Bound CO Is A Molecular Probe of Electrostatic Potential in the Distal Pocket of Myoglobin

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · OCTOBER 1999

Impact Factor: 3.3 · DOI: 10.1021/jp9918205

5 AUTHORS, INCLUDING:

George N Phillips
Rice University
315 PUBLICATIONS **12,953** CITATIONS
SEE PROFILE

John S. Olson
Rice University
246 PUBLICATIONS **12,300** CITATIONS
SEE PROFILE

FEATURE ARTICLE

Bound CO Is A Molecular Probe of Electrostatic Potential in the Distal Pocket of Myoglobin

George N. Phillips, Jr.,* Miguel L. Teodoro, Tiansheng Li, Benjamin Smith, and John S. Olson

W. M. Keck Center for Computational Biology and Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77005-1892

Received: June 3, 1999; In Final Form: July 26, 1999

Most recent experiments have indicated that distal pocket polarity rather than steric hindrance is the major factor governing the distribution of FeCO stretching frequencies (ν_{C-O}, ν_{Fe-CO}) in myoglobins and hemoglobins. Hydrogen bonding and other polar interactions have also been shown to play a key role in regulating O_2 and CO binding. To quantify the effects of polarity on ν_{C-O}, ν_{Fe-CO}, and ligand binding, we calculated electrostatic potential field distributions in the distal pockets of 18 different mutants and two wild-type forms of recombinant pig and sperm whale MbCO. The results were obtained using linearized Poisson−Boltzmann methods with coordinates from high-resolution structures determined experimentally by X-ray crystallography. The computed potential fields at the ligand atoms vary from +30 to −12 kcal/mol depending on the protein structure at the distal site. The electrostatic fields correlate inversely with ν_{C-O} and directly with ν_{Fe-CO}. In all our calculations, the distal histidine is modeled as the neutral Nϵ−H tautomer, regardless of which ferrous ligand is bound. If the neutral Nδ−H tautomer is used, the computed potentials at the bound ligand atoms are uniformly negative and show no correlation with ν_{C-O}, ν_{Fe-C}, and any ligand binding parameters. Although calculated using primarily MbCO structures, there is a linear, inverse relationship between the electrostatic field at the ligand binding site and the logarithm of the rate constant for O_2 dissociation. As a result, high O_2 affinity can be predicted semiquantitatively from a large positive potential field or from an experimentally low value of ν_{C-O}. Thus, the stretching frequency of bound CO serves as an empirical voltmeter that can be used to measure the polarity of the distal pocket and to predict the extent of electrostatic stabilization of bound O_2.

Introduction

For the past 10 years, the reactions of O_2, CO, and NO with myoglobin have been investigated systematically using a combination of site-directed mutagenesis, X-ray crystallography, and IR, RR, and NMR spectroscopic methods (for reviews see Springer et al.[1] and Phillips and Olson[2,3]). The results of these efforts have had a major impact on the general field of molecular biophysics. First, specific chemical mechanisms have been developed to explain how protein residues in the vicinity of bound ligands regulate O_2, CO, and NO affinity and various oxidative reactions. These mechanisms are now being used to design more efficient and safe recombinant heme proteins for use as O_2 delivery pharmaceuticals.[4,5] Second, the ideas and concepts developed from the myoglobin studies are being used to interpret the functional, structural, and spectroscopic properties of many other heme-containing proteins, including oxidases, peroxidases, and guanylyl cyclase.[6−8] Third, myoglobin serves as a well-characterized experimental system for testing the validity and utility of a wide variety of computational algorithms involving molecular mechanics and dynamics.[9−12] For example, quantum mechanical calculations have been used to estimate the effects of iron back-bonding, hydrogen bonding, and distortion on several heme−ligand coordination systems.[13−15]

Because of its triple bonding character, bound CO absorbs strongly in a region of the IR spectrum far from that of the rest of the protein and water, and its stretching frequency, ν_{C-O}, is readily measured.[16] As a result, ν_{C-O} has been used extensively to characterize different conformers of the bound state in MbCO[7,17−20] and substates of the photodissociated ligand.[21−23] The equilibrium IR absorbance spectrum of sperm whale MbCO shows four major ν_{C-O} peaks that were originally defined as A_0 (1965 cm^{-1}), A_1 (1947 cm^{-1}), A_2 (1942 cm^{-1}), and A_3 (1932 cm^{-1}).[17] The relative proportions of these peaks can be changed markedly by temperature, pH, or solvent conditions.[18] At neutral or slightly alkaline pH, the major band is derived from a combination of the A_1 and A_2 conformers, has a peak at 1945 cm^{-1}, and is normally designated simply as the A_1 state (see Figure 1).

Until recently, steric hindrance was thought to play an important role in determining the stretching frequency of bound CO; however, most recent work has suggested that polarity is the key determinant. By analyzing various mutants of sperm whale and pig myoglobin, Li et al.[24] concluded that electrostatic fields generated by polar distal pocket amino acids alter the electron distribution in the iron−carbonyl complex, changing the order of the CO bond and, accordingly, its IR stretching

* Corresponding author. Address: Department of Biochemistry-MS 140, Rice University, Houston, TX 77005-1892. Tel: (713)527-4910. Fax: (713)285-5154. E-mail: georgep@rice.edu.

10.1021/jp9918205 CCC: $18.00
Published on Web 09/30/1999

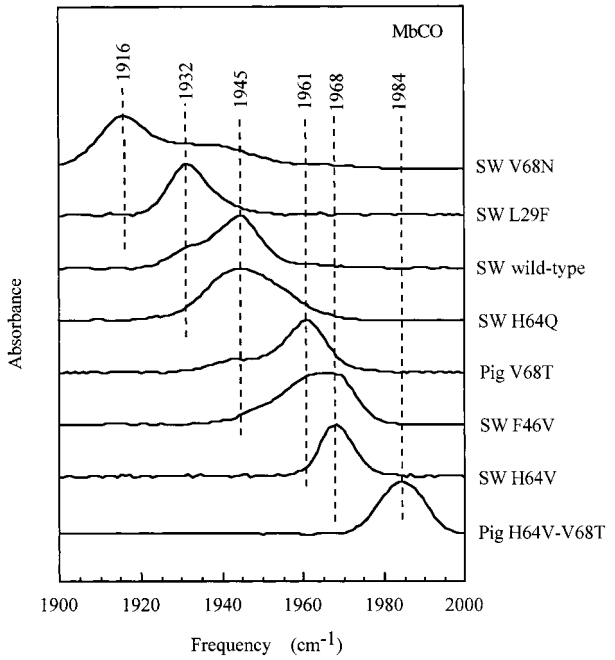

Figure 1. Iron−carbonyl IR spectra in the region of 1900−2000 cm^{-1} for sperm whale (SW) and pig myoglobins. The mutant proteins are designated by standard single letter abbreviations for the native residue, followed by the position in the primary sequence and the abbreviation for the mutant amino acid. The spectra are arranged from top to bottom in order of increasing frequency for the major peak. This order corresponds to decreasing electrostatic potential at the oxygen atom position of the bound CO ligand (Table 1).

frequency. In contrast, large bulky residues that cause marked decreases in CO affinity due to steric hindrance of the bound ligand have little effect on the IR spectrum. The results with mutant myoglobins confirmed the conclusions of Ray et al.[7] who showed that polar interactions rather than steric hindrance cause large variations in ν_{C-O} and ν_{Fe-CO} for a wide variety of model heme compounds and proteins. Stavrov, Ghosh, Spiro, and co-workers[13−15] have shown convincingly that the observed variations in IR spectra cannot be explained by geometrical bond distortions. Despite these extensive studies, the question of steric hindrance versus electrostatic effects remains a point of contention,[25] and on the basis of their 1.15 Å structure of native sperm whale MbCO, Kachalova et al.[26] have resurrected the hypothesis that steric hindrance is the main factor causing discrimination against CO binding.

The next logical step for evaluating the importance of polarity is to compute electrostatic fields in the distal pockets of wild-type and mutant myoglobins and then correlate the potentials quantitatively with both spectroscopic and ligand binding properties. We selected 20 different pig and sperm whale recombinant myoglobins for analysis based on large variations in spectroscopic and functional properties and the existence, in most cases, of high-resolution crystal structures. The ν_{C-O} and ν_{Fe-CO} values range from 1916 to 1984 and 478 to 526 cm^{-1}, respectively, and the K_{O_2} and K_{CO} values range from 15 to 0.02 and 1000 to 1 μM^{-1}, respectively (Tables 1−3).

Linearized Poisson−Boltzmann methods for calculating electrostatic fields in proteins have become standard and are now commonly displayed on molecular surfaces of structures determined by X-ray crystallography.[27] Calculation of fields inside proteins is a more recent application, and many papers describing the role of electrostatics in function have appeared.[28] In this work electrostatic potentials at the carbon and oxygen positions of bound CO were calculated using the focusing methods of Gilson et al.[29] The results were then analyzed by following the approach of Gilson and Honig,[30] which uses reciprocity to break down individual contributions to the potential at single sites in a macromolecule.

One key issue in computing electrostatic fields in the vicinity of the bound ligand is the tautomeric state of the neutral imidazole side chain of His64. Either a Nϵ hydrogen atom or the lone pair of electrons can point toward the FeCO complex, and clearly this choice will have a large effect on the electrostatic potential in the distal pocket. Unfortunately, X-ray crystallography does not "see" hydrogen atoms well, and protonation of the Nϵ versus Nδ atom must be determined or inferred by other methods. For methylimidazole in solution, there is no strongly preferred tautomer, and the equilibrium constant for Nϵ−H to Nδ−H isomerization is not far from unity in the absence of a protein environment.[31] The situation can be further complicated by the ability of the imidazole to rotate 180° about the Cβ−Cγ bond so that, in theory, either Nϵ or Nδ could point toward the ligand. However, in the high-resolution electron density maps of myoglobin reported by Berendzen and co-workers (PDB files 1A6G, 1A6K, 1A6M, and 1A6N), the ϵ position of the imidazole ring is pointing toward the iron atom, implying that the inward Nϵ conformer is preferred regardless of ligand and oxidation state. Their work has recently been confirmed by Kachalova et al.[26] NMR studies are available only for the cyanomet form of myoglobin but also show that the proton on the Nϵ atom is adjacent to the bound ligand.[32,33]

Using neutron diffraction crystallography, Schoenborn and colleagues have claimed to "see" the hydrogen atom on the outward-facing Nδ atom of His64 in CO−myoglobin and on the inward facing Nϵ atom in oxymyoglobin.[34−36] However, the effects of mutagenesis on ligand binding and FeCO stretching frequencies suggest equally strongly that the Nϵ−H tautomer occurs in both MbCO and MbO_2.[14,24] To resolve this controversy, we computed electrostatic potentials for both tautomeric states of the distal imidazole and then examined how the results correlate with ν_{C-O}, ν_{Fe-CO}, and rate constants for ligand dissociation from all of the proteins containing a distal histidine.

