Connection between the Taxonomic Substates and Protonation of Histidines 64 and 97 in Carbonmonoxy Myoglobin

Joachim D. Müller,* Benjamin H. McMahon,* Ellen Y. T. Chien,# Stephen G. Sligar,#§ and G. Ulrich Nienhaus*¶
Departments of *Physics, #Biochemistry and §Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801-3080 USA, and ¶Department of Biophysics, University of Ulm, D-89069 Ulm, Germany

ABSTRACT Infrared spectra of heme-bound CO in sperm whale carbonmonoxy myoglobin and two mutants (H64L and H97F) were studied in the pH range from 4.2 to 9.5. Comparison of the native protein with the mutants shows that the observed pH effects can be traced to protonations of two histidine residues, H64 and H97, near the active site. Their imidazole sidechains experience simple, uncoupled Henderson–Hasselbalch type protonations, giving rise to four different protonation states. Because two of the protonation states are linked by a pH-independent equilibrium, the overall pH dependence of the spectra is described by a linear combination of three independent components. Global analysis, based on singular value decomposition and matrix least-squares algorithms enabled us to extract the pK values of the two histidines and the three basis spectra of the protonating species. The basis spectra were decomposed into the taxonomic substates A_0, A_1, and A_3, previously introduced in a heuristic way to analyze CO stretch spectra in heme proteins at fixed pH (see for instance, Johnson et al., 1996, *Biophys. J.* 71:1563–1573). Moreover, an additional, weakly populated substate, called A_x, was identified. Protonation of H97 gives rise to a blue shift of the individual infrared lines by about 2 cm^{-1}, so that the *A* substates actually appear in pairs, such as A_0 and A_0^+. The blue shift can be explained by reduced backbonding from the heme iron to the CO. Protonation of the distal histidine, H64, leads to a change of the infrared absorption from the A_1 or A_3 substate lines to A_0. This behavior can be explained by a conformational change upon protonation that moves the imidazole sidechain of H64 away from the CO into the high-dielectric solvent environment, which avoids the energetically unfavorable situation of an uncompensated electric charge in the apolar, low-dielectric protein interior. Our results suggest that protonation reactions serve as an important mechanism to create taxonomic substates in proteins.

INTRODUCTION

For many proteins, transfer of protons to or from amino acid sidechains at the active site is key to their function. The free energy change associated with protonation of an amino acid sidechain can be significant in comparison with the overall stabilization energy of the folded protein, and therefore, conformational changes of the protein are frequently initiated by protonation of specific residues. Here we study the influence of protonations near the active site on the structure and function of sperm whale carbonmonoxy myoglobin (swMbCO). The infrared (IR) spectrum of MbCO shows multiple CO stretch bands, denoted A_0, A_1, and A_3, with vibrational frequencies $\nu(A_0) \approx 1967\ cm^{-1}$, $\nu(A_1) \approx 1945\ cm^{-1}$, and $\nu(A_3) \approx 1933\ cm^{-1}$, which have been assigned to three different conformations, called taxonomic substates (Frauenfelder et al., 1991; Nienhaus and Young, 1996). These bands offer a sensitive tool for studying protonations in myoglobin because of the large spectral changes of the CO absorption with pH. Understanding of how the structural differences among the *A* substates influence their specific properties may shed light on structure–function relations in proteins in general.

The *A* substates have been studied extensively with a wide variety of experimental techniques, such as IR spectroscopy (Braunstein et al., 1993; Li et al., 1994), nuclear magnetic resonance (Park et al., 1991), resonance Raman scattering (Ray et al., 1994), and x-ray crystallography (Yang and Phillips, 1996). It was shown that the equilibrium populations of the *A* substates depend on external conditions such as temperature, pH, solvent composition, and pressure (Fuchsman and Appleby, 1979; Shimada and Caughey, 1982; Ansari et al., 1987; Hong et al., 1990; Iben et al., 1989; Morikis et al., 1989; Frauenfelder et al., 1990; Zhu et al., 1992; Mourant et al., 1993; Johnson et al., 1996). The dynamics of interconversion between them has also been measured over a wide range of temperatures. At ≈180 K, conformational transitions between the *A* substates require hours or days to occur, whereas at room temperature, these transitions occur in nanoseconds or microseconds (Young et al. 1991; Johnson et al., 1996; Tian et al., 1996).

Structural models have been suggested to account for the pH dependence of the IR spectra (Fuchsman and Appleby, 1979; Shimada and Caughey, 1982). More recent Fourier transform infrared spectroscopy (FTIR) studies involving distal pocket mutants and theoretical studies have emphasized the importance of electrostatic interactions between the bound CO and the residues lining the heme pocket (Braunstein et al., 1993; Li et al., 1994; Kushkuley and Stavrov, 1996, 1997). In native MbCO, histidine 64 (H64)

Received for publication 13 August 1998 and in final form 12 April 1999.

Address correspondence to G. U. Nienhaus, Department of Biophysics, University of Ulm, D-89069 Ulm, Germany. Tel.: 49-731-502-3050; Fax: 49-731-502-3059; E-mail: uli@uiuc.edu.

B. H. McMahon's present address is Center for Nonlinear Studies, MS-B258, Los Alamos National Laboratory, Los Alamos, NM 87545.

E. Y. T. Chien's present address is Alliance Pharmaceutical Corp., 3040 Science Park Road, San Diego, CA 92121.

0006-3495/99/08/1036/16 $2.00

was shown to govern a variety of properties of the active site. Direct evidence of a conformational change accompanying protonation of H64 was provided by pH-dependent crystal structure analyses (Yang and Phillips, 1996).

The pH dependence of the ligand binding kinetics of MbCO was first noticed by Doster et al. (1982). Ansari et al. (1987) explained this effect by the different ligand binding properties of the individual *A* substates in conjunction with pH-induced changes in their relative populations. Champion and coworkers (Morikis et al., 1989; Tian et al., 1993, 1996) proposed a model to explain the pH dependence of the *A* substate populations, which involves an unprotonated or protonated H64 imidazole sidechain in either of two different protein conformations, termed open and closed, so that the model contains altogether four states. An appealing feature of their model is the explicit consideration of conformational changes upon protonation. However, the assumption implicit in the Champion model that protonation of the H64 sidechain only affects the population ratio between the open and closed conformation, but not the spectroscopic and functional properties of the two conformations, appears unrealistic in view of the experimental evidence that structural changes in the vicinity of the bound CO ligand markedly affect its bond order and thus the stretch frequencies (Braunstein et al., 1993; Li et al., 1994).

To clarify the interplay between conformational changes and protonation, we have revisited the pH dependence of the infrared spectra of native swMbCO. We have also used two mutant myoglobins, H97F and H64L, which enabled us to unambiguously assign the pH-dependent spectral changes to the protonation of two histidines, H97 and H64. A global data analysis method was used to fit the IR spectra without assuming a functional shape of the IR bands associated with the different protonation states and without the need to reach the low-pH endpoint of the titration. The analysis yields the basis spectra of the different conformational substates, characterized by the protonation states of H64 and H97 and conformational differences induced by protonation. Between pH 4.2 and 9.5, the pH dependence of the CO stretch bands can be modeled with the conceptually simplest model, two energetically uncoupled Henderson–Hasselbalch-type protonation reactions.

MATERIALS AND METHODS

Biochemical procedures

Native sperm whale myoglobin was purchased from Sigma Chemical Company (St. Louis, MO) as lyophilized powder and used without further purification. Mutations at positions 64 and 97 were generated using site-directed mutagenesis procedures described by Springer et al. (1989). Oligonucleotides were synthesized by the University of Illinois Biotechnology Center and purified by reverse-phase high performance liquid chromatography (Springer and Sligar, 1987). Mutant sequences were confirmed using the Sequenase double-stranded DNA sequencing reagents and protocol from United States Biochemicals (Cleveland, OH). Mutant sperm whale myoglobins were expressed in *Escherischia coli* using the synthetic gene of Springer and Sligar (1987). For the mutant at position 97, the amino acid at position 122 was corrected to aspartic acid (Phillips et al., 1990). DNA manipulations were essentially as described by Sambrook et al. (1989), using reagents from New England Biolabs (Beverly, MA) or Gibco-Bethesda Research Labs (Grand Island, NY). Radiochemicals were purchased from Amersham (Arlington Heights, IL). All other reagents were purchased from Sigma (St. Louis, MO).

The mutant proteins were purified as described by Springer and Sligar (1987), except for the following modifications. The cell lysate was adjusted to pH 5.9 with NaH_2PO_4. The solution was stirred for 30 min at 4°C, and precipitates were removed by centrifugation at 30,000 rpm for 30 min. The protein solution was diluted four-fold with water and then loaded on a CM-52 cellulose cation exchange column (Whatman, Hillsboro, OH) equilibrated with 10 mM Na_2PO_4, pH 6.0, and eluted with a linear pH gradient: 10 mM Na_2PO_4, pH 6.0, to 30 mM Na_2HPO_4. This procedure resulted in a homogeneous solution of Mb as monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis. Finally, the protein solutions were dialyzed against distilled water and lyophilized for subsequent use in the spectroscopic experiments.

Sample preparation

For measurements of the pH dependence of infrared spectra at room temperature (292 $\pm$ 2 K), lyophilized metmyoglobin powder was dissolved in water to yield a stock solution of 1 to 20 mM concentration. A fraction of the stock was stirred under a CO atmosphere and afterwards reduced with sodium dithionite solution under anaerobic conditions to give MbCO. Immediately before each measurement, a few μL of protein solution were mixed with a few μL of potassium phosphate (or glycine/NaOH for pH $>$ 8) buffer to yield the desired concentration of protein and ($\approx$300 mM) buffer. Approximately half of the sample was loaded into a micro flow cell with a path length of 100 μm (model 155, Wilmad, Buena, NJ), and IR spectra were acquired. Measurements were carried out at approximately 30 different pH values for each of the three proteins studied.

Immediately after taking the spectrum, the pH of the remaining half of the sample was determined using a calibrated combination electrode (model MI-413, Microelectrodes Inc., Londonderry, NH). Measurement of the pH in the concentrated protein solutions was complicated by adhesion of protein on the electrode surfaces, leading to continuous changes of its potential. To minimize these errors, the electrode was cleaned frequently, checked with standard pH solutions and recalibrated when necessary. We estimate an error of $\pm$0.1 units in the pH determination.

For experiments at cryogenic temperatures, MbCO was dissolved in 75% glycerol/25% potassium phosphate (or glycine/NaOH for pH $>$ 8) buffer solution (v/v). The protein solution was placed between two sapphire windows separated by a 75-μm thick mylar washer, which were screwed tightly onto a block of oxygen-free high-conductivity copper and mounted on the cold-finger of a closed-cycle helium refrigerator for IR data acquisition. The temperature was measured with a silicon temperature sensing diode and regulated with a digital temperature controller (Lake Shore Cryotronics, model DRC93C, Westerville, OH).

Spectroscopic methods

Transmittance spectra were collected on a Sirius 100 FTIR spectrometer (Mattson, Madison, WI) between 1800 and 2300 cm^{-1} with a resolution of 2 cm^{-1} using a photovoltaic InSb detector. Transmittance spectra of MbCO were referenced against met Mb samples to obtain absorbance difference spectra. Although great care was taken to prepare the reference met Mb samples under conditions identical to those of the MbCO samples, slight differences in sample composition and optical geometry made it still necessary to apply quadratic or cubic baseline corrections to the difference spectra. Typically, baselines of low-pH spectra needed most correction.

