70 mM $KP_i$ (pH 7.0), and 70% saturated AS. A marked color change was observed, indicating that reduction had taken place. To obtain MbCO crystals, crystals of unligated Mb were soaked in a solution pressurized with CO to ~50 bar. In the case of $MbO_2$ crystals, reduced unligated crystals were rinsed in a solution of 50 mM $KP_i$ at pH 7.0, 70% saturated AS, 10% glucose (w/v), and 10% sucrose (w/v), transferred into a pressure chamber, and exposed to 100 bar of $O_2$ for 30 min at 4°C. For both gas ligations the pressure was released over a period of ~30 s and the crystals were flash-cooled in liquid nitrogen within 1 min; it is possible that due to expansion cooling the crystals froze at a pressure higher than atmospheric. Data were collected at beam line X12C at the National Synchrotron Light Source using a MAP image plate detector and processed with the HKL suite of programs (Otwinoski and Minor, 1996).

### Refinement

The $MbO_2$ data extend to the highest resolution of the four data sets (1.0 Å) and provide the most complete and redundant data (Table 2). Refinement of this data set was used as a template for all other complexes and formed the basis for difference refinement. The 1.5-Å-resolution structure of sperm whale MbCO (Kuriyan et al., 1986) was used as a starting model for the refinement with the program XPLOR (Brünger et al., 1986), with the CO and solvent molecules omitted. Several steps of simulated slow-cool annealing (Brünger et al., 1990) were performed, followed by model rebuilding using the graphics program O (Jones et al., 1991). The annealed slow-cool protocol ($T = 4000$ K) was used to calculate unbiased omit maps of the heme pocket environment and other parts of the protein. We used the results of a Cambridge Structural Database (CSD) search of porphyrin fragments (Frazao et al., 1995) to modify the param19x.heme parameters for XPLOR in a way analogous to Engh and Huber (1991). Additionally, the $Fe–N_{porphyrin}$ distance restraint was suppressed. After inclusion of 172 solvent molecules and modeling of the thermal motion by individual isotropic Debye-Waller factors, the *R*-factor was 19.9% and $R_{free}$ was 23.8% for data between 10.0 and 1.5 Å. This model was further refined using conjugate gradient minimization in SHELXL (Sheldrick and Schneider, 1997). For this stage of refinement, the working set of measured intensities and their estimated standard deviations were used in the minimization formula with the standard SHELXL weighting scheme. $R_{free}$ was calculated from the reference reflection set after each round of SHELXL refinement.

**TABLE 2 Data collection statistics**

| | $MbO_2$ | MbCO | *aquomet*-Mb | Unligated Mb |
|---|---|---|---|---|
| Cell dimensions | | | | |
| a, Å | 63.80 | 63.80 | 63.90 | 63.76 |
| b, Å | 30.81 | 30.63 | 30.73 | 30.66 |
| c, Å | 34.35 | 34.42 | 34.36 | 34.31 |
| $\beta$, ° | 105.8 | 105.8 | 105.7 | 105.7 |
| Resolution, Å | 1.0 | 1.15 | 1.1 | 1.15 |
| Observations | 530931 | 413744 | 302654 | 378009 |
| Unique reflections | | | | |
| all refl. | 67676 | 42860 | 51794 | 45065 |
| $I > 2\ \sigma$ | 59682 | 32272 | 46823 | 40009 |
| Completeness, all reflections/$I > 2\ \sigma$ (%) | | | | |
| ∞–6.0 Å | 93/91 | 92/91 | 90/86 | 80/80 |
| 6.0–2.3 Å | 99/98 | 95/95 | 95/95 | 91/91 |
| 2.3–1.8 Å | 98/97 | 98/97 | 99/99 | 98/97 |
| 1.8–1.5 Å | 98/95 | 94/88 | 99/97 | 99/95 |
| 1.5–1.3 Å | 96/91 | 87/73 | 99/93 | 99/88 |
| 1.3–1.15 Å | 95/85 | 87/60 | 98/81 | 98/76 |
| 1.15–1.1 Å | 93/79 | | 97/71 | |
| 1.1–1.0 Å | 88/62 | | | |
| $R^*_{sym}$ (%) | 5.7 | 5.9 | 4.6 | 5.4 |
| Temperature, K | ≈100 | ≈100 | ≈90 | ≈100 |
| Crystal pH | 7.0 | 6.0 | 7.0 | 7.0 |

*$R_{sym} \equiv \Sigma|I_{hi} - \langle I\rangle_h|/\Sigma\ \langle I\rangle_h$; $I_{hi}$ is the scaled intensity of the *i*th symmetry-related observation of reflection *h* and $\langle I\rangle_h$ is the mean value.

The anti-bumping restraints were only applied during the first few steps of SHELXL refinement and then released. The heme group was parametrized using the mean values of the CSD analysis mentioned above (Frazao et al., 1995). The $O_2$ ligand was first restrained to the EXAFS values (Powers et al., 1984). The restraints on the heme and ligand were gradually removed, and the final stages of refinement restrained only the bond length and bond angles of the two propionic acids. Attempts to remove more restraints in the structure resulted in an unacceptably large number of significant deviations from accepted values. Diffuse solvent modeling using Babinet's principle was applied. In parallel with further solvent and protein side-chain remodeling, the resolution was increased to 1.0 Å. With the disorder still modeled isotropically, the *R*-factor dropped to 19.2% and $R_{free}$ to 22.3%.

At this stage, the $O_2$ ligand was modeled into a peanut-shaped density next to the iron. An anisotropic model for Debye-Waller factors was gradually accepted for ordered solvent and the protein, except for the heme, $O_2$ ligand, proximal, and distal histidine atoms. The difference $F_o - F_c$ electron density map revealed a very strong positive ring around the iron atom in the heme plane, clearly suggesting an anisotropy of the iron thermal motion. All atoms of the heme, $O_2$ ligand, and proximal histidine were visible as single peaks and were therefore modeled anisotropically at this stage.

The distal histidine is not well-ordered, exhibiting high elongation of the electron density along the plane of the imidazole ring. The density is not consistent with a 180° rotamer state about $\chi_2$ (Oldfield et al., 1991; Jewsbury and Kitagawa, 1994), which would result in $N^{\delta}$ pointing to the inside of the pocket. However, it can be modeled either by a single distal histidine side chain with highly anisotropic Debye-Waller factors (see Fig. 1) or, alternatively, by two conformations of the side chain with isotropic Debye-Waller factors. Both cases result in identical $R_{free}$ values.

Double conformations were observed for the side chains of 23 residues and 7 water molecules. The occupancies of all partial conformations were refined, as well as the occupancy of the $O_2$ ligand. These occupancies are Glu-4 (57/43), Gln-8 (57/43), Gln-26 (65/35), Arg-31 (58/42), Glu-41 (51/49), Lys-47 (68/32), Lys-56 (53/47), Leu-61 (66/34), Lys-63 (59/41), His-64 (50/50), Gln-91 (53/47), Lys-96 (55/45), Glu-109 (66/34), Leu-115 (52/48), His-116 (54/46), Arg-118 (52/48), Asp-126 (66/34), Gln-128 (56/44), Glu-136 (57/43), Phe-138 (51/49), Ile-142 (67/33), Tyr-146 (51/49), Tyr-151 (50/50), and $O_2$ ligand (100). One round of occupancy

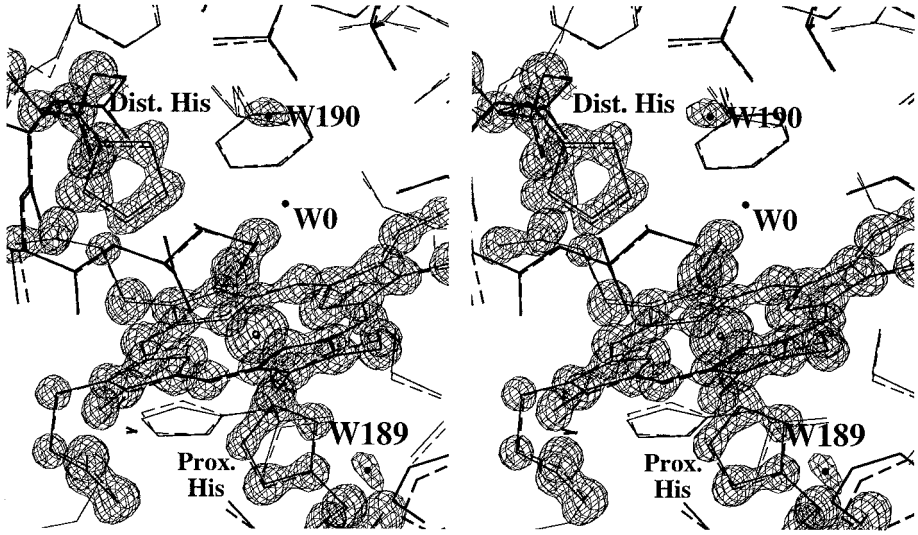


FIGURE 1 Stereo view of the final $2F_o - F_c$ electron density map of the $MbO_2$ complex (2.5 $\sigma$ level) with the single-histidine model (*solid line*) and a previously determined structure (1MBO; Phillips, 1980; *dashed line*) superimposed. For clarity, the density has been drawn only around the heme, proximal and distal histidine, and the newly observed density peaks near pocket modeled as W189 and W190. The position of a previously reported water molecule in the pocket, labeled W0, has no density in our structure and is likely due to a population of unligated Mb in the earlier structure.

refinement for solvent molecules was performed, then the occupancies were fixed at the resulting values until the end. Protons cannot be visualized from our data. The final refinement statistics are listed in Table 3.

For the remaining three complexes (*aquomet*-Mb, unligated Mb, and MbCO) the solvent network and side-chain conformations were altered from $MbO_2$ only where the density clearly indicated differences. Differences appeared at the ligand binding site and in the hydration network on the solvent side of the distal histidine. In the case of MbCO there were also differences reflected in alternative conformations of Arg-45, Phe-46, and the distal histidine. Otherwise, the model remained qualitatively equivalent for all four complexes. Fig. 2 shows electron density maps (both final and omit) and models of the active site of myoglobin in the four ligation states.

**TABLE 3 Final refinement statistics**

| | $MbO_2$ | MbCO | *aquomet*-Mb | Unligated Mb |
|---|---|---|---|---|
| Final refinement with experimental weighting | | | | |
| Resolution (Å) | 8.0–1.0 | 8.0–1.15 | 8.0–1.1 | 8.0–1.15 |
| $R$* (%) | 11.9 | 12.4 | 12.8 | 11.9 |
| (%) | 15.9 | 16.9 | 16.5 | 15.7 |
| r.m.s. bond distance dev. (Å) | 0.019 | 0.017 | 0.017 | 0.017 |
| r.m.s. angle dev. (°) | 0.036 | 0.036 | 0.035 | 0.034 |
| $\Delta r^{\#}$ (Å) | 0.05 | 0.07 | 0.07 | 0.06 |
| r.m.s. of $F_o - F_c$ map ($e$/Å$^3$) | 0.07 | 0.06 | 0.07 | 0.06 |
| Bayesian-weighted refinement | | | | |
| $R$* (%) | 12.2 | 12.6 | 13.0 | 12.1 |
| (%) | 15.8 | 16.8 | 16.4 | 15.3 |
| r.m.s. bond distance dev. (Å) | 0.016 | 0.015 | 0.016 | 0.015 |
| r.m.s. angle dev. (°) | 0.034 | 0.033 | 0.034 | 0.033 |
| $\Delta r^{\#}$ (Å) | 0.05 | 0.05 | 0.06 | 0.05 |
| r.m.s. of $F_o - F_c$ map ($e$/Å$^3$) | 0.05 | 0.04 | 0.04 | 0.04 |
| Bayesian-weighted difference refinement | | | | |
| $R$* (%) | — | 12.7 | 13.1 | 12.1 |
| $R_{Bayes}$[§] (%) | — | 8.9 | 8.7 | 8.2 |
| r.m.s. bond distance dev. (Å) | — | 0.015 | 0.017 | 0.016 |
| r.m.s. angle dev. (°) | — | 0.033 | 0.034 | 0.033 |
| $\Delta r^{\#}$ (Å) | — | 0.05 | 0.05 | 0.04 |
| r.m.s. of $F_o - F_c$ map ($e$/Å$^3$) | — | 0.03 | 0.03 | 0.03 |
| mean B-factor (Å$^2$) | | | | |
| solvent | 21.8 | 27.5 | 22.6 | 24.6 |
| all protein | 11.0 | 16.2 | 13.4 | 13.3 |
| main chain | 9.4 | 13.1 | 10.6 | 11.8 |
| heme | 9.1 | 13.4 | 9.8 | 12.3 |
| ligand | 11.3 | 13.0 | 9.7 | 21.2 |
| Fe | 6.9 | 11.4 | 7.8 | 9.9 |

*$R \equiv \Sigma|F_{o_h} - F_{c_h}|/\Sigma_h F_{o_h}$, where $F_{o_h}$ and $F_{c_h}$ are the observed and calculated structure factor amplitudes for reflection $h$. We used all reflections and a low-resolution limit of 8 Å for this calculation.

#$\Delta r$ is an average radial error of atomic position as estimated by a Luzatti plot (Luzatti, 1952).

§$R_{Bayes}$ was calculated as $R$, except the Bayesian-weighted difference structure factors were used. It is not directly comparable with $R$, since the correlated residuals were subtracted.

