Incorporating Protein Environments in Density Functional Theory: A Self-Consistent Reaction Field Calculation of Redox Potentials of [2Fe2S] Clusters in Ferredoxin and Phthalate Dioxygenase Reductase

Jian Li,* Melanie R. Nelson, Chun Y. Peng, Donald Bashford, and Louis Noodleman*

Department of Molecular Biology, TPC-15, The Scripps Research Institute, La Jolla, California 92037

Received: January 12, 1998; In Final Form: March 18, 1998

An approach to calculating molecular electronic structures of active-site clusters in the presence of protein environments has been developed. The active-site cluster is treated by density functional theory. The protein field, together with the reaction field arising mainly from solvent, is obtained from a finite-difference solution to the Poisson−Boltzmann equation with three dielectric regions, and then these are coupled to the density functional calculation by a self-consistent iterative procedure. The method is applied to compute redox potentials of ferredoxin from *Anabaena* 7120 and phthalate dioxygenase reductase (PDR) from *Pseudomonas cepacia*, both having similar $[Fe_2S_2(SR)_4]$ active-site clusters. The calculated redox potentials, −1.007 V and −0.812 V in 0.05 M ionic strength for ferredoxin and PDR, respectively, deviate significantly from experimental values of −0.440 and −0.174 V. However, the calculated data reproduce the experimental trend fairly well. The calculated redox potential for PDR is 195 mV more positive than that for ferredoxin, comparing very well with the experimental value of 266 mV. The energy decomposition scheme reveals that the protein field plays a key role in differentiating the redox potentials of these two proteins.

I. Introduction

Many chemical processes important to biological systems occur in the presence of protein and solvent. The enormous size and complexity of such reaction systems present a major challenge to theoretical models and simulations.[1] Since these processes usually involve bond breaking/forming, atom migration, protonation/deprotonation, and/or electron transfer, understanding these requires a quantum chemical treatment. The protein and solvent environment impose significant effects on the reactivities of reaction centers and are therefore inseparable parts of the systems that should not be excluded from the theoretical description. A common strategy, therefore, is to partition the system into a quantum region and a classical region, as shown in Figure 1a. The quantum portion contains only the active site or the most relevant part of the system and is treated by an appropriate quantum mechanical scheme. The remaining parts of the system, the protein environment and solvent, are considered as classical regions and are handled either by force field techniques[2] or as dielectric continua (as in this paper). Such a coupled quantum-classic approach has been widely employed to study chemical reactions in solution and catalytic processes in enzymes.[2]

We have recently developed a combined density functional and continuum dielectric theory to incorporate solvation effects into electronic structure calculations.[3] In this combined method, the solute molecule is computed by density functional theory in the presence of a solvent reaction field. The reaction field is evaluated from a finite-difference solution to the Poisson−Boltzmann equation and self-consistency between the reaction field potential and the electronic structure of solute is achieved by iteration. This method has been successfully applied to calculate solvation energies, acidities, and redox potentials of organic species, metal cations and clusters, and a model of an enzyme active site in aqueous solution.[3,4] Like many other approaches based on continuum dielectric theory,[5] this method only dealt with two dielectric constants, one for the solute (in a vacuum) and the another for the entire bulk surrounding. The remaining protein (where present), aside from the side chains in direct contact with the transition metal site, was omitted.

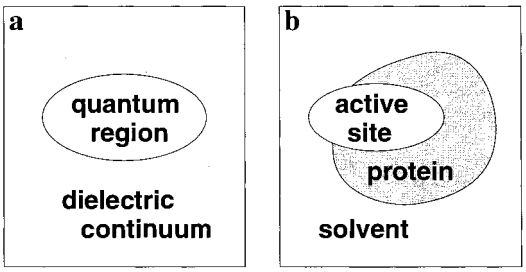


Figure 1. Partition of a molecular system: (a) two-dielectric constant partition; (b) multidielectric constant partition.

In this paper, we present our first work incorporating a protein environment into the combined density functional and continuum dielectric theory. Based on the macroscopic electrostatics with atomic detail (MEAD) model,[6] which solves the Poisson−Boltzmann equation for a multidielectric constant system, both protein field and solvent reaction field can be added to the molecular Hamiltonian in the self-consistent density functional calculation. This approach gives a more complete model for systems such as enzymes in which the active site is embedded in the protein but can also be accessed by solvent molecules (see Figure 1b). A similar scheme, the inhomogeneous self-consistent reaction field theory, has been explored by Tapia and Johannin in 1981.[2a] The method includes protein core effects in quantum chemical calculations at semiempirical CNDO−INDO level and was applied to study the proton relay system of liver alcohol dehydrogenase. As our first test, we

* Corresponding authors. Fax (619) 784-8896; Email lou@scripps.edu or lijian@scripps.edu.

S1089-5639(98)00753-1 CCC: $15.00
Published on Web 07/10/1998

applied the current method we developed in this paper to calculate redox potentials of [2Fe2S] clusters in ferredoxin and phthalate dioxygenase reductase (PDR).

Iron−sulfur proteins are important electron-transfer agents. They must be properly ordered by redox potentials within the electron transport chain in order to minimize energy loss and electron trapping along the electron-transfer pathway.[7] Redox potentials are dependent on the type of iron−sulfur clusters in these redox centers. However, the redox potentials are also modulated by the protein and solvent environment. As a result, proteins belonging to the same family may span a wide range of redox potentials and display different functions. An example of this is ferredoxin from *Anabaena* 7120 and phthalate dioxygenase reductase (PDR) from *Pseudomonas cepacia*.[8] The former primarily serves as a terminal electron acceptor from photosystem I and the latter delivers one electron from nicotinamide adenine dinucleotide (NADH) to phthalate dioxygenase. Both use the same type of [2Fe2S] center and anchor it to the protein with the same number of cysteine sulfur ligands. However, the redox potential for PDR is 266 mV more positive than that of ferredoxin. Correll et al.[8a] have investigated the differences in the protein environment that may be responsible for this shift.[8a] Our self-consistent reaction field calculations allow us to identify various factors contributing to the redox potential of this system. Our calculations clearly show how the protein environments determine the redox potential difference in these two proteins, and the role of solvent on redox potentials as well.

II. Computational Methods

1. Overview. The procedure to incorporate solvation effects in density functional calculations has been presented elsewhere.[3] After a brief review, the extension of this methodology to treat the active-site cluster (described quantum mechanically) in the presence of a complete protein plus solvent environment is presented.

In density functional theory (DFT), the electronic energy of a molecule can be expressed as[9]

$$E = \sum_i \langle \psi_i | h(1) | \psi_i \rangle + {}^1/_2 \int \rho(r) G(r,r') \rho(r')\, \mathrm{d}r\mathrm{d}r' + E_{\mathrm{XC}}[\rho_{\mathrm{el}}(r)] \quad (1)$$

where $h(1)$ is the one-electron Hamiltonian (but without the nuclear charge contribution), ψ_i are the Kohn−Sham orbitals, ρ is the total charge density (including the nuclear charges) and E_{xc} represents the exchange-correlation energy due to the electron density $\rho_{\mathrm{el}}(r)$. $G(r,r')$ stands for the Coulomb interaction operator in a dielectric medium, which is the screened Green function of the Poisson equation:

$$\nabla \epsilon(r) \nabla G(r,r') = -4\pi \delta(r,r') \quad (2)$$

where $\epsilon(r)$ is the dielectric constant and $\delta(r,r')$ is a Dirac delta function representing a unit charge at r'. As shown in Figure 1b, we set $\epsilon(r) = \epsilon_{\mathrm{i}}(r) = 1$ for the quantum region, $\epsilon(r) = \epsilon_{\mathrm{p}}(r) = 4$ for the protein region, and $\epsilon(r) = \epsilon_{\mathrm{s}}(r) = 80$ for bulk solvent. In gas phase, $G(r,r')$ is simply the vacuum Coulomb operator:

$$G^\circ(r,r') = \frac{1}{|r - r'|} \quad (3)$$

E in eq 1 becomes the normal gas-phase electronic energy, E_0:

$$E_0 = \sum_i \langle \psi_i | h(1) | \psi_i \rangle + {}^1/_2 \int \frac{\rho(r)\rho(r')}{|r - r'|}\, \mathrm{d}r\, \mathrm{d}r' + E_{\mathrm{XC}}[\rho_{\mathrm{el}}(r)] \quad (4)$$

When the active-site cluster resides in a combined protein and solvent environment, because of the three-region division in Figure 1b, we now are able to treat both protein and solvent environment effects as a correction to the gas-phase Coulomb interaction operator

$$G(r,r') = G^\circ(r,r') + G^*(r,r') \quad (5)$$

Denoting the charge densities of the active-site cluster and protein as $\rho_{\mathrm{c}}(r)$ and $\rho_{\mathrm{p}}(r)$, respectively, the Coulomb interaction energy term in eq 1 becomes

$$\begin{aligned} &{}^1/_2 \int \rho(r) G(r,r') \rho(r')\, \mathrm{d}r\mathrm{d}r' \\ &= {}^1/_2 \int [\rho_{\mathrm{c}}(r) + \rho_{\mathrm{p}}(r)]\, [G^\circ(r,r') + G^*(r,r')] \times \\ &\qquad [\rho_{\mathrm{c}}(r') + \rho_{\mathrm{p}}(r')]\, \mathrm{d}r\, \mathrm{d}r' \\ &= {}^1/_2 \int \rho_{\mathrm{c}}(r) G^\circ(r,r') \rho_{\mathrm{c}}(r')\, \mathrm{d}r\, \mathrm{d}r' + \\ &\qquad {}^1/_2 \int \rho_{\mathrm{c}}(r) G^*(r,r') \rho_{\mathrm{c}}(r')\, \mathrm{d}r\, \mathrm{d}r' + \\ &\qquad \int \rho_{\mathrm{c}}(r) [G^\circ(r,r') + G^*(r,r')]\, \rho_{\mathrm{p}}(r')\, \mathrm{d}r\, \mathrm{d}r' + \\ &\qquad {}^1/_2 \int \rho_{\mathrm{p}}(r) [G^\circ(r,r') + G^*(r,r')]\, \rho_{\mathrm{p}}(r')\, \mathrm{d}r\, \mathrm{d}r' \end{aligned} \quad (6)$$

The first term in the last equation is just the gas-phase Coulomb energy as in eq 4. The last term represents the intraprotein energy. When the protein structure and the corresponding charge distribution (charge set) are fixed, the last term is a constant, and it will be omitted from further analysis. We now define the reaction field potential and protein field potential:

$$\phi^*_{\mathrm{react}}(r) = \int G^*(r,r')\, \rho_{\mathrm{c}}(r')\, \mathrm{d}r' \quad (7a)$$

$$\phi_{\mathrm{prot}}(r) = \int [G^\circ(r,r') + G^*(r,r')]\, \rho_{\mathrm{p}}(r')\, \mathrm{d}r' \quad (7b)$$

the reaction field potential represents the potential induced by the cluster charge distribution acting over the 3 dielectric medium (dominated largely by the solvent polarization potential), and the protein potential is caused by the protein charge distribution acting over the 3 dielectric medium and includes screening by the solvent and protein dielectrics. Then, eq 6 can be rewritten as:

$$\begin{aligned} &{}^1/_2 \int \rho(r) G(r,r') \rho(r')\, \mathrm{d}r\, \mathrm{d}r' \\ &= {}^1/_2 \int \rho_{\mathrm{c}}(r) G^\circ(r,r') \rho_{\mathrm{c}}(r')\, \mathrm{d}r\, \mathrm{d}r' + {}^1/_2 \int \rho_{\mathrm{c}}(r) \phi^*_{\mathrm{react}}(r)\, \mathrm{d}r + \\ &\qquad \int \rho_{\mathrm{c}}(r) \phi_{\mathrm{prot}}(r)\, \mathrm{d}r \end{aligned} \quad (8)$$

Substituting eqs 4 and 8 into eq 1, the total electronic energy of the active-site cluster in the protein plus solvent system becomes

$$\begin{aligned} E &= E_0(\rho_{\mathrm{c}}) + E_{\mathrm{react}} + E_{\mathrm{prot}} \\ &= E_0(\rho_{\mathrm{c}}) + {}^1/_2 \int \rho_{\mathrm{c}}(r) \phi^*_{\mathrm{react}}(r)\, \mathrm{d}r + \int \rho_{\mathrm{c}}(r) \phi_{\mathrm{prot}}(r)\, \mathrm{d}r \end{aligned} \quad (9)$$

Since $\rho_{\mathrm{p}}(r)$ in eq 7b is not computed from the density functional calculation, it has been simplified to a set of point charges

centered on the atoms of the protein. Both $\phi^*_{react}(r)$ and $\phi_{prot}(r)$ can be obtained from solutions of the Poisson or Poisson−Boltzmann equation for the potentials due to the active-site cluster and protein charge distributions in a vacuum and protein plus solvent dielectric environments. On the basis of a variational treatment of eq 9, the reaction field and protein field potentials appear in the Kohn−Sham equation as additional potentials $[\phi^*_{react}(r) + \phi_{prot}(r)]$ to the electronic Hamiltonian of the cluster. The reaction field and protein field interaction energy, $E_{react} + E_{prot}$, in eq 9 is thus the total electrostatic interaction energy of the solvent-screened cluster−protein interaction and the cluster−reaction field interaction.