To assess the rebinding barriers in the low-temperature experiments, temperature derivative spectroscopy (TDS) was used. TDS is an experimental protocol designed to measure thermally activated rate processes with distributed barriers. The method has been described earlier in detail (Berendzen and Braunstein, 1990; Mourant et al., 1993; Nienhaus et al.,

1994), and therefore, we give only a brief summary here. The TDS measurement was started after completely photodissociating the MbCO sample with light from an argon ion laser at a temperature of 12 K, where rebinding is extremely slow on the experimental time scale. Subsequently, while ramping the sample temperature T_R linearly up in time, FTIR transmittance spectra were taken continuously. FTIR absorbance difference spectra, $\Delta\mathcal{A}(\nu, T_R)$, were calculated from transmittance spectra at successive temperatures,

$$\Delta\mathcal{A}(\nu, T_R) = \log \mathcal{J}(\nu, T_R - \tfrac{1}{2}\,\mathrm{K}) - \log \mathcal{J}(\nu, T_R + \tfrac{1}{2}\,\mathrm{K}). \quad (1)$$

The change in spectral area, $\Delta\mathcal{A}(\nu, T_R)$, is proportional to the population, ΔN, that rebinds between successive scans at temperature T_R. The TDS signal, $\Delta\mathcal{A}(\nu, T_R)$, approximates the distribution of enthalpy barriers for rebinding, $g(H)$. Given a first-order reaction and the Arrhenius relation with constant preexponential factor, the temperature T_R can be converted into an activation enthalpy H. The relation between T_R and H is linear to a good approximation, so the temperature axis can be treated as an enthalpy axis.

Data analysis

In our model, the IR spectrum at any given pH is a superposition of IR spectra associated with four different protein states connected by protonation reactions, together with random noise. We arrange the data in a matrix, **D**, of m different wavenumbers and n different pH values, which can be decomposed as

$$\mathbf{D} = \mathbf{S}\mathbf{F}^{\mathrm{T}}(\mathbf{p}) + \mathbf{E}, \quad (2)$$

where the rows of matrix **S** contain the absorbance spectra of the individual protein states that are accessible by varying the pH; they will be referred to as basis spectra or physical spectra. $\mathbf{F}^{\mathrm{T}}$ denotes the transpose of matrix **F**, the rows of which contain the pH-dependent amplitudes of the physical spectra, which depend on a set of fit parameters **p**, and the error matrix **E** contains the random errors. According to Beer's law, the coefficients or amplitudes of the basis spectra in matrix **F** are proportional to the concentration of the individual protein states.

To decompose the data matrix **D** according to Eq. 2, a thermodynamic model is needed that connects the pH with the population of protein states. In general, such a model depends on a set of control parameters **p** such as pK values of protonatable groups and other parameters that specify an equilibrium state of the sample. The model completely determines the population matrix **F**(**p**) as a function of the parameter set **p**. For the case of a single-site protonation reaction,

$$AH \overset{\mathrm{pK}}{\rightleftharpoons} A^- + H^+, \quad (3)$$

the normalized population of the protonated state, AH, is given by the Henderson–Hasselbalch equation,

$$f(\mathrm{pH}) = \frac{[AH]}{[A^-] + [AH]} = \frac{1}{1 + 10^{\mathrm{pH}-\mathrm{pK}}}, \quad (4)$$

which depends on only one parameter, the pK of the conjugate acid-base pair (AH/A^-).

Matrix least-squares fit

The decomposition of the matrix **D** (Eq. 2) allows the determination of the physical basis spectra **S** associated with the model matrix **F**. The best parameter set **p** is determined by minimizing the error function,

$$\phi = \|\mathbf{E}\| = \|\mathbf{D} - \mathbf{S}\mathbf{F}^{\mathrm{T}}(\mathbf{p})\|. \quad (5)$$

For any fixed model matrix **F**(**p**), ϕ is minimized by choosing $\mathbf{S} = \mathbf{D}(\mathbf{F}^{\mathrm{T}}(\mathbf{p}))^{+}$, where $(\mathbf{F}^{\mathrm{T}})^{+}$ is the pseudo-inverse of the matrix $\mathbf{F}^{\mathrm{T}}$ (Strang, 1980). The optimal parameter set $\hat{\mathbf{p}}$ can then be determined using a nonlinear least-squares fit program, and the quality of the model can be judged by the residual error,

$$\hat{\phi} = \|\mathbf{D} - \mathbf{D}(\mathbf{F}^{\mathrm{T}}(\hat{\mathbf{p}}))^{+}\mathbf{F}^{\mathrm{T}}(\hat{\mathbf{p}})\|. \quad (6)$$

This method had been termed the matrix least-squares (MLS) method by Hendler and Shrager (1994). MLS decouples the linear least-squares problem, which can be solved exactly, from the complex and computer-intensive nonlinear problem.

Singular value decomposition-based MLS

The MLS method alone is still quite computer intensive; each update of the parameter set **p** by the nonlinear least-squares fit requires the computation of an adjusted matrix **S** as part of the MLS algorithm, which involves all elements of the experimental data matrix **D** (in our case approximately 100 spectral data points × 30 pH values). A substantial reduction in the number of computations can be achieved by combining MLS with singular value decomposition (SVD), an algorithm that separates the noise from the information in the data (Shrager, 1986). The linearly independent signal components above the experimental noise level contain all the information present in the experimental data matrix **D**. The number of these independent linear components defines the effective rank of the data matrix **D**. The SVD algorithm is well suited to determine the effective rank from the data (Henry and Hofrichter, 1992; Press et al., 1992).

SVD decomposes the ($m \times n$) matrix **D** ($m \geq n$) into

$$\mathbf{D} = \tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{V}}^{\mathrm{T}}. \quad (7)$$

The columns of the ($m \times n$) matrix $\tilde{\mathbf{U}}$ form an orthonormal set of basis vectors, the ($n \times n$)-matrix $\tilde{\mathbf{V}}$ is orthonormal and $\tilde{\mathbf{S}}$ is an ($n \times n$) diagonal matrix with non-negative elements $\tilde{s}_k (k \in \{1, \ldots, n\})$. The diagonal elements $\tilde{s}_k$ are the singular values of the matrix **D** and are arranged in descending order ($\tilde{s}_1 \geq \tilde{s}_2 \geq \cdots \geq \tilde{s}_n \geq 0$). The first r singular values of $\tilde{\mathbf{S}}$ together with the first r columns of the matrices $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$, which will be referred to as $\tilde{\mathbf{S}}_r$, $\tilde{\mathbf{U}}_r$, and $\tilde{\mathbf{V}}_r$, constitute the best approximation of the matrix **D** of all matrices with rank r. In other words, the matrix $\mathbf{D}_r \equiv \tilde{\mathbf{U}}_r\tilde{\mathbf{S}}_r\tilde{\mathbf{V}}_r^{\mathrm{T}}$ has the smallest norm of $\|\mathbf{D} - \mathbf{D}_r\|$ among all matrices of rank r. In the case of the Frobenius norm, it is given by (Golub and Loan, 1983),

$$\|\mathbf{D} - \mathbf{D}_r\| = \sqrt{\tilde{s}_{r+1}^2 + \cdots + \tilde{s}_n^2}. \quad (8)$$

The deviation of the matrix $\mathbf{D}_r$ from the experimental data matrix **D** decreases continuously with increasing r; the error is determined by the remaining singular values according to Eq. 8. The first singular values decrease rapidly until the noise level is reached. Afterwards only a weak decrease of the singular values is observed because the noise is purely statistical; and thus, each noise component is independent and of almost equal weight (Wigner 1967; Trotter, 1984). The number of singular values above the noise level determines the effective rank of the matrix **D**.

Consequently, the experimental data matrix **D** with effective rank r is well approximated by the first r columns of the SVD matrices,

$$\mathbf{D} = \tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{V}}^{\mathrm{T}} \approx \tilde{\mathbf{U}}_r\tilde{\mathbf{S}}_r\tilde{\mathbf{V}}_r^{\mathrm{T}}. \quad (9)$$

The best parameter set $\hat{\mathbf{p}}$ of the model is then determined by minimizing the corresponding error function

$$\tilde{\phi} = \|\tilde{\mathbf{S}}_r(\tilde{\mathbf{V}}_r^{\mathrm{T}} - \tilde{\mathbf{V}}_r^{\mathrm{T}}(\mathbf{F}^{\mathrm{T}}(\hat{\mathbf{p}}))^{+}\mathbf{F}^{\mathrm{T}}(\mathbf{p}))\|. \quad (10)$$

Weighing of the matrix $\tilde{\mathbf{V}}_r$ with the singular values of $\tilde{\mathbf{S}}_r$ minimizes the differences between MLS and SVD-based MLS (Shrager, 1986).

Combining the MLS algorithm with SVD leads to an enormous reduction in the number of elements that need to be adjusted in the fit. In our case, instead of computing about 3,000 elements for each change of the

parameter set **p**, SVD-based MLS only updates $30 \times r$ elements, i.e., 90 in the case of two protonating residues, to determine the error function $\tilde{\phi}$.

RESULTS

pH dependence of the CO stretch spectra

The IR spectra of heme-bound CO in swMbCO are shown in Fig. 1 *a* in the range from pH 4.2 to 9.5. Above pH 7, the spectra are dominated by the A_1 and A_3 substates at 1945 and 1933 cm^{-1}. As the sample pH decreases, the A_1 and A_3 lines first shift by $\approx$2 cm^{-1} to the blue; subsequently, a band at 1966 cm^{-1}, A_0, grows at the expense of the other two. Below pH 4.2, this band broadens substantially, indicating the presence of acid denaturation of the protein sample. Because denaturation is cooperative, it has a sharp onset in pH and so does not influence data above pH 4.2. Here we limit ourselves to the pH range in which the native protein is observed and treat the denaturation transition separately (Müller et al., in preparation).

In the pH range covered by our study, the pH-dependent changes are caused by protonation of two histidine residues close to the active site (see below); H64, which is in the distal pocket and H97, which is stacked against the heme on the proximal side. To investigate the involvement of these two amino acids, we explored the pH dependence of the CO stretch bands of two mutants, H64L and H97F, each of which has one of the two histidines replaced by a nonprotonating amino acid. The heme group is covalently bound to the protein through H93, which is known not to protonate above pH 2.6 in MbCO (Han et al., 1990; Sage et al., 1991).

Replacement of H97 with a phenylalanine residue in the H97F mutant results in the IR spectra shown in Fig. 1 *b*. They are similar to those of the native protein, including the exchange from A_1 and A_3 to A_0 as the pH is lowered. The shifts of the A_1 and A_3 bands are absent, however, and a crisp isosbestic point exists, indicating that only a single protonation reaction influences the IR spectrum. Closer inspection also shows that the A_0 band in this mutant is