Methods

X-ray Structures, IR and RR Spectroscopy, and Ligand Binding Parameters. Atomic coordinates for the proteins used in this study were taken from the Brookhaven Protein Data Bank.[37,38] IR and RR spectroscopy results were taken from Li et al.,[24] Ling et al.,[39] Zhao et al.,[40] Anderton et al.,[41] Tomita et al.,[42] and Vogel et al.[43] Ligand binding parameters were determined as described in Rohlfs et al.[44] and taken from the appropriate primary references listed in Tables 2 and 3.

Electrostatic Calculations. Unless otherwise noted, the starting points for all calculations were actual high-resolution crystal structures (better than 2 Å). Hydrogen atoms were placed according to minimum energy configurations using the program XPLOR [45] and the topology and parameter files *topalh22x.pro* and *parallh22x.pro*. The following energies were included in the hydrogen placement algorithm: bond angle, covalent bond, improper dihedral angle, intramolecular electrostatic, and intramolecular van der Waals energies. For the HBUILD statement, the cutoff value for acceptors involving hydrogen bonds to waters was set at 7.5 Å and the stepsize for the dihedral angle search was set to 4.0°. Because of controversy surrounding the protonation state of the distal histidine residue, two sets of PDB files were created for each mutant with a distal histidine, one with Nϵ protonated His64 and one with Nδ protonated His64.

TABLE 1: Correlation between Calculated Electrostatic Potential and Area Averaged (av) IR and Raman Stretching Frequencies for Sperm Whale and Pig Myoglobin Mutants

pdb access 39 code	mutation	ligand	potential[a] at O position Nϵ−H (kcal/mol)	potential[a] at C position Nϵ−H (kcal/mol)	potential[a] at C position, Nδ−H (kcal/mol)	potential[a] at C position, Nδ−H (kcal/mol)	av ν_{CO} (cm^{-1})	av ν_{Fe-CO} (cm^{-1})	refs for (ν_{CO}, ν_{Fe-CO})
1MLQ[b]	V68N	C−O	29.31	16.83	−6.13	−6.71	1922	526	24, 41
2SPL	L29F	C−O	15.16	8.20	−8.18	−8.16	1932	525	24, 40
1MLQ	V68L	C−O	19.28	9.51	−15.93	−13.77	1938	n.d.[d]	24
1MLM	V68I	C−O	12.32	7.79	−10.70	−11.85	1938	n.d.	24
1MCY	H64Q, L29F	C−O	14.31	8.24			1938	513	24, 40
1MLJ	V68F	C−O	15.98	9.59	−13.69	−12.05	1940	n.d.	24
2MGK	Wild-type	C−O	13.54	8.41	−8.86	−10.50	1941	508	24, 39
1MLF	V68A	C−O	14.70	10.08	−11.59	−12.73	1943	n.d.	24
2MGF	H64Q	C−O	10.04	4.93			1945	507	24, 40
1YCA	V68T	C−O	5.98	3.23	−11.04	−12.09	1958	493	24, 41
1MYM	F46V	C−O	−0.029	−3.29	1.24	−2.32	1962	496	24, 41
1MLU	H64G, V68A	C−O	3.19	−0.39			1964	n.d.	68
2MGA	H64G	C−O	1.66	−1.37			1965	492	24, 68
2MGC	H64L	C−O	3.04	−0.28			1965	490	24, 39
1MOC	H64T	C−O	3.01	−0.88			1966	n.d.	24
1MNK[c]	H64V, V68T	C−O	−11.78	−9.93			1984	478	52, 41

[a] The electrostatic potentials were calculated at the O and C atoms of bound carbon monoxide using either the Nϵ−H or Nδ−H tautomer for the distal histidine when it was present. The potentials for mutants without a distal histidine are only listed in the columns for the Nϵ−H tautomer. [b] The sperm whale V68N MbCO structure was modeled as described in the text. [c] The pig H64V/V68T MbCO structure was also modeled. [d] n.d.: not determined

TABLE 2: Correlation between Calculated Electrostatic Potential, Average (av) IR Stretching Frequency, and Kinetic Parameters for NO−Myoglobin Mutants

mutation	ligand in crystal	potential at O position Nϵ−H (kcal/mol)	av ν_{N-O}[a] (cm^{-1})	k_{NO}[b] (s^{-1})	K_{NO} (pM^{-1})
V68N(SW model)	C−O	29.31	n.d.	0.00022	0.13
V68N(Pig obs)	O−O	24.65			
L29W	O−O	17.27	1615, 1612	0.00017	0.0094
L29F	C−O	15.16	1601	0.000021	1.3
	O−O	15.57			
V68L	C−O	19.28	n.d.	0.00015	0.31
V68I	C−O	12.32	n.d.	0.00073	0.036
H64Q, L29F	C−O	14.36	1614	0.000083	0.55
V68F	C−O	15.98	1605	0.000026	0.073
wild-type	C−O	13.54	1613	0.00010	0.22
	O−O	17.39			
V68A	C−O	14.70	1617	0.00017	0.24
H64Q	C−O	10.04	1619	0.00011	0.39
V68T	C−O	5.98	1632, 1628	0.00066	0.012
F46V	C−O	−0.29	1623	0.00090	0.073
H64G	C−O	1.66	1633	0.00080	0.28
H64L	C−O	3.04	1635, 1636	0.00013	1.5
H64T	C−O	3.01	n.d.	0.0013	0.25
H64V, V68T	C−O	−11.78	1631	0.0071	0.141
(Pig models)	O−O	−6.97			

[a] The first stretching frequencies for NO bound to H64G, H64L, and L29W were taken from Tomita et al.,[42] and the remaining values were taken from Vogel et al.[43] [b] The rate and equilibrium constants for NO binding were taken from Eich and co-workers[69,70] and unpublished data from T. Li, Y. Dou, R. F. Eich, and J. S. Olson (H64G, H64T, and H64V/V68T).

Following placement of hydrogen atoms, the electrostatic potentials at the carbon and oxygen positions of the CO ligand were calculated using the program UHBD (University of Houston Brownian Dynamics). For a detailed discussion of electrostatics calculations using UHBD, the reader should consult Davis et al.[46] and Madura et al.[47] The potential at the atom of interest is calculated by first solving the linearized Poisson−Boltzmann equation with only the atom of interest assigned a unit charge. The equation is then solved by "focusing", with calculations carried out iteratively at 1.0 Å intervals, then 0.5 Å intervals, and then 0.2 Å intervals. Because each calculation uses a 65 × 65 × 65 point grid, the $n - 1$ calculation supplies the boundary condition for the nth calculation. The parameters for all UHBD calculations were set as follows: protein dielectric constant = 2, solvent dielectric constant = 80, ionic strength = 150 mM, Stern layer thickness = 2 Å, maximum number of iterations = 200, nmap = 1.4, and nsph = 100. After solving the linearized Poisson−Boltzmann equation for the source atom, the charge of the source was then set to 0 and the charges of the protein amino acids were set according to the charmm22 parameter file. The electrostatic potential, E, at the location of the source atom was then calculated using

$$E = \sum_{i=1}^{N} q_i \phi_i \qquad (1)$$

where q_i is the charge of atom i, ϕ_i is the potential at atom i that is due to the source, and N is the number of atoms comprising the myoglobin structure.[48] The charges of the heme group were not included in calculating the electrostatic potentials used for comparison with observed stretching frequencies.

TABLE 3: Correlations between Calculated Electrostatic Potentials, Rate Constants for O_2 and CO Dissociation (k_{O_2} and k_{CO}), and Association Equilibrium Constants (K_{O_2} and K_{CO})[a]

mutation	ligand in crystal	potential at O position, Nϵ−H (kcal/mol)	k_{O_2} (s^{-1})	K_{O_2} (μM^{-1})	k_{CO} (s^{-1})	K_{CO} (μM^{-1})	ref for ligand binding
V68N(SW model)	C−O	29.31	0.54	3.5	0.0096	4.3	56
V68N(pig obs)	O−O	24.65					
V68N(pig obs)*	C−O	4.58					
L29W	O−O	17.27	8.5	0.029	0.008	0.48	59
L29F	C−O	15.16	1.4	15.0	0.006	37	71
	O−O	15.57					
V68L	C−O	19.28	6.8	3.4	0.011	48	1
V68I	C−O	12.32	14	0.22	0.24	2.1	1
H64Q, L29F	C−O	14.36	65	0.46	0.006	85	40
V68F	C−O	15.98	2.5	0.48	0.018	14	1
wild-type	C−O	13.54	15	1.1	0.019	27	1
	O−O	17.39					
V68A	C−O	14.70	18	1.2	0.021	56	1
H64Q	C−O	10.04	130	0.18	0.012	82	1
V68T	C−O	5.98	39	0.072	0.079	7.5	72
F46V	C−O	−0.29	100	0.20	0.050	7.0	64
H64G, V68A	C−O	3.19	94	1.20	0.014	3800	73
H64G	C−O	1.66	1600	0.090	0.038	150	1
H64L	C−O	3.04	4100	0.023	0.024	1100	1
H64T	C−O	3.01	6400	0.017	0.045	150	1
H64V, V68T	C−O	−11.78	4000	0.025	0.063	430	74
(pig model)	O−O	−6.97					

[a] The distal histidine in the low-temperature structure of pig V68N MbCO is in the out conformation with a $\Delta\chi 1$ angle >60°. This is the cause for the smaller electrostatic field in the vicinity of the CO atoms.

Numerical artifacts are caused by the high charge density around the heme and add significant noise to the calculations. Leaving out the heme charges is equivalent to saying that the contribution of the heme to the field is the same in all the distal pocket mutants studied and hence should not have a significant effect on the electrostatic field differences between them.

The sensitivity of the calculation to the value of the protein dielectric constant was briefly examined by repeating the calculation using values between 1 and 15. Only a very weak dependence of the correlation values was observed.

To gain more visual understanding of the effect of distal pocket amino acid mutations, UHBD was also used to calculate the complete electrostatic potentials in the active site of each protein. The method of focusing was employed as described above, with the exception that the heme charges were used and set to charmm22 values. In this case the field vector has the correct directionality, even if the magnitude is less accurate numerically. The potential grid was visualized using the software package GRASP,[49] which calculates and draws field lines within protein structures. The contributions from each atom used in eq 1 were visualized using the program MOLMOL.[50]

Computational Modeling. Actual crystal structures are available for most of the mutants used in the electrostatics calculations. The structures of two CO complexes were not available: sperm whale V68N and pig H64V/V68T myoglobins. For these mutants we constructed models based on the known structures of sperm whale V68L[51] and pig V68T MbCO,[52] respectively. Since the side chains of Leu and Asn are similar in size and geometry, the structure of sperm whale V68L MbCO was used for modeling V68N MbCO, in which the coordinates of Cβ and Cγ were retained while replacing the Cδ atoms with a carbonyl O and amide N. After the isobutyl group was changed to a propylamide, the side chain torsion angle $\chi 2$ was searched using MOLMOL[50] until a minimum distance (3.0 Å) between the Nδ atom of Asn68 and the O atom of the bound CO was obtained.

