VOLUME 114, NUMBER 27
DECEMBER 30, 1992

JOURNAL
OF THE
AMERICAN
CHEMICAL
SOCIETY

Molecular Dynamics Simulations on HiPIP from *Chromatium vinosum* and Comparison with NMR Data

Lucia Banci,[†] Ivano Bertini,*[,†] Paolo Carloni,[†] Claudio Luchinat,[‡] and Pier Luigi Orioli[†]

Contribution from the Department of Chemistry, University of Florence, Florence, Italy, and Institute of Agricultural Chemistry, University of Bologna, Bologna, Italy.
Received May 26, 1992

Abstract: The solution structure of HiPIP from *Chromatium vinosum* has been modeled through molecular dynamics (MD) calculations and compared with the X-ray structure. This is the first attempt to investigate open shell ions in a polymetallic system through MD. The results have been compared with available NOE constraints, which are definitely in better agreement with the MD model than with the X-ray structure. This contribution opens the route to the investigation of iron–sulfur proteins through the combined use of MD and NMR.

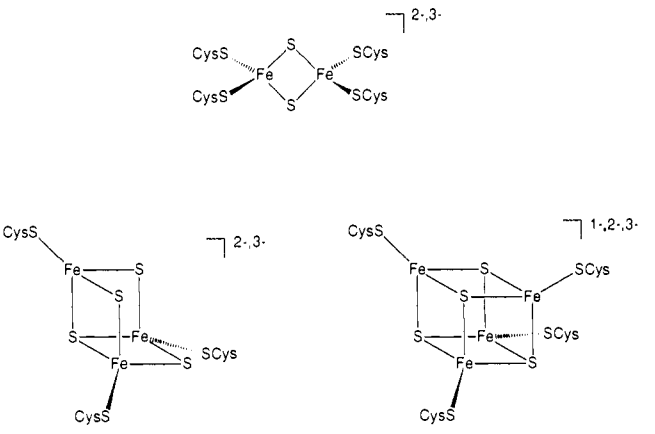
Introduction

In the past few years molecular dynamics (MD) studies have become increasingly useful as a tool for analyzing the structural features of polypeptides and proteins.[1,2] Furthermore, the available force field parameters for amino acid residues have already reached a noteworthy level of sophistication. They are capable of describing the dynamics at the picosecond level, of determining the solution structures with the use of NOE constraints, and of modeling the structure of a protein, given the X-ray structure of a homologous protein.[3-5]

The presence of metal ions coordinated to proteins is a major obstacle to MD studies because of the complexity of their electronic structure. Reliable parameters for a few specific cases, like the closed shell zinc(II) ion in zinc enzymes[6] and high-spin iron(III) in heme proteins,[7] have begun to appear. However, metalloproteins are still avoided by most groups involved in MD studies on biological molecules; if one recalls that about half of all known enzymes contain metal ions,[8] it is apparent that a large number of systems have been ignored.

The difficulties are even more serious in polymetallic systems where, besides metal–protein interactions, metal–metal interactions are also present. Iron–sulfur proteins are typical examples of this class.[9] They are electron transfer proteins containing Fe_2S_2, Fe_3S_4, or Fe_4S_4 clusters (Chart I). Among those containing the Fe_4S_4 cluster are the ferredoxins, displaying reduction potentials in the

Chart I



[†] University of Florence.
[‡] University of Bologna.

(1) (a) McCammon, J. A.; Harvey, S. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987. (b) Brooks, C. L., III; Karplus, M.; Petitt, B. M. *Proteins: A Perspective of Dynamics, Structure and Thermodynamics*; J. Wiley and Sons: New York, 1988.
(2) Kollman, P. A.; Merz, K. M. *Acc. Chem. Res.* **1990**, *23*, 246–252.

0002-7863/92/1514-10683$03.00/0

range −250 to −650 mV, and the so-called high-potential iron–sulfur proteins (HiPIP) with positive reduction potentials in the range +50 to +450 mV. It is known that in ferredoxins the redox pair is $[Fe_4S_4]^{2+}/[Fe_4S_4]^{+}$, whereas in HiPIPs it is $[Fe_4S_4]^{3+}/[Fe_4S_4]^{2+}$.

HiPIPs are particularly interesting as a class. They are widespread proteins whose biological function, at variance with ferredoxins, is still unclear; they have a surprisingly low degree of homology, redox potentials spanning over 400 mV for the same redox pair, and yet have very similar spectroscopic properties suggesting a similar overall fold around the metal cluster. The sequence is known for the proteins from *C. vinosum*, *R. gelatinosus*, *E. halophila I*, *E. halophila II*, *R. gracile*, and *R. tenue*,[10] and unreported sequences for *R. globiformis* and *E. vacuolata I* and *II*[11] are known. The X-ray structure is only available for *C. vinosum*[12] and *E. halophila I*,[13] at 2.0 and 2.5 Å resolution, respectively. If it were not for the presence of the iron–sulfur cluster, these proteins would have already constituted an ideal research area to attempt the modeling of the many unknown structures from the two existing structures.

Herein we investigate the possibility of applying MD calculations on HiPIPs in order to model their solution structure. We have chosen the HiPIP from *C. vinosum* because its X-ray structure is available.[12] We use the X-ray structure as the starting point and then use NMR data to test the correctness of our simulations. We will show that we can treat the Fe_4S_4 cluster by deriving a set of force field parameters which takes into account experimental data and values already used for other metal sites. To model the charge distribution in the cluster we have used recent ab initio calculations on Fe_2S_2 clusters.[14] These parameters are able to reproduce and maintain along all the MD trajectory the geometry of the cluster and therefore are suitable for characterizing the structure and dynamics of this protein. In metalloenzymes that contain catalytic metals, the metal center is solvent accessible and participates in the catalytic reactions, and therefore an accurate description of its electronic structure is necessary. HiPIPs are electron transport proteins where the cluster acts as electron donor/acceptor. In redox proteins, the variation in the oxidation number by one unit is not expected to sizably alter the structural features of the surrounding protein.[15] Therefore, an exact description of the electronic structures is not crucial, whereas it is important to pay attention to the geometric features immediately surrounding the cluster.

