J|A|C|S
ARTICLES
Published on Web 12/13/2002

Probing the Configurational Space of a Metalloprotein Core: An ab Initio Molecular Dynamics Study of Duo Ferro 1 Binuclear Zn Cofactor

Garegin A. Papoian,*,† William F. DeGrado,‡ and Michael L. Klein§

Contribution from the Department of Chemistry & Biochemistry, University of California at San Diego, 9500 Gilman Dr., La Jolla, California 92093-0371, Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6059, Center for Molecular Modeling & Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323

Received August 15, 2002; E-mail: gap5@cornell.edu

Abstract: We present three theoretical models of various degree of completeness to explore the chemical phase space available to the $Glu_4His_2Zn_2$ cofactor found in the four-helix bundle of de novo designed metalloprotein Duo Ferro 1. We have found that the planewave DFT geometry optimization of 94-atom Model I, which contains both the protein scaffold constraints as well as the second shell hydrogen bonding network, reproduces the crystal structure bonding with remarkable accuracy (0.34 Å). Surprisingly, the geometry optimization of 66-atom Model II (lacking the second shell hydrogen bonding) and 48-atom Model III (being also free of the protein scaffold constraints) still result in the fidelity with the crystallographic structure (RMSDs 0.29 and 0.34 Å, respectively). To examine whether these structures are close to the global minimum as well as to investigate various conformational transitions to which the di-Zn cofactor may be susceptible to, we have carried out a 10 ps Car−Parrinello Molecular Dynamics (CPMD) simulation of Model III. We suggest that weak hydrogen bonds between imidazole hydrogens and carboxylate oxygens modulate the dynamical behavior of the system. One part of the molecule was found to be rigid due to the particular H(imidazole)−O(carboxylate) interaction restricting both the motion of the imidazole ring as well as the terminal carboxylate conformational mobility. The second half of the system was very flexible demonstrating a coupling of a transient formation of H(imidazole)−O(carboxylate) bonds with the spinning of the imidazole ring and syn-anti isomerization of the terminal carboxylate group. In addition, two low-energy snapshots from the 10 ps CPMD run were quenched, and their geometries were optimized, leading to two new isomers 48 kJ/mol lower in energy than the one associated with the crystal structure. We suggest that periodic quenching of the CPMD simulation snapshots of a minimalist model may be used as an efficient method to generate a large number of competitive local minima, which may be consequently pruned by imposing the protein scaffold constraints as well as further tuned by the second shell hydrogen bonding network.

1. Introduction

Dimetal cofactors are found in a wide variety of proteins that catalyze hydrolytic reactions. For example, Di-Zn(II), di-Mn(II), di-Fe proteins are found in proteases,[1,2] arginases,[3,4] phosphatases,[5−7] and a major class of β-lactamases.[8,9] Di-Mn and di-Fe are also found in proteins that conduct oxygen-dependent chemistry, including methane monooxygenase (MMO),[10−14] the R2 subunit of ribonucleotide reductase (R2),[15] and Mn catalase.[16−18] The structures of the metal-binding sites of these proteins reflect a delicate interplay between often opposing requirements for tight binding of the cofactor, which drives toward static, coordinately saturated geometries, versus

† Department of Chemistry & Biochemistry, University of California at San Diego.
‡ Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine.
§ Center for Molecular Modeling & Department of Chemistry, University of Pennsylvania.

(1) D'Souza, V. M.; Holz, R. C. *Biochemistry* **1999**, *38*, 11 079−11 085.
(2) Stamper, C.; Bennett, B.; Edwards, T.; Holz, R. C.; Ringe, D.; Petsko, C. *Biochemistry* **2001**, *40*, 7035−7046.
(3) Christianson, D. W.; Cox, J. D. *Annu. Rev. Biochem.* **1999**, *68*, 33−57.
(4) Ash, D. E.; Cox, J. D.; Christianson, D. W. *Met. Ions. Biol. Syst.* **2000**, *37*, 407−428.
(5) Kimura, E. *Curr. Opin. Chem. Biol.* **2000**, *4*, 207−213.
(6) Knofel, T.; Strater, N. *J. Mol. Biol.* **2001**, *309*, 239−254.
(7) Lindqvist, Y.; Johansson, E.; Kaija, H.; Vihko, P.; Schneider, G. *J. Mol. Biol.* **1999**, *291*, 135−147.
(8) Wang, Z.; Fast, W.; Benkovic, S. J. *Biochemistry* **1999**, *38*, 10 013−10 023.
(9) Wang, Z.; Fast, W.; Valentine, A. M.; Benkovic, S. J. *Curr. Opin. Chem. Biol.* **1999**, *3*, 614−622.
(10) Merkx, M.; Kopp, D. A.; Sazinsky, M. H.; Blazyk, J. L.; Muller, J.; Lippard, S. J. *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 2782−2807.
(11) Lange, S. J.; Que, L., Jr. *Curr. Opin. Chem. Biol.* **1998**, *2*, 159−172.
(12) Nordlund, P.; Eklund, H. *Curr. Opin. Struct. Biol.* **1995**, *5*, 758−766.
(13) Waller, B. J.; D Lipscomb, J. *Chem. Rev.* **1996**, *96*, 2625−2657.
(14) Solomon, E. I. *Inorg. Chem.* **2001**, *40*, 3656−3669.
(15) Stubbe, J.; Riggs-Gelasco, P. *Trends. Biochem. Sci.* **1998**, *23*, 438−443.
(16) Ananyev, G. M.; Zaltsman, L.; Vasko, C.; Dismukes, G. C. *Biochim. Biophys. Acta.* **2001**, *1503*, 52−68.
(17) Barynin, V. V.; Whittaker, M. M.; Antonyuk, S. V.; Lamzin, V. S.; Harrison, P. M.; Artymiuk, P. J.; Whittaker, J. W. *Structure (Camb.)* **2001**, *9*, 725−738.
(18) Christianson, D. W. *Prog. Biophys. Mol. Biol.* **1997**, *67*, 217−252.

10.1021/ja028161l CCC: $25.00

function, which requires the cofactor to be poised to ligate substrates and undergo dynamic changes in configuration as the chemical reaction proceeds.

One of the most important questions concerning the activity of dimetal proteins is the extent to which the structures of their active sites reflect a low-energy geometric arrangement of the ligands, versus being strained to enhance their reactivity. To help address this issue, Lombardi and co-workers designed a model protein that reflect the minimal features required for folding and binding the dimetal cofactor.[19−21] This protein, DF1, was designed to idealize the diiron site observed in the structure of MMO and R2, and contains four Glu and two His residues as primary ligands, as well as a number of second-shell ligands.[20] The active site geometry of di-Zn(II) DF1 and the di-Mn(II) form of mutant of DF1 have been crystallographically characterized; both show a close resemblance to the diferrous and dimanganous forms of natural diiron proteins, indicating that DF1 provides a faithful geometric model for this class of proteins.

In light of the biological importance of four-helix bundle−dimetal core catalytic systems, understanding in better detail the potential energy surface (PES) for the dimetal cofactor might lead to more refined design of small protein models, as well as to additional insight into the mechanisms of various catalytic reactions. The steady advance of computational power has allowed modeling of dimetal cofactors with an ever increasing degree of accuracy.[22,23] Among the most large-scale calculations, Lippard, Friesner, and co-workers have carried out Density Functional Theory (DFT) modeling of the MMO Fe_2 active site using an approximately 100-atom description.[24−26] Miller, Lippard, Friesner, and co-workers have investigated the hydroxylation of alkanes by the MMO cofactor by reducing its full potential energy surface to a four-dimensional one, using certain chosen coordinates, and further carrying out semiclassical dynamics of the hydrogen abstraction on the reduced surface.[27] In a series of papers, Morokuma and co-workers have conducted detailed computational studies of the MMO catalytic cycle using small-size cofactor models.[28−32] The importance of electrostatic interactions between the cofactor and the host protein as well as the significance of magnetic interactions between metal atoms have been extensively studied as well.[33−36] A combination of quantum mechanical minimization and classical Molecular Dynamics techniques have been employed by Merz, Jr and co-workers to study the Zn_2 cofactor in β-lactamase.[37,38]

Although the ab initio works cited above have brought tremendous understanding to many aspects of the catalytic activity of dimetal cofactors in the MMO class of enzymes, they have also been largely confined to static geometry optimization of near-native local minima and transition state structures. When employing usual quantum chemical geometry optimization techniques, it is extremely hard to climb over barriers and explore larger regions of the *relevant* conformational phase space accessible to many-atom molecules. Quenching snapshots in classical Molecular Dynamics would appear as an appealing alternative; however, the presence of transition metal atoms as well as bond making/breaking during isomerization would preclude accurate determination of alternative structures and associated energies.