For modeling the structure of pig H64V/V68T MbCO, we used the structure of pig V68T MbCO in which the side chain of His64 was changed to that of Val and the orientation of the Val64 side chain was kept the same as in the structure of sperm whale H64V metmyoglobin.[51] The model used for electrostatic calculations on the oxy form of the H64V/V68T myoglobin mutant was constructed from the carbon monoxide model by replacing the linear FeCO complex with the bent FeO_2 system.

To determine the effect of the distal histidine conformation on the calculated CO stretching frequencies, we built a series of model structures based on the crystal structure of native MbCO.[51] The different models were obtained by rotating the dihedral angle $\chi 1$ of the distal histidine over a range of −30° to +95° relative to the original value. The −30° angle corresponds to the conformation where the imidazole is closer to the CO, and the 95° angle, to the swung-out conformation similar to that seen in the low-pH MbCO structure.[53] To remove steric clashes for conformations in which $\chi 1$ is changed by more than 30°, the alternative conformation for Arg45 determined for the low-pH structure of Mb was used.[53] The electrostatic potential was calculated at the carbon and oxygen positions of the CO molecule for all models as described above.

Results

Stretching Frequencies of CO Bound to Myoglobin. The IR spectrum of MbCO depends markedly on the structure and charge of the amino acid side chains surrounding the bound ligand.[16,17,24,52,54,55] The spectra in Figure 1 have been arranged from top to bottom in order of increasing frequency for the major peak; the average value of ν_{C-O} is given in Table 1. The same ordering shows decreasing frequencies for the Fe−CO bond, and the mean values for ν_{Fe-CO} are given in Table 1.

Cameron et al.,[52] Ray et al.,[7] and Li et al.[24] have suggested that the patterns shown in Figure 1 and Table 1 are due to differences in the sign and magnitude of electrostatic interactions with nearby distal pocket residues that alter the resonance structures of the FeCO complex. The extremes of behavior are shown schematically in Figure 2. When Val68 is replaced with an Asn, additional hydrogen bonding to the bound ligand occurs

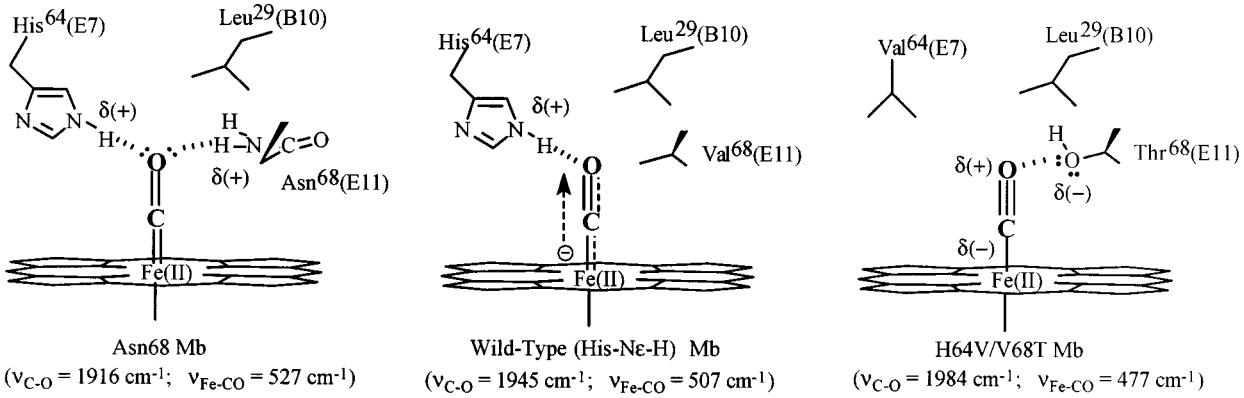

Figure 2. Proposed models for electrostatic interactions in the CO complexes of V68N, wild-type, and H64V/V68T myoglobins. As described in the text, the C−O bond order is ∼2.6 and the change in this order corresponding to the 70 cm^{-1} shift in ν_{C-O} is roughly 0.1. Thus, both the Fe═C═O structure for the V68N mutant and Fe−C≡O^+ for the H64V/V68T double mutant represent extreme resonance forms and not the dominant species.

and favors the neutral Fe═C═O resonance form with two sets of nonbonded electron pairs on the terminal O atom (left panel, Figure 2). As a result, the V68N mutation causes the position of the major IR peak to decrease from ∼1945 to 1916 cm^{-1}. At the other extreme, replacement of the distal histidine (His64) with aliphatic amino acids results in the appearance of a single IR band in the 1960−1970 cm^{-1} region. If, in addition, Val68 is replaced with Thr, the nonbonded electrons on the β-hydroxyl preferentially stabilize the Fe−C≡O^+ resonance form, increasing ν_{C-O} from ∼1945 cm^{-1} in the wild-type protein to 1984 cm^{-1} in the H64V/V68T double mutant (right panel, Figure 2). Qualitative interpretations of all the spectra presented in Figure 1, including the multiple bands, are given by Li et al.[24]

The resonance Raman peak for the Fe−CO stretching band also varies markedly, ranging from 478 cm^{-1} for H64V/V68T MbCO to 526 cm^{-1} for V68N MbCO (Table 1). In general, there is a strong inverse relationship between ν_{C-O} and ν_{Fe-CO}, and this correlation holds true for all of the mutants described here.[7,41]

Electrostatic Potentials in Myoglobin. Although consistent with the observed spectral properties, the interpretations illustrated in Figure 2 are qualitative at best. To quantify the importance of polarity on the bond orders and stretching frequencies of the iron carbonyl complex, we have computed electrostatic field strengths at both the O and C atoms of bound carbon monoxide. The calculations are based on known three-dimensional crystal structures, and the results are summarized in Table 1 along with the average values of ν_{C-O} and ν_{Fe-CO} measured in solution for the corresponding myoglobins. Except where noted in Figure 9, all computations were done with the neutral Nϵ−H tautomer of the distal histidine (see Discussion).

Stereodrawings of the contributions of the individual side chains to the electrostatic potentials in the active sites of V68N, wild-type, and H64V/V68T MbCO are shown in Figure 3. The atoms very close to bound CO (<5 Å) make large contributions to the calculated potential and dominate its value, as indicated by bright red or blue colors. The magnitudes of the electrostatic potentials at the bound C and O atoms are given in Table 1 and, for greater precision, exclude contributions from the heme group, which will be the same for the all the mutants (see Methods).

Vector representations of the fields and the effects of mutagenesis are shown in Figure 4 using the program GRASP. For this representation, the contributions of the heme group were taken into account explicitly to obtain the correct orientation of the field vector at a position equidistant between and the C and O atoms of bound carbon monoxide. The field lines are shown in green, and the field vector is given in yellow.

In the observed wild-type and modeled V68N sperm whale MbCO structures, the proton on the imidazole of His64 is very near the ligand and contributes a strong positive electrostatic field that acts on the bound ligand. In the wild-type protein, the magnitude of the electrostatic field is +13.5 kcal/mol at the O atom (minus contributions from the heme) and the vector is oriented at about 45° with respect to the C−O bond (Figure 4a). This positive field is enhanced significantly in magnitude by the amide protons of the asparagine side chain in the modeled structure of the V68N single mutant, and the field vector is more collinear with the Fe−C−O bonding system than in the wild-type protein (Figure 4b). Presumably, this is why the Asn68 mutant shows the largest decrease in ν_{C-O} (from ∼1945 to 1916 cm^{-1}) of any of the mutants examined.

The model of sperm whale V68N MbCO was based on the structure of V68L sperm whale myoglobin and is supported by the crystal structures of pig V68N met- and oxymyoglobin determined by Krzywda et al.[56] at low temperature in the presence of 25% sucrose and 10% glucose as cryoprotective solutes. These pig structures show that the side chain orientation of Asn68 is similar to that of Leu68 in sperm whale V68L metmyoglobin. However, the low-temperature structure of pig V68N MbCO shows the distal histidine in the up or out conformation with a $\Delta\chi 1$ angle >60° and entry of two water molecules into the active site. This conformation prescribes a smaller electrostatic field in the vicinity of the CO atoms (Table 3).

In solution at room temperature, both pig and sperm whale V68N MbCO show evidence for multiple conformations with major IR peaks at 1916 (65%) and 1936 (30%) cm^{-1} and a very small band centered at ∼1965 cm^{-1} (5%) (Figure 1[24]). The band at 1916 cm^{-1} almost certainly corresponds to the modeled structure of sperm whale V68N MbCO with His64 in the down conformation, which is seen in all native and wild-type sperm whale MbCO structures at room temperature in conventional solvents. The 1936 cm^{-1} band probably corresponds to the out His64 conformation seen in pig V68N MbCO at low temperatures in the presence of hydroscopic cryoprotectants. The high-frequency minor band may be due to the conformer with the O atom of the Asn68 pointing toward the bound ligand.[24]

When the distal histidine in MbCO is replaced with an apolar residue, the magnitude of the field vector becomes very small and ν_{C-O} increases to ∼1965 cm^{-1} (Table 1, H64L myoglobin

