

The Functional Form of Angular Forces around Transition Metal Ions in Biomolecules

A. E. Carlsson and S. Zapata

Department of Physics, Washington University, St. Louis, Missouri 63130-4899.

ABSTRACT A method for generating angular forces around σ -bonded transition metal ions is generalized to treat π -bonded configurations. The theoretical approach is based on an analysis of ligand-field and small-cluster Hamiltonians based on the moments of the electron state distribution. The functional forms that are obtained involve a modification of the usual expression of the binding energy as a sum of ligand–ligand interactions, which, however, requires very little increase in CPU time. The angular interactions have simple forms involving sin and cos functions, whose relative weights depend on whether the ligands are σ - or π -bonded. They describe the ligand-field stabilization energy to an accuracy of about 10%, and the interaction energy of covalently bonded systems to an accuracy of better than 4%. The resulting functional forms for the force field are used to model the structure of small clusters, including fragments of the copper blue protein structure. Large deviations from the typical square copper coordination are found when π -bonded ligands are present.

INTRODUCTION

Atomic-level simulations of proteins interacting with transition-metal ions have the promise of elucidating a wide variety of biophysical phenomena (Kaim and Schwederski, 1994). Such simulations can help establish both the native structure and reaction paths of metalloproteins and explain the roles of metal ions in protein folding and stability. Similar applications can be seen for simulations of metal ions interacting with DNA and RNA. However, the utility of such simulations depends critically on the availability of accurate but computationally tractable force fields for metals interacting with proteins. There are no quantitatively accurate force fields for these purposes, and it is probably not possible to construct quantitatively accurate force fields of a simple enough form to be useful. However, one can hope to generate force fields that include the basic physical effects at a level of accuracy sufficient to ascertain chemical trends as metals or ligands vary. The situation in this regard is best for “simple” metal ions that have no partly occupied shells. Here, an ionic picture based on electrostatics, supplemented by empirical repulsive forces, may hope to treat the most important physical effects. Even here, though, care should be taken, because the bond is not completely ionic but has some partial covalent character. For transition metals, the situation is much more difficult. The ligand-field splitting of the d -shell leads to important electronic contributions that cannot be ignored. These are manifested, for example, in the typical square or tetragonal coordination of Cu^{2+} complexes. Such effects cannot be described by radial interactions, but rather require the introduction of angular forces describing the energetics of the transition-metal d -

shell. Most existing methods for including angular forces in simulations of metal ions have used assumed functional forms (Comba et al., 1995; Allured et al., 1991; Timofeeva et al., 1995; Wiesemann et al., 1994; Sayle et al., 1995, 1997) based, for example, on the observed structures of small complexes. A recently developed method (Landis et al., 1998) treats the environment-dependent energetics of transition metals via a valence-bond approach. This method appears promising for covalently bonded metals. However, the bonding state of transition metals in proteins may be rather different from the sd^n hybridized configuration assumed in that work. In the present analysis, we seek to develop a method suitable for a broader range of chemical environments ranging from ionic to covalent.

The angular overlap model (AOM) (Gerloch et al., 1981; Jørgensen et al., 1963) provides a conceptual basis for treating ionically bound transition metals. This model provides a straightforward way of determining the parameters of an effective quantum-mechanical d - d Hamiltonian $h_d^{\mu\nu}$ in terms of the positions of the transition metal's ligand neighbors. The Hamiltonian parameters are given as simple angular functions of these positions. However, the total energy in the AOM does not have a simple analytic form. The ligand-field part of the energy is given as

$$E_{\text{LF}} = \sum_{\alpha} \epsilon_{\alpha}, \quad (1)$$

where the ϵ_{α} are the eigenvalues of $h_d^{\mu\nu}$ (and only the filled levels are included). The ϵ_{α} are not given as explicit functions of the $h_d^{\mu\nu}$, but are instead obtained using numerical matrix methods, except in certain highly symmetric cases. A total-energy method based on diagonalization of a d - d Hamiltonian obtained from the AOM has been combined with a classical molecular-mechanics method to study the structures of small molecules (Burton et al., 1995).

Our goal is to find an expression for the total energy that is simpler than that obtained by straightforward application of the AOM or more complete quantum-mechanical meth-

Received for publication 7 March 2000 and in final form 3 April 2001.

Address reprint requests to Anders E. Carlsson, Washington University, Department of Physics, Campus Box 1105, One Brookings Dr., Saint Louis, MO 63130. Tel.: 314-935-5739; Fax: 314-935-6219; E-mail: aec@howdy.wustl.edu.

© 2001 by the Biophysical Society

0006-3495/01/07/01/10 \$2.00

ods. We feel that this simplification is useful for four reasons:

1. Interpretability. The results of molecular-dynamics and energy-minimization calculations are much more useful if they can be easily interpreted in terms of the features of the underlying force fields. In standard methods, such as those based on the AOM, the matrix-diagonalization step provides a “black box” that obscures the energetic origins of the structural features that are found. By expressing the entire energy in a simple form, one can often associate characteristic atomic configurations with important minima in the real-space potential-energy function.
2. Ease of programming. One often wants to extend an existing force field to include transition-metal ions. The energy functions that we derive here are of a type not too different from those already present in standard force fields. For this reason, inclusion of these energy functions in such force fields is simpler than for straightforward applications of the AOM. We note that the present approach also yields some saving in CPU time, but, given that most biomolecules will contain transition metals only as a minority component, this savings is probably not a major factor.
3. Predicting structures of new molecules. A simple real-space picture of the force field allows one to have an a priori idea of what structures are likely to form around a transition-metal ion surrounded by a given combination of neighbors. In cases of a simple environment, such as a copper ion surrounded by four σ -bonded neighbors, one knows the likely structures even without examining the force field. However, in more complex environments, the constraints of the force field can provide very useful information.
4. Justification of classical descriptions of quantum-mechanical energetics. Many papers have been published that use classical energy descriptions to describe open-shell transition-metal systems. The type of analysis given in the present paper allows one to assess the accuracy of such calculations depending on the angular form of the interactions that are used.

We have recently shown (Carlsson, 1998) how the ligand-field splitting effects of σ -bonded transition-metal ions in ionic configurations can be described by a force field of a fairly simple form. One starts with an explicitly quantum-mechanical form for the electronic energy, and then solves the quantum-mechanical problem with systematic approximation methods to extract the real-space description of the electronic bonding energy. A second-order treatment of the hybridization terms between the ligand and metal orbitals is used to generate a ligand-field Hamiltonian for the d -electrons. The electronic bonding energy of this Hamiltonian is analyzed in terms of its moments. The moment analysis of the energy gives a “semiclassical” energy func-

tion with a simple trigonometric angular form. It is not precisely an additive function of ligand–ligand interactions, but is nearly as simple computationally. This energy function gives quite accurate energy results for σ -bonded ligands, with less than 10% error in the ligand-field stabilization energy (defined below).

