Journal of Chemical Theory and Computation

Density Functional Theory Calculation of Bonding and Charge Parameters for Molecular Dynamics Studies on [FeFe] Hydrogenases

Christopher H. Chang* and Kwiseon Kim

National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, Colorado 80401

Received August 18, 2008

Abstract: We have developed and tested molecular mechanics parameters for [FeS] clusters found in known [FeFe] hydrogenases. Bond stretching, angle bending, dihedral and improper torsion parameters for models of the oxidized and reduced catalytic H-cluster, [4Fe4S]^{+,2+}Cys₄, [4Fe4S]^{+,2+}Cys₃His, and [2Fe2S]^{+,2+}Cys₄, were calculated solely from Kohn—Sham density functional theory and Natural Population Analysis. Circumsphere analysis of the cubane clusters in the energy-minimized structure of the full *Clostridium pasteurianum* hydrogenase I showed the resulting metallocluster structures to be similar to known cubane structures. All clusters were additionally stable in molecular dynamics simulations over the course of 1.0 ns in the fully oxidized and fully reduced enzyme models. Normal modes calculated by quasiharmonic analysis from the dynamics data show unexpected couplings among internal coordinate motions, which may reflect the effects of the protein structure on metallocluster dynamics.

Introduction

The biological mechanisms for production of hydrogen gas are a current topic of great interest. 1,2 In particular, [FeFe] hydrogenases are understood to catalyze H_2 production more effectively than the reverse uptake reaction. As naturally occurring enzymes, they must catalyze reduction of protons at $\sim 10^{-7}$ M concentration, ambient or slightly elevated temperatures, and modest reduction potentials. Furthermore, the delivery of electrons from biological electron donors such as ferredoxin and protons from the cellular milieu must be coordinated. These properties make the [FeFe] hydrogenases attractive models for the development of engineered enzymatic catalysts, protein maquettes, 4 or chemical catalysts.

The most complex [FeFe] hydrogenase so far structurally characterized, hydrogenase 1 from *Clostridium pasteurianum*, contains not only the [2Fe]-[4Fe4S] H-cluster but also two auxiliary cubane-type [4Fe4S] clusters with tetracysteinate ligation, a single [4Fe4S] center with unique Cys₃His ligation, and a [2Fe2S]Cys₄ cluster. Given that *C. pasteurianum* is a strict anaerobe, it is reasonable to expect these clusters to function either as redox cofactors shuttling between their most reduced and next-most reduced states or as structural

In principle, the now widespread availability of userfriendly molecular dynamics and quantum chemistry packages together with communal supercomputing resources brings to bear unprecedented power to simulate and understand the catalytic dynamics of [FeFe] hydrogenases. However, the lack of a complete, consistent set of mechanical parameters derived under a single set of assumptions for the electron transfer and catalytic centers hampers more widespread study of these systems. [2Fe2S] and [4Fe4S] cluster parameters derived from crystallographic study are available (HIC-UP) but with arbitrary force constants. To our knowledge, a systematic effort to derive quantitative bond, angle, and dihedral parameters for the hydrogenase [2Fe], [4Fe4S]Cys₄, [4Fe4S]Cys₃His, and [2Fe2S]Cys₄ centers has not been reported. Here, we use *ab initio* quantum chemistry to generate the force parameters for the metalloclusters. Technical complexities such as the formally multireference

centers in one of those two states. Barring the example of the nitrogenase Fe protein which can achieve a remarkable [4Fe4S]⁰ state, in biological systems this implies cluster valence states of [2Fe]_H, [4Fe4S]^{2+,+}, and [2Fe2S]^{2+,+}, where Roman numerals refer to the formal valence of individual iron ions, and Arabic numerals to the overall valence of the inorganic cluster core.

^{*} Corresponding author e-mail: christopher_chang@nrel.gov.

character of antiferromagnetically coupled metalloclusters and the challenge of defining dihedral angles in cage structures make this process somewhat tedious. However, the *ab initio* calculation of force parameters for these systems has several attractions. Although the ring or cage structures of [2Fe2S] and [4Fe4S] centers make dihedral definition complicated, their lack of rotational freedom means no rotational isomers need consideration, and thus the behavior of dihedral potential functions away from their cluster-associated values may be reasonably neglected. Furthermore, few ad hoc decisions need to be made regarding particular data to fit. Finally, results may be systematically improved through the use of higher levels of theory, more complete basis sets, and fewer assumptions as needed.

In this spirit, we here report a first-generation set of molecular mechanical parameters and topology files for these centers in their expected functional valence states. Results of these parameters for energy minimization, molecular dynamics, and vibrational analysis are compared to quantum mechanical calculations as well as experimentally measured frequencies for the diatomic ligands on the H-cluster [2Fe] core.

Methods

Structural Models. Models for the catalytic center comprised the basic structure of the [2Fe]_H H-cluster core found in [FeFe] hydrogenase, with the sole cysteinate ligand replaced with methylthiol. The use of a proton to approximate [4Fe4S]_H on this core is precedented.⁵⁻⁷ The oxidized form of formal iron valence FeIFeII was modeled with an open ligation site on the distal Fe motivated by the unusually long Fedistal-X bond observed in the original crystallographic structure of the Clostridial enzyme⁸ and the apparent lack of a corresponding ligand in the Desulfovibrio desulfuricans [FeFe] hydrogenase structure.9 The reduced form was modeled with proton ligation to the distal Fe^I. Although this model is of questionable relevance to reaction mechanisms involving solely bridging hydride species, 7,10 it is consistent with other mechanisms that have been proposed,⁵ particularly in the presence of a DTMA bridging dithiolate. 11,12 Methylthiolate replaced cysteinate for all [2Fe2S] and [4Fe4S] clusters. For the special [4Fe4S]HisCys₃ model, histidine was ligated via N^ε, as modeled in the crystallographic structure for CpI (PDB code 1FEH).