Here, we use ab initio Molecular Dynamics calculations to determine the extent to which the protein structure restricts the geometry of the di-Zn(II) site of DF1 as well as to investigate cofactor's dynamical behavior. A series of successively simplified models are used, beginning with a complex model that considers both second shell ligands as well as the structural restraints of the protein backbone. At the other extreme, a model consisting only of unconstrained primary ligands is considered. These calculations show that the active site geometry of di-Zn(II) DF1 represents a local energy minimum whose geometry depends on neither the tethering of the ligands to the polypeptide backbone nor the presence of second shell ligands. However, a 10 ps dynamics calculation of the most minimal model revealed a number of lower-lying energy minima. These minima are not accessible to the native protein because of the restraints of the polypeptide backbone and second-shell ligands. These calculations also indicated that weak hydrogen bonding between imidazole CH groups and carboxylate oxygens helps stabilize specific conformations of the complexes, modulating in turn various longer-time scale cooperative molecular motions.

2. Computational Details

Car−Parrinello Molecular Dynamics[39] simulations were carried out both for the structure minimizations (0 K annealing; followed by the Steepest Descent minimization in some cases) as well as dynamical studies. Norm-conserving Martins−Troullier pseudopotentials were used to describe H, C, N, O, and Zn atoms.[40] The valence electrons for Zn atoms also included the 10 d electrons, a feature which was important for achieving the better convergence of the electronic structure optimizations and more faithful geometrical description of the metallocore complex. The Kleinman−Bylander decomposition was carried out for the nonlocal part of the pseudopotentials,[41] except the Gauss−Hermite decomposition was applied to Zn.[42]

To describe exchange and correlation effects, Becke, Lee, Yang, and Par (BLYP) GGA exchange and correlation func-

(19) Summa, C. M.; Lombardi, A.; Lewis, M.; DeGrado, W. F. *Current Opinion in Structural Biology* **1999**, *9*, 500.
(20) Lombardi, A.; Summa, C. M.; Geremia, S.; Randaccio, L.; Pavone, V.; Degrado, W. F. *Proc. Nat. Acad. Sci.* **2000**, *97*, 6298.
(21) Costanzo, L. D.; Wade, H.; Geremia, S.; Randaccio, L.; Pavone, V.; Degrado, W. F.; Lombardi, A. *J. Am. Chem. Soc.* **2001**, *123*, 12 749.
(22) Siegbahn, P. E. M. *Inorg. Chem.* **1999**, *38*, 2880.
(23) Siegbahn, P. E. M. *Curr. Opin. Chem. Biol.* **2002**, *6*, 227.
(24) Dunietz, B. D.; Beachy, M. B.; Cao, Y.; Whittington, D. A.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2000**, *122*, 2828.
(25) Gherman, B. F.; Dunietz, B. D.; Whittington, D. A.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2001**, *123*, 3836−3837.
(26) Guallar, V.; Gherman, B. F.; Lippard, S. J.; Friesner, R. A. *Curr. Opin. Chem. Biol.* **2002**, *6*, 236−242.
(27) Guallar, V.; Gherman, B. F.; Miller, W. H.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2002**, *124*, 3377−3384.
(28) Basch, H.; Mogi, K.; Musaev, D. G.; Morokuma, K. *J. Am. Chem. Soc.* **1999**, *121*, 7249.
(29) Basch, H.; Musaev, D. G.; Morokuma, K. *J. Phys. Chem. B* **2001**, *105*, 8452.
(30) Torrent, M.; Vreven, T.; Musaev, D. G.; Morokuma, K.; Farkas, O.; Schlegel, H. B. *J. Am. Chem. Soc.* **2002**, *124*, 192.
(31) Musaev, D. G.; Basch, H.; Morokuma, K. *J. Am. Chem. Soc.* **2002**, *124*, 4135.
(32) Torrent, M.; Musaev, D. G.; Basch, H.; Morokuma, K. *J. Comput. Chem.* **2002**, *23*, 59.
(33) Lovell, T.; Li, J.; Noodleman, L. *Inorg. Chem.* **2001**, *40*, 5251−5266.
(34) Lovell, T.; Li, J.; Noodleman, L. *Inorg. Chem.* **2001**, *40*, 5267−5278.
(35) Lovell, T.; Han, W. G.; Liu, T.; Noodleman, L. *J. Am. Chem. Soc.* **2002**, *124*, 5890−5894.
(36) Han, W. G.; Lovell, T.; Noodleman, L. *Inorg. Chem.* **2002**, *41*, 205−218.
(37) Diaz, N.; Suarez, D.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **2001**, *123*, 9867−9879.
(38) Suarez, D.; Brothers, E. N.; Merz, K. M., Jr. *Biochemistry* **2002**, *41*, 6615−6630.
(39) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
(40) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993.
(41) Kleinman, L.; Bylander, D. M. *Phys. Rev. Lett.* **1982**, *48*, 1425.
(42) Stich, I.; Car, R.; Parrinello, M.; Baroni, S. *Phys. Rev. B* **1989**, *39*, 4997.

tionals were employed.[43,44] The electronic wave functions were expanded in a plane-wave basis to 70 Ry energy cutoff. A fictitious mass of 800 au was assigned to electrons to couple the electronic and nuclear degrees of freedom.[39]

A cubic cell with a side of 20 Å was used for the largest Model I. An orthorhombic cell was employed for the smaller two models considered, the cell dimensions being 18.9, 14.2, and 15.2 Å for Model II and 17.3, 14.3, and 15.0 Å for Model III. These lattice parameters were chosen by adding 6 Å to the corresponding model molecule span in each of the three dimensions.

The time step of integration was chosen at 0.145 fs. For annealing runs, the system was initially prepared at 0 K, followed by an MD run, with 5−25% of the kinetic energy being discarded at every step. For the Molecular Dynamics study of the smaller 48-atom model, an equilibration run for 2 ps was followed by a 10 ps run at 300 K. The temperature was controlled by rescaling the velocities so to keep the system within a 40 K tolerance window.

All calculations were carried out with the CPMD program.[45]

3. Results and Discussion.

The four-helix bundle Glu_4His_2 binding motif housing simultaneously two metal ions is a recurring and very important theme in structural biology. Given the enormous size of the functionally active proteins, such as MMO, as well as various shortcomings associated with small inorganic models, there exists a niche for small protein frameworks which would faithfully describe the structural details and the catalytic activity of the larger enzymes. An important advantage of these models would be the much improved ability to conduct detailed structural, mechanistic, and catalytic studies on them, helping us better understand many important classes of enzymatic reactions.