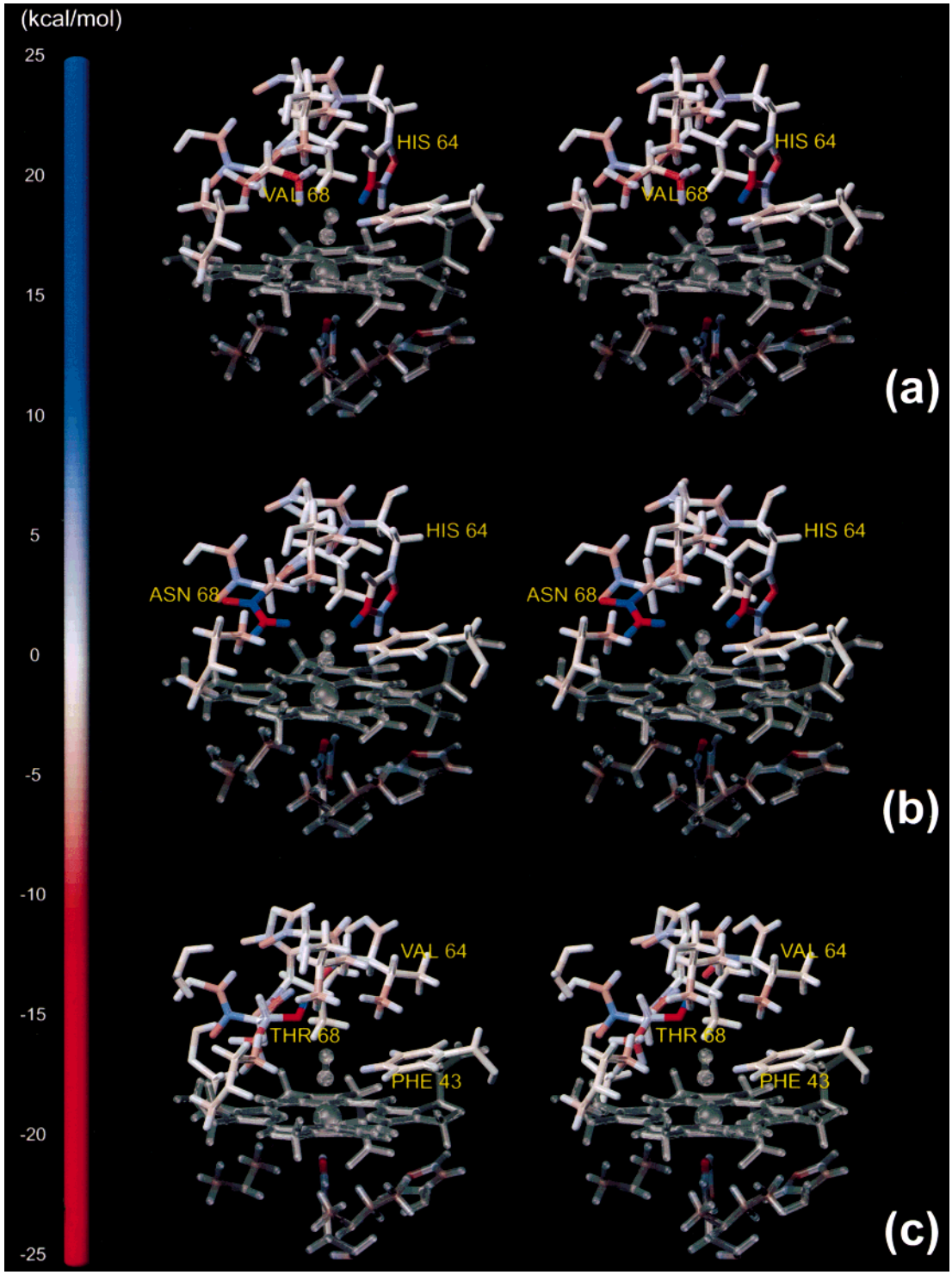

Figure 3. Stereoviews of myoglobin showing atomic contributions to the electrostatic potential at the bound ligand position. Atoms are colored from bright blue (strong positive) to bright red (strong negative), depending on their individual contributions to the electrostatic potential. The heme and ligand are colored vitreous white since they were not included for these calculations and the correlations with ν_{C-O} (see Methods). The structures on the proximal side of the heme group are also represented as vitreous since these contributions are the same for the wild-type and mutant myoglobins. The panels represent: (a) wild-type sperm whale MbCO; (b) V68N sperm whale MbCO; (c) H64V/V68T pig MbCO. Wild-type (or native) myoglobin represents an intermediate case in terms of potential at the bound ligand atoms. In V68N MbCO the overall potential is dominated by a very large positive contribution from the protons on the Nδ atom of the Asn68 side chain. In the H64V/V68T double mutant, the overall potential is negative due to a large negative contribution from the β-hydroxyl O atom of the Thr 68 side chain and the lack of a positive contribution from the residue at position 64.

in Figure 4d). When His64 is replaced with Val and Val68 is replaced with a Thr, the magnitude of the field becomes negative and the vector is almost perpendicular to the Fe−C−O bond system, explaining the very high value of ν_{C-O} (1984 cm^{-1}) observed for this double mutant (Figure 4c).

Correlations with ν_{C-O}, ν_{Fe-CO}, and ν_{N-O}. As shown in Table 1 and Figure 5A, there is a strong inverse linear relationship between the calculated electrostatic potentials at the bound ligand and the measured values of ν_{C-O} for all the proteins examined. The correlation coefficients are 0.93 for the

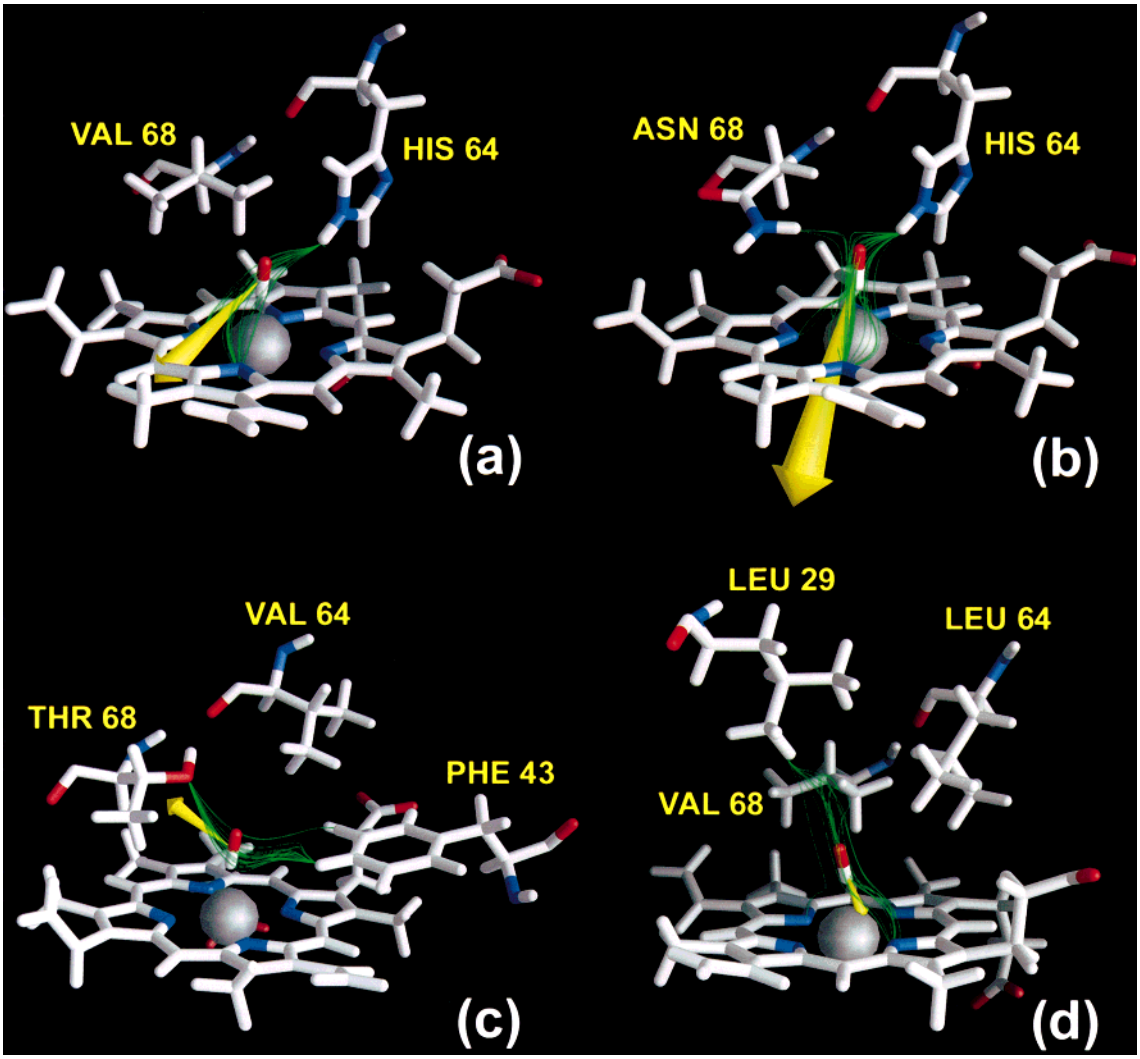

Figure 4. Calculated electrostatic field lines in the vicinity of the ligand binding site in myoglobin (green lines) and the field vector (yellow) at the bound ligand. The panels represent: (a) wild-type sperm whale MbCO, (b) V68N sperm whale MbCO; (c) H64V/V68T pig MbCO; (d) H64L sperm whale MbCO. Both wild-type and V68N MbCO show large positive field vectors pointing toward the iron atom, indicating that electrons in the FeCO bonding systems move in the opposite direction, i.e., toward the terminal O atom. The opposite situation occurs in the H64V/V68T double mutant with the field vector pointing away from the iron atom, indicating that electrons move toward the iron atom. In the H64L mutant, there are no highly polar residues near the bound ligand and as a result the field vector is very small.

straight line fits of potentials at either the C or O atoms versus ν_{C-O}. In principle, the effects on vibrational frequency and bond order should depend on the difference between the potential calculated at the O and that calculated at the C or iron atom (i.e., the projection of the protein electric field vector on the C−O bond). A plot of this difference between the O and C atoms versus ν_{C-O} also shows an inverse linear dependence with $R^2 = 0.77$ (Figure 5B). The correlations with both the absolute values of the potentials and the differences occur only when the Nϵ−H neutral imidazole tautomer is used to model the electrostatic characteristics of the distal histidine (see Discussion).

The resonance structures shown in the left and right panels of Figure 2 are exaggerations. The approximate value of ν_{C-O} for a true C═O double bond is ∼1600 cm^{-1}, whereas that for free C≡O is ∼2200 cm^{-1}. The observed value of 1945 cm^{-1} for wild-type protein indicates a bond order of ∼2.6, and the maximum observed change in peak positions, ∼60−70 cm^{-1}, indicates a change of only 10% in bond order. Thus, bound CO is a good sensor of electrostatic fields in the distal pocket without significantly influencing these fields by large changes in the distribution of its own resonance structures. For all the mutants examined, the FeCO system remains neutral, and the values of ν_{C-O} can be used to measure the electrostatic potential in the immediate vicinity of the iron atom. The slope of the correlation plot of potential versus ν_{C-O} is, in principle, related to the behavior of the CO vibration in applied as well as intrinsic electric fields (i.e., Stark tuning rate).[57]

The ν_{Fe-CO} peak is more difficult to assign in resonance Raman spectra and requires isotopic substitution of the bound ligand.[39,41] These bands are also harder to interpret due to coupling with heme bending modes.[15] As expected from the inverse relationship between ν_{C-O} and ν_{Fe-CO}, the Fe−CO stretching frequency does show a direct linear dependence on the potential at either the O or the C atom (Figure 6) and on the difference between them (not shown). The correlations of ν_{Fe-CO} with the absolute potentials are somewhat weaker ($R^2 = 0.85$) than those for ν_{C-O} but are still very good.