The calculation of E in eq 9 is then a multistep procedure. We first carry out a density functional calculation for the active-site cluster in the usual way in gas phase. The total charge density from this calculation is denoted as ρ_c^g. From ρ_c^g, a set of atomic point charges is fitted to best represent the molecular electrostatic potential (ESP) generated by the cluster charge density. The Poisson or Poisson−Boltzmann equation due to this set of ESP charges is solved to obtain reaction field $\phi^*_{react}(r)$, and the protein field $\phi_{prot}(r)$ is found from the protein atom charges. The reaction and protein fields are then added to the cluster Hamiltonian and the density functional calculation is repeated. This loop is iterated until the difference of E between the current cycle and previous cycle is smaller than a certain threshold. At this point, the self-consistent reaction field (SCRF) has converged. Now the protein and solvent effects enter E and $\rho_c(r)$ as in eq 9. Note that the final density ρ_c differs from the initial density ρ_c^g. Further aspects of the methodology and relevant parameters are summarized below.

2. Density Functional Calculations. The calculations used the Amsterdam Density Functional package[10] (ADF, version 113) with modifications to include a self-consistent protein field and reaction field. The local density approximation (LDA) for exchange and correlation are based on the parametrization of Vosko, Wilk, and Nusair.[11] Nonlocal corrections (NL), both Becke's gradient correction to exchange[12] and Perdew's correction to correlation,[13] were included in each self-consistent cycle. Convergence of the density functional calculation was achieved once the maximum element of the commutator of the Fock matrix and density matrix was smaller than 0.00003. The numerical integration scheme adopted was the polyhedron method developed by te Velde et al.[14] with the accuracy parameter ACCINT of 4.0. A set of uncontracted triple-ζ Slater-type orbitals (STO) was employed for the $(n + 1)$s, $(n + 1)$p, and nd valence orbitals of the transition metal atoms.[15ab] For the 2s and 2p orbitals of carbon, 3s and 3p orbitals of sulfur, and 1s orbital of hydrogen, the same quality basis set was used, augmented by extra d and p functions, respectively. The inner core shells were treated by the frozen core approximation, through Fe(3s, 3p), S(2s, 2p), and C(1s). A set of auxiliary s, p, d, f, and g STO functions, centered on all nuclei, was introduced to fit the molecular density and to represent Coulomb and exchange potentials accurately.[15c] The basis, core, and fit sets correspond to Basis Set IV of ADF 113. All calculations were done with a spin-unrestricted scheme. As discussed later, some additional single-point calculations were performed with a smaller frozen core Fe(2s, 2p) and with relativistic corrections (summarized below).

Geometry optimization of cluster models was done according to the analytic gradient method implemented at the LDA level by Versluis and Ziegler[16a] and at the NL level by Fan and Ziegler.[16] The optimization used the Newton−Raphson method and the Hessian was updated with the Broyden−Fletcher−Goldfarb−Shanno strategy.[17] Convergence was achieved when changes in coordinate values were less than 0.005 Å and the norm of all gradient vectors was smaller than 0.01. The relativistic correction was estimated by a quasirelativistic calculation that included the mass−velocity and Darwin terms in the first-order Hamiltonian and the induced density changes.[18] In this scheme, the core orbitals were replaced by the relativistic ones, which were obtained by numerical solution of the atomic Dirac equation.

3. Fitting of ESP Charges. A modified version of the CHELPG code of Breneman and Wiberg,[3,4,19] called Chargefit, was used to fit the point charges from the molecular electrostatic potentials (ESP) calculated by the ADF code. The total net charge of the molecule and the three Cartesian dipole moment components from density functional calculations were adopted as constraint conditions for the fit. The fitted points lay on a cubic grid between the van der Waals radius and the outer atomic radius with a grid spacing of 0.2 Å. The outer atomic radius for all atoms used was 5.0 Å and the van der Waals radii for $Fe^{2+/3+}$, S, C, and H were 1.5, 1.8, 1.7, and 1.2 Å, respectively. To minimize the uncertainties in the fitting procedure, the singular value decomposition (SVD) method[4,20] was introduced into the code to obtain a model with stable atomic charges and an accurate molecular dipole moment.

4. Solution of the Poisson or Poisson−Boltzmann Equation, Reaction and Protein Fields. The MEAD (Macroscopic Electrostatics with Atomic Detail) program suite developed by Bashford was employed to calculate the protein potential and the reaction field potential induced by the atomic ESP charges of the active-site molecule. These programs solve the Poisson or Poisson−Boltzmann equation by a numerical finite-difference method.[6] As shown in Figure 1b, the whole system is divided into three regions with dielectric constants of $\epsilon_i = 1$ for the active site (quantum region), $\epsilon_s = 80$ for the solvent region, and $\epsilon_p = 4$ for the protein. Compared to the purely electronic dielectric constant ($\epsilon = 2$), $\epsilon_p = 4$ adopted for the protein allows some mobility of the protein dipoles and accounts for some reorientational relaxation of the protein in an approximate way.[21] Support for the practical value of this model with $\epsilon = 4$ for protein and $\epsilon = 80$ for solvent comes from a variety of methodology and application papers (some quite recent)[6,21] on protein electrostatics and pH titrations beginning with early work of Tanford and Roxby[21h] and including also the protein dielectric experiments of Bone and Pethig[21j] (on dry protein powders, which emphasizes the internal protein dielectric constant, near $\epsilon = 4$). In other recent work, Simonson and Perahia[21d] found that the polarization energy from a "macroscopic model" (closely related to our protein/solvent model) with a protein dielectric of $\epsilon = 4$ combined with an aqueous solvent dielectric $\epsilon = 80$ represented the dipolar polarization energy of the protein well compared to the corresponding energy derived from a microscopic molecular dynamics model combined with linear response theory. Nonetheless, this problem of the proper internal protein dielectric constant to use is admittedly a difficult one, and we address some of the remaining difficulties associated with geometry changes and protein dynamics in the conclusion section.

For the purpose of defining the dielectric boundary, atomic radii of 1.5 Å ($Fe^{2+/3+}$), 1.80 Å (S), 1.7 Å (C), and 1.20 Å (H) were chosen, on the basis of our previous calculations for organic acids and for metal clusters in pure solvent.[3,4,22a] The radii we used for the atoms of the quantum cluster originated from Bondi.[22a] These radii are very similar to those in the PARSE (Parameters for Solvation Energy) paper[23] for hydration

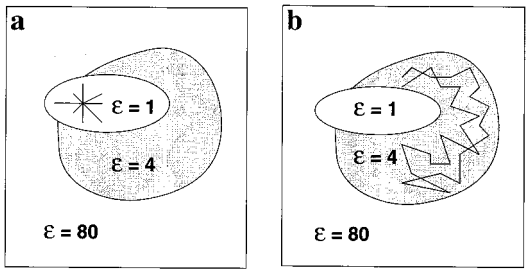


Figure 2. (a) Reaction field induced by the charge distribution in the active-site cluster; (b) protein field due to the charge distribution in protein.

free energies: our radii for sulfur (1.8 Å), carbon (1.7 Å), and H(1.2 Å), vs PARSE radii (1.85 Å for S, 1.7 Å for C, and 1.0 Å for H) (Sitkoff et al.).[23] These PARSE radii were also used in the representation of the protein atoms in the present paper and are widely used for protein electrostatics. The "solute interior" was defined as the region inaccessible to any part of a probe sphere of radius 1.4 Å rolling on the molecular surface of the atomic spheres. The boundary between the interior and exterior so defined is equivalent to Connolly's definition of the molecular surface.[22b] The geometry of the active-site clusters is such that, with the above radii, the solvent probe does not contact the Fe atoms, so that the dielectric boundaries do not depend on the choice of the Fe radius. The resulting Poisson equation was solved by using an over relaxation algorithm on successively finer grids of size 61^3, 61^3, and 81^3 with linear spacings of 1.0, 0.25, and 0.15 Å, respectively. The grids are centered on the active-site cluster, so that the finer grids are focused on the active site. This improves the potential and the subsequent quadrature on the ADF density functional grid.

The initial calculation is done with zero ionic strength. When counterion screening is taken into account, the Poisson equation was replaced by a linearized Poisson−Boltzmann equation:

$$\nabla\epsilon(r)\nabla\phi(r) - \epsilon(r)\kappa^2\phi(r) = -4\pi\phi(r) \qquad (10)$$

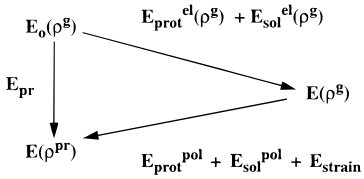
κ is the inverse salt-screening length, which is a function of ionic strength I.

To obtain the protein field and reaction field potentials defined in eq 7, three MEAD calculations are required. For reaction field potential, we first calculate a reference vacuum potential $\phi_{vac}(r)$ by setting all three dielectric constants ϵ_i, ϵ_s, and ϵ_p to 1. Then another calculation is carried out to determine $\phi_{dielectric}(r)$, the potential in the three dielectric environments, according to Figure 2a. In this calculation, only the cluster charge distribution is turned on and the protein charges are temporarily set to zero. The final reaction field potential is calculated as

$$\phi^*_{react}(r) = \phi_{dielectric}(r) - \phi_{vac}(r) \qquad (11)$$

The subtraction cancels out the vacuum Coulomb terms since $G^* = G - G^\circ$ (see eqs 3, 5, and 7). For protein field potential $\phi_{prot}(r)$, no reference vacuum state is needed and the calculation is done following the scheme as in Figure 2b. In this calculation of $\phi_{prot}(r)$, the cluster charge distribution is turned off. However, the electrostatic and structural properties of the protein are taken into account by assigning charges to the atoms in the protein region. The values of $\phi^*_{react}(r)$ and $\phi_{prot}(r)$ can be computed on the grid for numerical integration in the density functional calculations (ADF grid) by interpolation from the MEAD grid. Therefore, the computational implementation of eq 9 is straightforward.

SCHEME 1



5. Energy Decomposition Scheme. Once the self-consistent reaction field (SCRF) calculation converges, a new total charge density for the active-site cluster, ρ_c^{pr}, is obtained (where pr stands for protein plus reaction field), which includes the effects of protein and solvent environments and is different from the total charge density ρ_c^g calculated in a vacuum. The reaction field and protein interaction energy, $E_{react} + E_{prot}$ in eq 9, can be regarded as the main contributors to the free energy change to bring the active-site cluster from vacuum into the protein and solvent surroundings. However, as seen below, the change in total charge density from ρ_c^g to ρ_c^{pr} has an energetic cost on the gas-phase Hamiltonian energy of the cluster, E_{strain}. In Scheme 1, we introduce a hypothetical intermediate state with energy $E(\rho_c^g)$, in which both the protein field and reaction field have been included in the energy calculation but with the gas-phase charge density. (In Scheme 1, E_{sol} is the same as E_{react}.) Accordingly, the total reaction field and protein interaction energy, E_{pr}, can be partitioned into several components:

$$\begin{aligned} E_{pr} &= E_{prot} + E_{react} + E_{strain} \\ &= E_{prot}^{el}(\rho_c^g) + E_{prot}^{pol} + E_{react}^{el}(\rho_c^g) + \\ &\qquad E_{react}^{pol} + E_{strain} \end{aligned} \qquad (12)$$

where

$$E_{prot}^{el}(\rho_c^g) = \int \phi_{prot}(r)\, \rho_c^g(r)\, dr \qquad (13a)$$

$$E_{react}^{el}(\rho_c^g) = {}^1/_2 \int \phi^*_{react}(\rho_c^g, r)\rho_c^g(r)\, dr \qquad (13b)$$

are protein and reaction field energies of interaction with the gas-phase charge distribution in which the electronic polarization of the active-site molecule in the quantum region is neglected. Such molecular polarization contributions are accounted for by the terms

$$E_{prot}^{pol} = \int \phi_{prot}(r)\rho_c^{pr}(r)\, dr - \int \phi_{prot}(r)\rho_c^g(r)\, dr \qquad (14a)$$

$$\begin{aligned} E_{react}^{pol} = {}^1/_2 \int \phi^*_{react}(\rho_c^{pr}, r)\rho_c^{pr}(r)\, dr - \\ {}^1/_2 \int \phi^*_{react}(\rho_c^g, r)\rho_c^g(r)\, dr \end{aligned} \qquad (14b)$$