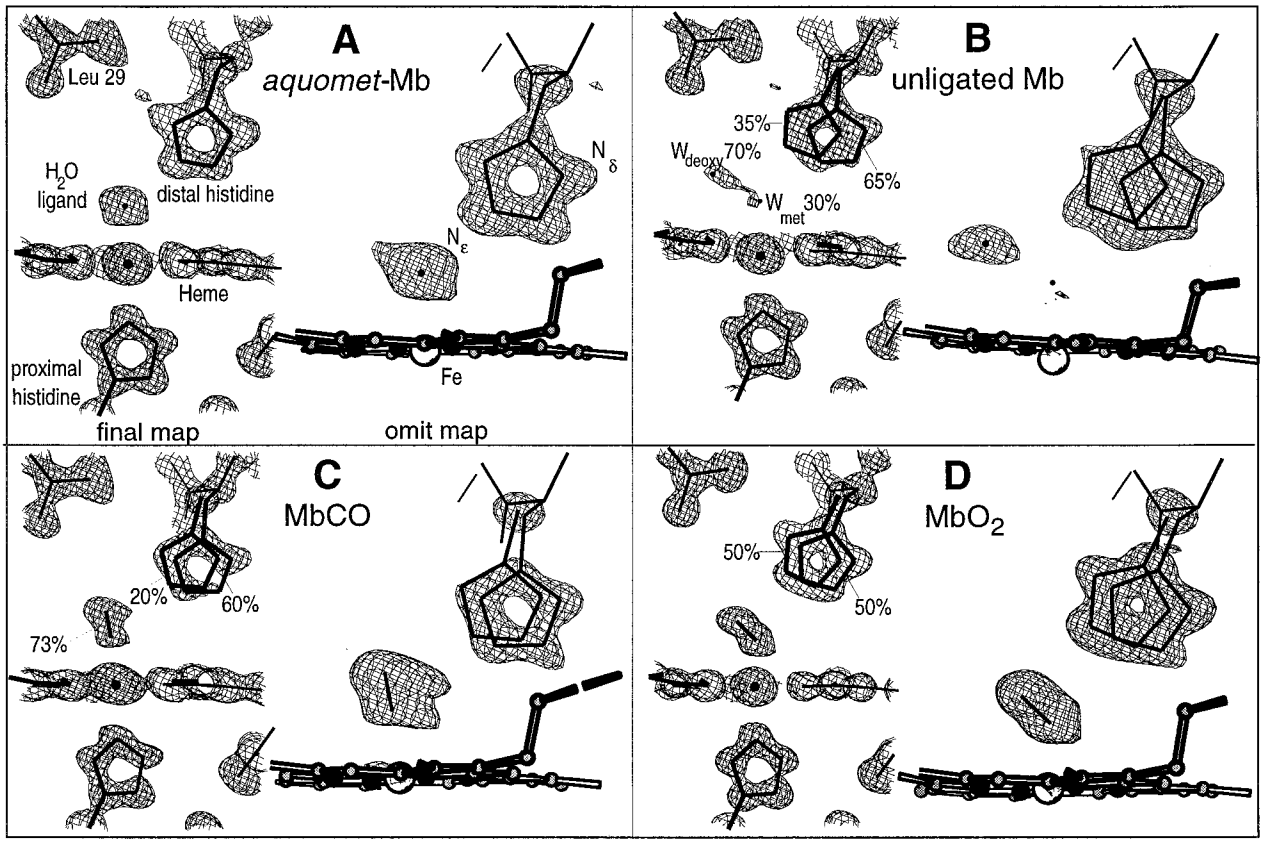


FIGURE 2 The final $2F_o - F_c$ electron density maps in the ligand binding area at 1.8 $\sigma$ level (*left*) and the simulated annealing omit $F_o - F_c$ difference electron density maps (*right*) for distal histidine and the ligand at the 3 $\sigma$ level for (*A*) *aquomet*-Mb, (*B*) unligated Mb, (*C*) MbCO, and (*D*) $MbO_2$. Simulated annealing to 3000 K was performed with the distal helix and central heme area (including a 3-Å surrounding) omitted from the model. Occupancies of the model that differ from unity are indicated in percent. The small density peak at the ligand site of the "unligated" Mb structure corresponds to a 30% *aquomet*-Mb contamination, as noted in the text.

The difference $F_o - F_c$ map of unligated myoglobin revealed a residual positive peak at the site of the *aquomet*-Mb ligand water (see Fig. 2 *B*). An alternative conformation of the distal histidine with a geometry identical to that in *aquomet*-Mb was needed to fit the density. This indicates a partial oxidation of the heme to the ferric $Fe^{III}$ state. This effect was treated by adding a fixed *aquomet*-Mb structure to the unligated model. This approach resulted in a lower $R_{free}$ by 0.7% as compared to the model with no *aquomet*-Mb part and 0.3% compared to the case where only the distal histidine and ligand water of *aquomet*-Mb were modeled. Several tests varying the occupancy ratio between the two parts of the structure resulted in an estimate of 20–40% *aquomet*-Mb contamination. Finally, the 70% unligated myoglobin and 30% *aquomet*-Mb model was accepted, as it gave the most favorable fit of the density peak at the ligand binding site. The *aquomet*-Mb contamination of the unligated Mb structure contributes significantly to the larger uncertainty estimates for this state.

At the last stage of the independent refinements, Bayesian weighting was implemented. Bayesian weights based on the mean-square residual errors were calculated on structure factors by the program HEAVY (Terwilliger and Berendzen, 1996a). These were transformed back to intensities via first-order expansion ($\sigma(F^2) = 2F \cdot \sigma(F)$) for the final step of SHELXL refinement; negative intensities were discarded from the refinement. $R$ and $R_{free}$ were calculated from the original work and reference data sets, respectively. The subsequent SHELXL refinement (Sheldrick and Schneider, 1997) resulted in a slightly better $R_{free}$ with lower bond and angle distance root-mean-square (r.m.s.) deviations from the target values, as listed in Table 3.

It has been demonstrated that independent refinement of structures with highly correlated errors in the atomic models (such as would be caused by failure to include, e.g., a few solvent molecules) can lead to exaggeration of the differences between them. We used a refinement strategy that makes better estimates of the *differences* between pairs of structures except where the data demand differences. To this end we used Bayesian difference refinement, in which an estimate of the correlated error between a "reference structure" and a "variant structure" is subtracted from the data before refinement and in which information about the residuals is used in determining the weighting (Terwilliger and Berendzen, 1996b). The $MbO_2$ data offered the best resolution and redundancy, so we selected this as our reference structure. Bayesian difference refinement of the remaining structures was carried out with weights and data calculated in HEAVY as mentioned above. The correlation coefficient between reference ($MbO_2$) and variant structure model errors were 0.72, 0.73, and 0.76, for unligated, MbCO, and *aquomet*-Mb, respectively. Bayesian difference refinement produced similar r.m.s. deviations from ideality and markedly lower r.m.s. differences among structures, and these were accepted as final. In order to obtain the estimated standard deviations of all refined parameters, a cycle of full-matrix least-square refinement with no shift of refined parameters was applied to the final models.

The measured intensities, final coordinates, and anisotropic Debye-Waller factors have been deposited at the Protein Data Bank. The entry names are 1A6G, 1A6M, 1A6N, and 1A6K for MbCO, $MbO_2$, unligated Mb, and *aquomet*-Mb, respectively.

## RESULTS

### Reference structure: $MbO_2$

The crystallographic literature on $MbO_2$ is not as extensive as that for MbCO. Of the five $MbO_2$ structures deposited at

the Protein Data Bank (Bernstein et al., 1977), only one corresponds to native sperm whale with iron in the heme center. The other structures are mutant proteins or have a cobalt ion in the heme center and represent a broad scale of experimental conditions, refinement techniques, temperatures, and crystal environment (Table 4). The most direct comparison that can be made is to the native $MbO_2$ x-ray crystal structure determined to 1.6 Å resolution by S.E.V. Phillips (1980). Overall, the two structures are very similar, with an r.m.s. deviation of 0.22 Å for main-chain and 0.51 Å for all protein atoms. The proximal histidine is nearly at the same position in both structures. When the distal histidine is modeled by a single conformation (with highly elongated thermal ellipsoids) it converges to the same conformation for both structures (see Fig. 2 *D*) with only 0.14 Å r.m.s. deviation. The $O_2$ ligand makes an angle of 58° with the mean heme plane in our structure. This is consistent with the previous determinations that show angles in the range 58–69°.

However, there are important local differences. The structures differ significantly in the iron position, namely a 0.19-Å out-of-plane deviation in the older structure versus 0.089 Å in our structure. This discrepancy is much larger than would be expected from the uncertainties in the individual iron positions. We find no water molecules in the ligand-binding pocket in $MbO_2$, unlike the previous determinations in wild-type $MbO_2$ (Phillips, 1980) and in the Asp-122–Asn mutant $MbO_2$ (Quillin et al., 1993) (see W0 in Fig. 1). Significantly, these are also the only two of the six $MbO_2$ structure determinations in the literature that find the iron atom out of the heme plane, suggesting an incomplete $O_2$ occupancy and partial occupancy of unligated Mb in those two structures.

A distinctive difference from previous results is also visible in the electron density near the heme pocket. There are two very well-ordered regions of density, one adjacent to the heme on the proximal side and one on the distal side near Leu-29, that can be modeled as water molecules (labeled W189 and W190, respectively, in Fig. 1). Even though the side chains close to the new density peaks have very similar conformations in the other structures listed in Table 4, neither of the two peaks has been observed before, except in xenon binding studies where these are the two most highly occupied binding sites (Tilton et al., 1984; Sauer et al., 1997). Because the crystals of the $MbO_2$ complex have been prepared at high $O_2$ concentration (100 bar), we believe it is likely that these peaks are not due to water, but rather to molecular oxygen that was trapped inside the protein by freezing after the pressure was released quickly. The electron densities at these sites are consistent with oxygen molecules with some amount of rotational disorder.

## Overall differences among ligation states

The influence of refinement strategy on the differences between various Mb ligation complexes can be seen in the example of $MbO_2$ and MbCO listed in Table 5. While two previous independent structure determinations of the MbCO and $MbO_2$ complexes find 0.57 Å r.m.s. deviation for all protein atoms, we find a difference of 0.28 Å after anisotropic SHELXL refinement and only 0.17 Å after Bayesian weighted difference refinement. We observe an excellent correlation between data statistics and r.m.s. differences in structure (Table 6). The smallest $R_{merge}$ (an estimate of the differences between two data sets, including errors) between two data sets is 9.6%, which is only slightly larger than the 8.0% r.m.s. $R_{sym}$ value (an estimate of the errors).

Consistent treatment of both samples and data and application of the advanced statistical methods mentioned in the Methods section has yielded a set of complexes with exceptionally low r.m.s. differences between the final models: the largest difference is 0.19 and 0.27 Å for main-chain and all side atoms, respectively, for MbCO versus unligated myoglobin. The highest r.m.s. deviations in our study are those of unligated myoglobin from all the other ligation states. The lowest r.m.s. deviations were observed between *aquomet*-Mb and $MbO_2$ states. A similar comparison of moderate- or high-resolution structures available in the Protein Data Bank, from different crystals, experiments, x-ray sources, refinement techniques, etc., resulted in r.m.s. deviations of 0.2–0.3 Å for main-chain and 0.5–0.7 Å for all protein atoms. This suggests that the independently derived

**TABLE 4 Some $MbO_2$ structures determined by diffraction methods**

| PDB | Year | Protein | pH | Sp.group | Source | Res., Å | Refinement | *R*, % | Temp., K |
|---|---|---|---|---|---|---|---|---|---|
| 1MBO | 1980 | native $MbO_2$ | 8.4 | $P2_1$ | x-ray | 1.6 | LS (Jack-Levitt) | 15.9 | 261 |
| XXXX* | 1981 | native $MbO_2$ | 8.4 | $P2_1$ | neutron | 1.5 | LS (Jack-Levitt) | 18.8 | 268 |
| 2SPN | 1992 | mutant L29F, D122N | n.a. | P6 | x-ray | 1.7 | LS (PROFFT) | 16.6 | n.a. |
| 2MGM | 1993 | mutant D122N | n.a. | P6 | x-ray | 1.9 | mol. dyn. (X-PLOR) | 15.0 | n.a. |
| 1LTW | 1996 | mutant L29W, D122N | 9.0 | P6 | x-ray | 1.7 | mol. dyn. (X-PLOR) | 15.8 | 295 |
| 1YOI | 1996 | cobalt $MbO_2$ | n.a. | $P2_1$ | x-ray | 1.7 | mol. dyn. (X-PLOR) | 16.2 | 295 |
| This work | 1998 | native $MbO_2$ | 7.0 | $P2_1$ | x-ray (SR)# | 1.0 | LS (SHELXL) | 12.8 | 100 |

1MBO (Phillips, 1980); XXXX (Phillips and Schoenborn, 1991); 2SPN (Carver et al., 1992); 2MGM (Quillin et al., 1993); 1LTW (Carver et al., 1992); 1YOI (E. A. Brucker, J. S. Olson, G. N. Phillips, Jr., Y. Dou and M. Ikeda-Saito, High-resolution crystal structures of the deoxy-, oxy-, and aquomet-forms of cobalt myoglobin, to be published).

*The data for this structure have not been deposited in the Brookhaven Protein Data Bank.

#SR: Synchrotron Radiation.

**TABLE 5** ***R*, $R_{free}$, and r.m.s. differences between MbCO and $MbO_2$ in the course of the refinement**

| Step | *R*, % | | $R_{free}$, % | | r.m.s. differences, Å | |
|---|---|---|---|---|---|---|
| | $MbO_2$ | MbCO | $MbO_2$ | MbCO | main | all |
| 1MBO × 1MBC* this refinement | 15.9 | 17.1 | | | 0.21 | 0.57 |
| X-PLOR[#] | 19.9 | 21.7 | 23.8 | 25.4 | 0.14 | 0.44 |
| SHEL isotr. | 18.9 | 21.3 | 22.2 | 24.6 | 0.15 | 0.22[§] |
| SHEL anis. | 11.9 | 12.4 | 15.9 | 16.9 | 0.12 | 0.28 |
| SHEL baywght. | 12.2 | 12.6 | 15.8 | 16.8 | 0.12 | 0.28 |
| SHEL baydiff. | 12.2[¶] | 12.7[¶] | 15.8 | | 0.12 | 0.17 |

*A comparison of existing structures from the PDB, 1MBO for $MbO_2$ (Phillips, 1980), and 1MBC for MbCO (Kuriyan et al., 1986). $R_{free}$ was not monitored for these structures.

[#]Only 105 waters were included in the MbCO model during this round, versus 180 for the others.

[§]Several disordered residues were changed to Ala in this round of refinement, making comparison of all atoms difficult.

[¶]$MbO_2$ was the reference structure, so the refinement did not change and these values are the same as for the Bayesian-weighted refinement.

structures in the PDB may have overestimated the size of structural changes accompanying ligation for atoms far from the active site by more than a factor of two. Comparison of the four structures shows that there are few significant differences outside the region of the ligand binding pocket. Significant differences are discussed in detail below.

## Ligand binding area and ligand geometry

To obtain an unbiased view of the ligand binding area we calculated simulated annealed omit maps with XPLOR (Brünger et al., 1990), a procedure common for medium- to high-resolution data but not standard for atomic resolution data. Features near the ligand obtained when omitting all atoms of the proximal and distal histidines, the ligand, within a shell of 3 Å around the ligand, and the porphyrin nitrogens, are shown in Fig. 2. The density of the ferric *aquomet*-Mb structure is the easiest to interpret. There are no multiple conformations in the ligand binding area (Fig. 2 *A*). In contrast, the ferrous complexes exhibit more complicated electron densities and require a detailed description of possible interpretations.

Particular care was devoted to modeling the ligand density for all ligation states. There is no doubt about a single binding geometry of the $O_2$ molecule, because it appears clearly as two separate peaks of electron density (see Fig. 2 *D*). The ligand occupancy was refined freely at the last stages and stayed at 100%. The $O_2$ ligand is bent significantly from the heme normal. The deviation is formed predominantly by the Fe–O–O bond angle and the tilt of Fe–O from the heme normal (see Table 7). The consistency of the $O_2$ ligand conformation in our structure and all previous crystallographic determinations is remarkable in view of the wide range of results on the ligand geometry in MbCO.