Quantum Chemistry. Geometry optimization and frequency calculations employed the Gaussian 2003 package, rev. C.2, and a BLYP/6-31+G* model chemistry. Coulomb fitting was employed to accelerate the SCF process. Higher symmetries possible for the [2Fe2S]Cys₄ and [4Fe4S]Cys₄ clusters were not enforced, i.e., all calculations were performed within C₁ symmetry. Geometries were optimized to default criteria (maximum and rms forces of 0.45 and 0.30 mHa/Bohr, displacements of 1.8 and 1.2 mBohr), and the rms change in the density matrix was set to 10^{-8} ("SCF=Tight") with the default pruned (75,302) integration grid. In certain cases, SCF nonconvergence observed for the iron-sulfur clusters was circumvented by constructing initial guesses from localized Natural Bond Orbitals¹³ generated from

loosely converged Kohn-Sham wave functions. In these cases, broken-spin-symmetry electronic configurations were input, with majority α - and β -spin ions distributed arbitrarily within the model cluster. Frequencies were calculated from analytical second derivatives. Output of the Hessian in internal coordinates was achieved using the Gaussian route flag IOp(7/32) and keyword option Freq=InternalModes. Hessian diagonal elements were scaled by (0.9945)², based on empirical frequency scaling studies. 14 The Urey-Bradley terms included as part of the CHARMM force field were not calculated, which may simply be thought of as neglecting anharmonicity in the angle bending coordinates. To convert force constants from atomic units (Hartree/Bohr² or Hartree/ radian²) to those used in the program package CHARMM (kcal/(mol·Å²) and kcal/(mol·rad²)), force constants were further scaled by 2242.3 (bonds) and 627.49 (bends and torsions).

Definition of Dihedral Angles. Where possible, the force constants calculated with density functional theory are used in harmonic potential functions, which differs from the common parametrization as a trigonometric function. This was the case for angles that could be defined with a single minimum-energy value, which in turn depends on both the dynamic rigidity in three-dimensional space as well as the chosen definition of atom types. Thus, although the cage structures of [FeS] clusters imply limited rotational flexibility, to make our force field as simple as possible we have chosen to limit cluster atom types to one per element, per [FeS] cluster type. In doing so, for the auxiliary electron transfer clusters (i.e., not the [2Fe]_H component of the H-cluster), optimum dihedral angles typically clustered as one or more $\pm \theta$ pairs; unfortunately, such a pattern is difficult to encode exactly with either a single harmonic or trigonometric function. Nevertheless, for these cases with multiple minima arising from degeneracy in our atom type definitions, we have chosen simplicity over precision by choosing the lowest multiplicity of a cosine function that places minima near all optimum θ values. A particularly pernicious example is the FEIR-SIR-FEIR-SIR dihedral associated with the iron and sulfide core of the reduced [4Fe4S]HisCys₃ cluster, with a multiplicity of 40 and phase angle of 180° to encompass true minima at approximately $\pm 9^{\circ}$ and $\pm 91^{\circ}$. The CHARMM potential function $V_{\text{dihedral}} = A[1 + \cos(n\theta + \varphi)]$, with multiplicity n and phase angle φ produces the requisite minima at 9 and 90 degrees, at the cost of introducing numerous spurious minima. However, the relatively rigid cage topology of this cluster type arising from effective dihedral constraints by bond and angle forces prevents sampling of these spurious dihedral minima, as illustrated in Figure 2. We have made a further approximation in taking the harmonic force constants from our *ab initio* calculations directly as cosine amplitudes. Again, however, the lack of physical transitions among the function's minima and the effective constraints from the stiffer bond and angle force constants makes the fine detail of the high-energy values mostly irrelevant.

Approximations for Fe-C-O/N Bending Parameters. The linear angles found in the Fe-C-(O/N) and *trans* S-Fe-C angles created two particular difficulties in parameter derivation. First, care was required to exclude any

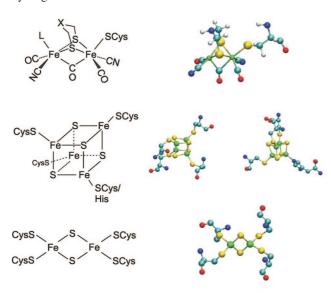


Figure 1. Schematic and CPK representations of [FeFe] hydrogenase electron transfer and catalytic centers considered.

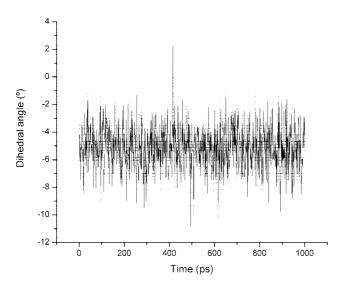


Figure 2. Dynamic stability of the Fe1-S1-Fe2-S2 dihedral angle from [4Fe4S]+HisCys₃ over 1 ns simulation time of the Cpl hydrogenase.

dihedral angle definitions involving three such atoms; the effects of the singularity produced by a linearity in an A-B-C-D system manifested as sudden atomic accelerations and immediate cessation of dynamics simulations. This observation is the primary driver behind our decision to abandon automatic generation of internal coordinates and the explicit definition of bonding topologies we have used. Second, linear bending coordinates are described along deformations in two orthogonal directions; however, the natural mechanism in classical biomolecular force fields is an improper torsion. The two linear bending force constants calculated quantum mechanically were used as constants to describe one improper torsion and one angle bend. By way of example, the FEL1-CLC-CLO1-OL4 (Fe_{prox}-C_{br}-CO_{prox}-O_{CO}) improper torsion force constant was assigned the DFTcalculated force constant for the linear FEL1-CLO1-OL4 bend within the FEL1/CLO1/CLC plane, and the FEL1-CLO1-OL4 bending constant was assigned the average of

the above linear bend constant and that for the FEL1-CLO1-OL4 bend perpendicular to the FEL1/CLO1/OL4 plane as a compromise to describe all possible FEL1-CLO1-OL4 bending motions relative to the FEL1/CLO1/OL4 plane.

van der Waals radii were primarily taken from standard CHARMM atom types, including "S" (sulfide and DTMA thiolate), "CT2" (DTMA methylene C), heme OM (carbonyl O), heme CM (carbonyl C), HA (DTMA methylene H) and HC (HCR hydride), wildcard N (cyanide N, DTMA nitrogen), and heme Fe. 15 We note that the Fe van der Waals interactions are neglected in CHARMM27 with an ε value of 0.0 in the standard Lennard-Jones 6–12 potential form.