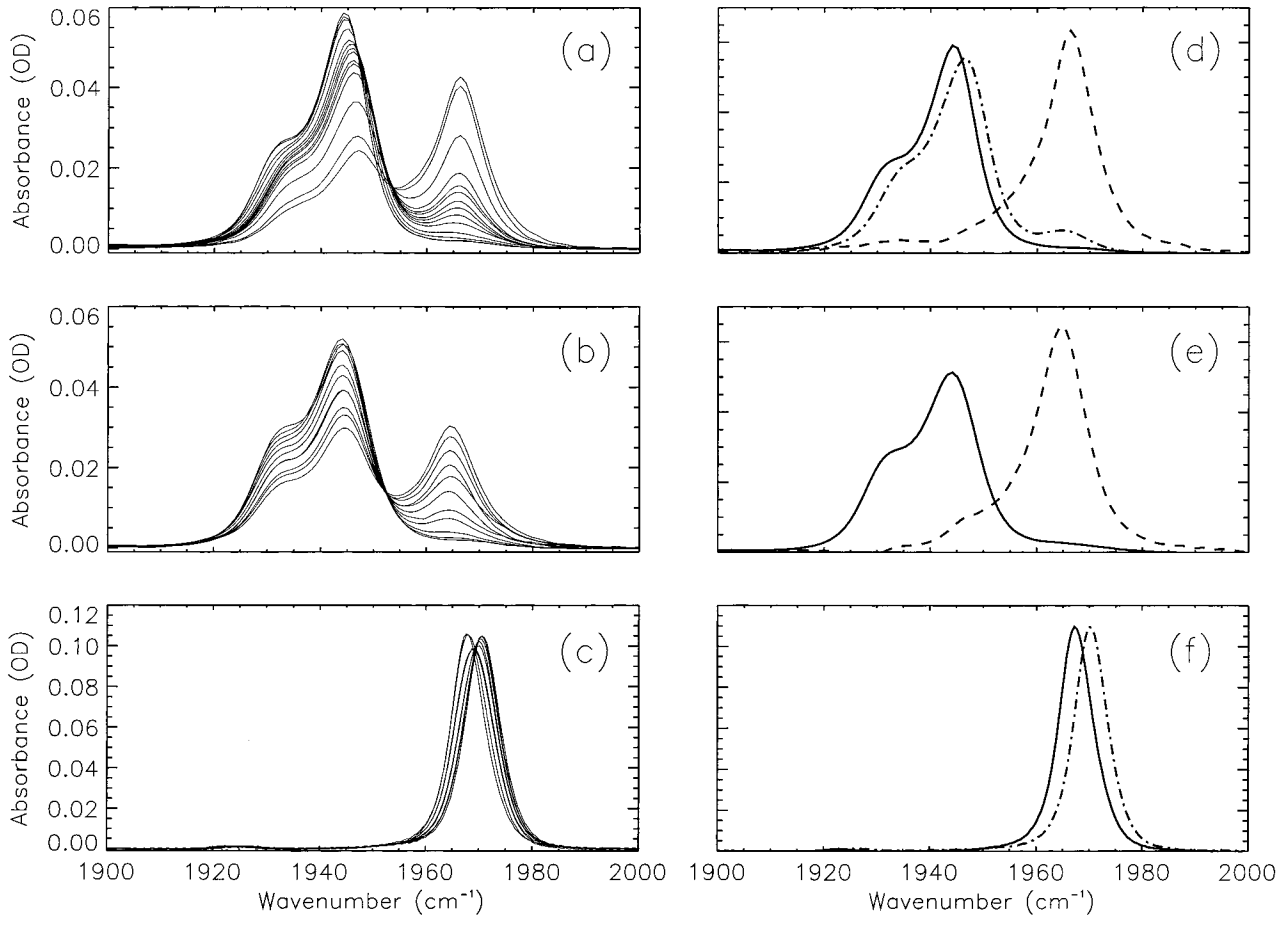

FIGURE 1 *Left column:* CO stretch infrared spectra in the pH range from 4.2 to 9.5 of (*a*) swMbCO, (*b*) proximal mutant H97F, and (*c*) distal mutant H64L, baseline corrected and normalized to an integrated area of 1 OD cm^{-1}. *Right column:* Basis spectra of the different protonation states as determined by least squares fits of the spectra to a model described in the text; (*d*) swMbCO, with neutral form S_{00} (——), mixed protonation state $S_{[01]}$ (– · – ·), and doubly protonated protein S_{11} (– – –); (*e*) H97F, with neutral H64, S_{0x}^{H97F} (——) and protonated H64, S_{1x}^{H97F} (– – –); (*f*) H64L, with neutral H97, S_{x0}^{H64L} (——), and protonated H97 S_{x1}^{H64L} (– · – ·).

red-shifted by about 2 cm^{-1} compared to the native protein, as expected from the unshifted form of this band.

Replacement of H64 by the aliphatic amino acid leucine in the H64L mutant causes a single band to appear in the wavenumber region of A_0 at all pH values (Fig. 1 *c*). This effect is also observed for other substitutions of H64 (Braunstein et al., 1993; Li et al., 1994). The band shifts to the blue upon lowering the pH, and an isosbestic point is again observed, implying that a single protonation reaction is responsible for the pH dependence of the IR band. The band is significantly narrower than the A_0 band of the native protein.

Several conclusions can be drawn directly from the spectra. First, the 2-cm^{-1} blue shift of the individual bands is absent in H97F but occurs in the mutant H64L and the native protein, both containing the residue H97. Thus, protonation of H97 results in a blue shift of the IR spectra by $\approx$2 cm^{-1}. Second, the splitting of the CO stretch spectrum into multiple bands is absent in H64L, and thus H64 is responsible for the conversion of the A_1/A_3 bands to the blue-shifted A_0 band upon protonation. Third, comparison of the mutants with the native protein shows that the protonation effects are largely additive.

Quantitative model

In this section, we will present a quantitative model that explains the entire set of pH-dependent spectra simultaneously without assuming a spectral form of each species. From our qualitative discussion in the previous subsection, it is evident that the model has to include four different states arising from the four combinations of two protonation states, protonated and unprotonated, regarding the two histidine residues, H64 and H97. The three (or four) heuristically introduced *A* substate bands (Ansari et al., 1987) are not suitable for this analysis. Instead, we denote the protonation state spectra by S_{00}, S_{01}, S_{10}, and S_{11}, where the indices are defined so that 0 represents an unprotonated and 1 represents a protonated residue, with the first index referring to H64 and the second one to H97. These states are represented by entire CO stretch spectra of the particular protonation species S_{ij}; the observed spectrum is represented by a superposition of the four spectra with the appropriate weight factors. We relate the S_{ij} to the well-known *A* substates below.

The model is shown schematically in Fig. 2 *a*. We assume that the protonation reactions of the two histidine sidechains are of simple Henderson–Hasselbalch type. Furthermore, although the two protonations simultaneously affect the IR spectra, it is assumed that the pK of H64 does not depend on the protonation state of H97 and vice versa, as indicated by the identical labels on opposing sides of the square in Fig. 2 *a*.

The degree of protonation of a sidechain, *f*(pH), is given by Eq. 4 and will be denoted with f_{64} and f_{97} for H64 and H97, respectively. The relative populations, f_{00}, f_{01}, f_{10}, and

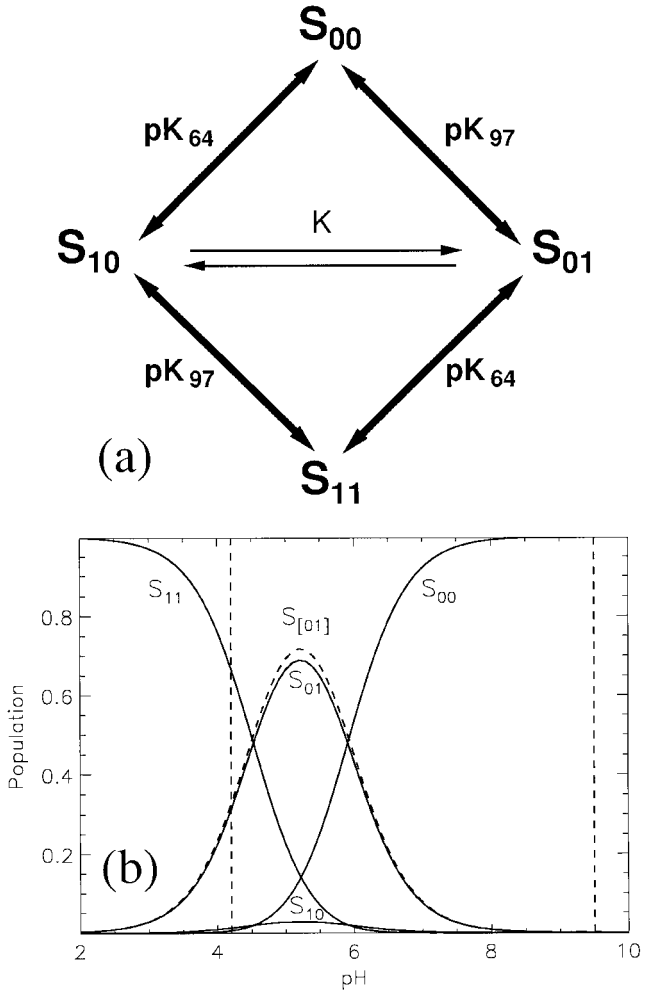

FIGURE 2 (*a*) Schematic representation of the model based on the independent titration of two histidine residues, H64 and H97, in swMbCO. The protein states associated with each protonation state are denoted by S_{00}, S_{01}, S_{10}, and S_{11}. The indices on S_{ij} are defined such that a 0 and a 1 represent an unprotonated and a protonated residue, respectively, and the first index refers to H64, and the second one to H97. Adjacent protein states are connected by simple Henderson–Hasselbalch type protonation reactions characterized by the pKs of either H97 or H64. The mixed protonation states S_{10} and S_{01} are in a pH-independent equilibrium and appear in the titration as a single species, which we denote as $S_{[01]}$. (*b*) Populations f_{ij} associated with each of the protonation states S_{ij}, shown as functions of pH for pK_{64} = 4.53 and pK_{97} = 5.91, as recovered by a least-squares fit to the data. The singly protonated species f_{01} and f_{10} are added to yield the population $f_{[01]}$ of the mixed species $S_{[01]}$ (*dashed line*). The vertical dashed lines at pH 4.2 and 9.5 show the experimental pH window over which data were collected and globally fitted according to the model in (*a*). Below pH 4.2, acid-induced denaturation of the protein occurs with a sharp onset in pH; thus only the populations above pH 4.2 are physically meaningful.

f_{11}, of the four protonation states S_{00}, S_{01}, S_{10}, and S_{11} are then given by

$$f_{00} = (1 - f_{64})(1 - f_{97}), \tag{11}$$

$$f_{10} = f_{64}(1 - f_{97}), \tag{12}$$

$$f_{01} = (1 - f_{64})f_{97}, \tag{13}$$

$$f_{11} = f_{64}f_{97}. \tag{14}$$

The scheme of Fig. 2 *a* leads to population ratios for each protonation state shown in Fig. 2 *b* assuming a $pK_{64} = 4.53$ and $pK_{97} = 5.91$.

Only three of the four states in our model are linearly independent because a pH-independent equilibrium exists between f_{01} and f_{10},

$$f_{01} \overset{K}{\rightleftharpoons} f_{10}, \tag{15}$$

with the equilibrium coefficient K given by

$$\frac{f_{01}}{f_{10}} = 10^{pK_{97} - pK_{64}}. \tag{16}$$

Intuitively, this behavior is expected because no protons are exchanged with the solvent in this reaction. Titration experiments cannot distinguish between the two singly protonated states S_{01} and S_{10}, so we denote the combined spectrum as $S_{[01]}$, which is a linear combination of S_{01} and S_{10},

$$S_{[01]} = (1 - f_M)S_{10} + f_M S_{01}, \tag{17}$$

with a pH-independent mixing coefficient f_M given by

$$f_M = \frac{1}{1 + 10^{pK_{97} - pK_{64}}}. \tag{18}$$

The fractional population of the mixed species is simply the sum of the two singly protonated populations,

$$f_{[01]} = f_{01} + f_{10}. \tag{19}$$

The pH dependence of $f_{[01]}$ is also included in Fig. 2 *b*. Thus, the overall spectrum at a particular pH is obtained as a superposition of the three basis spectra,

$$S = f_{00}S_{00} + f_{[01]}S_{[01]} + f_{11}S_{11}. \tag{20}$$

The model for the mutants has only to consider a single protonation reaction, with

$$S_{x0}^{H64L} + H^+ \overset{pK_{97}^{H64L}}{\rightleftharpoons} S_{x1}^{H64L} \tag{21}$$

for H64L, and

$$S_{0x}^{H97F} + H^+ \overset{pK_{64}^{H97F}}{\rightleftharpoons} S_{1x}^{H97F} \tag{22}$$

for H97F. Here, the indices referring to the protonation states of the mutants are analogous to that of the native protein, except that an *x* has been introduced to denote the mutated residue.

Global analysis of the IR spectra

The SVD-based MLS algorithm described under Materials and Methods, was used to globally fit an entire set of pH-dependent spectra according to the model in the previous subsection. The first twelve singular values of the swMbCO spectra, $\tilde{s}_i$, are shown in Fig. 3 *a*. Only the first three singular values are larger than 0.01. Moreover, for the first three singular values, a strong decrease is observed, whereas from the fourth one on, the decrease is much weaker, suggesting that only the first three components represent real features in the data, and the higher ones contain only noise. Consequently, we conclude that the effective rank of the data matrix equals three, consistent with the three linearly independent basis spectra S_{00}, $S_{[01]}$, and S_{11} from two protonations, as required by our model. Fig. 3 *b* shows the first four basis vectors of matrix $\tilde{\mathbf{U}}$, scaled by their respective singular values. Again, only the first three spectra of matrix $\tilde{\mathbf{U}}$ contribute to the signal of the IR absorption bands.