Table I. Force Field Parameters Used for the Iron–Sulfur Cluster[a]

Bending Constants	
S*–Fe–S = S*–Fe–S* = Fe–S*–Fe	230.1 kJ mol^{-1} rad^{-2} (55.0 kcal mol^{-1} rad^{-2})
C–S–Fe = LP–S–Fe = LP–S–C	627.6 kJ mol^{-1} rad^{-2} (150.0 kcal mol^{-1} rad^{-2})

Partial Atomic Charges

		oxidized HiPIP	reduced HiPIP
S	+0.31		
LP	−0.40		
C	−0.30		
Fe		+1.47	+1.34
S*		−0.91	−1.04

[a] S* = inorganic sulfur, S = cysteine sulfur, LP = electron lone pair for cysteine sulfur, C = cysteine β-carbon.

For the HiPIP from *C. vinosum*, besides the X-ray structure, NOE data have recently become available, providing NOE connectivities over a large part of the whole protein[16–18] and in particular in the immediate surroundings of the cluster.[16] The latter NOE data have been used, by taking the X-ray structure as a starting point, to assign all the protons of the coordinated cysteines and to therefore assign the individual oxidation states of the iron ions in the cluster.[16,17] The overall agreement with the X-ray data was acceptable; however, the intensity of the NOEs[16] sometimes did not follow the X-ray-derived interproton distances, suggesting that the actual structure in solution is somewhat different from the X-ray structure. We are thus in a position to perform a complete MD study on the HiPIP from *C. vinosum* by using the X-ray structure as a starting point and then to check the results against independent solution data. We show here that the results constitute a better description of the solution structure with respect to the crystal structure and open the way to model the solution structures of the other HIPIPs for which no structural data are available. Comparative analysis of these structures will help in understanding, for instance, the large variations in reduction potentials within the class of HIPIPs and between HIPIPs on one side and ferredoxins on the other and the distribution of charges inside the cluster, and it may provide some hints on the electron transfer pathways operative within the protein.

Computational Procedure

The simulations were carried out using the AMBER 4.0 program[19] running on an IBM RS/6000 520.

Parametrization of the Metal Active Site Parameters. (a) Nonbonded Interactions. It has been pointed out[6b,20] that an important point in MD calculations when dealing with metalloenzymes is to estimate the point charges for the atoms at the active site: it is well known[21] that electro-

(3) (a) Kaptein, R.; Zuiderweg, R. M.; Sheek, R. M.; Boelens, R.; van Gunsteren, W. F. *J. Mol. Biol.* **1985**, *182*, 179–182. (b) Clore, G. M.; Gronenborg, A. M.; Brünger, A. T.; Karplus, M. *J. Mol. Biol.* **1985**, *191*, 523–551. (c) Creiton, S.; Rudolph, B.; Lybrand, T.; Singh, U. C.; Shafer, R.; Brown, S.; Kollman, P. A.; Case, D. A.; Andrea, T. *J. Biomol. Struct. Dyn.* **1989**, *6*, 929–969. (d) Kessler, H.; Matter, H.; Gemmeker, G.; Kling, A.; Kottenham, M. *J. Am. Chem. Soc.* **1991**, *113*, 7550–7553. (e) Bassolino, D. A.; Hirata, F.; Kitchan, D.; Pardi, A.; Levy, R. M. *Int. J. Supercomput. Appl.* **1988**, *2*, 41–61. (f) Holak, T. A.; Prestegard, J. H.; Forman, J. D. L. *Biochemistry* **1987**, *26*, 4652–4660. (g) Braun, W. *Q. Rev. Biophys.* **1987**, *19*, 115–157. (h) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *298*, 992–1023. (i) Clore, G. M.; Gronenborn, A. M.; Nilges, M.; Ryan, C. A. *Biochemistry* **1987**, *26*, 8012–8023 and references therein.

(4) (a) Ovaska, M.; Taskien, J. *Proteins: Struct., Funct and Gen.* **1991**, *11*, 79–94. (b) Eklund, H.; Gleason, F. K.; Holmgren, A. *Proteins: Struct., Funct and Gen.* **1991**, *11*, 13–28. (c) Falconi, M.; Rotilio, G.; Desideri, A. *Proteins: Struct., Funct and Gen.* **1990**, *10*, 149–155. (d) Scully, J.; Evans, D. R. *Proteins: Struct., Funct and Gen.* **1991**, *9*, 191–206. (e) Struthers, R. S.; Kitson, D. H.; Hagler, A. T. *Proteins: Struct., Funct and Gen.* **1991**, *9*, 1–11.

(5) (a) Paulsen, M. D.; Ornstein, R. L. *Proteins: Struct., Funct. Genet.* **1991**, *11*, 184–204. (b) Nilsson, L.; Ahgren-Stålhandske, A.; Sjögren, A. S.; Hahne, S.; Sjöberg, B. M. *Biochemistry* **1990**, *29*, 10317–10322.

(6) (a) Merz, K. M., Jr.; Murko, M.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 4484–4490. (b) Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 406–411.

(7) (a) Case, D. A.; Karplus, M. *J. Mol. Biol.* **1978**, *132*, 343–368. (b) Lopez, M. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 6212–6222.

(8) Frausto da Silva, J. J. R.; Williams, R. J. P. *The Biological Chemistry of the Elements*; Clarendon Press: Oxford, 1991.

(9) (a) *Iron–Sulfur Proteins*; Spiro, T. G., Ed.; Wiley-Interscience: New York, 1982. (b) Thompson, A. J. In *Metalloproteins*; Harrison, P., Ed.; Verlag Chemie, Weinheim: FRG, 1985; p 79.

(10) (a) Tedro, S. M.; Meyer, T. E.; Bartsch, R. G.; Kamen, M. *J. Biol. Chem.* **1981**, *25*, 731–735. (b) Przysiecki, C. T.; Meyer, T. E.; Cusanovich, M. A. *Biochemistry* **1985**, *24*, 2542–2549.

(11) Ambler, R., personal communication.

(12) (a) Carter, C. W., Jr.; Kraut, J.; Freer, S. T.; Alden, R. A. *J. Biol. Chem.* **1974**, *249*, 6339–6346. (b) Carter, C. W., Jr.; Kraut, J.; Freer, S. T.; Xuong, N. H.; Alden, R. A.; Bartsch, R. G. *J. Biol. Chem.* **1974**, *249*, 4212–4225.

(13) Breiter, D. R.; Meyer, T. E.; Rayment, I.; Holden, H. M. *J. Biol. Chem.* **1991**, *266*, 18660–18667.

(14) Carloni, P.; Corongiu, G., in preparation.

(15) Freeman, H. C. In *Coordination Chemistry*; Laurent, J. P., Ed.; Pergamon Press: Oxford, 1980; p 29.

