Classical Molecular Electrostatics: Recognition of Ligands in Proteins and the Vibrational Stark Effect

Parminder K. Mankoo and Tom Keyes*

Department of Chemistry, Boston University, Boston, Massachusetts, 02215

Received: June 25, 2006; In Final Form: September 20, 2006

It is shown that classical electrostatics quantitatively describes both the binding of the diatomic ligands XO (X = C, N, O) to the heme group in myoglobin and the dependence of their vibrational frequencies upon an external field, the vibrational Stark effect. The key is a proper treatment of induced dipoles. The results suggest that ligand binding occurs via an "electrostatic bond", a generalization of the standard ionic bond to include induction, and, more generally, that classical electrostatics can replace quantum mechanics for a considerable simplification of some complex problems.

I. Introduction

The interactions of molecules with externally applied electric fields govern spectroscopy, and interactions with fields from neighboring molecules are central to chemical bonding. A complete description is found in quantum mechanics, but given the difficulty of quantum calculations for large, complex systems, it is desirable to determine the limits of applicability of classical electrostatics. Here we demonstrate that classical "molecular electrostatics" can accurately describe the binding to iron in myoglobin (Mb) of the diatomic ligands XO (X = C, N, O) and the dependence of their vibrational frequencies on an applied field, the vibrational Stark effect (VSE).[1]

Our primary result is that some ligand binding may be treated without quantum mechanics. We propose that it occurs by "electrostatic bonds" and suggest that the textbook scenario of "covalent bonds and ionic bonds" be replaced by "covalent bonds and electrostatic bonds", greatly extending the range of the classical, noncovalent alternative. Textbooks also state that molecular recognition of ligands by biomolecules is noncovalent, but the ionic bond is insufficiently robust to describe the process, and the electrostatic bond fills the gap. Electrostatic bonds are primarily a consequence of induction, also called polarization, the creation of dipoles by local electric fields.

The importance of induction in condensed systems, with no reference to bonding, is clear. Induced dipoles cause the increase of the mean dipole of a water molecule upon condensation and contribute[2] $^{18}/_{19}$ of the OH stretch intensity in the IR spectrum of liquid water. It has also been recognized[3] for some time that classical induction includes bonding effects that would nominally be considered covalent. Our proposal is to convert such observations into a quantitative theory for bonds that are not too strongly covalent.

We will discuss the VSE first to introduce and calibrate the model. The VSE of ligands in proteins is of considerable interest in its own right as a probe of structure and dynamics, since local fields are determined by the protein configuration. Recent studies have focused[4,5] on carbon monoxide[4−6] (CO), nitric oxide (NO), and diatomic oxygen (O_2) in proteins such as Mb and hemoglobin. Furthermore, the VSE has been invoked[7−10] to calculate the frequency fluctuations that determine linear and nonlinear IR spectra and to provide a structural basis for interpreting these experiments. The VSE, like binding, is ordinarily calculated quantum mechanically, and a classical theory represents a substantial simplification.

The paper begins with our molecular electrostatic model,[13] previously applied to CO, which provides an accurate representation of the dipole and quadrupole moment, the molecular polarizability tensor, and the dipole derivative with respect to bond length, and allows a quantitative classical calculation of the complete liquid-state infrared spectrum, including the vibrational line. The essential ingredient is a careful, but simple, description of induction.

Induced dipoles are routinely allowed to polarize their bonded neighbors when calculating the response to an external field to obtain the molecular or whole-system polarizability. Polarization of bonded neighbors, which is central to the model and which we call "bonded neighbor induction" by permanent charges and bonded Coulomb interactions, is usually excluded to avoid overcounting electrostatic interactions already built into the bond potential. However, field-dependent contributions to the energy are not overcounting, since bond potentials are determined at zero field and bonded neighbor induction cannot be neglected in the presence of a field; correcting for overcounting at zero field is not difficult.

II. Classical Electrostatic Calculations

A. Molecular Electrostatic Model. Each atom is assigned atomic permanent dipoles in addition to atomic charges and atomic scalar polarizabilities. The induced dipoles obey $\mu_i^{\mathrm{ind}} = \alpha_i \mathbf{E}_i$, where $\mathbf{E}_i$ is the total local field from external sources, charges, permanent dipoles, and other induced dipoles, including contributions from bonded atoms, and α_i is the atomic polarizability. For sufficiently large separations, dipole fields may be calculated with the dipole tensor as usual. However, when the interatomic distance approaches $(4\alpha_i\alpha_j)^{1/6}$, unphysically large values for the induced dipole moments are obtained. For this reason, we employ Thole's damping;[14] modified dipole tensors and Coulomb fields result, which increase less strongly with decreasing r_{ij} for $r_{ij} < s_{ij}$, where $s_{ij} = \gamma(\alpha_i\alpha_j)^{1/6}$ and γ_{ij} is the damping constant.

* Author to whom correspondence should be addressed. E-mail: keyes@bu.edu.

