

# Molecular Dynamics Study of *Desulfovibrio africanus* Cytochrome $c_3$ in Oxidized and Reduced Forms

Céline Bret,\* Michel Roth,<sup>†</sup> Sofie Nørager,<sup>‡</sup> E. Claude Hatchikian,<sup>§</sup> and Martin J. Field\*

\*Laboratoire de Dynamique Moléculaire and <sup>†</sup>Laboratoire de Cristallographie et Cristallogénèse des Protéines, Institut de Biologie Structurale J. P. Ebel, 38027 Grenoble, France; <sup>‡</sup>Centre for Crystallographic Studies, University of Copenhagen, DK-2100 Copenhagen, Denmark; and <sup>§</sup>Unité de Bioénergétique et Ingénierie des Protéines, Institut de Biologie Structurale et Microbiologie, 13402 Marseille, France

**ABSTRACT** A 5-ns molecular dynamics study of a tetraheme cytochrome in fully oxidized and reduced forms was performed using the CHARMM molecular modeling program, with explicit water molecules, Langevin dynamics thermalization, Particle Mesh Ewald long-range electrostatics, and quantum mechanical determination of heme partial charges. The simulations used, as starting points, crystallographic structures of the oxidized and reduced forms of the acidic cytochrome  $c_3$  from *Desulfovibrio africanus* obtained at pH 5.6. In this paper we also report structures for the two forms obtained at pH 8. In contrast to previous cytochrome  $c_3$  dynamics simulations, our model is stable. The simulation structures agree reasonably well with the crystallographic ones, but our models show higher flexibility and the water molecules are more labile. We have compared in detail the differences between the simulated and experimental structures of the two redox states and observe that the hydration structure is highly dependent on the redox state. We have also analyzed the interaction energy terms between the hemes, the protein residues, and water. The direct electrostatic interaction between hemes is weak and nearly insensitive to the redox state, but the remaining terms are large and contribute in a complex way to the overall potential energy differences that we see between the redox states.

## INTRODUCTION

Cytochromes  $c_3$  are periplasmic 15–16-kDa tetraheme proteins found in sulfate-reducing bacteria. These bacteria are mainly anaerobes. The principal function proposed for these proteins is to be the redox partner of hydrogenases with which they exchange two electrons and two protons, in a pH-dependent oxidation-reduction reaction (Louro et al., 1997). A comparison of the amino acid sequence of cytochromes  $c_3$  from different bacterial species reveals a general low homology of ~30%. Their redox properties are, however, very similar; namely, they evince low redox potentials (ranging between –400 and –100 mV) and show relatively low specificity in their interaction with hydrogenases of different species. Each of the four prosthetic heme groups are covalently bound to the polypeptide chain by two cysteines that are separated by three or four amino acids in the primary sequence, and the fifth and sixth axial ligands of the iron are histidines. The four hemes form a kind of cluster in the  $c_3$  molecule around which the polypeptide chain is wrapped. The three-dimensional atomic structure of this cluster constitutes, with a few small pieces of the polypeptide chain, the part of the molecule that is the most conserved among the cytochromes  $c_3$  (Nørager et al., 1999). This cluster is thus characteristic of all cytochromes  $c_3$ , if one excludes a new type of cytochrome  $c_3$  recently discov-

ered in *Shewanella frigidimarina* bacteria (Gordon et al., 2000; Pessanha et al., 2001).

The *Desulfovibrio africanus* acidic cytochrome  $c_3$ , called *Dafri* in the following text, was discovered by Pieulle et al. (1996) and further investigated by Magro et al. (1997). It is characterized by its poor reactivity with hydrogenases, as compared with the other known cytochromes  $c_3$ . It exhibits a few structural features that distinguish it from the latter (Nørager et al., 1999), such as an exposed heme I surrounded by an acidic surface area and a heme IV lacking most of the surface lysine patch that is suggested to be the site of hydrogenase interaction in typical cytochromes  $c_3$ , and these may explain its change of reactivity with its redox partners. As another tetraheme cytochrome, homologous to *Dafri*, has been isolated from the membrane fraction of *Desulfovibrio vulgaris* by Valente et al. (2001), the authors proposed classifying the acidic cytochromes  $c_3$  in a new type of monomeric tetraheme cytochrome group termed type II cytochromes  $c_3$  (TpII- $c_3$ ), the other cytochromes  $c_3$  being classified in the group termed type I cytochromes  $c_3$  (TpI- $c_3$ ). In a recent study using NMR and visible spectroscopy, an analysis of the oxido-reduction properties of the hemes in *Dafri* was carried out, leading to the conclusion that this protein may have a functional role distinct from the basic cytochromes  $c_3$  (Pereira et al., 2002).

The present study was initiated as a continuation of the work on the comparison of the x-ray three-dimensional atomic crystal structures of *Dafri*, in the oxidized and reduced forms by Nørager et al. (1999). More recently, a few other comparative structural studies on cytochromes  $c_3$  have been published, applied to *Desulfovibrio gigas* by NMR (Brennan et al., 2000) and to *Desulfovibrio desulfuricans*

Submitted April 7, 2002, and accepted for publication September 6, 2002.

Address reprint requests to Dr. M. J. Field, Institut de Biologie Structurale J. P. Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France. Tel.: 33-4-38785919; Fax: 33-4-38785122; E-mail: mjfield@ibs.fr.

© 2002 by the Biophysical Society

0006-3495/02/12/3049/17 \$2.00

ATTC 27774 by x-ray crystallography (Louro et al., 2001). There have also been a number of theoretical studies of these proteins. The sole dynamical simulation to date was published by Soares and co-workers on *Desulfovibrio vulgaris* Hildenborough cytochrome  $c_3$  (Soares et al., 1998), who used the GROMOS force field and program (van Gunsteren and Berendsen, 1990). However, the study suffered from the problem that the authors had to constrain the  $C_\alpha$  positions to conserve the structure of the protein during the (relatively short) simulation. Other studies have focused on predicting the protonation and redox properties of various cytochromes  $c_3$  using continuum electrostatic methods. Examples include the works of Baptista et al. (1999), Ullmann (2000), Louro et al. (2001), and Teixeira et al. (2002).

The aim of the current work has been to try to obtain a more detailed understanding of the properties of *Dafri* in the fully oxidized and fully reduced states. To do this we have performed 4- and 5-ns molecular dynamics simulations of the reduced and oxidized forms of the cytochrome  $c_3$  at pH 5.6, and in parallel, we have carried out an x-ray crystal structure determination of *Dafri* at pH 8 in both redox states (data and results in the Appendix). In our simulation analyses, we have concentrated upon the structural and hydration differences between the two states, the dynamical properties of the various structures, and the interaction energies inside the protein and between the protein and solvent. Our ultimate goal is to identify the factors that contribute to the changes in free energy that occur upon oxidation or reduction of the protein, but we do not tackle this point directly here.