Test Simulations. Energy minimization, molecular dynamics, and analysis calculations of the protein system with the exception of frequency estimations were carried out using the NAMD¹⁶ and VMD¹⁷ program packages. Metallocluster normal modes and frequencies were calculated from the trajectory data using the CHARMM program package, version 34. 18 After projecting out translational and rotational motions of the clusters, modes and frequencies were obtained by quasiharmonic analysis. 19 To estimate localized "C-O" and "C-N" stretching frequencies for the H-cluster diatomic ligands, the normal modes were projected onto the appropriate localized mode and vibrational frequency of the localized mode calculated from the weighted rms over the set of normal-mode frequencies²⁰

$$\Lambda = \hat{L}^{\dagger} \hat{H}_{int} \hat{L}
\vec{m}^{t} \cdot \Lambda \cdot \vec{m} = (\vec{m}^{t} \cdot \hat{L}^{\dagger}) \hat{H}_{int} (\hat{L} \cdot \vec{m})
\lambda_{m} = \sum_{i} \lambda_{i} m_{i}^{2} \lambda = 4\pi^{2} v^{2}
\nu_{m} = \sqrt{\sum_{i} \nu_{i}^{2} m_{i}^{2}} = \sqrt{\sum_{i} \nu_{i}^{2} w_{i}}$$
(1)

where L is the matrix of eigenvectors (i.e., normal modes) of the Hessian in internal coordinates H_{int} , L^{\dagger} is its adjoint, Λ is the diagonal matrix of eigenvalues λ_i , \vec{m} is the localized mode vector of interest, v_i is the frequency of the i^{th} normal mode, ν_m is the frequency of the localized mode of interest, and $w_i = \vec{m}_i^t \cdot \vec{m}_i$ is the weight of the i^{th} normal mode in the localized mode.

Simulation parameters other than the force constant and atomic charges considered here were as previously described.²¹ Circumsphere plots of iron and sulfide angular coordinates were made with the Python plotting module Plothon version 0.1.2 (now SVGFig²²). Ellipsoid plots were constructed by determining the appropriate rotations to achieve the desired perspective and then applying these to a reference dynamics frame prior to calculation of anisotropic temperature factors using an in-house VMD script implementing the appropriate calculations.²³ PDB files including the calculated ANISOU records were constructed and used as input to the Raster3D package.²⁴

Results and Discussion

Atom Type Definitions. We have limited our efforts to the two most relevant valence states of each metallocluster considered here, given the requirements for reduction of

Table 1. Residue Names, Atom Type Labels, and Atom Names for Metalloclusters Considered

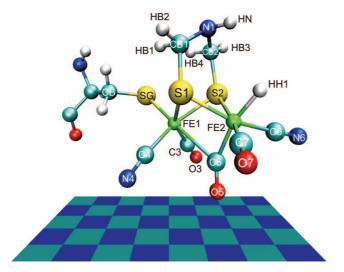
structure	residue	atom types	atom names		
[2Fe2S]Cys ₄ [4Fe4S]Cys ₄ [4Fe4S]Cys ₃ His [2Fe] _H	F2(O/R) F4(O/R) FH(O/R) HC(O/R)	FEK(O/R), SK(O/R) FEJ(O/R), SJ(O/R) FEI(O/R), SI(O/R) FE(M/L)1, FE(M/L)2, S(M/L)1, S(M/L)2, N(M/L)M, O(M/L)4, N(M/L)5, O(M/L)6, C(M/L)O1, C(M/L)N1, C(M/L)O2, C(M/L)N2, C(M/L)C, HC(M/L)1, HC(M/L)2, CC(M/L)1	FE1, FE2, S1, S2 FE1FE4, S1S4 FE1FE4, S1S4 FE1, FE2, S1, S2, N1, (O3 or O7),(N4 or N6), O5, C3, C4, C7, C6,C5, (HB1 or HB4), (HB2 or HB3), (CB1 or CB3)		

protons to H₂ and typical redox potentials in anaerobic bacterial cytoplasm. For the H-cluster [2Fe] core, we have likewise considered two forms; however, given the greater bonding flexibility as compared to [FeS] cluster cofactors, we were forced to make several choices regarding the nature of the open coordination site found on Fedistal. First, although the nature of the (μ_2,μ_2) -bridging dithiolate ligand has been variously proposed as 1,3-propanedithiolate, 9 di(thiomethyl)amine, 11,25 or di(thiomethyl)ether, 26 we have chosen di(thiomethyl)amine (DTMA) due to its attraction as a proton transfer intermediary and the presence of Cys299 within hydrogen bonding distance of the central atom of this bridging ligand. This choice in turn favors a proton transfer mechanism involving a terminal binding mode to Fedistal, at least in the initial stages of proton reduction at [2Fe]_H. Thus, for the "active oxidized" Fe^IFe^{II} form (as compared to the "inactive oxidized" diferrous form found in certain as-isolated enzymes, e.g., Desulfovibrio vulgaris27,28), the distal coordination site was left open based on the unusually long Fe-O bond found in the crystallographic structure and on the need for an open site to permit terminal hydride binding. The fully reduced [2Fe]_H unit was modeled as a terminally protonated Fe^IFe^I species.

Our atom type nomenclature in this report is summarized in Table 1. There are three noncatalytic electron carrier cluster types, each with two potential redox states. The oxidized or reduced state is signaled through a terminal "O" or "R", respectively, on the atom type name for Fe and inorganic sulfide atoms. The nature of the cluster is identified through an index character immediately following

the element name, with $(I, J, K) = ([4Fe4S]Cys_3His,$ [4Fe4S]Cys₄, [2Fe2S]Cys₄). The bonding changes between oxidized and reduced forms of [2Fe]_H were associated with different nomenclature. Atoms of the oxidized cluster are denoted with "M" following the elemental character, with the exceptions of "HCM1", "HCM2", and "CCM1" denoting DTMA methylene hydrogen atoms pointing toward Fe_{proximal}, DTMA methylene hydrogen atoms pointing toward Fe_{distal}, and DTMA methylene carbons, respectively. Terminal carbonyl oxygen atoms are denoted by an index number of "4", the bridging carbonyl oxygen atom by "6", and the terminal cyano nitrogen atoms by "5": thus, "OM6" is the oxygen atom of the bridging CO group. Diatomic ligand carbon atoms bound to Fe_{proximal} have a numeric index of "1"; those bound to Fedistal an index of "2", leading to "CMN2" corresponding to the terminal cyano carbon bound to Fe_{distal}, for example. Nomenclature for reduced [2Fe]_H is identical save for the replacement of "M" with "L" and the terminally bound proton denoted by atom type "HH". The atom naming and typing system is illustrated in Figure 3 for the reduced [2Fe]_H center.