10.1021/jp063971v CCC: $33.50
Published on Web 11/16/2006

The dipole fields are then calculated as

$$\mathbf{E}_i^{\mathrm{p,ind}} = -\sum_{j\neq i} \mathbf{T}_{ij}^{\gamma}\boldsymbol{\mu}_j^{\mathrm{p,ind}} \tag{1}$$

where $\mathbf{T}_{ij}^{\gamma}$ is the modified dipole tensor

$$\mathbf{T}_{ij}^{\gamma} = \frac{1}{r_{ij}^{\,3}}\left(\mathbf{I}(4v_{ij}^{\,3} - 3v_{ij}^{\,4}) - \frac{3\mathbf{r}_{ij}\mathbf{r}_{ij}(v_{ij}^{\,4})}{r_{ij}^{\,2}}\right) \tag{2}$$

$$v_{ij} = r_{ij}/s_{ij} \quad \text{if} \quad r_{ij} < s_{ij}$$
$$= 1 \quad \text{otherwise} \tag{3}$$

and $s_{ij} = \gamma(\alpha_i\alpha_j)^{1/6}$. The Coulomb field at atom i from the charge q_j on atom j becomes

$$\mathbf{E}_{\mathrm{C}}^{i} = \sum_j \frac{(4v_{ij}^{\,3} - 3v_{ij}^{\,4})q_j\mathbf{r}_{ij}}{r_{ij}^{\,3}} \tag{4}$$

The modified fields may be regarded as partial infinite resummations of the multipole expansion.

The parameters to be determined at the equilibrium geometry are two atomic polarizabilities, charges, permanent dipoles, and γ. The charges and total atomic dipoles are taken from ab initio calculations plus[15] "distributed multipole analysis"; note that we do employ quantum mechanics but only for the relatively simple problem of parametrizing the ligand.

The atomic polarizabilities, permanent dipoles, and γ are fit to the gas-phase mean molecular polarizability and polarizability anisotropy,[13,16,17] total atom dipoles,[13,18] and experimental dipole derivative with respect to bond length[13] (polarizability anisotropy derivative for O_2). The dipole moment of the ligand is given by the charge contribution plus the sum of all induced and permanent atomic dipoles and thus by construction agrees with the ab initio results. Model properties are given in Tables 1 and 2. For all three XO ligands the atom dipoles point inward; that is, an atom dipole points toward its bonded neighboring atom.

The Morse potential is chosen (Table 3) as the bond potential, $U_{\mathrm{bond}}(r)$, for NO and O_2, with equilibrium bond lengths r_{e} of 1.151 and 1.207 Å, respectively. For CO, we use the same Huffaker potential[19] as in our earlier study,[13] with $r_{\mathrm{e}} = 1.128$ Å. The zero field harmonic approximations to the frequencies of CO, NO, and O_2 using these anharmonic potentials at their equilibrium bond distances are 2170, 1832, and 1658 cm^{-1}, respectively.

B. Electric-Field-Dependent Energies. We consider XO and FeXO with and without the remainder of the protein as a source of Coulomb fields. The total energy U in an external field is given by the XO bond potential, U_{bond}, plus the electrostatic energy, U_{el}, further divided into the contributions of the permanent charges and dipoles, U_{cp}, and the induction energy, $U_{\mathrm{ind}} = -{}^1/_2\sum_i \boldsymbol{\mu}_i^{\mathrm{ind}}\cdot(\mathbf{E}_i^{\mathrm{c}} + \mathbf{E}_i^{\mathrm{p}} + \mathbf{E}_i^{\mathrm{ext}})$, where the fields are due to permanent charges, permanent dipoles, and external sources, respectively, and the sum runs over X, O, and, if present, Fe. All $\mathbf{E}^{\mathrm{ext}}$-independent contributions of an isolated XO are already in the bond potential, so

$$U(r) = U_{\mathrm{bond}}(r) + (U_{\mathrm{el}}^{(0)}(r) - U_{\mathrm{XO,\,el}}^{(0)}(r)) + \mathbf{E}^{\mathrm{ext}}\cdot(\mathbf{U}_{\mathrm{cp}}^{(1)}(r) + \mathbf{U}_{\mathrm{ind}}^{(1)}(r)) + \mathbf{E}^{\mathrm{ext}}\cdot\mathbf{U}_{\mathrm{ind}}^{(2)}(r)\cdot\mathbf{E}^{\mathrm{ext}} \tag{5}$$

where $\mathbf{r}$ is the OX bond vector and the subscript XO indicates a property of isolated XO.

TABLE 1: Atomic Charges δ, Atomic Polarizabilities α (Å^3), and Damping Coefficient γ for Molecular Electrostatic Models

ligand	δ_{X}	δ_{O}	α_{X}	α_{O}	γ
CO	0.382	−0.382	1.79	0.173	1.65
O_2	0.000	0.000	0.774	0.774	1.52
NO	0.226	−0.226	1.45	0.325	1.56

We take the point of view that an accurate molecular potential already includes all the electrostatic energy, and the subtraction in the parentheses is the correction for overcounting. Thus, for XO (no Fe), at zero field eq 5 reads $U(r) = U_{\mathrm{bond}}(r)$, as desired. For FeXO, the Fe−XO interaction is entirely electrostatic, so it does not contribute a bond potential or require a subtraction. The protein environment is included through the Coulomb field from nearby atoms acting as an external field on Fe−XO. We emphasize that $\mathbf{U}^{(1)}{}_{\mathrm{ind}}$, which dominates the VSE, is nonzero, only because we consider bonded neighbor induction. U depends on the complete geometry, not just r. The other dependences are implicit, but in practice, for the XO stretch, we require only the r dependence with R_{FeX} and the Fe−X−O angle fixed at the experimental values.