The CO ligand also shows peaks for individual atoms. However, their occupancy converged to 73% and noticeable extensions of the density are visible (see Fig. 2 *C*). The iron is in the plane of the heme, which would indicate it is unlikely that the missing 27% of the CO density could be explained by the presence of unligated Mb. However, the data are consistent with the presence of a weakly occupied CO conformation not included in our model. More hints as to the nature of this conformation can be seen in the residual peaks that show up at low $\sigma$ level in an $F_o - F_c$ difference electron density map on the final anisotropically refined model. The geometry of these residual peaks would suggest that the "missing" minority conformation is bent at a very similar angle to the majority, but displaced slightly in the general direction of the distal histidine. Consistency of Fe–C distances would require shifting the heme to model this conformation, similar to what is seen in the pH 4 structures (Yang and Phillips, 1996). However, since attempts to include a second conformation did not produce stable refinements, it was not included in our final model.

We observe a deviation of only 18 ± 3° from the normal of the heme plane for the majority conformation of the CO ligand. This contrasts with previous structures of MbCO using the monoclinic crystal form, which find angles of 30–60° (Kuriyan et al., 1986; Cheng and Schoenborn, 1991; Yang and Phillips, 1996). The CO deviation from the heme normal in our structure is formed equally by a 9° tilt of the Fe–C bond and a 9° bend of Fe–C–O. In contrast, $MbO_2$ has a tilt angle of only 0.3°.

## Proximal and distal histidines

The geometry of the proximal histidine with respect to the heme is, to within experimental uncertainty, the same in the $MbO_2$ and MbCO structures. In the unligated structure, the Fe–$N^{\epsilon}$ bond distance is larger and there is 2.4° more tilt off axis, but the azimuthal angle is not significantly different. Azimuthal differences in the proximal histidine between previous ferrous Mb structures seem to have been exaggerated by independent refinement by different methods.

The variability of the distal histidine position within the set of four ligation complexes is significantly higher than the mean r.m.s. deviation of the whole protein. Higher resolution than previous structures and clear densities associated with hydration allows us to rule out 180° rotamer states about $\chi_2$ of the distal histidine (Oldfield et al., 1991) from being occupied to any significant degree. The multiple conformations of the distal histidine that we see (with the exception of the swung-out conformer in MbCO associated with a doubly protonated imidazolate) are all the rotamer with $N^{\epsilon}$ on the inside of the protein (in the binding pocket) and $N^{\delta}$ pointing out toward the solvent and communicating with a complex hydration network consisting of Arg-45, Thr-67, and individual solvent molecules. There are signif-

**TABLE 6 Data and model differences upon ligation**

| | Data *R*-factors,* % | | | | Model r.m.s. differences,# Å | | | |
|---|---|---|---|---|---|---|---|---|
| | *aquomet*-Mb | MbCO | $MbO_2$ | unligated Mb | *aquomet*-Mb | MbCO | $MbO_2$ | unligated Mb |
| *aquomet*-Mb | — | 13.2 | 9.6 | 11.7 | — | 0.13 | 0.09 | 0.15 |
| MbCO | 7.5 | — | 11.7 | 16.8 | 0.21 | — | 0.11 | 0.19 |
| $MbO_2$ | 7.3 | 8.2 | — | 14.0 | 0.16 | 0.17 | — | 0.15 |
| Unligated Mb | 7.1 | 8.0 | 7.8 | — | 0.25 | 0.27 | 0.23 | — |

#The r.m.s. deviations were calculated after superimposing the $C_\alpha$ atoms of a whole protein for the two particular ligation states. The values above the diagonal correspond to main-chain atoms, below the values diagonal to all-protein atoms.

*The values below the diagonal correspond to the geometric mean value of the $R_{sym}$ of two data sets, above the diagonal to their $R_{merge}$. The unligated Mb data set is thought to be 30% *aquomet*-Mb-contaminated.

icant differences in the hydration on the solvent side of the distal histidine near $N^\delta$.

*aquomet*-Mb is the only state that shows a distal histidine in a single well-ordered conformation. This conformation is stabilized by hydrogen bonds between the ligand water and $N^\epsilon$ on the pocket side and between $N^\delta$ and a fully occupied sulfate molecule located on the solvent side. Since sulfate at pH 7.0 can only serve as an electron donor to this hydrogen bond, this indicates that the distal histidine in *aquomet*-Mb is fully in the $HN^\delta$ tautomer. However, the sulfate ion shows a 60/40 mixture of two conformations that apparently represent two different possible connectivities of the hydration network.

In the ferrous ligation states the density of the distal histidine is significantly more disordered and the tautomer assignments are less clear. After correction for 30% *aquomet*-Mb contamination, unligated myoglobin shows two well-separated conformations of the distal histidine with equal occupancies, one at the same position as in *aquomet*-Mb and one protruding much further into the pocket. In the inward conformation, the distal histidine is hydrogen-bonded to a water molecule (the $N^\epsilon - H_2O$ dis-

**TABLE 7 Heme, ligand, and histidine geometries**

| | $MbO_2$ | MbCO | *aquomet*-Mb | unligated Mb |
|---|---|---|---|---|
| Ligand* | | | | |
| Occupancy (%) | 100 | 70 (10) | 100 | 70 (10) |
| IR ∠# (°) | 57 (1) | 18 (3) | | |
| bend ∠§ (°) | 122 (1) | 171 (3) | | |
| Tilt ∠¶ (°) | 0.3 | 9.0 | 6.8 | 30.6 |
| Fe—C (Å) | 1.81 (1) | 1.82 (2) | | |
| C—O (Å) | 1.24 (2) | 1.09 (2) | | |
| Fe—O (Å) | 2.68 (2) | 2.91 (2) | 2.13 (1) | 3.53 (5) |
| C—distal His $N^{\epsilon 2}$ (Å) | 3.07 (3)/3.02 (2) | 3.42 (3)/3.18 (8)/6.95 (7) | | |
| O—distal His $N^{\epsilon 2}$ (Å) | 2.96 (3)/2.67 (2) | 3.16 (4)/2.74 (8)/6.58 (7) | 2.67 (2) | 3.89 (5)/2.76 (4) |
| Heme | | | | |
| Fe—⟨$N_p$ plane⟩ (Å) | −0.024 (6) | −0.001 (9) | −0.106 (7) | −0.363 (11) |
| Fe—⟨heme plane⟩ (Å) | −0.089 (3) | −0.048 (5) | −0.138 (4) | −0.390 (6) |
| ⟨Fe—$N_p$⟩ (Å) | 2.01 (2) | 1.98 (2) | 2.03 (2) | 2.07 (3) |
| ⟨$N_p$—$N_p$⟩ (Å) | 2.84 (2) | 2.81 (3) | 2.87 (2) | 2.89 (3) |
| plane doming (°) | 1.0 | 1.8 | 3.1 | 3.0 |
| Proximal histidine | | | | |
| Fe—$N^{\epsilon 2}$ (Å) | 2.06 (1) | 2.06 (2) | 2.14 (1) | 2.14 (2) |
| Tilt ∠¶ (°) | 3.4 | 3.4 | 2.4 | 5.8 |
| dihedral $N_pA$–Fe–$N^{\epsilon 2}$–$C^{\epsilon 1}$ (°) | 1.9 (1.4) | 0.2 (1.7) | 8.6 (1.5) | 2.7 (2.0) |
| Distal histidine | | | | |
| $\chi_1$ (°) | −174 (2)/−178 (2) | −157 (2)/−166 (3)/−90 (4) | −170 (2) | −156 (3)/−178 (3) |
| $\chi_2$ (°) | 67 (3)/61 (3) | 61 (3)/58 (7)/72 (5) | 64 (2) | 68 (4)/64 (4) |
| $\chi_3$ (°) | −178 (2)/−175 (2) | −179 (2)/179 (2)/−179 (2) | −178 (1) | −179 (2)/−179 (2) |
| Occupancy (%) | 50/50 | 60/20/20 | 100 | 35/35‖ |

Multiple entries refer to alternative distal histidine conformations. Angle brackets indicate a mean value averaged over several atoms. The estimated standard deviation values are listed in parentheses.

*For the $MbO_2$ structures, the equivalent positions for an $O_2$ ligand are listed. For the *aquomet*-Mb and unligated Mb these entries refer to water molecules at the active site.

#The IR angle lies between the C—O bond (or O—O bond) and the normal to the mean heme plane.

§The bend angle is between the iron, the nearer ligand atom, and the farther ligand atom.

¶The tilt angle lies between the line Fe—ligand and the normal to the mean heme plane.

‖A 30% occupancy fixed distal histidine conformation has been subtracted. This conformation, corresponding to *aquomet*-Mb contamination, is identical to the first entry in the table.

tance is 2.8 Å) that sits near the "docking site" in the pocket where carbon monoxide has been shown to go after photolysis at low temperature (Schlichting et al., 1994).

The electron density of the distal histidine in MbCO is strongly elongated in the plane of the imidazole. It can be modeled either as one conformer with highly elongated thermal ellipsoids or as two histidine conformers. We exclude the single-conformer model for steric reasons, because the carbonyl oxygen atom would be too close to the histidine in the middle of the required range. Refinement of the two-conformer model yielded conformers with distances from the ligand oxygen atom to the distal histidine $N^{\epsilon}$ of 3.2 Å and 2.7 Å and occupancies of 60% and 20%, respectively. The distance for the majority conformation is slightly longer than would be expected if a hydrogen bond were present. The electron density of the hydrogen-bonding partner of $N^{\delta}$ appears to be more consistent with a water molecule than a sulfate as in *aquomet*-Mb. Neutron diffraction data on deuterated MbCO (collected from the same crystal form at room temperature and pD 5.7) also find a water molecule in this position with a fully occupied $HN^{\delta}$ tautomer (Cheng and Schoenborn, 1991). There is no inconsistency between the absence of the sulfate and the presence of the $HN^{\epsilon}$ tautomer, since there are significant differences in the distal histidine position between MbCO and *aquomet*-Mb (r.m.s. difference 0.52 Å) and around Arg-45 that may prevent binding of sulfate. We revisit the question of tautomer assignments for these conformations in the Discussion section.

In the last steps of the refinement of MbCO, a third conformation of distal histidine was observed in the difference electron density map (see Fig. 3) at a site nearly identical to the swung-out conformation observed in a diffraction study of MbCO at pH 4 (Yang and Phillips, 1996). The presence of alternative conformations for Phe-43, Arg-45, and disorder of the heme propionic acid and its hydration pattern further support the identity of this substate with the conformer observed at full occupancy at pH 4. This substate is doubly protonated, based on the $pK_a$ value of 4.6 of the distal histidine (Wilbur and Allerhand, 1977; Fuchsman and Appleby, 1979). In refinement, the occupancy of this conformation converged to 20%. The uncertainty in the distal histidine occupancies is ~10%.

The electron density of the distal histidine in $MbO_2$ is also strongly elongated in the plane of the imidazole. In contrast to MbCO, in $MbO_2$ the single-conformer model is not excluded by steric interference with the ligand. Refinement of a single-conformer model leads to extended thermal ellipsoids in the plane of the histidine side chain, and a 2.8 Å distance from $N^{\epsilon}$ to the terminal oxygen of $O_2$, as has been seen in neutron diffraction studies (Phillips and Schoenborn, 1981). However, we prefer the two-conformer model because spectroscopic results on $MbO_2$ show two distinct O–O stretch bands at neutral pH that reflect different electrostatic interactions with the surroundings (Tsubaki and Yu, 1981; Potter et al., 1987; Jeyarajah et al., 1994; Miller and Chance, 1995). Nothing else nearby in the structure (including the $O_2$ itself) shows multiple conformations that could be associated with these bands. Refinement of the two-conformer model in $MbO_2$ leads to equal occupancies for the two distal histidine conformations. The first conformer forms a 2.7-Å hydrogen bond from $N^{\epsilon}$ to the terminal oxygen of $O_2$. The second conformer is identical to the single conformer seen in *aquomet*-Mb and the distance to the terminal ligand atom is 3.0 Å, suggesting a weaker interaction with the $O_2$. The precise assignment of the density at the hydrogen-bonding partner of $HN^{\delta}$ is not clear. The electron density peak is slightly too strong and nonspherical to fit a single water molecule, and a pair of disordered water molecules does not fully match the den-

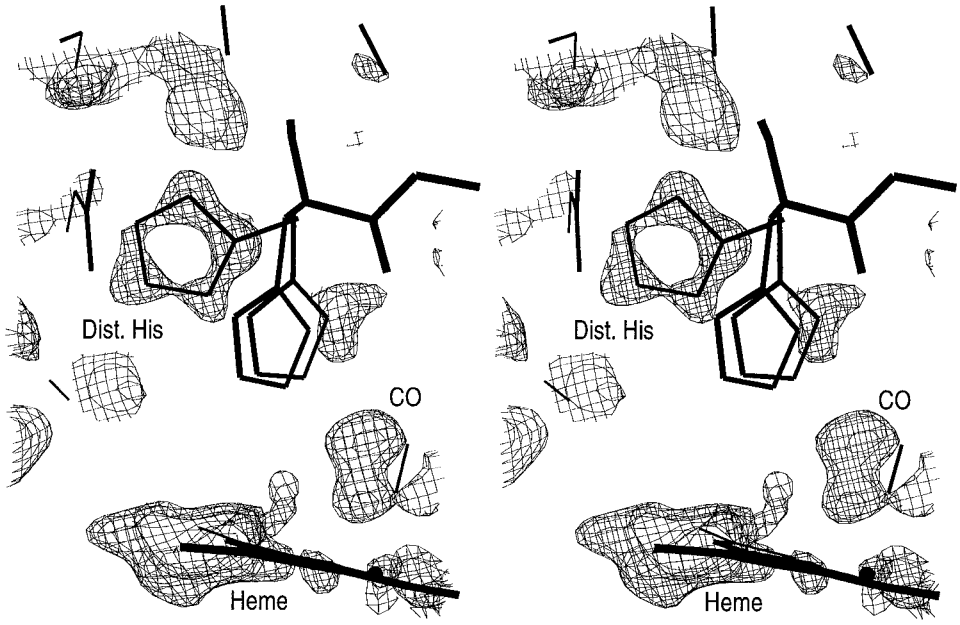


FIGURE 3 Stereo view of the $F_o - F_c$ electron density map in the ligand binding area of MbCO at 1.7 $\sigma$ level showing the "swung-out" conformer of the distal histidine.

sity. We believe that there may be a combination of a disordered sulfate and a water molecule, each corresponding to one partially occupied histidine side-chain orientation, although we did not include this model in our final refinement. A neutron-diffraction structure of deuterated $MbO_2$ (collected from 2 crystals at 268 K at pD 8.4) indicates that the proton is on $N^{\epsilon}$ (Phillips and Schoenborn, 1981) but the likely presence of a hydrogen-bonded sulfate in our data would indicate partial occupancy of the $HN^{\delta}$ tautomer.