This approach is extended in the present paper in three ways. First, we develop energy functions that treat π -bonded ligands as well. This results in new and distinct angular forms. Second, we extend the analysis to treat covalent environments. Finally, we present results for the lowest-energy structures of small clusters, which are a guide to understanding the bonding preferences of transition metals in proteins. The organization of the remainder of the paper is as follows. The next section develops the formalism underlying the angular forces. The following section gives tests of the functional form of the angular forces by comparing with results from diagonalization of simple cluster Hamiltonians. The subsequent section gives the small-cluster results for ideal examples and for an analog of the copper environment in blue-copper proteins.

Model

We model the electronic structure of the environment of a single transition-metal ion in terms of orbitals $|L, \mu\rangle$ localized on the ligands and $|M, \nu\rangle$ localized on the single metal ion. The orbitals are taken to be orthogonal for simplicity of calculation. Here, L denotes a particular ligand atom and the $|L, \mu\rangle$ are distinct orbitals on that atom. The Hamiltonian takes the form,

$$\hat{H} = \sum_{L, \mu} \varepsilon_{L, \mu} |L, \mu\rangle \langle L, \mu| + \sum_{\nu} \varepsilon_{M, \nu} |M, \nu\rangle \langle M, \nu| + \sum_{L, \mu, \nu} [h_{LM}^{\mu\nu} |L, \mu\rangle \langle M, \nu| + h_{ML}^{\nu\mu} |M, \nu\rangle \langle L, \mu|]. \quad (2)$$

Here, the ε terms are on-site energies for the orbitals, and the h terms are hybridization energies between the ligand orbitals and the metal orbitals. We ignore electron-interaction terms and explicit electrostatic effects. The energy associated with this Hamiltonian can then be obtained by diagonalizing its matrix as given in the $|L, \mu\rangle - |M, \nu\rangle$ basis, and taking the sum of the energies of the occupied eigenstates.

To obtain a real-space description of the energetics of this Hamiltonian, we treat the ionic and covalent cases separately. In the ionic case, we make the simplifying approximation that the dimensionless ratios $|h_{LM}^{\mu\nu}|/(\varepsilon_{M, \nu} - \varepsilon_{L, \mu})$ are small and can be used as expansion parameters. This allows the application of ligand-field theory, as in the AOM, which gives an effective Hamiltonian for the d -shell,

$$\hat{H}_d = \sum_{\mu, \nu} h_d^{\mu\nu} |M, \mu\rangle \langle M, \nu|, \quad (3)$$

where

$$h_d^{\mu\nu} = \sum_{L, \eta} (\varepsilon_d - \varepsilon_{L, \eta})^{-1} h_{ML}^{\mu\eta} h_{LM}^{\eta\nu}. \quad (4)$$

Note that we have assumed all of the d -orbitals on the transition-metal ion to have the same energies before ligand-field effects are “turned on.” The angular overlap model provides formulas for the $h^{\mu\nu}$ in terms of the angular and radial coordinates of the ligands.

In a fully quantum-mechanical ligand-field-theory calculation, one would numerically diagonalize the matrix \hat{H}_d , as in Burton et al. (1995). To obtain a force field with a simpler form, we instead work with the moments of \hat{H}_d , defined in terms of the traces of its powers, as follows. The first moment is the average energy of the d -complex, or

$$\begin{aligned}\bar{\varepsilon} &= \frac{1}{5} \text{Tr } \hat{H}_d \\ &= \frac{1}{5} \sum_{L,\eta} (\varepsilon_d - \varepsilon_{L,\eta})^{-1} \sum_{\mu} h_{ML}^{\mu\eta} h_{LM}^{\eta\mu} \\ &= \frac{1}{5} \sum_{L,\eta} (\varepsilon_d - \varepsilon_{L,\eta})^{-1} g_{\eta\eta},\end{aligned}\quad (5)$$

where

$$g_{\eta,\eta'} = \sum_{\mu} h_{L'M}^{\eta'\mu} h_{ML}^{\mu\eta}. \quad (6)$$

Thus $\bar{\varepsilon}$ is given as a sum of independent contributions from the ligands. Because the d -shell by itself has spherical symmetry, the contribution from each ligand is a radial function (no angular dependence) of the metal–ligand distance. Our major interest is in the angular terms resulting from the ligand-field splitting, so we do not consider the $\bar{\varepsilon}$ term further.

The width W of the d -complex corresponding to the ligand-field splitting is, in the simplest picture, described by the second moment or variance $\delta\varepsilon^2$ of the eigenvalues of \hat{H}_d . We expect that $W \propto \sqrt{\delta\varepsilon^2}$. We can obtain a simple real-space form for $\delta\varepsilon^2$ as follows. First, we note that $\delta\varepsilon^2 = \frac{1}{5} \sum_n (\varepsilon_n - \bar{\varepsilon})^2 = \frac{1}{5} \text{Tr}(\hat{H}_d - \bar{\varepsilon}\hat{I})^2$, where the ε_n are the eigenvalues of \hat{H}_d . From Eqs. 3, 4, and 6, we see that

$$\begin{aligned}\text{Tr } \hat{H}_d^2 &= \sum_{\mu,\nu} h_d^{\mu\nu} h_d^{\nu\mu} \\ &= \sum_{L,\eta,L',\eta'} (\varepsilon_d - \varepsilon_{L,\eta})^{-1} (\varepsilon_d - \varepsilon_{L',\eta'})^{-1} \\ &\quad \times \sum_{\mu} h_{L'M}^{\eta'\mu} h_{ML}^{\mu\eta} \sum_{\nu} h_{LM}^{\eta\nu} h_{ML}^{\nu\eta'} \\ &= \sum_{L,\eta,L',\eta'} (\varepsilon_d - \varepsilon_{L,\eta})^{-1} (\varepsilon_d - \varepsilon_{L',\eta'})^{-1} g_{\eta\eta'}^2.\end{aligned}\quad (7)$$

Then the variance is

$$\begin{aligned}\delta\varepsilon^2 &= \frac{1}{5} \text{Tr}(\hat{H}_d - \bar{\varepsilon}\hat{I})^2 = \frac{1}{5} (\text{Tr } \hat{H}_d^2 - 5\bar{\varepsilon}^2) \\ &= \frac{1}{5} \sum_{L,\eta,L',\eta'} (\varepsilon_d - \varepsilon_{L,\eta})^{-1} (\varepsilon_d - \varepsilon_{L',\eta'})^{-1} [g_{\eta\eta'}^2 - (\frac{1}{5})g_{\eta\eta}g_{\eta'\eta'}].\end{aligned}\quad (8)$$

We will focus on the ligand-field stabilization energy, which is defined as the sum of the energies of the occupied orbitals relative to $\bar{\varepsilon}$, or $E_{\text{LFSE}} = \sum_n' (\varepsilon_n - \bar{\varepsilon})$, where the sum

is over only the occupied eigenfunctions. We assume that E_{LFSE} is proportional to the ligand-field splitting W , and also identify W with the standard deviation $\sqrt{\delta\varepsilon^2}$. Thus, our expression for the ligand-field stabilization energy has the form

$$\begin{aligned}E_{\text{LFSE}} &= - \left\{ A/5 \sum_{L,\eta,L',\eta'} (\varepsilon_d - \varepsilon_{L,\eta})^{-1} (\varepsilon_d - \varepsilon_{L',\eta'})^{-1} \right. \\ &\quad \times \left. [g_{\eta\eta'}^2 - \frac{1}{5} g_{\eta\eta} g_{\eta'\eta'}] \right\}^{1/2},\end{aligned}\quad (9)$$

where A is a dimensionless constant that will later be used as a fitting parameter.