The individual electrostatic contributions for free ligands are

$$U_{\mathrm{ind}}^{(0)} = -\frac{1}{2}(\mathbf{E}^{\mathrm{c}} + \mathbf{E}^{\mathrm{p}})\cdot\boldsymbol{\alpha}_{\mathrm{mol}}\cdot(\mathbf{E}^{\mathrm{c}} + \mathbf{E}^{\mathrm{p}}) \quad U_{\mathrm{ind}}^{(1)} = -(\mathbf{E}^{\mathrm{c}} + \mathbf{E}^{\mathrm{p}})\cdot\boldsymbol{\alpha}_{\mathrm{mol}}$$
$$U_{\mathrm{cp}}^{(1)} = -(\boldsymbol{\mu}^{\mathrm{c}} + \boldsymbol{\mu}^{\mathrm{p}}) \qquad U_{\mathrm{ind}}^{(2)} = -\frac{1}{2}\boldsymbol{\alpha}_{\mathrm{mol}} \tag{6}$$

where $\boldsymbol{\alpha}_{\mathrm{mol}}$ is the molecular polarizability tensor and $\boldsymbol{\mu}^{\mathrm{c}}$ and $\boldsymbol{\mu}^{\mathrm{p}}$ are the charge and permanent dipoles, respectively; $U_{\mathrm{el}}^{(0)}$ is the usual energy of charges and dipoles. These bond-vector-dependent quantities are written in the $3N$-dimensional space indexed by Cartesian direction and atom number.

C. Vibrational Stark Effect. The frequency shift of an oscillator in an electric field, $\Delta\bar{\nu}$, in cm^{-1}, is expressed quantum mechanically as

$$hc\Delta\bar{\nu} = -\Delta\mu\cdot\mathbf{E} - \mathbf{E}\cdot\frac{1}{2}\Delta\boldsymbol{\alpha}^{t}\cdot\mathbf{E} \tag{7}$$

where $\mathbf{E}$ is the local field, h is Planck's constant, c is the speed of light and $\Delta\mu$ and $\Delta\boldsymbol{\alpha}^t$ are the transition dipole and polarizability, the differences of the molecular dipole and polarizability, respectively, in the excited and ground vibrational states, and $\Delta\mu \equiv |\Delta\mu|$ is called the Stark tuning rate.

One calculates the IR spectrum in the presence of a field by noting that every molecule has a frequency and also a transition dipole, determined by its orientation with respect to the field. The spectrum is obtained as an average over the orientational distribution, which is isotropic in a glassy system, perturbed from isotropy by the field in a liquid or gas, and constrained in a protein cavity. The result is a sum of terms proportional to the zero-field spectrum and its frequency derivatives,[1,11,12] where the coefficient of the second derivative term is simply proportional to $\Delta\mu^2$, and thus fitting yields the Stark tuning rate.

With the energy in hand, the classical frequency in the presence of a field is calculated as $((\mathrm{d}^2U_{\mathrm{XO}}/\mathrm{d}r^2)_{r_{\mathrm{e}}(E)}/M)^{1/2}$, where M is the reduced mass and $r_{\mathrm{e}}(E)$ is the field-dependent bond length. A quantum anharmonicity correction could be applied, but it does not influence frequency shifts to a first approximation, so we will proceed with harmonic frequencies. The field dependence is a sum of a linear and a quadratic term, leading to a pure downward parabola peaked at a nonzero field and with an asymmetry about zero field. Our simple calculation yields results similar to those found[20] for the O−H frequency

TABLE 2: Electrostatic Parameters for Free Ligands[a]

ligand	$\|\mu^t\|$	$\alpha_{\|\|}$	$\alpha_\perp$	intensity	μ^{ind}	μ^c	μ^p
CO	0.323 (0.112)	2.32 (2.32)	1.79 (1.79)	1.51 (1.21−1.54)	2.546	−2.068	−0.801
NO	0.289 (0.157)	2.30 (2.30)	1.46 (1.46)	2.143 (2.098−2.325)	1.183	−1.252	−0.220
O_2	0.000 (0.000)	1.52 (1.60)	1.24 (1.10)	3.89 (3.89)	0.000	0.000	0.000

[a] Dipoles are given in debye (D), and polarizabilities are given in Å^3, with experimental values from from refs 16 and 17. Intensity denotes the squared dipole (polarizability anisotropy for O_2) derivative with respect to bond length (in Å), with the experimental values in parentheses from refs 35, 36, and 37. μ^{ind} denotes the induced dipole, μ^p the permanent dipole, and μ^c the charge dipole. The sign convention is that X is on the −*z*-axis and O is on the +*z*-axis.

TABLE 3: Morse Potential Parameters[40,41] for O_2 and NO[a]

ligand	D_e	β	r_e
O_2	120.2	2.78	1.207
NO	152.6	2.64	1.151

[a] Dissociation energy D_e, including zero-point energy, is given in kcal/mol, inverse range β is given in Å^{-1}, and equilibrium internuclear distance r_e is given in Å.

in a HOD molecule by performing ab initio calculations including an external field. The frequency decreases if a field is applied parallel to $\Delta\mu$, and thus, we determine that for the two heteronuclear XO molecules it points toward X. By symmetry, $\bar{\nu}(E)$ is peaked at zero field for free homonuclear O_2, and the Stark tuning rate vanishes. However, $\bar{\nu}(E)$ has a finite slope at finite field.