(16) Bertini, I.; Capozzi, F.; Ciurli, S.; Luchinat, C.; Messori, L.; Piccioli, M. *J. Am. Chem. Soc.* **1992**, *114*, 3332–3340.

(17) Nettesheim, D. G.; Harder, S. R.; Feinberg, B. A.; Otvos, J. D. *Biochemistry* **1992**, *31*, 1234–1244.

(18) Gaillard, J.; Albrand, J.-P.; Moulis, J.-M.; Wemmer, D. E. *Biochemistry* **1992**, *31*, 5632–5639.

(19) AMBER 4.0 by Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Siebel, G.; Singh, U. C.; Weiner, P.; Kollman, P. A., University of California, 1986–1991.

(20) (a) Kollman, P. A. In *The Chemistry of Enzyme Action*; Page, M. I., Ed.; Elsevier Press: Amsterdam, 1983; pp 53–71. (b) Shen, J.; Wong, C. F.; Subrasmaniam, S.; Albright, T. A.; McCammon, J. A. *J. Comp. Chem.* **1990**, *30*, 346–350.

static interactions play a fundamental role in determining the nonbonded interactions, which, in proteins, are key factors which determine the tertiary structure. The use of formal charges has been demonstrated to be unrealistic, as the metal ion charge is actually spread out partially on metal ion ligands.[22] Indeed, also in the present system, the use of the formal charge on the iron ions (+2 or +3) and sulfide ions (-2) produces a collapse of the cluster, even during energy minimization.

Computation of atomic charges for metalloenzymes can be accomplished using quantum mechanical procedures. For example, this has been done for the zinc metalloenzymes carbonic anhydrase[23] and carboxypeptidase,[24] giving satisfactory results. These systems are closed shell systems; however, one of the oxidation states in the present systems is open shell and this poses more problems in obtaining atomic point charges.[25]

In order to treat the point charges in the cluster we have thus used recent ab initio SCF model calculations of plant type 2Fe-2S ferredoxins,[14] containing the Fe_2S_2 cluster (Chart I). The oxidized protein bears two ferric iron ions, while the reduced form can be described best as a Fe(II)-Fe(III) system.[26] Carloni and Corongiu[14] have carried out a Mulliken population analysis[27] for two compounds of the formula $[Fe_2S_2(SCH_3)_4]^{2-/3-}$, which were chosen to model the active site of oxidized and reduced ferredoxin, respectively. As already mentioned in the introduction, reduced HiPIPs are characterized by two ferric and two ferrous iron ions, bridged by inorganic sulfurs in a cubane-type arrangement. This cluster is equivalent, in terms of type and number of atoms and of total number of electrons, to twice the reduced Fe_2S_2 cluster. At variance with the Fe_2S_2 cluster, however, we know from Mössbauer and NMR data that the extra electrons of the reduced iron ions are delocalized over all the cluster, thus making all the iron ions equivalent. Thus, we have taken the charges of the iron ions in the reduced form of the Fe_4S_4 cluster equal to the average of the Mulliken charge values calculated for the two iron ions in the reduced Fe_2S_2 ferredoxin cluster model (Table I). As far as the point charges of the inorganic sulfurs and those of the sulfurs and β carbons of the coordinated cysteines are concerned, the same Mulliken charges obtained in the Fe_2S_2 model also were used in the present system (Table I).

For the oxidized form of HiPIP we have simply removed 0.125 electron from each atom (iron and inorganic sulfur) present in the cluster. We performed calculations also on a system in which the extra electron has been removed from the cluster in an asymmetric way (see below). Passing from the reduced to the oxidized form, we did not change the charges of the cysteine sulfurs since, from the ab initio computation,[14] it appears that the charges on the thiolate sulfurs binding the iron atoms are not significantly affected by removing one electron from the reduced form.

As far as the Lennard-Jones (LJ) parameters are concerned, for the iron atoms we have used the values used by Karplus for iron in myoglobin,[7a] while for the inorganic sulfur atoms the all atom[28] cysteine sulfur values are used.

(b) Bond Interactions. During all the simulations, the SHAKE algorithm[29] has been used for all the bonds of the cluster, thus constraining bond lengths to their starting X-ray values.

As far as bending constants are concerned, suitable values can, in principle, be derived from experimental data, as, for example, from the analysis of vibrational spectra.[1] However, we have found that force constant values two to three times higher (around 200 kJ mol^{-1} rad^{-2} or 50 kcal mol^{-1} rad^{-2}) than the experimental ones[30] (75.3 kJ mol^{-1} rad^{-2} or 18.0 kcal mol^{-1} rad^{-2}) are needed for the Fe-S-Fe, S-Fe-S, and Fe-S-C bending constants to maintain the cubane geometry of the cluster. Indeed, it has been found for other metalloproteins that the experimental values, obtained from the analysis of vibrational spectra, may not be able to model correctly the metal ion coordination and that force constant values about two times higher than the experimental values are needed.[31] The bending parameters we found suitable for describing the present systems are very close to the values used for the N-Fe-N bending constant in myoglobin,[32] i.e. for an iron atom in a rigid, fixed geometry. It is worth noting that these values are still lower than most standard bending constants used in AMBER, still leaving some mobility to the cluster. As an example, AMBER uses a cysteine C-S-C bending constant of 284.7 kJ mol^{-1} rad^{-2} (68 kcal mol^{-1} rad^{-2}).

Our parameters are able to reproduce the geometry of the cluster without keeping it as a rigid moiety. The equilibrium angle values were those of the starting X-ray structure. The torsion angle constants of the atoms forming the cubane were set to zero, and no "lone pair" was added to the inorganic sulfurs. On the other hand, the four bonded cysteines were treated as standard except for the charges of the sulfur and Cγ carbon, for which the ab initio charges were used. The sum of the cysteine and cubane charges results in the correct net charge of -1 for the oxidized and -2 for the reduced form. The parameters used in the present system to describe the iron-sulfur cluster for both the reduced and the oxidized forms are summarized in Table I.

We have therefore phenomenologically built a set of parameters which satisfactorily models the properties of the cluster and which allows us to characterize the structural and dynamical properties of the protein residues around the cluster itself.

Parametrization of the Rest of the Protein. For the residues within a 15 Å sphere centered on the center of mass of the cluster we have used the all-atom AMBER force field parameters, developed by Kollman and co-workers,[28] while for the rest of the protein the united-atom[28] force field has been used.