Force constant values and NPA charges are given in their entirety in the Supporting Information. We have generated CHARMM-style topology files with explicit definitions of bond, angle, dihedral angle, and improper torsion angle definitions, allowing the general user to avoid certain problems associated with linear angles found in the [2Fe]_H center and to map cleanly to the available force parameters. In addition, explicit internal coordinate definitions corresponding to the density functional theory-optimized struc-



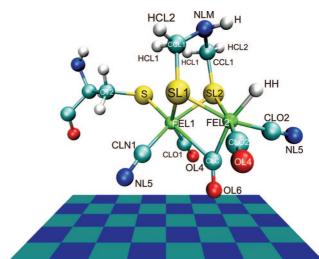


Figure 3. Atom names (left) and types (right) associated with the force field calculated for the reduced [2Fe]_H core of [FeFe] hydrogenases. Oxidized types may be trivially derived from the reduced types by replacing "L" with "M" and neglecting bound hydride "HH".

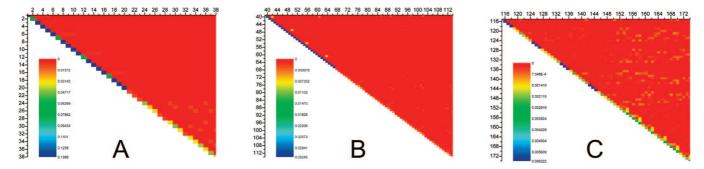


Figure 4. Color maps of Hessian sections calculated for the oxidized [4Fe4S] cluster model. The maximum of the dynamic range is set to the average of the diagonal values for (A) bonds, (B) angles, and (C) dihedrals. Only coordinate couplings within each type of coordinate are shown. Axes are labeled by the internal coordinate indices.

tures for [2Fe]_H are included for automatic regeneration of the complete cluster structure from partial experimental structures, with direct user intervention limited to minor text editing of the Protein Data Bank file. Beyond general build practices for proteins, user responsibility is limited to ensuring proper atom naming in the input PDB file consistent with that in Figure 3 and our convention of naming the Fe atom bound to histidine in [4Fe4S]Cys₃His as FE1.

Quantum Chemistry. Geometry optimizations and charge calculations were straightforward. The three terminal hydrogen charges of methylthiolate were adapted to the two methylene hydrogen atoms of cysteine simply by redistributing the sum of the former three atoms over two, thereby preserving overall the integer charge of the system. Calculation of force constants as diagonal Hessian elements often required repeated calculations with manually specified internal coordinates, as program-generated redundant internal coordinates did not always include those desirable for a molecular mechanics force field. However, after initial storage of force constants in the checkpoint file, such repetitions were very brief, requiring only conversion of Cartesian force constants to internal coordinates and output of the Hessian including the manually specified internal coordinates.

As expected, the orthogonality of coordinates decreases as one moves from two-body bonds, through three-body angles, to four-body dihedral and improper torsions. This is illustrated in Figure 4, where a color map of the upper triangular Hessian for the oxidized [4Fe4S]²⁺ cluster is plotted. The dynamic range maximum was set to the average diagonal value of the bonds (4A), angles (4B), or dihedral angles (4C) to allow visual evaluation of the relative diagonal dominance for each class of parameter. The diagonal dominance is clear in all three cases, with some coordinate coupling becoming apparent for the weakest force constants (dihedrals and impropers). It should be noted that the ringand cage-type structures of the clusters considered constrain the space of achievable geometries. Thus, errors arising from neglected coupling among weak four-body force constants should be "drowned out" in simulation practice by the much stronger bonding interactions. Nevertheless, for the sake of completeness and to include as much physics as possible in this first-generation force field, we have included these fourbody parameters.

Due to the requirement for a complete [2Fe]-[4Fe4S]-(methylthiolate)₄ second derivatives calculation to derive the angular parameters governing the linkage between [2Fe]_H and [4Fe4S]_H, we have chosen to neglect the four-body terms between these entities and to approximate the Fe-S(Cys)-Fe bending force constant with an arbitrary 500 kcal/(mol·rad²) value, with the optimum angle left as that observed in the crystallographic structure for Clostridium pasteurianum hydrogenase I. The treatment of the [6Fe4S] H-cluster as structurally separable [2Fe]_H and [4Fe4S]_H clusters for the purposes of molecular mechanical/dynamical calculations raises the question of their actual physical autonomy. Recent computational and spectroscopic data have revealed a degree of electronic coupling^{29,30} as expected for such a coordinatecovalently linked metallocluster, and careful examination of ⁵⁷Fe hyperfine couplings in the H_{ox} state confirmed exchange coupling between [2Fe]_H and [4Fe4S]_H in this redox state.³¹ The observed exchange coupling (spin polarization) in oneelectron theory is distinct from electron "delocalization" between the clusters in the sense of extended molecular orbitals, which was proposed based on examination of the calculated difference electron density between isolated and structurally connected clusters, canonical (i.e., those diagonalizing the approximate one-electron Hamiltonian) frontier orbital shape, and nonadditivity of open-shell iron character in the sulfur K-edge X-ray absorption bands.³⁰ However, direct physical interpretation of one-electron molecular orbitals, particularly those derived from Kohn-Sham theory,³² can be physically ambiguous, especially among near-degenerate orbitals that may be mixed without dramatically changing the wave function. Self-interaction error has been raised as a particular source of excessive delocalization in canonical Kohn-Sham orbitals of odd-electron systems. 33,34 Spectral analysis of [2Fe]_H, [4Fe4S], and [6Fe4S]_H model complexes shows that the assigned sulfide-to-high-spin-Fe transition intensity in a combined [6Fe4S] model arises from contributions of [4Fe4S] peak intensity, [2Fe]_H low-spin-Fe(I)-to-thiolate tail intensity, and cooperative effects potentially including delocalization;³⁰ however, it is not clear whether the electronic changes expected from exchange coupling alone could account for the observed nonadditivity, or whether electron delocalization is necessary. The pattern and contour values of difference density suggested that the bulk of electronic reorganization between the isolated and combined cluster models occurred among orbitals on indi-