DeGrado and co-workers have de novo designed, synthesized, and structurally characterized two such model metalloproteins−Duo Ferro 1 (DF1) and Duo Ferro 2 (DF2).[20,21] A number of crucial factors that went into the design of DF1 (PDB code 1EC5) are the following: (1) an approximate D_2 symmetry was built into the model, as inferred from the structural analysis of biologically relevant proteins;[19] (2) the Glu-Xxx-Xxx-His motif was included to ligate the metal ions; (3) satisfaction of side chain packing requirements was ensured; (4) the hydrophobic and polar residues were carefully placed so to greatly stabilize the native state. As designed, DF1, consisting of two helix-loop-helix monomers (21 residues long), associates with non-covalent interactions into a stable dimer even in the absence of metal cofactors with a free energy stabilization of −12.8 kcal/mol.[20]

DF1 was initially crystallized as the di-Zn(II) complex, which is the object of the current study. The protein bound the metal ions in the expected manner (Figures 1 and 2) with two Glu carboxylates bridging in a μ-1,3-syn,syn manner (consult Figure 3 for the notation), and two terminal Glu carboxylates bound to individual Zn(II) ions in a chelating bidentate coordination. Finally, two His side chains complete the coordination sphere

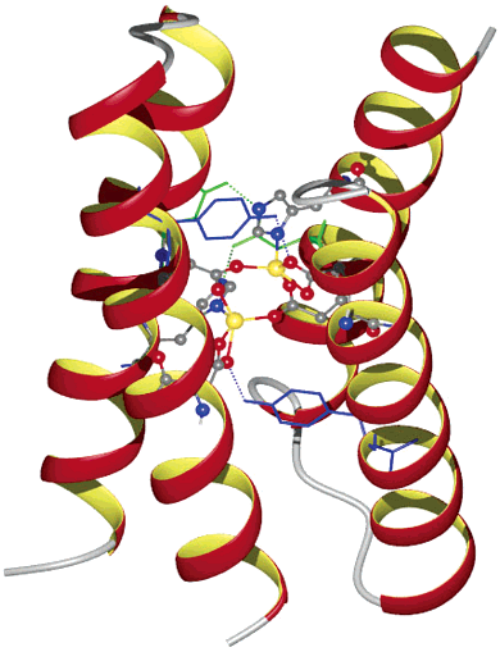

Figure 1. Four-helix bundle hosting the di-zinc core in the artificial Due Ferro 1 protein (asymmetric structure). Two Tyr residues forming H-bonds to Glu side chains are shown in blue. Two Asp residues forming H-bonds with His are shown in green. Zn: large golden spheres. N: blue spheres. O: red spheres. C: gray spheres.

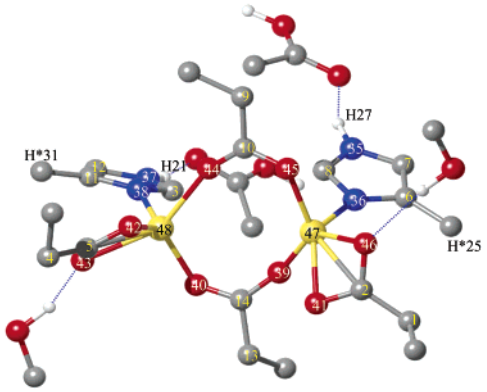

Figure 2. The di-zinc core of the asymmetric four-helix bundle. Imidazole methylene carbons are denoted as H* in the numbering scheme to indicate the forthcoming substitution in the smallest computational model.

around each Zn(II) ion, rendering the Zn(II) ions pent-coordinate. Second shell ligands were also included to stabilize metal-binding site[46] and allow the modulation of the electrochemical midpoint potential and chemical reactivity of the metal ions.[47−50] These second shell ligands include an Asp residue, which receives a hydrogen bond from the His ligand, and a Tyr, which donates a hydrogen bond to the terminal carboxylate.

The di-Mn(II) complex of a single-site variant of DF1 with a more solvent-accessible active site has also been crystallographically characterized. The arrangement of primary and secondary ligands in this protein is very similar to di-Zn(II) DF1, although a bridging solvent molecule is found between

(43) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(44) Lee, C.; Yang, W.; Parr, R. C. *Phys. Rev. B* **1988**, *37*, 785.
(45) Hutter, J.; Alavi, A.; Deutch, T.; Bernasconi, M.; Goedecker, S.; Marx, D.; Tuckerman, M.; Parrinello, M. CPMD, MPI für Festkörperforschung and IBM Zurich Research Laboratory, 1995−1997.
(46) Christianson, D. W.; A Fierke, C. *Acc. Chem. Res.* **1996**, *29*, 331−339.
(47) Schwartz, A. L.; Yikilmaz, C. K.; Vance, S.; Vathyam, R. L.; F Miller, A. *J. Inorg. Biochem.* **2000**, *XXX*, 247−256.
(48) Vance, C. K.; Miller, A.-F. *J. Am. Chem. Soc.* **1998**, *120*, 461−467.
(49) Valentine, J. S.; Sheridan, R. P.; Allen, L. C.; Kahn, P. C. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1009−1013.
(50) Gooden, D. B.; McRee, D. E. *Biochemistry* **1993**, *32*, 3313−3324.

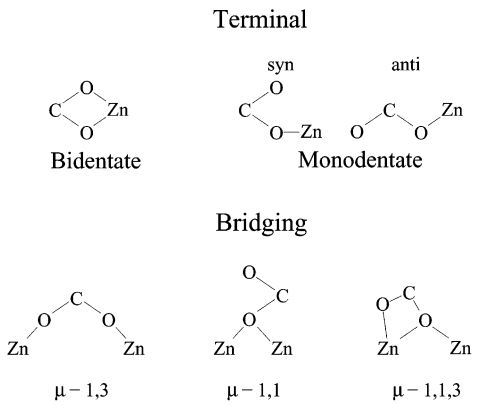

Figure 3. Explanation for the notation associated with various modes of carboxylate bonding.

the two Mn(II) ions. Here, we focus on the di-Zn(II) complex, because of the relative ease of handling this d^{10} metal ion, and also because crystallographic data were available only for the di-zinc protein at the time that we initiated these studies.

Because much or the following discussion entails a comparison of the crystallographic and the computed active site structures, it is important to consider the resolution of the crystallographic data. The structure was solved to only modest resolution (2.5 Å, $R_{conv} = 0.24$ and $R_{free} = 0.30$), which introduces significant coordinate error. For example, in the reported model the terminal carboxylates bind in a relatively symmetrical, chelating manner. However, it is also possible that the carboxylates are dynamically or statically averaging between two different conformations in which one carboxylate oxygen is at a significantly greater distance than the other from the Zn(II). Also, the protein crystallized in a unit cell with two distinct dimers. Two of the monomers form a dimer that lacks crystallographic symmetry (this *asymmetric* dimer is the main reference structure for this paper) while a third monomer forms a crystallographic dimer (the *symmetrical* dimer). The two dimers show minor but significant structural differences (RMSD = 0.51 Å for the Zn(II) and active site ligands), the most significant of which is the angle of the planes formed by the two bridging carboxylate ligands. In the asymmetric dimer the planes are nearly parallel with a crossing angle of approximately 15°, whereas in the asymmetric dimer these planes cross at an angle of approximately −58°. Similar structural differences have frequently been observed between the individual metal centers of different monomers in the unit cells of natural diiron proteins. Presumably, this structural heterogeneity reflects the flexibility of the metal sites, which is indeed necessary to allow dynamic changes in geometry that occur as a part of the reaction cycle.

3.1 Choosing a Model. Ideally, one would like to treat the whole protein with a high level (DFT, MP2) large basis set quantum mechanical methods, but this is going to be out of reach for quite some time. Therefore, the choice of model to represent the active site is crucial. With the current generation of computer power calculations of models with nearly 100 atoms are becoming feasible, as demonstrated by Lippard and Friesner and co-workers[24] and by Parrinello and co-workers.[51,52] However, these calculations still require many months of supercomputer time, thus prompting an evaluation of more minimal models for each particular system. In the current work, we have built and compared three models of various degrees of completeness: (1) a 94-atom model which consists of complete amino acid residue side chains as well as second shell hydrogen bonding network (Model I); (2) a 66-atom model which is lacking the second shell hydrogen bonding network but otherwise is similar to the previous model (Model II); (3) a 48-atom minimalist model where Glu side chains are replaced by acetates and His side chains by imidazoles and no second shell hydrogen bonding ligands are present (Model III).