Protein Model. As our starting system, we have taken the X-ray structure of oxidized HiPIP from *Chromatium vinosum* (2 Å resolution) whose coordinates are available in the Protein Data Bank.[13] By using the interactive molecular graphics package SYBYL,[33] we have added a few atoms whose coordinates were not available in the PDB file, according to the amino acid sequence, and made some amino acid changes in agreement with a subsequent primary structure determination.[10a] We have chosen to simulate a neutral pH solution. Therefore, the only histidine residue present in the molecule was taken as neutral in agreement with its $pK_a < 7$ suggested by potentiometric[34] and NMR titrations.[35] We have protonated it on the Nϵ position; being a surface residue not engaged in obvious hydrogen bonding nets, the choice of which ring nitrogen is protonated is irrelevant.

The protein charge (-4 at neutral pH) was neutralized by adding four sodium ions far from the active site, and they were not placed near salt bridges so as not to interfere with them. The 75 crystallographic water molecules were added to the protein. The whole system was solvated by a sphere of TIP3P waters[36] with the SOL option in EDIT. Water was added up to 9 Å around the protein, which resulted in the addition of 1500 water molecules. The final system was composed of about 5500 atoms.

Molecular Dynamics Calculations. We have used the same procedure for both oxidized and reduced HiPIP.

A 9-Å cutoff was chosen for the nonbonded interactions. The number of resulting pair interactions was about 1 600 000, the pair list being updated every 10 steps of minimization or dynamics. The time step used in the dynamics was 1.5 fs; a dielectric constant value of 1 was chosen. The temperature of the system was maintained constant at 300 K by coupling it with the thermal bath.[37] The values of the energies and the coordinates were saved every 100 steps. All the water molecules were equilibrated by minimizing the root-mean-square energy gradient of 0.42 kJ mol^{-1} (0.10 kcal mol^{-1}) followed by 9 ps of MD. Then, a minimization on the whole system was carried out, the convergence criterion being as above. Finally, MD was performed for 75.6 ps on the whole system. During this step, the water molecules in a shell between 24 and 30 Å from the center of mass of the cluster were constrained by a very low harmonic

(21) (a) Smith, J. C.; Karplus, M. *J. Am. Chem. Soc.* **1992**, *114*, 801-812. (b) Ahlström, P.; Teleman, O.; Jönsson, B.; Forsén, S. *J. Am. Chem. Soc.* **1987**, *109*, 1541-1551.
(22) Hoops, S. C.; Kenneth, W. A.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 8262-8270.
(23) Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 406-411.
(24) Banci, L.; Schroeder, S.; Kollman, P. A. *Proteins: Struct., Funct. Genet.* **1992**, *13*, 288-305.
(25) Ohanessian, G.; Brusich, M. J.; Goddard, W. A., III. *J. Am. Chem. Soc.* **1990**, *112*, 7179-7189.
(26) (a) Palmer, G.; Dunham, W. R.; Fee, J. A.; Sands, R. H.; Izuka, T.; Yonetani, T. *Biochim. Biophys. Acta* **1971**, *115*, 711-715. (b) Dugad, L. B.; La Mar, G. N.; Banci, L.; Bertini, I. *Biochemistry* **1990**, *29*, 2263-2271.
(27) Mulliken, R. S. *J. Chem. Phys.* **1955**, *39*, 1.
(28) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
(29) Van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311-1327.
(30) Czernuszawicz, R. S.; Macor, K. A.; Johnson, M. K.; Gevirth, A.; Spiro, T. S. *J. Am. Chem. Soc.* **1987**, *109*, 7178-7187.
(31) (a) Bernhart, P. V.; Comba, P., *Inorg. Chem.*, in press. (b) Merz, K. M., Jr., personal communication.
(32) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657-1666, as quoted in ref 19.
(33) SYBYL, Molecular Modelling Software Package, Version 5.2, Tripos Associates Inc.
(34) Mizrahi, I. A.; Wood, F. E.; Cusanovich, M. A. *Biochemistry* **1976**, *15*, 343-348.
(35) Nettesheim, D. G.; Meyer, T. E.; Feinberg, B. A.; Otvos, J. D. *J. Biol. Chem.* **1983**, *258*, 8235-8239.
(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.
(37) H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684-3690.

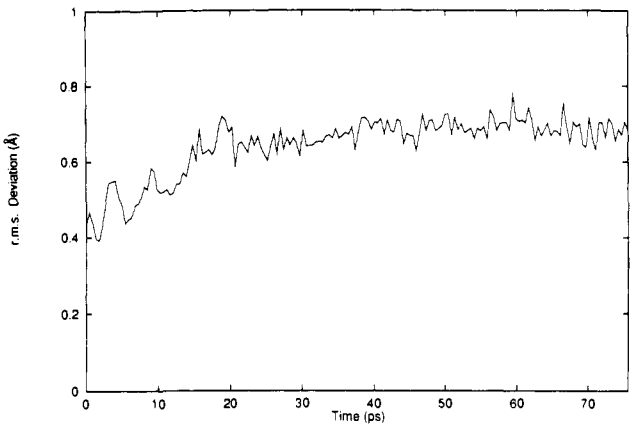


Figure 1. Root-mean-square deviation between the instantaneous MD and the crystal structures of HiPIP from *C. vinosum* for the backbone atoms plotted as function of time. The rms values are plotted every 0.45 ps of dynamics.

potential (0.42 kJ mol^{-1} or 0.10 kcal mol^{-1}) in order to prevent their evaporation without considerably affecting their motion. In the first 12.6 ps the system was gradually heated by performing MD runs of 1.8 ps at the following temperatures: 10, 50, 100, 150, 200, 250, and 300 K. In all simulations the energy reached a constant value after a fraction of a picosecond. The coordinates of another 27 ps of dynamics performed at 300 K were discarded.

Data Analysis. The trajectories of the final 36 ps were collected for analysis and from them an averaged structure was generated by using the MDANAL module of AMBER. The root-mean-square (rms) fluctuations at time t_s, averaged over the protein atoms involved in the MD and over time, from the starting time step (t_j = 27 ps) to the sth time step ($t_j = t_s$) were calculated as

$$\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{s}\sum_{j=1}^{s}|\mathbf{r}_i(t_j) - \langle\mathbf{r}_i\rangle|^2\right\}^{1/2}$$

where $\mathbf{r}_i$ is the position of the ith atom, t_j the jth time step, and n the total number of atoms. The rms deviations of the average structure with respect to the X-ray structure were calculated as $((1/n)\sum_{i=1}^{n}|\Delta\mathbf{r}_i|^2/n)^{1/2}$, where $\Delta\mathbf{r}_i$ is the displacement of an atom in the average MD structure with respect to the X-ray structure, and the sum is calculated over all the atoms, over all the heavy atoms, and over all the backbone atoms of the protein.