The accuracy of this form will be demonstrated in the next section through specific numerical experiments. At present, we will show how this form for E_{LFSE} results in an expression for the energy in terms of simple angular interactions between the ligands. We thus need to develop analytic forms for the $g_{\eta\eta'}$. First, we note that we can write

$$g_{\eta\eta'} = \langle L, \eta | \hat{H}_d \hat{P}_d^2 \hat{H}_d | L', \eta' \rangle = \langle \Psi | \Psi' \rangle, \quad (10)$$

where $\hat{P}_d = \sum_{\mu} |M, \mu\rangle \langle M, \mu|$ is the projection onto the d -subspace of the transition-metal ion. Here, we define $|\Psi\rangle = \hat{P}_d \hat{H} |L, \eta\rangle$, and we have used the relation $\hat{P}_d^2 = \hat{P}_d$, which holds because \hat{P}_d is a projection operator.

Consider first the case (Carlsson, 1998) where both the $|L, \eta\rangle$ and $|L', \eta'\rangle$ orbitals have σ -character with respect to the metal ion. Then they couple only to the d -orbitals $|M, \sigma\rangle$ and $|M, \sigma'\rangle$ that have σ -character with respect to the bond axes, so that $|\Psi\rangle = h_{\sigma} |M, \sigma\rangle$ and $|\Psi'\rangle = h_{\sigma'} |M, \sigma'\rangle$, where h_{σ} and $h_{\sigma'}$ are the appropriate coupling strengths. Thus $g_{\eta\eta'} = h_{\sigma} h_{\sigma'} \langle M, \sigma | M, \sigma' \rangle$, and the angular dependence is contained in the last inner product. However, this is simply a matrix element of a rotation about M , which carries L into L' . We note that $|M, \sigma\rangle$ is equivalent to $|M, m=0\rangle$, where $m=0$ denotes the angular dependence of the spherical harmonic Y_{2m} . Choosing our coordinate system so that L is along the z -axis and L' is in the z - x plane, we find that

$$\langle M, \sigma | M, \sigma' \rangle = D_{00}^{(2)}(0, \theta, 0) = (3 \cos^2 \theta - 1)/2, \quad (11)$$

where the D -term is a matrix element of the $l=2$ representation of the rotation group, and the second equality follows from the explicit formulas of the D -terms given in Wigner (1959). In summary, for σ -bonded ligands,

$$g_{\eta\eta'} = h_{\sigma} h_{\sigma'} (3 \cos^2 \theta - 1)/2. \quad (12)$$

For the case of a π -bonded orbital L and a σ -bonded orbital L' , we consider only the case in which the axis of the orbital lies along the circle connecting L and L' ; if it is perpendicular to this circle, L and L' have different inversion symmetries so their coupling vanishes. We write $|\Psi\rangle = h_{\pi} |M, \pi\rangle$ where $|M, \pi\rangle = (1/\sqrt{2})(-|M, m=1\rangle + |M, m=-1\rangle)$ and m is the usual azimuthal angular momentum index for the spherical harmonics. (The minus sign comes from

the definition of the spherical harmonics). Then, again choosing a coordinate system in which L is along the z -axis and L' is in the z - x plane, and following reasoning parallel to the σ -bonded case, we see that

$$\begin{aligned}\langle M, \pi | M, \sigma' \rangle &= (1/\sqrt{2})[-D_{10}^{(2)}(0, \theta, 0) + D_{-10}^{(2)}(0, \theta, 0)] \\ &= \sqrt{3} \sin \theta \cos \theta.\end{aligned}\quad (13)$$

and

$$g_{\eta\eta'} = \sqrt{3}h_{\pi}h'_{\sigma}\sin \theta \cos \theta. \quad (14)$$

Finally, we turn to the case of two π -bonded orbitals. To obtain the subsequent results, it is sufficient to consider the case in which the orbitals are parallel to the arc connecting L and L' , and the case in which they are perpendicular to it. By reasoning similar to that above, one sees that, in the first case,

$$\begin{aligned}\langle M, \pi | M, \pi' \rangle &= \frac{1}{2}[D_{11}^{(2)}(0, \theta, 0) + D_{-1-1}^{(2)}(0, \theta, 0) \\ &\quad - D_{1-1}^{(2)}(0, \theta, 0) - D_{-11}^{(2)}(0, \theta, 0)] \quad (15) \\ &= \cos 2\theta\end{aligned}$$

and

$$g_{\eta\eta'} = h_{\pi}h'_{\pi}\cos 2\theta. \quad (16)$$

In the second case,

$$\begin{aligned}\langle M, \pi | M, \pi' \rangle &= \frac{1}{2}[D_{11}^{(2)}(0, \theta, 0) + D_{-1-1}^{(2)}(0, \theta, 0) \\ &\quad + D_{1-1}^{(2)}(0, \theta, 0) + D_{-11}^{(2)}(0, \theta, 0)] \quad (17) \\ &= \cos \theta\end{aligned}$$

and

$$g_{\eta\eta'} = h_{\pi}h'_{\pi}\cos \theta. \quad (18)$$

Thus, combining Eqs. 9, 12, 14, 16, and 18, we find the following explicit form for E_{LFSE} as a sum of ligand–ligand interactions:

$$E_{\text{LFSE}} = - \sqrt{\sum_{L, \eta, L', \eta'} U(L, \eta; L', \eta')}, \quad (19)$$

where the ligand–ligand interaction U is given as:

For σ – σ interactions,

$$U(L, \eta; L', \eta') = (A/5)e_{\sigma}(r)e_{\sigma}(r')u_{\sigma\sigma}(\theta); \quad (20)$$

For π – σ interactions,

$$U(L, \eta; L', \eta') = (A/5)e_{\pi}(r)e_{\sigma}(r')u_{\pi\sigma}(\theta); \quad (21)$$