We believe that Model I (see Figure 4) provides a quite complete description of the di-zinc active site. By replacing the side chain C_β methylene groups with methyl groups, and pinning the positions of the latter carbons to their crystallographic values, we have essentially reproduced the steric constraints imposed by the protein scaffold on the dimetal core. In addition, we have replaced second-shell Tyr and Asp hydrogen bonding side chains by methanol and acetic acid (capping singly the latter carboxylates by protons to preserve charge neutrality), thus faithfully representing the most important effects of the protein environment on the di-Zn cofactor. To even more accurately represent the system, we envision adding the remaining four-helix bundle residues in a combined QM/MM treatment, which is work in progress.

By alternating the 0 K annealing using Car−Parrinello Molecular Dynamics procedure with planewave DFT geometry optimization, the forces on atoms in the Model I (see Figure 4 for superposition with the crystal structure) were brought to less than 0.05 eV/Å. The RMSD deviation between two structures is only 0.34 Å indicating that our model reproduces the detailed geometry of the di-zinc core. The majority of bond lengths and angles are quite similar between the two structures. For instance, the Zn−Zn distance is calculated to be 3.88 Å compared with 3.91 Å in DF1 crystal structure (the modulation of a similar Fe−Fe distance is thought to play an important role in the catalytic activity of the di-iron containing enzymes). Two hydrogen bonds from second shell methanol H to terminal carboxylate O (1.80 Å, 1.95 Å) as well as two additional hydrogen bonds from from second shell acetic acid O to His ϵ-H (1.79 Å, 1.88 Å) help to stabilize the structure, as expected.

Although the calculated and the crystal structures are almost superimposable, there are a few interesting differences. Left terminal carboxylate (TC1) is found to be in the bidentate coordination, although somewhat distorted, compared with nearly ideal bidentate coordination in the crystal structure. Right terminal carboxylate (TC2), on the other hand, underwent a transition to the monodentate-syn mode (the calculated Zn−O distance are 1.96 and 2.60 Å as opposed to 2.05 and 2.13 Å in the crystal structure). The simultaneous occurrence of bidentate TC1 and monodentate TC2 is indicative of a soft potential energy surface for *carboxylic shifts*, as discussed in earlier works.[53,54] Finally, bridging carboxylates are found in a slightly twisted conformation, compared with the crystal structure.

The apparent success of the 94-atom model in describing the structural details of the di-Zn cofactor should be taken with some caution. Given the soft potential energy surface (to be discussed below) for the Zn-ligand bonding, the geometry optimization

(51) Rovira, C.; Parrinello, M. *Chem. Eur. J.* **1999**, *5*, 250.
(52) Rovira, C.; Carloni, P.; Parrinello, M. *J. Phys. Chem. B* **1999**, *103*, 7031.
(53) Rardin, R. L.; Tolman, W. B.; Lippard, S. J. *New J. Chem.* **1991**, *15*, 417−430.
(54) Ryde, U. *Biophys. J.* **1999**, *77*, 2777.

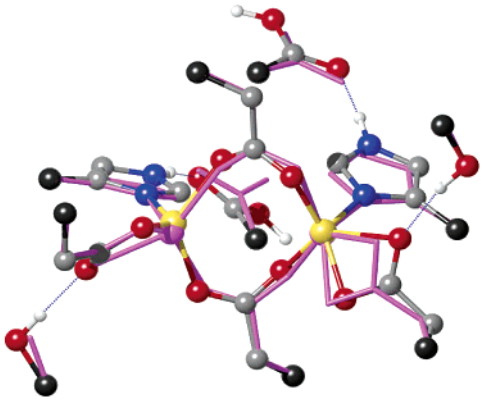

Figure 4. Superposition of the original crystal structure (magenta sticks) with the fully optimized structure. Carbon atoms constrained to their original positions during the geometry optimizations of Models I and II are shown as black spheres. Zn: large golden spheres. H-bonds are indicated by broken blue lines. N: blue spheres. O: red spheres. C: gray spheres. H: white small spheres.

of Model I took an unusually long time to converge (a similar experience was reported by Ryde[54] for other Zn-carboxylate complexes). Thus, we expect the complete exploration of its energy landscape to search other local minima to be a difficult undertaking. Therefore, we are not certain whether the optimized structure of Model I (Figure 4) is the global minimum of the potential energy surface. However, we can make some progress in answering this question by utilizing smaller models, that still faithfully represent the intrinsic chemistry of the coordination sphere around Zn ions. A multitude of local minima competing in energy may be obtained with these models, thus enabling us to chart much wider regions of the cofactor's conformational phase space. This is the task which we take up next.

3.2 Examining the Viability of Smaller Models. The geometrical optimization of Model II led to essentially the same structure as Model I for the 66 atoms represented. The RMSD between the Model II optimized structure and the crystal structure slightly improved from 0.34 to 0.29 Å; however, this change is not significant given the limited resolution of the crystal structure as well as the limited accuracy of theoretical methods employed. Two terminal carboxylates were found in the same mixed bidentate TC1 and monodentate TC2 modes. This calculation suggests that second-shell hydrogen bonding does not effect the conformational details for the particular local minimum to which the crystal structure belongs. However, one would certainly expect changes in the relative ordering in energy among various competing local minima.

Next, in the smallest model, Model III, we have replaced Glu and His side chains by acetates and imidazoles respectively, reducing the number of atoms to 48 (see Figure 5a). We believe that this is the most basic model which reflects the intrinsic chemical character of the Zn_2 cofactor. For instance, it is not clear whether a molecule where imidazoles are further replaced by ammonia and acetates by formates would be very relevant to the study of the parent dimetal core (see ref 24 for the further discussion).

In contrast to the previous models, no constraints were imposed on the positions of any atoms in Model III, thus allowing a full conformational flexibility for the molecule. This is an important advantage of the minimalist model; it provides for a much better exploration of the intrinsic energy landscape

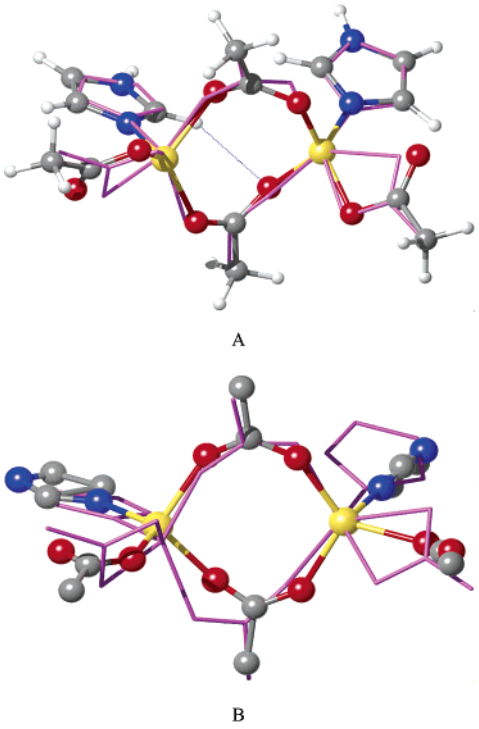

Figure 5. Minimal 48-atom model of the Zn core. (A) A superposition of the original crystal structure (magenta sticks) with the fully optimized model structure. (B) A superposition of the original crystal structure (magenta sticks) with the average structure from the 10 ps CPMD run.

Table 1. Relative Energies of the Various Local Minima Structures for the Minimalist Model. the Average Energy of the CPMD Simulation Was Subtracted from Absolute Energies

	A	C	D
energy (kJ/mol)	−122.4	−170.1	−170.4

of the underlying metal−ligand system, which in turn may suggest the specific roles played by the protein scaffold and second shell hydrogen bonding play in shaping the final geometry of the cofactor.

In addition to the optimized structure A (Figure 5a) which was reached from the initial crystal structure coordinates, two additional local minima were found by quenching from a subsequent Car-Parrinello Molecular Dynamics run (to be discussed below). The corresponding structures C and D are shown in Figure 6a and 6b. One additional conformation is given in Figure 5b, which is the average structure calculated from the CPMD run. Finally, the numbering scheme required for the following discussion was given in Figure 2.