## MATERIALS AND METHODS

All molecular modeling calculations were done with the CHARMM program, version 25b1 (Brooks et al., 1983) and the CHARMM22 force field and parameter set (MacKerell et al., 1998). The dynamics models were based on the crystal structures of *Dafri* at pH 5.6 determined and analyzed by Nørager et al. (1999) and deposited in the Protein Data Bank (Berman et al., 2000) with references 3CAO and 3CAR for the oxidized and reduced forms, respectively.

### Heme modeling and special features of cytochrome $c_3$ modeling

As far as hemes are concerned, CHARMM22 parameters exist only for the reduced heme of hemoglobin. These parameters are not appropriate for our bi-histidinyl heme in either the oxidized or the reduced state, and so we reparameterized the models. The heme partial charges were computed with the quantum chemistry program JAGUAR (Schrödinger, Portland, OR), a density functional theory method with the B3LYP functional (Friesner, 1991; Becke, 1993a,b) and the large LACVP\*\*++ (Hay and Wadt, 1985) basis set that includes diffuse and polarization functions. The partial charges were derived by fitting to the electrostatic potential (Chirlian and Francel, 1987; Woods et al., 1990; Breneman and Wiberg, 1990). Because of computational cost and convergence difficulties, we decided to restrict the charge calculation to the truncated heme model (shown in Fig. 1), where all the side chains, the covalently bound cysteines, and the axial histidines were excluded. The final parameters are given in Table 1.

Previous charge models of bi-histidinyl hemes include those of Martel et al. (1999) who used a Mulliken population analysis and open-shell Hartree-Fock calculations with a minimal STO-3G basis set (Pople and Beveridge, 1970) to evaluate the atomic charges and that of Ullmann (2000) who used a similar approach to ours although on a fuller heme model.

We completed the heme model by adding the excluded side chains and the histidine-heme interactions to the truncated heme model. We took most of the relevant parameters (bond, angle, dihedral, charge, and Lennard-Jones) for these groups from the CHARMM22 hemoglobin model, assuming that they are not redox dependent. The exceptions were the charges and Lennard-Jones parameters for the cysteinyl sulfurs ligating the heme, which we determined by fitting against the results of quantum mechanical calculations on model compounds. We obtained values equal to  $-0.08e$ ,  $0.38 \text{ kcal mol}^{-1}$ , and  $3.950 \text{ \AA}$  for the partial charge, Lennard-Jones well depth, and Lennard-Jones radius, respectively. The parameters for the cyclic N-terminus of the protein were computed in a similar way to those of the heme. Thus, the charges were calculated by fitting to the electrostatic potential generated by quantum mechanical calculations, and the remaining parameters were taken from the CHARMM22 force field.

### Initialization of the model structures from the crystal structures

The completion of the models based on the crystal structures was made in two steps, first by protonation of some titratable residues according to the value of their pKa and second by addition of all H atoms lacking in the crystal structures.

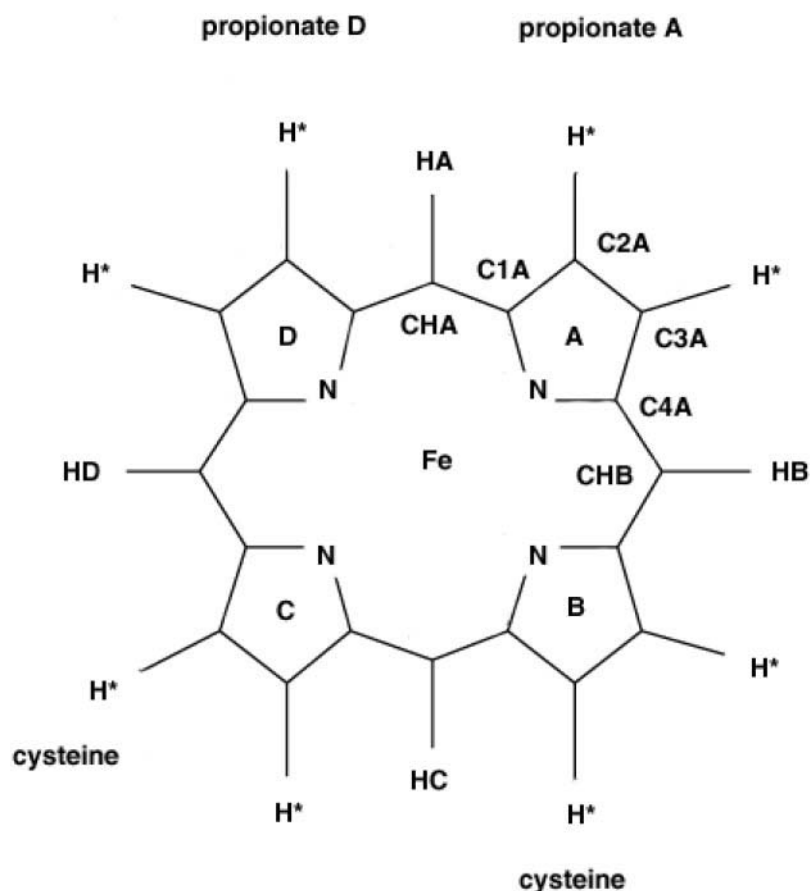
### Protonation state of the titratable residues

We carried out a pKa calculation of all titratable residues of the molecules (Bret, 2000), including the carboxyl groups of the heme propionates and of the residue termini, in both redox states, using a program developed by L. David (Institut de Biologie Structurale, 1995, personal communication) that combined the cluster method of Gilson (1993) and the ionization energy calculation method of Antosiewicz et al. (1994). As far as the acidic residues are concerned, the pKa values ranged from 3.2 to 4.7 and from 4.2 to 5.2 for Asp and Glu, respectively, with minor differences ( $<0.3$ ) between redox states. The pKa values of the propionate carboxyl groups ranged from 4.7 to 6.1, and from 5.2 to 6.5, in oxidized and reduced forms, respectively. Because many of these values are rather close to the pH value of 5.6 of the reference crystallographic structure, they do not give a clear-cut criterion for the protonation state assignment of these acidic groups at this pH. The inspection of the crystal structures at both pH's does not help very much in this respect (see Appendix). No significant difference can be found between the structures at pH 5.6 and pH 8, as far as the chemical coordination between the carboxylic oxygens and protein neighbor atoms is concerned, which presumably means either that there is no change of carboxyl protonation state at the different pHs or that the structures do not contain any strong evidence concerning their protonation states. We assumed therefore that all carboxylic groups were deprotonated, even at pH 5.6, as very likely they are at pH 8. However, in the light of very recent results by Teixeira et al. (2002), this conclusion appears now to be possibly wrong because these authors showed by more accurate pKa calculations that some propionates may have pKa values as high as 8 in *Dafri*. As far as nonacidic residues are concerned, the pKa of the single free histidine His6 was 7.8, and the pKa values of the cyclic N-terminus, of the Tyr78 and of the basic residues were all larger than 10. We therefore assumed that these residues were protonated in our pH 5.6 model.