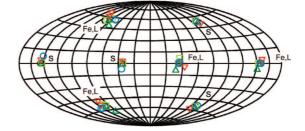


Figure 5. Circumsphere plots of oxidized (left) and reduced (right) Cpl [FeFe] hydrogenase [4Fe4S] cluster coordinates, together with maximally symmetric coordinates. Black circle, perfect symmetry; chartreuse square, [4Fe4S]₅₈₁; green triangle, [4Fe4S]₅₈₂; red inverted triangle, [4Fe4S]₅₈₃; blue circle, Cys₃His[4Fe4S]₅₈₄. Cluster numbering is as in PDB entry 1FEH.

Table 2. Experimental and Geometry-Optimized Circumsphere Radii for CpI [4Fe4S] Clusters

		Fe			sulfide			protein ligands		
	Ехр	Comp _{ox}	Comp _{red}	Ехр	Comp _{ox}	Comp _{red}	Ехр	Comp _{ox}	Comp _{red}	
HYDA ^a	1.663	1.488	1.590	2.244	2.232	2.468	3.949	3.812	4.118	
HYDB	1.689	1.488	1.602	2.245	2.226	2.453	3.974	3.807	4.133	
HYDC	1.688	1.484	1.612	2.242	2.222	2.446	3.965	3.804	4.153	
HYDD	1.675	1.400	1.753	2.232	2.266	2.324	3.880	3.571	4.087	

^a [4Fe4S] hydrogenase residues are here indexed alphabetically, with increasing lexical value reflecting increasing distance from the catalytic [2Fe]_H H-cluster in the *C. pasteurianum* enzyme structure.

vidual ions, rather than between clusters, a result which could arise from atomic orbital near-degeneracy combined with numerical factors in the SCF procedure as well as small perturbations of the cluster electronic structures induced by exchange coupling. The inclusion of the cubane in a [2Fe]_H computational geometry optimization resulted in a weakening of the Fe_{prox}-Cys bond and a shift of the bridging carbonyl group toward the proximal Fe²⁹ relative to the protonated cysteine thiol [2Fe]_H models commonly used, which might be explained by differential charge polarization within the [2Fe]_H cluster induced by a proton versus an exchangecoupled [4Fe4S] center. Overall, the existing evidence suggests that direct electron delocalization between clusters is a small effect and that the dominant direct effect of [4Fe4S] linkage is a modest perturbation of the charge distribution on the [2Fe] center as compared with a cysteine thiol model. The observation of discrete [2Fe]I,II-[4Fe4S]+ ("H_{trans}") and [2Fe]^{I,I}-[4Fe4S]²⁺ ("H_{ox}") valence states during activation in the enzyme from Desulfovibrio desulfuricans35 further suggests that the two subclusters retain some electronic autonomy in the enzyme's more oxidized states. Lack of an observable reduced [4Fe4S]⁺ moiety during catalytic turnover may arise from rapid and directional electron transfer from [4Fe4S]_H to [2Fe]_H, as opposed to inextricable electronic structure in a nondegenerate ground electronic state. Given the evidence, we assert that for the purposes of deriving molecular mechanical parameters, isolated cluster calculations are suitable as an initial approximation. Nevertheless, the effect of electronic coupling between these subclusters on structure and properties will be an area for future refinement and study of the force field.

Test Simulations—Optimized Geometries. The geometry of the *Clostridium pasteurianum* hydrogenase I was optimized in fully oxidized and fully reduced states with the parameters reported. In order to evaluate the modeled energy-minimized geometries, we compare the angular positions of the iron, sulfide, and bonded ligand atoms to an ideal cluster

geometry using the circumsphere methodology of Fee and co-workers.³⁶ Figure 5 shows overlays of atomic positions associated with all [4Fe4S] clusters in the geometry-optimized CpI hydrogenase in its fully oxidized and reduced states, with an idealized cluster comprising three Platonic tetrahedra circumscribed about a common origin. Comparison to this Platonic ideal and to the clusters examined in ref 28 shows excellent angular overlap, confirming that the derived [4Fe4S] parameters preserve the expected angular distribution of Fe and S ions, while still allowing adjustment to the protein environment.

Average calculated Fe-sulfide and Fe-protein ligand bond lengths were 2.23 \pm 0.012 and 2.29 \pm 0.086 and 2.43 \pm 0.032 and 2.50 ± 0.11 for the minimum-energy oxidized and reduced models, respectively, averaged over the [4Fe4S] clusters in CpI [FeFe] hydrogenase. As compared with experimental values of 2.31 \pm 0.029 for Fe-sulfide and 2.27 \pm 0.066 for Fe-protein ligand bonds in this protein, the bond lengths in the energy-minimized oxidized CpI hydrogenase model offer satisfactory agreement with experimentally measured values. Circumsphere radii for the Fe, sulfide, and protein ligand atoms are shown in Table 2 for the [4Fe4S] clusters found in CpI hydrogenase. Most notably, the Fe circumsphere radii in the oxidized state are contracted by ~ 0.2 Å relative to the experimental structure, and the protein ligand circumsphere by 0.14-0.31 Å. The oxidized sulfide circumsphere matches the experimental structure quite closely, with a small relative ~ 0.02 Å contraction. Save for the [4Fe4S]Cys₃His cluster (HYDD), in silico reduction leads to minimum-energy clusters with a slightly ($\sim 0.1 \text{ Å}$) contracted Fe sphere and expanded sulfide (0.2 Å) and ligand (0.3 Å) spheres. The Fe circumsphere of the His-ligated cluster actually expands upon reduction; the sulfide and protein ligand circumspheres also expand but less so than the all-cysteinate-ligated iron-sulfur clusters. Expansion upon reduction has been noticed previously³⁷ and is evident in our gas-phase energy-optimized structures (Supporting



Figure 6. ORTEP-style 50% ellipsoid plots generated from a 1.0 ns molecular dynamics simulation on the CpI hydrogenase model with all clusters reduced. Left, [2Fe]-[4Fe4S]_H with cysteinate ligands; middle, the [4Fe4S]Cys₃His auxiliary electron transfer cluster; right, the [2Fe2S]Cys4 auxiliary cluster. Selected group labels and bonds have been drawn into the H-cluster diagram where appropriate to assist in orienting the eye.