For π – π interactions,

$$U(L, \eta; L', \eta') = (A/5)e_{\pi}(r)e_{\pi}(r')u_{\pi\pi}(\theta); \quad (22)$$

where

$$e_{\sigma}(r) = h_{\sigma}^2/(\varepsilon_d - \varepsilon_{L, \eta}), \quad (23)$$

$$e_{\pi}(r) = h_{\pi}^2/(\varepsilon_d - \varepsilon_{L, \eta}), \quad (24)$$

$$u_{\sigma\sigma}(\theta) = (9 \cos^4 \theta - 6 \cos^2 \theta + \frac{1}{5})/4, \quad (25)$$

$$u_{\pi\sigma}(\theta) = (-3 \cos^4 \theta + 3 \cos^2 \theta - \frac{2}{5}), \quad (26)$$

$$u_{\pi\pi}(\theta) = (4 \cos^4 \theta - 3 \cos^2 \theta + \frac{1}{5}). \quad (27)$$

In each case, U is given as a product of radial terms involving the two ligands and a simple angular function. (Note that, although denoted an interaction here, U does not have units of energy because of the square root in Eq. 19). In the calculation of the terms involving π – σ interactions, each term involves a sum over two π -orbitals, parallel and perpendicular to the arc connecting the two ligands. In the latter case, $g_{\eta\eta'}$ in Eq. 9 vanishes, but $g_{\eta\eta}$ and $g_{\eta'\eta'}$ do not. In the calculation of the terms involving π – π interactions, one has a similar scenario except that one sums over two pairs of π -orbitals.

These forms are plotted out in Fig. 1 *A*. Note that the σ – σ interactions are fairly similar in form to the π – π interactions, both having pronounced minima at 180° (as well as the physically irrelevant one at 0°) and a shallower minimum at 90° . The σ – π interaction is complementary to these, having minima at 45° and 135° . We shall see later that these differences lead to large differences in ground-state structures of small clusters. For comparison, we show, in Fig. 1 *B*, an empirical angular interaction curve (Comba et al., 1995), assumed in some previous calculations of small transition-metal structures. The angular dependence is based on the observed square structure of small complexes of the transition metals of interest, and is proportional to $\sin^2 2\theta$. Because the metals that have square coordination generally have predominantly σ -bonds to their neighbors, the most relevant comparison is to the σ – σ curve in Fig. 1 *A*. We see that the behavior is quite different. The empirical curve has equivalent minima at 180° and 90° , whereas, in the theoretical curve, the 180° minimum is much deeper.

The above analysis provides theoretical support for the use of the form of Eq. 9 for the ligand-field stabilization energy in the ionic limit where the electronic structure can be described accurately by an effective Hamiltonian for the d -complex. We now derive an energy function of equal simplicity for covalently bonded systems. We begin with a Hamiltonian of the form of Eq. 2, in the covalent limit in which $\varepsilon_{L, \mu} = \varepsilon_{M, \nu} = 0$. (The choice of the common energy to be zero affects no physical results and is taken for algebraic convenience.) We derive the energy function using a moment analysis, which has been used previously to develop energy functions for elemental metals (Carlsson, 1990). The moments of the electronic density of states, projected on the d -shell, are given as

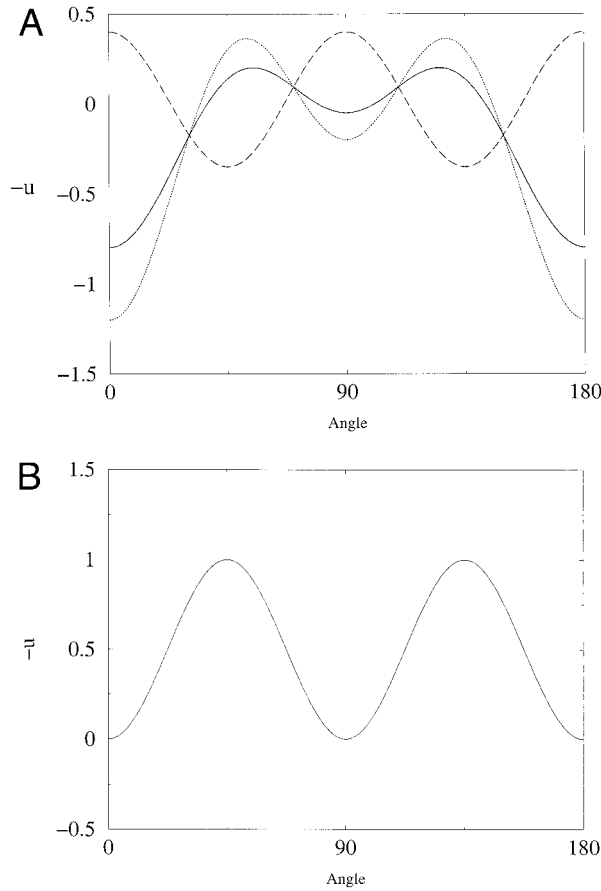


FIGURE 1 (A) Interaction between ligands of central transition metal ion. The dimensionless quantity u (cf. Eqs. 19–21) is either $u_{\sigma,\sigma}$ (solid line), $u_{\pi,\sigma}$ (dashed line), or $u_{\pi,\pi}$ (dotted line). The quantity $-u$ is plotted so that negative values will correspond to lower energies. (B) Empirical force field from Comba et al. (1995).

$$\langle \hat{H}^n \rangle_d = \frac{1}{5} \text{Tr } H_d^n = \frac{1}{5} \sum_{\alpha} \epsilon_{\alpha}^n \langle \alpha | \hat{P}_d | \alpha \rangle, \quad (28)$$

where the $|\alpha\rangle$ are the eigenfunctions of \hat{H} and the factor of $1/5$ simplifies the algebra. In the completely covalent case considered here, all of the eigenfunctions have equal squared amplitudes on the ligand and d -states, so that

$$\langle \alpha | \hat{P}_d | \alpha \rangle = \frac{1}{2}. \quad (29)$$

For the special case of no ligand-field splitting, in which all of the bonding states have the same energy $-\epsilon$, and all of the antibonding states have the energy ϵ , there are altogether 10 eigenfunctions (plus potentially other ligand orbital combinations that are not coupled to the metal ion). Using Eq. 29, one sees that, in this case,

$$\epsilon = \langle \hat{H}^2 \rangle_d^{1/2} = \langle \hat{H}^4 \rangle_d^{1/4}. \quad (30)$$

It has been established (Cyrot-Lackmann, 1968) that the moments can be written as a real-space sum over neighbor positions as

$$\langle \hat{H}^2 \rangle_d = \frac{1}{5} \sum_{L,\mu,\nu} h_{ML}^{\nu\mu} h_{LM}^{\mu\nu} \quad (31)$$

$$\langle \hat{H}^4 \rangle_d = \frac{1}{5} \sum_{L,L',\mu,\nu,\delta,\eta} h_{ML}^{\nu\mu} h_{LM}^{\mu\delta} h_{ML'}^{\delta\eta} h_{L'M}^{\eta\nu}. \quad (32)$$

The moments of odd order vanish when both on-site energies vanish. These expressions involve functions quite similar in form to those that appear in the ionic limit. Specifically,

$$\langle \hat{H}^2 \rangle_d = \frac{1}{5} \sum_{L,\eta} g_{\eta,\eta}, \quad \langle \hat{H}^4 \rangle_d = \frac{1}{5} \sum_{L,\eta,L',\eta'} g_{\eta,\eta'}^2. \quad (33)$$