### Addition of hydrogen atoms

Hydrogen atoms were added to the molecule with the HBUILD module of CHARMM, and the molecule was relaxed by energy minimization with a

FIGURE 1 Truncated heme model used for partial charge calculations with the quantum chemistry program JAGUAR in the oxidized and reduced heme state. Results are given in Table 1. The atoms are named according to the Protein Data Bank convention (Berman et al., 2000).  $H^*$  are fictive H atoms that were substituted for the missing side chains.



conjugate gradient method so as to remove unfavorable steric contacts. This also helps the structure adapt to the CHARMM force field, which is different from the force field of the program used for the refinement of the crystal structures (program REFMAC (CCP4, 1994)). In the first phases of the minimization, the positions of all the atoms, except the hydrogens, were fixed. In subsequent steps, however, all the atoms were free to move although their positions were constrained to remain close to their initial values using harmonic position constraints with progressively decreasing force constant values.

### Solvation of the cytochrome

The different protein models were solvated in a cubic box of dimension 56.6 Å that contained 5832 water molecules. The density of the water was the same as that of liquid water. After superposition of the protein and water box, all water molecules closer than 2.8 Å from the protein were removed. A Langevin dynamics simulation of 55-ps duration was then carried out on the total system following a scheme similar to that of the minimization described above. Thus, at first, the protein atoms were fixed and only the water was allowed to move, but then the protein atoms were free although harmonic position constraints of decreasing force constant were applied. After the simulation, the total system was again superposed with the original water box so as to fill any cavities that had appeared. All new water molecules closer than 2.8 Å from the protein and from the first water molecules were removed. This protocol of simulation and superimposition was repeated until no more water molecules could be added, leading to 4937 and 4929 water molecules in the reduced and oxidized models, respectively. To ensure that the total charge of the box was zero,

we added 21  $Na^+$  ions to the water box for the reduced model and 17 for the oxidized one.

### Molecular dynamics simulations

To simulate the dynamics of the system in the canonical (NVT) ensemble, we used a Langevin algorithm with a reservoir temperature of 300 K and a friction coefficient of  $100 \text{ ps}^{-1}$  for all the atoms. The time step was 1 fs. Periodic boundary conditions were applied. The Lennard-Jones interactions were truncated at 12 Å using a shifted smoothing function whereas the electrostatic interactions were calculated with the particle mesh Ewald method (Essmann et al., 1995), with a 32-point grid for each dimension of reciprocal space and charge densities modeled by Gaussians of width 0.32 Å. Two dynamics simulations were performed, one of 4.03 ns for the reduced state and one of 5.04 ns for the oxidized form. Copies of the time-dependent structures were saved along the simulations every 0.5 ps, giving ~8,000 and 10,000 saved instantaneous structures for the reduced and oxidized models, respectively.

### Time-dependent average structures

To analyze the evolution of the molecular models during the dynamical simulation, we calculated a time-averaged structure of the molecule at different times of the run. The time-averaged structures were calculated as follows. As instantaneous coordinates of the atoms of the molecule were saved in a file every 0.5 ps during the run of the dynamics, we used these data to calculate a moving time-average of the structure, with a moving

**TABLE 1** Partial changes of the oxidized and reduced hemes derived from quantum chemical calculations

Atom	Oxidized	Reduced
FE	1.18	1.14
NA	-0.57	-0.50
NB	-0.73	-0.65
NC	-0.57	-0.50
ND	-0.73	-0.65
C1A	0.46	0.35
C2A	-0.03	-0.08
C3A	-0.01	-0.08
C4A	0.37	0.30
C1B	0.51	0.42
C2B	-0.06	-0.11
C3B	-0.04	-0.09
C4B	0.56	0.44
C1C	0.46	0.35
C2C	-0.03	-0.08
C3C	-0.01	-0.08
C4C	0.37	0.30
C1D	0.51	0.42
C2D	-0.06	-0.11
C3D	-0.04	-0.09
C4D	0.56	0.44
CHA	-0.53	-0.50
HA	0.22	0.20
CHB	-0.44	-0.45
HB	0.20	0.18
CHC	-0.53	-0.50
HC	0.22	0.20
CHD	-0.44	-0.45
HD	0.20	0.18

The atom names follow the PDB standard.

time window length of 128 structures, i.e., 56 ps. Time averaging was applied to the coordinates of each atom of the molecule, after having aligned each new instantaneous structure with respect to the current running time-averaged structure. Only a few atoms of the whole molecule were selected for the alignments, namely, a subset of 124 nonvariable heme atoms from the tetraheme cluster. This selection excluded the atoms of the propionate arms and a few other peripheral heme atoms. This same alignment technique was used by Nørager et al. (1999) for structural analyses of cytochrome  $c_3$  structures, and it is based on the remarkable stability of the tetraheme cluster in the cytochrome  $c_3$  protein family, which we also observed in our dynamical simulations.

### Method of localization of ions and water molecules

The localization of the solvent components was performed in the following way. With the help of two three-dimensional 1-Å grids, one for the water molecules and one for the  $\text{Na}^+$  ions, superimposed on the simulation box, the number of water molecules or  $\text{Na}^+$  ions found in each voxel of the grid over a given period of time was counted. Two series of  $\sim 1000$  structures (corresponding to a time window of  $\sim 500$  ps for each series) located toward the end of the simulation were scanned. The water molecule and  $\text{Na}^+$  ion three-dimensional distributions, represented by these filled grid voxels, were then Fourier transformed and back-transformed (at 1.5-Å resolution) to get time-averaged smoothed water molecule and  $\text{Na}^+$  ion density maps that could then be peak-searched to find the locations of water molecules and ions. For each simulation, a site was classified as a

permanent water or ion site, if it was occupied in at least two different series of averaging.

## RESULTS AND DISCUSSION

### Structural aspects

#### Crystal structures

The main data and new results on the pH 8 structure determination are given and discussed in the Appendix. No major structural difference between pH 5.6 and pH 8 homologous forms are found. The comparison of pH 5.6 and pH 8 structures demonstrates the stability of the intramolecular interactions by changing the pH and the orientational variability, in the solvent region, of a number of solvent-exposed charged side chains. The crystal structures that will be used as a reference for comparison with the simulated structures are the structures at pH 5.6.