Information), consistent with a net Fe-Fe antibonding contribution for the "active" electron, i.e., the electron that enters the oxidized or leaves the reduced cluster form. The substantial compression of the Fe circumspheres in the oxidized models relative to both the gas-phase density functional calculations from which the force constants were derived as well as the experimental protein structure points to a significant role of the protein structure, potential limitations to the classical harmonic potential approximation as applied to redox-active metalloclusters, and possibly subtle effects not yet captured in the current force field. The differences in circumsphere distances involved are satisfyingly small, however, considering the underlying comparison being made between the experimental or gas-phase density functional models on the one hand and the classical mechanical bonding model with protein present on the other.

In order to evaluate the performance of the derived parameters in a broader context, we calculated the geometric properties of [4Fe4S] clusters contained in the PDB database. All structures containing the residue name "SF4" were filtered to eliminate structures containing (1) clusters with multidentate or missing external ligands, (2) clusters with questionable structures, (3) clusters with dual refined positions, (4) clusters falling on a crystallographic symmetry element, such that not all atomic coordinates are explicitly defined, and (5) clusters with bond lengths to external ligands longer than 2.7 Å. The resulting 235 crystallographic protein structures contained 635 [4Fe4S] clusters. We expect these structures to be representative of the oxidized member of the relevant redox couple, due to oxidation during crystallization and limited photoreduction during crystallographic data collection, although this latter assumption is uncertain³⁸ and depends on specific data acquisition conditions. Average bond lengths, distances between circumcenters, and circumcenter radii were calculated, and the angular circumcenter coordinates were plotted as histograms, as shown in the Supporting Information (Table S2 and Figures S1 and S2). Average circumcenter radii were very close to the experimental [FeFe] hydrogenase values shown in Table 2 as well as to the sulfide and ligand circumsphere radii calculated for the oxidized hydrogenase model. The circumsphere radius calculated for the oxidized hydrogenase [4Fe4S] clusters is \sim 0.2 Å less than either the DFT-optimized value or the mean experimentally observed value. Angular coordinates of the crystallographic cluster structures cluster around very similar circumsphere theta and phi values to the minimum-energy hydrogenase [4Fe4S] cluster structures shown in Figure 5.

Test Simulations—Dynamics Stability. Figure 6 shows 50% thermal ellipsoid representations for the fully oxidized and reduced models of CpI hydrogenase over 1 ns of molecular dynamics simulation. The Fe and S atoms exhibit narrower distributions than the lighter atoms, as expected. More motion is evident among the second-row atoms of the [2Fe]_H cluster, comparable in magnitude to C/N/O motion in the cysteinate ligands. rms deviations of atomic positions in the protein only, metalloclusters only, and protein + metalloclusters over 1000 frames (1 ns simulation time) were $0.825 \pm 0.007 \text{ Å}$, $0.482 \pm 0.007 \text{ Å}$, and $0.824 \pm 0.007 \text{ Å}$ for the oxidized enzyme and 0.835 ± 0.008 Å, 0.363 ± 0.004 Å, and 0.833 ± 0.008 Å for the reduced enzyme, respectively. The reported parameters thus yield dynamically stable structures over at least 1.0 ns of simulated time, with comparable but more restricted motion than the bulk polypeptide.

Test Simulations—Frequencies. By virtue of possessing triply bonded cyanide and carbonyl ligands bound to the H-cluster, [FeFe] hydrogenases show distinctive infrared spectroscopic absorptions well separated from the spectroscopic region associated with amino acid vibrations. Fourier transform infrared spectra of the [FeFe] hydrogenase in its active oxidized state from Desulfovibrio vulgaris showed bands in its air-exposed form at 1848, 1983, 2008, 2087, and 2106 cm⁻¹, assigned to CO_{bridge} , 2 × $CO_{terminal}$, and 2 × CN_{terminal}, in order of increasing energy.³⁹ Bands seen for the D. desulfuricans enzyme at 1802, 1940, 1965, 2079, and 2093 cm⁻¹ were assigned to C-X stretches in the ligands C5O5, C7O7, C3O3, C6N6, and C4N4, using our atom naming convention (see Figure 3).⁴⁰ Chen, et al. found a similar vibrational manifold for the active oxidized Clostridium pasteurianum [FeFe] hydrogenase, with analogous frequencies at 1802, 1948, 1971, 2072, and 2086 cm⁻¹.41 Notably, the frequencies shift dramatically depending on the redox state of the enzyme and the presence of exogenous $CO^{.39}$

To test the parameter performance in the context of full protein dynamics, we calculated by quasiharmonic analysis the normal modes and frequencies associated with the hydrogenase metalloclusters from the molecular dynamics trajectories. Although the analysis was done only on the metalloclusters and immediately bonded ligand atoms, the

