Abstract: The effect of internal and applied external electric fields on the vibrational stretching frequency for bound CO (ν_{CO}) in myoglobin mutants was studied using density functional theory. Geometry optimization and frequency calculations were carried out for an imidazole—iron—porphine—carbon monoxide adduct with various small molecule hydrogen-bonding groups. Over 70 vibrational frequency calculations of different model geometries and hydrogen-bonding groups were compared to derive overall trends in the C—O stretching frequency (ν_{CO}) in terms of the C—O bond length and Mulliken charge. Simple linear functions were derived to predict the Stark tuning rate using an approach analogous to the vibronic theory of activation.\(^1\)

Potential energy calculations show that the strongest interaction occurs for C—H or N—H hydrogen bonding nearly perpendicular to the Fe—C—O bond axis. The calculated frequencies are compared to the structural data available from 18 myoglobin crystal structures, supporting the hypothesis that the vast majority of hydrogen-bonding interactions with CO occur from the side, rather than the end, of the bound CO ligand. The ν_{CO} frequency shifts agree well with experimental frequency shifts for multiple bands, known as A states, and site-directed mutations in the distal pocket of myoglobin. The model calculations quantitatively explain electrostatic effects in terms of specific hydrogen-bonding interactions with bound CO in heme proteins.

Introduction

The electrostatic environment is a key determinant of both binding affinity and reactivity of proteins. A central issue in the current understanding of electrostatic effects is the distinction between specific short-range hydrogen-bonding interactions, on one hand, and long-range interactions from internal fields, on the other. Internal fields arise from a sum of smaller interactions at a distance. Given the long-range nature of Coulombic interactions, internal fields that arise from specific protein geometries may be significant. Enzymatic catalysis and molecular recognition involve highly specific electrostatic interactions that include salt bridges and hydrogen bonds. To address the relative magnitude of these effects, an experimental reporter such as carbon monoxide bound to heme in myoglobin provides a useful probe of the protein electrostatic environment.\(^{1–22}\) The stretching frequency of the CO molecule, ν_{CO}, serves as a probe that can be studied by structural and spectroscopic methods to provide correlations of electrostatic environment and function.\(^{4,19,22}\) The iron—carbon stretching frequency, ν_{Fe—CO}, is correlated with ν_{CO} and provides complementary information on the structure of the distal pocket (Figure 1). In wild-type myoglobin, the CO frequency shifts agree well with experimental frequency shifts for multiple bands, known as A states, and site-directed mutations in the distal pocket of myoglobin. The model calculations quantitatively explain electrostatic effects in terms of specific hydrogen-bonding interactions with bound CO in heme proteins.


Figure 1. Depiction of the iron–imidazole–carbonmonoxophor apoduct studied. In this figure, two hydrogen-bonding groups are shown as an example. The CH3 and CH2N (Im) hydrogen-bonding distances are 2.0 and 1.92 Å, respectively. The angle θ is defined as the Fe−O−H angle.

<table>
<thead>
<tr>
<th>Table 1. Listing of Closest Atoms in Myoglobin Mutant Crystal Structuresa</th>
</tr>
</thead>
<tbody>
<tr>
<td>mutant</td>
</tr>
<tr>
<td>2MGK</td>
</tr>
<tr>
<td>H64-Ne</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>1MNN</td>
</tr>
<tr>
<td>T68-O</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>1MOC</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>T64-C</td>
</tr>
<tr>
<td>2MGC</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>L64-C</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>2MGA</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>H20</td>
</tr>
<tr>
<td>2H2O</td>
</tr>
<tr>
<td>1MLU</td>
</tr>
<tr>
<td>A68-C</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>H2O</td>
</tr>
<tr>
<td>1MYM</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>1YCA</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>H64-Ne</td>
</tr>
<tr>
<td>F43-C</td>
</tr>
<tr>
<td>2MGF</td>
</tr>
<tr>
<td>V68-C</td>
</tr>
<tr>
<td>Q64-N</td>
</tr>
<tr>
<td>Q64-C</td>
</tr>
</tbody>
</table>

a The distances and Fe−O−X angles are reported on the basis of the X-ray crystal structure data.

that the electrostatic effect of, at most, a few amino acid residues is likely to play a major role in the observed frequency shifts of νFe−CO and νCO bands in mutants of myoglobin. In some cases, a single amino acid dominates the interaction because of its proximity to CO. For example, the available evidence indicates that the conformations of the closest residue in wild-type myoglobin, H64, play a dominant role in the hydrogen-bonding environment of CO.25–44 The distribution of Fe−O−X angles, defined in Figure 1 as the angle, θ, is dominated by angles of 110° or less. Out of 60 amino acids listed in Table 1 that contain hydrogen atoms within a radius of 4.5 Å, only 15 have the defined angle θ > 110°, and only nine have θ > 120°, suggesting that there is a preferred geometry of interaction. These considerations indicate that specific hydrogen-bonding interactions are important and, moreover, that there is an orientational preference for hydrogen bonding to CO.