#### Evolution of the structure during dynamics

As described in Materials and Methods, the long Langevin dynamic simulations for the oxidized and reduced systems were preceded by preparatory simulations. First, energy minimization of the crystal structures in the CHARMM force field was performed and, then as part of the solvation procedure, three Langevin dynamics runs, each of 55 ps, were carried out to equilibrate the system. The dynamic structures resulting from these simulations showed a global atom coordinates root-mean-square deviation (RMSD) with the crystal structures of  $\sim 1$  Å. An increase of the radius of gyration of 2% was also noticed, which seems to correspond to a homogenous swelling of the structure and to a rearrangement of external loops.

The evolution of the RMSDs of both molecular forms during the subsequent long dynamic runs is presented in Fig. 2, *A* and *B*. From the beginning of the dynamic runs (time  $t = 0$  corresponds to the state of the system after the preliminary runs), the tetraheme cluster seems to have reached a rather stable state characterized by an average RMSD from the crystal structures of  $\sim 0.7$ – $0.8$  Å, with random fluctuations of  $\sim 0.1$  Å around this value. In contrast, the RMSD of the backbone and side-chain atoms are less stable. An initial transient period ranging from 500 to 1000 ps exists, followed by long period fluctuations of the RMSD of  $\sim 0.2$  Å, in addition to the short period fluctuations already observed for the hemes. Similar evolution with time of the RMSD has also been obtained by Sterpone et al. (2001) on lysozyme. The average value of the RMSD reached in both molecular forms, is of the order of magnitude of 1 Å for the backbone and 1.5–1.6 Å for the side chains. (The RMSD of the atoms of the heme cluster, the backbone, or the side chains is calculated after having aligned the heme cluster, backbone, or side-chain atoms of the dynamical model against the corresponding atoms of the



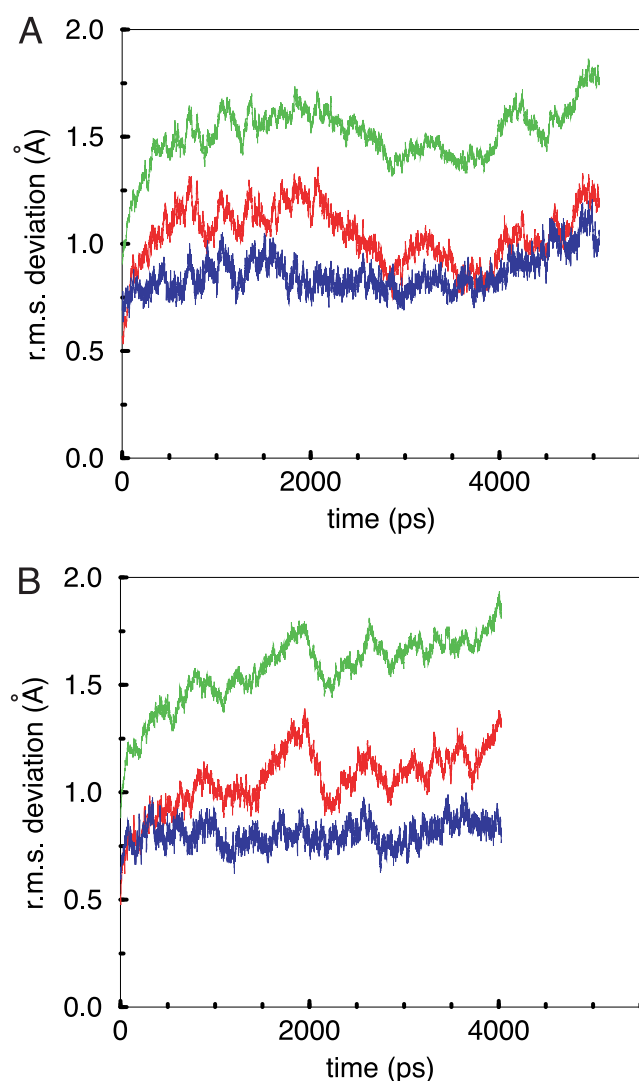


FIGURE 2 Evolution during the dynamics, of the RMSD between the position of homologous atoms of the dynamical model of *D. africanus* cytochrome  $c_3$  and of the crystal structure in the oxidized form (A) and the reduced form (B) (unit = 1 Å). The red, green, and blue curves correspond to the RMSD of the backbone atoms only, the side chain atoms only, and the heme cluster atoms only, respectively. The RMSDs of the selected atoms were calculated after alignment of the model and crystal structures with respect to the selected atoms only.

crystallographic structure.) During the dynamics, fluctuations of the radius of gyration of  $\sim 0.2\%$  could also be observed, probably related to the fluctuations of the conformation of external loops.

The evolution of the molecular models during the dynamical simulation was analyzed in more detail by considering the time-averaged structures of the molecule calculated as described in Materials and Methods. Comparing the time-averaged model structures to the crystal structure on one hand (see Fig. 3 A) and the time-averaged structures with respect to the last averaged one in the dynamics on the other