We now derive an approximate energy function from these quantities. In the covalent case, the ligand-field stabilization energy is not as well defined as in the ionic case, because the d -complex is not well separated from the ligand orbitals. For this reason, we focus on the total interaction energy E_{int} between the metal ion and the ligands. With our choice of energy zero, this is simply the sum of the eigenvalues of all of the occupied states. To derive the energy function, we note that the assumption of only one bonding and one antibonding eigenvalue, described by Eq. 30, corresponds to the case in which $\langle \hat{H}^4 \rangle_d = \epsilon^4 = \langle \hat{H}^2 \rangle_d^2$. Thus a natural form for the energy would be one that begins with Eq. 30 and adds a correction based on $\langle \hat{H}^4 \rangle_d - \langle \hat{H}^2 \rangle_d^2$. The simplest form for the total energy based on dimensional analysis is

$$E_{\text{int}} = -A \sqrt{\langle \hat{H}^2 \rangle_d} - B \sqrt{[\langle \hat{H}^4 \rangle_d - \langle \hat{H}^2 \rangle_d^2] \langle \hat{H}^2 \rangle_d}, \quad (34)$$

where A and B are coefficients to be determined. The second term is taken to have a square-root dependence because, in systems with narrow bonding and antibonding complexes, the width of these complexes is proportional to $\sqrt{\langle \hat{H}^4 \rangle_d - \langle \hat{H}^2 \rangle_d^2}$. (One readily shows this by taking a model density of states consisting of two rectangular subbands, each of width W , centered at energies $+\epsilon_0$ and $-\epsilon_0$. In this case, straightforward calculation shows that $\langle \hat{H}^4 \rangle_d - \langle \hat{H}^2 \rangle_d^2 = 4\epsilon_0^2 W^2 + 4W^4/45$, which reduces to $4\epsilon_0^2 W^2$ for narrow bands.) We also find that this provides an excellent fit to the quantum-mechanical results for small clusters given in the following section.

Using Eqs. 33, we have for the total energy

$$E_{\text{int}} = -A \left\{ \frac{1}{5} \sum_{L,\eta} g_{\eta,\eta} \right\}^{1/2} - B \left\{ \sum_{L,\eta,L',\eta'} [g_{\eta,\eta'}^2 - \frac{1}{5} g_{\eta,\eta} g_{\eta',\eta'}] / \left[\sum_{L,\eta} g_{\eta,\eta} \right] \right\}^{1/2}. \quad (35)$$

This form is quite similar to Eqs. 9 and 19 but differs in two ways. First, the term based on $\langle \hat{H}^2 \rangle_d$ makes a contribution involving a sum of single-ligand contributions. This term is not sensitive to the angular character of the local environment, because it is determined entirely by the distances and

types of the ligands. Second, the $\langle \hat{H}^4 \rangle_d$ term is divided by a factor involving the $\langle \hat{H}^2 \rangle_d$ term. Computationally, these differences incur no significant extra costs. We note that, in distinguishing potential structures with similar distances and types of neighbors, the energy functions, Eqs. 9 or 19 and 35, are equivalent, because $(\sum_{L,\eta} g_{\eta\eta})$ is determined by the neighbor distances and types. Thus, the dependence of the structure on the angular aspects of the environment is given by a function of the form Eq. 19 in either case, and this is used in the next section.

A physical system will be somewhere between the limits of complete ionicity and complete covalency. A suitable function for interpolating between these two limits is obtained by replacing the quantity $(\sum_{L,\eta} g_{\eta\eta})$ in the first term and in the denominator of the second term of Eq. 35 by $(\sum_{L,\eta} g_{\eta\eta}) + (\epsilon_d - \epsilon_{L,\eta})^2$. With this form, one sees that, in the ionic limit (large $\epsilon_d - \epsilon_{L,\eta}$), the second term of Eq. 35 becomes equivalent to Eq. 9, and the first term simply describes the average energy of the d -complex. One can also show (Carlsson, unpublished) that, for the case in which the widths of the bonding and antibonding complexes go to zero (no ligand-field splitting), the first term of Eq. 35 thus modified gives the correct interaction energy for all values of the ionicity.

TESTS OF FUNCTIONAL FORM

To evaluate the accuracy of the semiclassical form of Eq. 19 for the ligand-field stabilization energy, we have performed explicit tests for small clusters. These clusters consist of a central transition-metal ion with four neighbors placed at random orientations at random distances relative to the central ion. We consider only the minority-spin orbitals, because these determine E_{LFSE} for high-spin late-transition metals. Four electrons are placed in these states, corresponding to Cu^{2+} ; other band-filling values give similar results. The random distances are taken into account by varying the couplings h_σ and h_π uniformly over a finite interval ranging from zero to h_{max}^σ or h_{max}^π . We take $h_{\text{max}}^\pi = 0.5h_{\text{max}}^\sigma$. The energies are obtained by explicit diagonalization of a tight-binding Hamiltonian for this cluster. All of the ligand orbitals are taken to have the same value of $\epsilon_{L,\eta}$. The only dimensionless variable that enters the results is then $\gamma = |h_{\text{max}}/(\epsilon_d - \epsilon_{L,\eta})|$. For small values of γ , the bonding is primarily ionic, and, for larger values, it acquires more covalent character.

In our tests of the energy function for ionic systems, we use $\gamma = 0.1$. We fit the energies of the clusters to the semiclassical form for the energy,

$$E_{\text{LFSE}}^2 = \sum_{L,\eta,L',\eta'} U(L, \eta; L', \eta') + B, \quad (36)$$

involving two parameters A and B (where U contains A). We use a database of 10,000 clusters to determine the parameters, and then test them on a set of 1000 clusters not

included in the “training” set. Typical results are shown in Fig. 2, which shows the semiclassical energies versus exact energies for 1000 clusters. Figure 2 *A* corresponds to three σ -bonded ligands and one π -bonded one, whereas Fig. 2 *B* corresponds to four π -bonded ligands. The rms errors in E_{LFSE} for these two cases (and for the cases of the two and three π -bonded ligands) are between 9% and 10%. This is to be compared with rms errors of 25–30% that are obtained with empirical force fields (Carlsson, 1998).

In the tests for covalent systems, we use the same clusters but with $\gamma = \infty$ corresponding to vanishing energy differences. In this case, we use an energy function of the form given by Eq. 35, and vary A and B to obtain the best description of E_{int} . For clusters with zero, one, two, three, and four π -bonded ligands, the rms errors in E_{int} are 2.6%, 3.3%, 3.2%, 3.8%, and 3.6%, respectively. Figure 2 *C* and *D*, show the energy data for the cases of three σ -bonded ligands and one π -bonded one, and four π -bonded ones. In comparing these results with those for the ionic limit, one should note that, although the percent errors are smaller, the values of E_{int} are, in general, larger in magnitude than those of E_{LFSE} , so that the absolute errors may not necessarily be smaller. Nevertheless, an accuracy of a few percent in E_{int} is quite encouraging. To compare the errors in E_{LFSE} and E_{int} , consider a case in which the eigenvalues coming from the ligand and metal orbitals are uniformly spread out from the top to the bottom of the complex. If one considers the d -complex as being the top half of this complex, and the holes are concentrated in the top quarter of the complex, then the magnitude of E_{LFSE} will be roughly a third of E_{int} . The absolute errors in the covalent and ionic limits cases are then roughly comparable.