There are four νCO stretching frequencies observed by infrared spectroscopy in many forms of wild-type myoglobin and

hemoglobin. These bands have been designated as A₀, A₁, A₂, and A₃ bands, and it has been suggested that their frequencies depend on different conformational states of myoglobin, hemoglobin, and other heme proteins. However, studies of myoglobin mutants in the H64, V68, F43, F46, and L29 positions often exhibit a single νCO frequency. In particular, among the H64 mutants, only H64A and H64G show evidence for multiple νCO bands. Disordered H₂O in the distal pocket has been hypothesized to lead to a broad set of IR bands. The A₁ and A₃ conformers are observed in wild-type Mb at pH 7.0 at νCO ≈ 1945 cm⁻¹ and a less intense IR band at νCO ≈ 1932 cm⁻¹, respectively. The A₁ conformer is present in a number of different species of myoglobin (MbCO) and hemoglobin (HbCO), but is much smaller in adducts of the H93G mutant of SWMb. The IR band corresponding to the A₀ state at νCO ≈ 1967 cm⁻¹ is observed to grow in intensity as the pH is lowered to pH 4.5. The A₀ state, also known as the "open" state, has a lower binding affinity and a more rapid CO recombination rate. The X-ray crystal structure of a low pH form of myoglobin corresponds to the structure of the A₀ substate. The X-ray crystal structure and single-crystal Raman data of the A₀ conformer are consistent with a protonated H64 that has moved to the surface of the protein so that it is exposed to solvent, which increases the volume of the distal pocket in the A₀ substate. Direct experimental evidence for specific hydrogen-bonding interactions in the A₁ substate has been obtained from deuteration shifts of the νFe–CO Raman band.

The origin of A states of myoglobin categorized by the νCO and νFe–CO stretching frequency of observed IR bands is still in need of an explanation that ties together structure and function. The reversible binding of oxygen essential for the function of myoglobin must occur in competition with endogenous CO. The origin of the discrimination against CO binding has been considered in terms of both an electrostatic13,53 and a steric role. A molecular understanding of the electrostatic effect requires determination of the relative strength of hydrogen bonding as opposed to other internal fields.14,15 Hydrogen bonding by the distal histidine may favor the binding of O₂ over that of CO.15,53,55 In the hydrogen-bonding model, the open state has lower O₂ binding affinity and may represent a physiological trigger for O₂ release.28 Because the νCO frequency is correlated with the open (A₀) and closed (A₁) conformers, it reports on the difference in electrostatic environment that has functional consequences for diatomic ligands other than CO.25 The vibrational Stark effect on νCO demonstrates that frequency shifts can emerge from an internal field.

The vibrational Stark effect arises from an electric field interaction with a change in the CO dipole moment, ∆µCO = µCO(excit.) − µCO(ground), known as the Stark tuning rate. Internal or matrix electric field effects can arise due to specific molecular interactions such as hydrogen bonding of a group CO–H–X, where X = C, N, or O. The concerted motion of several groups can also be viewed as a steric mechanism because the motion is coupled to ß-helices that shift in position upon photolysis. The goal of the present theoretical study is to quantitatively relate these two phenomena. The ansatz of this study is that both the Stark tuning rate and the effect of hydrogen bonding should be connected to structure through the change in the CO bond length.

The present study investigates the effect of an applied electric field and of hydrogen bonding to explain the νCO frequency shifts of CO in mutants and conformational substates of myoglobin. The vibrational frequency of νFe–CO and νCO has been calculated for a range of different hydrogen-bond partners CH₃, NH₃, H₂O, H⁺, imidazole, imidazolium, pyrrole, and combinations of these at angles from 180° to 109° and over a range of distances. Density functional theory (DFT) was used to determine the frequencies. DFT is an appropriate method for the calculation of frequencies and potential energy surfaces. Comparisons have been made between two functionals, the BLYP functional and the GGA functional, as was done in a previous study of proximal effects on the νFe–CO and νCO frequencies. The calculated geometries and Mulliken charges for over 70 structures are compared to obtain a correlation plot that represents the vibronic effect of hydrogen bonding with the Fe–C–O oscillator. The approach is conceptually related to the vibronic theory of activation (VTA), although Mulliken charge is used for the correlation instead of the population of specific molecular orbitals. This study shows the consistency of the hydrogen-bonding hypothesis with vibrational Stark effect measurements, solutions of the Poisson–Boltzmann equation, and the VTA approach to obtain the magnitude of electrostatic effects.

**Methods**

The optimized ground-state geometries were obtained using both the generalized gradient approximation (GGA) of Perdew and Wang and the BLYP functional as implemented in DMol3 (Molecular Dynamics) software.
The procedure for the vibrational frequency calculation is given in
the following. The geometry optimizations were carried out until
the energy difference was less than $10^{-6}$ au on subsequent iterations.
Following geometry optimization, the Hessian matrix, $V$, was con-
structed by finite differences of analytic gradients. The finite differ-
encing proceeds from atom to atom. An energy calculation is carried
out for displacements in the $x$, $y$, and $z$ directions of each atom
in the molecule. $V$ has dimension $3N \times 3N$, where $N$ is the number
of atoms in the molecule. The Cartesian second derivatives in energy
(force constants), which are the elements of $V$, are listed in order of $x_i,
$y_i$ and $z_i$ for each atom $j$ successively. The vibrational frequencies
are obtained by matrix diagonalization of the resulting Hessian matrix in
mass-weighted Cartesian coordinates. Mass weighting is applied by
the transformation $M = M^{-1/2}VM^{-1/2}$, where $M^{-1/2}$ is a diagonal matrix
of dimension $3N \times 3N$, where the diagonal elements are $1/\sqrt{m_i}$ for $j = 1
$to $3N$. The elements of $M^{-1/2}$ are also in order of $x_i$, $y_i$, and $z_i$ for
each atom $j$ successively. The matrix equation for this procedure is
$S^{-1}HS = \Lambda$, where $S$ is the matrix of eigenvector coefficients, and $\Lambda$
is the diagonal matrix of eigenvalues for $j$ modes $\lambda_j$ for $j = 1$ to $3N$.$^{74}$
The dimensions of $S$ are $3N$ rows by $N$ columns. The Cartesian matrix
diagonalization procedure gives six eigenvalues (for a nonlinear
molecule) that correspond to translations and rotations of the center
of mass of the molecule and $3N - 6$ eigenvalues that correspond to
vibrational normal modes. The vibrational frequencies, $\nu$ in units
of wavenumbers (cm$^{-1}$), are obtained from the eigenvalues by
$\nu_j = 4\pi^2 \sqrt{\lambda_j}$. The infrared intensities are calculated using the finite difference
procedure to estimate the difference dipole moment derivative $\partial \mu / \partial Q_j$
along a given normal mode, where the normal coordinates for each of
the $j$ modes correspond to each of three consecutive rows of $S$ ($x_i$, $y_i$, $z_i$, etc.). The transition intensity is obtained within the harmonic
approximation from the square of the vibrational transition moment.$^{75}$
AVS (Advanced Visual System) and insightII (Molecular Simulations)
programs were used for visualization of the results.