A comparison of structures A, C, and D reveals a striking feature—C and D are found to be nearly degenerate and 48 kJ/mol lower in energy than A (see Table 1)—suggesting in turn that the protein scaffold and second shell hydrogen bonding do play a crucial role in stabilizing structure A. Although no other quenches were carried out in the current work, it appears likely that many more local minima would have been found by this procedure. Therefore, a systematic exploration of other alternative structures remains a challenge to be addressed in the future.

The calculated structure A also turns out to be very similar to the crystal structure, with RMSD of 0.34 Å. Conformations C and D are quite distinct, however, with RMSD of 1.38 and 1.62 Å, respectively. The visual inspection of Figures 5a, 6a,

Table 2. Comparison of the Zn−N Bond Lengths, the Zn−Zn Distance and the (HIS)C−N−Zn−Zn Dihedral Angle between the Experimental and Calculated Values

bond	exp. (asym/sym)	A	B	C	D	note
Zn48−N38 (Å)	1.93/2.02	2.10	2.00	2.01	2.05	HIS1
C11−N38−Zn48−Zn47 (deg)	−161.7/−156.8	−158.0	7.3	36.6	−38.4	HIS1
Zn47−N36 (Å)	2.12/2.02	2.06	2.04	2.13	2.01	HIS2
C8−N36−Zn47−Zn48 (deg)	31.4/23.1	15.8	123.4	12.9	168.4	HIS2
Zn47−Zn48 (Å)	3.91/3.94	3.92	4.05	3.80	4.22	

and 6b reveals that additional terminal carboxylate O−imidazole H secondary interactions (three in C and four in D as opposed to two in A and two in the crystal structure) may be responsible for the additional stabilization of structures C and D. In the real protein DF1, two imidazole hydrogens would be substituted by side chain methylenes, thus rendering these types of interactions somewhat weaker.

To study the conformational similarities and differences between structures A, B, C, and D, selected structural data were compiled into Tables 2−4. As may be inferred from Table 2, Zn-N (Imidazole) distances do not vary much between the calculated and the experimental structures. The relative orientation of imidazole rings also does not distinguish much the structures from each other (note that 0° and 180° orientations of the imidazole ring with respect to the Zn−Zn axes are equivalent for Model III). The metal−metal distance, which is thought to be an important factor in various catalytic mechanisms, was found to vary from 3.8 Å in C to 3.9 Å in A and in the crystal structure to 4.2 Å in C, thus being quite elastic.

When inspecting the bonding modes for the terminal carboxylates in the calculated structures, one finds that a chelating bidentate mode (Figure 3) is not favored; only A has a single distorted bidentate carboxylate. As for the monodentate mode of bonding, the syn orientation is preferred, only one example of monodentate-anti carboxylate is found in C. The formation of the latter was caused by the underlying dynamics of the CPMD run, as discussed below. Finally, we infer from Figure 6a, 6b, and Table 3 that the terminal carboxylates are oriented in such a way so to maximize O−H(Im) interactions (notice, however, that in the complete protein model there exist additional hydrogen bonds which would compete with these interactions). A remarkable feature of structure B, while being an average over many hundreds of snapshots of the 10 ps CPMD run, is that it provides a reasonable structural descriptions of terminal carboxylates 1 (TC1) and two bridging carboxylates (additionally discussed below).

Although there are quite a few minor differences in the orientation of bridging carboxylates among structures A, B, C, and D, only one conformational alternative was found−BC1 in D adopts a μ-1,3-anti,syn orientation (the Zn−O−C−O dihedral angle of 106 degrees). Although the μ-1,1,3 modes of bonding were observed transiently during the course of CPMD simulation (discussed below), all four structures A, B, C, and D contain bridging carboxylates ligated in solidly μ-1,3 manner.

In summary, two newly generated local minima C and D are significantly (48 kJ/mol) stabilized compared to the near native structure A. Although the terminal carboxylates are ligated mainly in the monodentate-syn manner and the bridging ones in the μ-1,3-syn,syn manner, examples of terminal monodentate-anti and bridging μ-1,3-anti,syn modes were observed as well. This static analysis, while being very important, provides only a limited description of the potential energy surface−a Molecular Dynamics simulation is helpful in uncovering many more details.

3.3 Dynamical Behavior of the Smaller Model. The main details for protocol of the 10 ps Car−Parrinello Molecular Dynamics simulation of Model III were given in the Computational Details section. Optimized structure A (Figure 5) was used as a starting configuration. The Kohn−Sham energy for the molecule as a function of simulation time is shown in Figure 7a. Because the simulation temperature was allowed to fluctuate by 40 K around the average temperature of 300 K, this translates into approximately 50 kJ/mol fluctuations for the kinetic energy as well as the potential energy. This is indeed the magnitude of the short-time scale energy fluctuations observed in Figure 7a. However, one may also interpret the energy plot in Figure 7a as suggestive of a low-frequency potential energy oscillation caused by some slow relaxation motion. We have indeed observed several-picosecond time scale motions for Model III, as discussed below in great detail; however, much longer simulation times are necessary to characterize the exact nature of the potential energy low-frequency modulation observed in Figure 7a.

As alluded to earlier, we have chosen two low-lying configurations at 5.5 and 9.3 ps for the geometry optimization, which result in structures C and D, having corresponding energies of −170.1 kJ/mol and −170.4 kJ/mol (compared to −122 kJ/mol for structure A, Figure 5). It is reasonable to expect that many more alternative low-energy conformations would be found by quenching from other snapshot configurations. Because any two conformations would belong to the same basin of attraction if not separated by sufficiently long time window, there should exist an optimal sampling rate to quench the simulation snapshots.

The calculated RMSD between the CPMD trajectory configurations and both the crystal structures and structure A (Figure 7b) indicate that the crystal structure has "unfolded" already in the 2 ps equilibration phase, the RMSD never dropping below 1 Å during the entire simulation. This is easily understood a posteriori when considering at least the 48 kJ/mol excitation energy required to reach the "native" A-like state, resulting in the low Boltzmann weight for the A state's occupation (unless some excess entropy is associated with state A, which is not very likely). The peak RMSD of 2 Å during the course of the simulation for such a small molecule is suggestive of quite efficient sampling of the conformational space, implying low barriers between multitude of local minima.

To further investigate the various aspects of the phase dynamics during the CPMD simulation, we have followed the individual motion of two imidazole, two terminal carboxylate and two bridging carboxylate ligands. See Figure 2 for the numbering scheme: imidazole 1 (Im1), terminal carboxylate 1 (TC1) and imidazole (Im2), terminal carboxylate 2 (TC2) refer to the left and right sides of the molecule respectively, whereas BC1 and BC2 refer to lower and upper bridging carboxylates.

The rotation of the imidazole rings with respect to the Zn−N bond axis is one of the important factors determining the overall

Table 3. Zn−O Bonds Lengths, O(Glu)−H(His) Bond Lengths (not shown if less than 3.0 Å), and Zn−O−C−O Dihedral Angles for Terminal Carboxylates TC1 and TC2[a]

bond or dihedral	exp. (asym/sym)	A	B	C	D	note
Zn48−O42 (Å)	2.13/1.96	2.00	1.86	2.00	1.95	TC1
Zn48−O43 (Å)	2.05/1.93	2.45	2.57	2.64	2.57	TC1
O(Glu)−H(C;His) (Å)	none	**2.79**	none	**2.64**	**2.89**	TC1
O42−C5−O43 (deg)	114.1/114.0	121.3	109.8	122.9	122.0	TC1
Zn48−O42−C5−O43 (deg)	31.4/−4.9	7.4	−0.2	9.2	4.0	TC1
Zn48−O43−C5−O42 (deg)	−32.5/4.9	−6.1	n/a	n/a	n/a	TC1
Zn47−O41 (Å)	1.98/1.96	1.96	3.01	4.11	4.06	TC2
Zn47−O46 (Å)	2.19/1.93	2.65	2.04	1.90	2.03	TC2
O(Glu)−H(C;His) (Å)	none	none	none	**2.75**	**2.28,2.53**	TC2
O41−C2−O46 (deg)	113.0/114.0	123.0	n/a	122.7	124.7	TC2
Zn47−O41−C2−O46 (deg)	0.0/−4.9	−6.8	n/a	n/a	n/a	TC2
Zn47−O46−C2−O41 (deg)	0.0/4.9	n/a	n/a	174.5	−82.6	TC2
terminal carboxylate 1	bident./bident.	bidentate	mono-syn	mono-syn	mono-syn	
terminal carboxylate 2	bident./bident.	mono-syn	N/A	mono-anti	mono-syn	

[a] Some entries are shown as n/a when considered ill-defined.