For free CO, the field-dependent potential is shown in Figure 1 a. The shift to a larger bond length with increasing field and the softening corresponding to the red shift are clear. The quantum Stark tuning rate can be expressed as the sum[21] of contributions for fixed geometry and for the change in geometry. Both contributions are included in the classical calculation presented here. The classical VSE is demonstrated for the three free diatomics in Figure 1b. For CO, the classical electrostatic calculation yields $\Delta\mu = 0.63$ cm^{-1}/(MV/cm), falling in the range of available experimental values, 0.69[5] and 0.43.[6] With the O_2 frequency function peaked at zero field, the vanishing of the tuning rate may be seen from the parabolic field dependence. We have not been able to find experimental data for NO and O_2. Our results are comparable to those of an ab initio[22] study of CO and NO (Table 4). The $\Delta\alpha^t$ values for CO, NO, and O_2 are 0.066, 0.090, and 0.140 a.u., respectively.

With an accurate bond potential and dipole derivative, the correct tuning rate of the free diatomics is expected.[6,9] This means that bonded neighbor induction dominates the VSE by providing the correct dipole derivative or, more generally, a simple and accurate dipole surface.[23] Other classical models[23] can do this, but we believe that intramolecular induction, particularly between bonded neighbors, should be included for the most physically motivated approach.

D. VSE of Ligands Bound to Iron. The VSE of diatomic ligands bound to the heme group serves[4,5,7−9] as a probe of electric fields in the protein, which might be further related to structure. Due to the complexity and importance of these systems, successful classical calculations would be particularly useful. As a first approximation, we include the heme iron only.

Experimentally,[4,5] CO bound to heme iron has a Stark tuning rate of $(1.8/f)-(2.4/f)$ cm^{-1}/(MV/cm), where f is a local field correction factor relating the local electric field at the chromophore to the applied field. Local field corrections pose an interesting problem due to the heterogeneity of proteins. The local field around the CO is rapidly varying and quite different from the applied field due mainly to the induced dipole in the iron; this constitutes a short-range local field effect and is included in the theory. It also makes sense to correct the field

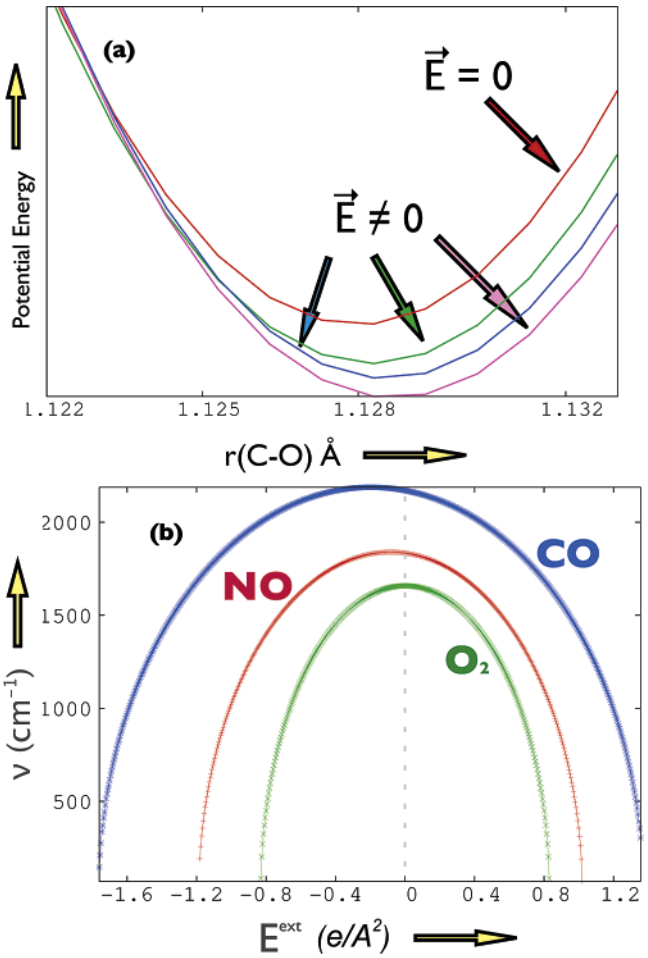

Figure 1. (a) Bond potential changes upon interaction of a molecule with an applied electric field. The fields range from 0 to 0.01 *e*/Å^2 and point from O to X. (b) Complete dependence of CO, NO, and O_2 vibrational frequencies upon electric field along the bond at their equilibrium bond lengths.

acting on the entire chromophore with a long-range dielectric cavity-type correction. Estimates of f lie[24] between 1.1 and 1.3. Given the uncertainties, to avoid introducing another parameter, we will ignore f ($f = 1$) in modeling FeCO and assume a tuning rate of 2.4 cm^{-1}/(MV/cm), corresponding to the more recent experiment[5] and to an increase of $\Delta\mu$ on the order of 4 times with respect to free CO.

To obtain the electrostatic energy of Fe−CO we require the geometry plus three additional parameters: the Fe charge and polarizability and the damping parameter γ for the Fe−C interactions. Since our molecular dynamics calculations on Mb use the CHARMM force field,[25,26] we retain the corresponding values for MbCO of $R_{Fe-C} = 1.92$ Å, $R_{C-O} = 1.17$ Å, and an Fe−C−O bond angle[27] of 180°; the Fe charge is +0.24.