These two terms represent an energy stabilization. The reaction field potential is a functional of the cluster charge distribution, whether this arises from the initial cluster charge distribution $\rho_c^g(r)$ or from the final SCRF charge distribution $\rho_c^{pr}(r)$. By contrast, the protein potential is entirely determined by the protein charge distribution and the three dielectric regions, considered to be fixed. The energy cost due to charge redistribution or electronic deformation (electronic strain) in the active site molecule is

$$E_{strain} = E_0(\rho_c^{pr}) - E_0(\rho_c^g) \qquad (15)$$

This decomposition scheme is very instructive for the analysis

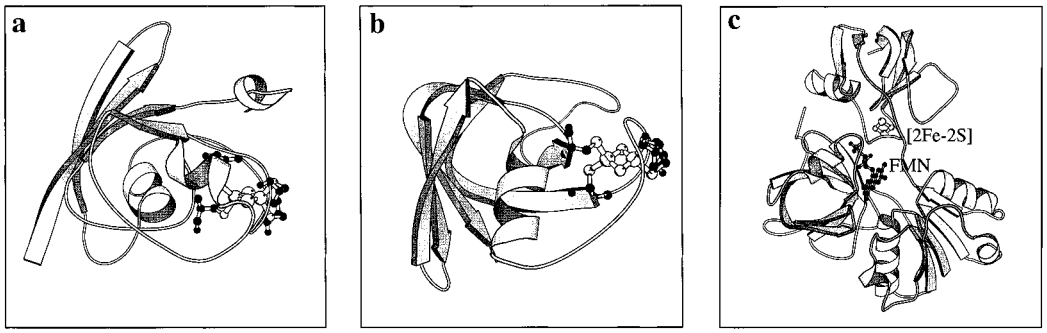


Figure 3. Protein structures of (a) ferredoxin from *Anabaena* 7120. (b) [2Fe2S] domain of phthalate dioxygenase reductase (PDR) from *Pseudomonas cepacia*, and (c) full protein (3 domains) of PDR with FMN and [2Fe2S] core represented by ball-and-sticks. This figure and Figures 4, 6, and 7 were prepared with MOLSCRIPT: Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946.

of various factors contributing to the protein interaction and reaction field energies.

III. Protein and Active-Site Model Structures

1. Protein Structures. The X-ray crystallographic structures of [2Fe2S] ferredoxin (Fd) from *Anabaena* 7120[24] and phthalate dioxygenase reductase (PDR) from *Pseudomonas cepacia*[25] have been reported. For *Anabaena* [2Fe2S] ferredoxin, only the structure of oxidized form was determined to a resolution of 2.5 Å. The protein contains only one domain with 98 residues, in which the side chains of cysteines 41, 46, 49, and 79 coordinate to [2Fe2S] core via terminal cysteine sulfur atoms (Figure 3a). PDR is a multidomain protein in which [2Fe2S] cluster and other electron transfer centers, flavin mononucleotide (FMN) and nicotinamide adenine dinucleotide (NADH) are bound to distinct domains. The three domains are brought together near a central cleft in the molecule, with only 4.9 Å separating the flavin 8-methyl and cysteine sulfur ligated to iron. The structure of the oxidized form was determined at 2.0 Å resolution and the reduced form at 2.7 Å resolution. We used the oxidized 2.0 Å resolution structure for PDR, determined in the absence of NADH or NAD^+ for all of our calculations. This also corresponds to the conditions for the equilibrium redox potential titrations with dithionite.[8,25,36] In the [2Fe2S] domain, the sulfur atoms from side chains of cysteines 272, 277, 280, and 308 ligate to irons (Figure 3b,c). We use the oxidized structures in our calculation, which implies an assumption that the backbones and the side chains of the proteins are kept rigid during the one-electron reduction of [2Fe2S] system. However, some protein side chain mobility is taken into account implicitly by using a protein dielectric $\epsilon = 4$ which is larger than the purely electronic protein dielectric $\epsilon \approx 2$.

The coordinates of the two proteins were downloaded from the Protein Data Bank (PDB) of Brookhaven National Laboratory (PDB ID codes: ferredoxin from *Anabaena* 7120, 1FXA; PDR, 2PIA). Hydrogen atoms were added at pH 7 and the positions were optimized with the Discover module in Insight II (Biosym/MSI), with all heavy atoms fixed. The PARSE charge and radii set for protein atoms[23] was employed to determine the dielectric interface and charge distribution of the proteins. Since the flavin mononucleotide (FMN) cofactor in PDR plays an important role in the biological reduction−oxidation process and is close to the [2Fe2S] cluster (Figure 3c),[25] we calculated a set of ESP charges to replace the PARSE charge set. This set of ESP charges is based on density functional calculations of two fragments of FMN in the oxidized form: a cyclic portion and a sugar phosphate chain. In the redox titration of the $[2Fe2S]^{2+/1+}$ cluster from PDR, the FMN is in a nearly equal mixture of coupled oxidation/protonation states, specifically $(FMN)_{ox}$ and $(FMNH^{\bullet})$, based on the nearly equal redox potentials observed for this couple and for the $[2Fe2S]^{2+,1+}$ cluster.[25] We then used the $(FMN)_{ox}$ charge state for this cofactor, while an average of $(FMN)_{ox}$ and $(FMNH^{\bullet})$ would be more appropriate. However, this reaction is electroneutral and so should not exert a significant effect on the [2Fe2S] redox energy.

2. $[Fe_2S_2(SCH_3)_4]$ Active-Site Model. The active-site model, which was treated rigorously by density functional methods, is $[Fe_2S_2(SCH_3)_4]$ where the ligated cysteines are simplified as methyl thiolates. The resolution of X-ray structures for the proteins is not high enough to distinguish the fine structural details of the [2Fe2S] core. Therefore, we used active-site models in which the [2Fe2S] core geometry is based on the structures of synthetic analogues of $[Fe_2S_2(SR)_4]$.[26] Following a model presented earlier,[27] the Fe−Fe and Fe−S* (S* is the sulfur in the [2Fe2S] core) distances were set to 2.69 and 2.21 Å, respectively, in the oxidized form. In reduced form, the Fe−Fe distance was increased to 2.73 Å and the Fe−S* distance was expanded by 0.07 Å at the reduced Fe site. All Fe−S (S in SCH_3) bonds were set at 2.31 Å. The [2Fe2S] cores built in such a way were then fit back into the $(SCH_3)_4$ frames, in which the S and C atoms are placed in the positions they occupy in the proteins and the four α-hydrogens H_α are aligned with $C_\beta-H_\alpha$ collinear to the four α-carbons ($C_\beta-C_\alpha$) of the ligating cysteines. The reduced Fe site, as established experimentally from the spin distribution by EPR for PDR and NMR for *Anabaena* ferredoxin,[28] is closer to the surface (about 5 Å) of the proteins. In the calculations, the reduced site was defined following this observation. The structures of $[Fe_2S_2(SCH_3)_4]$ are depicted in Figure 4, which clearly shows the different orientation of SCH_3 groups in two models. These model geometries were used for the subsequent SCRF calculations.

The $[Fe_2S_2(SCH_3)_4]$ active site model can be regarded as a small part of the protein terminating with one H_α in place of C_α for the four ligated cysteines. H_α is part of the quantum mechanical cluster and C_α represented electrostatically belongs to the classical region. In the self-consistent reaction field calculation, the cluster and protein must be carefully coupled at the junction atoms H_α and C_α. In the charge fitting step, H_α is included in the construction of the fitting grid surrounding the whole cluster, but it is excluded from the charge fitting procedure by setting the H_α charge to 0. When defining the dielectric boundary, C_α, rather than H_α, is used. The charge of C_α is also set to zero (as in the PARSE charge set). This "dual boundary" approach, as shown in Figure 5, guarantees charge conservation of the active-site cluster and avoids the nonphysical charge interaction between H_α and C_α.

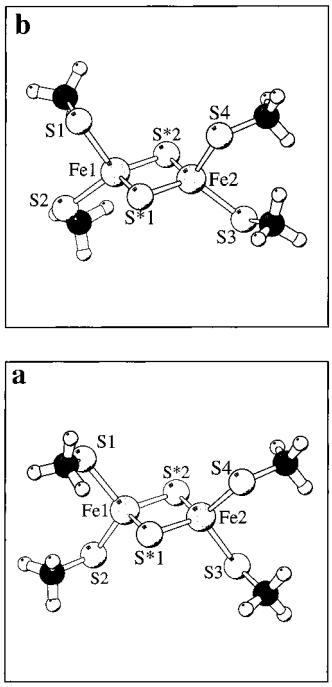


Figure 4. Structures of $[Fe_2S_2(SH_3)_4]$ clusters in (a) ferredoxin from *Anabaena* 7120 and (b) phthalate dioxygenase reductase (PDR) from *Pseudomonas cepacia*.

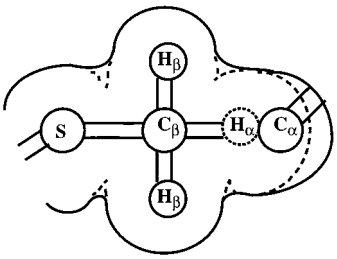


Figure 5. "Dual boundary" for coupling $[Fe_2S_2(SH_3)_4]$ cluster and protein environments. The dashed line represents the boundary for charge fitting and the solid line is the dielectric boundary.

IV. Results and Discussion

1. Gas-Phase Ionization Potentials. We first calculate the gas-phase ionization potential, IP_{red}, of $[Fe_2S_2(SCH_3)_4]^{2-}$, which is closely related to the redox potential in gas phase. IP_{red} is defined as the gas-phase ground-state energy difference between $[Fe_2S_2(SCH_3)_4]^{2-}$ and $[Fe_2S_2(SCH_3)_4]^{3-}$ (the oxidized minus the reduced state energy). Calculation of the ground-state energy of a dinuclear metal cluster by density functional methods is complicated by the spin coupling of the unpaired electrons in the metal centers.[27,29] Such a spin coupling can be determined by a Heisenberg Hamiltonian

$$H_{spin} = J\, S_1 \cdot S_2 \tag{16}$$

where S_1 and S_2 refer to spin quantum numbers on two iron centers and J is the spin coupling parameter. For the oxidized form $[Fe_2S_2(SCH_3)_4]^{2-}$, the corresponding quantum numbers S_1 and S_2 are both $^5/_2$. In contrast, the reduced form $[Fe_2S_2(SCH_3)_4]^{3-}$ is a mixed valence dimer with $S_1 = {}^5/_2$ and $S_2 = 2$. Furthermore, the extra electron can be delocalized (at least potentially). This effect is most prominent for large total spin S. In this case, another term has to be added to the spin Hamiltonian:

$$H_{spin} = J_{red}\, S_1 \cdot S_2 \pm B\,(S + {}^1/_2) \tag{17}$$

TABLE 1: Spin Hamiltonian Parameters and Ground-State Energies of $[Fe_2S_2(SCR)_4]^{2-/3-}$

oxidized form $[Fe_2S_2(SCR)_4]^{2-}$	reduced form $[Fe_2S_2(SCR)_4]^{3-}$
$S_1 = S(Fe^{3+}) = {}^5/_2$	$S_1 = S(Fe^{3+}) = {}^5/_2$
$S_2 = S(Fe^{3+}) = {}^5/_2$	$S_2 = S(Fe^{2+}) = 2$
	$B = {}^1/_{10}[e(HS)_u - e(HS)_g]^a$
$J_{ox} = {}^2/_{25}[E_0(HS) - E_0(BS)]$	$J_{red} = {}^1/_{10}[E_0(HS)_g - E_0(BS) + 5B]$
$E_0(GS) = E_0(BS) - {}^5/_2 J_{ox}$	$E_o(GS) = E_o(BS) - 2J_{red}$

[a] $e(HS)_u$ and $e(HS)_g$ are orbital energies corresponding to u and g components of metal d orbitals split by resonance delocalization.