The first four amplitude vectors of matrix $\tilde{\mathbf{V}}$ are plotted in Fig. 4. The first three vectors (panels *a*–*c*) show a systematic dependence on pH, whereas the fourth (panel *d*) and higher components no longer show a correlation with pH. Consistent with the singular value analysis, this observation also supports the presence of three linearly independent components. Indeed, the determination of the effective rank

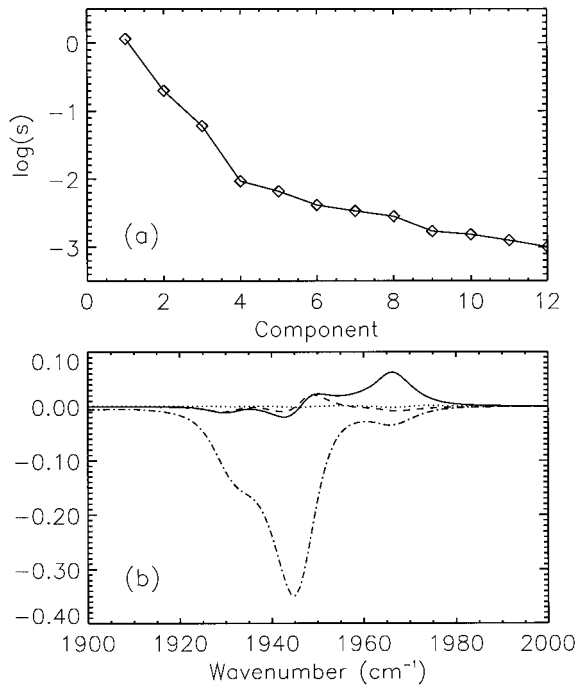

FIGURE 3 (*a*) The first 12 singular values $\tilde{s}_i$ of the swMbCO data, shown on a semilogarithmic plot. Only the first three singular values are above 0.01 and decrease strongly, which suggests an effective rank of three for the data matrix. (*b*) The first four basis spectra of the SVD matrix $\tilde{\mathbf{U}}$, scaled by their respective singular values. Line style order: −·−·, ——, – – –, · · · · · (starting with the first basis spectrum). Only the first three basis spectra contribute appreciably to the swMbCO spectra.

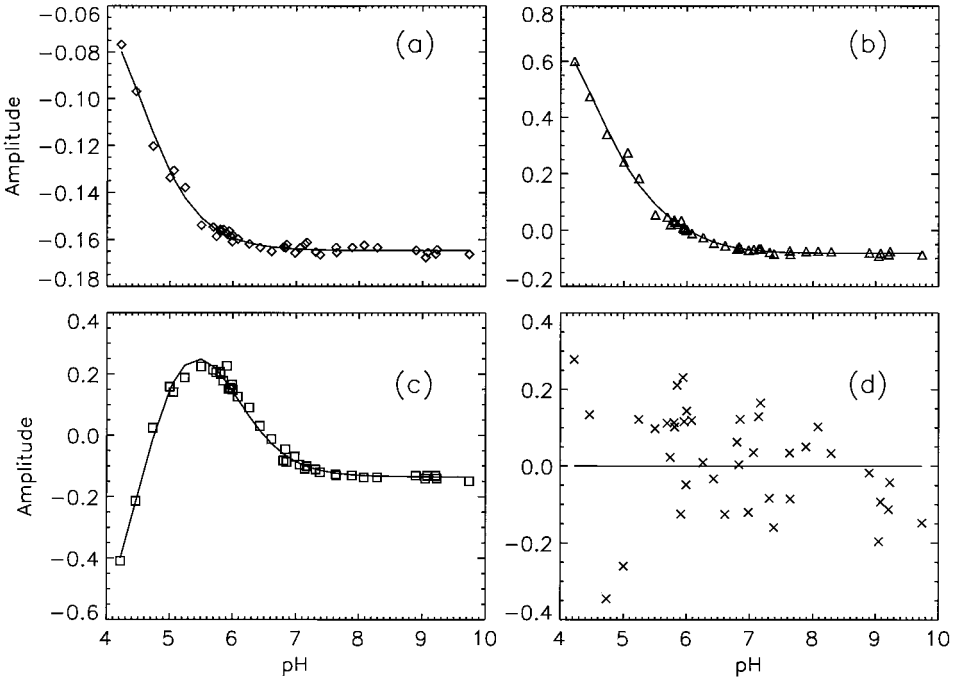

FIGURE 4 Symbols represent the first four amplitude vectors of the SVD matrix $\tilde{\mathbf{V}}$ determined from the swMbCO spectra. Only the first three amplitude vectors (*a*)–(*c*) show a systematic dependence on pH, whereas the fourth component (*d*) is no longer correlated with pH, consistent with an effective rank of three of the data matrix. Amplitude vectors calculated with the fitted parameters pK_{64} = 4.53 and pK_{97} = 5.91 are plotted as solid lines.

of the data matrix should not be based entirely on the behavior of the singular values, but also take into account the presence of correlations in the matrix $\tilde{\mathbf{V}}$ (Shrager and Hendler, 1982).

The MLS algorithm is needed to transform the orthogonal spectra in Fig. 3 *b* to physical spectra of the different protonation states S_{ij}. The least-squares fit of our thermodynamic model to the data yields pK_{64} = 4.53 and pK_{97} = 5.91. With these parameters, the best approximation to the $\tilde{\mathbf{V}}_r$ matrix by the model was calculated as $(\mathbf{F}(\hat{\mathbf{p}})\mathbf{F}(\hat{\mathbf{p}})^+)\tilde{\mathbf{V}}_r$; the three amplitude vectors of this matrix are plotted in Fig. 4 as solid lines on top of the experimental amplitude vectors of the matrix $\tilde{\mathbf{V}}$. According to Eq. 10, comparison of the $\tilde{\mathbf{V}}$ matrix vectors of the experimental data with the amplitudes based on the least-squares fit demonstrates the successful modeling of the data.

Fig. 1 *d* shows the physical spectra of the three distinguishable species of swMbCO, S_{00}, $S_{[01]}$, and S_{11}, and Fig. 2 *b* the relative weights, f_{00}, $f_{[01]}$, and f_{11} as a function of pH. The deprotonated (high-pH) form, S_{00}, consists essentially of A_1 and A_3. These two *A* substates do not respond to pH changes with adjustment of their relative populations. For the doubly protonated form, S_{11}, we expect the A_0 substate band. The S_{11} spectrum peaks around 1966 cm^{-1} as expected for the A_0 line, but the shape of the band is clearly asymmetric with a tail toward lower wavenumbers. The mixed species, $S_{[01]}$, appears in a first approximation as a blue-shifted copy of the unprotonated state, S_{00}.

The same analysis was also applied to the mutants H97F and H64L. In contrast to the native protein, only two linearly independent components were uncovered by SVD, and the pH dependence of the first two amplitude vectors of matrix $\tilde{\mathbf{V}}$ followed simple Henderson–Hasselbalch behavior (Müller, 1997).

The physical basis spectra of the mutants H97F and H64L are shown in Fig. 1, *e* and *f*, respectively. The basis spectra of H97F resemble S_{00} and S_{11} of the native protein, whereas H64L exhibits narrow, symmetric bands. The pK values obtained for all three proteins are compiled in Table 1.

Shapes and shifts of the basis spectra

Here we take a closer look at the effect of protonation on the shape and position of the individual basis spectra. The influence of protonation of H97 on the IR stretch bands can most easily be seen in the H64L mutant. Fig. 5 *a* shows the spectra of the deprotonated form, S_{x0}^{H64L}, and the protonated form, S_{x1}^{H64L}. The dashed line is S_{x0}^{H64L}, shifted by 2.8 cm^{-1} to the blue, producing a band almost identical to that of S_{x1}^{H64L}. The differences, aside from the band shift, are shown by the dotted line in Fig. 5 *a*, and indicate that the extent of conformational heterogeneity, which determines the width and shape of the band, is not appreciably affected by the protonation state of H97.

In the native protein, the protonation of H64 makes the effect of H97 protonation less obvious than in the H64L mutant (Fig. 1 *a*). Based on the behavior of the H64L mutant, we expect that protonation of H97 causes a blue shift of the individual *A* bands, so that the three *A* substates

TABLE 1 pK values as obtained from the global analysis of the IR spectra

	pK_{97}	pK_{64}
SwMb	5.91	4.53
H97F	—	4.49
H64L	6.33	—

Errors in the pKs are estimated to be ±0.1.

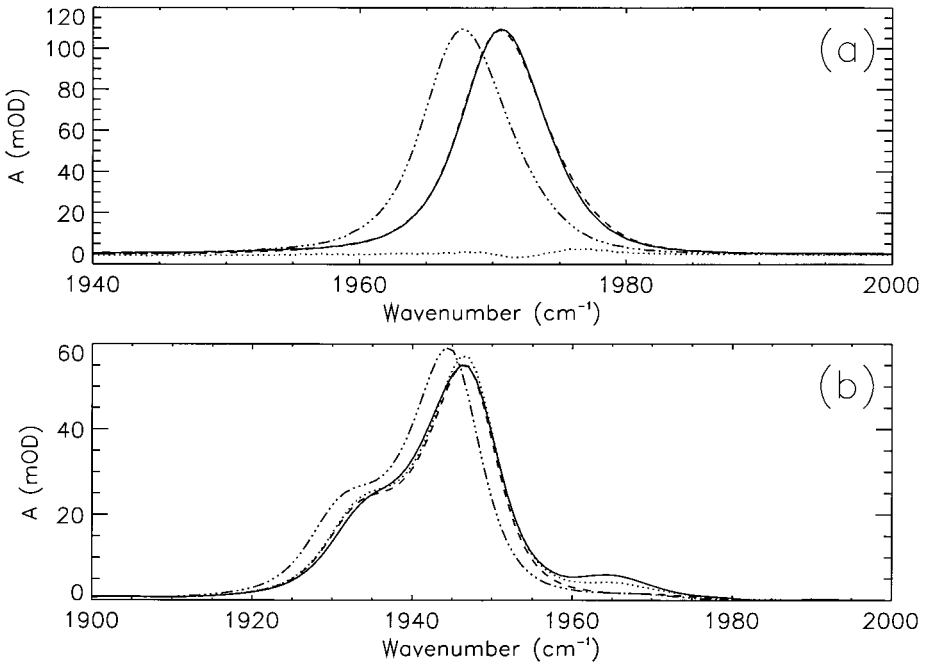

FIGURE 5 (*a*) Basis spectra of the distal mutant H64L; S_{x0}^{H64L} (–···–) associated with a neutral H97 residue; S_{x0}^{H64L} (– – –) shifted by 2.8 cm^{-1} to the blue; S_{x0}^{H64L} (——) associated with the protonated H97; difference spectrum (· · · · ·) of shifted S_{x0}^{H64L} minus S_{x1}^{H64L}. (*b*) Basis spectra of native swMbCO; S_{00} (–···–); $S_{[01]}$ (——); S_{00} (– – –) shifted by 2.0 cm^{-1} and scaled to the absorption maximum of $S_{[01]}$; fitted spectrum (·····): best fit resulting from a superposition of the $S_{[01]}$ spectrum with high pH form S_{00} and the low pH form S_{11}, which were allowed to shift to the blue and to the red, respectively, to mimic the influence of the protonation of H97 on the *A* substates. For details, see the text.

occur as pairs, A_0/A_0^+, A_1/A_1^+, and A_3/A_3^+, where the superscript + denotes the protonated state of H97.