For the Fe polarizability, we fit to the tuning rate of 2.4 cm^{-1}/(MV/cm), obtaining, with no Fe−C damping, $\alpha_{Fe} = 4.39$ Å^3. This large polarizability gives a large induced dipole, which is

TABLE 4: Calculated Vibrational Stark Effect for Free Ligands and Experimental[5,6] Values for CO (in 2Me-THF) and Recent Ab Initio[22] Calculations for CO and NO[a]

ligand	$\Delta\mu$	$\Delta\mu$ (exptl)	$\Delta\mu$ (ab initio)
CO	0.63	0.43[r1], 0.69[r2]	0.51[r3]
O_2	0.00		
NO	0.53		0.36[r3]
Fe−CO	2.40	1.8−2.4[r4]	1.3[r3]

[a] Stark tuning rates ($\Delta\mu$) are given in units of cm^{-1}/(MV/cm). Blank cells indicate no data. The superscipt "r1" refers to experiment,[6] the superscript "r2" also refers to experiment,[5] and the superscipt "r3" refers to the ab intio study.[22] The polarizability of iron $\alpha_{Fe} = 4.39$, in $Å^3$, has been adjusted to fit the experimental Stark tuning rate from ref 5. The superscipt "r4" refers to experiment.[5] For the experimental geometry, the charge for Fe is taken from CHARMM[25,26] for MbCO with $R_{C-O} = 1.17$ Å, $R_{C-Fe} = 1.92$ Å, $\theta_{Fe-X-O} = 180°$, and $\delta_{Fe} = 0.24$ *e*.

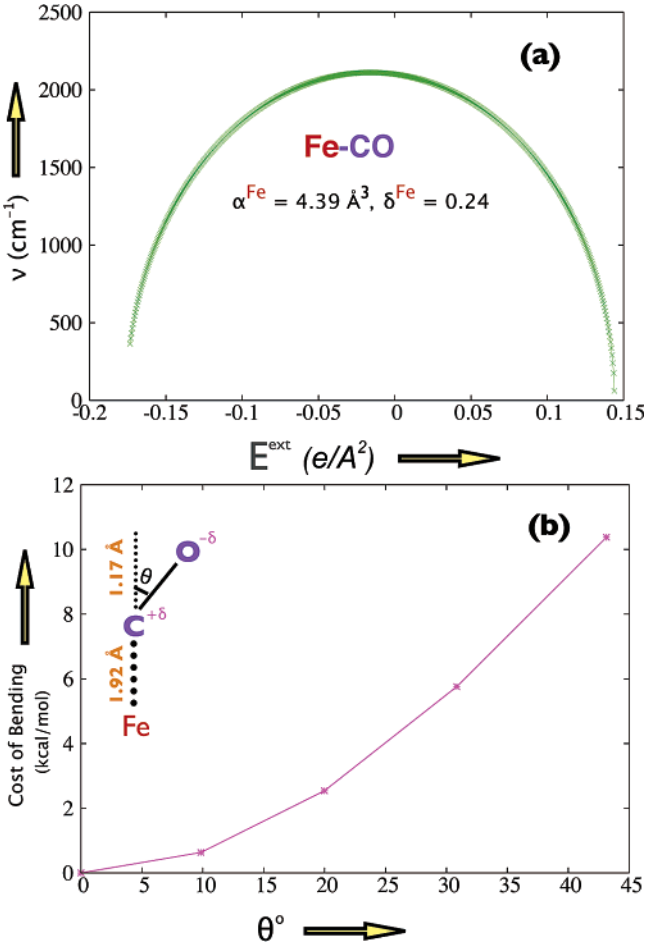

Figure 2. (a) Variation of frequency with field for CO bound to Fe at the experimental Fe−C−O geometry. (b) Electrostatic energy rises upon bending Fe−C−O away from linear (0°), providing a classical explanation for the true linear geometry.

the source of the enhancement of the Stark response, showing yet again the importance of induction. The dipole-tensor singularity then appears at 1.77 Å, which is well below the experimental[25] bond length of 1.92 Å of the Fe−CO bond in Mb (at 1.71 Å < 1.81 Å for Fe−NO and at 1.54 Å ≪ 1.73 Å for Fe−O_2). Therefore, again attempting to be parameter-lean, we simply retain undamped Fe−C interactions. The results are summarized in Table 4 and Figure 2a. The tuning rate for FeCO is listed for completeness; it is not a prediction.

While we take $\Delta\mu$, the zero-field slope of $\bar{\nu}(E)$, from experiment, in return we obtain the entire function $\bar{\nu}(E)$. Calculations of IR spectra[7−9] have invoked the linear low-field form, but we now have the means to treat frequency fluctuations due to large fluctuating electric fields in protein interiors. For an estimate[28] of the average field along the CO bond in mutant H64V MbCO of 0.015 $e/Å^2$, the calculated red shift exceeds that from the low-field linear approximation by ∼10 cm^{-1}. In wild-type MbCO, where the local fields should be larger,[28] the spectral range of the A_0, A_1 and A_3 features in the linear IR is[9] ∼60 cm^{-1} (peaks separated by ∼35 cm^{-1}), so the nonlinear effects seem significant. Calculated frequency fluctuations should be influenced by the derivative, $d\bar{\nu}/dE$, and at $E = 0.015$ $e/Å^2$ it is about about twice the zero-field slope, $\Delta\mu$. Equations 5 and 6 also describe the response to an inhomogeneous field that may be different at each atom and, away from zero field, to a field perpendicular to the bond. The presence of the Fe modifies the Stark response itself, in a manner that will be very sensitive to configurational fluctuations. Thus, the theory is well suited to describe ligands in proteins.