TABLE 2: Calculated Spin Hamiltonian Parameters, Ground-State Energies, and Ionization Potentials

		B (cm^{-1})	J (cm^{-1})	$E_0(GS)^a$ (eV)	IP_{red}^{gas} (eV)
ferredoxin[b]	$[Fe_2S_2(SCH_3)_4]^{2-}$		868	−121.371	
	$[Fe_2S_2(SCH_3)_4]^{3-}$	912	884	−116.357	−5.014
PDR[b]	$[Fe_2S_2(SCH_3)_4]^{2-}$		841	−121.244	
	$[Fe_2S_2(SCH_3)_4]^{3-}$	790	828	−116.265	−4.979

[a] The total energy calculated by ADF package is the bonding energy with respect to spherical restricted atoms. [b] Parameters calculated for model 1 geometries.

where S is the total spin quantum number, $|S_1 - S_2| \le S \le S_1 + S_2$ and B is the resonance delocalization parameter. For small S, the resonance energy B term is usually quenched by vibronic, solvation, and protein environment effects.[29,30] Even with complete quenching (trapped valence), the B-term energy must be evaluated to calculate J_{red}. An efficient scheme has been developed to estimate B, J, and energies of various spin states based on density functional calculations.[4,27,29] Following this scheme, we first carried out a calculation on the highest spin (maximum S) state with energy $E_0(HS)$. For the reduced case, $E_0(HS)$ was calculated for the lowest such state with a bonding delocalized electron in a *gerade*-type orbital. In addition, a so-called broken-symmetry state was introduced. This state corresponds to a spin-unrestricted determinant in which the spin-up electrons are predominantly on onehalf of the dimer while the spin-down electrons are on the other half. The energy of the broken-symmetry state is calculated to be $E_0(BS)$. Table 1 shows all equations needed to calculate B, J, and ground-state energy $E_0(GS)$ based on $E_0(HS)$ and $E_0(BS)$. These equations are consistent with the Hamiltonians in eqs 16 and 17. Further discussion and derivations of these equations can be found elsewhere.[4,27,29]

Table 2 summarizes the calculated parameters of B, J, and energies $E_0(GS)$ for $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters in ferredoxin and PDR. The calculated J value for the oxidized ferredoxin and PDR and J and B values for the reduced complexes are larger than those in previous calculations,[4a] where a symmetrized $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ (C_{2v} symmetry) was used ($J_{ox} = 763\ cm^{-1}$, $J_{red} = 514\ cm^{-1}$, $B = 394\ cm^{-1}$). Experimentally, J values are determined to be 366 and 298 cm^{-1} for oxidized spinach ferredoxin[31] and a synthetic analogue,[32] respectively. For the *Spirulina maxima* 2Fe2S protein, the experimental values are $J_{ox} = 364\ cm^{-1}$, and $J_{red} = 196\ cm^{-1}$.[33,34] Overall, the calculations predict correct antiferromagnetic behavior but tend to overestimate J values.

The calculated values of gas-phase ionization potential, IP_{red}^{gas}, are included in Table 2 as well. It is worthwhile noting that the absolute IP_{red}^{gas} value of the $[Fe_2S_2(SCH_3)_4]^{3-}$ cluster in PDR is only about 0.035 eV more positive than that in ferredoxin, which implies that the electronic structures of the $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters in the two proteins are almost same. This is not surprising since the $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters were constructed to contain the same [2Fe2S] core geometries. The

TABLE 3: Geometrical Parameters of [2Fe2S] Core[a]

		model 1 geometries[b,c]		gas-phase constrained optimization in PDR		gas-phase constrained optimization in ferredoxin	
			$[Fe_2S_2(SCH_3)_2]^{2-}$ (oxd)				
Fe_1-Fe_2		2.688		2.752		2.742	
$Fe_1-S^*_1$	$Fe_1-S^*_2$	2.208	2.208	2.265	2.269	2.246	2.253
$Fe_2-S^*_1$	$Fe_2-S^*_2$	2.208	2.208	2.258	2.228	2.257	2.251
Fe_1-S_1	Fe_1-S_2	2.308	2.308	2.375	2.388	2.401	2.408
Fe_2-S_3	Fe_2-S_4	2.308	2.308	2.391	2.388	2.381	2.394
			$[Fe_2S_2(SCH_3)_2]^{3-}$ (red)				
Fe_1-Fe_2		2.728		2.860		2.878	
$Fe_1-S^*_1$	$Fe_1-S^*_2$	2.277	2.277	2.356	2.372	2.357	2.353
$Fe_2-S^*_1$	$Fe_2-S^*_2$	2.209	2.209	2.268	2.237	2.258	2.267
Fe_1-S_1	Fe_1-S_2	2.310	2.310	2.446	2.488	2.442	2.442
Fe_2-S_3	Fe_2-S_4	2.306	2.306	2.496	2.512	2.490	2.500
$\Delta E = E(\text{oxd})_{BS} - E(\text{red})_{BS}$[d]		−4.91 (PDR)	−4.96 (Fd)	−4.78		−4.81	

[a] Bond lengths are given in angstrom. [b] Same core geometries for PDR and ferredoxin. These model geometries were used in the SCRF calculations. [c] For all the reduced clusters, Fe_1 is the reduced Fe site and Fe_2 is the oxidized Fe site. [d] Energies are given in electron volts.

only difference between the two clusters, as can be seen in Figure 4, is in the conformation of four SR^- groups, which is determined by the orientation of the side chains of the four cysteines coordinated to the [2Fe2S] core. The energetic difference associated with such changes in torsion angle of side chains should not be dramatic.

To obtain a more accurate estimate of gas-phase ionization potential for $[Fe_2S_2(SCH_3)_4]^{3-}$, some correction terms can be added. First, since the model [2Fe2S] geometries used in the calculations were constructed according to the structures of synthetic analogues (but with altered side-chain orientations as in the proteins), the real [2Fe2S] structures in proteins could differ due to this orientation of ligated cysteine side chains, which imposes a constraint on the $Fe_2S_2(SR)_4$ active site. This effect should be larger for the reduced form in which the electronic relaxation upon reduction is coupled with structural changes. (Further, the reduced model structure is an extrapolation from known $[Fe_2S_2(SR)_4]^{2-}$ and $[Fe(SR)_4]^{1-,2-}$ synthetic structures, since no synthetic $[Fe_2S_2(SR)_4]^{3-}$ reduced X-ray structure is available.) To evaluate the correction from these geometrical effects, we carried out a set of gas-phase constrained geometry optimizations on $Fe_2S_2(SR)_4$, $R = CH_3$ in *Anabaena* ferredoxin and PDR, for both the reduced and oxidized forms in the broken-symmetry state. In each optimization, $Fe_2S_2(SR)_4$ was allowed to change size and reorient, but all the dihedral angles associated with $FeSCH_3$, S*FeSC, and FeS*FeS were fixed. The optimized geometrical parameters are summarized in Table 3, together with the parameters from the models (model 1 geometries) used in the SCRF calculations for comparison. It follows from Table 3 that although the optimized geometries differ only moderately from those used in SCRF calculations, the broken-symmetry state energy difference between oxidized and reduced states is indeed shifted to more positive values by about 0.13−0.15 eV, due to stabilization of the reduced form. This shift will be observed in the gas-phase ionization potentials as well and we added this correction, ΔE_{geom}, to IP_{red}^{gas}, as shown in Table 4. We expect that this term is still underestimated since we used the same protein structure from the oxidized form to build constraint geometry of $-SCH_3$ moieties for both $[Fe_2S_2(SCH_3)_4]^{2-}$ and $[Fe_2S_2(SCH_3)_4]^{3-}$ clusters, which is unfavorable to the reduced form. More improvements can be achieved if we are able to use different geometries of $-SCH_3$ moieties for reduced and oxidized forms. This requires either an accurate crystallographic structure for the reduced form of the proteins or a computational optimization scheme that can include the whole or part of the protein and solvent environment in the molecular force field, or alternatively an extended quantum cluster including amide−S hydrogen bonding.

TABLE 4: Several Correction Terms to Gas-Phase Ionization Potentials[a]

	IP_{red}^{gas} (1)	ΔE_{R+C}	ΔE_{geom}	IP_{red}(final)
ferredoxin (Fd)	−5.014	−0.126	+0.150	−4.990
PDR	−4.979	−0.118	+0.130	−4.967

[a] Energies are given in electron volts. IP_{red} (final) = IP_{red}^{gas} (1) + ΔE_{R+C} + ΔE_{geom}. IP_{red}^{gas} (1) refers to calculations using the model 1 geometries cited in Tables 2 and 3. Corrections for relativistic terms plus inclusion of Fe(3s,3p) in the valence space give ΔE_{R+C}, while independent calculations using constrained optimized geometries give ΔE_{geom}.

Although relativity is not critical in first-row transition metal complexes, it can still affect ionization potentials to some extent. For highly charged complexes such as $[Fe_2S_2(SCH_3)_4]^{2-/3-}$, the frozen core approximation, especially with a large frozen core on Fe, will introduce errors in energy calculations. To estimate these errors, we carried out single-point calculations using quasirelativistic DFT[19] for $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters (model 1 geometries) including iron (3s, 3p) in the valence space. These effects tend to stabilize the oxidized form and in total shift down the gas-phase ionization potential by about 0.1 eV; see term ΔE_{R+C} in Table 4.

Another potential source of error in the ionization potential involves the too large J values resulting from the calculations compared to the experimental values (as discussed above). In earlier work,[4a] we presented a detailed argument that the J contribution to the redox potential can be determined by using the spin barycenter states (spin degeneracy weighted average) as reference states. The associated redox contribution is $\Delta JT = 7J_{red} - 9J_{ox}$ (where ΔJT = the J term difference), and the error estimate based on the experimental versus calculated J values is simply $(\Delta JT)_{error} = (\Delta JT)_{exp} - (\Delta JT)_{calc}$ (experimental J values from *Spiralina maxima* 2Fe2S protein were used). The resulting $(\Delta JT)_{error}$ estimate varies from +0.075 to −0.016 eV. On average, a small positive shift for ionization potentials is expected of about +0.03 eV. This error estimate also shows that the redox potential is fairly insensitive to errors in calculated versus experimental J values. In view of the semiempirical character of the $(\Delta JT)_{error}$ term, we have omitted this term from IP_{red}(final) in Table 4.

2. Protein Interaction and Solvation Energies. After an SCRF calculation is converged, we can obtain the protein interaction and reaction field energy of the model cluster in the

TABLE 5: Calculated Protein Interaction and Reaction Field Energies of $[Fe_2S_2(SCH_3)]^{2-/3-}$ Clusters in Proteins[a]

		$I = 0.00$ M						$I = 0.05$ M					
		E_{prot}^{el}	E_{prot}^{pol}	E_{react}^{el}	E_{react}^{pol}	E_{strain}	E_{pr}	E_{prot}^{el}	E_{prot}^{pol}	E_{react}^{el}	E_{react}^{pol}	E_{strain}	E_{pr}
ferredoxin	red	−59.1	−6.9	−308.3	−20.8	12.9	−382.2	−62.0	−6.9	−308.8	−20.7	12.9	−385.5
	oxd	−41.4	−4.6	−139.3	−10.2	6.0	−189.5	−43.3	−4.6	−139.5	−10.0	6.0	−191.5
PDR	red	−87.2	−8.6	−303.0	−17.7	11.0	−405.6	−88.7	−8.6	−303.4	−17.6	11.0	−407.4
	oxd	−62.5	−6.2	−136.1	−8.8	5.5	−208.2	−63.5	−6.2	−136.4	−8.8	5.5	−209.4

[a] Energies are given in kilocalories per mole.

TABLE 6: SCRF Calculated Redox Potentials of $[Fe_2S_2(SCH_3)_4]^{2-}$ Cluster in Proteins[a]

		$I = 0.00$ M		$I = 0.05$ M		
	IP_{red}	ΔE_{pr}	ΔE_{redox}°	ΔE_{pr}	ΔE_{redox}°	exp
ferredoxin (Fd)	−4.990	8.359	−1.061	8.413	−1.007	−0.440[b]
PDR	−4.967	8.560	−0.837	8.585	−0.812	−0.174[c]
Δ(Fd − PDR)	−0.023	−0.201	−0.225	−0.172	−0.195	−0.266

[a] Energies are given in electron volts. The redox potentials are calculated according to eq 18. [b] Reference 35. [c] Reference 36.

protein plus solvent environment. These values are listed in Table 5 for *Anabaena* ferredoxin and PDR, decomposed into a sum of terms as in eq 12. Two ionic strength conditions were used in these calculations. An ionic strength of 0.05 M corresponds to a 100 mM HEPES buffer, pH 8, which was used for a variety of kinetic measurements on PDR.[35] $I = 0.05$ M is then compared to ionic strength $I = 0$ in our calculations. The ionic strength seems only to shift the protein interaction and reaction field energies by a few kilocalories per mole for each oxidation state. The net effect on the total redox potentials are even less, amounting to 0.025−0.05 eV (0.58−1.15 kcal/mol) as shown in Table 6. The experimental redox titrations on native *Anabaena* ferredoxin were conducted at low ionic strength, $I = 0.012$ M,[35] while we have not found the ionic strength conditions for the PDR redox potentials cited in Gassner et al.[36] Our calculations indicate low sensitivity to ionic strength conditions.