SMALL CLUSTER MINIMUM-ENERGY STRUCTURES

In this section, we describe some of the implications of the angular forms developed above for the structure of small model clusters consisting of a metal atom and four ligands. Such cluster calculations cannot treat a protein environment accurately; this would await parameterization and incorporation of the force field into protein codes, which is in progress in our group. However, from the small-cluster calculations it is possible to see the structural preferences of the angular forces by themselves. These can have a large impact on the choice of binding sites by particular ligands, and on the corresponding affinities. We emphasize that the goal of these calculations is to obtain the simplest description of the energetics that is both reasonably accurate and physically justifiable. Therefore, we expect that more elaborate methods, such as a complete application of the AOM, would give results similar to the present ones.

To keep the calculations as simple as possible, we place the ligands at frozen bond lengths from the central metal ion. Here “frozen” means that they are fixed at a given set

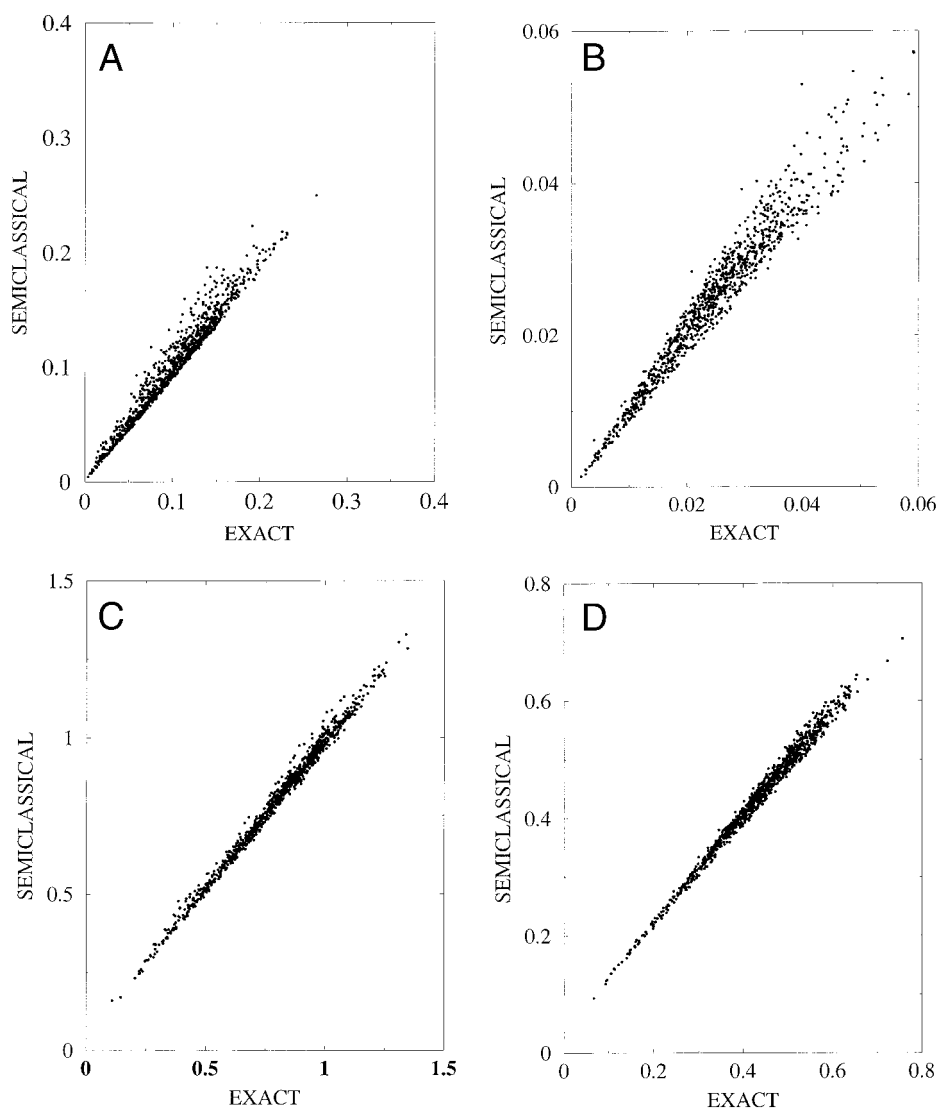


FIGURE 2 Energies obtained by semiclassical force field versus exact energies for small model clusters. Energies given in units of h_{\max} , maximal coupling strength between ligands and transition metal. (A) One π -bonded ligand and three σ -bonded ligands, ionic system. (B) Four π -bonded ligands, ionic system. (C) One π -bonded and three σ -bonded ligands, covalent system. (D) Four π -bonded ligands, covalent system.

of values, but these values are not necessarily the same for all of the ligands. The energy terms include the ligand-field stabilization energy as described by Eq. 19, and a repulsive radial interaction between ligands forbidding close approaches. (We note that, because the bond lengths are fixed, the $\langle \hat{H} \rangle^2$ terms in Eq. 34 do not vary during the simulation, and, as pointed out in the second section, Eq. 19 can then be used to describe both covalent and ionic systems.) The repulsive ligand–ligand term is needed because otherwise spurious structures involving 45° bond angles can appear when π -bonding ligands are present. To evaluate the magnitudes of the couplings, we assume a value of 100 kJ/mol = 1.04 eV for the ligand-field stabilization energy of a cluster with four σ -ligands in square coordination; this number would correspond to experimental values for the

more strongly stabilized complexes (Cotton and Wilkinson, 1972). For clusters with unequal bond lengths, we assume an exponential decay, so that $e_\sigma(r) = e_\sigma(r_0)\exp[-\kappa(r - r_0)]$, where r_0 is the reference distance and κ is a decay parameter. Because A multiplies the couplings e_σ and e_π , any changes in A can be compensated for by changing the overall magnitudes e_σ and e_π . For simplicity, we therefore take $A = 1$. Using Eq. 19 and the above value of E_{LFSE} , we then find that $e_\sigma(r_0) = 0.79$ eV. In the blue-copper protein fragment that we treat, the π -bonding effects are quite strong, and, in the absence of reliable estimates for e_π , we take $e_\sigma = e_\pi$. The effects of varying the couplings are discussed below. Because the most prominent case that we consider of a distant ligand is sulfur, we identify κ with the spatial decay rate coming from the measured first ionization

energy E_1 of sulfur, 10.4 eV (= 1000 kJ/mol), using the formula $\hbar^2\kappa^2/2m = E_1$. This yields $\kappa = 1.65 \text{ \AA}^{-1}$. The ligand–ligand terms contain an exponential term taken as the repulsive part of the van der Waals interactions as given in the MM2 force field (Sprague et al., 1987). In our model

clusters, we use parameters and bond lengths typical for copper or nickel interacting with nitrogen or sulfur ligands. This is because copper and nickel have the largest ligand-field energies among the $3d$ transition metals, and typical ligands for these metals are nitrogen and sulfur. We have