Table 3. Summary of Vibrational Analysis for Isolated Quantum Mechanical H-Cluster Model and Whole-Protein Classical Mechanical Metalloclusters^a

oxidized				reduced			
QH frequency	quasiharmonic description	QM frequency	QM stretch ^b	QH frequency	quasiharmonic description	QM frequency	QM stretch ^b
2271	$\frac{CO_{prox} + CN_{prox}}{CO_{br}(S)} +$	2088	0.9940CN _{prox}	2522	CN _{term}	2111	0.9935CN _{term}
2232	${\color{red}CO_{term}} + {\color{red}CN_{term}}(AS) + {\color{red}CO_{br}}(AS)$	2076	$0.9870\text{CN}_{\text{term}}$	2463	$CO_{br} + Fe-H (S)$	2088	0.9990CN _{prox}
2136	${\stackrel{CO_{prox}}{CO_{br}}} + {\stackrel{CN_{prox}}{CO_{prox}}} (S) + {\stackrel{CO_{br}}{CO_{term}}} (AS)$	1944	0.0659CO _{prox} , 0.0520CO _{br} , 0.8742CO _{term}	2412	$CO_br + Fe-H \; (AS)$	1965	0.0386CO _{prox} , 0.1938CO _{br} , 0.7559CO _{term}
1973	CN_{prox} - $CO_{br}(S)$ - $CO_{term}(S)$	1911	0.8825CO _{prox} , 0.0176CO _{br} , 0.0910CO _{term}	2351	$CO_br + Fe-H \; (AS)$	1927	0.1870CO _{prox} 0.2763CO _{br} , 0.1389CO _{term}
1726	CN _{term} stretch and bend	1825	0.0383CO _{prox} , 0.9265CO _{br} , 0.0321CO _{term}	2115	CO _{term}	1904	0.3954CO _{prox} , 0.1744CO _{br} , 0.0040CO _{term}
				2074 2054 1956	${{\sf CO}_{\sf prox}} \atop {{\sf CO}_{\sf prox}} + {\sf CN}_{\sf prox}$ (AS)		

^a Frequencies are given in cm⁻¹. Two or more terms imply coupled vibration, specified relative to the first listed local stretch as symmetric (S) or antisymmetric (AS). Vibrations are stretches unless otherwise noted. ^b Coefficients for QM stretches are summed squares of Cartesian eigenvector components of C and O/N atoms involved in the mode. If the mode vector is comprised only of diatomic ligand atom motions, these will sum to 1.0000; deviation from 1.0000 reflects motion distributed among other atoms in the H-cluster model.

dynamics trajectory was extracted from that of the entire enzyme. Hence, the nature of the vibrational modes will not necessarily match the quantum mechanical modes of the isolated cluster, to the extent that polypeptide motions couple with the H-cluster or otherwise lead to mixing of localized C-X stretches in the normal modes. Table 3 summarizes results from both density functional theory calculations on the isolated cluster model and those from quasiharmonic analysis of the dynamics trajectories. Frequencies from the quantum mechanical calculations on the isolated H-cluster show isolated C-N stretches but mixing among the possible C-O stretches with more mixing evident in the reduced H-cluster model. None of the classical modes shows predominantly C-X stretching motions alone-DTMA bridge motions in both H-cluster redox forms and Fe-hydride stretches in the reduced system coupled noticeably with the localized C-X stretches. In addition, C-X stretches coupled among themselves (see the Supporting Information for animations of vibrational modes). The observed mixing in the classical dynamics may reflect the effects of the protein environment on the metallocluster vibrations, particularly electrostatic interactions (e.g., CN_{prox} H-bonding to the Ser232 backbone amide N-H and the CN_{term} Lys358 chargedipole interaction). The quasiharmonic modes containing significant C-X stretching show frequency ranges for the oxidized and reduced models that bracket the DFT-calculated values for the isolated cluster. Quantitative agreement is not expected at this stage of parameter development, and careful refinement of nonbonded parameters, an account of the protein environment's effect on hydrogen bonding and electronic anisotropy of the cluster⁴² via more expensive QM/ MM calculations, and possibly a more explicit treatment of static and dynamic electron correlation should be necessary to achieve such agreement. Nevertheless, the data presented support the utility of the calculated parameters for molecular

dynamics simulations of known [FeFe] hydrogenase enzymes and related proteins with similar metallocluster species.

Conclusions

We have presented a set of molecular mechanical parameters relevant to [2Fe2S]^{2+,+}, [4Fe4S]^{2+,+}Cys₄, [4Fe4S]^{2+,+}Cys₃His, and [2Fe]_H^{II,I;I,I} metalloclusters found in known [FeFe] hydrogenase enzymes. Modeled minimum-energy hydrogenase structures are consistent with those found experimentally, and the cluster dynamics are stable, while still permitting as much flexibility as is allowed by the quantum mechanical force constants. Calculated vibrational frequencies associated with the catalytic [2Fe]_H CO and CN ligands agree semiquantitatively with those measured experimentally and calculated with density functional theory. It is our hope that the consistent derivation procedure for all four cluster types in both oxidized and reduced states will permit both high-quality simulations of hydrogenase molecular dynamics, particularly of the protein environment immediately around the clusters, as well as allow systematic improvement of these parameters by the modeling community should shortcomings be found. It is expected that the pragmatic interest in alternative fuels combined with fundamental scientific questions of electron transfer and proton dynamics and reduction in hydrogenases will benefit from the availability of these parameters.

Acknowledgment. The work is supported by U.S. Department of Energy under Contract No. DE-AC36-99GO10337. Calculations were done on the "Lester" cluster at the National Renewable Energy Laboratory and the "Jacquard" cluster at the National Energy Research Supercomputing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Supporting Information Available: CHARMM-format topology files for cluster reconstruction and system

setup, CHARMM-format parameter files for simulations, an example of a VMD/PSFGen build script, a table of circumsphere radii for DFT-optimized model clusters, summary data for PDB [4Fe4S] cluster survey, and trajectory (32-bit binary DCD) and structure (text PSF) files for visualizing H-cluster diatomic vibrational modes. This material is available free of charge via the Internet at http://pubs.acs.org.