Results

The effect of various chemical groups that hydrogen bond to
carboxyl- CO was considered systematically as a function of
distance and angle. The geometry, electronic structure, and
frequencies of the iron–carboxyl–CO system were compared
for four different distal groups, $H_2 O$ (as a model for serine and
threonine), $NH_3$ (as a model for unprotonated lysine), $CH_4$ (as
a model for leucine, isoleucine, and valine), and imidazole (as
a model for histidine). These groups span most of the chemical
interactions of biologically relevant amino acids in heme
proteins. The potential energy of interaction and optimal
hydrogen-bond length for each of the groups are both given in

Figure 2. Hydrogen-bond energy and bond length for four hydrogen-
bonding groups to Im–phosphine–CO as a function of the Fe–O–H bond
angle, $\theta$. (A) The hydrogen-bond energy is shown as calculated using a
potential energy surface relative to a reference energy calculated at a distance
of $10 \text{ Å}$ for each hydrogen-bonding group. (B) The optimum hydrogen-
bonding distance determined from the potential energy surface (PES). The PES
for $CH_4$ is relatively shallow, and thus there is a larger variability in
the minimum position than for the other hydrogen-bonding groups.
Figure 3. Dependence of the Fe–C and C–O bond lengths on the distance and angle of hydrogen-bonding groups. The effects of both H2O and imidazole (Im) are shown for two angles each. The series of points represent distances of 2.0, 2.2, 2.4, 2.6, 2.8, and 3.0 Å from left to right across the figure. Results of both calculations are shown on the same plot. Open and filled circles represent the BLYP calculation for H2O at 180° and 100°, respectively. Open and filled squares represent the GGA calculation for H2O at 180° and 100°, respectively. Open and filled diamonds represent the BLYP calculation for Im at 180° and 110°, respectively. Open and filled bowties represent the GGA calculation for Im at 180° and 110°, respectively.

The Carbonmonoxyl Stretching Band of Myoglobin

1.72 Å\textsuperscript{77} for nonporphine model systems. Porphine–Fe–CO calculations similar to those performed here gave values of $d(Fe–C)$

\begin{equation}
1.72 \text{ Å} \quad \text{and} \quad d(Fe–C) \approx 1.78 \text{ Å}\textsuperscript{80} \text{by DFT methods. The DFT calculated C–O bond length is } d(C–O) \approx 1.17 \text{ Å in a}
\end{equation}

number of other studies.\textsuperscript{76,77} Hydrogen bonding by imidazole (Im) results in calculated bond length changes of $-0.019$ and $+0.006$ Å, respectively, for $d(Fe–C)$ and $d(C–O)$.\textsuperscript{76} These compare well with values of $-0.017$ and $+0.006$ Å obtained in the present study. DFT calculations are also in reasonable agreement with the most recent high-resolution X-ray crystal structure that gives the Fe–C and C–O bond lengths as $d(Fe–C) \approx 1.73 \text{ Å}$ and $d(C–O) \approx 1.12 \text{ Å}$.\textsuperscript{54} A discrepancy in the C–O bond length is the major difference between the DFT calculations\textsuperscript{53,63,75,77} and the X-ray structural data.\textsuperscript{54,78–80} The calculated geometries are in closer accord with the Fe d$_6 \rightarrow$ CO π\textsuperscript{*} back-bonding interaction model that predicts longer bond lengths for iron-bound CO than for free CO commensurate with the trends in the frequency data.\textsuperscript{78,81}

To calculate the effects of hydrogen bonding, geometric constraints were imposed on the geometry optimization procedure. The angle dependence for various hydrogen-bonding groups was determined at a fixed hydrogen-bond distance of 1.92 Å with fixed positions for atoms Fe, O, and H. Thus, the carbon in Fe–C–O–H is free to vary as were all remaining atoms. Details are given in the Supporting Information.

In all of the calculations, the iron–pyrrole-nitrogen (Fe–Np) bond lengths changed little with hydrogen-bonding distance or geometry. In the more than 70 calculated structures considered, the total change in the Fe–Np distance was less than 0.0016 Å for calculations using the GGA functional. The Fe–Np distance is nearly constant for the hydrogen-bonding groups H2O, NH3, and CH4. On the other hand, the changes in the Fe–

$\text{Np distance are nearly 3 times as large for changes in proximal ligation. A recent study of hydrogen bonding to the Nb–H}
\end{equation}

hydrogen found significant core size effects (i.e., effect on Fe–

$\text{Np bond length, decreased.}\textsuperscript{67} \text{Fe–Ne bond lengths are relatively constant for all of the structures studied here except for a decrease of 0.05–0.1 Å in Fe–Ne bond length for all hydrogen-bonding groups for } 100 < \theta < 140° \text{ (results not shown).}
\end{equation}

The effect of hydrogen-bond distance on the Fe–C–O geometry is shown in Figure 3 for H2O and Im interaction with bound CO. The range of CO...H distances studied was 2.0–3.0 Å. Figure 3 shows an inverse relationship between the Fe–C and C–O bond lengths as a function of the hydrogen-bond distance. The effect of both H2O and Im on the C–O bond length is larger for smaller Fe–O...H angles. In summary, based on the GGA calculation, a side hydrogen-bonding interaction results in a 3-fold greater increase in C–O bond length than an end interaction. Numerical values are given in the Supporting Information.