There is no direct experimental data which can be used to compare and calibrate the calculated energy terms in Table 5. However, a few observations can be made. The absolute values of both the calculated E_{prot} and E_{react} are larger for reduced forms than for oxidized forms. This is understandable because the reduced cluster bears one more negative charge than the oxidized one. For the same reason, the reduced cluster should be softer and more easily polarized. This is reflected in E_{strain}, which is larger for the reduced form than for the oxidized, although E_{strain} itself is a small component. E_{strain} is opposite in sign but follows the trend in magnitude of $(E_{prot}^{pol} + E_{react}^{pol})$. In general, E_{prot} is significantly smaller than E_{react}. However, E_{prot} is subject to a considerable change in the two different protein environments.

3. Redox Potentials. From the calculated gas-phase ionization potential IP_{red} in Table 4 and E_{pr} in Table 5, the redox potentials can be computed as

$$\Delta E^{\circ}_{redox} = IP_{red} + \Delta E_{pr} + \Delta SHE \qquad (18)$$

where ΔSHE represents the standard hydrogen electrode potential of −4.43 eV.[37] E_{pr} is the difference of E_{pr} given by the oxidized minus reduced state energies. The calculated redox potential values ΔE°_{redox}, together with the experimental data, are tabulated in Table 6.

The redox potentials are calculated under two ionic strength condition, $I = 0$ and $I = 0.05$ M. Since the experimental redox potentials were determined in the presence of certain salt concentrations but some important details have not been published for PDR, we shall in the following use the calculated results at $I = 0.05$ M as a principal reference concentration to compare with experiment ($I = 0$ results are quite similar as shown). A direct comparison of the absolute values between calculated and experimental redox potentials shows a substantial deviation of about 0.5−0.6 V. However, we note that our calculations predict the correct order of redox potentials for PDR and ferredoxin and that the experimental redox potential difference ΔE°_{red}(Fd − PDR) of −0.266 V is fairly well reproduced. The calculated difference, −0.195 V, is only 0.071 V more positive than the experimental value.

Experimentally, the redox potentials for some synthetic analogue clusters have been reported. These values are consistently more negative than the ones measured in proteins, for example, −1.25 V for $[Fe_2S_2(S_2\text{-o-xylyl})_2]^{2-}$ and −0.85 V for $[Fe_2S_2(SPh)_4]^{2-}$.[38] It is interesting to note that our calculated ΔE°_{redox} values are closer to these experimental values for synthetic analogues. In our previous work[4a] on 2Fe2S synthetic models in high dielectric solvent ($\epsilon = 37$) using the simpler MEAD method, the calculated redox potential was −1.38 eV, compared to −1.25 to −0.85 eV for the synthetic systems. The error is somewhat smaller, even without geometry optimization of the clusters. The experimental 2Fe2S protein redox potentials, in the range of −0.18 to −0.44 eV, are both considerably more positive than those in synthetic systems and similarly more positive than the calculated values. The errors of 0.5−0.6 eV in our calculated redox potentials may have a number of causes. We will emphasize the two effects we consider to be most physically significant. First, the quantum cluster $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ has a large net charge, particularly in the reduced form, and forms charged NH−S and NH−S* hydrogen bonds with the surrounding protein residues. These energies are evaluated by using the electrostatic interaction of the quantum cluster with the protein, but a proper description of S → HN or S* → HN charge transfer would include a quantum treatment of the amide group. We think that the strength of these terms is underestimated when the protein is described electrostatically. Further, the protein oxidized geometry was used throughout the oxidized and reduced cluster-protein calculations (since the oxidized structures are known most accurately), but NH−S and NH−S* bonds should shorten and strengthen on reduction of the cluster. Calculations including the geometry change are possible, but the internal protein geometry and energy change on reduction will also contribute to the energy difference. This is difficult to calculate accurately but is probably feasible with high-quality force fields and accurate cluster charge distributions.

Accurate calculations of redox potentials of molecules that can be compared quantitatively with experimental measurements are challenging. Some successful calculations have been reported by our group[4] and other groups,[39,40] with typical errors of 100−400 mV and maximum errors of about 1000 mV. Recently, Zhang and Friesner[40] carried out a ab initio SCRF calculation on bacteriochlorophyll (Bchl) and bacteriopheophytin (BPh) molecules at HF level. The calculated redox potentials

TABLE 7: Energy Decomposition of ΔE_{pr} from SCRF Calculations[a]

	ΔE_{prot}^{el}	ΔE_{prot}^{pol}	ΔE_{react}^{el}	ΔE_{react}^{pol}	ΔE_{strain}	ΔE_{pr}
			$I = 0.00$ M			
ferredoxin (Fd)	0.768	0.100	7.329	0.460	−0.299	8.359
PDR	1.072	0.105	7.235	0.388	−0.240	8.560
Δ(Fd − PDR)	−0.304	−0.005	0.094	0.072	−0.059	−0.201
			$I = 0.05$ M			
ferredoxin (Fd)	0.811	0.100	7.342	0.464	−0.299	8.413
PDR	1.093	0.105	7.246	0.380	−0.239	8.585
Δ(Fd − PDR)	−0.282	−0.005	0.096	0.084	−0.060	−0.172

[a] Energies are given in electron volts.

TABLE 8: ESP Charges[a] and Dipole Moments (μ) of $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ Clusters[b]

		in gas phase					in (protein + solvent)				
		Fe	S*	SCH_3		μ	Fe	S*	SCH_3		μ
ferredoxin	red	+0.812	−0.848	−0.826	−0.769	2.576	+0.816	−0.860	−0.824	−0.774	4.757
		+0.846	−0.812	−0.717	−0.685		+0.838	−0.841	−0.687	−0.669	
	oxd	+0.694	−0.607	−0.598	−0.536	1.039	+0.699	−0.613	−0.599	−0.529	2.470
		+0.692	−0.539	−0.555	−0.550		+0.687	−0.556	−0.559	−0.531	
PDR	red	+0.805	−0.831	−0.729	−0.723	3.559	+0.804	−0.856	−0.718	−0.697	4.846
		+0.808	−0.863	−0.838	−0.629		+0.803	−0.887	−0.842	−0.606	
	oxd	+0.669	−0.572	−0.516	−0.571	1.807	+0.670	−0.589	−0.504	−0.556	2.540
		+0.681	−0.614	−0.592	−0.485		+0.680	−0.629	−0.596	−0.476	

[a] The ESP charges in the first line for each entry are for Fe(1), S*(1), S(1)CH_3 and S(2)CH_3; in the second line, for Fe(2), S*(2), S(3)CH_3 and S(4)CH_3. Fe(1), S(1)CH_3 and S(2)CH_3 correspond to the reduced site for the reduced complex. All systems correspond to model 1 geometries.[b] At ionic strength $I = 0.05$ M. Dipole moments are given in debyes.

deviated from the experimental values of about 1 V. However, this calculation also predicted a correct trend of redox potentials for Bchl and BPh.

4. Effects of Protein Field and Reaction Field. As can be seen from Tables 5−7, the protein and solvent environments make major contributions to the redox potentials. In contrast to the negative IP_{red} values of $[Fe_2S_2(SCH_3)_4]^{3-}$, which reflect the difficulty of adding an extra electron to $[Fe_2S_2(SCH_3)_4]^{2-}$, the large positive ΔE_{pr} values imply that the protein interaction and solvation preferentially stabilize the reduced form of $[Fe_2S_2(SCH_3)_4]^{3-}$ over the oxidized form. As a whole, the ionization potentials in the protein plus solvent environment, $IP_{ps} = IP_{red} + \Delta E_{pr}$, are positive, which is the physical requirement for the stability of the 3^- cluster in the complete environment. From eq 18, we also find $IP_{ps} = \Delta E^\circ_{redox} - \Delta SHE = \Delta E^\circ_{redox} + 4.43 \geq 0$.

The protein field and reaction field also play an important role in distinguishing the redox potentials of $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters in different proteins. According to eqs 12−15, we can decompose ΔE_{pr} into several terms, as shown in Table 7. It can be concluded from Table 7 that the dominant component of ΔE_{pr} is the reaction field contribution ΔE_{react}^{el}. This term along with the smaller ΔE_{react}^{pol} shifts the net ionization potential IP_{ps} (and ΔE°_{redox}) to a more positive value but is not responsible for the difference seen between proteins. By contrast, the protein contribution ΔE_{prot}^{el} is a small portion of ΔE_{pr}. However, it is this smaller term that is significantly different in the two proteins. For instance, at 0.05 M ionic strength, the difference between the reaction field contributions ΔE_{react}^{el} for ferredoxin and PDR is only 0.096 eV, with an opposite sign to the ΔE_{pr}. The ΔE_{prot}^{el} difference is, however, −0.282 eV, even larger than ΔE_{pr} of −0.172 eV. This indicates that the electrostatic interaction between protein field and $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters preferentially stabilizes the 3^- cluster form over the 2^- from more strongly in PDR than in *Anabaena* ferredoxin. These ΔE_{prot}^{el} data clearly indicate in a quantitative way that the protein environments can tune the redox potentials via electrostatic interactions between the protein environment and active-site clusters.

The other terms, ΔE_{react}^{pol}, ΔE_{prot}^{pol}, and ΔE_{strain}, make smaller contributions to ΔE_{pr} and are not as sensitive to the change of protein environment. This is not surprising because these terms arise from the polarization interaction, which is relatively weaker compared to the direct first-order electrostatic interaction between the charged cluster and the protein and solvent environment.

Table 8 shows the ESP charges of the $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters in gas phase and in the protein plus solvent environment. It is interesting to note that the charges change to a minor extent when the cluster is brought from gas phase into the protein and solvent. This observation is understandable since the electrostatic interaction with protein and solvent should not perturb the electronic structure and electron distribution of the cluster dramatically. However, this altered charge distribution affects the dipole moments of the clusters. For both the reduced and oxidized forms of the clusters, when the protein and reaction fields are added, the dipole moments increase by about 0.8−1.5 D. The 6−12 kcal/mol in E_{strain} energy in Table 5 can be considered as the associated energy cost of the increased dipole and higher moments in the protein/solvent environment. An unexpected feature of the ESP charges is that the Fe atoms are more positive in the reduced form [for both reduced and oxidized Fe sites Fe(1) and Fe(2)] than in the oxidized form. The added electron in the reduction process is mainly distributed to the bridging sulfur atoms and terminal thiolate groups. All S, S* atoms have increased charges. This has been observed in previous calculations[4a] and is associated with large relaxation effects upon reduction.

5. MEAD Calculations. A full SCRF calculation with several iterative steps is computationally intensive although the convergence is usually well behaved. Since the ESP charges, as shown in Table 8, change to a small degree after SCRF convergence, it can be expected that calculations based on only the gas-phase ESP charges would give a reasonable estimate

TABLE 9: Comparison of Protein Interaction and Reaction Field Energies Calculated by SCRF and MEAD Methods[a]

		SCRF		MEAD	SCRF		MEAD	SCRF	SCRF	MEAD
		E_{prot}	E_{prot}^{el}	E_{prot}	E_{react}	E_{react}^{el}	E_{react}	E_{strain}	E_{pr}	E_{pr}
					$I = 0.00$ M					
ferredoxin	red	−66.0	−59.1	−61.0	−329.1	−308.3	−324.4	12.9	−382.2	−385.3
	oxd	−46.0	−41.4	−42.6	−149.5	−139.3	−148.6	6.0	−189.5	−191.1
PDR	red	−95.8	−87.2	−84.9	−320.7	−303.0	−318.7	11.0	−405.6	−403.7
	oxd	−68.7	−62.5	−60.0	−144.9	−136.2	−145.7	5.5	−208.2	−205.7
					I = 0.05 M					
ferredoxin	red	−68.9	−62.0	−63.3	−329.5	−308.8	−324.7	12.9	−385.5	−388.0
	oxd	−47.9	−43.3	−44.1	−149.5	−139.5	−148.7	6.0	−191.5	−192.8
PDR	red	−97.3	−88.7	−87.3	−321.0	−303.5	−318.9	11.0	−407.4	−406.2
	oxd	−69.7	−63.5	−61.6	−145.2	−136.4	−145.3	5.5	−209.4	−206.9

[a] Energies are given in kilocalories per mole.