Recently, Spiro, Kitagawa, and co-workers have measured the stretching frequency for N−O bound to a series of myoglobin mutants.[42,43] A summary of their results are shown in Table 2 along with the calculated potentials at the bound O atom in MbCO and MbO_2 structures of the corresponding myoglobins. A plot of the potential at the O atom versus ν_{N-O}

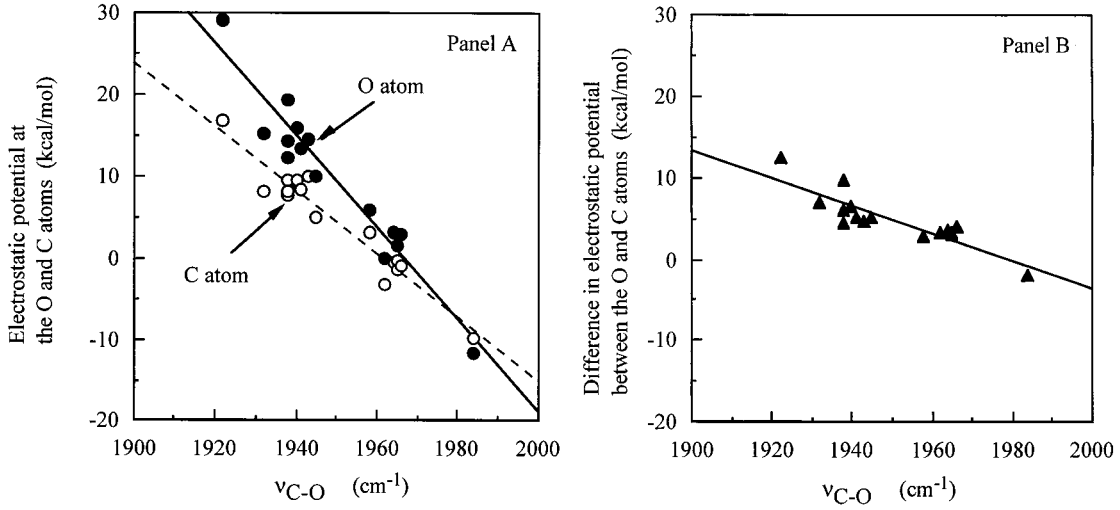

Figure 5. Correlation between electrostatic potential and ν_{C-O} for sperm whale and pig myoglobin mutants. Panel A: ν_{C-O} is plotted versus the electrostatic potential calculated at the O atom (filled circles) and at the C atom (open circles) of bound carbon monoxide. The results were fitted to straight lines, and correlation coefficients of 0.93 were obtained for potentials at either atom position. Panel B: ν_{C-O} is plotted versus the difference in calculated electrostatic potential at the O minus C positions. The correlation coefficient for the linear least-squares fit to the difference in potentials is 0.77.

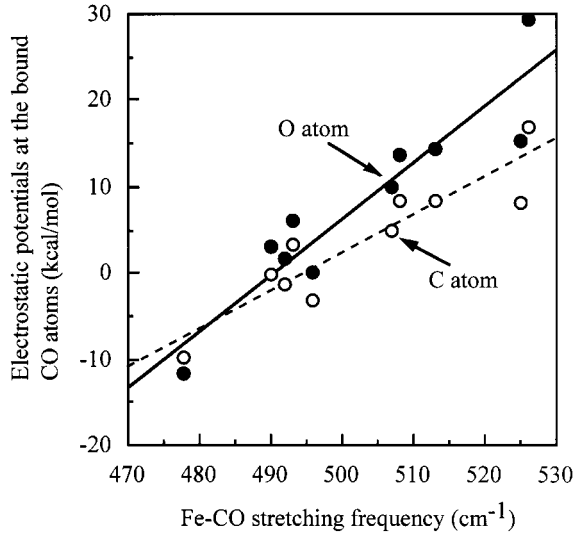

Figure 6. Correlation between electrostatic potential and ν_{Fe-CO} for sperm whale and pig myoglobin mutants. ν_{Fe-CO} is plotted versus calculated electrostatic potential at the O atom position (filled circles) and at the C atom (open circles) of bound CO. Least-squares fits to a straight line gave correlation coefficients of 0.84 and 0.86, respectively, for the data at the O and C atoms.

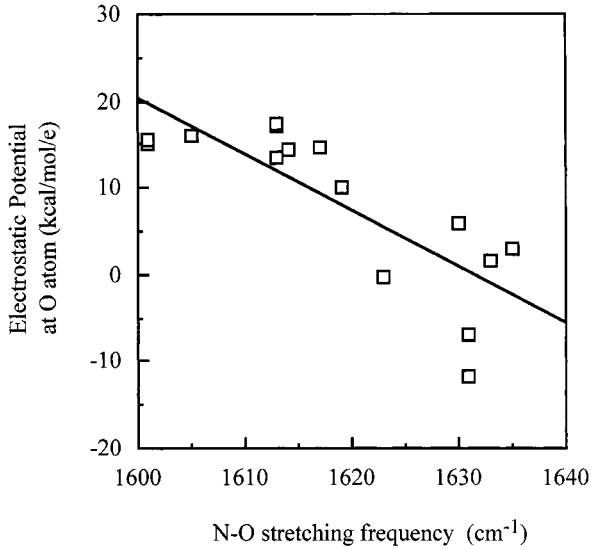

Figure 7. Correlation between electrostatic potential and ν_{N-O} for sperm whale and pig myoglobin mutants. ν_{N-O} is plotted versus calculated electrostatic potential at the O atom position based on MbCO crystal structures. The least-squares fit to a straight line gave a correlation coefficient of 0.64.

is shown in Figure 7. As with ν_{C-O}, there is an inverse relationship between the electrostatic potential and the N−O stretching frequency, but the correlation is weaker ($R^2 = 0.64$). The poorer correspondence is due both to the lack of authentic NO structures for more exact calculations and to the smaller range of changes in bond order allowed for the FeNO complex.[58] Unlike the situation for bound CO, there is no correlation between ν_{Fe-NO} and ν_{N-O}.[42,58] Very little ν_{O-O} data are available for the oxy complexes of mutant myoglobins, but Hirota et al.[59] have shown that ν_{Fe-O_2} is little affected by the H64L, L29F, and L29W mutations. Thus, only the vibrational properties of the iron carbonyl complex are suitable for defining electrostatic fields spectroscopically.

Discussion

Nϵ−H versus *Nδ*−H Tautomeric States of His64. Over 15 years ago, Schoenborn and co-workers published a series of papers[34−36] comparing the crystal structures of MbO_2 and MbCO using neutron diffraction in an attempt to locate the position of exchangeable protons on the distal histidine. Although these studies were technically very challenging and difficult to interpret, the results suggested strongly that a hydrogen atom is on the Nϵ atom of His64 when O_2 is bound and on the Nδ atom when CO was bound to the iron atom. Thus, most workers have assumed that tautomerization of the distal histidine occurs when O_2 is replaced with CO. However, effects of mutagenesis on the IR spectrum of MbCO suggest equally strongly that the distal histidine is contributing a positive electrostatic field in the vicinity of both bound ligands.[7,24] Other evidence to support this Nϵ−H model comes from spectroscopic studies that indicate that there is only a single conformation for bound CO and that the different stretching frequencies originate from conformational changes surrounding the ligand.[22,60,61]

Tautomerization of the distal histidine must occur in ferric myoglobin during the transition from water coordination at low

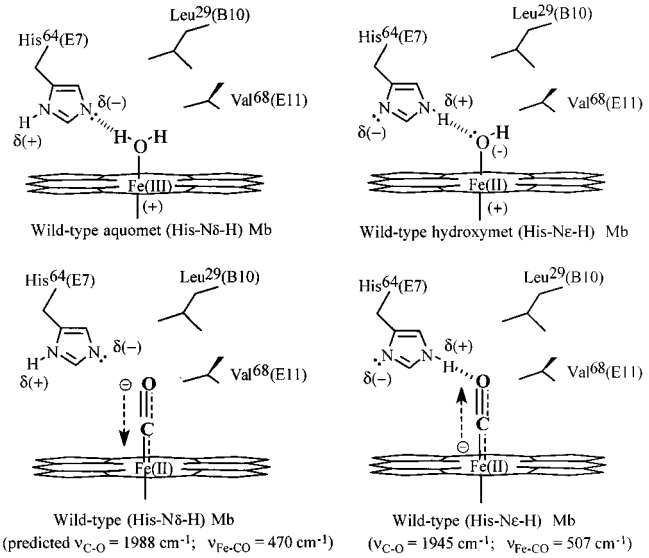

Figure 8. Drawings of Nδ−H and Nε−H tautomers of the distal histidine in metMb and MbCO. A water molecule is coordinated to the iron atom in the neutral pH form of aquometmyoglobin (upper left drawing). The Nδ−H tautomer allows stabilization of the bound ligand with the Nε atom acting as a proton acceptor. A hydroxide anion is bound to the iron atom in the high pH form of metmyoglobin. In this case the Nε−H tautomer stabilizes the bound ligand by donation of proton from the Nε atom (upper right drawing). The corresponding His64 tautomers of wild-type MbCO are shown in the lower drawings along with predicted or observed stretching frequencies.

pH to hydroxide binding at high pH (Figure 8). Hydrogen atom donation by the Nε−H tautomer is not feasible at low pH since a hydronium cation would be formed adjacent to the positively charged ferric atom. Similarly, it is unlikely that the nonbonded electrons of a deprotonated Nε atom would be adjacent to a coordinated hydroxide anion. In the case of ferrous deoxymyoglobin, His64 probably occurs as the Nδ−H tautomer since the lone electron pair of the Nε atom appears to be accepting a hydrogen atom from a noncoordinated, presumably neutral water molecule present in the distal pocket.[51] In the case of oxymyoglobin, the Nε−H tautomer appears to be the dominant form, donating a hydrogen atom to the negatively charged bound O atoms. The situation for bound CO is the least clear and depends on an interpretation of the vibrational properties of the bound ligand (Figure 8).

To examine this problem more quantitatively, we recalculated the electrostatic fields in all of the mutants containing a distal histidine using the Nδ−H tautomer, and the results are shown in Table 1 and Figure 9. The presence of the lone pair of electrons on the Nε atom of the Nδ−H tautomer creates a negative field in the vicinity of the bound ligand that dominates the calculated potentials. As a result, the computed values at both the O and C atoms are very negative and show no correlation with the observed C−O stretching frequencies (Figure 9A). Similarly, the differences between the potentials at the O and C atoms are small and do not correlate at all with ν_{C-O} (Figure 9B). In contrast, the potentials and differences calculated using the Nε−H tautomer correlate well with the observed stretching frequencies for His64 containing proteins and with the potentials calculated for those mutants without a distal histidine. The results in Figure 9 are compelling and indicate very strongly that the Nε−H form is the dominant tautomer in all of the CO complexes containing a distal histidine.