Mulliken Charge. The magnitude of charge displacement from the iron to CO increases as $\theta$ approaches 100° for all of the distal hydrogen-bonding groups studied. The Fe → CO charge displacement was estimated by the sum of the C and O Mulliken charges, $q_C + q_O$. Because the total Mulliken charge $q_C + q_O$ is zero for free CO, the magnitude of this sum is a measure of the net charge displacement from the iron–porphine complex onto bound CO. For the hydrogen-bonding ligands studied here, there is between $-0.02$ and $-0.11$ units of negative charge displaced onto CO. The Mulliken charge sum $q_C + q_O$ grows more negative by $>-0.01$ charge unit for H2O, NH3, CH4, and Im as the angle decreases down to an angle of 140°. For Im and CH4, the largest negative value of Mulliken charge is reached at 100°. For H2O and NH3, the charge displacement onto CO is greatest for $\theta$ = 140°.

Frequencies. The $v_{Fe–CO}$ and $v_{CO}$ stretching frequencies range from 522 to 541 cm$^{-1}$ and 1955–2043 cm$^{-1}$, respectively, for the GGA DFT calculation as shown in Figure 4. The range represents the angle and distance dependence of the principal four hydrogen-bonding groups studied, H2O, NH3, CH4, and Im, in proximity to bound CO. The values of $v_{Fe–CO}$ and $v_{CO}$ for a non-hydrogen-bonded CO are 498 and 2013 cm$^{-1}$, respectively, using the GGA DFT method.\textsuperscript{67} The values are comparable to the $v_{Fe–CO}$ and $v_{CO}$ stretching frequency calculated by other DFT methods. For example, $v_{Fe–CO}$ and $v_{CO}$ in a porphine model system are shifted from 586 and 1907 cm$^{-1}$ to 573 and 1927 cm$^{-1}$, respectively, for an interaction with the lone pair of NH3 as compared to CO alone.\textsuperscript{53} The shift of $v_{Fe–CO}$ to lower values and of $v_{CO}$ to higher values is consistent with lone pair interactions as shown in Table 2. A second comparison with other DFT calculations is provided by the shift of $v_{CO}$ from 1967 cm$^{-1}$ for Fe–CO to 1935 cm$^{-1}$ for $v_{CO}$ with a hydrogen bond to imidazole (Fe–CO...Im).\textsuperscript{67} Using the BLYP functional, the current method gives 464 and 1995 cm$^{-1}$ for $v_{Fe–CO}$ and $v_{CO}$, respectively, without any hydrogen bonding and 515 and 1955 cm$^{-1}$ for a hydrogen bond to imidazole.\textsuperscript{67} Tabulated numerical results that correspond to Figure 4 are given in the Supporting Information.

For all of the hydrogen-bonding groups, the $v_{Fe–CO}$ and $v_{CO}$ stretching vibrations increase in parallel with increasing angle $\theta$ and inversely with increasing distance (Figure 4). The $v_{CO}$
frequency decreases in the series CH₄ > NH₃ > H₂O > Im as can be seen by comparing the frequencies in Figure 4B. This decrease in νCO is expected on the basis of the order of the 
hydrogen-bond strength Im > H₂O > NH₃ > CH₄ (Figure 2).

The effect of hydrogen bonding on the νFe-CO vibration is less pronounced. However, it is generally the case that stronger hydrogen-bonding groups such as H₂O (see Figure 2) have lower νFe-CO frequencies. The values for imidazole show an interesting deviation from this trend. Although imidazole has a stronger interaction with CO than the nonaromatic groups H₂O, NH₃, and CH₄, the νFe-CO frequency of imidazole is closest to that of NH₃. The interaction of the π system of imidazole presumably contributes to the interaction energy with bound CO, but does not affect the νFe-CO frequency as greatly. On the other hand, the νFe-CO and νCO frequencies show a π-back-bonding correlation as a function of the distance of the hydrogen-bonding group at a fixed angle. The νFe-CO and νCO frequencies given in Figure 4 for two different angles, θ, of the hydrogen-bonding 
group Fe–O...H show a negative correlation with the hydrogen-bonding distance.

Table 2 provides frequency data for a number of additional hydrogen-bonding groups including imidazolium and the lone pairs of H₂O and Im. Table 2 shows that imidazole protonation leads to a significant frequency lowering of νCO. Imidazolium hydrogen bonding at a distance of 1.916 Å leads to a lengthening of the C=O bond to >1.17 Å for two different θ angles. For θ = 110°, the Fe–C bond is lengthened to 1.771 Å, leading to a much lower νFe-CO frequency than that in neutral imidazole.

The pyrrole model is a test of the effect of substitution of the Nδ nitrogen by carbon. Table 2 shows that the frequencies of both νFe-CO and νCO are increased for a pyrrole model at θ = 110°, indicative of significantly weaker hydrogen bonding for C=H than for Ne=H of imidazole.