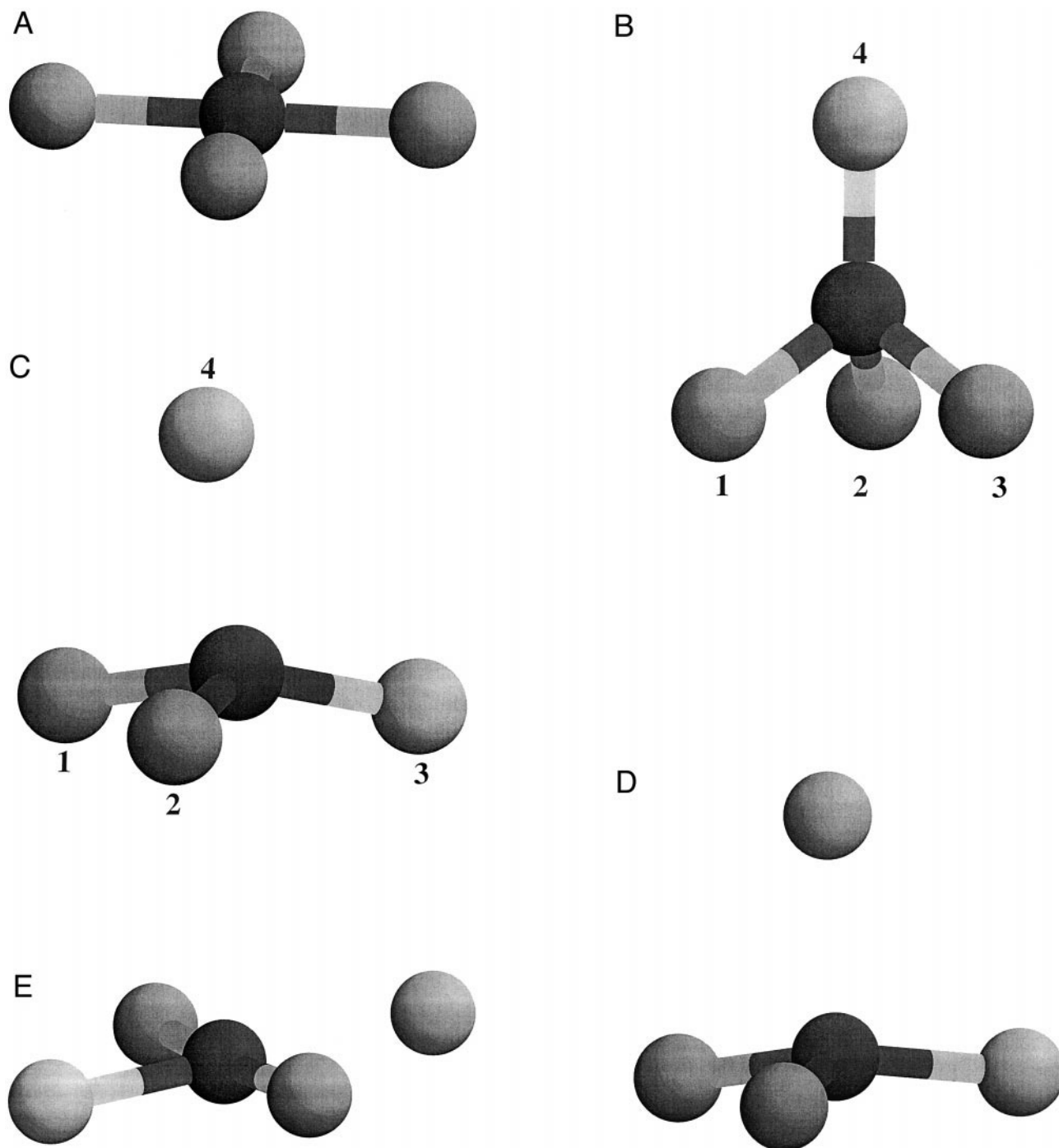


FIGURE 3 Lowest-energy structures obtained in relaxations of five-atom cluster using angular forces. (A) Four equivalent σ -bonded neighbors. (B) Three σ -bonded neighbors and one π -bonded neighbor 4, all at the same distance. (C) Two σ -bonded nitrogen neighbors 1 and 2, a near π -bonded sulfur neighbor 3, and a far σ -bonded neighbor 4. (D) Structure for blue-protein model obtained by quantum-chemical calculations (Ryde et al., 1996). (E) Three near σ -bonded neighbors and a far σ -bonded neighbor.

examined three simple geometries; the corresponding minimum-energy clusters are shown in Fig. 3 *A–C*.

1. A cluster with all four ligands σ -bonded and placed at equal distances of 2.0 Å. We take repulsion parameters appropriate for nitrogen. The result is a square-planar cluster (cf. Fig. 3 *A*), consistent with the observed structures of many copper and nickel complexes. With this parameterization, the square structure is quite strongly favored over the tetrahedral one, by 0.31 eV or 7.0 kcal/mole. The square structure remains stable under increases of the ligand–ligand repulsion strengths of up to a factor of over three. We note, however, that the stability of the square structure is strongly sensitive to the choice of repulsive terms. For example, we find that, when the van der Waals term from the OPLS-II force field (Jorgensen and Tirado-Rives, 1988) is used instead of the MM2 form, an increase of the repulsion strengths of only 20% is enough to stabilize the tetrahedral structure. We expect that input from fully quantum-mechanical total-energy calculations will be necessary to precisely pin down the competition between the electronic energy and the repulsion terms in determining the structural energy differences.
2. A cluster similar to that of 1, but with one π -bonded ligand. This leads to a substantial deviation from planarity, as shown in Fig. 3 *B*, where the π -bonded ligand is atom 4. As expected from the maximum of the π – σ potential in Fig. 1, the angle between atom 4 and atoms 1–3 is greater than 90°, about 123°. The angle among the σ -bonded atoms 1, 2, and 3, is slightly greater than the ideal value of 90°, because of the repulsive energy term.
3. A cluster with two σ -bonded nitrogen ligands 1 and 2, one π -bonded sulfur ligand 3, and a σ -bonded sulfur ligand 4. The distances are 2.04 Å for the nitrogen ligands, 2.18 Å for the π -bonded sulfur ligand, and 2.64 Å for the σ -bonded sulfur ligand. This geometry is motivated by the observed geometry of copper sites in blue-copper proteins. In these proteins, the nitrogen ligands and a cysteine sulfur ligand are close to the copper, and it is believed that the cysteine sulfur ligand is primarily π -bonded. An additional σ -bonded methionine sulfur ligand is farther from the copper. We note that experiments and calculations indicate that the copper binding in this system has a high degree of covalency (Solomon et al., 2000). However, because Eqs. 19 and 35 are equivalent for energy minimization with fixed bond lengths, this is not a problem. The values of the ligand distances here are taken from cluster calculations for blue-protein models (Ryde et al., 1996). As above, the van der Waals parameters are taken from the OPLS-II parameter set, but we do not have a reliable procedure for determining the differences between the ligand–ion interactions of the nitrogen and the sulfur ligands. Because the greater distance of the sulfur at 2.18 Å from the