References

- (1) Melis, A.; Seibert, M.; Ghirardi, M. L. Hydrogen Fuel Production by Transgenic Microalgae. In *Transgenic Microalgae as Green Cell Factories*; Leon, R., Gavan, A., Fernandez, E., Eds.; Landes Bioscience: 2007.
- (2) Vignais, P. M.; Billoud, B. Chem. Rev. 2007, 107, 4206– 4272.
- (3) van Haaster, D. J.; Hagedoorn, P.-L.; Jongejan, J. A.; Hagen, W. R. Biochem. Soc. Trans. 2005, 33, 12–14.
- (4) Jones, A. K.; Lichtenstein, B. R.; Dutta, A.; Gordon, G.; Dutton, P. L. J. Am. Chem. Soc. 2007, 129, 14844–14845.
- (5) Cao, Z.; Hall, M. B. J. Am. Chem. Soc. 2001, 123, 3734–3742.
- (6) Liu, Z.-P.; Hu, P. J. Am. Chem. Soc. 2002, 124, 5175-5182.
- (7) Zhou, T.; Mo, Y.; Liu, A.; Zhou, Z.; Tsai, K. R. *Inorg. Chem.* **2004**, *43*, 923–930.
- (8) Peters, J. W.; Lanzilotta, W. N.; Lemon, B. J.; Seefeldt, L. C. Science 1998, 282, 1853–1858.
- (9) Nicolet, Y.; Piras, C.; Legrand, P.; Hatchikian, C. E.; Fontecilla-Camps, J. C. *Structure* **1999**, *7*, 13–23.
- (10) Zhou, T.; Mo, Y.; Zhou, Z.; Tsai, K. Inorg. Chem. 2005, 44, 4941–4946.
- (11) Fan, H.-J.; Hall, M. B. J. Am. Chem. Soc. **2001**, 123, 3828–3829.
- (12) Greco, C.; Bruschi, M.; De Gioia, L.; Ryde, U. *Inorg. Chem.* 2007, 46, 5911–5921.
- (13) Weinhold, F. Natural Bond Orbital Methods. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F. I., Schreiner, P. R., Eds.; Wiley Interscience: New York, 1998; Vol. 3, pp 1792–1813.
- (14) Scott, A. P.; Radom, L. J. Phys. Chem. 1996, 100, 16502– 16513.
- (15) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J. Phys. Chem. B 1998, 102, 3586–3616.
- (16) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. J. Comput. Chem. 2005, 26, 1781–1802.
- (17) Humphrey, W.; Dalke, A.; Schulten, K. J. Mol. Graph. 1996, 14, 33–38.
- (18) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J. Comput. Chem. 1983, 4, 187–217.

- (19) Karplus, M.; Kushick, J. N. Macromolecules 1981, 14, 325– 332.
- (20) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*; Dover Publications: New York, 1980.
- (21) Chang, C. H.; King, P. W.; Ghirardi, M. L.; Kim, K. Biophys. J. 2007, 93, 3034–3045.
- (22) Pivarski, J. *SVGFig, 1.1.6*; Google Code: 2009. http://code.google.com/p/svgfig/ (accessed March 2009).
- (23) Burden, C. J.; Oakley, A. J. Phys. Biol. 2007, 4, 79-90.
- (24) Merritt, E. A.; Bacon, D. J. Methods Enzymol. 1997, 277, 505–524.
- (25) Nicolet, Y.; de Lacey, A. L.; Vernède, X.; Fernandez, V. M.; Hatchikian, E. C.; Fontecilla-Camps, J. C. *J. Am. Chem. Soc.* 2001, 123, 1596–1601.
- (26) Pandey, A. S.; Harris, T. V.; Giles, L. J.; Peters, J. W.; Szilagyi, R. K. J. Am. Chem. Soc. 2008, 130, 4533–4540.
- (27) Patil, D. S.; Moura, J. J. G.; He, S. H.; Teixeira, M.; Prickril,
 B. C.; DerVartanian, D. V.; Peck, H. D., Jr.; LeGall, J.; Huynh,
 B. H. J. Biol. Chem. 1988, 263, 18732–18738.
- (28) Pierik, A. J.; Hagen, W. R.; Redeker, J. S.; Wolbert, R. B. G.; Boersma, M.; Verhagen, M. F. J. M.; Grande, H. J.; Veeger, C.; Mutsaers, P. H. A.; Sands, R. H.; Dunham, W. R. *Eur. J. Biochem.* **1992**, 209, 63–72.
- (29) Fiedler, A. T.; Brunold, T. C. Inorg. Chem. 2005, 44, 9322–9334.
- (30) Schwab, D. E.; Tard, C.; Brecht, E.; Peters, J. W.; Pickett, C. J.; Szilagyi, R. K. Chem. Commun. 2006, 3696–3698.
- (31) Silakov, A.; Reijerse, E. J.; Albracht, S. P. J.; Hatchikian, E. C.; Lubitz, W. J. Am. Chem. Soc. 2007, 129, 11447– 11458.
- (32) Stowasser, R.; Hoffmann, R. J. Am. Chem. Soc. 1999, 121, 3414–3420.
- (33) Mantz, Y. A.; Gervasio, F. L.; Laino, T.; Parrinello, M. J. Phys. Chem. A 2007, 111, 105–112.
- (34) Kuleff, A. I.; Dreuw, A. J. Chem. Phys. 2009, 130, 034102.
- (35) Roseboom, W.; De Lacey, A. L.; Fernandez, V. M.; Hatchikian, E. C.; Albracht, S. P. J. J. Biol. Inorg. Chem. 2005, 11, 102–118.
- (36) Fee, J. A.; Castagnetto, J. M.; Case, D. A.; Noodleman, L.; Stout, C. D.; Torres, R. A. J. Biol. Inorg. Chem. 2003, 8, 519–526.
- (37) Torres, R. A.; Lovell, T.; Noodleman, L.; Case, D. A. J. Am. Chem. Soc. 2003, 125, 1923–1936.
- (38) Carugo, O.; Carugo, K. D. *Trends Biochem. Sci.* **2005**, *30*, 213–219.
- (39) Pierik, A. J.; Hulstein, M.; Hagen, W. R.; Albracht, S. P. Eur. J. Biochem. 1998, 258, 572–578.
- (40) De Lacey, A. L.; Stadler, C.; Cavazza, C.; Hatchikian, E. C.; Fernandez, V. M. J. Am. Chem. Soc. 2000, 122, 11232– 11233.
- (41) Chen, Z.; Lemon, B. J.; Huang, S.; Swartz, D. J.; Peters, J. W.; Bagley, K. A. *Biochemistry* **2002**, *41*, 2036–2043.
- (42) Dey, A.; Roche, C. L.; Walters, M. A.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *Inorg. Chem.* 2005, 44, 8349–8354.

CT800342W