It has been proposed that lone pair interactions increase νCO by virtue of a polarization of CO opposite to that in hydrogen-bonding interactions. As an example of such a lone pair interaction, Table 2 gives the results of a vibrational frequency calculation for the Ne-tautomer of imidazole. In this tautomer, the hydrogen bond is eliminated, and the Nδ lone pair is the only interaction with CO. The consequence is that νCO shifts from ~1959 to ~2023 cm⁻¹ (Table 2). For the interaction of a water lone pair, CO=O_H₂ results in a νCO shift from ~1973 to ~2023 cm⁻¹. For both H₂O and Im lone pair interactions, the increase in νCO frequency is accompanied by a downshift of νFe-CO.
It is apparent from the X-ray crystal structure data given in Table 1 that specific interactions of groups in the distal pocket with bound CO involve multiple interactions. Calculations were carried out for doubly hydrogen-bonded species as well (Table 2). Particular emphasis is placed on double hydrogen-bonding geometries that include CH₄, where methane serves as a model for aliphatic amino acids. These geometries account for the mutants of myoglobin that involve substitution of aliphatic amino acids in the distal pocket. Examination of Table 1 indicates that the majority of close contacts involve C–H hydrogen bonding as one of a number of interactions with bound CO. Table 2 shows that the addition of a CH₄ hydrogen bond at a distance of 2.0 Å to a first hydrogen bond from CH₄, H₂O, Im, or Im⁺ results in a frequency lowering of νₜ₀ by 24, 17, 19, or 23 cm⁻¹, respectively, relative to the singly hydrogen-bound species. The hydrogen bonding of water in the distal pocket can have a profound effect. For a single CH₄ group at a distance of 1.92 Å, νₜ₀ ≈ 2017 cm⁻¹. This frequency shifts down by ~60 cm⁻¹ to a value of ν₂₀ ≈ 1957 cm⁻¹ when an additional H₂O is hydrogen bonded at a distance of 2.0 Å opposite CH₄. A series of double hydrogen-bond geometries consisting of Im⁺ and CH₄ were calculated as models of the “open” or A₀ state of myoglobin and has a protonated histidine. As a result of protonation, the histidine swings out toward the surface of the globin and is located further from CO than neutral imidazole. Because the interaction of imidazolium with bound CO is quite strong, it is reasonable to consider the effect of long-range interactions in this case. As shown in Table 2, the combination of Im⁺ and CH₄ gives rise to relatively low frequencies (and long C–O bond lengths) even when the hydrogen of Im⁺ is as far as 3.8 Å from the oxygen of bound CO.

A Vibronic Approach. The vibronic theory of activation (VTA) is based on the idea that the bonding of bound ligands is altered by charge transfer from a metal center that leads to a change in orbital occupancy. The change in orbital occupancy can be calibrated by calculation of molecule properties (e.g., bond length and frequency) for a small molecule along with its cation and anion. The calculation for the isolated ligand serves to establish the correlation of change in bond length and frequency as charge is injected or withdrawn from the molecule. The vibronic theory has been applied to CO using semiempirical methods. The difference in the present approach is that a change in Mulliken charge is taken as an indicator of the strength of the Fe–CO interaction instead of the occupancy of the 5σ and 2σ⁺ orbitals. While the correlation with occupancy of the 5σ and 2σ⁺ orbitals is of interest, the correlation that emerges in DFT calculations is complicated by the fact that these orbitals are involved in a number of MOs in iron porphine. The current approach employs C–O bond lengths (Figure 3), frequencies (Figure 4), and Mulliken charges given in Table 2 to study CO polarization.

A linear correlation for bond length versus frequency for four species, neutral CO, CO cation, CO anion, and CO excited state, is shown in Figure 5. A linear correlation with frequency is the same as Badger’s rule, because a plot of the bond length versus force constant similar to Figure 5 is still linear. The bond length–frequency correlation obtained from over 70 vibrational frequency calculations in Table 2 and in the Supporting Information is shown in Figure 5. A fit of these calculated values to a line yields the following relation:

\[
ν_{CO} = 2168 - 8677(R_{CO} - R_{CO}^0) \text{ cm}^{-1}
\]  

(1)

where \(R_{CO}\) is obtained from a geometry optimization, and \(R_{CO}^0\) is the equilibrium bond length of neutral CO, \(R_{CO}^0 = 1.141 \text{ Å}\). Figure 5 shows that the slope of the line fit to the calculated metalloporphine CO frequencies is larger than the slope for the unbound CO neutral, cation, anion, and singlet excited state. For this reason, the free CO values are not used for quantitative applications and are shown in Figure 5 only for comparison. The bond length–frequency correlation given in eq 1 can be used to estimate the CO frequency after a geometry optimization of a particular hydrogen-bonding group or groups.

The correlation of the frequency with the Mulliken charge is shown in Figures 6 and 7. The sum of the Mulliken charges on C and O of the CO ligand is shown as \(q_C + q_O\). The linear correlation obtained is as far as 3.8 Å from the oxygen of bound CO.
to the strength of the interaction, this Mulliken method has the same significance, but loses the distinction between $\sigma$- and $\pi$-effects in the original VTA theory. The line in Figure 6 is given by

$$v_{\text{CO}} = \{2094 + 1535(q_C + q_O)\} \text{ cm}^{-1} \quad (2)$$

There are several outlying points that are well below the line. These are the species Im OH$_2$ at $q_C + q_O = -0.064$, $v_{\text{CO}} = 1953.5$ cm$^{-1}$, the Ne tautomer of Im at $q_C + q_O = -0.022$, $v_{\text{CO}} = 2023.2$ cm$^{-1}$, and the dimethane hydrogen-bonded species at $q_C + q_O = -0.042$, $v_{\text{CO}} = 1993.9$ cm$^{-1}$, and other double hydrogen-bonded species. These deviations appear to lie on a second line shifted 30 cm$^{-1}$ to lower energy. If all of the calculated frequencies in this study are included, the slope is 1433 cm$^{-1}$/Mulliken charge, and the intercept is the same.