TABLE 5: Binding Energies of Ligands to the Heme in Mb (kcal/mol)[a]

ligand	BE (calcd)	BE (ab initio)	BE (exptl)	R_{X-O}	R_{Fe-X}	θ_{Fe-X-O}
CO	29.3, 16.8(P)	26[r3a], 35[r3b]	10.8[r1], 21.4[r2]	1.17	1.92	180°
O_2	6.82	9[r3a], 15[r3b]	10.2[r1], 18.1[r2]	1.30	1.73	133°
NO	25.3	23.4[r4], 35[r3a]		1.16	1.81	139°

[a] The calculated values include Fe only except (P) denotes a treatment of all protein residues within 10 Å of the iron. Distances are given in Å. The Fe−X−O geometry is taken from CHARMM[25,26] for CO, from experiment[36] for O_2, and from ab initio calculations[31] for NO. The superscript "r1" refers to experiment,[33] the superscript "r2" also refers to experiment,[32] the superscipts "r3a" and "r3b" refer to ab initio calculation,[34] where "3a" is ligand + heme and "3b" is ligand + heme + imidazole, and the superscript "r4" refers to ab initio calculation,[31] ligand + heme.

E. Electrostatic Bonds and Binding Energies: Recognition of Ligands in Proteins. Discrimination among ligands in binding is physiologically relevant and constitutes a key to the relationship between structure and function. Although the physiological function of Mb is O_2 binding, CO is also a biologically significant ligand for Mb and other heme proteins because it is a poison. The oxygenated form of Mb readily reacts with NO and may prevent NO from inhibiting cytochrome c oxidase, making the protein a protector of cellular respiration.[29] Several physical effects contribute to binding: covalent bonding, hydrogen bonding, steric hindrance, and electrostatic interactions. The electrostatic energy is particularly sensitive to changes in local protein environment through the variation of the local fields. Now, having constructed and parametrized the FeXO models, we will show that binding is dominated by electrostatics.

First we treat the Fe only, identifying the binding energy as the difference between the energy of Fe−XO with the ligand removed to infinity and at a geometry determined from the literature (Table 5). Results from the bare-bones approximation lie in the range of literature values from both ab initio[30,31] calculations and experiment,[32,33] all reported in Table 5. Comparison is most apt to calculations including the fewest atoms of the protein. For example, an ab initio[30] calculation on heme + ligand finds binding energies in (kcal /mol) of NO(35) > CO(26) ≫ O_2(9), while we have, CO(29.3) ≳ NO(25.3) ≫ O_2(6.82). Given the variation in the energies obtained in published quantum calculations, with different basis sets and models of the protein, this amounts to agreement.

The cost of bending the Fe−C−O complex away from linear is shown in Figure 2b. The energy increases by about 0.7 kcal /mol for a departure from linearity of 10° and reaches 6 kcal/mol at 30°, in reasonable agreement with a recent ab initio