If the results in Figure 9 are accepted, then the neutron diffraction studies need re-interpretation. In both the MbCO and MbO_2 studies, the crystals were kept near room temperature for ~3 months to allow complete D_2O exchange, but there was little precaution to prevent autoxidation. In our hands, solutions of MbO_2 autooxidize completely in 1−2 days at room temperature even at high pH. Samples of MbCO oxidize in 1−2 weeks, particularly at low pH, unless there is continuous sparging with CO and the addition of reducing agents. Thus, the crystals were probably partially or completely oxidized. The MbO_2 sample was examined at high pH where the oxidized species is the low spin hydroxide form of metmyoglobin which has a red color and a visible absorption spectrum similar to authentic oxymyoglobin. As shown in Figure 8, the hydroxide met form would show a deuterium atom attached to Nε of the distal histidine. In addition, it would be hard to distinguish the OD^- ligand from bound O_2 since the O and D atoms have very similar neutron scattering strengths, 5.8 and 6.6 Fermis, respectively. If the low pH MbCO sample were oxidized, the aquomet form would be observed, and there would be no deuterium associated with the Nε atom of His64 (Figure 8). Again, it would be difficult to distinguish between bound CO and bound D_2O by neutron diffraction, particularly since the positions of the water deuterium atoms are not well fixed.

Correlations with Rates of O_2, CO, and NO Dissociation. For the past 10 years, we have argued that discrimination in favor of O_2 and against CO binding is due primarily to differential stabilization of the bound ligand by hydrogen bonding interactions with the distal histidine.[1,3] Summaries of the key features of the bound O_2, NO, and CO complexes of myoglobin are shown in Figure 10. The highly polar FeO_2 complex is stabilized ~1000-fold by the positive field of the Nε−H tautomer of His64, whereas there is much less stabilization of the apolar FeNO and FeCO complexes. Quantitative evidence in favor of this electrostatic theory for ligand discrimination is shown in Figure 11 where the logarithms of the O_2, CO, and NO dissociation rate constants are plotted versus the computed electrostatic potentials at the O atom in the corresponding MbCO or MbO_2 mutant structures (Tables 2 and 3).

As shown in the left panel of Figure 11, there is strong linear inverse correlation between $\log(k_{O_2})$ and the computed electrostatic potential at the O ligand atom, even though most of the potentials were based on CO structures (closed circles). Interestingly, when the four MbO_2 structures were analyzed separately (open circles), the correlation coefficient for the linear relationship increased from 0.71 to 0.98. Thus, the correspondence between the dissociation rate constant and the field at the bound ligand might be even greater if more mutant MbO_2 structures were available. It is important to note that the value of k_{O_2} decreases over 4 orders of magnitude, from ~5000 to 0.5 s^{-1} as the electrostatic field at the second bound O atom increases from ~−10 to +30 kcal/mol, respectively. Similar, but somewhat smaller changes occur in the overall oxygen association equilibrium constant, with positive fields enhancing O_2 affinity and negative fields decreasing it (Table 3). More detailed descriptions of the contributions of electrostatics, water displacement, and steric hindrance to the kinetics and equilibria of ligand binding are given by Olson and Phillips.[3]

The rate constants for CO and NO dissociation show weak and almost no dependence, respectively, on surrounding electrostatic fields. The equilibrium constants for the binding of these ligands also show little or no correlation with the computed potentials (see Tables 2 and 3[3,24]). In the case of CO dissociation, the rate constant decreases only 10-fold going from negative to positive fields. This small variation is due to the apolar nature of the FeCO complex and the small (<10%) changes in bond

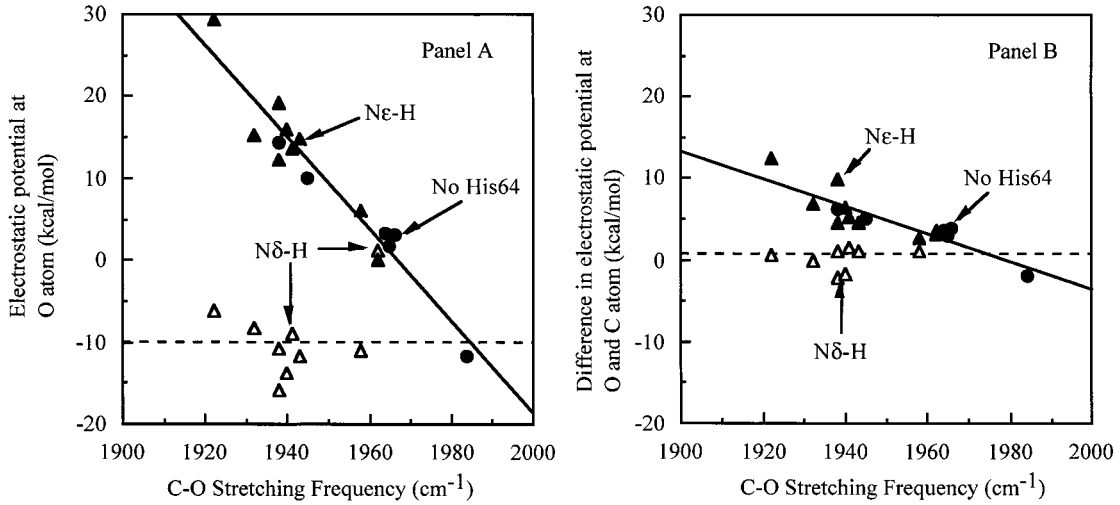

Figure 9. Correlations between ν_{C-O} and electrostatic potentials calculated for the Nϵ−H versus Nδ−H tautomers of His64 in MbCO complexes. In panel A the electrostatic potential was calculated at the O atom of bound CO, whereas in panel B the difference in potential between the O minus C positions was computed. Values for the Nϵ−H and Nδ−H tautomers are represented by filled and open triangles, respectively. The mutants that do not have a histidine residue at position 64 are represented by filled circles. The lines for the Nϵ−H tautomer data were taken from Figure 5, and strong correlations are observed. In contrast, there is no correlation between the potentials calculated for the Nδ−H tautomers and ν_{C-O} ($R^2 \approx 0.0$).

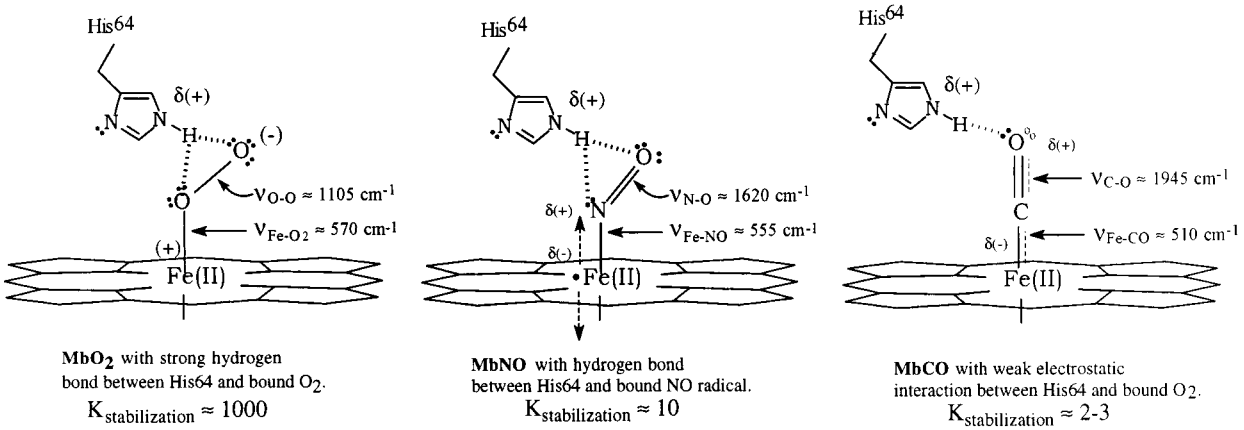

Figure 10. Drawings of MbO_2, MbNO, and MbCO showing bond orders, stretching frequencies, and estimated stabilization of the bound ligand by hydrogen bonding to His64. These structures and the stabilization parameters are taken from Olson and Phillips.[3]

order produced by mutagenesis. Thus, although ν_{C-O} is a good measure of electrostatic fields, it cannot be used to predict CO affinity because the free energy of stabilization of bound carbon monoxide is little affected by polar interactions.

Somewhat larger (~40-fold) changes in the NO dissociation rate constant are observed, but the correlation with electrostatic potentials is very poor. Part of the reason for the poor correlation may be the lack of MbNO crystal structures for more exact calculations. However, the major cause is that the overall rate constant for NO dissociation from myoglobin is not governed by the rate of thermal Fe−NO bond breakage. Instead, k_{NO} is determined by the product of the equilibrium constant for going from the bound to internal geminate states and the rate of ligand movement out of the protein.[10] Both of the latter parameters are significantly influenced by the size and accessibility of the distal pocket, which vary greatly among the mutants examined but are often unrelated to the polarity of the active site.[3] In contrast to NO, the overall rate constants for carbon monoxide and oxygen dissociation are more dependent on the rate of Fe−ligand bond breakage which, in the case of bound O_2, is dramatically inhibited by stabilizing electrostatic or hydrogen bonding interactions with adjacent protein residues (for a more complete discussion see Olson and Phillips[2,3]).

His64 Conformers, Tautomers, and the A States. Recent ultrahigh-resolution refinements of native MbCO at 85 K in two crystal forms by Romo[62] and Berendzen and co-workers (PDB files 1A6G, 1A6K, 1A6M, and 1A6N[37,38]) show discrete disorder of the distal histidine, revealing several distinct conformers of the imidazole side chain. Although not discussed, evidence for multiple orientations of His64 is also seen in the room temperature, high-resolution electron density maps of MbCO presented by Kachalova et al.[26] This conformational variability is almost certainly the cause of the multiple peaks seen in the IR spectrum of native MbCO. The idea that multiple orientations and tautomeric forms of the distal histidine alter the stretching frequency of bound CO is not new and was first proposed in detail by Oldfield et al.[63] on the basis of ^{13}CO NMR experiments. Li et al.[24] used this type of interpretation to explain the multiple bands seen in the IR spectra of a variety of mutant myoglobins, most notably F46V Mb, which shows multiple conformations of the distal histidine in its various crystal forms.[64] Kushkuley and Stavrov[65] have carried out theoretical studies of the effects of these conformers on the bond orders of the FeCO complex in native myoglobin.