Table 4. Zn−O Bonds Lengths, O(Glu)−H(His) Bond Lengths (not shown if less than 3.0 Å), and Zn−O−C−O Dihedral Angles for Bridging Carboxylates BC1 and BC2[a]

bond or dihedral	exp. (asym/sym)	A	B	C	D	note
Zn48−O40 (Å)	2.09/2.23	1.97	1.99	2.12	2.03	BC1
Zn47−O39 (Å)	2.12/2.14	2.13	1.94	1.97	1.98	BC1
Zn48−O39 (Å)	3.15/2.82	3.25	3.50	3.63	3.85	BC1
Zn47−O40 (Å)	3.93/4.09	3.63	3.25	3.14	3.09	BC1
O(Glu)−H(C;His) (Å)	**2.73**/	**2.73**	none	**2.65**	none	BC1
O39−C14−O40 (deg)	124.0/121.1	125.2	121.8	123.2	123.0	BC1
Zn47−O39−C14−O40 (deg)	36.2/49.4	54.9	0.7	−6.0	11.8	BC1
Zn48−O40−C14−O39 (deg)	16.8/45.0	23.2	−57.1	−69.4	−105.8	BC1
Zn48−O44 (Å)	1.80/2.14	2.15	1.99	2.12	2.13	BC2
Zn47−O45 (Å)	2.10/2.23	2.00	1.97	2.02	1.95	BC2
Zn48−O45 (Å)	3.69/4.09	3.75	3.51	3.31	3.95	BC2
Zn47−O44 (Å)	3.28/2.82	3.12	3.38	3.32	3.29	BC2
O(Glu)−H(C;His) (Å)	**2.84**/	none	none	none	**2.56**	BC2
O44−C10−O45 (deg)	125.0/121.1	123.7	125.0	122.8	123.2	BC2
Zn47−O45−C10−O44 (deg)	−43.2/45.0	−4.3	2.6	50.3	−4.0	BC2
Zn48−O44−C10−O45 (deg)	73.1/49.4	82.2	42.8	35.0	73.8	BC2
bridging carboxylate 1	syn,syn/syn,syn	syn,syn	syn,syn	syn,syn	**anti,syn**	
bridging carboxylate 2	syn,syn/syn,syn	syn,syn	syn,syn	syn,syn	syn,syn	

[a] Bridging carboxylates are found in the μ-1,3 conformation for all structures listed in this table.

flexibility of the molecule because its hydrogen atoms (bonded to carbons) form weak secondary interactions with the carboxylate oxygens (since imidazole rings are connected to the protein backbone in the experimental structure, the protein scaffold is expected to reduce somewhat their rotational mobility). Although the latter secondary interactions are discussed in greater detail below, we observe from Figure 8 that there is a dynamical asymmetry between Im1 and Im2: Im1 undergoes mainly a rocking motion, whereas Im2 is spinning with a periodicity of approximately 3 ps. This asymmetry defines the basic theme of Model III dynamics during the course of the simulation, with the left side of the molecule being relatively rigid and the right side having a floppy character.

The reason for the apparent rigidity of Im1 becomes more obvious when the secondary bonding between Im1 hydrogens (attached to carbons) and TC1 oxygens is monitored (Figure 9a). H21 (Im1) is locked almost permanently into O43 (TC1), restricting the conformational motion for both the imidazole and the carboxylate. To follow the evolution of the bonding mode for the latter ligand during the simulation, the corresponding O−Zn distances are plotted in Figure 9b. In addition, the orientation of the carboxylate plane with respect to the shortest O−Zn bond as a function of simulation time is shown in Figure 10. The latter two figures indicate that TC1 is ligated overwhelmingly in a monodentate-syn fashion, the transient bidentate orientation being observed at 1.6, 7.25, 8.6, and 9.2 ps.

In contrast to the left side of the molecule, the right side undergoes extensive conformation rearrangements during the course of the simulation. The spinning motion of the Im2 ring (see Figure 8) is well correlated with the sequential hydrogen bonding by imidazole H25 and H27 with the terminal carboxylate oxygens (Figure 11a). Notice that three out of four possible H(Im2)−O(TC2) hydrogen bonds are periodically formed and broken, with an average oscillation frequency of 4 ps for bonds O46−H25 and O46−H27. At the beginning and the end of the simulation, two oxygens of the TC2 group hydrogen bond simultaneously to the same imidazole hydrogen (H27; see Figure 11a). The optimized structure D (quenched from a 9.3 ps snapshot) is representative of such bonding arrangement (Figure 6).

Terminal carboxylate 2, on the other hand, undergoes interesting isomerizations. While it is almost always found in a monodentate mode (see Figure 11b), the initial O41 bond to Zn is swapped with O46 at 2.5 ps. In addition, the original syn orientation of the carboxylate plane suddenly turns into anti at 4.2 ps followed by a switchback to syn at 9.4 ps (Figure 10). The TC2 in structure C (annealed from a 5.5 ps snapshot) is a representative example of a carboxylate monodentate-anti

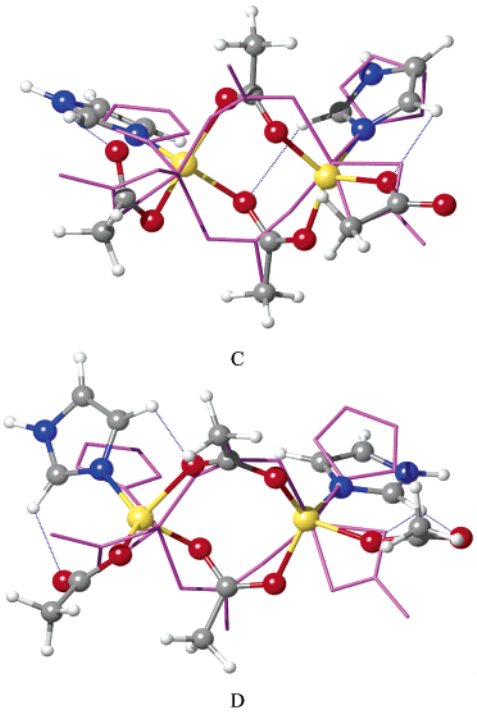

Figure 6. Minimal 48-atom model of the Zn core. (C) A superposition of the original crystal structure (magenta sticks) with one of the local minimum structures as determined from a 5.5 ps snapshot of a 10 ps CPMD run. (D) A superposition of the original crystal structure (magenta sticks) with another local minimum structures as quenched from a 9.3 ps snapshot of a 10 ps CPMD run.

ligation (Figure 6). Although TC2 in structure D (annealed from a 9.3 ps snapshot) is found formally in the syn conformation (Figure 6), the Zn−O−C−O dihedral angle is very close to being a right angle (Table 3), indicating a transition midpoint between the anti and syn conformations (Figure 10).

Finally, the conformational modes of bridging carboxylates 1 and 2 (lower and upper in Figure 2), have been followed with the help of the corresponding O−Zn distances (Figure 12). As may be inferred from Figure 12, bridging carboxylates are found overwhelmingly in the μ-1,3 mode, with transient switchover to μ-1,1,3 mode at 1.9 and 9.4 ps for BC1 and 5.5 and 9.0 ps for BC2. Interestingly enough, optimized structures C and D quenched from the 5.5 and 9.3 ps snapshots respectively do not show any sign of μ-1,1,3 ligation (Figure 6, Table 4), providing an additional indication that this mode is not favored. BC1 of structure D was found in the μ-1,3-anti,syn orientation, pointing to an additional source of flexibility (Table 4).