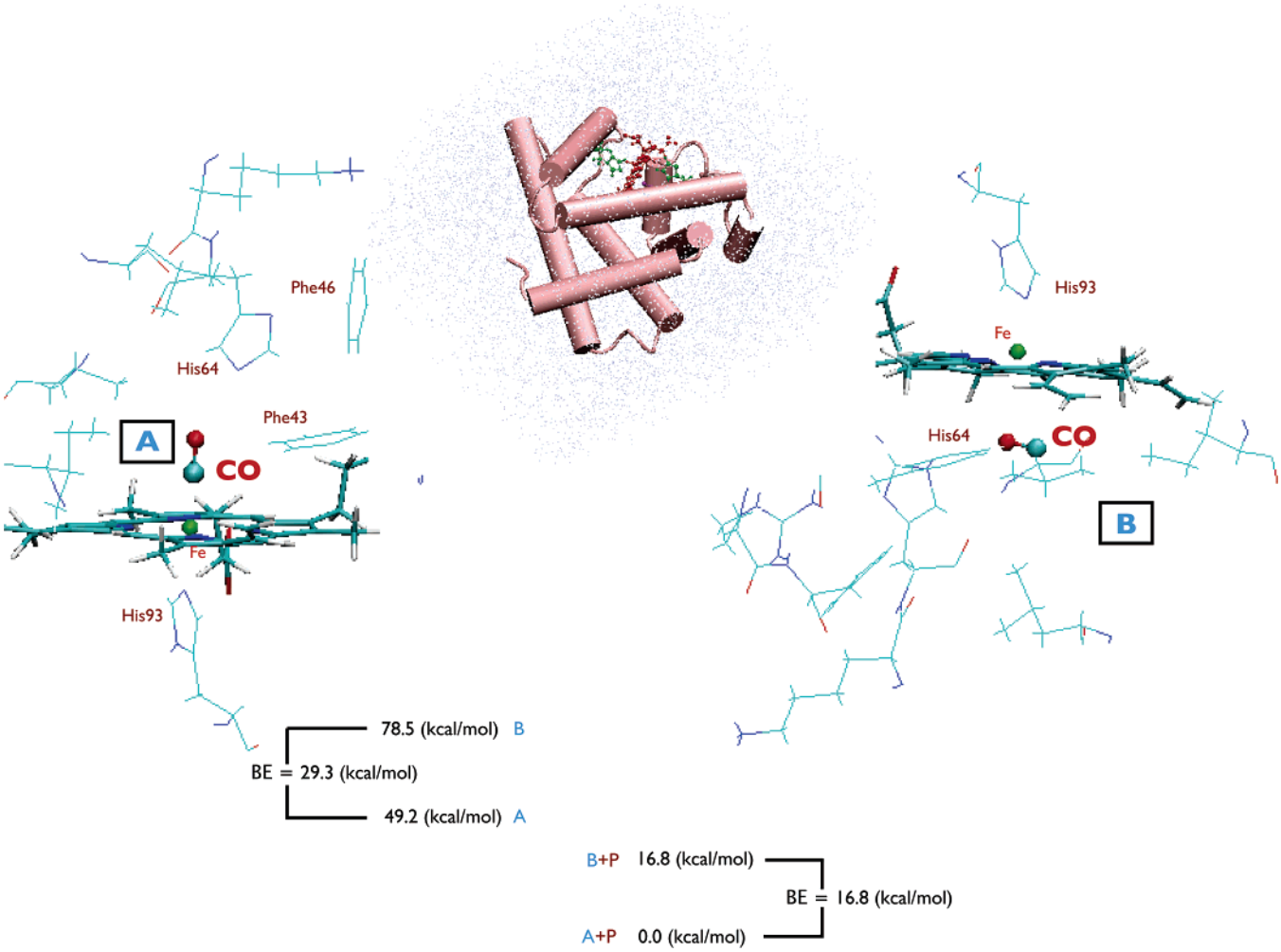

Figure 3. Myoglobin in an orthorhombic box of water molecules. An enlargement of the active site illustrates the CO ligand in (A) bound and (B) unbound states, with important nearby residues. The energy level diagram depicts the energies of bound and unbound states in the absence and presence of the protein (P) environment.

study,[34] which shows an energy increase of approximately 1 kcal/mol for a departure from linearity of 10° and approximately 9 kcal/mol for a departure of 30°.

Next we include the contribution of the protein environment to CO binding. Starting with the X-ray structure of MbCO, molecular dynamics simulations with explicit water and energy minimizations with backbone constraints are used to generate well-equilibrated configurations, representative (Figure 3) of the bound "A" state and, switching to parameters for five-coordinate iron,[35] the unbound "B" state.

The environment influences the electrostatic energy via the external electric field from the CHARMM partial charges. We include the residues (heme, Leu29, Phe43, Arg45, Phe46, Lys63, His64, Val68, His93, and Ile107), within 10 Å of the ligand (Figure 3). The result is a substantial decrease in binding energy to 16.8 (kcal/mol), which lies in the range of experimental results.[32,33]

III. Discussion

The detailed electrostatic origin of the VSE redshift is as follows. When an electric field is applied parallel to $\Delta\mu$ of an XO molecule in the equilibrium geometry, pointing toward partially positive X from partially negative O, the field-induced dipoles are parallel to the field. Both charge- and field-induced dipole interactions are unfavorable; i.e., they raise the energy. Recalling that the atom dipoles point inward (toward the bonded neighbor), the $(\mu_X)-$(field-induced dipole on O) interaction is also unfavorable, leaving a single favorable $(\mu_O)-$(field-induced dipole on X) term. With three unfavorable and one favorable contributions, the bonded neighbor induction energy is positive and falls as the bond length is increased, partially canceling the rise in the bond potential, resulting in a softening and red shift.

Electrostatic bonds provide a simple physical picture of binding, without quantum mechanics. Ligands bind to Fe primarily through induction, even overcoming a Coulomb repulsion for positively charged carbon or nitrogen. Induction by a charge is attractive, since the induced dipole is in the same direction as the Coulomb field, lowering the energy. The partially positive X in XO, not the negative O, bonds to the positive Fe, and O_2 has the lowest BE, because (Table 1) $\alpha_X \gg \alpha_O$. The electrostatic energy of FeOC ($R_{Fe-O} = 1.92$ Å, linear) is 35 kcal /mol higher than that of FeCO, despite the favorable Coulomb attraction.

Coulomb and induction attraction to the Fe of the far oxygen and atom dipoles pointing toward the Fe favor a bent geometry. The complex will be linear only if atom dipoles pointing away from the Fe overcome these effects. The magnitudes of the XO ligand atom dipoles are $\mu_C = 0.37$, $\mu_N = 0.30$, and $\mu_X \gg \mu_O$, and they point inward. Upon bending from linearity, the largest initial effect is the increase of the interaction energy of the atom dipole closer to the Fe, in the higher field. For O_2 and NO, the induction attraction to the outer O takes over, and the energy

begins to fall before an isosceles triangle is reached, which is roughly as far as we can go without van der Waals repulsions. In the case of O_2, the energy at that point is below that of the linear complex, so we predict, correctly, that Fe−O−O is bent, although we cannot obtain the angle without the rest of the potential. For NO, the energy remains higher, and we can reach no conclusion. The large C dipole on CO causes the energy to rise monotonically, and we predict, again correctly, that Fe−C−O is linear.

Our long-term aim is to describe the structure, dynamics, and spectroscopy of biomolecules with maximal exploitation of classical methods. In support of this approach, we have shown that classical theories of the VSE and of ligand binding and bending result in gratifying agreement with ab initio calculations and experiment, if induction is adequately included. The VSE follows from bonded neighbor induction, and binding usually described in terms of back-bonding of a metal ion to a ligand, can be described in a simpler way through mutual induction of the metal and the ligand.

The resulting picture of binding does not involve a covalent bond. The textbook alternative is an "ionic bond" but with Coulomb repulsion that is impossible: We propose that the true alternative is a far richer "electrostatic bond", capable of describing molecular recognition. A theory of complex biochemical systems including a systematic, self-consistent treatment of induction will demonstrate the broad applicability of classical mechanics, including some aspects of chemical bonding, and correct, environment-dependent vibrational frequencies.

Acknowledgment. This research was supported by the National Science Foundation (Grant No. CHE-0352026). We thank Dr. Lewyn Li and Dr. Christian Burnham for suggestions on the manuscript.