The electrostatic calculations and correlations shown in Figures 3−6 provide a simple straightforward interpretation of the A states as defined by their ν_{C-O} values. The electrostatic field at the active site will depend markedly on the distance between the Nϵ−H proton and the bound CO atoms. In Figure 12, the calculated electrostatic field at the O atom is plotted as

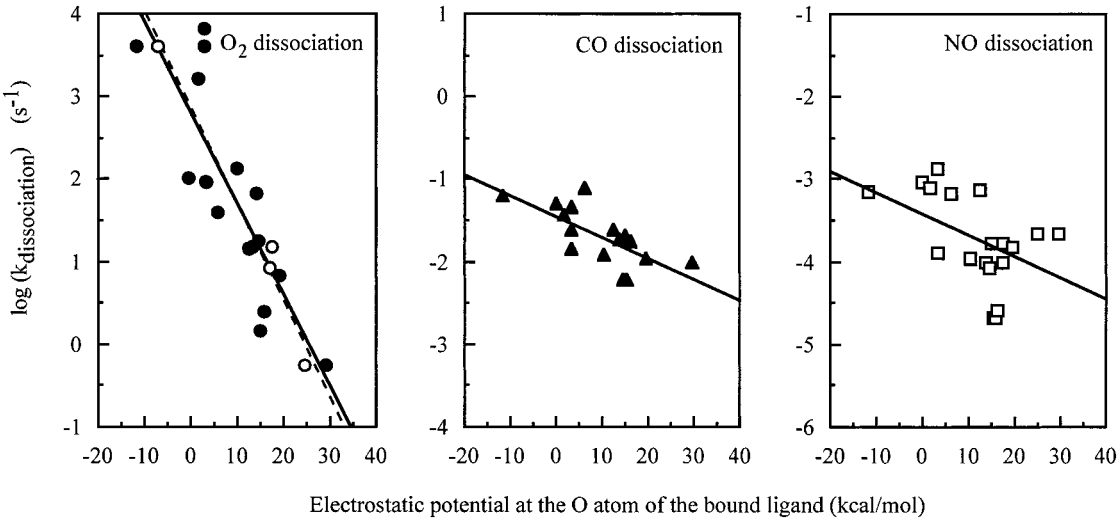

Figure 11. Correlations of the electrostatic field with O_2, CO, and NO dissociation rate constants (k_{O_2}, k_{CO}, and k_{NO}). Left panel: $\log(k_{O_2})$ versus electrostatic potential at the O atom calculated using MbO_2 (open circles) or MbCO crystal structures (filled circles). Straight line fits to the MbO_2 and MbCO data gave correlation coefficients (R^2) of 0.98 and 0.71, respectively. Middle panel: $\log(k_{CO})$ versus electrostatic potential at the O atom of bound CO based only on MbCO crystal structures or models. In this case, the correlation for a straight line fit was poorer ($R^2 = 0.50$). Right panel: $\log(k_{NO})$ versus electrostatic potential at the O atom based on MbCO structures. In this case, the fit to a straight line is very poor and little or no correlation is observed ($R^2 = 0.18$).

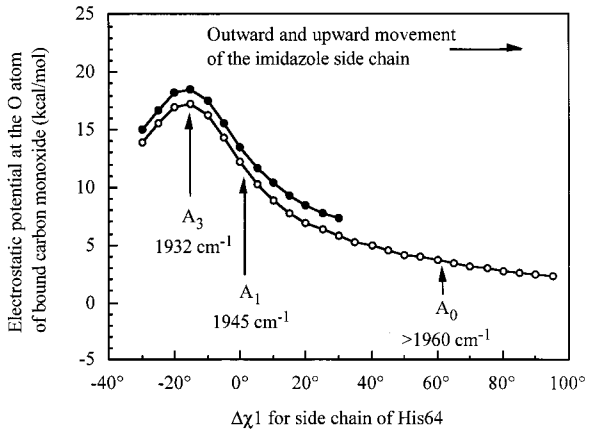

Figure 12. Calculated dependence of the electrostatic potential at the O atom of bound CO on the $\chi1$ angle of the distal histidine. The angle variation, $\Delta\chi1$, represents the change in this angle from its original value. Two alternative conformations for Arg45 were used in these calculations (see Methods). The filled circles represent potentials computed using the Arg45 conformation observed in the high pH structure of native sperm whale MbCO, whereas the open circles correspond to potentials calculated using the alternative Arg45 conformation observed at low pH.[53] The latter conformation allows for a greater range of motion of the His64 side chain. These changes in electrostatic potential with rotation of the distal histidine provide a simple interpretation of the bound A states, which were originally defined on the basis of peak positions seen in the IR spectrum of native MbCO (see Discussion).

a function of the variation in Cα-Cβ torsion angle from its original value (−168°), $\Delta\chi1$, for two situations. In the first case (closed circles), Arg45 is kept in the major position seen in the pH 7 structure of native MbCO, and the His64 side chain is rotated until it contacts the guanidino side chain. In the second case (open circles), the Arg45 side chain was put into the major solvent exposed conformation seen at low pH,[53] and His64 was allowed to relax to a position similar to that in the native or wild-type MbCO crystal structures. Then $\Delta\chi1$ was allowed to vary over a much wider range, from −30° to +95°, and the electrostatic fields at the bound CO atoms were calculated.

Since there is a linear relationship between the magnitude of the electrostatic potential at the O atom and ν_{C-O}, the A state stretching frequencies define the field strengths at the ligand atoms which, in turn, determine the $\chi1$ angles of the corresponding His64 conformers. As shown in Figure 12, the A_3, A_1, and A_0 states in native MbCO can be explained by distal histidine conformations with $\Delta\chi1 = -15°$, 0°, and >60°, respectively.

Similar analyses can be used to define the conformers in all of the mutants containing a His64 residue. In our view, a continuum of A states is possible for the different mutant myoglobins, depending on the free energy landscape of rotation about the $\chi1$ and $\chi2$ angles of the distal histidine and the presence or absence of partial charges at the other positions in close proximity to the bound ligand (i.e., residues 29(B10), 43(CD1), and 68(E11) in Figures 2−4).

Support for our interpretation in Figure 12 comes from the three major conformations of the distal histidine reported in the rapidly frozen crystal structure of MbCO at atomic resolution by Vojtechovsky et al. (PDB file 1A6G [37,38]). The highest occupancy (60%) conformer is similar to the one observed in previous lower resolution structures and has a $\chi1$ angle of −157°.[51] In the second conformer (20% occupancy), the imidazole ring moves toward the CO ligand and shows a $\Delta\chi1$ angle of −9°. The third conformation (20% occupancy) is similar to the swung-out position observed in the low pH native MbCO structure with $\Delta\chi1 = 76°$.[53] These angles correspond to those assigned to the A_1, A_3, and A_0 conformers in Figure 12 based on stretching frequencies measured at room temperature and a linear inverse relationship between electrostatic potential and ν_{C-O}. It is satisfying that the most populated conformer in the X-ray structure predicts the stretching frequency of the largest peak (A_1) seen in the IR spectra of native MbCO at both low and room temperature (Figure 1).

The correlations between the conformers observed by high-resolution crystallography and those predicted by our electrostatic calculations add further support to the view that the Nϵ atom of the distal histidine in the down (A_1 and A_3) conformations is protonated. If the Nδ−H tautomer were the major species, the electrostatic potential at the O atom would increase with increasing $\Delta\chi1$ and predict changes in ν_{C-O} from ∼1988 to ∼1960 cm^{-1} as this angle increased from −10° to +70°,

respectively. The latter results do point out an ambiguity in interpreting the conformers with ν_{C-O} values in the 1960−1980 cm^{-1} range. The cause of the higher stretching frequency is always a small or negative electrostatic field, but the change in field could be due to the distal histidine either swinging away from the bound ligand or tautomerizing to the Nδ−H isomer. Most crystallographic data, either at low pH or low temperature, suggest that outward movement of the distal histidine causes the appearance of the A_0 state in the native protein. On the basis of molecular dynamics simulations, Jewsbury and Kitagawa[66] have suggested that formation of the Nδ−H tautomer may actually cause outward movement of the distal histidine at neutral pH.

Conclusions

The peaks in the IR spectrum of carbon monoxide bound to myoglobin are clearly a sensitive gauge of electrostatic fields near the ligand binding site. The various A states in myoglobins containing a distal histidine are readily explained by changes in the conformation of the imidazole side chain that bring the Nϵ proton closer to or farther away from the bound ligand. In effect, the C−O bond acts as vibrating reed galvanometer to measure the sign and magnitude of the voltage in the active site of the protein. Since the potential field determines the extent of discrimination in favor of or against O_2 binding, the measured value of ν_{C-O} can also be used to predict physiologically relevant ligand binding parameters.

These conclusions are not limited to studies with myoglobin. For example, most peroxidases have positively charged or proton donating residues in their distal pockets.[67] Consequently, low ν_{C-O} values are both predicted and observed for the CO complexes of these enzymes.[7] Similarly, protonation of a residue in the active site of cytochrome oxidase appears to decrease the ν_{C-O} values of CO bound to either heme a_3 or Cu_B.[6] In the absence of a crystal structure, the value of ν_{C-O} can be used as an experimental measure of the electrostatic field in the distal pocket. For example, Deinum et al.[8] estimated that ν_{Fe-CO} and ν_{C-O} are ∼470[1] and ∼2000 cm^{-1}, respectively, for soluble guanylyl cyclase. They then used these values to argue that the very low affinity of guanylyl cyclase for O_2 is due both to proximal strain and to the presence of a negative electrostatic field in the distal pocket of the enzyme. As shown in Figure 11, a large negative field could readily increase the oxygen dissociation rate constant to a value precluding binding under physiological conditions but have little effect on NO binding. Reduced guanylyl cyclase does not bind O_2 in air but reacts readily with nitric oxide.[8] Thus, electrostatic analyses of ligand binding and ν_{C-O} appear to be applicable in general to heme proteins, subject only to changes in frequency scale when the proximal ligands are different.[7]

Acknowledgment. We are thankful to Michael K. Gilson for help, advice, and many stimulating discussions. This work was supported by United States Public Health Service Grants GM14276 (Q.H.G.), AR40252 (G.N.P.), GM35649 (J.S.O.), and HL47020 (J.S.O.); Grants C-1142 (G.N.P.) and C-612 (J.S.O.) from the Robert A. Welch Foundation; and the W.M. Keck Center for Computational Biology.

References and Notes

(1) Springer, B. A.; Sligar, S. G.; Olson, J. S.; Phillips, G. N., Jr. *Chem. Rev.* **1994**, *94*, 699−714.

(2) Olson, J. S.; Phillips, G. N., Jr. *J. Biol. Chem.* **1996**, *271*, 17593−17596.

(3) Olson, J. S.; Phillips, G. N., Jr. *J. Biol. Inorg. Chem.* **1997**, *2*, 544−552.

(4) Doherty, D. H.; Doyle, M. P.; Curry, S. R.; Vali, R. J.; Fattor, T. J.; Olson, J. S.; Lemon, D. D. *Nat. Biotechnol.* **1998**, *16*, 672−676.

(5) Hartman, J. C.; Argoudelis, G.; Doherty, D.; Lemon, D.; Gorczynski, R. *Eur. J. Pharmacol.* **1998**, *363*, 175−178.

(6) Mitchell, D. M.; Shapleigh, J. P.; Archer, A. M.; Alben, J. O.; Gennis, R. B. *Biochemistry* **1996**, *35*, 9446−9450.

(7) Ray, G. B.; Li, X.-Y.; Ibers, J. A.; Sessler, J. L.; Spiro, T. G. *J. Am. Chem. Soc.* **1994**, *116*, 162−176.

(8) Deinum, G.; Stone, J. R.; Babcock, G. T.; Marletta, M. A. *Biochemistry* **1996**, *35*, 1540−1547.