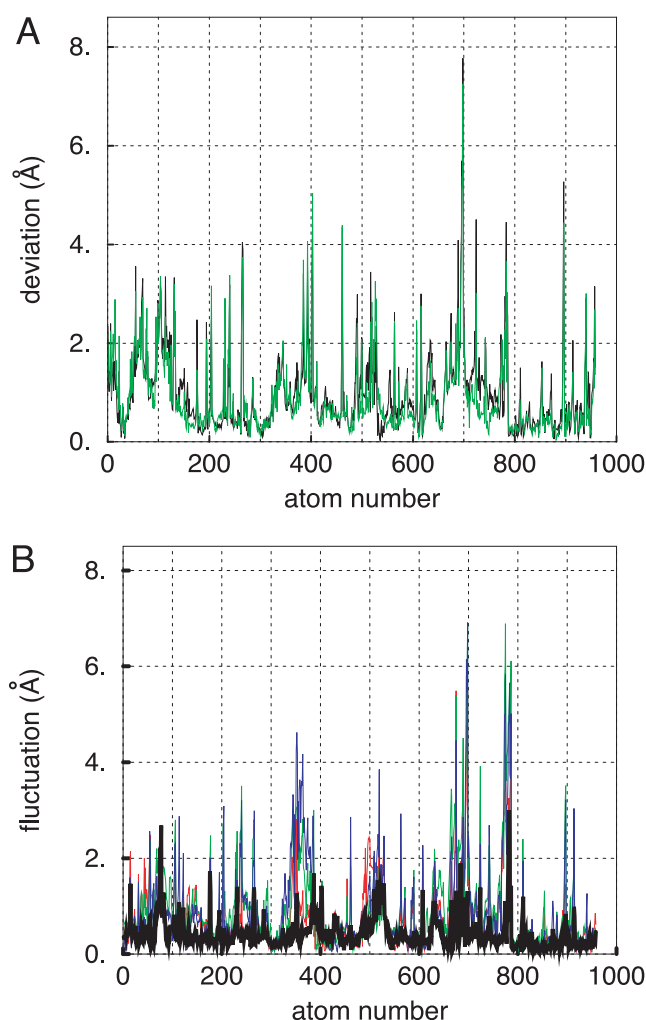


FIGURE 3 (A) Distance in Å for each atom between its position in the crystal structure and its position in the dynamic model (averaged over  $\sim 500$  ps). We present here the results obtained for the oxidized model. The green and black curves give the deviation after  $\sim 3.5$  and 4 ns, respectively, showing that the system has stabilized with respect to the crystal structures. The atom numbering follows that of the structures deposited in the Protein Data Bank (Berman et al., 2000). The atoms of the tetraheme cluster are those on the right side of the diagram, between atoms 787 and 958, and correspond to residues 104–107. The distances, called deviation in the text, were calculated after alignment of the molecules using a restricted set of stable atoms of the tetraheme cluster. (B) Distance in Å between the positions of the homologous atoms in the dynamic oxidized model averaged over  $\sim 500$  ps, after 0.5 ns (red), 1.5 ns (green), 2.5 ns (blue), 3.5 ns (black), and its quasi-stable mean position (bold black) reached after  $\sim 4$  ns. The distances, called fluctuation in the text, were calculated after alignment of the molecules using a restricted set of stable atoms of the tetraheme cluster. The progressive decrease of these fluctuations shows the dynamical stabilization of the system.

hand (see Fig. 3 B) revealed the following results, which are valid for both oxidized and reduced forms.

First, the atom position differences between the time-averaged model and crystal structure (we use the term deviation for this difference from now on) are not uniformly distributed along the molecular chain and on the hemes.

There are regions of high deviations alternating with regions of low deviations (see Fig. 3 *A*). These regions are distributed identically in both redox forms.

Second, these deviations tend to diminish with time and seem to stabilize at the end of the simulation runs (see Fig. 3 *A*), which is true especially in the low-deviation regions. Some residues, however, in the low-deviation regions keep showing large side-chain deviations.

Third, the regions of high and low deviation of the model with respect to crystal structure are also regions of high and low fluctuations of the time-averaged structure (denoted fluctuations from now on) during the dynamics. Deviations and fluctuations are of the same order of magnitude as those in the transient regime. They are much larger than the differences existing between the four crystal structures. These fluctuations tend to decrease toward the end of the simulation run (see Fig. 3 *B*), indicating that a quasi-equilibrium of the structure seems to have been attained after  $\sim 4$  ns. This fact was not clear by considering only the overall RMSD of the whole structure shown in Fig. 2, *A* and *B*.

The behavior of our model during the dynamics is thus characterized by a rather long transient period with large structural fluctuations before stabilization. As will be shown later when looking at the time evolution of the global cytochrome-solvent interaction energy, this transient regime seems to be related essentially to the evolution of the interactions between cytochrome and water molecules, which seem to need much more time to stabilize than was expected a priori. Thus, the programmed preliminary stabilization runs happened to be much too short.

#### Structure of the stabilized dynamical models

The deviations between the model and the crystal structures do not correspond to random atomic displacements. Their amplitude varies continuously from one place to another in the molecular structure. They can be seen as resulting from the application to the crystal structure of a geometrically continuous displacement-distortion field. Their local amplitude reflects the degree of flexibility or rigidity of the local packing of the structure.

As shown in Fig. 4, the low-deviation regions of the molecule between model and crystal structures constitute the kernel of the cytochrome  $c_3$ , formed by the tetraheme cluster and by four protein regions, located in the dihedrals formed by heme III/heme I, heme I/heme II, and heme II/heme IV. It can be seen quite easily that the low-deviation regions correspond, with the exception of the heme IV binding sequence, to the parts of the sequence that are highly conserved in the cytochrome  $c_3$  family (see alignment in Nørager et al., 1999). They contain most of the secondary structure elements of the molecule, namely, the  $\alpha$ - and  $3_{10}$ -helices, with the exception of the  $3_{10}$ -helix Thr9 - Gly13, and the C-terminus  $\alpha$ -helix, which has been shifted

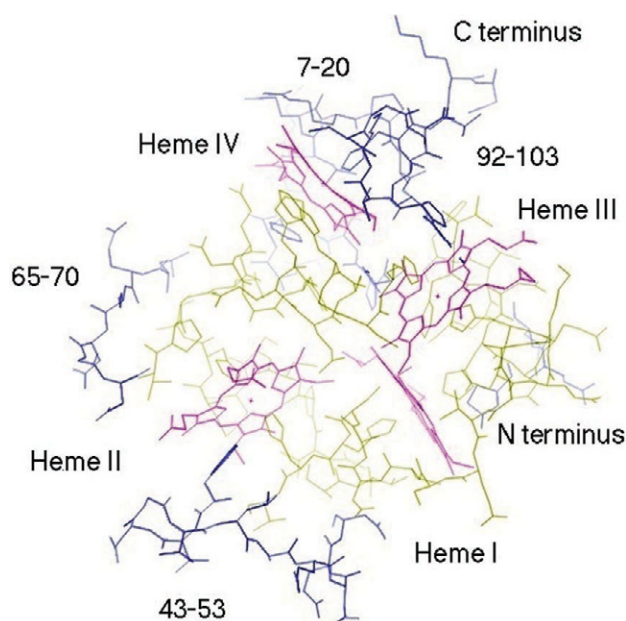


FIGURE 4 Localization on the oxidized form of the dynamical model, of the regions of low and high deviation with respect to the oxidized crystal structure. The low deviation regions (yellow) constitute the kernel of the molecule and occupy the space between the hemes (purple). All high-deviation regions (blue) are represented except residues 87–91, which are not shown and would lie on the front of the picture. The C-terminus region corresponds to the region 92–103.

in the model (Nørager et al., 1999). The deviation is of the order of 0.5 Å if one omits a few of the side chains.

On the contrary, the high-deviation parts constitute peripheral fragments that are in intimate contact with the solvent and for which the deviation of the backbone atoms is larger than 1 Å. These high-deviation regions do not seem to have any particular dominant acidic or basic character. The high-deviation regions are the loops connecting the preceding regions, and the N and C termini. They show large structural fluctuations during the transient regime, proof of their greater flexibility. Being peripheral, they are more sensitive to fluctuations of the interaction with the solvent water molecules and ions.