We test this supposition in Fig. 5 *b*, where we examine in detail the changes between S_{00} and $S_{[01]}$. The spectrum of the mixed protonation states, $S_{[01]}$, consists of ≈95% species S_{01}, based on the pK difference of ≈1.4 between H97 and H64. Thus the $S_{[01]}$ spectrum reflects, in a good approximation, the changes in the S_{00} spectrum caused by protonation of H97. Analogous to the quantitative analysis of the H64L basis spectra, the S_{00} spectrum was shifted by 2.0 cm^{-1} (*dashed line* in Fig. 5 *b*) and scaled to the absorption maximum of the $S_{[01]}$ spectrum, to better compare the spectral changes associated with protonation of H97. The shifted spectrum approximates the mixed protonation spectrum $S_{[01]}$ well below 1950 cm^{-1}. As in the H64L mutant, marked changes in the band shapes are absent. However, the A_3 band clearly shifts further to the blue than the A_1 band upon protonation of H97F.

At wavenumbers above ≈1950 cm^{-1}, deviations between the shifted S_{00} and $S_{[01]}$ spectrum are observed, which arise from the contribution of S_{10} to $S_{[01]}$. S_{10} is expected to contribute a spectrum similar to that of the doubly protonated species, S_{11}, but red-shifted by ≈2 cm^{-1} due to the deprotonation of H97. If the protonation reaction of H97 merely shifts the *A* bands without changing their relative populations, then we should be able to represent the $S_{[01]}$ spectrum by

$$\tilde{S}_{[01]} = (1 - f_M)S_{00}(\nu + \Delta\nu_1) + f_M S_{11}(\nu - \Delta\nu_2), \qquad (23)$$

in which the basis spectra of the high pH form S_{00} and the low pH form S_{11} are allowed to shift by $\Delta\nu_1$ to the blue and by $\Delta\nu_2$ to the red to mimic the influence of the protonation of H97 on the *A* substates. The dotted line in Fig. 5 *b* gives the best representation of the $S_{[01]}$ spectrum based on a least-squares fit of the above equation. The coefficient f_M was calculated for a pK difference of 1.4 between H64 and H97 and kept fixed during the fit. The dotted line leads to a fair representation of the mixed protonation spectrum $S_{[01]}$. The difference in area between the fit and the $S_{[01]}$ spectrum in the region around 1964 cm^{-1} can be avoided by allowing the mixing coefficient f_M to vary during the fit. The mixing coefficient f_M of such a fit leads to a pK difference of 1.2 between H97 and H64, which is 0.2 pH units less than the difference determined by the global analysis.

The differences between the fit (*dotted line* of Fig. 5 *b*) and the basis spectrum are thus within the uncertainty of our pK determination by the global analysis. From these results we conclude that protonation of H97 causes a blue shift of the *A* bands without obvious population changes in the A_0, A_1, and A_3 substates; therefore, the free energy differences among these substates are only weakly influenced by the protonation of H97.

Basis spectra and taxonomic (*A*) substates

Figure 6 *a* shows the high-pH basis spectrum, S_{00} (symbols: ◇), together with a heuristic fit using three Voigtian lines (convolutions of Lorentzians and Gaussians). The two predominantly populated lines at 1944.5 cm^{-1} and 1933.1 cm^{-1}, assigned to A_1 and A_3, appear as a single species in our protonation scheme, because they do not respond to pH changes with adjustment of their relative populations for a deprotonated H97 and H64.

Close inspection of S_{00} (Fig. 6 *a*) reveals an additional line with a relative population of ≈1% at 1968 cm^{-1}, shown in an expanded view in Fig. 7. This line, which we denote

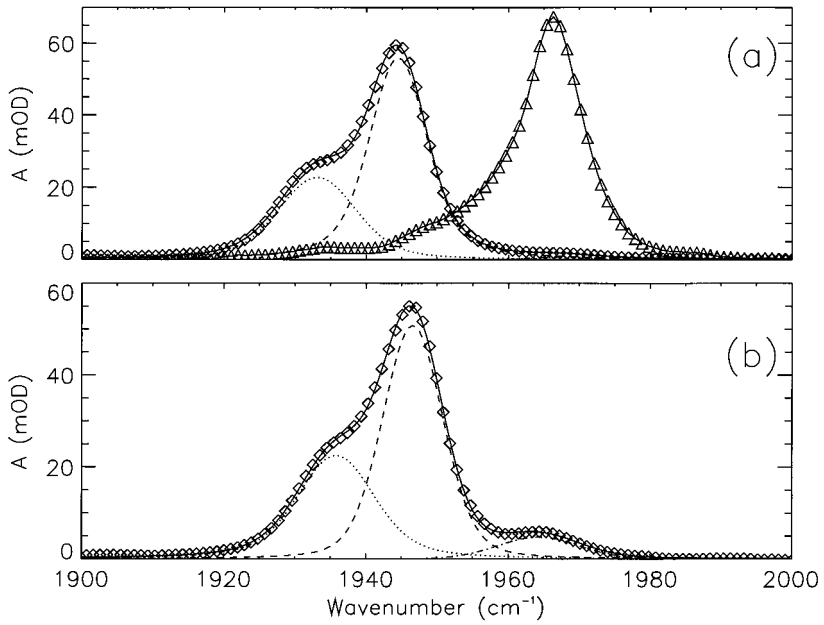

FIGURE 6 Fit of the basis spectra of swMbCO with discrete Voigtian lines, (*a*) the S_{00} spectrum (◇), fitted with three Voigtian bands to isolate the A_1 (– – –), A_3 (· · · · ·), and the A_x (– · – ·) substates. Solid line: sum of the three lines. The A_x substate with a population of ~1% is negligible on this scale. (See Fig. 7 for an expanded view). Basis spectrum of the doubly protonated form S_{11} (△, solid line to guide the eye) is asymmetric and needs at least two Voigtian bands for a fit to discrete lines (see Table 2). We identify the asymmetric S_{11} spectrum with the A_0 substate. (*b*) $S_{[01]}$ spectrum (◇), fitted with Voigtian bands to isolate the A_1^+ (– – –) and A_3^+ (· · · · ·) substates. The band (– · – ·) at 1964 cm^{-1} contains contributions from the A_x^+ and A_0 substate as discussed in the text. Solid line: sum of the three lines.

with A_x, must not be confused with either A_0 at 1964 cm^{-1}, or A_0^+ at 1966 cm^{-1}. The A_0 state (with neutral H97) should be red-shifted by ≈2 cm^{-1} from the low pH form (with protonated H97) to 1964 cm^{-1} and is 4 cm^{-1} from the A_x peak position. This suggests that A_x is a minority species that is not related to the A_0 or A_0^+ conformations, which are characterized by a H64 imidazole outside of the heme pocket. Further evidence of this assertion comes from ligand binding studies of the A_x substate (see below). The population ratio $A_x/A_1 = \exp[-\Delta G/(RT)] \approx 0.012$ corresponds to a difference in free energy ΔG between the two substates of ≈11 kJ/mol at 300 K.

In Fig. 6 *b*, a fit of separate Voigtian lines to the $S_{[01]}$ spectrum is plotted. The main contribution to the spectral area arises from two lines at 1946.6 and 1935.8 cm^{-1}.

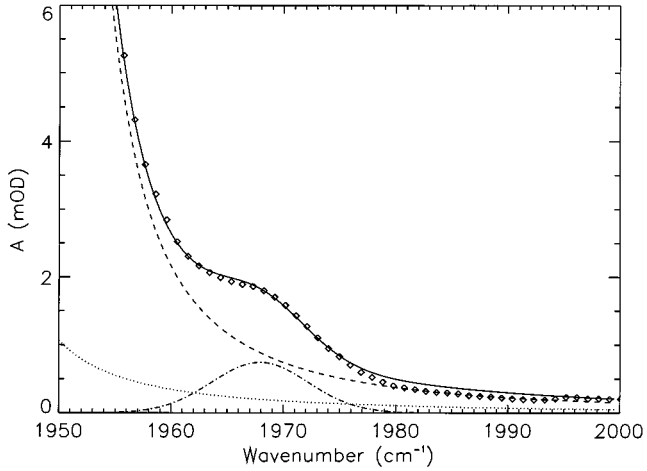

FIGURE 7 Expanded view of S_{00} in the region above 1950 cm^{-1} to display the A_x substate at 1968 cm^{-1} (symbols and linestyles as in Fig. 6 *a*).

Because $S_{[01]}$ contains predominantly the species S_{01} (protonated H97), we assign the two bands to A_1^+ and A_3^+. The presence of both A_1 and A_1^+ at intermediate values of pH may have led to the initial assignment of an A_2 state in swMbCO (Ansari et al., 1987). The high wavenumber part of $S_{[01]}$ not only contains A_x^+, but also an additional contribution from the S_{10} protonation state, which is appropriately assigned to A_0. As S_{10} contributes ≈6% to the total area of the spectrum and A_x^+ only ≈0.8%, the peak position of this line at 1964.3 cm^{-1} is essentially determined by the A_0 substate. The features of the A_x band are completely obscured in the $S_{[01]}$ basis spectrum by the dominance of the A_0 substate; therefore no further deconvolution of the line at 1964.3 cm^{-1} into two separate bands was attempted.

Also shown in Fig. 6 *a* is the basis spectrum S_{11} (symbols: △) of the doubly protonated form, which peaks at 1966.3 cm^{-1} and is therefore assigned to the A_0^+ substate. This band is asymmetric, with a substantial broadening toward the red side. Analysis of S_{1x}^{H97F} (Fig. 1 *e*) reveals a similar broadening. Because acid denaturation begins to occur at pH 4.2, it is not possible to physically isolate this species, global analysis is required to obtain the spectrum of doubly protonated, but still native, myoglobin. Furthermore, a functional line fit will always assign the red tail to one of the other substates, leading to misinterpretation of both the line shapes and the pH dependence of the IR spectrum. To get a reasonable agreement of S_{11} with a fit based on model functions, at least two Voigtian lines have to be included. This, however, is merely an exercise in line fitting, because the individual bands cannot currently be associated with different species.

Close inspection of the shoulder in the S_{00} and $S_{[01]}$ spectrum around 1935 cm^{-1} demonstrates that a fit of the A_1/A_1^+ and A_3/A_3^+ bands with Voigtian lineshapes cannot reproduce this feature accurately. Instead of assigning the

deviation between data and fit to an additional substate, for which we have no further evidence at this time, we note that the assumptions of simple model bands is most likely only an approximation of the true lineshape of the A_1/A_1^+ and A_3/A_3^+ substates as in the case of the asymmetric A_0^+ band.

The ratio of the A_1 and A_3 population was calculated from the results of the Voigtian deconvolution (see Table 2) for the S_{00} and the corresponding ratio A_1^+/A_3^+ for the $S_{[01]}$ basis spectrum. The ratio changes from 0.5 for the S_{00} state to 0.6 for the $S_{[01]}$ state. The relative change of 0.1 is just within the error margin based on the uncertainty of the Voigtian fit parameters, but we still like to comment on this relative increase in A_3^+ population with respect to A_3. Protonation of H97 on the proximal side may cause slight conformational changes on the distal side. A change in the relative free energy between A_1/A_1^+ and A_3/A_3^+ by only 0.4 kJ/mol upon protonation of H97 would explain the increase in the population ratio.

In summary, the protonation states of our model, S_{ij}, are related to the individual A substate lines in the following way:

$$S_{00} = f_1A_1 + f_3A_3 + f_xA_x \tag{24}$$

$$S_{01} = f_1^+A_1^+ + f_3^+A_3^+ + f_x^+A_x^+ \tag{25}$$

$$S_{10} = A_0 \tag{26}$$

$$S_{11} = A_0^+ \tag{27}$$

$$S_{[01]} = f_M(f_1^+A_1^+ + f_3^+A_3^+ + f_x^+A^+) + (1 - f_M)f_0A_0. \tag{28}$$

Here, the f_i denote the fractional populations of the A substates. The parameters obtained from fits of Voigtian lines to the protonation states are compiled in Table 2.