(9) Elber, R.; Karplus, M. *J. Am. Chem. Soc.* **1990**, *112*, 9161−9175.

(10) Gibson, Q. H.; Regan, R.; Elber, R.; Olson, J. S.; Carver, T. E. *J. Biol. Chem.* **1992**, *267*, 22022−22034.

(11) Andrews, B. K.; Romo, T.; Clarage, J. B.; Pettitt, B. M.; Phillips, G. N., Jr. *Structure* **1998**, *6*, 587−594.

(12) Beroza, P.; Case, D. A. *Methods Enzymol.* **1998**, *295*, 170−189.

(13) Ghosh, A.; Bocian, D. F. *J. Phys. Chem.* **1996**, *100*, 6363−6367.

(14) Kushkuley, B.; Stavrov, S. S. *Biophys. J.* **1997**, *72*, 899−912.

(15) Spiro, T. G.; Kozlowski, P. M. *J. Am. Chem. Soc.* **1998**, *120*, 4524−4525.

(16) Alben, J.; Caughey, W. *Biochemsitry* **1968**, *7*, 175−183.

(17) Shimada, H.; Caughey, W. S. *J. Biol. Chem.* **1982**, *257*, 11893−11900.

(18) Ansari, A.; J, B.; Braunstein, D.; Cowen, B.; Frauenfelder, H.; Hong, M.; Iben, I.; Johnson, J.; Ormos, P.; Suake, T.; Scholl, R. S.; A.; Steinbach, P.; Vittitow, J.; Young, R. *Biophys. Chem.* **1987**, *26*, 337−355.

(19) Ormos, P.; Braunstein, D.; Frauenfelder, H.; Hong, M. K.; Lin, S.-L.; Sauke, T. B.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8492−8496.

(20) Sage, J. T.; Morikis, D.; Champion, P. M. *Biochemistry* **1991**, *30*, 1227−1237.

(21) Ansari, A.; Berendzen, J.; Bowne, S.; Frauenfelder, H.; Iben, I.; Suke, T.; Shyamsunder, E.; Young, R. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5000−5004.

(22) Lim, M.; Jackson, T. A.; Anfinrud, P. A. *Science* **1995**, *269*, 962−966.

(23) Ma, J.; Shuanghong, H.; Straub, J. *J. Am. Chem. Soc.* **1997**, *119*, 2541−2551.

(24) Li, T.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *Biochemistry* **1994**, *33*, 1433−1446.

(25) Slebodnick, C.; Ibers, J. A. *J. Biol. Inorg. Chem.* **1997**, *2*, 521.

(26) Kachalova, G. S.; Popov, A. N.; Bartunik, H. D. *Science* **1999**, *284*, 473−476.

(27) Warshel, A.; Papazyan, A. *Curr. Opt. Struct. Biol.* **1998**, *8*, 211−217.

(28) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219−10225.

(29) Gilson, M. K.; Sharp, K.; Honig, K. *J. Comput. Chem.* **1988**, *9*, 327−335.

(30) Gilson, M.; Honig, B. *Proteins* **1988**, *3*, 32−52.

(31) Reynolds, W. F.; Peat, I. R.; Freedman, M. H.; Lyerla, J. R., Jr. *J. Am. Chem. Soc.* **1973**, *95*, 328−331.

(32) Bhattacharya, S.; Sukits, S. F.; MacLaughlin, K. L.; Lecomte, J. T. *Biophys. J.* **1997**, *73*, 3230−3240.

(33) Bhattacharya, S.; Lecomte, J. T. *Biophys. J.* **1997**, *73*, 3241−3256.

(34) Norvell, J. C.; Nunes, A. C.; Schoenborn, B. P. *Science* **1975**, *190*, 568−570.

(35) Hanson, J. C.; Schoenborn, B. P. *J. Mol. Biol.* **1981**, *153*, 117−146.

(36) Cheng, X.; Schoenborn, B. P. *J. Mol. Biol.* **1991**, *220*, 381−399.

(37) Abola, E. E.; Sussman, J. L.; Prilusky, J.; Manning, N. O. *Methods Enzymol.* **1997**, *277*, 556−571.

(38) Sussman, J. L.; Lin, D.; Jiang, J.; Manning, N. O.; Prilusky, J.; Ritter, O.; Abola, E. E. *Acta Crystallogr. D Biol. Crystallogr.* **1998**, *54*, 1078−1084.

(39) Ling, J.; Li, T.; Olson, J. S.; Bocian, D. F. *Biochim. Biophys. Acta* **1994**, *1188*, 417−421.

(40) Zhao, X.; Vyas, K.; Nguyen, B.; Rajarathnam, K.; La Mar, G.; L, T.; Phillips, G. J.; Eich, R.; Olson, J.; Ling, J.; Bocian, D. *J. Biol. Chem.* **1995**, *270*, 20763−20774.

(41) Anderton, C. L.; Hester, R. E.; Moore, J. N. *Biochim. Biophys. Acta* **1997**, *1338*, 107−120.

(42) Tomita, T.; Hirota, S.; Ogure, T.; Olson, J. S.; Kitagawa, T. *J. Phys. Chem.*, in press.

(43) Vogel, K. M.; Rush, T. S.; Kozlowski, P. M.; Spiro, T. G.; Dou, Y.; Ikeda-Saito, M.; Olson, J. S.; Zgierski, M. Z. Manuscript in preparation.

(44) Rohlfs, R. J.; Mathews, A. J.; Carver, T. E.; Olson, J. S.; Springer, B. A.; Egeberg, K. D.; Sligar, S. G. *J. Biol. Chem.* **1990**, *265*, 3168−3176.

(45) Brunger, A. T. *X-PLOR Version 3.1: A system for X-ray crystallography and NMR*; Yale University Press: New Haven, CT, 1992.

(46) Davis, M.; Madura, J.; Luty, B.; McCammon, J. A. *Program. Comput. Phys. Commun.* **1990**, *62*, 187−197.

(47) Madura, J.; Davis, M.; Gilson, M.; Luty, B.; Wade, R.; McCammon, J. *Rev. Comput. Chem.* **1993**, *5*.
(48) Gilson, M.; Honig, B. *Proteins* **1988**, *4*, 7−18.
(49) Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins* **1991**, *11*, 281−296.
(50) Koradi, R.; Billeter, M.; Wuthrich, K. *J. Mol. Graph.* **1996**, *14*, 51−55, 29−32.
(51) Quillin, M. L.; Arduini, R. M.; Olson, J. S.; Phillips, G. N., Jr. *J. Mol. Biol.* **1993**, *234*, 140−155.
(52) Cameron, A. D.; Smerdon, S.; Wilkinson, A.; Habash, J.; Helliwell, J. R.; Li, T.; Olson, J. *Biochemistry* **1993**, *32*, 13061−13070.
(53) Yang, F.; Phillips, G. N., Jr. *J. Mol. Biol.* **1996**, *256*, 762−774.
(54) Caughey, W. S.; Alben, J. O.; McCoy, S.; Boyer, S. H.; Charache, S.; Hathaway, P. *Biochemistry* **1969**, *8*, 59−62.
(55) Balasubramanian, S.; Lambright, D. G.; Boxer, S. G. *Proc. Nat. Acad. Sci. U.S.A.* **1993**, *90*, 4718−4722.
(56) Krzywda, S.; Murshudov, G. N.; Brzozowski, A. M.; Jaskolski, M.; Scott, E. E.; Klizas, S. A.; Gibson, Q. H.; Olson, J. S.; Wilkinson, A. J. *Biochemistry* **1998**, *37*, 15896−15907.
(57) Park, E.; Andrews, S.; Hu, R.; Boxer, S. *J. Phys. Chem.*, in press.
(58) Vogel, K. M.; Kozlowski, P. M.; Zgierski, M. Z.; Spiro, T. G. Manuscript in preparation.
(59) Hirota, S.; Li, T. S.; Phillips, G. N. J.; Olson, J. S.; Kitagawa, T. *J. Am. Chem. Soc.* **1996**, *118*, 7845−7846.
(60) Ivanov, D.; Sage, J. T.; Keim, M.; Powell, J. R.; Asher, S. A.; Champion, P. M. *J. Am. Chem. Soc.* **1994**, *116*, 4139−4140.
(61) Sage, J. T.; Jee, W. *J. Mol. Biol.* **1997**, *274*, 21−26.
(62) Romo, T. Ph.D. Thesis, Rice University, 1998.
(63) Oldfield, E.; Guo, K.; Augspurger, J. D.; Dykstra, C. E. *J. Am. Chem. Soc.* **1991**, *113*, 7537−7541.
(64) Lai, H.; Li, T.; Lyons, D.; Phillips, G. J.; Olson, J.; Gibson, Q. *Proteins* **1995**, *22*, 322−339.
(65) Kushkuley, B.; Stavrov, S. S. *Biophys. J.* **1996**, *70*, 1214−1229.
(66) Jewsbury, P.; Kitagawa, T. *Biophys. J.* **1994**, *67*, 2236−2250.
(67) Finzel, B. C.; Poulos, T. L.; Kraut, J. *J. Biol. Chem.* **1984**, *259*, 13027−13036.
(68) Braunstein, D.; Ansari, A.; Berendzen, J.; Cowen, B. R.; Egeberg, K. D.; Frauenfelder, H.; Hong, M. K.; Ormos, P.; Sauke, T. B.; Scholl, R.; Schulte, A.; Sligar, S. G.; Springer, B. A.; Steinbach, P. J.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8497−8501.
(69) Eich, R. F.; Li, T.; Lemon, D. D.; Doherty, D. H.; Curry, S. R.; Aitken, J. F.; Matthews, A. J.; Johnson, K. A.; Smith, R. D.; Phillips, G. N., Jr.; Olson, J. S. *Biochemistry* **1996**, *35*, 6976−6983.
(70) Eich, R. F. Ph.D. Thesis, Rice University, 1997.
(71) Carver, T. E.; Brantley, R. E., Jr.; Singleton, E. W.; Arduini, R. M.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *J. Biol. Chem.* **1992**, *267*, 14443−14450.
(72) Smerdon, S. J.; Dodson, G. G.; Wilkinson, A. J.; Gibson, Q. H.; Blackmore, R. S.; Carver, T. E.; Olson, J. S. *Biochemistry* **1991**, *30*, 6252−6260.
(73) Egeberg, K. D.; Springer, B. A.; Sligar, S. G.; Carver, T. E.; Rohlfs, R. J.; Olson, J. S. *J. Biol. Chem.* **1990**, *265*, 11788−11795.
(74) Smerdon, S. J.; Krzywda, S.; Brzozowski, A. M.; Davies, G. J.; Wilkinson, A. J.; Brancaccio, A.; Cutruzzolà, F.; Allocatelli, C. T.; Brunori, M.; Li, T.; Brantley, R. E., Jr.; Carver, T. E.; Eich, R. F.; Singleton, E.; Olson, J. S. *Biochemistry* **1995**, *34*, 8715−8725.