A third correlation is observed for the polarization of the CO bond. The interaction of Fe with CO removes electron density from the carbon by interaction with the 5o orbital and increases electron density on oxygen by $\pi$-back-bonding to the 2$\pi^*$ orbital. Of course, $\pi$-back-bonding tends to replenish some of the charge on oxygen. Nonetheless, the difference in the Mulliken charges on carbon and oxygen is a measure of this polarization that does not require detailed examination of the occupancies of the 5o and 2$\pi^*$ orbitals in a number of different MOs where they interact with metal, porphine, and trans ligand orbitals. The correlation for the polarization $\Delta q_{\text{CO}} = q_C - q_O$ with frequency is excellent as shown in Figure 7. All of the calculations for neutral hydrogen-bonding ligands that were carried out in this paper are included in Figure 7. The most appropriate fit is to a linear function that allows the frequency to approach the free CO value as $\Delta q_{\text{CO}}$ approaches zero.

$$v_{\text{CO}} = \{2274 - 605\Delta q_{\text{CO}}\} \text{ cm}^{-1} \quad (3)$$

The polarization $\Delta q_{\text{CO}}$ is defined so that it will be a positive number representing the extent of charge displacement within bound CO.

**Calculation of the Stark Tuning Rate.** The effect of applied electric field on the C–O bond length and Mulliken charge was calculated by geometry of optimization of the imidazole–iron–porphine–CO system at four values of an applied electric field perpendicular to the plane of the porphine. The simplified VTA approach described above provides a means to use geometry-optimized structures to correlate distance or Mulliken charge presented in Table 3 with $v_{\text{CO}}$ frequency. The application of VTA is shown in Table 3 for four values of the applied electric field spanning $\pm1$ MV/cm. Using C–O bond distances and Mulliken charges, the $v_{\text{CO}}$ frequency was calculated using eqs 1 and 2. The calculated Stark tuning rates using eqs 1 and 2 are $1.6 \times 10^{-4}$ and $2.6 \times 10^{-4}$ cm$^{-1}/(\text{MV/cm})$, respectively. These values compare extremely well with the experimental value of $2.4 \times 10^{-4}$ cm$^{-1}/(\text{MV/cm})$.

$^4$The Stark tuning rate for free CO is calculated to be $0.5 \times 10^{-3}$ cm$^{-1}/(\text{MV/cm})$, which is also in reasonable agreement with experiment. The bond-length frequency correlation is validated by this comparison.

**Discussion**

The DFT calculations provide a set of frequencies and intensities that span conformational and chemical interactions relevant to the CO bands of heme proteins. Hydrogen bonding in the distal pocket emerges as a dominant factor that gives rise to an electrostatic effect on the $v_{\text{Fe–CO}}$ and $v_{\text{CO}}$ frequencies. The A states of wild-type Mb and the effects of mutation can be categorized on the basis of the calculated frequency shifts in $v_{\text{CO}}$ presented in Table 2 and further studied using the VTA approach based on the geometry of Table 1.

The **Hydrogen-Bonding Model for v_{CO} Frequency Shifts.** The hypothesis of the current study is that specific hydrogen-bonding environments are responsible for the observed electrostatic effects on $v_{\text{CO}}$. The calibration of the calculated effect using an applied electric field, the Stark tuning rate, in terms of a change in C–O bond length presents a comprehensive view of the origin of the vibrational Stark effect. Current theories of the vibrational Stark effect are based on the anharmonicity of the potential energy surface$^{66-68,85,86}$ and a bond polarization term.$^{38}$ The current approach is based on the field-induced perturbation of the geometry. Anharmonic terms have been shown to be consistently too small to account for the observed Stark tuning rate.$^{58}$ According to the DFT calculations, the small change in bond length of $3 \times 10^{-4}$ Å/(MV/cm) (Table 3) is the primary effect giving a calculated Stark tuning rate of $2.5 \times 10^{-3}$ cm$^{-1}/(\text{MV/cm})$ using eq 1. Given a structural Stark shift of $3 \times 10^{-4}$ Å/(MV/cm), an internal field acting at a distance within the protein the magnitude of the field would require $10^7$ V/cm to give the frequency shift between the A$_0$ and A$_1$ substate of $\Delta v_{\text{CO}} \approx 22$ cm$^{-1}$.

The **$\pi$-Back-Bonding Correlation.** The $\pi$-back-bonding correlation observed in many heme proteins consists of an inverse correlation between the bands, $v_{\text{Fe–CO}}$ and $v_{\text{CO}}$. The $\pi$-back-bonding correlation for distal hydrogen-bonding groups holds only if the Fe–Np bond length is nearly constant in a series of structures. Equatorial $\pi$-back-bonding competes with axial $\pi$-back-bonding and complicates the correlation as shown in a previous study of proximal effects.$^{67}$ In the present study, the change in Fe–Np bond length is less than $1.6 \times 10^{-3}$ Å. The resulting small change in the core size has little effect on

---

Table 3. Effect of an Applied Electric Field on Bond Length and Mulliken Charge Using the GGA Functional\textsuperscript{a}