TABLE 10: MEAD Calculated Redox Potentials of $[Fe_2S_2(SCH_3)_4]^{2-}$ Cluster in Proteins[a]

		$I = 0.00$ M		$I = 0.05$ M		
	IP_{red}	ΔE_{pr}	$\Delta E_{redox}°$	ΔE_{pr}	$\Delta E_{redox}°$	exp
ferredoxin (Fd)	−4.990	8.422	−0.998	8.465	−0.955	−0.440
PDR	−4.967	8.583	−0.814	8.643	−0.754	−0.170
Δ(Fd − PDR)	−0.023	−0.161	−0.184	−0.178	−0.201	−0.266

[a] Energies are given in electron volts. The redox potentials are calculated according to eqs 18 and 19.

of protein interaction and reaction field energy E_{pr}. Such a calculation can be done with only the MEAD package as discussed in section II.4, and the reaction field and protein interaction energy E_{pr} in eq 12 is simplified to:

$$E_{pr} = E_{prot} + E_{react}$$
$$= {}^1/_2\sum_i q_i\phi^*_{react}(i) + \sum_i q_i\phi_{prot}(i) \qquad (19)$$

where q_i are the ESP charges of the $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters and $\phi^*_{react}(i)$ and $\phi_{prot}(i)$ are the reaction field and protein field potentials at the atomic nuclei as obtained from the solution to the Poisson−Boltzmann equation. This procedure approximates the initial electrostatic interaction from the SCRF calculation, without the iteration loop and without the integrations in eqs 13 and 14. Table 9 compares E_{prot}, E_{react}, and E_{pr} obtained by SCRF and MEAD calculations. For protein interaction energy E_{prot}, the MEAD results are close to E_{prot}^{el} calculated by SCRF, while for reaction field energy, E_{react}(MEAD), is close to E_{react}-(SCRF) (within about 5 kcal/mol). The total E_{pr} values obtained from two approaches agree very well with a difference of about 2−3 kcal/mol. The redox potentials calculated from the MEAD data, as shown in Table 10, compare very well with $\Delta E_{redox}°$ from SCRF calculations in Table 6. The SCRF total energies for E_{prot} and E_{react} are larger than those from MEAD, but there is partial compensation from the E_{strain} term (of opposite sign), which is absent in the simpler MEAD approach. E_{strain}(MEAD) is equal to zero since no cluster polarization is allowed with the single-step MEAD method.

Once the gas-phase ESP charges are fitted from density functional electrostatic potentials of the cluster, the MEAD calculation only takes a small fraction of the CPU time of a full SCRF calculation and therefore provides an economical way to estimate protein interaction and reaction field energies. To evaluate the comparative contribution to the interaction between the $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ clusters and the residues in PDR and ferredoxin proteins, we carried out a series of MEAD calculations with two or nine closest peptide dipoles removed following the earlier work of Correll et al.[8] This can be done by setting the charges of H, N, C, and O atoms in the peptide backbone

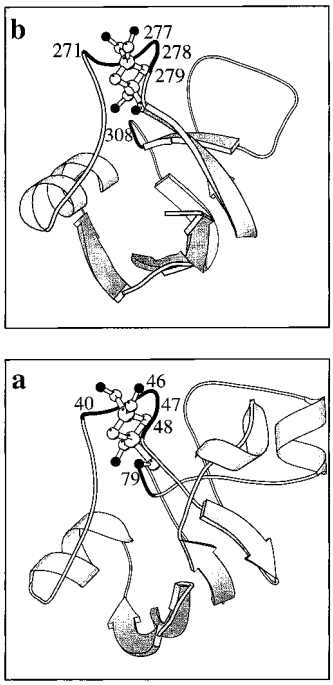


Figure 6. Nine (two) peptide backbone dipoles closest to $[Fe_2S_2(SH_3)_4]$ clusters are from (a) residues 40−48 and 79 (47 and 48) in ferredoxin from *Anabaena* 7120. (b) Residues 271−279 and 308 (277 and 278) in phthalate dioxygenase reductase (PDR) from *Pseudomonas cepacia*. These residues are marked by black coils.

of selected residues to zero. The two closest peptide dipoles are from residues 46 and 47 in ferredoxin and residues 277 and 278 in PDR. The nine closest peptide dipoles include the loop regions from residues 40−48 and 79 in ferredoxin and residues 271−279 and 308 in PDR. These residues are sketched in Figure 6. Such a procedure can be regarded as a "computational mutation", from which we are able to identify the residues important in the interaction of the protein with the active-site clusters. However, these computational techniques also give us access to parts of the protein that cannot be easily and predictably changed in traditional mutagenesis, such as the peptide backbone. Table 11 summarizes the MEAD-calculated ΔE_{prot} with either the two or nine closest peptide dipoles removed in the two proteins. Specifically, the charges are set to zero but the atom positions and corresponding low-dielectric regions are retained. The calculated data show that the protein interaction energy difference between ferredoxin and PDR, ΔE_{prot}(Fd − PDR), dramatically decreases if the nine closest peptide dipoles are removed. The nine closest dipoles have a strong stabilizing effect on both PDR and ferredoxin, but the main effect on the difference in ΔE_{prot} comes from the two

TABLE 11: Effects of Peptide Dipoles on Protein Interaction Energy ΔE_{prot}[a]

		MEAD[b] ($I = 0.05$ M)	MEAD[b] ($I = 0.15$ M)	DelPhi[c] ($I = 0.15$ M)
full protein	ferredoxin(Fd)	0.833	0.844	0.741
	PDR	1.114	1.123	0.968
	Δ(Fd − PDR)	−0.281	−0.279	−0.227
protein with two closest dipoles removed	ferredoxin (Fd)	0.813	0.828	0.751
	PDR	0.906	0.913	0.870
	Δ(Fd − PDR)	−0.093	−0.085	−0.119
protein with nine closest dipoles removed	ferredoxin (Fd)	0.132	0.148	0.272
	PDR	0.206	0.214	0.358
	Δ(Fd − PDR)	−0.074	−0.066	−0.086

[a] Energies are given in electron volts. [b] This work. [c] Reference 8.

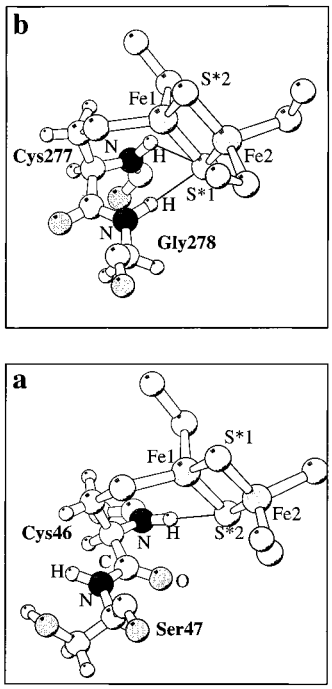


Figure 7. Hydrogen bonds connecting [2Fe2S] core and two closest residues in (a) ferredoxin from *Anabaena* 7120, only one such NH−S* hydrogen bond, and in (b) phthalate dioxygenase reductase (PDR) from *Pseudomonas cepacia*, two NH−S* hydrogen bonds.

closest dipoles. These two closest peptide dipoles from Cys46 and Ser47 in ferredoxin and Cys277 and Gly 278 in PDR contribute about two-thirds of ΔE_{prot}(Fd − PDR). The origin of this difference can be further attributed to different H-bond patterns between the active-site clusters and the two closest residues in the proteins. As can be seen in Figure 7, there are two NH−S* hydrogen bonds in PDR connecting [2Fe2S] core and Cys277 and Gly278, respectively, while there is only one such hydrogen bond in *Anabaena* ferredoxin. The H−N bond of Ser47 in ferredoxin points away from [2Fe2S] core. Given the highly negative charge on S*, especially in reduced form, such an NH−S* hydrogen bond is rather strong. The same conclusion has been achieved by Correll *et al.*[8] by an electrostatics calculation using the DelPhi program[41] with CHARMM charges[42] for proteins and the density functional ESP charges[4a] for the cluster. Their calculated data are included in Table 11 as well for comparison. By comparison with Table 12, ΔE_{prot} for either Fd or PDR is a small fraction (about 10−15%) of the total protein and reaction field energy ΔE_{pr}, but ΔE_{prot} accounts for the main part of the change in ΔE_{pr} from Fd to PDR, Δ(Fd − PDR). Table 7 shows that the full SCRF calculation gives the same relationship. Table 12 shows a breakdown of ΔE_{pr} into $\Delta E_{react} + \Delta E_{prot}$ for MEAD, similar to the SCRF breakdown in Table 7.

TABLE 12: Energy Decomposition of ΔE_{pr} from MEAD Calculations[a]

	ΔE_{prot}	ΔE_{react}	ΔE_{pr}
	$I = 0.00$ M		
ferredoxin (Fd)	0.798	7.624	8.422
PDR	1.081	7.502	8.583
Δ(Fd − PDR)	−0.283	0.122	−0.161
	$I = 0.05$ M		
ferredoxin (Fd)	0.833	7.632	8.465
PDR	1.114	7.528	8.643
Δ(Fd − PDR)	−0.281	0.104	−0.178

[a] Energies are given in electron volts.

6. Previous Calculations of Redox Potential Shifts. A number of recent papers contain redox potential shift calculations for iron−sulfur proteins.[43−47] These often use charge models from our previous density functional calculations[4a,48] to calculate redox shifts within a set of structurally related 1Fe, 2Fe2S, 3Fe4S, or 4Fe4S proteins. We will focus on recent work on 2Fe2S proteins.

Stephens et al.[43] have calculated redox shifts for a variety of iron−sulfur proteins using their protein dipoles Langevin dipoles (PDLD) method in combination with molecular dynamics (MD) studies. They used iron−sulfur cluster charge models from our earlier (1985) X_α scattered wave method[48] charge partitioning instead of our more recent ESP charges from ADF calculations.[4a] This could have a significant[4a,44] but probably not a major effect on the reaction and protein field energies. In previous work, Jensen et al.[44] found from PDLD that total protein field plus reaction field (ΔE_{pr}) energies changed by about −300 mV comparing X_α scattered wave and ESP-derived charges in 4Fe4S ferredoxins, while relative redox shifts varied by −30 to −50 mV. Similarly, we[4a] found differences in solvation energies of −120 and −50 mV comparing Mulliken and ESP charges for $[Fe_2S_2(SR)_4]^{2-,3-}$ and $[Fe_4S_4(SR)_4]^{2-,3-}$ clusters in pure solvent.

In the PDLD method,[43] the Langevin dipoles represent solvent water molecules. In the MD-PDLD studies, a more detailed model is used compared to the static PDLD calculations. Atomic level water molecules and protein atoms are used in the dynamics, with the atomic level waters replacing the Langevin dipoles within a 12 Å sphere centered on the iron−sulfur cluster. Between 12 and 18 Å, Langevin dipoles represent the solvent, and a continuum water representation is used beyond 18 Å. Since they do not predict absolute redox potentials, the fit to a set of redox potentials for a given protein structural-redox class involves a free parameter. MD improves the calculated accuracy of redox shift predictions for all classes of iron−sulfur proteins compared to the PDLD method for static structures, and the final RMS errors are less than 50 mV for all classes. In their selected set of 2Fe2S proteins, there were four examples (including *Anabaena* ferredoxin, but not PDR) but all have experimental redox potentials within a very narrow range with only a span of 60 mV from −380 to −440 mV. It

is then difficult to assess the significance of the predicted shifts; the RMS deviation from experimental shifts is 40 mV and the calculated span of potentials is 248 mV compared to 60 mV (exp). The 3Fe4S and 4Fe4S proteins they have included span a broader experimental redox potential range and provide better evidence for the significance of the theoretical shift calculations.

In the MD-PDLD methodology, all acid or basic side chains in the protein are made neutral; later, the correction for the expected charges of side chains is calculated separately by using a macroscopic dielectric screening ($\epsilon = 80$). In the MD-PDLD results, these terms are smaller than -170 mV (absolute), and relative redox shifts are less than -70 mV (differences within the class) for these 2Fe2S systems. We have calculated the interaction of these charged side chains with the 2Fe2S complex directly. These charged side chains are well screened by solvent, and as shown in Table 11, the major contribution to the protein field energy ΔE_{prot} comes from the nine closest dipoles in both 2Fe2S proteins.