In summary, H(imidazole)−O(carboxylate) weak hydrogen bonds play a key role in modulating the dynamics of the system. The nearly permanent locking of the H21 (Im1)−O43(TC1) secondary bond imposes a strong constrain on the left side of the molecule, leading to a rocking motion of Im2 and overwhelmingly monodentate-syn mode of the TC2 ligation. The fluid formation of H(Im2)−O(TC2) hydrogen bonds, on the other hand, is correlated well with the spinning of the Im2 ring and the TC2 syn-anti isomerization, suggesting, perhaps surprisingly, that quite a rich chemistry is available for exploration at the 10 ps time scale. From the general symmetry considerations, we expect that the observed left-right *structural* asymmetry would be smoothened if a sufficiently long time simulation were to be carried out (i.e., in a sense of time-averaged structural

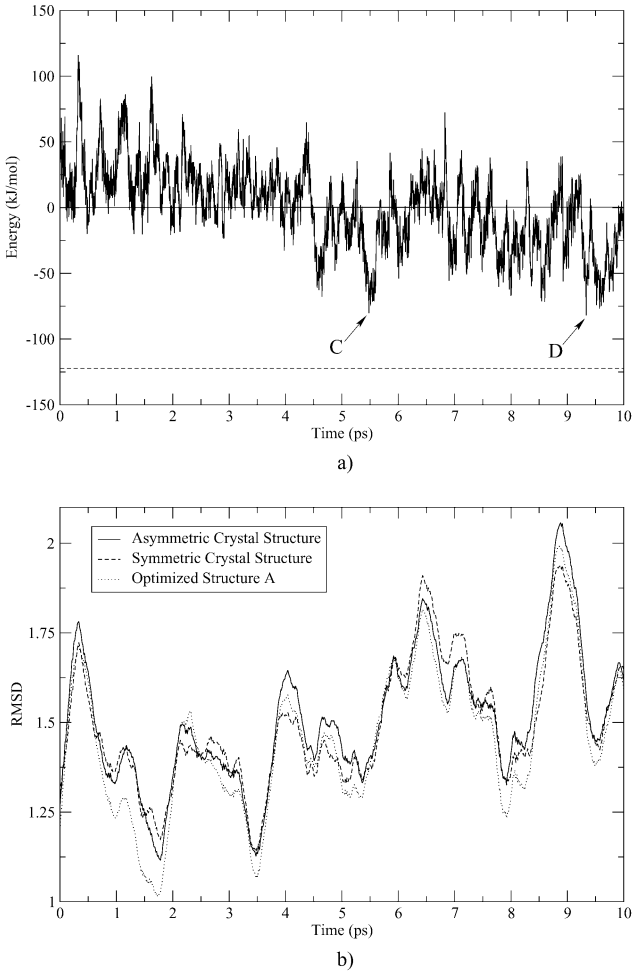

Figure 7. (a) Energy of the minimal 48-atom model as a function of simulation time. (b) Root Mean Square Displacement (RMSD) between the simulation conformation and the asymmetric crystal structure (solid line), the symmetric crystal structure (dashed line), and the optimized structure A (dotted line).

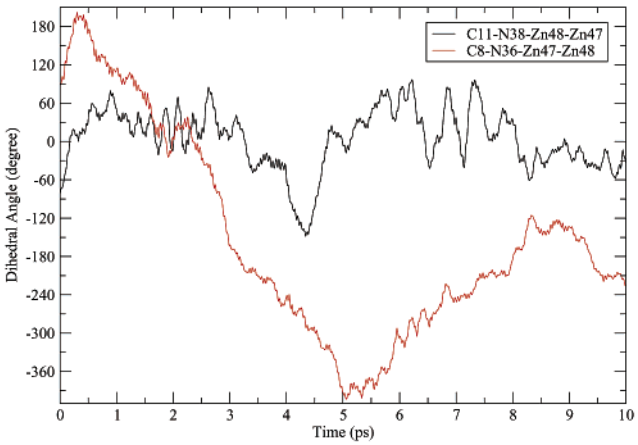

Figure 8. C−N−Zn−Zn dihedral angle for Imidazole 1 (I1) (black) and Imidazole 2 (I1) (red) as a function of simulation time.

elements). As for the *dynamical* left-right asymmetry, our study suggests that at any given time one side of the molecule is expected to be more labile than the other. However, it is quite likely that the left and right clusters in the molecule would switch back and forth from exhibiting a floppy behavior to being conformationally restrained.

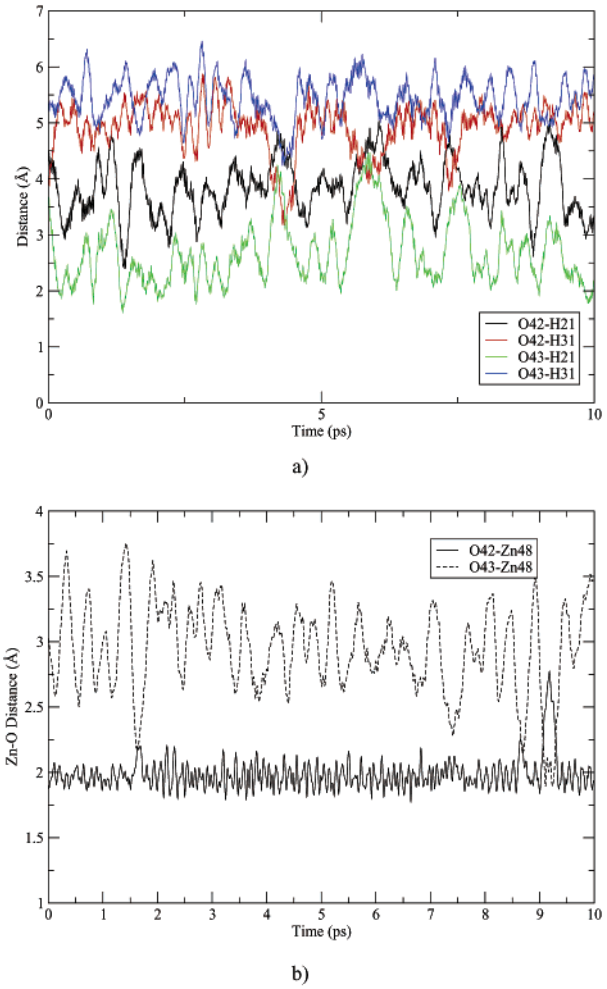

Figure 9. Terminal carboxylate 1 (TC1) Zn−O(Glu) and O(Glu)−H(His) distance as a function of simulation time.

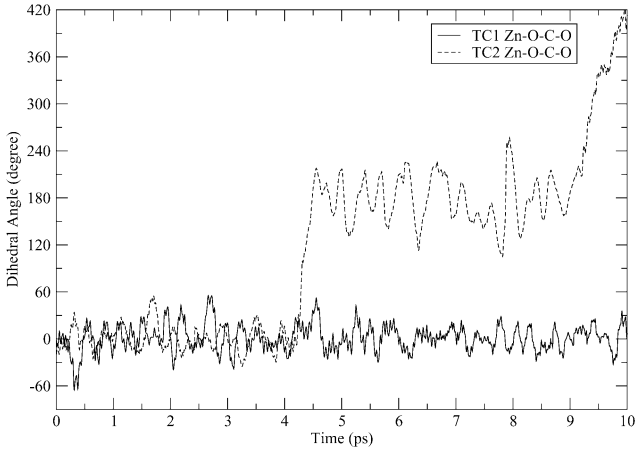

Figure 10. Dihedral angle which defines the syn/anti (0 or 360 vs 180) mode of the terminal carboxylate groups as a function of simulation time. Terminal carboxylate 1 (TC1) is shown as a solid line, terminal carboxylate 2 (TC2) is shown as a dashed line.