Crystal and model structures of the tetraheme cluster superimpose rather well. After alignment, the RMSD of the tetraheme cluster of the model with respect to the crystal structure is 0.35 Å if one considers only the 124 heme atoms used for the alignment and 0.7 Å if one considers all heme atoms. A closer examination of the hemes in the model structures reveals, however, that the porphyrin rings are slightly more curved than in the crystal structures. The concavity of this cylindrical deformation is directed toward one of the Fe ligand histidines, its axis being parallel to the plane of that histidine. In addition, the Fe is found to lie asymmetrically between the two histidine Ne2 atoms. The Fe-Ne2 distance is 0.06 Å longer or shorter on the inner or the outer side of the cylinder, respectively, although the

average distance is close to the experimental one. This asymmetry is never observed in the crystal structures. The cylindrical deformation is variable from one heme to another and might be due to strains induced by the deformation of the surrounding protein, revealing thus possibly a lack of heme planar rigidity in our model parameterization.

The propionate conformations are approximately the same in the crystal and the model structures, even if some weak progressive distortion leads to deviations from crystal structures larger than 1–1.5 Å at the end of the propionate arms. There are nevertheless some differences concerning the orientation of the carboxyl groups of propionates A of hemes III and IV in the oxidized form and of heme II in the reduced form. More generally, some aspects of the propionate-protein interaction are different. Thus, as far as heme II is concerned, the propionate carboxyl oxygen atoms O1A and O1D are located, in the crystal structures, within H-bond distance from Trp43 Nε1 and Asp70 OD2 atoms, respectively. In the models, these distances are larger, far beyond H-bond distance. The same is true for the interaction between atom O2D of heme I and atom N of Cys59, which has also disappeared in the models. On the contrary, in the case of heme IV, all but one of the H-bond interactions in the crystal are conserved in the models. For heme III, the models reveal an interaction between the heme carboxyl groups and the neighboring lysines, Lys30 and Lys91, which is not observed in the crystal structures where the propionates interact only with water molecules.

These discrepancies between models and crystal structures might result from an inadequate protonation modeling of some of the propionate carboxyl oxygens (see Material and Methods). Testing this hypothesis would need additional series of dynamic simulation runs.

Calculated Debye-Waller factors (B-factors) of the atoms in the dynamical model  $B_{\text{calc}}$  were deduced from the mean square deviation  $\langle r^2 \rangle$  between the instantaneous and the time-averaged positions of every atom of the structure by the relation (Lonsdale, 1985):

$$B_{\text{calc}} = \frac{8\pi^2}{3} \langle r^2 \rangle \quad (1)$$

The experimental B-factors  $B_{\text{exp}}$  were obtained by subtracting from the crystal B-factors of each atom  $B_{\text{cryst}}$  a constant value  $B_{\text{static}}$  according to:

$$B_{\text{exp}} = B_{\text{cryst}} - B_{\text{static}} \quad (2)$$

The quantity  $B_{\text{static}}$  represents the contribution to the atomic disorder of the static disorder of the molecules in the crystal (molecular or mosaic disorder) and was determined empirically by a best fit of  $B_{\text{calc}}$  against  $B_{\text{exp}}$ . Its value was 6 Å<sup>2</sup> and 16 Å<sup>2</sup> for the oxidized and reduced crystal forms, respectively. With this correction, a quite good agreement is observed between model and measured B-factors (see Fig. 5), except for a small number of solvent-exposed side

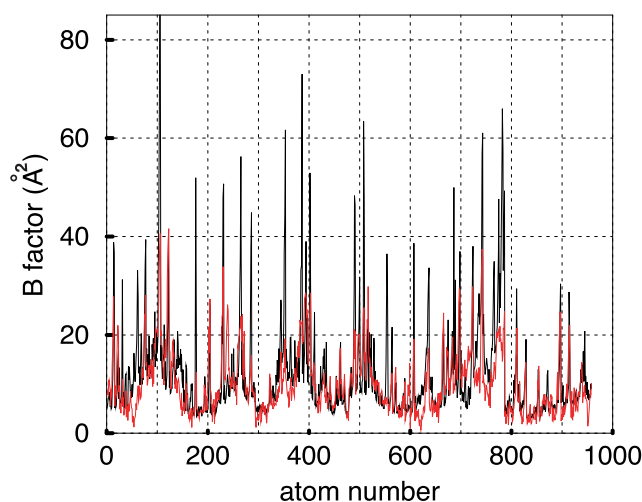


FIGURE 5 B-factors of the atoms in the dynamical model  $B_{\text{calc}}$  (black curve) and crystal structure  $B_{\text{exp}}$  (red curve) (unit = 1 Å<sup>2</sup>) in the oxidized form. The abscissa is the atom number in the structure. The same good agreement between model and measured B-factors is found in the reduced form (data not shown). The agreement fails for a small number of solvent-exposed side chains, where  $B_{\text{calc}} \gg B_{\text{exp}}$  in the oxidized form: Met4, Pro8, Lys14, Asn23, Glu29, Glu34, Asn37, Asn45, Glu50, Glu52, Leu65, Gln67, Leu74, Gln80, Trp83, Gln89, Lys98, Val101, Lys102, and Asn103, and in the reduced form: Glu2, Met4, Asp25, Glu29, Glu34, Asn37, Asn45, Glu50, Leu65, and Gln67.

chains, corresponding almost entirely to acidic and other polar side chains.

By looking at a series of time-averaged structures of the oxidized form, regularly sampled every 50 ps over 500 ps at the end of the dynamics, rather slow oscillations of the structure of some short peripheral residue sequences on this time scale can be seen (see Fig. 6). They concern mainly the sequences Pro8-Asp10, Val44-Glu52, and the C terminus Val101-Asn103, which are part of the flexible and high-deviation regions. These movements are not included in the B-factor evaluation because they correspond to a time scale larger than the time averaging window used for the B-factor evaluation and their amplitude (2–3 Å) would correspond to unrealistically large B values. (As a rule of thumb, one may consider that a region with fluctuations larger than 1.5 Å corresponds to a crystallographically disordered region ( $B$  60 Å<sup>2</sup>).) The significance of these large-scale oscillations is discussed in the Conclusions. Fig. 6 demonstrates also that as far as the dynamics of the molecule is concerned, coherent domain movements in the structure do not exist, if one excludes the peripheral movements just mentioned.