From their Table 4, Stephens et al.[43] find that the total protein plus reaction field contribution to the redox potential for $[Fe_2S_2(SR)_4]^{2-,3-}$ clusters in 2Fe2S proteins is about 10.0−10.3 eV (with the smallest value for *Anabaena* ferredoxin). This is about 20% larger than the reaction field plus protein field energy we calculate ($\Delta E_{\text{pr}} = 8.41$, 8.58 eV in Table 6), but this is still fairly comparable. (The 2Fe2S proteins we have studied span a wider range of redox potentials since both PDR and *Anabaena* ferredoxin are included.) From our calculated IP_{red} and the experimental $\Delta E^{\circ}_{\text{redox}}$ and using eq 18, we can estimate experimental values for ΔE_{pr} giving +8.98 and +9.23 eV for *Anabaena* ferredoxin and PDR. These are 7−8% above our calculated total protein−reaction field energies.

Banci et al.[46] recently performed DelPhi electrostatic calculations on the algal 2Fe2S ferredoxin from *Spirulina platensis* to evaluate the comparative energies for reduction of alternate Fe sites of this protein (labeled Fe_A and Fe_B depending on which Cys residues are bound; Cys41 and Cys46 are bound to Fe_A). In this paper, the comparative redox potentials for a number of 4Fe4S high-potential proteins (HIPIP's) were also evaluated and compared with experimental redox shifts; further, for two different 4Fe4S proteins, the relative energies for different locations of the mixed-valence versus ferric pair were evaluated. For most of these systems, only net redox shift data were reported, but a more detailed breakdown was provided for the electrostatic contributions to the redox potential for *Spirulina platensis* 2Fe2S ferredoxin. Total electrostatic redox free energy differences $\Delta G^{\text{pr}}_{\text{red}}$ (or ΔG_{red} in their notation) and breakdowns into solvation $\Delta\Delta G_{\text{solv}}$ (presumably the reaction field redox energy) and charge and dipole contributions (ΔG_{ch}, ΔG_{dip}) to the protein field energy were reported. [It is useful to remember that the redox free energy is related to the redox potential by $\Delta G = -nF\ \Delta E^{\circ}_{\text{redox}}$, where F = the Faraday constant and $n = 1$ (number of moles of electrons transferred).] The reported total $\Delta G^{\text{pr}}_{\text{red}}$ are only equivalent to +0.74 and +0.84 eV for sites Fe_A and Fe_B, which are both far too small compared to the ΔE_{pr} we have found (see also Stephens et al.[43] or our previous work[4a] on 2Fe2S clusters in solvent environments). Further, it appears that the sign is also incorrect, since with a positive $\Delta G^{\text{pr}}_{\text{red}}$, the corresponding ΔE_{pr} would be negative. The origin of this discrepancy is not clear. The DelPhi protein field energy from the work of Correll et al.[8] is quite similar to our calculated result (see Table 11), and there is no reason the reaction field energy should not also be comparable to ours or that of Stephens. (Another indication of a problem is the very small values reported for $\Delta G_{\text{ch}} + \Delta G_{\text{dip}}$ of -0.04 and -0.02 eV for Fe_A and Fe_B. Further, the finest grid spacing used was evidently 0.67 Å, which is very coarse; our finest grid spacing is 0.15 Å, but it is unclear whether this is the source of the problem.) It is potentially relevant to note that in calculations of the electrostatic contributions to the redox potential, partial charges on the entire $Fe_2S^*_2(SR)_4$ unit must be included in the energy evaluation and that the entire charge distributions of the oxidized and reduced clusters must be used to evaluate the reaction field potentials and energies [the set of $(q_i)_{\text{ox}}$, $(q_i)_{\text{red}}$ must be used and not just (Δq_i)]. A reexamination of these issues with new calculations would be appropriate.

V. Conclusions

We have developed a method for incorporating the protein/solvent environment in density functional electronic structure calculations and applied this to redox potential calculations for two 2Fe2S proteins. This extends our previous work that included solvent effects alone.[3,4] Here, we deal with a heterogeneous environment consisting of multiple dielectric regions (with partial charges on the atoms of the protein), and using the finite-difference solution of the classical linearized Poisson−Boltzmann equation. Ionic strength effects are also included, but are small for these systems. The resulting reaction field and protein field potentials can be incorporated in the quantum DFT Hamiltonian and give cluster-field interaction energies of various types that are valuable both for energy calculations and for detailed analysis. The more complex self-consistent reaction field (SCRF) method is compared with the simpler one-cycle MEAD method. Both global energy analysis (reaction field vs protein field, direct electrostatic effects vs polarization effects) and analysis of the charge effects of different amino acid residues (by analysis of the effects of setting certain atom charges to zero) are quite useful. Direct electrostatic effects from the reaction field and protein field−cluster interaction are much larger than terms arising from cluster electronic polarization and electronic strain. The ease with which charges can be turned on/off in the energy analysis, particularly in the simpler one-cycle MEAD method, gives considerable versatility to this form of computational mutagenesis. This approach provides a set of tools for analyzing metalloprotein energetics and for planning new experiments. The same methodology can be applied to other biomolecular environments.

The redox potentials calculated by the current method deviate from experimental values by about 0.5−0.6 V in absolute magnitude. (This error is only about 6−7% of the total reaction field plus protein field interaction energy, ΔE_{pr}, showing the difficulty of obtaining highly accurate absolute redox potential predictions.) However, the calculations reproduce the experimental shift in redox potential between the two 2Fe2S proteins with good accuracy. The calculated $\Delta E^{\circ}_{\text{redox}}$ for phthalate dioxygenase reductase (PDR) from *Pseudomonus cepacia* is predicted 195 mV more positive than that of ferredoxin from *Anabaena* 7120, comparing well with the experimental value of 266 mV. The calculations also reveal that this redox potential difference arises mainly from the protein environment and not from the solvent.

Energy analyses of the composition of the redox potentials are given in Tables 6, 7, and 10−12. These analyses lead to the following conclusions: (1) The electronic structures of the active-site clusters of PDR and *Anabaena* ferredoxin are very similar (both before and after geometry optimization), as reflected in the nearly equal cluster ionization energies (IP_{red}) for the isolated $Fe_2S_2(SCH_3)_4$ clusters; (2) The reaction field

contributions to the redox potentials are extremely large (about 7.8 eV), but nearly equal for PDR and *Anabaena* ferredoxin (within about 2%). A large stabilizing reaction field energy is necessary for the physical stability of the reduced cluster in the protein/solvent environment ($IP_{ps} > 0$). Since solvation is the dominant energy term here, the near invariance implies that the solvent access to the active-site iron−sulfur cluster is also very similar. This is not so evident from the two protein structures; these are fairly structurally homologous in the 2Fe2S binding domain, but PDR has three domains while *Anabaena* ferredoxin has only one. Evidently, there is considerable solvent access to the PDR 2Fe2S active site via the hinge region between the 2Fe2S binding domain (domain 1) and the other two domains (see Figure 3b,c). (Further evidence for this conclusion is that the iron−sulfur cluster redox potential is rather insensitive to the FMN atom charges.) Although there are only two examples of 2Fe2S proteins in the present study, we expect on the basis of our results that solvent access to the active site and the solvation contribution to the redox potential is tightly controlled in most 2Fe2S proteins (and within any set of iron−sulfur proteins of a particular structural class and redox couple). This is necessary to obtain redox potentials within an appropriate physiological range; PDR and *Anabaena* ferredoxin represent the extreme limits of known 2Fe2S protein redox potentials.[7] Overall, the reaction field term (mainly from the solvent−cluster interaction) can be considered to provide "coarse tuning" of the redox potential, while the protein field interaction with the iron−sulfur cluster provides the "fine-tuning" (as shown below). Even coarse tuning is difficult, since the reaction field must be controlled to roughly 2%. (3) The protein field contributions to the redox potentials, while only a small fraction (10−15% or 0.9−1.2 eV) of the corresponding reaction field terms, are critical for differentiating the redox potentials of these two proteins. For PDR and *Anabaena* ferredoxin, the protein interaction energy is dominated by the nine closest peptide dipoles with most of these in H-bonding contact with the 2Fe2S cluster via thiolate and bridging sulfurs. (4) The main part of the redox potential difference between PDR and *Anabaena* ferredoxin is due to the two closest peptide dipoles (see Figure 7), and in particular to the main-chain amide (N−H)−S* hydrogen bond that is eliminated when the conformation of Gly278 in PDR is compared to that of Ser47 in *Anabaena* ferredoxin. These results for the protein field−cluster interaction energy agree with those of Correll et al.[8] (Table 11), who used the DelPhi method and our ESP charges for model $[Fe_2S_2(SCH_3)_4]^{2-,3-}$ complexes[4] (only the 2Fe2S binding domain of PDR was used in the calculations of Correll et al.[8])

Several factors have not been taken into account at the current stage of the calculations. First, due to the partition of quantum/classical regions, charge transfer in the hydrogen bonds that connect the active site (quantum region) with the (classical) protein environment was not properly treated. While there are fundamental difficulties at any quantum−classical interface (since charge transfer and Pauli repulsion are quantum mechanical in nature), this problem can be effectively reduced by enlarging the quantum region to include the hydrogen bonds. This is planned in future work. Second, we ignored geometry change and relaxation of the proteins upon reduction, although the dielectric constant ($\epsilon = 4$) adopted for the proteins implicitly allows some mobility of the protein dipoles (compared with a purely electronic dielectric constant $\epsilon = 2$). Again, this problem will be reduced but not eliminated by expanding the quantum region. The highly charged active site (2−,3− charges) represents a difficult case for both of these effects. These two factors are expected to be the main sources of the calculational errors in absolute redox potentials. Inclusion of these two factors in the calculations will probably stabilize the reduced form and as a result shift up redox potentials to some extent (making these more positive). The last factors omitted are associated with the molecular dynamics of the quantum active site and the protein and solvent environment along with zero-point vibrational energy. There are both enthalpic and entropic terms of this type at finite temperatures. In the protein, the dielectric constant ($\epsilon = 4$) accounts for the free energy effects of dipole mobility, but only approximately. The same can be said for the average macroscopic dielectric constant ($\epsilon = 80$) for the aqueous solvent. Some solvent will be immobilized near charged surface groups, especially near the numerous carboxylate side chains. Our treatment of the quantum problem and the associated electrostatics of the protein/solvent environment provides a basis for at least an approximate examination of dynamics. For example, one could try to separate the electrostatic part of the problem and related dynamics from the quantum electronic structure and charge-transfer issues, using ESP charges for the quantum active site and force-field charges for the protein for molecular mechanics and dynamics studies.

One significant advantage of our approach is that the major energetic contributions to the absolute redox potential are all calculated. This allows us to assess the comparative sizes of the different energy terms entering into the redox potential, which provides an understanding of orders of magnitude and sensitivity to different physical terms. For example, the final redox potentials are fairly insensitive to changes in Heisenberg spin coupling parameters but are sensitive to eliminating one charged hydrogen bond (NH−S*). The effects of successive protein shells and of solvent access can be similarly evaluated, or the sensitivity to changes in protein structures or differences in atom positions from different X-ray protein refinement models.

Acknowledgment. This work was supported by NIH Grants GM39914 to D. A. Case and GM45607 to D. Bashford and by NSF Fellowship DGE9616174 to M. R. Nelson. We thank D. A. Case, C. L. Fisher, and M. Ludwig for many useful discussions and E. J. Baerends and the Amsterdam group for the use of the ADF package. The calculations were carried out on the HP 9000/735 workstations at The Scripps Research Institute.

Supporting Information Available: Five tables listing Cartesian coordinates and ESP charges of all [2Fe-2S] clusters and FMN cofactor (5 pages). Ordering information is given on any current masthead page.

References and Notes

(1) (a) Warshel, A. *Computer Modelling of Chemical Reactions in Enzymes and Solutions;* Wiley: New York, 1991. (b) Warshel, A. *Curr. Opin. Struct. Biol.* **1992**, *2*, 230. (c) Merz, K. M., Jr. *Curr. Opin. Struct. Biol.* **1993**, *3*, 234. (d) Mulholland, A.; Grant, G. H.; Richards, W. G. *Protein Eng.* **1993**, *6*, 133.

(2) (a) Tapia, O.; Johannin, G. *J. Chem. Phys.* **1981**, *75*, 3624. (b) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523. (c) Singh, O. J.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718. (d) Field, M. J.; Bash, P. A.; Karplus, M. A. *J. Comput. Chem.* **1990**, *11*, 710. (e) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199. (f) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298. (g) Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1995**, *99*, 11266.