Low-temperature studies

At low temperature, where interconversion between A states does not occur, it is possible to determine the enthalpy barrier to each A state using TDS. Because multiple A substates are measured at the same time in the same sample, the measurement is sensitive to slight differences between substates. TDS studies were done with two samples, one at pH 6.8, and the other one at pH 8.5, both pH values as measured at room temperature. In general, the pKs of protonable groups depend on temperature (Douzou, 1977). Below 160 K, however, the solvent is frozen, so both pK and pH are constant. Additionally, the protonation states can be inferred from the spectra at 12 K, shown in Fig. 8 *a*. The solid line shows the spectrum of the sample at pH 8.5. It agrees with the spectrum of the doubly deprotonated species except that the lines are narrower and spectrally better resolved at 12 K, with A_1 at 1946 cm^{-1} and A_3 at 1925 cm^{-1}. The A_x substate at 1970 cm^{-1} is well separated from the tail of A_1.

The dashed line in Fig. 8 *a* represents the spectrum at pH 6.8. The band at 1966 cm^{-1} consists predominantly of A_0^+. From the fraction of A_0^+, the population of the protonation states S_{00}, $S_{[01]}$, S_{11}, can be estimated as 10, 75, and 15%, respectively. Therefore, the spectrum below 1950 cm^{-1} consists mainly of A_1^+ and A_3^+. Consistent with room temperature data, these bands are slightly blue-shifted from A_1 (by 2 cm^{-1}) and A_3 (by 3.5 cm^{-1}) due to H97 protonation. Fig. 8 *b* shows a TDS contour plot of the pH 8.5 sample. A_1 peaks at 47 K and A_3 at 70 K. The maximum of the A_x TDS signal appears at 45 K, close to, but clearly distinct from, that of A_1, whereas maximum recombination of A_0^+ occurs at 42 K, as seen in the contour map of the pH 6.8 sample in Fig. 8 *c*. Consequently, A_x is both kinetically and spectroscopically distinct from the other A bands and therefore represents a separate substate.

Error analysis

The dominant sources of error are the pH determination of the samples and the baseline correction of the IR spectra. Both of these errors are systematic in nature, and their

TABLE 2 Parameters of discrete Voigtian *A* substate lines fitted to the spectra of the protonation species S_{ij} of swMbCO and the mutants

	Line 1			Line 2			Line 3		
	Peak [cm^{-1}]	FWHMF [cm^{-1}]	Area [%]	Peak [cm^{-1}]	FWHM [cm^{-1}]	Area [%]	Peak [cm^{-1}]	FWHM [cm^{-1}]	Area [%]
S_{00}	$1933.1^{+0.4}_{-0.2}$	12.9 ± 0.7	33.8 ± 2.4	1944.5 ± 0.1	9.7 ± 0.3	$65.5^{+1.2}_{-4.5}$	1967.9 ± 1.5	10.2 ± 3.5	0.8 ± 0.2
$S_{[10]}$	1935.8 ± 0.4	12.9 ± 0.9	35.3 ± 3	1946.6 ± 0.1	10.1 ± 0.25	$57.9^{+4.0}_{-5.5}$	$1964.3^{+1.0}_{-0.0}$	$13.6^{+2.5}_{-2.0}$	6.8 ± 1.1
S_{11}	1957.3	17.7	37.5	1966.5	9.2	62.5	—	—	—
S_{x0}^{H64L}	1968.0	7.8	100	—	—	—	—	—	—
S_{x1}^{H64L}	1970.8	7.6	100	—	—	—	—	—	—
S_{0x}^{H97F}	1932.4	11.7	30.1	1944.3	11.4	68.7	1968.0	10.0	1.2
S_{1x}^{H97F}	1953.3	23.2	17.2	1964.8	10.8	82.8	—	—	—

Peak positions, full widths at half maximum, and relative areas of Voigtian lines were determined by nonlinear least-squares fits. The standard deviation of the fitting parameters was determined for the basis spectra S_{00} and $S_{[01]}$. For the weak A_x band of S_{00} at 1967.9 cm^{-1}, the errors in the fitting values were determined with fixed A_3 parameters. The two Voigtian bands used to approximate the basis spectra S_{11} and S_{1x}^{H97F} are mere parameterizations of the asymmetric bands peaking at 1966.3 and 1964.7 cm^{-1}, respectively. The errors associated with the deconvolution of the S_{0x}^{H97F} substate are similar to those of the corresponding bands of the S_{00} spectrum.

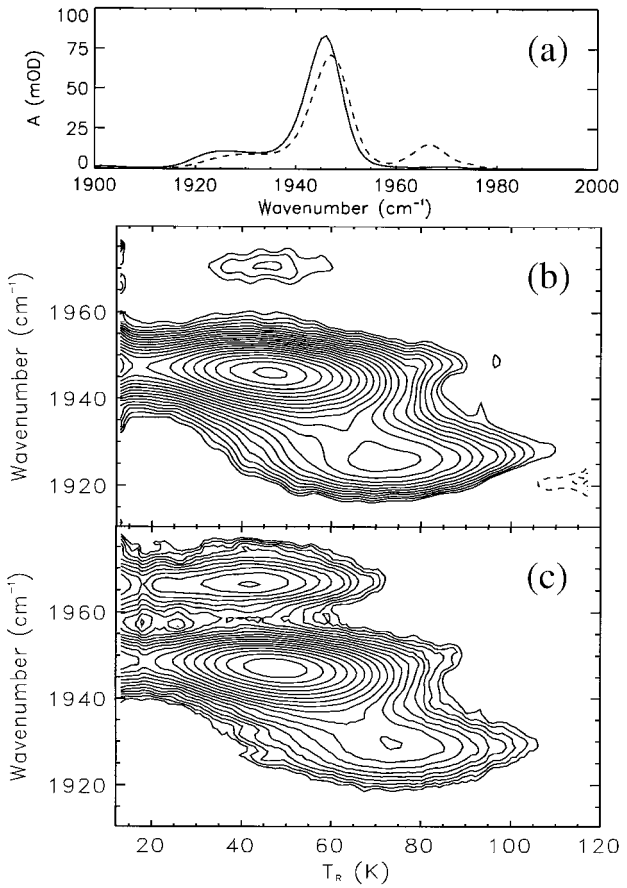

FIGURE 8 Low temperature FTIR difference spectra and TDS contour maps of swMbCO. (*a*) Spectrum at $T = 12$ K for a sample at pH 8.5 (——) and at pH 6.8 (– – –), with total areas normalized to 1 OD cm^{-1}, pH as determined at room temperature. (*b*) TDS contour map of the *A* substate region for the pH 8.5 sample. (*c*) TDS contour map of the *A* substate region for the pH 6.8 sample. The contour lines are spaced logarithmically.

influence on the precision of the basis spectra and pK parameters are difficult to quantify. Estimates of the errors of the pK values were obtained by comparison of the parameters with those obtained from a second data set and also from an error analysis analogous to the Student's *F* test (that actually assumes the presence of statistical errors). In both cases, the estimated errors in the pK determinations were ± 0.1 pH units.

The uncertainty in the pKs leads to an uncertainty in the determination of the physical spectra of the different protonation species. Fig. 9 *a* shows the best-fit $S_{[01]}$ spectrum together with the two spectra that constitute the extreme forms that can be obtained by changing the pK values in the range of ± 0.1 pH units. The extreme forms where determined by fitting the data while restricting either the pK of H64 or H97 to fixed values within 0.1 pH units around their optimal value. The basis spectra of this forced fit were compared to the original set by computing the least-squares deviation. The two most strongly deviating spectra, which encompass all the other cases, are shown in Fig. 9 *a* to convey the uncertainty connected with the determination of the basis spectra.

Although the low wavenumber range of the $S_{[01]}$ spectrum is almost unaffected, the region around 1966 cm^{-1} is sensitive to uncertainties. The spectrum S_{11} is most strongly dependent on the choice of the pK parameters (Fig. 9 *b*). Its relatively small population and strong overlap with the $S_{[01]}$ and S_{00} spectra at low wavenumbers leads to a significant uncertainty below 1950 cm^{-1}, whereas the high-wavenumber region is again well determined. The spectrum S_{00} of the deprotonated form is directly observed at high pH and therefore not affected by uncertainties in the pK parameters.

We found the shapes of the basis spectra to be a good marker from which to judge the quality of a model. Systematic errors in the data, such as those caused by a drifting pH-electrode, typically resulted in basis spectra that were clearly unphysical (e.g., negative at some wavenumbers). We considered models other than the one sketched in Fig. 2 *a* but rejected them as they yielded poor fits or unphysical basis spectra.

DISCUSSION

Protonation of H97 and π back-bonding

The pH-dependent shifts of the A_1 and A_3 bands by a few wavenumbers were noted by Shimada and Caughey (1982) for bovine MbCO. They concluded that a group with a pK ≈ 6 must be responsible for this change. Here we have identified this protonating group as H97. The shift can be understood from the nature of the chemical bond between the CO and the heme iron, which can be represented as a resonant hybrid of three electronic configurations,

$$\underset{(1)}{M^{\delta(-)} - C \equiv O:^{\delta(+)}} \rightleftharpoons \underset{(2)}{M = C = \dot{O}:} \rightleftharpoons \underset{(3)}{M^{\delta(+)} = C = \ddot{O}:^{\delta(-)}}. \tag{29}$$

In terms of molecular orbitals, the major contributions to the Fe–CO bond are provided by a σ donation from the weakly antibonding 5σ molecular orbital to the iron 4*s* and $3d_z^2$ orbitals and π back-bonding from the iron $3d_\pi$ orbitals to the strongly antibonding CO $2\pi^*$ orbitals (Kushkuley and Stavrov, 1996). Consequently, with increased back-bonding, the Fe–C bond becomes stronger and the C–O bond weaker, and an inverse correlation between the stretch frequency, ν, of the Fe–C and the CO bond is expected. Indeed, on the basis of data on many heme proteins, Li and Spiro (1988) determined the following inverse correlation empirically,

$$\nu_{FeC}[cm^{-1}] = 1935 - 0.73\ \nu_{CO}[cm^{-1}]. \tag{30}$$

In the heme molecule, the iron *d* orbitals are coupled to the extended delocalized π electron system of the heme, and the extent of back-bonding depends on the properties of the heme macrocycle. It was shown that electron-withdrawing groups at the heme periphery decrease π back-bonding and thus increase ν_{CO} by a few wavenumbers (Alben and

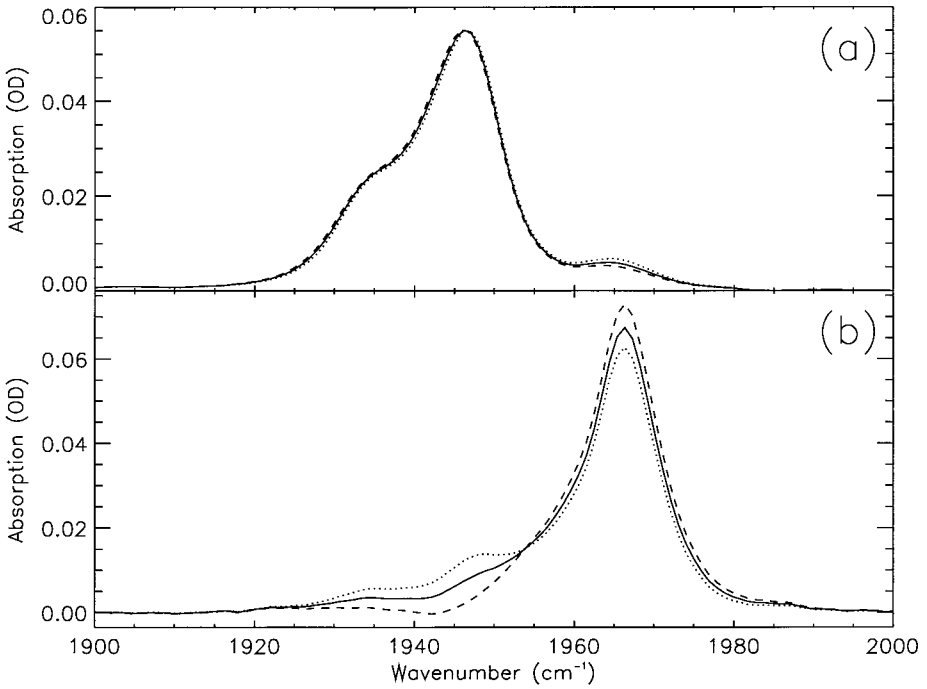

FIGURE 9 Best-fit spectra (*a*) $S_{[01]}$ and (*b*) S_{11}, shown as solid lines together with two spectra (*dashed* and *dotted lines*) that constitute the largest deviations that can be obtained by changing the pK values in the range of the estimated error of ±0.1 pH units. The extreme forms where determined by fitting the data while restricting either the pK of H64 or H97 to fixed values within 0.1 pH units around their optimal value. The resulting spectra of this forced fit were compared to the original set by computing the least-squares deviation.