#### Ions and water molecules in the dynamic models

Ions and water molecules play an important role in the model, and so a close examination of their location in the structure deserves attention. Hydration shell and permanent



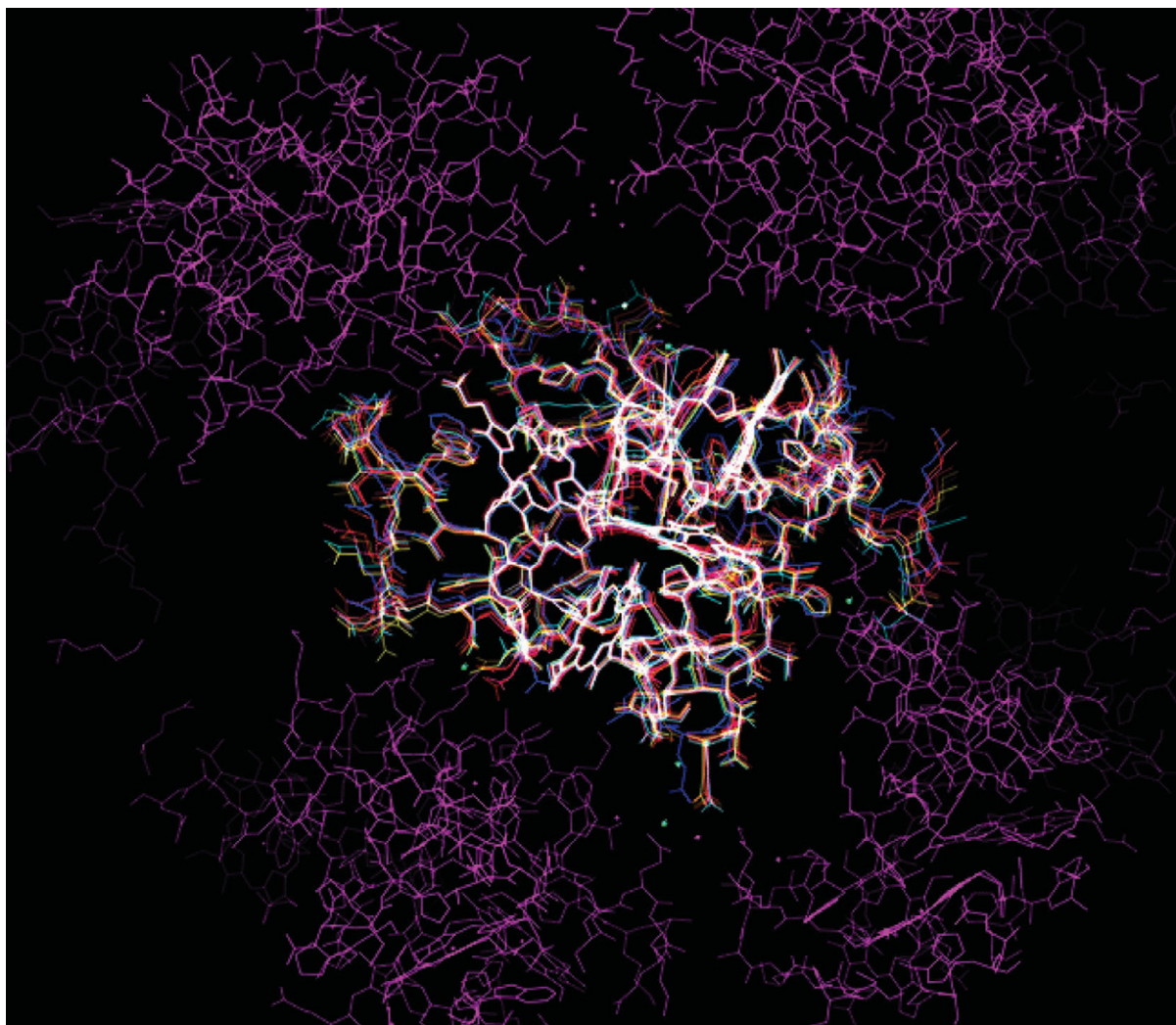


FIGURE 6 Dynamical model and crystal-packing: superimposition of 10 structures of the dynamical model of the oxidized form at the end of the dynamics, each averaged over  $\sim 50$  ps and at successive 50-ps intervals, on top of one of the crystal structure molecules surrounded by its symmetry-related neighbor molecules. All structures were aligned using selected atoms of the tetraheme cluster. The small green spheres represent the Zn and As ions in the crystal structure, which stabilize the crystal packing. This picture shows the slow swinging movements of some peripheral loops and side chains in the dynamical model. Each superimposed structure has been given a different color, and consequently the molecular framework is white where the structures superimpose almost exactly and colored elsewhere. The neighbor molecules of the crystal structure are drawn in purple. It can be seen that some of the mobile regions correspond to intermolecular contact regions in the crystal packing.

water molecules were determined with a technique described in Materials and Methods.

During the transient regime of the dynamics, it was found that the number of water molecules within a given distance from the  $c_3$  molecule increased with time. At the same time, the radial distribution function of water molecules calculated with respect to the center of mass of the  $c_3$  molecule shows a shift of the curve toward the center of the molecule (see Fig. 7). Both effects reflect probably the progressive build-up of the hydration shell and the increase of the solvent-accessible surface (SAS) of the cytochrome (see below). Actually, it is quite conceivable that the technique of inserting the cytochrome in a water box by punching out

water molecules creates a slightly depleted water region at the solvent-cytochrome boundary. The time for water penetration into a protein was shown to be of the order of 1.5 ns in cytochrome *c* by Garcia and Hummer (2000). The long water equilibration time seems thus to be an artifact of the model-building technique.

The number of permanent water molecule sites found in the oxidized and reduced dynamical models are given in Table 2, together with the number of crystal water molecule positions. Of these sites, only  $\sim 10$  coincide (i.e., are at less than 1.2 Å) between the crystal and dynamical structures in the oxidized and reduced forms. Most of these common sites are located in superficial pits of the structures, although



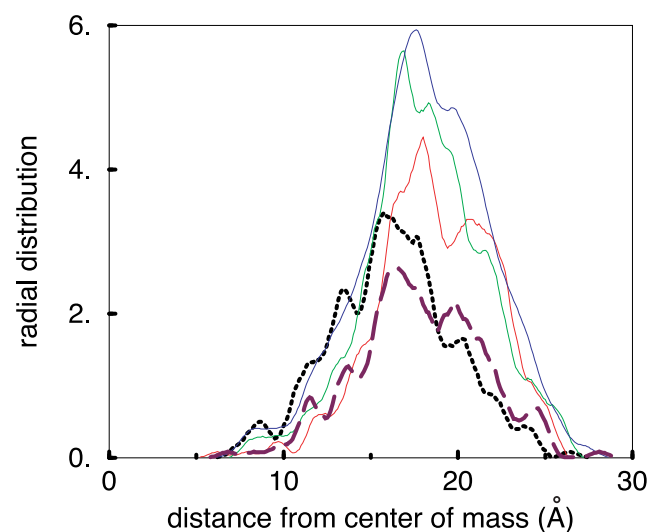


FIGURE 7 Radial distribution functions (RDFs) of the water molecules with respect to the center of the oxidized protein, calculated at different times of the dynamics. The RDF calculation was limited to the water molecules located closer than 2.5 Å from the protein atoms. The diagram represents the RDF of water molecules in the crystal structure (*dotted black*); the instantaneous RDF in the model at the beginning of the dynamics (*red*), after ~0.4 ns (*green*), and after ~4.5 ns (*blue*); and the RDF of permanent water molecule sites at the end of the dynamics (*dashed violet*). It can be seen that there is a progressive penetration of the model by water molecules, until the RDFs of the model and crystal structures coincide. The number of permanent sites of water molecules, however, remains lower in the model than in the crystal structure, because of higher water molecule mobility.

two are buried sites. Radial distribution functions of water molecules, calculated from single instantaneous structures of the dynamics as a function of the distance to the Fe atoms of the hemes, show the occurrence of labile water molecules at the same short distances (5–6 Å) from the Fe, as fixed water molecules were found in the crystal (Nørager et al., 1999) (data not shown). But in the models, fewer permanent sites are found at distances shorter than ~6 Å from the Fe atoms.