References and Notes

(1) Liptay, W. *Angew. Chem., Int. Ed. Engl.* **1969**, *8*, 177.
(2) Ahlborn, H.; Space, B.; Moore, P. B. *J. Chem. Phys.* **2000**, *112*, 8083.
(3) (a) Felder, C.; Applequist, J. *J. Chem. Phys.* **1981**, *75*, 2390. (b) Atkins, P.; Jones, L. *Chemical Principles: The Quest for Insight*; W. H. Freeman: New York, 2004; Chapter 2.14.
(4) Park, E. S.; Andrews, S. S.; Hu, R.; Boxer, S. G. *J. Phys. Chem. B* **1999**, *103*, 9813.
(5) Park, E. S.; Boxer, S. G. *J. Phys. Chem. B* **2002**, *106*, 5800.
(6) (a) Lambert, D. K. *Phys. Rev. Lett.* **1983**, *26*, 2106. (b) Lambert, D. K. *Solid State Commun.* **1984**, *51*, 297. (c) Lambert, D. K. *J. Chem. Phys.* **1988**, *89*, 3847.
(7) Williams, R. B.; Loring, R.; Fayer, M. D. *J. Phys. Chem. B* **2001**, *105*, 4068.
(8) Merchant, K.; Xu, Q.; Thompson, D.; Fayer, M. D. *J. Phys. Chem. A* **2002**, *106*, 8839.
(9) Merchant, K. A.; Noid, W. G.; Akiyama, R.; Finkelstein, I. J.; Goun, A.; McClain, B. L.; Loring, R.; Fayer, M. D. *J. Am. Chem. Soc.* **2003**, *125*, 13804.
(10) Eaves, J.; Tokmakoff, A.; Geissler, P. *J. Phys. Chem. A* **2005**, *109*, 9424.
(11) Hush, N. S.; Reimers, J. R. *J. Chem. Phys.* **1995**, *99*, 15798.
(12) Reimers, J. R.; Hush, N. S. *J. Phys. Chem. A* **1999**, *103*, 10580.
(13) Mankoo, P. K.; Keyes, T. *J. Chem. Phys.* **2006**, *124*, 204503.
(14) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.
(15) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: New York, 1997.
(16) Bridge, N. J.; Buckingham, A. D. *Proc. R. Soc. London, Ser. A* **1966**, *295*, 334.
(17) Voisin, C.; Cartier, A. *J. Mol. Struct. (THEOCHEM)* **1993**, *286*, 35.
(18) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497.
(19) (a) Huffaker, J. N. *J. Chem. Phys.* **1976**, *64*, 3175. (b) Huffaker, J. N. *J. Chem. Phys.* **1976**, *64*, 4564.
(20) Hermansson, K. *J. Chem. Phys.* **1993**, *99*, 861.
(21) Brewer, S. H.; Franzen, S. *J. Chem. Phys.* **2003**, *119*, 851.
(22) Dalosto, S. D.; Verkooi, J. M.; Sharp, K. A. *J. Phys. Chem. B* **2004**, *108*, 6450.
(23) (a) Burnham, C. J.; Xantheas, S. S. *J. Chem. Phys.* **2002**, *116*, 5115. (b) Fanourgakis, G.; Xantheas, S. *J. Chem. Phys.* **2006**, *124*, 174504.
(24) Andrews, S. S.; Boxer, S. G. *J. Phys. Chem. A* **2000**, *104*, 11853.
(25) Kuriyan, J.; Wilz, S.; Karplus, M.; Petsko, G. *J. Mol. Biol.* **1986**, *192*, 133.
(26) Kuczera, K.; Kuriyan, J.; Karplus, M. *J. Mol. Biol.* **1990**, *213*, 351.
(27) Lim, M.; Jackson, T. A.; Anfinrud, P. A. *Science* **1995**, *269*, 962.
(28) Loring, R. F. Personal communication.
(29) Moller, J. K. S.; Skibsted, L. H. *Chem. Rev.* **2002**, *102*, 1167.
(30) Rovira, C.; K. Kune, K.; Hutter, J.; Ballone, P.; Parinello, M. *J. Phys. Chem. A* **1997**, *101*, 8914.
(31) Nutt, D. R.; Karplus, M.; Meuwly, M. *J. Phys. Chem. B* **2005**, *109*, 21118.
(32) Keyes, M. H.; Falley, M.; Lumry, R. *J. Am. Chem. Soc.* **1971**, *93*, 2035.
(33) Projahn, H.-D.; Eldik, R. *Inorg. Chem.* **1991**, *30*, 3288.
(34) Rovira, C. *J. Phys.: Condens. Matter* **2003**, *15*, S1809.
(35) Meuwly, M.; Becker, O. M.; State, R.; M. Karplus, M. *Biophys. J.* **2002**, *98*, 183.
(36) Jameson, G. B.; Rodley, G. A.; Robinson, W. T.; Gagne, R. R.; Reed, C.; Collman, J. P. *Inorg. Chem.* **1978**, *17*, 850.
(37) Yamaguchi, Y.; Frisch, M.; Gaw, J.; Schaefer, H. F., III; Binkley, J. S. *J. Chem. Phys.* **1986**, *84*, 2262.
(38) Bilingsley, F. P. *J. Chem. Phys.* **1975**, *84*, 864.
(39) Drake, M. C. *Opt. Lett.* **1984**, *7*, 435.
(40) Konowalow, D. D.; Hirschfelder, J. O. *Phys. Fluids* **1961**, *4*, 637.
(41) Nutt, D. R.; Meuwly, M. *ChemPhysChem* **2004**, *5*, 1710.