## Heme position

When comparing the structures of the four complexes presented here it is noticeable that the heme itself shifts. To estimate these shifts upon ligand binding, we aligned the backbone atoms of the structures to each other, then estimated the shifts using the 28 central atoms of the heme (removing from consideration the iron, propionic acids, and vinyl groups).

The shifts are relatively small, although well outside the noise. The r.m.s. deviations between the heme atoms of the ligated states are $<0.15$ Å, whereas the deviations between ligated and unligated states range from 0.25 to 0.31 Å. The relative shifts of the heme with respect to unligated Mb in the internal molecular coordinate system are listed in Table 8. The largest component of the shift is approximately in the direction of the "D" pyrrole ring, which lies roughly perpendicular to the plane of proximal histidine and (more crudely) in the direction of the distal histidine. The component of shift in this direction is 0.23 and 0.21 Å for $MbO_2$ and MbCO from the unligated state, respectively. Shifts in the direction perpendicular to the heme plane range from 0.07 to 0.15 Å from the unligated state toward the proximal side.

## Hydration

Differences in the solvent network observed between various complexes in previous studies have been linked to different data processing and refinement strategies (Phillips and Pettit, 1995). To avoid this problem, we carefully analyzed the hydration of the reference $MbO_2$ complex first and then checked for differences in the other structures; 185 of 190 observed waters are common to all four structures. All observed changes are limited to the proximity of the heme and the proximal and distal histidines. There are significant differences in hydration on the solvent side of the distal histidine near $N^{\delta}$, as mentioned in the discussion of the distal histidine.

In MbCO, the solvent network near the heme propionic acid is disturbed in the conformation with the swung-out distal histidine. Two solvent molecules (W24, W149) have double conformations in MbCO that have not been seen in other neutral-pH structures, but are identical to those observed in the low-pH study (Yang and Phillips, 1996).

Our MbCO and $MbO_2$ complexes have no water molecules inside the heme pocket, with the possible exception of the W189 density at the proximal xenon binding sites in $MbO_2$ mentioned earlier. The *aquomet*-Mb complex and the unligated state do have water in the distal pocket that directly interacts with the distal histidine via a hydrogen bond (Takano, 1977a, b), but we do not see evidence for two water molecules in the pocket in either of the latter ligation states. Spectroscopic studies of the C–O stretch bands indicate there are likely to be small differences in the structure near the CO between physiological temperatures and frozen solution, although these seem primarily due to changes in the protonation of His-97 (Müller, 1997) and may be too small to be visualized in structures of less than atomic resolution. There may also be changes in the hydration shell upon crossing the water-ice transition.

**TABLE 8 Heme shift relative to unligated Mb**

| | Shift, Å | | | |
|---|---|---|---|---|
| | *x* | *y* | *z* | Residual, Å r.m.s. |
| *aquomet*-Mb | 0.23 | 0.02 | 0.13 | 0.063 |
| MbCO | 0.21 | 0.09 | 0.07 | 0.066 |
| $MbO_2$ | 0.13 | 0.03 | 0.15 | 0.066 |

The coordinate system used here has the origin at the iron, the *x*-axis points toward ND of the pyrrole ring, the *y*-axis approximately in the NC direction, and the *z*-axis out of the plane toward the proximal histidine.

## Comparison with EXAFS results

Results from EXAFS experiments are characterized by considerable precision, although it is known that EXAFS analyses that do not include the effects of anharmonicity can show systematic errors (Crozier et al., 1988). We compared the Fe-neighbor distances and uncertainties with those obtained by EXAFS. As mentioned in Methods, we refined the positions of the iron atom, central heme atoms, and ligands without geometric restraints. As one can see from the estimated standard deviations listed in Table 9, the precision of the atomic resolution x-ray data matches that of EXAFS studies.

All values for $MbO_2$ agree within one sigma. The EXAFS data give two alternatives for the calculated $O_2$ orientation, and the bent structure (Powers et al., 1984) fits our data. In the case of *aquomet*-Mb, the distances from the distal histidine and porphyrin ring nitrogens to the iron atom match the EXAFS data to within one sigma. However, in our data the water in the *aquomet* state is 0.25 Å further away from the iron, a discrepancy well outside experimental uncertainties.

Serious disagreements were also observed in the MbCO complex for the Fe–C distance (five estimated standard deviations) and proximal-histidine-to-Fe distance (four estimated standard deviations). It is interesting that our data deviate in the opposite direction from the neutron data (Cheng and Schoenborn, 1991). Those distances were significantly longer than the ones presented here, 2.12 and 2.26 Å for the ligand-to-Fe and proximal-histidine-to-Fe dis-

**TABLE 9 Comparison with distances from EXAFS**

| | $MbO_2$ Structure | | MbCO Structure | | *aquomet*-Mb Structure | | Unligated Mb | |
|---|---|---|---|---|---|---|---|---|
| | This work | EXAFS | This work | EXAFS | This work | EXAFS | This work | EXAFS |
| Fe—ligand (Å) | 1.81 (1) | 1.80 (2) | **1.82 (2)** | **1.93 (2)** | **2.13 (1)** | **1.88 (2)** | — | — |
| $\langle Fe—N_p \rangle$ (Å) | 2.01 (2) | 2.02 (2) | 1.98 (2) | 2.01 (2) | 2.03 (2) | 2.04 (2) | 2.07 (3) | 2.06 (2) |
| $Fe—N^{\epsilon 2}$ (Å) | 2.06 (1) | 2.06 (2) | **2.06 (2)** | **2.20 (2)** | 2.14 (1) | 2.11 (2) | 2.14 (2) | 2.12 (2) |

EXAFS data are from Powers et al., 1984. Bold entries indicate disagreement.

tances, respectively. Analysis of mean-square displacement amplitudes obtained in small-molecule crystallography has shown that Fe-C distances in metal carbonyls can differ between neutron and x-ray refinement and between isotropic and anisotropic refinement (Braga and Koetzle, 1987, 1988). It may be that there are systematic errors in the EXAFS analysis due to significant anharmonicity in the Fe–CO and Fe–$N^{\epsilon}$ bonds in MbCO. However, the EXAFS results are an average over all conformations, while our refinements only account for 73% of the CO occupancy. It is possible that the conformations not represented in our refinements have longer Fe–CO and Fe–$N^{\epsilon}$ distances, although the very long distances this model would imply (2.2 and 2.5 Å, respectively) make it unlikely that this is the only effect that comes into play.

## Rigid-body analysis of anisotropic Debye-Waller factors

The anisotropic analysis of atomic disorder which is enabled by the availability of high-resolution data and by the capabilities of a refinement program such as SHELXL allow additional questions about the dynamics of the molecule to be addressed within the framework of a harmonic description. More or less sophisticated analyses can be made about correlations in atomic disorder (and therefore presumably rigidity) among any groups of atoms in the protein. Correlations in disorder of parts of small molecules can be rigorously elucidated from the anisotropic displacement parameters by analysis of the $\Delta$ matrix (Dunitz et al., 1988), in which calculations are made of differences in the components of anisotropic displacement parameters of two atoms along the vector joining them. Such analyses are common in small-molecule crystallography but are not yet common in macromolecular crystallography.

As a first attempt at considering the wealth of information represented by the anisotropic disorder information, we attempted a simpler analysis that considers the protein as a collection of helices and a heme group, and asked which of these groups plausibly behave as rigid bodies in their conformational disorder at cryogenic temperature. This was done by fitting a translation-libration-screw (TLS) model of the rigid groups to the experimentally derived Debye-Waller factors (Schomaker and Trueblood, 1968). Agreement of the model with the data gives a measure of the applicability of the rigid-body model, and in cases of good agreement the three resulting tensors describe the character of the protein disorder.

We performed Schomaker and Trueblood rigid-body disorder analysis for the heme, proximal (F), and distal (E) helix using the program package PLATON (Spek, 1992) on data from *aquomet*-Mb. The equivalent analysis with $MbO_2$ data resulted in qualitatively similar results. Only main-chain atoms C, CA, and N were used to define the helices and only planar atoms were used for the heme.

As shown in Fig. 4, the agreement of the observed anisotropic disorder with the rigid-body model is good in the case of the heme and the proximal F helix, which suggests that they may indeed move as approximately rigid bodies. Although its average disorder is smaller than that of the proximal F helix, the anisotropic disorder of atoms in the distal E helix is less well described by a rigid-body model. It seems to suggest that this helix may flex and unfold rather than move as a single group. This appears to conflict with conclusions of an NMR study of collective helix motion in cyanometmyoglobin (Tolman et al., 1997), in which an analysis based on differences between calculated and observed dipole couplings concluded that both the proximal and distal helices were described within experimental uncertainty by a model using motion of rigid helices on a cone. The translation and libration tensors are listed in Table 10 and represented graphically in Fig. 5. The two largest translational modes of the heme are nearly in-plane, with eigenvalues more than twice as high as the out-of-plane direction. It is interesting to note that these translation vectors lie approximately along the direction seen for heme displacements between the deoxy and ligated states. The preference in librational movement is not as strong, but the librations around the in-plane axis are larger than the librations around the perpendicular axis.

## DISCUSSION

Myoglobin has been and still is the focus of a vast number of biophysical studies aimed at understanding the relationship between protein structure and function. The structures of Mb determined so far were obtained at lower resolutions that those obtainable today. Fine comparisons at the level of detail needed to explore the physics and chemistry of ligand binding are also hampered by the differing experimental conditions and refinement protocols employed in previous independent studies of Mb. We set out to re-determine the

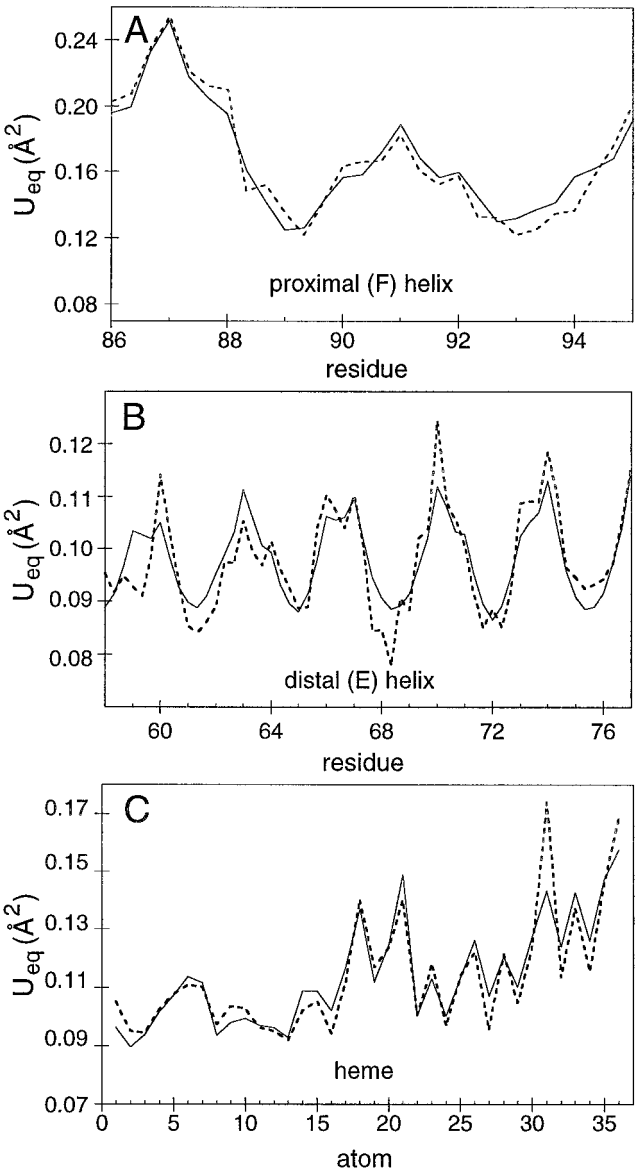


FIGURE 4 Comparison of the observed equivalent isotropic thermal factors (*dashed line*) and the ones derived from the vibration tensor of a rigid body motion (*solid line*) for C, CA, and N atoms of (*A*) the proximal (F) helix, (*B*) the distal (E) helix, and (*C*) the planar heme atoms.

crystal structures of the ferric *aquomet*-Mb complex and of the ferrous complexes of unligated, dioxygenated, and carbonmonoxynated Mb at atomic resolution, using basically the same experimental protocol to generate the four complexes and to collect and refine their diffraction data. Advanced statistical approaches were used to minimize differences not demanded by the data, thereby greatly reducing the noise in the differences between structures. As expected, the structures are very similar overall, with r.m.s. differences of $<0.25$ Å on all atoms. However, significant local differences are observed. The structures described in this paper allow one to address structural questions about Mb in different ligation states with an accuracy not previously possible (Ray et al., 1994; Phillips and Pettit, 1995; Olson and Pillips, 1996). To the best of our knowledge, these are first myoglobin complexes where the planar part of the heme, the iron, and the ligands were refined as free atoms. In the Discussion, we attempt to take the geometric information we have obtained and make a synthesis with results from other diffraction and spectroscopic studies to produce a coherent picture.

**TABLE 10 Results of TLS analysis of the anisotropic disorder in *aquomet*-Mb**

| | Translational Tensor | | Librational Tensor | |
|---|---|---|---|---|
| | Eigenvector | Value, Å$^2$ | Eigenvector | Value, deg$^2$ |
| Heme | [−0.24, −0.97, 0.08] | 0.12 | [0.01, −0.19, −0.98] | 7.8 |
| | [0.18, −0.12, −0.98] | 0.10 | [−0.58, −0.80, 0.14] | 6.4 |
| | [0.96, −0.22, 0.20] | 0.04 | [−0.82, 0.57, −0.12] | 4.5 |
| F helix | [−0.20, −0.25, 0.95] | 0.18 | [0.07, −0.16, 0.98] | 27.2 |
| | [0.06, −0.97, −0.24] | 0.09 | [0.19, −0.97, −0.17] | 4.3 |
| | [0.98, 0.01, 0.21] | 0.08 | [0.98, 0.20, −0.04] | 0.6 |

The coordinate system of these vectors is the one defined by the PDB convention.