(3) (a) Chen, J. L.; Noodleman, L.; Case, D. A.; Bashford, D. *J. Phys. Chem.* **1994**, *98*, 11059. (b) Richardson, W. H.; Peng, C.; Bashford, D.; Noodleman, L.; Case, D. A. *Int. J. Quantum Chem.* **1997**, *61*, 207.

(4) (a) Mouesca, J.-M.; Chen, J. L.; Noodleman, L.; Bashford, D.; Case, D. A. *J. Am. Chem. Soc.* **1994**, *116*, 11898. (b) Fisher, C. L.; Chen, J. L.; Li, J.; Bashford, D.; Noodleman, L. *J. Phys. Chem.* **1996**, *100*, 13498. (c)

Li, J.; Fisher, C. L.; Chen, J. L.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1996**, *35*, 4694. (d) Li, J.; Noodleman, L. in *Spectroscopic Methods in Bioinorganic Chemistry*, Solomon, E. I.; Hodgson, K. O., Eds.; American Chemical Society, Washington, DC, 1998.

(5) (a) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305. (b) Liotard, D. A.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **1995**, *16*, 422. (c) Karelson, M. M.; Zerner, M. C. *J. Phys. Chem*. **1992**, *96*, 6949. (d) Furuki, T.; Sakurai, M.; Inoue, Y. *J. Comput. Chem.* **1995**, *16*, 378. (e) Wang, B.; Ford, G. P. *J. Chem. Phys*. **1992**, *97*, 4162. (f) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys*. **1981**, *55*, 117. (g) Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc*. **1991**, *113*, 4776. (h) Marcos, E. S.; Pappalardo, R. R.; Rinaldi, D. *J. Phys. Chem*. **1991**, *95*, 8928. (i) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringbalda, M.; Goddard, W. A., III; Honig, B. *J. Am. Chem. Soc*. **1994**, *116*, 11875. (j) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans.* **1993**, 799. (k) Truong, T. N.; Stefanovich, E. V. *Chem. Phys. Lett.* **1995**, *240*, 257. (l) Ruiz-Lopez, M. F.; Bohr, F.; Martins-Costa, M. T. C.; Rinaldi, D. *Chem. Phys. Lett.* **1994**, *221*, 109. (m) Adamo, C.; Lelj, F. *Chem. Phys. Lett.* **1994**, *223*, 54. (n) Fortunelli, A.; Tomasi, J. *Chem. Phys. Lett.* **1994**, *231*, 34. (o) Baldridge, K.; Fine, R.; Hagler, A. *J. Comput. Chem*. **1994**, *15*, 1217. (p) Hall, R. J.; Davidson, M. M.; Burton, N. A.; Hillier, I. H. *J. Phys. Chem*. **1995**, *99*, 921.

(6) (a) Bashford, D. *Curr. Opin. Struct. Biol.* **1991**, *1*, 175. (b) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473. (c) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610. (d) Bashford, D.; Case, D. A.; Dalvit, C.; Tennant, L.; Wright, P. E. *Biochemistry* **1993**, *32*, 8045.

(7) (a) Cammack, R. *Adv. Inorg. Chem.* **1992**, *38*, 281. (b) Johnson, M. K. in *Encyclopedia of Inorganic Chemistry*; King, R. B., Ed.; John Wiley: New York, 1994; and references within.

(8) Correll, C. C.; Ludwig, M. L.; Bruns, C. M.; Karplus, P. A. *Protein Sci.* **1993**, *2*, 2112.

(9) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

(10) (a) Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem. Phys*. **1973**, *2*, 41. (b) Ravenek, W. in *Algorithms and Applications on Vector and Parallel Computes*; te Riele, H. J. J., Dekker: Th. J., van de Vorst, H. A., Eds.; Elsevier: Amsterdam, 1987.

(11) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys*. **1980**, *58*, 1200.

(12) Becke, A. D. *J. Chem. Phys.* **1986**, *84*, 4524.

(13) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822. Erratum, *ibid.* **1986**, *34*, 7406.

(14) (a) Boerrigter, P. M.; te Velde, G.; Baerends, E. J. *Int. J. Quantum Chem*. **1988**, *33*, 87. (b) te Velde, G.; Baerends, E. J. *J. Comput. Phys.* **1992**, *99*, 84.

(15) (a) Snijders, J. G.; Baerends, E. J.; Vernooijs, P. *At. Nucl. Data Tables* **1982**, *26*, 483. (b) Vernooijs, P.; Snijders, J. G.; Baerends, E. J. *Slater Type Basis Functions for the Whole Periodic System;* Internal report: Free University of Amsterdam: The Netherlands, 1981. (c) Krijn, J.; Baerends, E. J. *Fit Functions in the HFS-method;* Internal report (in Dutch); Free University of Amsterdam, The Netherlands, 1984.

(16) (a) Versluis, L.; Ziegler, T. *J. Chem. Phys.* **1988**, *88*, 322. (b) Fan, L.; Ziegler, T. *J. Chem. Phys.* **1991**, *95*, 7401.

(17) (a) Schlegel, H. B. In *Ab initio Methods in Quantum Chemistry I*; Lawley, K. P., Ed; Advances in Chemistry and Physics, Vol. 67; Wiley: New York, 1987. (b) Head, J. D.; Zerner, M. C. *Adv. Quantum Chem.* **1988**, *20*, 1.

(18) (a) Ziegler, T.; Tschinke, V.; Baerends, E. J.; Snijders, J. G.; Ravenek, W. *J. Phys. Chem*. **1989**, *93*, 3050. (b) Snijders, J. G.; Baerends, E. J. *Mol. Phys*. **1979**, *36*, 1789. (c) Snijders, J. G.; Baerends, E. J.; Ros, P. *Mol. Phys*. **1979**, *38*, 1909.

(19) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.

(20) (a) Press, W. H.; Flanery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes*, The Art of Scentific Computing; Cambridge University Press: Cambridge, England, 1986. (b) Francl, M. M.; Carey, C.; Chirlian, L. M.; Gange, D. M. *J. Comput. Chem.* **1996**, *17*, 367.

(21) (a) Harvey, S. C. *Proteins* **1989**, *5*, 78. (b) Sharp, K.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem*. **1990**, *19*, 301. (c) Simonson, T.; Brooks, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452. (d) Simonson, T.; Perahia, D. *J. Am. Chem. Soc.* **1995**, *117*, 7987. (e) Gilson, M. K.; Honig, B. H. *Biopolymers* **1986**, *25*, 1097. (f) Smith, P. E.; Brunne, R. M.; Mark, A. E.; van Gunsteren, F. *J. Phys. Chem.* **1993**, *97*, 2009. (g) Beroza, P.; Fredkin, D. R.; Okamura, M. Y.; Feher, G. *Biophys. J.* **1995**, *68*, 2233. (h) Tanford, C.; Roxby, R. *Biochemistry* **1972**, *11*, 2192. (i) Beroza, P.; Case, D. A. *J. Phys. Chem.* **1996**, *100*, 20156. (j) Bone, S.; Pethig, R. *J. Mol. Biol.* **1985**, *181*, 323.

(22) (a) Bondi, A. *J. Chem. Phys.* **1964**, *64*, 441. (b) Connolly, M. L. *Science* **1983**, *221*, 709.

(23) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978.

(24) Rypniewski, W. R.; Breiter, D. R.; Benning, M. M.; Wesenberg, G.; Oh, B.-H.; Markley, J. L.; Rayment, I.; Holden, H. M. *Biochemistry* **1991**, *30*, 4126.

(25) Correll, C. C.; Batie, C. J.; Ballou, D. P.; Ludwig, M. *Science* **1992**, *258*, 1604.

(26) (a) Mayerle, J. J.; Denmark, S. E.; De Pamphilis, B. V.; Ibers, J. A.; Holm, R. H. *J. Am. Chem. Soc*. **1975**, *97*, 1032. (b) Holm, R. H. *Acc. Chem. Res*. **1977**, *10*, 427. (c) Nakamura, A.; Ueyama, N. In *Encyclopedia of Inorganic Chemistry*; King, R. B., Ed.; John Wiley: New York, 1994.

(27) (a) Norman, J. G. Jr.; Ryan, P. B.; Noodleman, L. *J. Am. Chem. Soc*. **1980**, *102*, 4279. (b) Noodleman, L.; Baerends, E. J. *J. Am. Chem. Soc.* **1984**, *106*, 2316.

(28) (a) Bertrand, P.; More, C.; Camensuli, P. *J. Am. Chem. Soc.* **1995**, *117*, 1807. (b) Dugad, L. B.; La Mar, G. N.; Banci, L.; Bertini, I. *Biochemistry* **1990**, *29*, 2263. (c) Skjeldahl, L.; Westler, W. M.; Oh, B.-H.; Krezel, A. M.; Holden, H. M.; Jacobson, B. L.; Rayment, I.; Markley, J. L. *Biochemistry* **1991**, *30*, 7363.

(29) (a) Noodleman, L.; Case, D. A. *Adv. Inorg. Chem*. **1992**, *38*, 423. (b) Noodleman, L.; Peng, C. Y.; Case, D. A.; Mouesca, J. M. *Coord. Chem. Rev.* **1995**, *144*, 199.

(30) (a) Noodleman, L.; Case, D. A.; Mouesca, J. M.; Lamotte, B. *J. Bio. Inorg. Chem.* **1996**, *1*, 177. (b) Blondin, G.; Girerd, J. J. *Chem. Rev.* **1990**, *90*, 1359.

(31) Palmer, G.; Dunham, W. R.; Fee, J. A.; Sands, R. H.; Izuka, T.; Yonetani, T. *Biochim. Biophys Acta* **1971**, *240*, 201.

(32) Gillum, W. O.; Frankel, R. B.; Foner, S.; Holm, R. H. *Inorg. Chem.* **1976**, *15*, 1095.

(33) Petersson, L.; Cammack, R.; Rao, K. K. *Biochim. Biophys. Acta* **1980**, *622*, 18.

(34) Palmer, G. In *Iron*−*Sulfur Proteins*; Lovenberg, W., Ed.; Academic Press: New York, 1973.

(35) Hurley, J. K.; Salamon, Z.; Meyer, T. E.; Fitch, J.; Cusanovich, M. A.; Markley, J.; Cheng, H.; Xia, B.; Chae, Y. K.; Medina, M.; Gomez-Moreno, C.; Tollin, G. *Biochemistry* **1993**, *32*, 9346.

(36) Gassner, G.; Wang, L.; Batie, C.; Ballou, D. P. *Biochemistry* **1994**, *33*, 12184.

(37) Reiss, H.; Heller, A. *J. Phys. Chem.* **1985**, *89*, 4207.

(38) Holm, R. H.; Ibers, J. A. In *Iron-Sulfur Proteins*; Lovenberg, W., Ed.; Academic Press: New York, 1977.

(39) (a) Wheeler, R. A. *J. Am. Chem. Soc.* **1994**, *116*, 11048. (b) Raymond, K. S.; Grafton, A. K.; Wheeler, R. A. *J. Phys. Chem. B* **1997**, *101*, 623. (c) Beveridge, A. J.; Williams, M.; Jenkins, T. C. *J. Chem. Soc., Faraday Trans*. **1996**, *92*, 763. (d) Wu, J. H.; Reynolds, C. A. *J. Am. Chem. Soc.* **1996**, *118*, 10545.

(40) Zhang, L. Y.; Friesner, R. A. *J. Phys. Chem.* **1995**, *99*, 16479.

(41) Gilson, M. K.; Sharp, K. A.; Honig, B. *J. Comput. Chem.* **1988**, *9*, 327.

(42) Brooks, B. R.; Bruccoleri, R. E.; Alafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(43) Stephens, P. J.; Jollie, D. R.; Warshel, A. *Chem. Rev.* **1996**, *96*, 2471.

(44) Jensen, G. M.; Warshel, A.; Stephens, P. J. *Biochemistry* **1994**, *33*, 10911.

(45) Langen, R.; Jensen, G. M.; Jacob, U.; Stephens, P. J.; Warshel, A. *J. Biol. Chem.* **1992**, *267*, 25625.

(46) Banci, L.; Bertini, I.; Savellini, G. G.; Luchinat, C. *Inorg. Chem.* **1996**, *35*, 4248.

(47) Smith, E. T.; Tomich, J. M.; Iwamoto, T.; Richards, J. H.; Mao, Y.; Feinberg, B. A. *Biochemistry* **1991**, *30*, 11669.

(48) Noodleman, L.; Norman, J. G., Jr.; Osborne, J. H.; Aizman, A.; Case, D. A. *J. Am. Chem. Soc.* **1985**, *107*, 3418.