Caughey, 1968; Ray et al., 1994). Although H97 is not a sidegroup of the heme, it is in the immediate vicinity of it and maintains a hydrogen bond with the heme-7-propionate. Upon protonation, the positive charge on H97 attracts electron density from the π electron system of the heme, and the π donor character of the heme iron decreases. Consequently, the extent of back-bonding decreases, and the IR stretch frequency increases. The fact that the size of the shift depends on the particular *A* substate line (Table 3) shows that back-bonding is sensitive to the different conformations of the taxonomic substates.

Variations in the Fe–C distance lead to strong changes in π back-bonding. Model calculations by Kushkuley and Stavrov (1996) indicate that a minute decrease in the Fe–C bond distance by 0.01 Å, which requires less than 0.1 kJ/mol in their calculation, reduces back-bonding markedly and leads to shifts of about 5 cm^{-1} in ν_{CO}. Consequently, small conformational changes around the active site between the different substates can produce significant shifts.

TABLE 3 Frequency shifts of the *A* substate bands upon protonation of the H97 imidazole sidechain

	swMbCO			H64L
	$A_0 - A_0^+$ [cm^{-1}]	$A_1 - A_1^+$ [cm^{-1}]	$A_3 - A_3^+$ [cm^{-1}]	$A_{H64L} - A_{H64L}^+$ [cm^{-1}]
S_{ij} shift	1.4–2.2	2.0–2.1	2.5–2.7	2.8
Voigtian	2.0	2.1	2.7	2.8

The shifts of the *A* substates upon protonation of H97 were estimated by either shifting two basis spectra S_{ij} with respect to each other to match the individual bands or by calculating the difference in their peak positions based on a Voigtian deconvolution of the basis spectra (see Table 2). The A_0 substate was assumed to be well represented by the band at 1964.3 cm^{-1} of the $S_{[01]}$ spectrum, neglecting the small contribution of the A_x substate.

The reduction of the π donor character of the heme iron upon protonation of H97 causes an overall increase in ν_{CO} for each substate, but the influence exerted upon ν_{CO} depends on the Fe–C bond length, thus influencing each *A* substate differently. Moreover, other effects such as subtle differences in the tilt angle of the FeCO unit could also affect the back-bonding and contribute to the observed variations in the shift of the individual *A* bands (Kushkuley and Stavrov, 1996).

The inverse correlation between ν_{FeC} and ν_{CO} would predict a shift of the Fe–C Raman line of the A_1 substate near 510 cm^{-1} (Li and Spiro, 1988; Ramsden and Spiro, 1989) upon protonation of H97 by 1.5 cm^{-1} to the red. However, electrical charges close to the edge of the porphyrin ring, as is the case for a protonated H97, can lead to rather complicated relationships between ν_{CO} and ν_{FeC} (Kushkuley and Stavrov, 1996).

Protonation of H64 and electrostatic effects

When H64 is replaced by an aliphatic amino acid, a single *A* substate band appears in the region near 1970 cm^{-1} (Braunstein et al., 1993; Li et al., 1994). H64L is typical for this class of mutants (see Fig. 1 *c*) with relatively weak interactions between the CO dipole and the distal heme pocket. The IR line position is similar to that of A_0 and A_0^+ of the native protein, which implies that the H64 imidazole sidechain only weakly interacts with the CO in these substates. Indeed, the x-ray structure analysis of MbCO at various pH values reveals that the H64 sidechain can assume at least two distinctly different conformations (Yang and Phillips, 1996). Fig. 10 *a* shows parts of the active site region at pH 6, with the H64 imidazole in the interior of the

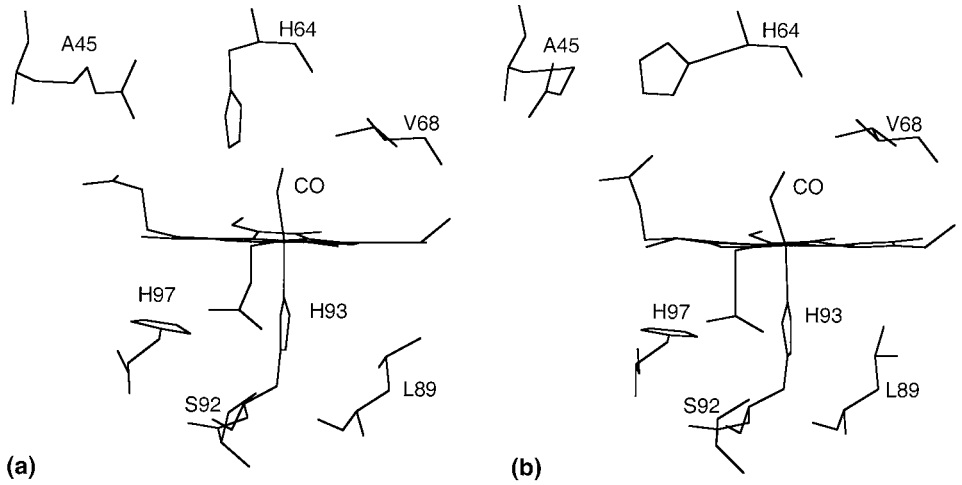

FIGURE 10 Heme pocket environment of swMbCO at (*a*) pH 6 and (*b*) pH 4. The data were taken from the Protein Data Bank (accession numbers 1spe, 1vxf).

protein in close vicinity to the bound CO. At pH 4, however, the H64 sidechain has rotated around the C_α—C_β bond so that its ring points out of the distal heme pocket into the solvent (Fig. 10 *b*). The distance to the bound CO has become much larger, so that much weaker electrostatic interactions are expected, consistent with the CO stretch band at a similar wavenumber as in the aliphatic H64 mutants.

The observation that the pH dependence of the H97F mutant is described by a simple Henderson–Hasselbalch titration implies that the conformational change observed in the x-ray structures must be coupled to the protonation reaction; protonation of the H64 sidechain is the reason for the observed conformational change. Uncompensated charges in the low-dielectric interior of the protein carry a large energetic penalty. Transfer of a charge from the solvated environment at the protein surface to the low-dielectric interior costs in excess of 10 kJ/mol in free energy (Nakamura, 1996), which is a significant fraction of the entire stabilization energy. Upon protonation of the H64 sidechain, the protein therefore responds with a conformational change so that the imidazole becomes exposed to the high-dielectric solvent medium. This sidechain motion can be interpreted as a local unfolding of the protein structure. Whereas protonation of H64 is still accommodated by a largely native protein molecule, pH values below 4.2 lead to acid-induced denaturation, a global unfolding transition, during which the majority of the interior histidines are protonated. The unfolded structure of the protein allows effective shielding of the charges by the solvent.

Protonation of H64 requires a conformational change, or local unfolding, and so is related to the type of changes occurring in hydrogen exchange experiments by Englander and coworkers (Englander et al., 1992; Bai et al., 1995). They have given clear evidence that the exchange of amide hydrogens under mildly denaturing conditions is governed by concerted local unfolding of small protein segments, whereas a global, cooperative unfolding transition is induced by harsh conditions such as high temperature, high denaturant concentrations, or extreme pH values.

Structural properties of the *A* substates

Although vibrational spectra cannot give structural information directly in terms of atomic coordinates, they are sensitive to local interactions and can—once these interactions are understood—indirectly provide structural information. The observed variations in ν_{CO} (and ν_{FeC}) have led to a variety of structural models, including electron donor-acceptor interactions between the N_ϵ of H64 and the bound CO ligand (Maxwell and Caughey, 1976; Makinen et al., 1979; Fuchsman and Appleby, 1979; Ray et al., 1994), distortions of the Fe–CO complex (Maxwell and Caughey, 1976; Makinen et al., 1979; Li and Spiro, 1988; Potter et al., 1990), and electrostatic interactions (Li and Spiro, 1988; Morikis et al., 1989; Oldfield et al., 1991; Balasubramanian et al., 1993; Ivanov et al., 1994; Li et al., 1994; Springer et al., 1994). Kushkuley and Stavrov (1996, 1997) concluded from quantum-chemical calculations that electrostatic interaction is the most important factor determining the CO stretch frequencies. The specific assignments are, however, still ambiguous. In the following, we discuss some new insights that can be gained from the basis spectra of the different protonation states.

The high-pH (neutral) species, S_{00}, consists mainly (99%) of the A_1 and A_3 substates, which both have their imidazole side chain inside the heme pocket, as seen in the x-ray (Kuriyan et al., 1986; Quillin et al., 1993), neutron (Cheng and Schoenborn, 1991), and NMR (Ösapay et al., 1994) structures. In all these structures, A_1 and A_3 cannot be distinguished, and unfortunately, the assignment of their structural differences on the basis of spectroscopic data is not conclusive either. For a discussion of various suggestions, we refer to Oldfield et al. (1991), Ray et al. (1994), and Kushkuley and Stavrov (1997).

Here we only wish to comment on the suggestion of an additional hydrogen bond to $N_{\delta 1}$ by a solvent molecule outside of the pocket (Kushkuley and Stavrov, 1997). This suggestion is quite appealing, as it is consistent with our observation that the dynamics of interconversion between

A_1 and A_3 is very fast at room temperature (≈nanoseconds) and sets in near the glass transition of the solvent (Johnson et al., 1996). Moreover, the ligand binding properties of A_3 in crystals are modified (Nienhaus et al., 1998), which could arise from a different solvent structure near the H64 imidazole sidechain.

Besides A_1 and A_3, the S_{00} species also contains the weakly populated A_x substate at 1968 cm^{-1}. In this substate, the H64 imidazole sidechain is neutral, so it is expected to reside inside the heme pocket, but in such a way that it interacts minimally with the bound CO. Olfield and coworkers (1991) showed that the $H64^{\epsilon 2,180}$ conformation interacts minimally with the CO stretch at 1967 cm^{-1}. They assigned A_0 to this conformation before it became clear that the distal histidine ring points out of the pocket in the A_0 conformation. Kushkuley and Stavrov (1997) also obtained a small blue shift of this species of 3.5 cm^{-1} from its value in the absence of electrostatic interactions. We can use the position of A_0 at 1964.3 cm^{-1}, as extracted from the $S_{[10]}$ basis state (Table 2), as the reference state, because the imidazole side chain in A_0 is expected not to interact with the CO. This yields a red shift of A_x of 3.7 cm^{-1}, in quantitative agreement with the theoretical predictions by Kushkuley and Stavrov (1997). Consequently, we suggest that the A_x substate may represent the $H64^{\epsilon 2,180}$ conformation.

The S_{11} basis spectrum corresponds to the A_0^+ substate with the protonated H64 imidazole swung out of the heme pocket. The S_{11} spectrum, which peaks at 1966.3 cm^{-1}, exhibits an asymmetric bandshape with a substantial tail toward the red. Such a broadening has also been observed in the H64G and H64A mutants, whereas H64V and H64L yield symmetric, narrower lines (Braunstein et al., 1993; Li et al., 1994). Replacement of H64 with glycine or alanine leaves enough room in the pocket to accommodate an additional water molecule in the CO-bound state, as has been observed by x-ray crystallography (Quillin et al., 1993). It has been suggested that the interaction between the CO and the additional water molecule gives rise to the red tail in these mutants. The fact that the crystal structure of the native protein at pH 4 does not show a water molecule in the distal pocket does not necessarily disprove this suggestion, because a water molecule with a low occupancy of 10%, as judged from the area under the red tail, will not be resolved in the electron density maps.