copper will be compensated to some extent by the greater size of sulfur relative to nitrogen, we simply assume that the three close ligands have the same coupling strength to the central atom. We have varied the ratio of the sulfur-to-nitrogen coupling strengths by up to 30% in both directions and found changes of only a few degrees in the bond angles of the cluster. For the far ligand, we use the scaling procedure described above, which leads to a coupling that has 50% of the strength of the close ligands. We have varied this coupling from 0% to 75% of the close-ligand value, and again found bond-angle changes of only a few degrees. The lowest-energy structure for this cluster is indicated in Fig. 3 *C*. The geometry is trigonal, with bond angles of 125° between the cysteine sulfur and the nitrogens; the nitrogen–nitrogen bond angle is 91°. By comparison, the optimal values for protein models obtained by quantum-chemical calculations (Ryde et al., 1996) are 125° for the cysteine sulfur–nitrogen bond angle and 103° for the nitrogen–nitrogen bond angle. The results obtained by the present method, of course, do not have quantitative accuracy, but the overall structure of the coordination shell is quite similar to that obtained by Ryde et al. (1996), which is shown in Fig. 3 *D*. The formation of this structure is not due to the bond lengths that we used as input. To demonstrate this, we have considered a cluster with the same bond lengths as the blue-protein fragment, with all ligands σ -bonded. In this case, the structure is planar, as shown in Fig. 3 *E*. Thus the formation of the trigonal structure is directly related to the special character of the angular interactions associated with π -bonding.

CONCLUSION

We have seen that it is possible to extract the functional form of angular forces around transition metals in biomolecules by using an approximate treatment of the ligand-field Hamiltonian in ionic systems or the cluster Hamiltonian for covalent systems. The resulting angular forms have simple trigonometric forms, which are very different for π -bonding as compared to σ -bonding. Ligand-field energies in ionically bonded systems are represented well by these angular forces, and much better than by force fields with assumed functional forms. The interaction energies of covalently bonded systems are also obtained accurately. Future work should aim to include the functional forms derived here in widely used biomolecular simulation packages.

This work was supported by the National Science Foundation under Grant DMR-9971476. We appreciate informative conversations with Garland Marshall and Jay Ponder.

REFERENCES

- Allured, V. S., C. M. Kelly, and C. R. Landis. 1991. Shapes empirical force field: new treatment of angular potentials and its application to square-planar transition-metal complexes. *J. Am. Chem. Soc.* 113:1–12.
- Burton, V. J., R. J. Deeth, C. M. Kemp, and P. J. Gilbert. 1995. Molecular mechanics for coordination complexes—the impact of adding *d*-electron stabilization energies. *J. Am. Chem. Soc.* 117:8407–8415.
- Carlsson, A. E. 1990. Beyond pair potentials in transition metals and semiconductors. In *Solid State Physics: Advances in Research and Applications*, Vol. 43. H. Ehrenreich and D. Turnbull, (editors). Academic, New York. 1–91.
- Carlsson, A. E. 1998. Angular forces around transition metals in biomolecules. *Phys. Rev. Lett.* 81:477–480.
- Comba, P., T. W. Hambley, and M. Ströhle. 1995. The directionality of *d*-orbitals and molecular-mechanics calculations of octahedral transition-metal compounds. *Helv. Chim. Acta.* 78:2042–2047.
- Cotton, F. A., and G. Wilkinson. 1972. *Advanced Inorganic Chemistry*. Interscience, New York. 595.
- Cyrot-Lackmann, F. 1968. Sur le calcul de la cohésion et de la tension superficielle des métaux de transition par une méthode de liaisons fortes. *J. Phys. Chem. Solids.* 29:1235–1243.
- Gerloch, M., J. H. Harding, and R. G. Woolley. 1981. The context and application of ligand field theory. *Struct. Bond. (Berlin)* 46:1–46.
- Jørgensen, C. K., R. Pappalardo, and H. H. Schmidtke. 1963. Do the “ligand-field” parameters in the lanthanides represent weak covalent binding? *J. Phys. Chem.* 39:1422–1430.
- Jørgensen, W. L., and J. Tirado-Rives. 1988. The OPLS potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.
- Kaim, W., and B. Schwederski. 1994. *Bioinorganic Chemistry: Inorganic Elements in the Chemistry of Life*. Wiley, New York.
- Landis, C. R., T. Cleveland, and T. K. Firman. 1998. Valence bond concepts applied to the molecular mechanics description of molecular shapes. 3. Applications to transition metal alkyls and hydrides. *J. Am. Chem. Soc.* 120:2641–2649.
- Ryde, U., M. H. M. Olsson, K. Pierloot, and B. O. Roos. 1996. The cupric geometry of blue copper proteins is not strained. *J. Mol. Biol.* 261: 586–596.
- Sayle, D. C., M. A. Perrin, P. Nortier, and C. R. A. Catlow. 1995. Simulation study of copper(I) and copper(II) species in ZSM-5 zeolite. *J. Chem. Soc. Chem. Commun.* 9:945–947.
- Sayle, D. C., C. R. A. Catlow, J. D. Gayle, M. A. Perrin, and P. Nortier. 1997. Computer modeling of the active-site configurations within the NO decomposition catalyst CU-ZSM-5. *J. Phys. Chem.* 101:3331–3337.
- Solomon, E. I., D. W. Randall, and T. Glaser. 2000. Electronic structures of active sites in electron transfer metalloproteins: contributions to reactivity. *Coord. Chem. Rev.* 200:595–632.
- Sprague, J. T., J. C. Tai, Y. Yuh, and N. L. Allinger. 1987. The MMP2 calculational method. *J. Comput. Chem.* 8:581–603.
- Timofeeva, T. V., J. H. Lii, and N. L. Allinger. 1995. Molecular mechanics explanation of the metallocene bent sandwich structure. *J. Am. Chem. Soc.* 117:7452–7459.
- Wiesemann, F., S. Teipel, B. Krebs, and U. Howeler. 1994. Force-field calculations on the structures of transition-metal complexes. 1. Application to copper(II) complexes in square-planar coordination. *Inorg. Chem.* 33:1891–1898.
- Wigner, E. P. 1959. *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*. Academic Press, New York, 166–170.