## Bound CO geometries

We are unable to account quantitatively for 27% of the CO density in MbCO. The observation of a highly planar heme makes the possibility of significant unligated Mb contamination unlikely. Instead, the marked extensions in the density that we observe probably indicate the presence of a second CO conformation that we have not modeled. The extensions of the CO density that we observe can only belong to the minority CO conformation that is associated with the swung-out histidine conformation, because the other minority histidine conformation would be too close to the CO. These extensions lie in the direction of the position of the CO ligand in the pH 4 MbCO structure (Yang and Phillips, 1996). In the 2.0 Å low-pH structure, the CO is shifted in the general direction of the distal histidine from our majority conformation. Therefore, we identify the missing occupancy as due to a conformation associated with the doubly protonated histidine conformation. In this conformation, the CO angle is similar to that of the majority conformer and there is probably a slight shift of the heme, as seen in the pH 4 crystal structure (Yang and Phillips, 1996). The differences in orientation of this minority CO conformation must be fairly small to agree with the extensions that we see; IR linear dichroism studies report that the projection of the CO transition dipole moment onto the {001} face of monoclinic crystals for this substate differs by only 1.9° from the other substates (Sage, 1997). It is possible that previous diffraction determinations of CO geometries have been seriously affected by the presence of this conformer (particularly those at pH values $< 6$), since modeling the minority conformations without including the shifts of the heme could cause tilting of the CO in order to satisfy

restraints. Our structure determination did not restrain the iron and CO geometry.

The angle we obtain between the C–O axis and the heme plane, 18 ± 3°, is much smaller than those reported by earlier diffraction studies on monoclinic MbCO crystals, but it is identical to within experimental uncertainties to a previous determination using hexagonal crystals (Quillin et al., 1993) (see Table 1). Time-resolved IR photoselection spectroscopy experiments report the angle between the IR C–O stretch transition dipole moment and the heme transition dipole moment to be perpendicular within an uncertainty of 7° in solution (Lim et al., 1995). Measurements of IR linear dichroism in monoclinic crystals find the same angle to be 6.7 ± 0.9° (Sage, 1997). Our uncertainty estimates for this structure allow us to identify the differences between the spectroscopic and diffraction results as a 3.6-$\sigma$ discrepancy. However, a recent density-functional theory study has cast doubt on an underlying assumption of the IR work, namely that the C–O stretch transition dipole moment lies along the C–O bond axis. Taking the measured transition dipole angle from IR crystallographic measurements and the oxygen displacement from the iron atom measured

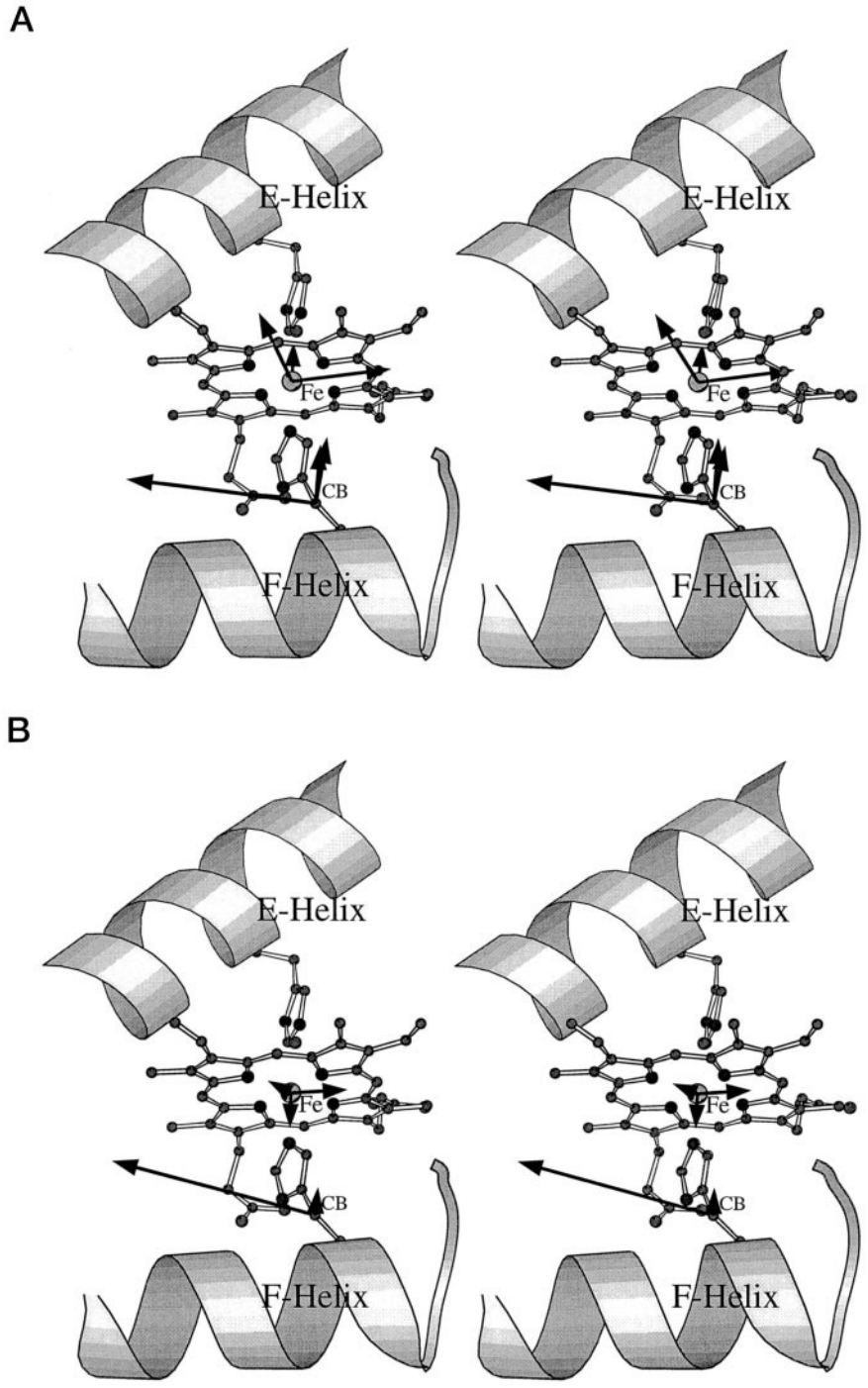


FIGURE 5 A stereo representation of the eigenvectors representing the translational (A) and librational (B) tensors of the rigid-body motion. The vectors obtained for planar heme atoms and the proximal helix are positioned into the heme iron and the proximal histidine CB atom, respectively. The length of each eigenvector is proportional to its eigenvalue.

along the heme plane from x-ray crystallographic studies, the density-functional theory calculations find a minimum-energy geometry for the CO with a tilt angle ($\tau$) of 9.5° and a bend angle ($\beta$) of 5.8° (Spiro and Kozlowski, 1998). Although the total IR angle of 15.3° agrees to within experimental uncertainty with the value we determined of 18 ± 3°, the deviation is primarily in the well-determined bend angle ($\beta$) + 9° rather than in the less-certain tilt angle of 9 ± 3°. Results combining NMR and Mössbauer measurements with density-functional theory calculation reach a slightly different conclusion of ($\tau$) = 4° and ($\beta$) = 7° (McMahon et al., 1998). It would appear that the large discrepancies among results of various techniques for the CO binding geometry in MbCO are mostly resolved. However, the accurate calculations of functional properties of MbCO being attempted now by quantum chemists depend critically on this geometrical parameters. Our MbCO coordinates were refined without geometrical constraints and they include anisotropic B factors and uncertainty estimates. They should enable more accurate calculations on the relationship between structure and function in heme proteins than was previously possible.

Structures of several heme protein model compounds (Scheidt et al., 1981, Kim et al., 1989; Kim and Ibers, 1991; Tetreau et al., 1994; Slebodnick et al., 1996a,b) all exhibit proximal histidine tilt and bond (FE–NE2–CE2) angles within 1.4 and 4.0°, respectively, of our structure. Examination of the structures of encumbered model compounds shows that upon CO binding the porphyrin ring either ruffles or domes, or both; the proximal ligand tilts [The most frequently cited standard for an ideal CO-binding geometry, Fe(TPP)(Py)(CO) (Peng and Ibers, 1976) has a 10° distortion of the proximal pyridine substituent]; or the distal cavity expands to allow a nearly perpendicular geometry. Some of these model compounds have a very similar CO binding geometry to MbCO, as shown in Table 11. Both the Fe(C2-Cap)(CO)(1-MeIm) and Fe(PocPivP)(CO) (1-MeIm) systems discriminate against CO (Slebodnick et al., 1996a,b) compared to an unencumbered system (Kim and Ibers, 1991). The structural and vibrational data on the PocPiv and $C_2$-Cap model complexes have been used to calculate steric and electronic energies from which it was concluded that the affinity decrease resulted mainly from steric interactions (Ray et al., 1994).

## Conformational substates in MbCO

The infrared spectrum of MbCO contains three stretch bands of the CO bound to the heme iron, denoted $A_0$ (1969 $cm^{-1}$), $A_1$ (1945 $cm^{-1}$), and $A_3$ (1927 $cm^{-1}$) that have been attributed to different conformational substates (Alben et al., 1982; Frauenfelder et al., 1991). The relative intensities of the bands have been shown to vary with temperature (Ansari et al., 1987), pressure (Frauenfelder et al., 1990), pH (Fuchsman and Appleby, 1979), and ionic strength (Müller, 1997), with $A_0$ favored at low pH. Each band has a different kinetic barrier for rebinding of CO to the heme after photolysis (Alben et al., 1982; Frauenfelder et al., 1991). Therefore, understanding their structural origin is relevant to understanding the relationship between structure and function in heme proteins.

Initially it was believed that the bands were caused by different orientations the CO relative to the heme normal (Ormos et al., 1988). This interpretation was supported by diffraction studies that showed two orientations for the bound CO (Kuriyan et al., 1986; Cheng and Schoenborn, 1991). More recently, it has been suggested based on results from simulation (Jewsbury and Kitagawa, 1994), x-ray diffraction (Yang and Phillips, 1996), spectroscopy (Li and Spiro, 1988; Park et al., 1991; Ray et al., 1994), and mutant studies (Li et al., 1994) that different distal histidine conformers, each producing a different local electric field, give rise to the three infrared CO bands. Our near-atomic resolution MbCO structure allows us to test these and other hypotheses about the origins of the A substates.

The proximal histidine shows a single well-ordered conformation in MbCO, suggesting that the differences between $A_1$ and $A_3$ do not arise from a subtle *trans* effect from H97. Similarly, the suggestion that the substates correspond to differently tilted CO orientations does not agree with our data. The single bound CO geometry in our MbCO structure refined to 73% occupancy and the residuals in the electron density allow only slight differences for the CO geometry of the remainder of the population. Moreover, the observed extensions of the CO electron density must correspond to the swung-out substate, because they would be too close to the histidine in the other conformers. Oldfield et al. (1991) followed by Jewsbury and Kitagawa (1994) have suggested that $A_3$ might be a tautomer of the singly protonated substate of the distal histidine with the imidazole ring rotated

**TABLE 11 CO tilt and bend angles in MbCO and model compounds**

| Compound | Reference | Fe—CO tilt, ° | Fe—C—O bend angle, ° |
|---|---|---|---|
| MbCO | This work | 9.0 | 9.0 |
| Fe($OC_3$OPor)(CO)(1-MeIm) | Slebodnick et al., 1996 | 7.7 | 6.1 |
| Fe($C_2$-Cap)(CO)(1-MeIm) | Kim and Ibers, 1997 | | |
| molecule 1 | | 5.5 | 7.2 |
| molecule 2 | | 4.1 | 4.1 |
| Fe($\beta$-PocPivP)(CO)(1,2$Me_2$Im)* | Kim et al., 1989 | 6.1 | 7.6 |

*The CO adduct of this compound has a porphyrin that is strongly ruffled.

by 180° about $\chi_2$, resulting in $N^{\delta}$ pointing into the pocket as opposed to out toward the solvent, as in the $A_1$ substate. However, we can rule out rotamers of the distal histidine as an explanation of the A substates, since they would require changes in the position of the hydration partner of the outward nitrogen that are not consistent with the observed density.

The only structural feature that is correlated with the infrared A substates is the conformation of the distal histidine. The simultaneous observation of three histidine conformers in this study agrees with what had been expected from mutant studies, theory, and simulation about the underlying cause of the A substates. We identify $A_1$ with the majority distal histidine conformation ($N^{\epsilon}$ − carbonyl distance = 3.2 Å) (Sage, 1997; Nienhaus et al., 1998). The swung-out conformation of the distal histidine is essentially the same as seen in the low-pH MbCO structure (Yang and Phillips, 1996), which leads to the identification of this conformer with the $A_0$ substate. This leaves the identification of the minority conformer with the $N^{\epsilon}$–carbonyl oxygen distance of 2.7 Å as the $A_3$ substate. This identification makes sense in terms of the barriers to ligand rebinding, since the increased steric hindrance of the ligand with the distal histidine in this conformation would lead to the higher barrier seen for $A_3$. The occupancies we find for the three histidine conformers in the crystal at pH 6.0 of 60/20/20 are similar to the relative occupancies of the IR bands in similarly prepared monoclinic crystals at pH 7.0 of 85% $A_1$, 10% $A_3$, and 5% $A_0$ (Nienhaus et al., 1998). Lowering the pH favors populating $A_0$, and cooling rates can make minor differences in population ratios as well (Chu et al., 1993). This agrees to within the uncertainties in our population estimates.

The distances from the CO oxygen to the distal histidine $N^{\epsilon}$ of 3.2 Å in $A_1$ and 2.7 Å in $A_3$, respectively, suggest that a hydrogen bond is present in $A_3$, but is much weaker or not present in $A_1$. A hydrogen bond would require a protonated $N^{\epsilon}$ tautomer. Despite the quality of our data and model, the data do not allow us to assign the tautomer state. Direct data on the protonation state for the distal histidine comes from neutron crystallography of deuterated MbCO and shows a fully occupied $N^{\delta}$ deuterium with Debye-Waller factors close to the average for the structure (Cheng and Schoenborn, 1991). The size of the error in the occupancy of the $DN^{\delta}$ tautomer may be obtained based on IR spectroscopy of deuterated MbCO solutions; Hong (1989) measured an $A_0$ population of 25% in a pD 5.7 sample.

An inverse correlation between the majority IR C–O stretch frequency and the Raman Fe–CO stretch frequency is observed across a wide range of heme proteins and model compounds with imidazole as fifth ligands (Li and Spiro, 1988; Ray et al., 1994). This inverse correlation is evidence for back-bonding between the iron and CO, since withdrawing electrons from the CO system (e.g., through interaction with a nearby positive charge) results in transfer of more electron density from the iron $d_{\pi}$ orbital to the CO $\pi^*$ orbital, which is strongly anti-bonding. This weakens the C–O bond and strengthens the Fe–C bond, thereby downshifting the C–O stretch frequency while upshifting the Fe–CO stretch frequency. Any explanation of the A substates in MbCO needs to be consistent with this picture, in which the majority $A_1$ substate at neutral pH and the majority $A_0$ substate at low pH exhibit the same correlation as the model compounds. The Fe–CO stretch frequency corresponding to $A_3$ is somewhat ambiguous, but one assignment suggests that $A_3$ may not exhibit the same correlation (Ray et al., 1994).