4. Conclusions

Dimetal protein cofactors housed in four-helix bundles catalyze a diverse group of "difficult" chemical reactions, playing an important role in the biological world. To mimic such large and complex enzymes, DeGrado and co-workers have *de novo* designed, synthesized, and structurally characterized

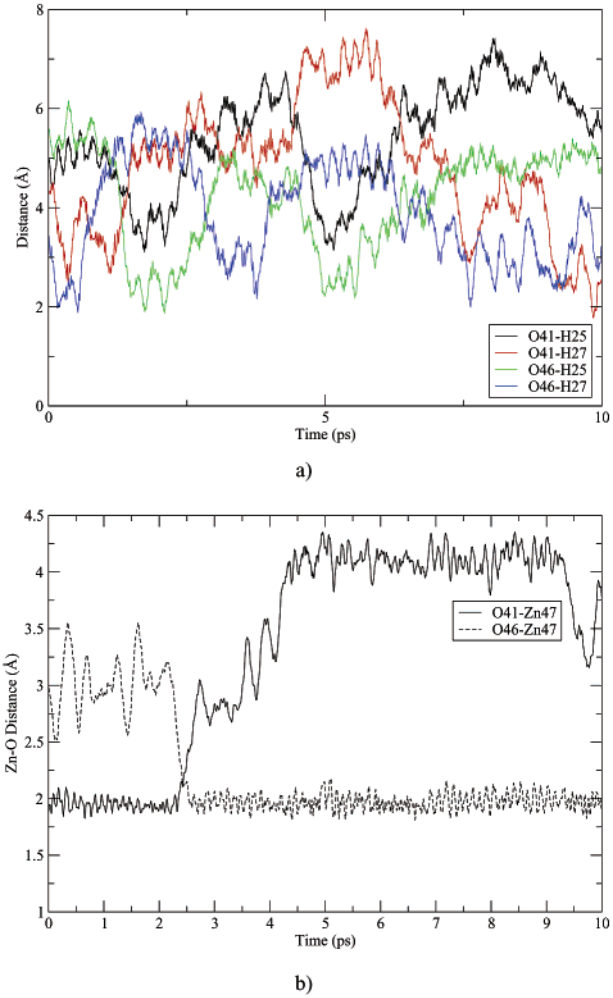

Figure 11. Terminal carboxylate 2 (TC2) Zn−O(Glu) and O(Glu)−H(His) distance as a function of simulation time.

two related small proteins, DF1 (Zn cofactor) and DF2 (Mn cofactor), which are structurally very similar to natural enzymes. A systematic experimental and theoretical investigation of these proteins should afford a better understanding of the catalytic action by natural dimetal enzymes, as well as may lead to novel proteins with tailored catalytic properties.

In this paper, we have analyzed the di-Zn core of DF1 employing three models of decreasing levels of complexity. 94-atom Model I, the most complete one, was derived from DF1 by replacing ligating (4 Glu and 2 His) side chain C_β methylene groups by methane, and pinning the positions of the corresponding carbon atoms to simulate the constraining effect of the protein scaffold. In addition, second shell Tyr and Asp residues which hydrogen bond to terminal Glu and His respectively, were modeled by methanol and acetic acid, respectively. In 66-atom Model II, the latter second shell hydrogen bonding ligands were dropped. Finally, in 48-atom minimalist Model III, Glu sidechains were replaced by acetic acids and His side chains by imidazole, lifting in addition all constrains on the motion of atoms.

The geometry optimization of all three Models by a plane-wave DFT method lead to a surprising conclusion−all models reproduced the crystal structure with a remarkable accuracy (RMSDs: 0.34 Å, 0.29 Å, and 0.34 Å for Models I, II, and III). The only minor deviation from crystal structure bonding

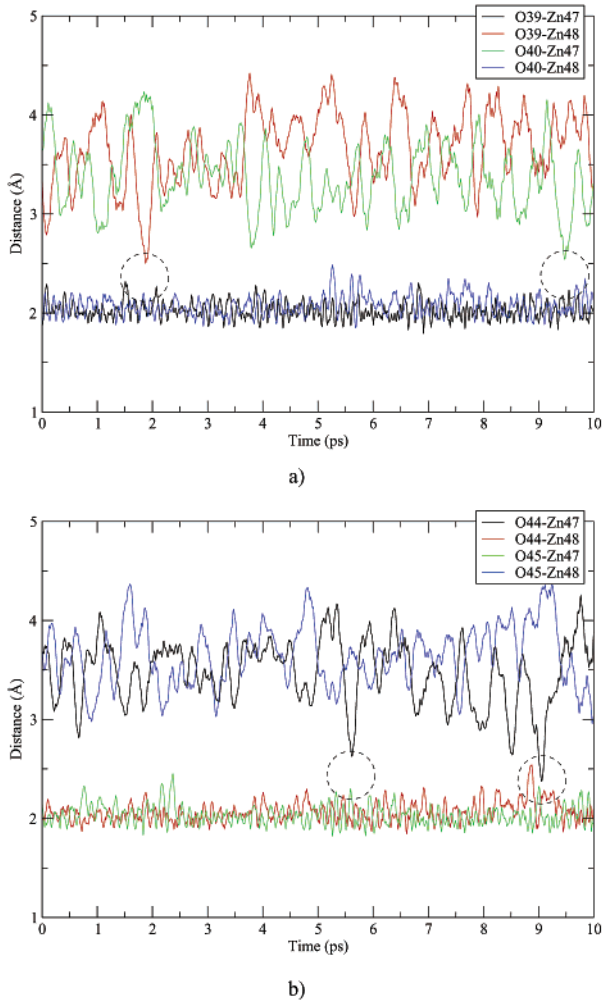

Figure 12. Zn−O distance as a function of simulation time. (a) Bridging Carboxylate I (BC1); (b) Bridging Carboxylate 2 (BC2).

arrangement was the change in the ligating mode for one of the terminal carboxylates from a bidentate to monodentate-syn conformation. Because it was not clear whether these structures were the only low-lying minima for their respective models, we have carried out a 10 ps Car−Parrinello Molecular Dynamics simulation for Model III, to explore the intrinsic chemical phase space of the di-Zn core.

By optimizing the geometries of two low-energy snapshots of the CPMD simulation we have arrived at structures C and D which turned out to be 48 kJ/mol lower in energy than the native local minimum structure A. Additional secondary interactions between imidazole hydrogens (attached to carbons) and carboxylate oxygens may be responsible for the lower energy of these structures. The carboxylates were found in somewhat different orientations in C and D, one of C's terminal carboxylates being flipped from syn to anti orientation, while one of D's bridging carboxylates being changed from μ-1,3-syn,syn to μ-1,3-anti,syn. We have also calculated the average structure of the CPMD run (structure B), which reproduces well the geometry and bonding of those ligands which were not moving too excessively during the simulation.

To analyze the CPMD trajectory, the molecule is imagined as being partitioned into left and right halves. We have suggested that H(imidazole)−O(carboxylate) secondary interactions play a critical role in fine-tuning the dynamics of the system. Left half imidazole hydrogen and carboxylate oxygen atoms were locked in a particular nearly permanent interaction, leading to the rocking motion of the imidazole ring and very few orientational transitions for the carboxylate group (overwhelmingly monodentate-syn mode). On the contrary, hydrogen bonding for the right half of the molecule was very transient, correlating well with spinning of the imidazole ring and syn−anti−syn isomerization of the terminal carboxylate group. These results bring certain hope that short time-scale CPMD simulations may help us to uncover significant details of the dimetal core chemical phase space.

In the future work, the local minima found by quenching the snapshots from the CPMD simulation of the minimalist model may be screened for the compatibility with protein scaffold constraints as well as for the effect of the second shell hydrogen bonding network. We anticipate that a CPMD simulation for the full model may show slow dynamics due to the steric constraints and strong hydrogen bonding network. However, we expect it to be very useful for studying the mechanisms of certain catalytic reactions where no significant rearrangement of the dimetal cofactor is thought to be important. Otherwise, the quenching procedure described in this paper may be used to generate possible candidates for the catalytic intermediates.

Acknowledgment. We gratefully acknowledge the National Institute of Health for its generous support of this work. We thank Dawne A. Yarne, Simone Raugei, Dongsup Kim, Preston B. Moore, Carme Rovira, and Michael Eastwood for help and stimulating discussions.

JA028161L