<table>
<thead>
<tr>
<th>MV/cm</th>
<th>Fe-C</th>
<th>C-O</th>
<th>q\textsubscript{Fe}</th>
<th>q\textsubscript{C}</th>
<th>q\textsubscript{Fe} + q\textsubscript{C}</th>
<th>q\textsubscript{Fe} - q\textsubscript{C}</th>
<th>V\textsubscript{CO eq 1}</th>
<th>V\textsubscript{CO eq 2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>1.7864</td>
<td>1.1569</td>
<td>0.190</td>
<td>-0.237</td>
<td>-0.047</td>
<td>0.427</td>
<td>2021.4</td>
<td>2021.8</td>
</tr>
<tr>
<td>-0.5</td>
<td>1.7865</td>
<td>1.1572</td>
<td>0.190</td>
<td>-0.238</td>
<td>-0.048</td>
<td>0.428</td>
<td>2018.7</td>
<td>2020.3</td>
</tr>
<tr>
<td>0</td>
<td>1.7861</td>
<td>1.1572</td>
<td>0.190</td>
<td>-0.239</td>
<td>-0.049</td>
<td>0.429</td>
<td>2018.7</td>
<td>2018.9</td>
</tr>
<tr>
<td>0.5</td>
<td>1.7861</td>
<td>1.1573</td>
<td>0.190</td>
<td>-0.240</td>
<td>-0.050</td>
<td>0.430</td>
<td>2016.2</td>
<td>2017.3</td>
</tr>
<tr>
<td>1</td>
<td>1.7858</td>
<td>1.1575</td>
<td>0.190</td>
<td>-0.240</td>
<td>-0.050</td>
<td>0.430</td>
<td>2016.2</td>
<td>2017.3</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Columns 7 and 8 are calculated values of the CO stretching frequency based on the VTA correlations in the text. The equations used are eqs 1 and 2, respectively. Geometry optimizations were carried out at applied electric fields in the range of ±2 × 10\textsuperscript{-4} au. Using the conversion for applied electric field, 10\textsuperscript{4} au = 0.5142 MV/cm, to a good approximation the range is ±1 MV/cm as reported in the table.

The orbital symmetry helps to explain why the \( \pi \)-back-bonding effect is so much larger in the case of side-on hydrogen bonding than for end hydrogen bonding. The interaction of the hydrogen-bond group with the \( \sigma \)-bonding group depends on the electronegativity of the element X in the X–H donor. Figure 4 shows that there are two regimes. Near 100°, \( \nu_{\text{Fe-CO}} \) and \( \nu_{\text{CO}} \) have an inverse relationship, and near 180°, \( \nu_{\text{Fe-CO}} \) and \( \nu_{\text{CO}} \) both increase in parallel in the order CH\textsubscript{4} > NH\textsubscript{3} > H\textsubscript{2}O. The parallel change in \( \nu_{\text{Fe-CO}} \) and \( \nu_{\text{CO}} \) frequency in this series arises from a change from \( \pi \)-type bonding at angles near 90° to \( \sigma \)-type bonding at angles near 180°. The crossover occurs near 140°.

The Correlation of Structure and Spectroscopy in A\textsubscript{0} and A\textsubscript{1} States. The multiple CO stretching bands observed in many myoglobins and hemoglobins can be explained by comparison of imidazole and imidazolium hydrogen bonding calculated by the DFT methods in this study. On the basis of the wild-type myoglobin X-ray crystal structure (2MGK), the Im–H distance of H64 is <2.1 Å, and there are simultaneous hydrogen bonds from V64 to 2.2 Å and F43 at 2.6 Å.\textsuperscript{35} The model calculation for the two hydrogen-bond groups, CH\textsubscript{4} and Im (Table 2, row 30), has a calculated frequency of \( \nu_{\text{CO}} = 1939 \text{ cm}^{-1} \). The experimental values for the largest (A\textsubscript{1}) substate of myoglobin are \( \nu_{\text{Fe-CO}} \approx 507 \text{ cm}^{-1} \) and \( \nu_{\text{CO}} \approx 1943 \text{ cm}^{-1} \). The A\textsubscript{0} state is represented by X-ray crystal structures 1VXA, 1VXC, and 1VXF at pH 4.0, 5.0, and 6.0, respectively.\textsuperscript{37} In these structures, H64 has swung out of the distal pocket to varying degrees and is nearer to the surface of the globin.\textsuperscript{37} The pK\textsubscript{a} of H64 is 6.0, which suggests that H64 is in the imidazolium form in the low pH structures.\textsuperscript{66} The distance of the imidazolium ImH\textsuperscript{+} from the CO is greater than that of Im in wild-type Mb.\textsuperscript{37} The geometries in these structures were modeled in DFT calculations given in Table 2 (rows 33–36) for a fixed CH\textsubscript{4} distance representing V68 and a variable ImH\textsuperscript{+} distance. The calculations in Table 2 show that the \( \nu_{\text{CO}} \) frequency varies from 1897 to 1971 cm\textsuperscript{-1} as the ImH\textsuperscript{+} hydrogen-bonding distance increases from 2.0 to 3.8 Å. The experimental frequencies for the A\textsubscript{0} substate are \( \nu_{\text{Fe-CO}} \approx 493 \text{ cm}^{-1} \) and \( \nu_{\text{CO}} \approx 1676 \text{ cm}^{-1} \). The calculated \( \nu_{\text{Fe-CO}} \) frequencies corresponding to A\textsubscript{0} and A\textsubscript{1} substates are 508 and 535 cm\textsuperscript{-1}, in agreement with the experimental trend that A\textsubscript{1} is larger than A\textsubscript{0}. The DFT calculations are consistent with a weak interaction of imidazolium with bound CO. The best agreement is found for a CO–H distance of 3.4 Å, which is long enough to no longer consider a hydrogen bond. The corresponding 4.4 Å distance between N6 of ImH\textsuperscript{+} and the oxygen of CO is approximately 1.2 Å larger than that observed in the pH 5.0 X-ray structure. It is possible that solvation effects change the structure slightly as the pH is lowered or that experimental data are affected by the equilibrium between the protonated and unprotonated forms.