There are three alternative identifications of the A substates, each with an accompanying interpretation of our structures, the neutron diffraction results, and the spectroscopic results:

1. As suggested by Ray et al. (1994), the $A_1$ substate could be the $HN^{\epsilon}$ tautomer that is hydrogen-bonded to the CO, resulting in a downshift of the CO stretch frequency from the $A_0$ state. In this model, the $A_3$ substate is the $HN^{\delta}$ tautomer (which was proposed by Ray et al. to be the majority population in the crystals used for the neutron study), and the $N^{\epsilon}$ lone pair makes a donor interaction with the $\pi^*$ orbital of CO, as suggested by Maxwell and Caughey (1976). The donor interaction changes the back-bonding and causes the $A_3$ substate to deviate from the inverse correlation between $\nu_{Fe-CO}$ and $\nu_{C-O}$ seen for $A_0$ and $A_1$. This model can explain the spectroscopic data, but requires that $A_3$ be the dominant substate for the two neutron structures (Hanson and Schoenborn, 1981; Cheng and Schoenborn, 1991). While orthorhombic MbCO crystals can have an $A_3$ population that is as large as $A_1$ (Makinen et al., 1979; Mourant et al., 1993), this effect has not been seen in monoclinic crystals by any high-quality IR measurements. Instead, the IR data on similarly prepared crystals indicate that the substate ratio is approximately the same in $P2_1$ crystals and solutions at similar pH or pD (Sage, 1997; Nienhaus et al., 1998), with $A_1$ dominant by a factor of almost 10. Moreover, explanation of our structures by this model requires either the donor interaction to have shorter $N^{\epsilon}$–ligand distances than the hydrogen-bonded interaction, or it would require our structure to have $A_3/A_1$ ratios of 3:1. The first is unlikely from chemical arguments, and the second has never been seen in IR studies of any crystal form or solutions.
2. $A_1$ and $A_3$ could both be the $HN^{\epsilon}$ tautomer, both downshifted from $A_0$ through back-bonding. The difference between the stretch frequencies of $A_1$ and $A_3$ in this picture would be due to a stronger interaction with the $N^{\epsilon}$ proton in the $A_3$ state, which is reflected in the shorter O–$N^{\epsilon}$ distance seen in our structure. This stronger hydrogen bond might decrease the bond order of C–O and enhance the occupancy of the non-bonded $sp^2$ orbitals of the oxygen atom. Presumably the difference is ascribed to alternative connectivity of the hydrogen-bonding network extending from $N^{\delta}$, as seen in the *aquomet*-Mb density. This model completely disagrees with the results

of the neutron diffraction experiments on MbCO that locate the deuteron on $N^{\delta}$ only. A possible explanation of this apparent discrepancy would be to assume that the MbCO crystals used in the neutron study were at least partly oxidized, resulting in the $N^{\delta}$ tautomer. Oxidation might have occurred even under CO atmosphere due to the long time period (many weeks) for equilibration against $D_2O$ and data acquisition with the crystals kept at relatively low pD (pD 5.7). In this interpretation, the progressive decrease seen in $\nu_{CO}$ as additional proton donors are added to the pocket by the V68T and V68D mutants causes the $A_1$ peak to shift from 1945 to 1932 to 1916 $cm^{-1}$, respectively. In the V68T and V68D mutants, the peak increases to 1960–1970 $cm^{-1}$ (Decatur and Boxer, 1995; Li et al., 1994; Anderton et al., 1997).

3. $A_3$ could be the $HN^{\epsilon}$ tautomer with a hydrogen bond to the CO, and $A_1$ could be the $HN^{\delta}$ tautomer with a lone-pair $N^{\epsilon}$ near the CO. This model is consistent with the diffraction data, but requires explanation with regard to the spectroscopic data. One might expect that the negative environment of the lone-pair $N^{\epsilon}$ in $A_1$ in this model would upshift the CO stretch frequency and that the positive environment hydrogen bond in the $A_3$ substate would downshift the CO stretch frequency with respect to the apolar $A_0$ environment, but this contradicts the experimental observation that $A_1$ and $A_3$ are both downshifted from $A_0$. Our data suggest that the underlying premise—that the only significant effect of the substates is through different electrostatic environments at the ligand—may not be true. The displacement of the CO and heme group in $A_0$ suggested by the extensions on the CO electron density and by the low-pH structure (Yang and Phillips, 1996) imply that $A_0$ is likely to have significant steric differences in the CO, proximal histidine, and heme from $A_1$ and $A_3$. Small differences in geometry can have effects on the stretch frequencies that are as large as or larger than the effects of the electric field at the ligand (Kushkuley and Stavrov, 1997). In this model, $A_1$ and $A_3$ have the same heme-CO geometry, and the interaction of the lone-pair nitrogen $N^{\epsilon}$ with the CO $\pi^*$ would result in upshifting of $A_1$ relative to $A_3$ where a hydrogen bond is present. Thus, $A_0$ would be upshifted from $A_1$ not by the change in electrostatic interaction (the sign is incorrect for this argument) but by significantly different steric interactions that would probably involve the proximal histidine. This model provides qualitative agreement with all of the spectroscopic and diffraction data of which we are aware. Conversion between the $HN^{\epsilon}$ and $HN^{\delta}$ tautomers would presumably require a large-scale fluctuation of the distal histidine similar in size to the change between the swung-in and swung-out conformations. This view is consistent with results from dynamics studies that show the exchange among $A_0$, $A_1$, and $A_3$ is viscosity-dependent and involves rearrangement of many bonds (Young et al., 1991). The fact that $A_0$ and $A_1$ lie on the same line of $\nu_{Fe-CO}$ vs. $\nu_{C-O}$ is attributable to the linearity of back-bonding in this regime and does not distinguish between steric and electrostatic contributions to back-bonding. Detailed calculations will be needed to see if the agreement can be made quantitative. It is illustrated in Fig. 6.

Recent Raman measurements of the Fe–CO stretch frequency show a shift of $+1$ $cm^{-1}$ upon H/D exchange in wild-type MbCO, but no shift in mutants that cannot hydrogen-bond to the CO (Unno et al., 1998). This was interpreted in terms of a hydrogen bond from the CO to the distal histidine in the $A_1$ state. However, measurements of the IR stretch bands of bound CO upon H/D exchange show that the isotope shift seen in the $A_1$ substate is small ($\approx$1 $cm^{-1}$) and the same for $A_1$ and the $A_0$ substate, which cannot hydrogen-bond to the CO (Hong, 1989). This result is therefore not evidence for a hydrogen bond to the CO in $A_1$, but rather evidence of the effects of deuterium substitution on the hydrogen-bonding network between the distal histidine and the heme, and therefore would have no bearing on model 3.

## Conformational substates in $MbO_2$ and the photolysis yield

Our data for the $MbO_2$ complex are the highest quality of any of the ligation states, and the structure was determined to the highest resolution (1.0 Å). The electron density for the $O_2$ ligand is peanut-shaped and can be modeled with a single $O_2$ orientation. The electron density for the distal histidine can either be modeled by a single orientation with highly extended disorder in the plane of the imidazole ring, or by two 50%-occupied conformers. In the two-conformer model, the first conformer is characterized by a 2.7 Å distance from the distal histidine $N^{\epsilon}$ to the ligand terminal oxygen, a distance consistent with a hydrogen bond. The second conformer has a 3.0 Å distance to the terminal oxygen. We prefer the two-conformer model because IR spectra of $MbO_2$ with various isotopes of $O_2$ show two bands ($\approx$1150 and $\approx$1135 $cm^{-1}$) that have been ascribed to conformational substates with different hydrogen bonding (Potter et al., 1987; Miller and Chance, 1995). The size of the shift between these bands is large (15 $cm^{-1}$) when compared with the shift of model compounds in hydrogen-bonding versus non-hydrogen-bonding solvents ($\approx$5 $cm^{-1}$) (Potter et al., 1987).

A recent FTIR study on the photolysis yield of horse oxymyoglobin finds only 50% photolyzable at neutral pH and assigns the photolyzable fraction to the 1150 $cm^{-1}$ band and the non-photolyzable (or rapidly rebinding) fraction to the 1135 $cm^{-1}$ band (Miller and Chance, 1995). The apparent photolysis yield decreases at low pH, showing a titration behavior that parallels the pH dependence of the $A_0$ population in MbCO (Miller et al., 1996). This strongly suggests that the doubly protonated, swung-out conformer of the distal histidine (which cannot hydrogen-bond to the $O_2$) has low photolysis yield (Miller et al., 1996). The apparent

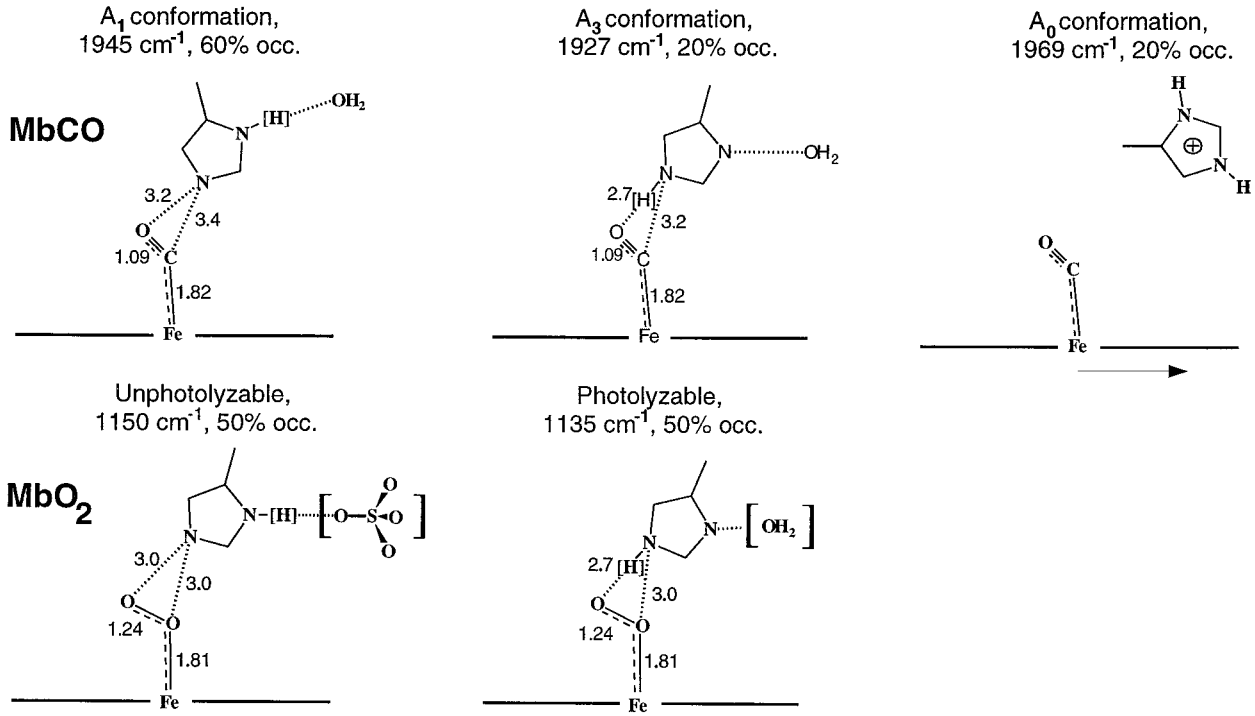


FIGURE 6 A schematic view of the conformations of the distal histidine. Shown is one possible tautomer state of the distal histidine (model 3, see text for details and the other models), with connections to the spectroscopic conformational substates of MbCO and $MbO_2$, respectively.

photolysis yield at low pH reaches an asymptotic value of ~10% photolyzable, even though the population of the hydrogen-bonded fraction goes to zero. FTIR spectra of MbCO at pH 4 show an $A_0$-like band with an asymmetric feature that has been suggested to arise from a 10%-occupied water in the pocket (Müller, 1997). This suggests that the difference between the photolyzable conformation and the non-photolyzable conformation is the presence or absence of a hydrogen bond, respectively.

The neutron structure of $MbO_2$ shows a fully occupied $HN^{\epsilon}$ tautomer (Phillips and Schoenborn, 1981). As with MbCO, there are three alternative identifications of the spectroscopic substates of $MbO_2$:

1. The two O–O stretch bands could represent different orientations of Fe–O–O and the proximal histidine imidazole planes, as suggested by Potter et al. (1987). This explanation would be consistent with the neutron results, but our findings of a single conformation of the ligand and proximal histidine make this explanation unlikely. To be consistent with the electron density we observe, the possible range between conformations would be too small to have such a large effect on the O–O stretch frequency.
2. The two O–O stretch bands could both arise from interactions with the $HN^{\epsilon}$ tautomer of the distal histidine, with the difference between them due to a weaker interaction with the $N^{\epsilon}$ proton in the unphotolyzable state (3.0 Å $N^{\epsilon}$–O distance). As in the equivalent proposal for MbCO, presumably this difference would be ascribed to alternative connectivity of the hydrogen-bonding network extending from $N^{\delta}$, as seen in the *aquomet*-Mb density. For $MbO_2$, this explanation agrees with the neutron results, but it does not explain why the spectroscopic shift between the substates (representing a fractional change in hydrogen bonding) should be three times larger than that seen between model compounds in hydrogen-bonding and non-hydrogen-bonding solvents (Potter et al., 1987).
3. The photolyzable fraction could correspond to the $HN^{\epsilon}$ tautomer of the distal histidine (which corresponds to the $A_3$ substate in MbCO, see model 3) and it forms a 2.7-Å hydrogen bond to $O_2$. The non-photolyzable fraction would correspond to the $HN^{\delta}$ tautomer (which corresponds structurally to the $A_1$ substate in MbCO), as illustrated in Fig. 6. This is the only model in which the size of the spectroscopic shift between substates can be accounted for (due to the differing signs of the interaction for the two conformations) and because the assignment of a 50%-occupied $HN^{\delta}$ tautomer with a 50%-occupied sulfate molecule bound to it from the solvent side would explain the observed electron density at the water/sulfate position.

However, there are two points that need explanation in interpretation 3). The first is the neutron data, which show a fully occupied $HN^{\epsilon}$ tautomer. As with the MbCO data, a conformation with up to 30% occupancy would probably not be seen because of phase noise in the map, but 50% seems too high. One possible explanation of the full occu-

pancy of the $HN^{\epsilon}$ tautomer might be the formation of some amount of alkaline *aquomet*-Mb during the course of data collection (Austin and Drabkin, 1935).