Results

MD Calculations on the Oxidized HiPIP. The rms deviation of the average MD structure with respect to the starting X-ray structure of the heavy atoms is 0.96 Å. The rms fluctuations over the whole trajectory, with respect to the MD average structure, averaged over the heavy atoms and the backbone atoms are 0.65 and 0.46 Å, respectively. These very small values indicate that the present protein is characterized by high rigidity and that the average obtained from MD simulations is relatively close to the crystallographic one. Figure 1 shows the rms deviation of the MD structure with respect to the initial structure, relative to the backbone atoms. After initial fast equilibrium, the rms values stabilize to constant values, indicating that the simulated structure is already substantially well equilibrated after about 20 ps. The observed rms value is relatively small for a protein of the present molecular weight.

Figure 2 compares the backbone atoms for the X-ray and the MD average structures. This protein is characterized by the presence of secondary structure motifs (α helixes and β sheets); all these motifs are completely maintained along the entire trajectory.

Figure 3a shows the rms deviations, averaged per residue (considering only the non-hydrogen atoms), between the starting and the average MD structure, while Figure 3b shows the fluctuations, averaged per residue (again considering only the non-hydrogen atoms), with respect to their positions in the average MD structure. In Figure 3a the amino acid sequence is also reported. Figure 3a shows that the residues in the average MD structure, as already mentioned, do not experience a large deviation from their position in the starting structure, except for the residues in the 51–64 range; this region mainly forms a hydrogen-bonded secondary structure, namely strands of twisted antiparallel β-sheets. In any case, the changes in the structure occur during the equilibration; indeed, Figure 3b shows that the fluctuations of each residue around the equilibrium position are all very small and very similar one another.

Finally, the structural arrangement of the residues near the cluster is essentially maintained during the dynamics. Figure 4 shows a comparison between the starting X-ray and the average MD structures for the iron–sulfur cluster and most of the surrounding residues whose proton–proton NOEs have been observed. We find that the backbone structure is not greatly changed in the average structure with respect to the starting one, although the side chains, especially for some residues, experience some significant displacement.

Some of the displacements observed for residues close to the cluster are clearly already detected after energy minimization of the system, which indicates that they are mainly due to the initial presence of bad contacts. Indeed, along the trajectory, after a few picoseconds of equilibration, these residues reach the equilibrium position and then experience only small fluctuations. Some other displacements occur during the MD, but again they are completed within the time range of the simulation. The overall time dependence of the rms for the present system suggests that its increase with time—despite some contribution due to errors in the force field potential—is due to real relaxation of the protein.

Mössbauer studies[38,39] have shown that the four iron atoms in reduced HiPIPs are all equivalent, with mean oxidation states of +2.5. However, in oxidized HiPIPs the extra electron is not evenly distributed over the four centers but is delocalized on only two iron atoms, thus constituting a mixed valence pair. NMR studies have shown that the latter two ions are those coordinated to Cys 63 and Cys 77.[16,17] In our model, discussed above, we have kept a symmetrical electron distribution in both the reduced and oxidized forms. We have then checked the effect of an unsymmetrical charge distribution by performing a second MD calculation where a smaller positive point charge (+1.27) was assigned to the two iron ions belonging to the mixed valence pair and a larger point charge (+1.67) to the two ferric iron atoms. The charge difference of 0.4 unit between the two pairs of irons is again taken from Carloni and Corongiu's calculations[14] on the reduced Fe_2S_2 system. We estimate this to be an upper limit for the charge difference in this case, since we are dealing with two Fe^{3+} versus two $Fe^{2.5+}$ ions. These calculations gave very similar results to those observed in the case of the symmetric system. In particular, all the movements observed for several of the residues close to the iron–sulfur cluster using the symmetrical charge distribution are again observed in these calculations. Furthermore, the average positions are very similar in both cases. Hence, we can conclude that minor changes in the charge distribution on the cluster have little influence on the MD results. We recall, however, as pointed out in the computational procedure, that the use of full formal charges causes a collapse of the cluster.

The above data yield a picture of the structure of oxidized *C. vinosum* HiPIP which is presumably closer to the solution structure than the X-ray structure as far as the bulk of the protein is concerned. However, the ultimate goal of our approach is to learn about the structural properties of the residues in the immediate surrounding of the cluster. To test the validity of the approach in this particular region of the protein we make use of the NMR data recently obtained by us.[16]

Comparison with the NMR Data. Estimates of proton–proton distances in solution can be obtained from ^{1}H NOE or NOESY experiments. These techniques are now routinely applied to provide the intramolecular distance constraints necessary to solve protein structures in solution.[40]

(38) Dickson, D. P. E.; Johnson, C. E.; Cammack, R.; Evans, M. C. W.; Hall, D. O.; Kao, K. K. *Biochem. J.* **1974**, *139*, 105–108.
(39) Middleton, P.; Dikson, D. P. E.; Johnson, C. E.; Rush, J. D. *Eur. J. Biochem.* **1980**, *104*, 289–296.

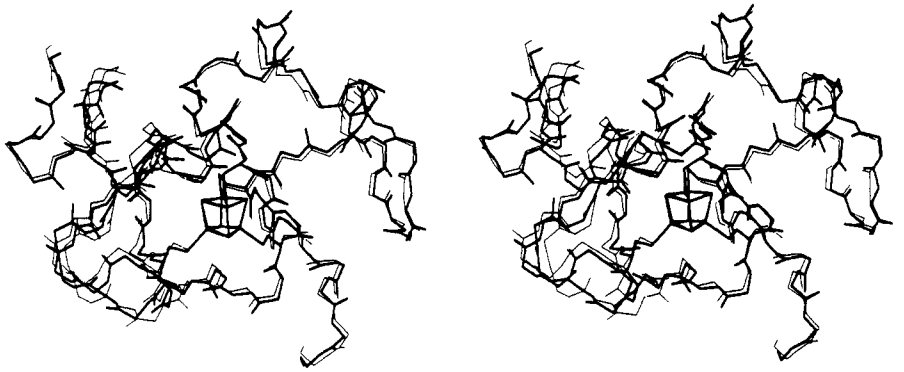


Figure 2. Comparison between the backbone of the X-ray (thin line) and the MD average (bold line) structures of oxidized HiPIP from *C. vinosum*. The latter has been calculated for the oxidized form with a symmetrical charge distribution on the iron ions (see text).