In summary, the model gives a reasonable representation of the type of \( \nu_{\text{CO}} \) frequency shift that is expected for a structural change associated with A\textsubscript{0} and A\textsubscript{1} conformers.

The Effect of Mutation on the \( \nu_{\text{CO}} \) Frequency. The results of the VTA correlation (eqs 1–3) can be used to address specific hydrogen-bonding geometries in mutants of myoglobin. A test of the VTA model is carried out by geometry optimization of each of the structures considered as a model for the mutant in question. The geometries and Mulliken charges and \( \nu_{\text{CO}} \) frequencies for three examples are given in Table 4. Mutations in the H64 position have a profound effect on the \( \nu_{\text{CO}} \) stretching frequency. The calculations suggest that a combination of CH\textsubscript{4} and H\textsubscript{2}O hydrogen bonding gives rise to a frequency in the range observed, that is, \( \nu_{\text{CO}} \approx 1957 \text{ cm}^{-1} \) (Table 2, rows 31 and 32). This value of \( \nu_{\text{CO}} \) is in the correct range for H\textsubscript{2}O hydrogen
Table 4. Models for Myoglobin Mutants Based on the Vibronic Theory of Activation a

<table>
<thead>
<tr>
<th>mutant/model</th>
<th>C–O</th>
<th>q_c</th>
<th>p_i</th>
<th>q_c + p_i</th>
<th>q_c - p_i</th>
<th>v_CO (eq 1)</th>
<th>v_CO (eq 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H64L</td>
<td>1.163</td>
<td>0.239</td>
<td>-0.282</td>
<td>-0.043</td>
<td>0.521</td>
<td>1967.5</td>
<td>1967.9</td>
</tr>
<tr>
<td>CH_3 CH_3 CH_4</td>
<td>1.157</td>
<td>0.220</td>
<td>-0.242</td>
<td>-0.022</td>
<td>0.462</td>
<td>2017.9</td>
<td>2000.2</td>
</tr>
<tr>
<td>H64L/V68T</td>
<td>1.164</td>
<td>0.225</td>
<td>-0.288</td>
<td>-0.063</td>
<td>0.513</td>
<td>1963.2</td>
<td>1967.3</td>
</tr>
<tr>
<td>CH_3 CH_3 OH_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH_3 CH_3 CONH_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a The distances and Mulliken charges obtained from a geometry optimization are used to calculate the v_CO frequency for comparison with experiment.

We equate the change in the sum of the square of the occupation numbers with the absolute value of the change in Mulliken charge Δp^2 = |q_c + q_o|. The vibronic model becomes

\[
\nu = \nu_0 \sqrt{1 - \frac{|q_c + q_o|}{v_CO}}^2 \nu_CO \Delta
\]

where f_Co is the vibronic matrix element for an iron CO charge transfer state. This model is equivalent to the empirical fit to a line (eq 2) because an expansion of eq 2 yields

\[
\nu = \nu_0 \left(1 - \frac{f_Co^2}{2v_CO^2} |q_c + q_o| \right)
\]

and the slope is f_Co^2/2ν_COΔ. Alternatively, eq 6 can be used to fit the data in Figure 6, which yields a value of νo = 2091 cm^−1 and f_CO^2/ν_COΔ = 1.4. Using the parameters of the fit of the values in Figure 6 in eq 7, the frequency and slope in a linear fit are estimated to be 2091 and 1464 cm^−1 as compared to 2093 and 1535 cm^−1 obtained using eq 2. The value of χ^2 for the fits differs by only 1%

Conclusion

The effect of hydrogen bonding on the vibrational frequency of CO in iron–porphine adducts has been compared to the effect of an applied electric field, the vibrational Stark effect. The magnitude of the applied electric field is consistent with a correlation of frequency with bond length and Mulliken charge on bound CO obtained from a large number of DFT calculations that include many different hydrogen-bonding environments. The calculations have been interpreted with a VTA approach that unites the experimentally observed vibrational Stark effect and observed hydrogen-bonding shifts on ν_CO. A change in the C–O bond length occurs both for hydrogen bonding and in the presence of an applied electric field. A quantitative interpretation of the dependence of the bond length on an externally applied field provides a structural interpretation for the Stark tuning rate. The calculations are consistent with the observation that hydrogen bonding occurs mainly from the side at Fe–⋯H angles of θ = 100–120°. The π-back-bonding correlation is maintained for changes in distance at a fixed angle θ and to a certain extent even for chemical changes in the hydrogen-bonding group for the same geometry. The calculations provide further evidence that specific hydrogen-bonding interactions dominate the electrostatic effects observed in the vibrational spectra of heme proteins. In particular, the dominant effect on the ν_CO frequency in myoglobin arises from specific hydrogen-bonding effects.

Acknowledgment. S.F. acknowledges support through NSF grant MCB-9874895 and a grant from the North Carolina

The Carbonmonoxyl Stretching Band of Myoglobin

Supercomputer Center (NCSC). Thanks is given to Lee Bartolotti of the NCSC for many helpful suggestions.

Supporting Information Available: Information on the basis set. Detail information on energies, frequencies, and Mulliken charge as a function of hydrogen-bond geometry (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.