The second point that needs to be addressed is whether a 3.0 Å distance from a lone pair on $N^{\epsilon}$ to the bound dioxygen (which may be highly charged) is a very high-energy conformation that should not be observed. There has been a great deal of controversy over the last 30 years about whether dioxygen bound in heme proteins is superoxide-like or neutral. Seeming evidence for the former arises from the electronic spectrum (Weiss, 1964), the frequency of the IR O–O stretch bands (Barlow et al., 1973), the change in the x-ray absorption spectra upon oxygenation (Bianconi et al., 1985), and the large asymmetry in the charge distribution on the iron as measured by Mössbauer studies (Lang and Marshall, 1966; Weissbluth and Maling, 1967). Evidence that seems more consistent with a neutral dioxygen model comes primarily from the diamagnetism of the heme-$O_2$ as seen in magnetic susceptibility (Pauling and Coryell, 1936). However, as several authors have pointed out, arguments over correspondence of experimentally determined quantities with oxidation state formalism do little to advance our understanding of the problem (Momenteau and Reed, 1994). For our present purpose, the most important question is the size of the interaction energy of the dioxygen with the protonated or unprotonated histidine. None of the experiments directly probe the electrostatic charge on the dioxygen, and at the moment none of the theoretical calculations of the charge includes the effects of the environment in a realistic way. The most recent quantum-mechanical calculations of electrostatic charge on the bound $O_2$ range from 0.34 electrons (Bertran et al., 1991), to nearly zero (Nakatsuji et al., 1996), to 0.26 electrons (Rovira et al., 1997). A realistic treatment of the interactions among the heme, bound dioxygen, and proximal and distal histidines with reference to the latest structural and spectroscopic data is still a matter for future work. At our present level of understanding, electrostatic interactions with the bound dioxygen (which is probably only slightly negatively charged) does not rule out the possibility of a 50%-occupied $HN^{\delta}$ tautomer.

However, $MbO_2$ low-temperature photolysis yield-data on the mutant myoglobin His-64–Leu, in which the distal histidine is replaced by a leucine that is not capable of forming a hydrogen bond, pose a serious challenge to model 3. The apparent $O_2$ low-temperature quantum yield in the His-64–Leu mutant is near unity (Miller et al., 1998), while the x-ray structure of the CO-bound complex shows high correlation between the leucine position and that of the histidine in the native structure and no crystallographically bound water in the pocket (Quillin et al., 1993; Christian et al., 1997). How do we reconcile this seemingly contradictory evidence from model 3 and the IR and Raman data? We propose that the quantum yield for the His-64–Leu $MbO_2$ mutant is high because there is a water in the pocket of the $O_2$ adduct of this mutant (most likely near the position of $N^{\epsilon}$ of His-64 in wild-type $MbO_2$) which hydrogen bonds to the bound oxygen molecule. This would explain the observation that in this mutant the $\nu_{Fe-CO}$ Raman band is displaced by 40 $cm^{-1}$ from the wild-type, while $\nu_{Fe-O2}$ is changed by only 2 $cm^{-1}$ (Hirota et al., 1996). Cavity analysis of a model of the His-64–Leu $MbO_2$ mutant shows that there is indeed a small cavity between CD2 of Leu-32 and the $O_2$. Potential-of-mean-force analysis of this region shows a slight water peak that is not found in the CO structure, which indicates the possibility of a water molecule at this site in the $O_2$ adduct (Gerhard Hummer, unpublished results).

The reasons why a hydrogen-bonded conformation should rebind with a barrier similar to that of MbCO while a non-hydrogen-bonded conformation rebinds without barriers is not clear. Miller and Chance (1994) suggest that the difference is due to the trajectory of the photodissociated $O_2$, with the hydrogen-bonded substate giving a "kick" to the photodissociated $O_2$. Another possibility is that the hydrogen bond can affect the spin state of the photodissociated $O_2$ or heme, perhaps through formation of highly reactive singlet $O_2$ species in the non-hydrogen-bonded conformation. Spectroscopic measurements on $MbO_2$ on subpicosecond time scales are needed to discriminate between the possibilities.

## Dynamics

High-resolution diffraction data present an opportunity for making models of correlated motions in proteins at a higher level than that taken in typical molecular dynamics simulations. Knowledge about correlated disorder among sets of atoms in a structure can be used to build a model more like an engineer's model of a bridge, where "girders" and other rigid mechanical elements are identified and abstracted with properties such as Young's moduli. However, motions in proteins have liquid-like properties. Anisotropic refinement of Debye-Waller factors (as implemented in SHELXL), though a great improvement over isotropic refinement, is rooted firmly in a harmonic description (García et al., 1997). The increasing number of atomic-resolution structures (there are now ~40 distinct protein structures with resolution better than 1.2 Å in the PDB) warrant development of analysis methods that can extract more of the information in such data, with the goal of moving beyond structure to addressing crucial problems of dynamics and function. The information present in our high-resolution data could be re-analyzed outside a harmonic framework and compared with similarly analyzed molecular dynamics trajectories, see, for example, van Aalten et al. (1997).

Our analysis of the anisotropic Debye-Waller factors that we obtained was crude, but produced an interesting result: the proximal (F) helix and heme show disorder characteristic of rigid bodies, while the distal (E) helix does not. Evidence from rebinding kinetics of wild-type and mutant myoglobins and from fluorescence quenching of internal residues by $O_2$ indicates there is no single pathway for ligand entry and escape, but rather a multitude of them

(Yedgar et al., 1991; Huang and Boxer, 1994; Frauenfelder et al., 1998). Rather than thinking of ligand diffusion through a fluctuating protein as transport through one or a few "gates," one might instead put the motions that control this process on a continuum that includes harmonic motions around local minima at one end and unfolding of the protein at the other end. From our data we would expect that a satisfactory model of ligand entry and exit would show the F helix as a rigid body, while the E helix significantly bends or unfolds. There are, however, substantial technical barriers to be overcome before simulations of fully hydrated Mb will reach the time scale of the fluctuations that accompany ligand exit from the pocket.

## CONCLUSIONS

These structures are derived from data of high resolution and high quality, and for the moment the 1.0 Å $MbO_2$ structure is the highest-resolution structure available of a heme protein. Consistent treatment of samples and data and the latest technical improvements for collecting and analyzing diffraction data have yielded a set of four structures of myoglobin in different ligation states with exceptionally low noise in the differences between the structures. For this reason, we expect that these structures will form the basis for many future calculations about the dynamics and function of the molecule, even after higher-quality individual Mb structures become available (although an unligated Mb structure that is free of *aquomet*-Mb contamination would be of significant benefit).

We have produced structural models of MbCO and $MbO_2$ that make a bridge between results from diffraction and spectroscopic techniques. We also were able to comment about the dynamics of the molecule from the anisotropic Debye-Waller factors we obtained. Both of these efforts will need to be deepened and expanded in order to gain more useful information about this important model system for protein reactions.

Near-atomic resolution structures of unligated Mb (1.15 Å), *aquomet*-Mb (1.2 Å), and Mb.CO (1.15 Å) were reported recently by Bartunik and co-workers (Kachalova, G. S., A. N. Popov, and H. Bartunik. 1999. A steric mechanism for inhibition of CO binding to heme proteins. 1999. *Science*. 284:473–476) at ambient temperature. Their results differ significantly from ours (listed in parentheses) in terms of the CO binding geometry (tilt angle 4.8 ± 0.9° (9.0 ± 3.0°), bend angle 7.4 ± 1.9° (9.0°); Fe–C distance 1.73 ± 0.03 Å (1.82 ± 0.02 Å)), the lack of observation of multiple conformations of either the distal histidine, or the bound CO. In addition, we do not observe the strong rotational motion of the E and F helixes between the unligated and the CO-bound state. Further experiments will be needed to establish the reasons for these discrepancies.

We would like to point out that one of us (J.B.) very strongly favors the models involving a $N^{\delta}$ tautomer (model 3 for the CO A-states and the non-photolyzable $O_2$ fraction, respectively).

We are grateful to Hans Frauenfelder, Angel García, Gerhard Hummer, Christiane Jung, Ben McMahon, Jose Onuchic, Benno Schoenborn, Tom Terwilliger, and Elizabeth Wilkinson for help, advice, and many stimulating discussions. We thank Thomas Schneider for advice with SHELXL and TLS calculations, Christoph Kratky and Oliver Sauer for providing their xenon derivative data, George Phillips for critical discussion of the manuscript, and the reviewers for their helpful and productive comments.

This research was supported in part by the International Human Frontiers Organization and the LDRD program of Los Alamos National Laboratory. The $MbO_2$ diffraction data were collected at the Integrated Structural Biology Resource of LANL.

## REFERENCES

Alben, J. O., D. Beece, S. F. Bowne, W. Doster, L. Eisenstein, H. Frauenfelder, D. Good, J. D. McDonald, M. C. Marden, P. P. Moh, L. Reinisch, A. H. Reynolds, E. Shyamsunder, and K. T. Yue. 1982. Infrared spectroscopy of photodissociated carboxymyoglobin at low temperatures. *Proc. Natl. Acad. Sci. USA.* 79:3744–3748.

Anderton, C., R. Hester, and J. Moore. 1997. A chemometric analysis of the resonance Raman spectra of mutant carbonmonoxy myoglobin. *Biochim. Biophys. Acta*. 1338:107–120.

Ansari, A., J. Berendzen, D. Braunstein, B. R. Cowen, H. Frauenfelder, M. K. Hong, I. E. T. Iben, B. Johnson, P. Ormos, T. B. Sauke, R. Scholl, A. Schulte, P. J. Steinbach, J. Vittitow, and R. D. Young. 1987. Rebinding and relaxation in the myoglobin pocket. *Biophys. Chem.* 26: 337–355.

Antonini, E., and M. Brunori. 1971. Hemoglobin and Myoglobin in Their Reactions with Ligands. North-Holland, Amsterdam.

Austin, R. H., K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus. 1975. Dynamics of ligand-binding to myoglobin. *Biochemistry*. 14:5355–5373.

Austin, J. H., and D. L. Drabkin. 1935. Spectrophotometric studies III. methemoglobin. *J. Biol. Chem.* 112:67–78.

Barlow, C. H., J. C. Maxwell, W. J. Wallace, and W. S. Caughey. 1973. Elucidation of mode of binding of oxygen to iron in oxyhemoglobin by infrared spectroscopy. *Biochem. Biophys. Res. Commun.* 55:91–95.

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.

Bertran, J., M. F. Ruiz-Lopez, and D. Rinaldi. 1991. An INDO study of environmental effects on the dioxygen position of an oxymyoglobin model. *J. Mol. Struct.* 232:337–347.

Bianconi, A., A. Congui-castellano, M. Dell'Ariccia, A. Giovannelli, E. Burattini, and P. J. Durham. 1985. Increase of the Fe effective charge in hemoproteins during oxygenation process. *Biochem. Biophys. Res. Commun.* 131:98–102.

Braga, D., and T. F. Koetzle. 1987. Treatment of light atoms in X-ray structural studies on metal-carbonyl clusters: a critical view. *J. Chem. Soc., Chem. Commun.* 3:144–146.

Braga, D., and T. F. Koetzle. 1988. A mean-square displacement amplitude analysis of terminally bound CO groups in transition-metal clusters. *Acta Crystallogr. B.* 44:151–155.

Brünger, A. T. 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*.355:472–475.

Brünger, A. T., A. Krukowski, and J. W. Erickson. 1990. Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. A.* 46:585–593.

Brünger, A. T., J. Kuriyan, and M. Karplus. 1986. Crystallographic R factor refinement by molecular dynamics. *Science*. 235:458–460.

Carver, T. E., R. E. Brantley, Jr., E. W. Singleton, R. M. Arduini, M. L. Quillin, Jr., and J. S. Olson. 1992. A novel site-directed mutant of myoglobin with an unusually high $O_2$ affinity and low autooxidation rate. *J. Biol. Chem.* 267:14443–14450.

Chance, M. R., S. H. Courtney, M. D. Chavez, M. R. Ondrias, and J. M. Friedman. 1990. $O_2$ and CO reactions with heme proteins: quantum yields and geminate recombination on picosecond time scales. *Biochemistry.* 29:5537–5545.

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

Christian, J. F., M. Unno, J. T. Sage, P. M. Champion, E. Chien, and S. G. Sligar. 1997. Spectroscopic effects of polarity and hydration in the distal heme pocket of deoxymyoglobin. *Biochemistry*. 36:11198–11204.

Chu, K., G. U. Nienhaus, and R. Philipp. 1993. Structural heterogeneity in proteins at cryogenic temperatures: cooling rate dependence. *Chem. Phys. Lett.* 216:275–280.

Crozier, E. D., J. J. Rehr, and R. Ingalls. 1988. X-Ray Absorption. D. C. Koningsberger and R. Prins, editors. Wiley, New York.

Decatur, S. M., and S. G. Boxer. 1995. A test of the role of electrostatic interactions in determining the CO stretch frequency in carbonmonoxy-myoglobin. *Biochem. Biophys. Res. Commun.* 212:159–164.

Dickerson, R. E., and I. Geis. 1983. Hemoglobin. Benjamin/Cummings, Menlo Park, CA.

Dunitz, J. D., V. Shomaker, and K. N. Trueblood. 1988. Interpretation of atomic displacement parameters from diffraction studies of crystals. *J. Phys. Chem.* 92:856–867.

Engh, R. A., and R. Huber. 1991. Accurate bond and angle parameters for x-ray protein-structure refinement. *Acta Crystallogr. A.* 47:392–400.

Feitelson, J., and S. Yedgar. 1991. The effect of the solvent viscosity on the migration of small molecules through the structure of myoglobin. *Biorheology*. 28:99–105.

Frauenfelder, H. 1989. The Debye-Waller factor. *Int. J. Quant. Chem.* 35:711–715.

Frauenfelder, H., N. A. Alberding, A. Ansari, D. Braunstein, B. R. Cowen, M. K. Hong, I. E. T. Iben, J. B. Johnson, S. Luck, M. C. Marden, and J. R. Mourant. 1990. Proteins and pressure. *J. Phys. Chem.* 94: 1024–1037.

Frauenfelder, H., B. McMahon, B. P. Stojković, and K. Chu. 1998. Is myoglobin like a Swiss watch? Proc. Adriatico Res. Conf. L. Mattsson, ed. World Scientific, New Jersey. 145–164.

Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science*. 254:1598–1603.

Frazao, C., C. M. Soares, M. A. Carrondo, E. Pohl, Z. Dauter, K. S. Wilson, M. Hervas, J. A. Navarro, M. A. De la Rosa, and G. M. Sheldrick. 1995. Ab initio determination of the crystal structure of cytochrome $c_6$ and comparison with plastocyanin. *Structure*. 3:1159–1169.