The problems encountered in obtaining accurate distance estimates have been thoroughly discussed.[40,41] They can be summarized as follows: (1) difficulties in determining the selective T_1 values for steady-state NOE measurements, (2) spin diffusion contributions to steady-state NOEs, (3) uncertainty in the reorientational correlation time, (4) anisotropic reorientation, and (5) internal motions. The problems connected with points 1 and 2 can be virtually eliminated by reducing the irradiation time (in 1D experiments) or the mixing time (in 2D experiments) below the relaxation times of the protons involved.[42] In this way, the NOE or NOESY intensities are proportional to the product of the irradiation (mixing) time times the cross-relaxation rate, which is in turn proportional to $1/r^6$. Point 3 can be overcome by using intensity ratios instead of absolute numbers. As far as points 4 and 5 are concerned, the roughly spherical shape of the protein[13] rules out anisotropic rotation as an important source of error; large internal motions are also unlikely in the vicinity of the cluster, as shown by MD calculations (with the possible exception of methyl groups, see below).

A comment is due on the influence of the paramagnetic cluster on the NMR parameters of the protein protons and hence on the reliability of the observed NOEs in the light of the above considerations. In general, paramagnetism decreases the T_1s of the signals and the extent of the NOEs, thereby decreasing the accuracy. On the other hand, smaller NOEs reduce the problem of spin diffusion and make distance estimates more reliable. The only problem may be that if T_1s are too short, then the time region over which buildup is linear in NOESY spectra may be too short to be practical. However, if only one of the dipolar coupled pair of protons has a short T_1, 1D NOE measurements can be performed by irradiating this proton. In this case the buildup is linear up to $\sim T_1$ of the other proton on which NOE is observed. In this respect, the presence of the paramagnetic cluster turns out to be a big advantage: we have as many as 9 cysteine protons outside the diamagnetic region. These protons are well isolated and can be selectively saturated in order to detect NOEs to other protons surrounding the cluster. The latter protons, although still slightly affected by the paramagnetic cluster, have T_1s of the order of 100 ms or more. Indeed, even the paramagnetically shifted α-CH proton of Cys 77 has a T_1 as long as $\sim$40 ms.[17,43]

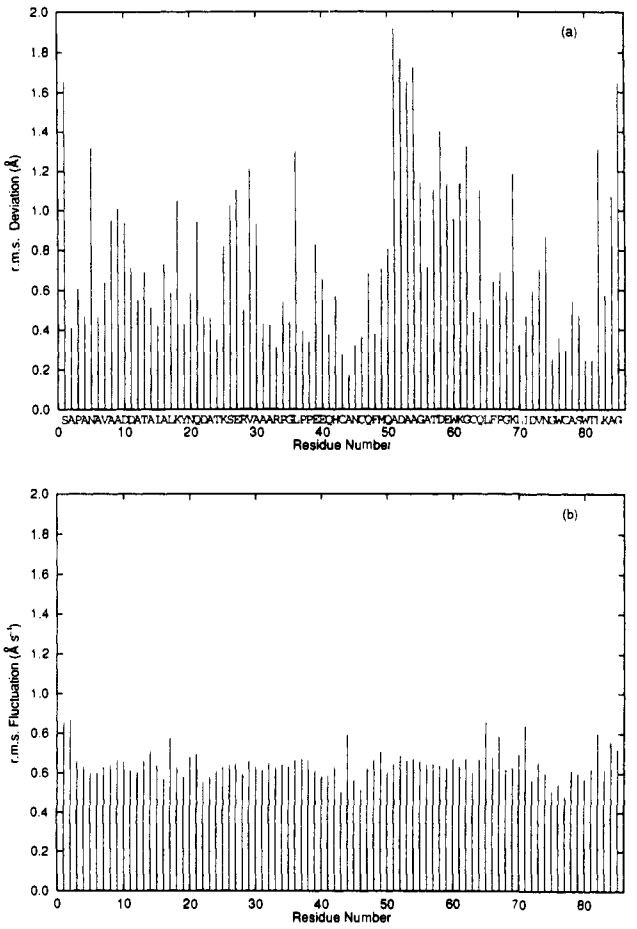


Figure 3. (a) Root-mean-square deviation (together with the amino acid sequence) of the MD average structure of HiPIP from *C. vinosum* with respect to the X-ray-structure and (b) root-mean-square fluctuation per residue with respect to the MD average structure of HiPIP from *C. vinosum*.

In the light of the above considerations we can safely make use of the already available NOE data,[16-18] which have been recorded using the truncated NOE mode.[16] In addition, we have performed some specific checks to establish the linearity between NOE intensities and irradiation time. We have found that for all the relevant NOEs the response is practically linear up to irradiation times as long as 80 ms, hence we do not sacrifice the NOE intensities more than necessary. In this respect, truncated NOEs

(40) (a) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986. (b) Wüthrich, K. *Science* **1989**, *243*, 45-50. (c) Weber, C.; Wider, G.; von Freiberg, B.; Traber, R.; Braun, W.; Widmer, H.; Wütrich, K. *Biochemistry* **1991**, *30*, 6563-6574. (d) Wütrich, K. In *Methods In Enzymology*; Academic Press Inc.: San Diego, CA, 1989; pp 125-131. (e) Sheek, R. M.; van Gunsteren, W. F.; Kaptein, R. In *Methods In Enzymology*; Academic Press Inc.: San Diego, CA, 1989; pp 204-218. (f) Wagner, R. G. *Prog. NMR Spectrosc.* **1990**, *22*, 101-139. (g) Bax, A. *Annu. Rev. Biochem.* **1989**, *58*, 223-256.

(41) (a) Borgias, B. A.; Gochin, M.; Kerwood, D.; James, T. L. *Prog. NMR Spectrosc.* **1990**, *22*, 83-100. (b) Banci, L.; Bertini, I.; Luchinat, C.; Piccioli, M. In *NMR and Biomolecular Structure*; Bertini, I., Molinari, H., Niccolai, N., Eds.; VCH Publishers: Weinheim, FRG, 1991; pp 31-68. (c) James, T. L.; Borgias, B. A.; Bianucci, A. M.; Zhou, N. In *NMR and Biomolecular Structure*; Bertini, I., Molinari, H., Niccolai, N., Eds.; VCH Publishers: Weinheim, FRG, 1991; pp 87-111. (d) Kaptein, R.; Boelens, R.; Koning, T. M. G. In *NMR and Biomolecular Structure*; Bertini, I., Molinari, H., Niccolai, N., Eds.; VCH Publishers: Weinheim, FRG, 1991; pp 113-139. (e) Majundar, A.; Hosur, V. *Prog. NMR Spectrosc.* **1992**, *24*, 109-158.