Asymmetric broadening may even appear in the absence of water molecules in the native protein if the positively charged H64 imidazole fluctuates significantly around its mean position outside of the heme pocket. Coulomb interactions are long-ranged, and fluctuations, which decrease the distance between the imidazole ring and the ligand CO, can produce significant red shifts, as indicated by theoretical calculations (Kushkuley and Stavrov, 1996). Because the electric field strength at the CO position does not change linearly with histidine position, an asymmetric band can result from spatially symmetric fluctuations.

The Champion model

Champion and coworkers have proposed a four-well model that treats conformational change and protonation of H64 (Morikis et al., 1989; Zhu et al., 1992; Tian et al., 1993, 1996). The conformation characterized by the H64 imidazole in the pocket was termed closed and associated with the A_1 and A_3 substates, whereas the one with the imidazole swung out of the heme pocket was called open. Schematically, the model can be represented by

$$\begin{array}{ccc} A_{1,3} & \overset{pK_{64}^{closed}}{\longleftrightarrow} & A_{1,3}^{p} \\ K \updownarrow & & \updownarrow K^{p} \\ A_0 & \underset{pK_{64}^{open}}{\longleftrightarrow} & A_0^{p} \end{array} . \qquad (31)$$

In this model, $A_{1,3}$ and $A_{1,3}^p$, denote the closed state, and A_0 and A_0^p the open state with unprotonated and protonated H64 sidechain, respectively. The pK of H64 depends on the conformation and is denoted by pK_{64}^{open} and pK_{64}^{closed}; thus protonation affects the equilibrium between the conformations, quantified by K and K^p. Furthermore, the model assumes that the spectroscopic properties are governed by the conformation, but not by the protonation states. In other words, A_0 and A_0^p and $A_{1,3}$ and $A_{1,3}^p$, are assumed to have the same spectra. For the closed conformation, i.e., the pair $A_{1,3}/A_{1,3}^p$, this appears unlikely, as a charge in the interior of the protein, close to the bound CO, should yield a large shift of the CO stretch band from its position in the unprotonated form.

The model rests on the assumption that an equilibrium exists between A_0^p and $A_{1,3}^p$ at low pH (pH $<$ 4), with significant population of both species, and between A_0 and $A_{1,3}$ at high pH (pH $>$ 7), again with population of both species. Although it is very difficult to perform a clean analysis at low pH due to the appearance of the acid-denatured form, there is no such problem at high pH. At high pH, we have shown, however, that the band at high wavenumber, A_x, has distinctly different spectroscopic and ligand-binding properties so that it should not be confused with A_0.

The pH dependence predicted by the Champion model disagrees with the experimental data on MbCO in Fig. 1 *a*. In contrast, our conceptually simpler model, involving two energetically uncoupled protonations, successfully describes the data to within the statistical error.

CONCLUSIONS

To investigate the pH dependence of conformational substates in MbCO and two mutants, H97F and H64L, a global analysis technique was performed on infrared spectra of the CO stretch bands in MbCO, which assumes no particular functional form for the component spectra yet allows ther-

modynamic models involving protonating residues (H97 and H64) to be quantitatively tested. It provides the spectra of the individual protonation species even when they cannot be isolated in pure form. These basis spectra are much more readily related to structure and theory than is the complete IR spectrum.

The analysis shows that protonation of histidine 97 shifts the CO stretch frequency by ≈ 2 cm^{-1} to the blue. Unprotonated H64 produces distinct ν_{CO} bands at 1933 and 1945 cm^{-1} with distinctly different ligand-binding properties. Moreover, it was shown that a minor species at 1968 cm^{-1} exists that is spectroscopically and kinetically distinct from the conformation that appears upon protonation of H64, associated with a single line at 1966 cm^{-1}. No other pH-dependent changes were observed in the spectrum between pH 4.2 and 9.5. In the mutants, protonation effects of either of the two histidines on the spectra were observed individually, whereas they occur in the native protein in an additive fashion, implying that their energetic coupling is sufficiently weak (less than 1 kJ/mol) to remain unresolved in the data.

While the present work clarifies some aspects of conformational heterogeneity in heme proteins related to protonations near the active site, more work remains to be done to characterize the multitude of conformational substates in the lower tiers of the substate hierarchy (Nienhaus and Young, 1996).

We thank H. Frauenfelder and Solomon S. Stavrov for helpful discussions. This work was supported by grants from the National Science Foundation (PHY95-13217) and the National Institutes of Health (GM 18051).

REFERENCES

Alben, J. O., and W. S. Caughey. 1968. An infrared study of bound carbon monoxide in the human red blood cell, isolated hemoglobin, and heme carbonyls. *Biochemistry.* 7:175–183.

Ansari, A., J. Berendzen, D. Braunstein, B. R. Cowen, H. Frauenfelder, M. K. Hong, I. E. T. Iben, J. 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.

Bai, Y., T. R. Sosnick, L. Mayne, and S. W. Englander. 1995. Protein folding intermediates: native-state hydrogen exchange. *Science.* 269: 192–197.

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

Berendzen, J., and D. Braunstein. 1990. Temperature-derivative spectroscopy: a tool for protein dynamics. *Proc. Natl. Acad. Sci. USA.* 87:1–5.

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

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

Doster, W., D. Beece, S. F. Bowne, E. E. DiIorio, L. Eisenstein, H. Frauenfelder, L. Reinisch, E. Shyamsunder, K. H. Winterhalter, and K. T. Yue. 1982. Control and pH dependence of ligand binding to heme proteins. *Biochemistry.* 21:4831–4839.

Douzou, P. 1977. Cryobiochemistry. Academic Press, New York.

Englander, S. W., J. J. Englander, R. E. McKinnie, G. K. Ackers, G. J. Turner, J. A. Westrick, and S. J. Gill. 1992. Hydrogen exchange measurements of the free energy of structural and allosteric change in hemoglobin. *Science.* 256:1684–1687.

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, J. R. Mourant, P. Ormos, L. Reinisch, R. Scholl, A. Schulte, E. Shyamsunder, L. B. Sorenson, P. J. Steinbach, A. Xie, R. D. Young and K. T. Yue. 1990. Proteins and pressure. *J. Phys. Chem.* 94:1024–1038.

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

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.

Golub, G., and C. V. Loan. 1983. Matrix computations. Johns Hopkins University Press, Baltimore, MD.

Han, S., D. L. Rousseau, G. Giacometti, and M. Brunori. 1990. Metastable intermediates in myoglobin at low pH. *Proc. Natl. Acad. Sci. USA.* 87:205–209.

Hendler, R. W., and R. I. Shrager. 1994. Deconvolution based on the singular value decomposition and the pseudoinverse: a guide for beginners. *J. Biochem. Biophys. Meth.* 28:1–33.

Henry, E. R., and J. Hofrichter. 1992. Singular value decomposition: Application to analysis of experimental data. *In* Methods in Enzymology, Vol. 210. Academic Press, New York. 129–193.

Hong, M. K., D. Braunstein, B. R. Cowen, H. Frauenfelder, I. E. T. Iben, J. R. Mourant, P. Ormos, R. Scholl, A. Schulte, P. J. Steinbach, A. H. Xie, and R. D. Young. 1990. Conformational substates and motions in myoglobin: external influences on structure and dynamics. *Biophys. J.* 58:429–436.

Iben, I. E. T., D. Braunstein, W. Doster, H. Frauenfelder, M. K. Hong, J. B. Johnson, S. Luck, P. Ormos, A. Schulte, P. J. Steinbach, A. H. Xie, and R. D. Young. 1989. Glassy behavior of a protein. *Phys. Rev. Lett.* 62:1916–1919.

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.

Johnson, J. B., D. C. Lamb, H. Frauenfelder, J. D. Müller, B. McMahon, G. U. Nienhaus, and R. D. Young. 1996. Ligand binding to heme proteins. VI. Interconversion of taxonomic substates in carbonmonoxymyoglobin. *Biophys. J.* 71:1563–1573.

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

Kushkuley B., and S. S. Stavrov. 1996. Theoretical study of the distal-side steric and electrostatic effects on the vibrational characteristics of the FeCO unit of the carbonylheme proteins and their models. *Biophys. J.* 70:1214–1229.

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.

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

Li, X. Y., and T. G. 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.

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 hexaphospate induced cleavage of proximal histidine to iron bonds. *Biochemistry.* 15:388–395.

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

Mourant, J. R., D. 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. M. 1997. Charakterisierung von Struktur und Funktion taxonomischer Konformationssubzustände in CO-ligandiertem Myoglobin. Ph.D. thesis, Technische Universität München, Germany.

Nakamura, H. 1996. Roles of electrostatic interaction in proteins. *Quart. Rev. Biophys.* 29:1–90.

Nienhaus, G. U., J. R. Mourant, K. Chu, and H. Frauenfelder. 1994. Ligand binding to heme proteins. The effect of light on ligand binding in myoglobin. *Biochemistry.* 33:13413–13430.

Nienhaus, G. U., and R. D. Young. 1996. Protein dynamics. *In* Encyclopedia of Applied Physics, Vol. 15. G. L. Trigg, editor. VCH Publishers, New York. 163–184.

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

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.

Park, K. D., K. Guo, F. Adebodun, 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 vibration 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.

Phillips, G. N., Jr., R. M. Arduini, B. A. Springer, and S. G. Sligar. 1990. Crystal structure of myoglobin from a synthetic gene. *Proteins Struct. Funct. Genet.* 7:358–365.

Potter, W. T., J. H. Hazzard, M. G. Choc, M. P. Tucker, and W. S. Caughey. 1990. Infrared spectra of carbonyl hemoglobins: characterization of dynamic heme pocket conformers. *Biochemistry.* 29:6283–6295.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical Recipes in C. 2nd ed. Cambridge University Press, New York.

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.

Ramsden, J., and T. G. Spiro. 1989. Resonance Raman evidence that distal histidine protonation removes the steric hindrance to upright binding of carbon monoxide to myoglobin. *Biochemistry.* 28:3125–3128.

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

Sage, J. T., D. Morikis, and P. M. Champion. 1991. Spectroscopic studies of myoglobin at low pH: heme structure and ligation. *Biochemistry.* 30:1227–1237.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shimada, H., and W. S. Caughey. 1982. Dynamic protein structures. *J. Biol. Chem.* 257:11893–11900.

Shrager, R. I., and R. W. Hendler. 1982. Titration of individual components in a mixture with resolution of difference spectra, pKs, and redox transitions. *Anal. Chem.* 54:1147–1152.

Shrager, R. I. 1986. Chemical transitions measured by spectra and resolved using singular value decomposition. *Chemom. Intell. Lab. Syst.* 1:59–70.

Springer B. A., and S. G. Sligar. 1987. High-level expression of sperm whale myoglobin in *Escherichia coli. Proc. Natl. Acad. Sci. USA.* 84:8961–8965.

Springer B. A., K. D. Egeberg, S. G. Sligar, R. J. Rohlfs, A. J. Mathews, and J. S. Olson. 1989. Discrimination between oxygen and carbon monoxide and inhibition of autooxidation by myoglobin. *J. Biol. Chem.* 264:3057–3060.

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

Strang, G. 1980. Linear Algebra and Its Applications. 2nd ed. Academic Press, New York.

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

Tian, W. D., J. T. Sage, P. M. Champion, E. Chien, and S. G. Sligar. 1996. Probing heme protein conformational equilibration rates with kinetic selection. *Biochemistry.* 35:3487–3502.

Trotter, H. F. 1984. Eigenvalue distributions of large Hermitian matrices; Wigner's semi-circle law and a theorem of Kac, Murdock and Szegö. *Adv. Math.* 54:67–82.

Wigner, E. 1967. Random matrices in physics. *SIAM Rev.* 9:1–23.

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- and temperature dependence of large-scale conformational transitions in myoglobin. *Chem. Phys.* 158: 315–328.

Zhu, L., J. T. Sage, A. A. Rigos, D. Morikis, and P. M. Champion. 1992. Conformational interconversion in protein crystals. *J. Mol. Biol.* 224: 207–215.