Fuchsman, W. H., and C. A. Appleby. 1979. CO and $O_2$ complexes of soybean leghemoglobins: pH effects upon infrared and visible spectra. Comparisons with CO and $O_2$ complexes of myoglobin and hemoglobin. *Biochemistry*. 18:1309–1321.

García, A. E., J. A. Krumhansl, and H. Frauenfelder. 1997. Variations on a theme by Debye and Waller: from simple crystals to proteins. *Proteins: Struct., Funct., Genet.* 29:150–153.

Garman, E., and T. R. Schneider. 1997. Macromolecular Cryocrystallography. *J. Appl. Crystallogr.* 30:211–237.

Ghosh, A., and D. F. Bocian. 1996. The CO tilting and bending potential-energy surface of carbonmonoxyhemes. *Biophys. J.* 70:388.

Hanson, J. C., and B. P. Schoenborn. 1981. Real-space refinement of neutron-diffraction data from sperm wale carbonmonoxy-myoglobin. *J. Mol. Biol.* 153:117–146.

Hirota, S., T. Li, G. N. Phillips, Jr., J. S. Olson, M. Mukai, and T. Kitagawa. 1996. Perturbation of the Fe–$O_2$ bond by nearby residues in the heme pocket: observation of $\nu_{Fe-O2}$ Raman bands for oxymyoglobin mutants. *J. Am. Chem. Soc.* 118:7845–7846.

Hofrichter, J., and W. A. Eaton. 1976. Linear dichroism of biological chromophores. *Annu. Rev. Biophys. Bioeng.* 5:511–560.

Hong, M. K. 1989. FTJR studies of environmental effects on carbonmonoxy myoglobin. Thesis. University of Illinois at Urbana-Champaign.

Huang, X. H., and S. G. Boxer. 1994. Discovery of new ligand-binding pathways in myoglobin by random mutagenesis. *Nat. Struct. Biol.* 1:226–229.

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., and T. Kitagawa. 1995. Distal residue CO interaction in carbonmonoxy myoglobins: a molecular-dynamics study of three distal mutants. *Biophys. J.* 68:1283–1294.

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 heme prosthetic unit. *J. Am. Chem. Soc.* 116:11586–11587.

Jeyarajah, S., L. M. Proniewicz, H. Bronder, and J. R. Kincaid. 1994. Low-frequency vibrational modes of oxygenated myoglobin, hemoglobins, and modified derivatives. *J. Biol. Chem.* 269:31047–31050.

Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A.* 47:110–119.

Kim, K., J. Fettinger, J. L. Sessler, M. Cyr, J. Hugdahl, J. P. Collman, and J. A. Ibers. 1989. Structural characterization of a sterically encumbered iron(II) porphyrin CO complex. *J. Am. Chem. Soc.* 111:403–405.

Kim, K., and J. Ibers. 1991. Structure of a carbon monoxide adduct of a capped porphyrin Fe-C-2-CAP-CO-1-methylimidazole. *J. Am. Chem. Soc.* 113:6077–6081.

Kuriyan, J., S. Wilz, M. Karplus, and G. Petsko. 1986. X-ray structure and refinement of carbon-monoxy (FeII)-myoglobin at 1.5 Å resolution. *J. Mol. Biol.* 192:133–154.

Kushkuley, B., and S. S. Stavrov. 1997. Theoretical study of the electrostatic and steric effects on the spectroscopic characteristics of the metal-ligand unit of heme proteins. 2. C-O vibrational frequencies, $^{17}O$ isotropic chemical shifts, and nuclear quadrupole coupling constants. *Biophys. J.* 72:899–912.

Lang, G., and W. Marshall. 1966. Mössbauer effect as seen in some haemoglobin compounds. *J. Mol. Biol.* 18:385–404.

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

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

Lim, M., T. A. Jackson, and P. A. Anfinrud. 1995. Binding of CO to myoglobin from a heme pocket docking site to form a nearly linear Fe-C-O. *Science*. 269:962–966.

Luzatti, P. V. 1952. Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr.* 5:802–810.

Makinen, M. W., R. A. Houtchens, and W. S. Caughey. 1979. Structure of carboxymyoglobin in crystals and in solution. *Proc. Natl. Acad. Sci. USA*. 76:6042–6046.

Maxwell, J. C., and W. S. Caughey. 1976. An infrared study of NO bonding to heme B and hemoglobin A. Evidence for inositol hexaphosphate induced cleavage of proximal histidine to iron bonds. *Biochemistry*. 15:388–395.

McMahon, M. T., A. C. deDios, N. Godbout, R. Salzmann, D. D. Laws, H. Le, R. H. Havlin, and E. Oldfield. 1998. An experimental and quantum chemical investigation of CO binding to heme proteins and model systems: a unified model based on $^{13}C$, $^{17}O$, and $^{57}Fe$ nuclear magnetic resonance and $^{57}Fe$ Mössbauer and infrared spectroscopies. *J. Am. Chem. Soc.* 120:4784–4797.

Miller, L., and M. Chance. 1994. Probing conformational changes upon photolysis: FTIR studies of the low temperature liganded and photoproduct states of oxy and carbonmonoxymyoglobin. *J. Am. Chem. Soc.* 116:9662–9669.

Miller, L. M., and M. R. Chance. 1995. Structural and electronic factors that influence oxygen affinities: a spectroscopic comparison of ferrous and cobaltous oxymyoglobin. *Biochemistry*. 34:10170–10179.

Miller, L. M., Q. He, and M. R. Chance. 1998. Examination of the low-temperature photoproduct yield and dioxygen stretching frequency in mutant myoglobins. *Biophys. J.* 74:82*a*. (Abstr.).

Miller, L. M., M. Patel, and M. R. Chance. 1996. Identification of conformational substates in oxymyoglobin through the pH-dependence of the low-temperature photoproduct yield. *J. Am. Chem. Soc.* 118:4511–4517.

Momenteau, M., and C. A. Reed. 1994. Synthetic heme dioxygen complexes. *Chem. Rev.* 94:659–698.

Moore, J. N., 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.

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.

Müller, J. D. 1997. Charakterisierung von Struktur und Funktion taxonomischer Konformations subzustände in CO-ligandiertem Myoglobin. Thesis, Technical University of Munich.

Nakatsuji, H., J. Hasegawa, H. Ueda, and M. Hada. 1996. Ground and excited states of oxyheme: SAC/SAC-CI study. *Chem. Phys. Lett.* 250: 379–386.

Nienhaus, G. U., K. Chu, and K. Jesse. 1998. Structural heterogeneity and ligand binding in carbonmonoxy myoglobin crystals at cryogenic temperatures. *Biochemistry*. 37:6819–6823.

Oldfield, E., K. Guo, J. Augspurger, and C. Dykstra. 1991. A molecular model for the major conformational substates in heme proteins. *J. Am. Chem. Soc.* 113:7537–7541.

Olson, J. S., and G. N. Phillips, Jr. 1996. Kinetic pathways and barriers for ligand-binding to myoglobin. *J. Biol. Chem.* 271:17593–17596.

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

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

Otwinowski, Z., and W. Minor. 1996. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276:307–326.

Park, K. D., K. Guo, F. Abedoun, M. L. Chiu, S. G. Sligar, and E. Oldfield. 1991. Distal and proximal ligand interactions in heme proteins: correlations between C–O and Fe–C vibrational frequencies, oxygen-17 and carbon-13 nuclear magnetic resonance chemical shifts, and oxygen-17 nuclear quadrupole coupling constants in $C^{17}O$ and $^{13}CO$-labeled species. *Biochemistry*. 30:2333–2347.

Pauling, L., and C. D. Coryell. 1936. The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonmonoxyhemoglobin. *Proc. Natl. Acad. Sci. USA*. 22:210–216.

Peng, S.-M., and J. A. Ibers. 1976. Stereochemistry of carbonylmetalloporphyrins. The structure of (pyridine)(carbonyl)(5, 10, 15, 20-tetraphenylporphinato)iron(II). *J. Am. Chem. Soc.* 98:8032–8036.

Phillips, S. E. V. 1980. Structure and refinement of oxymyoglobin at 1.6 Å resolution. *J. Mol. Biol.* 142:531–554.

Phillips, G. N., Jr., and B. M. Pettit. 1995. Structure and dynamics of the water around myoglobin. *Protein Sci.* 4:149–158.

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

Potter, W. T., M. P. Tucker, R. A. Houtchens, and W. S. Caughey. 1987. Oxygen infrared-spectra of oxyhemoglobins and oxymyoglobins: evidence of 2 major liganded $O_2$ structures. *Biochemistry*. 26:4699–4707.

Powers, L., J. L. Sessler, G. L. Woolery, and B. Chance. 1984. CO bond angle changes in photolysis of carboxymyoglobin. *Biochemistry.* 23: 5519–5523.

Quillin, M. L., R. M. Arduini, 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 and steric effects in heme proteins and models. *J. Am. Chem. Soc.* 116:162–176.

Rovira, C., K. Kunc, J. Hutter, P. Ballone, and M. Parrinello. 1997. Equilibrium geometries and electronic structure of iron-porphyrin complexes: a density functional study. *J. Phys. Chem. A*. 101: 8914–8925.

Sage, J. T. 1997. Myoglobin and CO: structure, energetics and disorder. *J. Bioinorgan. Chem.* 2:537–543.

Sage, J. T., and W. Jee. 1997. Structural characterization of the myoglobin active site using infrared crystallography. *J. Mol. Biol.* 274:21–26.

Sauer, O., A. Schmidt, and C. Kratky. 1997. Freeze-trapping isomorphous xenon derivatives of protein crystals. *J. Appl. Crystallogr.* 30:476–486.

Scheidt, R. W., K. J. Haller, M. Fons, T. Mashiko, and C. A. Reed. 1981. A (carbonmonoxy)heme complex with a weak proximal bond. Molecular stereochemistry of carbonyl(deuteroporphinato) (tetrahydrofuran)-iron(II). *Biochemistry.* 20:3653–3657.

Schlichting, I., J. Berendzen, G. N. Phillips, Jr., and R. M. Sweet. 1994. Crystal structure of photolyzed carbonmonoxymyoglobin. *Nature*. 371: 808–812.

Schomaker, V., and K. N. Trueblood. 1968. On the rigid-body motion of molecules in crystals. *Acta Crystallogr. B.* 24:63–76.

Sheldrick, G. M., and T. R. Schneider. 1997. SHELX: high-resolution refinement. *Methods Enzymol.* 277:319–343.

Slebodnick, C., M. L. Duval, and J. A. Ibers. 1996a. Structural characterization of $OC_3OPor$ capped porphyrins: $H_2OC_3OPor$, $Fe(OC_3OPor)(Cl)$, $Fe(OC_3OPor)(CO)(1\text{-MeIm})$, and $Fe(OC_3OPor)(CO)(1, 2\text{-}Me_2Im)$. *Inorg. Chem.* 35:3607–3613.

Slebodnick, C., J. C. Fettinger, H. B. Peterson, and J. A. Ibers. 1996b. Structural characterization of 5 sterically protected porphyrins. *J. Am. Chem. Soc.* 118:3216–3224.

Spek, A. L. 1992. PLATON-92. University of Utrecht, The Netherlands.

Spiro, T. G., and P. M. Kozlowski. 1998. Discordant results on FeCO deformability in heme proteins reconciled by density functional theory. *J. Am. Chem. Soc.* 120:4524–4525.

Springer, B. A., S. G. Sligar, J. S. Olson, and G. N. Phillips, Jr. 1994. Mechanisms of ligand recognition in myoglobin. *Chem. Rev.* 94: 699–712.

Takano, T. 1977a. Structure of myoglobin refined at 2.0 Å resolution. I. Crystallographic refinement of metmyoglobin from sperm whale. *J. Mol. Biol.* 110:537–568.

Takano, T. 1977b. Structure of myoglobin refined at 2.0 Å resolution. I. Structure of deoxymyoglobin from sperm whale. *J. Mol. Biol.* 110: 569–577.

Terwilliger, T. C., and J. Berendzen. 1996a. Bayesian weighting for macromolecular crystallographic refinement. *Acta Crystallogr. D.* 52: 743–748.

Terwilliger, T. C., and J. Berendzen. 1996b. Bayesian difference refinement. *Acta Crystallogr. D.* 52:1004–1011.

Tetreau, C., D. Lavalette, M. Momenteau, J. Fischer, and R. Weiss. 1994. Structure-reactivity relationship in oxygen and carbon monoxide binding with some heme models. *J. Am. Chem. Soc.* 116:11840–11848.

Tilton, R. F., I. D. Kuntz, and G. A. Petsko. 1984. Cavities in proteins: structure of a metmyoglobin-xenon complex solved to 1.9 Å. *Biochemistry.* 23:2849–2857.

Tolman, J. R., J. M. Flanaga, M. A. Kennedy, and J. H. Pretegard. 1997. NMR evidence for slow collective motions in cyanometmyoglobin. *Nat. Struct. Biol.* 4:292–297.

Tsubaki, M., and N. T. Yu. 1981. Resonance Raman investigation of dioxygen bonding in oxycobalt-myoglobin and oxycobalt-hemoglobin: structural implication of splittings of the bound O-O stretching vibration. *Proc. Natl. Acad. Sci. USA*. 78:3581–3585.

Unno, M., J. F. Christian, J. S. Olson, J. T. Sage, and P. M. Champion. 1998. Evidence for hydrogen bonding effects in the iron ligand vibrations of carbonmonoxy myoglobin. *J. Am. Chem. Soc.* 120:2670–2671.

van Aalten, D. M. F., D. A. Conn, B. L. de Groot, H. J. C. Berendsen, J. B. C Findlay, and A. Amadei. 1997. Protein dynamics derived from clusters of crystal structures. *Biophys. J.* 73:2891–2896.

Weiss, J. J. 1964. Nature of the iron-oxygen bond in oxyhaemoglobin. *Nature*. 202:83–84.

Weissbluth, M., and J. E. Maling. 1967. Interpretation of quadrupole splittings and isomer shifts in hemoglobin. *J. Chem. Phys.* 47: 4166–4172.

Wilbur, D. J., and A. J. Allerhand. 1977. Titration behavior of individual tyrosine residues of myoglobins from sperm whale, horse, and red kangaroo: application of natural abundance C-13 nuclear magnetic-resonance spectroscopy. *J. Biol. Chem.* 252:4968–4975.

Yang, F., and G. N. Phillips, Jr. 1996. Crystal structures of CO-, deoxy- and met-myoglobins at various pH values. *J. Mol. Biol.* 256:762–774.

Young, R. D., H. Frauenfelder, J. B. Johnson, D. C. Lamb, G. U. Nienhaus, R. Phillip, and R. Scholl. 1991. Time-dependence and temperature-dependence of large-scale conformational transition in myoglobin. *Chem. Phys.* 158:317–327.