(42) Neuhaus, D.; Williamson, M. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH Publishers: Weinheim, FRG, 1989.

(43) Sola, M.; Cowan, J. A.; Gray, H. B. *Biochemistry* **1989**, *28*, 5261-5268.

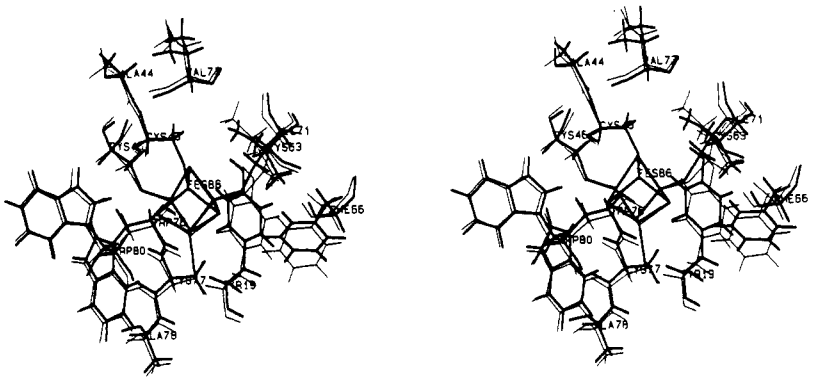


Figure 4. Comparison between the iron–sulfur cluster and some surrounding residues in the X-ray structure (thin line) and the MD average structure (thick line) of HiPIP from *C. vinosum*.

Table II. Comparison between Experimental NOEs in the Neighborhood of the Fe_4S_4 Cluster of *Chromatium vinosum* and the Expectations from the X-ray Structure and the MD Refined Structure[a]

irradiated signal	observed NOE	dist (Å) (X-rays)	NOE[a] (X-rays)	NOE[b] (exp)	NOE[a] (MD)	dist (Å) (MD)
Hα Cys-77	Hα Tyr-19	2.2	vs	s	s	2.4
$H\beta_1$ Cys-77	Hα Tyr-19	2.7	m	m	m	2.6
$H\beta_2$ Cys-77	Hα Tyr-19	2.7	m	s	s	2.4
$H\beta_2$ Cys-77	Hδ Tyr-19	2.0	vs	m	s	2.4
$H\beta_1$ Cys-77	$H\beta_1$ Tyr-19[c]	2.1	vs	w	m	2.9
$H\beta_1$ Cys-77	$H\beta_2$ Tyr-19[c]	2.3	s	m	m	2.6
$H\beta_1$ Cys-77	HN Ala-78[c]	2.8	m	s	s	2.4
Hα Cys-77	HN Ala-78[c]	2.3	s	m	m	2.5
$H\beta_1$ Cys-43	Hα Val-73	2.5	m	vs	vs	2.2
$H\beta_2$ Cys-43	Hα Val-73	3.5	w	w	w	3.1
$H\beta_1$ Cys-63	$H\beta_2$ Phe-66	3.4	w	vw	vw	4.7
$H\beta_2$ Cys-63	$H\beta_2$ Phe-66	1.8	vs	m	m	3.0
$H\beta_1$ Cys-63	$H\delta_2$ Phe-66	3.1	w	vw	vw	4.0
$H\beta_2$ Cys-63	$H\delta_2$ Phe-66	2.2	vs	m	m	2.5
$H\beta_1$ Cys-46	$H\epsilon_1$ Trp-80	3.1	w	m	m	3.0
$H\beta_1$ Cys-46	$H\delta_1$ Trp-80	2.5	m	m	m	2.5
Hα Cys-77	$H\epsilon_3$ Trp-76	2.8	m	s	m	2.8
Hα Cys-77	$H\zeta_3$ Trp-76	3.6	w	w	w	3.5
$H\beta_1$ Cys-43	HN Ala-44	2.9	m	m	m	2.7
$H\beta_1$ Cys-43	HN Cys-43	2.8	m	m	m	2.8
$H\beta_1$ Cys-43	Hα Cys-43	3.0	m	w	m	3.0[d]
$H\beta_2$ Cys-43	Hα Cys-43	2.3	s	m	s	2.4[d]
$H\beta_1$ Cys-46	Hα Cys-43	2.9	m	s	s	2.4
$H\beta_2$ Cys-46	Hα Cys-43	2.6	m	m	m	2.6
$H\beta_1$ Cys-63	$H\gamma_1$ Ile-71	1.9[e]	vs	m	m	2.8[e]
$H\beta_2$ Cys-63	$H\gamma_1$ Ile-71	3.4[e]	w	m	m	2.9[e]

[a] ≤2.2 Å = very strong (vs); 2.3–2.4 Å = strong (s); 2.5–3.0 Å = medium (m); 3.1–3.6 Å = weak (w); >3.7 Å = very weak (vw). [b] The signal intensities for which the agreement with experimental NOEs is improved by MD are shown in boxes. In no case is a worse agreement obtained. [c] Assignment taken from ref 17. [d] The distance after MD does not show large variations due to the constraints imposed. [e] Distance from the closest proton of the methyl group.

are in a more favorable situation than NOESY because of their monotonic increase with time, while NOE intensities in NOESY maps pass through a maximum for mixing times of the order of the selective T_1.[41,42]

In order to minimize possible systematic errors arising from the choice of a correlation time, we have only considered *relative* NOE intensities (see below). Moreover, the comparison with X-ray on the one hand and MD-derived distances on the other is not made by translating NOEs into distances but rather by translating distances into predicted NOE intensities and searching for the best match. This procedure is becoming popular among researchers in the field of NMR-derived structures.[41a,c,d]

Table II collects X-ray and MD-derived distances together with experimental NOE intensities (which, as discussed above, in our conditions are essentially proportional to $1/r^6$ through a single proportionality constant). The intensities are divided into five classes (very strong, strong, medium, weak, and very weak), which is an even finer subdivision than presently used in most structural NMR studies.[40f,44] Distances are then translated into these five categories as detailed in Table II; one point to be settled in this process is obviously the scaling involved in this translation. As stated previously, we prefer not to use estimates of the reorientational correlation time as the scaling factor, but rather we use the whole set of data to search for the best match. A posteriori we can state that we do not bias the results in favor of the MD structure over the X-ray structure, since the deviations are roughly equally distributed between over- and underestimates of NOE intensities. In any case, we have also recalculated the reorientational correlation time at 25 °C from the predicted 1D NOE intensities and found a very plausible value of 5×10^{-9} s.

By examining Table II in detail some striking features are clearly evident: for instance, the NOE between $H\beta_2$ Cys 63 and $H\beta_2$ Phe 66 is sizably smaller than that expected on the basis of the X-ray structure distance (1.8 Å). However, in the MD averaged structure the interproton distance is calculated to be 3.0 Å, in agreement with the experimental value. Furthermore, the NOE observed between $H\beta_1$ Cys 43 and Hα Val 73 is too large to be consistent with the H–H distance observed in the X-ray

structure (2.5 Å), and indeed the latter residue during the dynamics moves closer to $H\beta_1$ Cys 43, the averaged distance value being 2.2 Å. Other significant changes involve the geometric relationships between Cys 77 and Tyr 19, or between Cys 63 and Ile 71 (Table II).

For other protons the changes are smaller and, in some cases, within the indetermination of both the X-ray and MD data. In this respect, we may consider that differences smaller than 0.3 Å are not significant.

In any case, it is particularly meaningful that in all circumstances, the MD results are in better or at least equal agreement with the experimental NOEs (Table II). In no case was the agreement found to be worse.

These findings can be critically evaluated in light of the considerations made at the beginning of this section. First, we feel that in some cases even relatively small distance changes are meaningful, although on a relative rather than on an absolute scale. For example, the movement of $H\alpha$ Tyr-19 is small but it is still enough to *increase* one distance and *decrease* another from the same coordinated cysteine. Analogous considerations hold for HN Ala-78 or $H\gamma_1$ Ile-71. In the latter case the agreement may be more qualitative because of the presence of methyl groups with possible internal motions and/or uncertainty in the proton positions. However, it is meaningful that the simulation tends to *equalize* the distances and therefore the NOE intensities, as experimentally observed. In other cases the changes are more striking, as in the case of the Phe-66 residue. Here we observe a large increase of the interresidue Cys-63–Phe-66 distance, which is in excellent agreement with the experimental NOE intensities.

Hence, by comparing the observed NOE and the averaged MD structure we can conclude that the latter is in excellent agreement with the experimental NMR observations. This demonstrates the validity of using solution-phase MD simulations to obtain a better representation of the protein in solution, even as far as the immediate vicinity of the cluster is concerned.

The above results provide strong experimental evidence for the robustness of our MD analysis. We have in our hand a structure of *C. vinosum* HiPIP that represents the solution conformation quite well. Hence, it should be possible to reproduce the backbone structure of other HIPIPs or to predict their structure when unavailable.

The full set of atomic coordinates generated in the present research is available from the authors as a pdb file through E-mail (BERTINI @ IFICHIM.BITNET).

MD Simulation on the Reduced HiPIP. We have also performed a test on the reduced form of HiPIP. In this case, the charge of the extra electron had been equally distributed on the four ions, still with a symmetrical charge distribution. As stated before, an equal electron distribution is consistent with all the available data on the reduced protein.[38,39,45]

We have performed MD calculations with the cluster charges relative to the reduced state starting from the X-ray structure of the oxidized species. In this simulation, the average MD structure remains close to the starting one. This is confirmed by rms deviation and fluctuation values that are very similar to those of the oxidized form (0.91-Å deviation with respect to the starting X-ray structure and 0.65 and 0.46 Å for the rms fluctuations with respect to the MD average structure averaged over the heavy atoms and the backbone atoms, respectively). This insensitivity of the protein structure to the overall charge of the cluster is a further indication that our approach is not critically dependent on this parameter.

Discussion

The present calculations have shown that with a proper set of parameters, MD simulations can be performed with success on metal-cluster-containing proteins. The structures of reduced and oxidized HiPIP show very small differences, and no sizable conformational change is induced by the addition of one electron to the metal cluster. This fact is completely consistent with the function of the protein, i.e. electron transfer.[15] This class of proteins should accept and release one electron with minimal structural perturbations which could interfere in the interaction with other proteins, to which the electron is exchanged.

As mentioned in the Results section, we have chosen somewhat high bending constant values within the cluster. This probably prevented the cluster-bound cysteines, and particularly the β-CH_2 groups, to rearrange completely to the conformation assumed in solution, even though they still experience some mobility. This partial rigidity may have prevented the detection of the fluxionality of the β-CH_2 of Cys 43 which was predicted on the basis of the temperature dependence of the chemical shifts of the corresponding protons.[16,45] This is the only experimental feature which has not been reproduced in the present calculations, while all the other experimental observations fit nicely with the MD simulation. We should stress again that the model of the metal center we use is phenomenological, and furthermore, the time scale of our simulation is relatively short. However, the agreement between our MD distances and those found by NMR is remarkably good. This strongly suggests that refining an X-ray structure with solution-phase MD simulations gives a better representation of the protein in solution.

We expect that the combined use of NMR and MD calculations will further our understanding of the factors determining the conformational arrangement of proteins. Furthermore, we should be able to better correlate observed structural features with function. The success of this treatment suggests that we can predict the structure of homologous proteins for which only the primary structure is available, provided that one X-ray structure is available. NMR can further be used either to refine the structure or to verify the predictions in certain regions of the protein.

Acknowledgment. Thanks are expressed to Prof. P. A. Kollman for his competent comments and to Prof. K. M. Merz for thoroughly reading and discussing the manuscript and for his many suggestions.

(44) (a) Clore, G. M.; Wingfild, P. T.; Gronenborn, A. M. *Biochemistry* **1991**, *30*, 2315–2323. (b) Molinari, H.; Pastore, A.; Lu-yun, L.; Hawkes, G. E.; Sales, K. *Biochemistry* **1990**, *29*, 2271–2277. (c) Havel, F. T.; Wütrich, K. *J. Mol. Biol.* **1985**, *182*, 281.

(45) Bertini, I.; Briganti, F.; Luchinat, C.; Scozzafava, A.; Sola, M. *J. Am. Chem. Soc.* **1991**, *113*, 1237–1245.