Some very interesting properties of the system are revealed by a look at the water density map itself in Fig. 8. The 1 $\sigma$  contour level of the map shows that the whole cytochrome molecule is embedded in a region of higher water density than the bulk solvent, in the sense that the

probability of finding a water molecule there is statistically higher than in the average bulk solvent. This contour delineates a diffuse hydration shell. The 6 $\sigma$  contour level reveals a labyrinthine volume running along the surface of the molecule and represents a kind of water molecule network. Nodes in the network correspond to preferred water molecule sites. The network by itself indicates the presence of a degree of ordering between the water molecules. Many of the numerous water molecules found in the crystal structures, especially in the oxidized form, are found in the model hydration network. Indeed, the water density map at the 6 $\sigma$  contour level compares rather well to the crystal electron density 2Fo-Fc map that was used for finding the crystal water molecules. The latter shows, however, a higher nodosity than the former, i.e., a better definition of the water molecule sites. These results are very similar to the results published by Lounnas and co-workers (Lounnas et al., 1994; Lounnas, 1999), except that in the present case no depletion of the solvent density map is found between the hydration shell and the bulk solvent region.

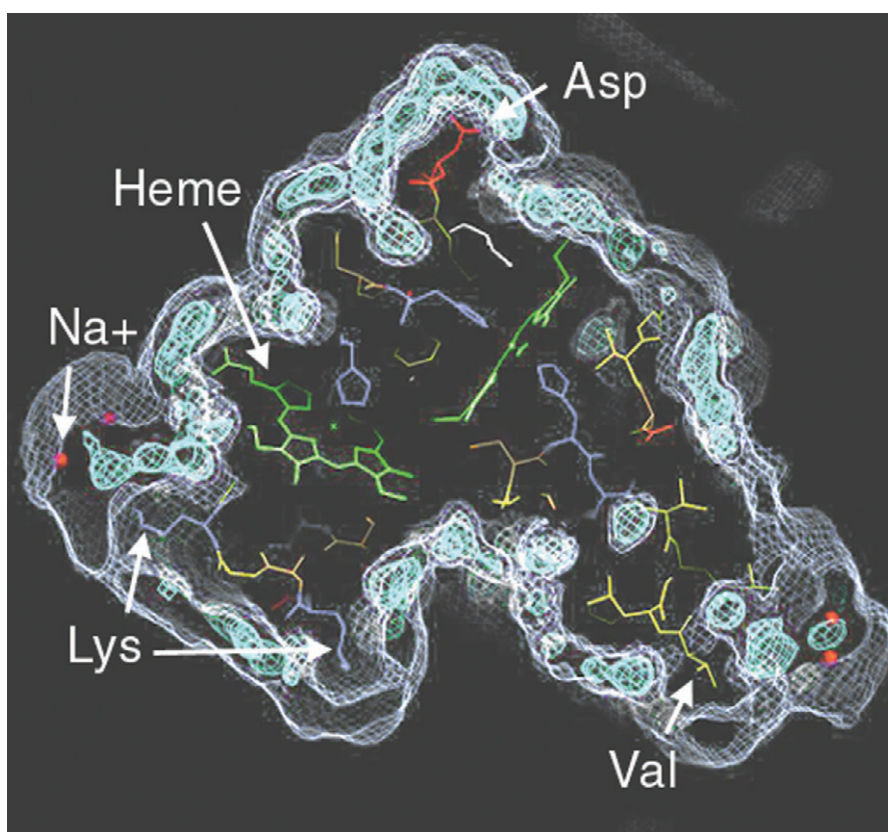
It seems that the covering of the molecule by the hydration network is not uniformly distributed. That does not mean necessarily that the hydration of the molecule is not uniform but more likely that the degree of water ordering is variable. Some residue side chains induce water ordering, whereas others do not. An examination, residue by residue, reveals that the water molecule network is absent in the vicinity of either hydrophobic surface residues (Pro18, Val42, Val44, Val47, Leu48, Ala49, Val55, Gly56, Leu65, Val94, Met95, and Val101) or near solvent-exposed lysines (Lys14, Lys91, and Lys102) and a few other residues (Glu34, Asn45, and Glu66). It seems that the lysines have little effect on water ordering, perhaps because of their long flexible side chains and monopolar terminal group. Sterpone et al. (2001) found also in their dynamical simulations of lysozyme that the density of the hydration shell is lower in the vicinity of hydrophobic residues.

No permanent Na<sup>+</sup> ion sites were found. The ions seem to constitute in the model a rather mobile and disordered system. Most of the peaks in the ion density maps are nevertheless found in a shell between 6 and 10 Å from the molecule. That does not mean that all ions are distributed in this shell but that the likelihood of ion localization is higher in this shell. In the oxidized form, this likelihood seems to be lower in the neighborhood of basic residues, but it is not certain that this result is statistically significant. Only very few ions are found close to carboxylic groups, including the heme propionates, and if so, the distance Na<sup>+</sup>-O<sup>-</sup> is no shorter than 4 Å, with one exception of 2.5 Å. No ions are found at the Zn ion positions of the crystal structures. An interesting fact is that a significant degree of hydration of the Na<sup>+</sup> ions (at 3 $\sigma$ , in the water density map at least; see Fig. 8) is observed around almost all ions.

**TABLE 2** Number of permanent water molecule sites found in the oxidized and reduced forms in the dynamics models and in the crystallographic structures

	Number of sites
Crystallographic oxidized form, pH 5.6	265
Crystallographic oxidized form, pH 8	169
Dynamical model oxidized form	213
Crystallographic reduced form pH 5.6	161
Crystallographic reduced form pH 8	72
Dynamical model reduced form	197

FIGURE 8 Images of the hydration shell around a selected region of the protein with a few  $\text{Na}^+$  counterions, from the water molecule probability density map. The  $3\sigma$  contour of the map (light blue) constitutes a sheet that envelopes completely the molecule and the  $\text{Na}^+$  ions (shown as small red spheres). The  $6\sigma$  contour (light green) has a ramified tubular structure with nodules and represents the hydration network, i.e., what is understood as a network of more or less ordered water molecules. This network is seen in the vicinity of the carboxyl groups of acidic amino acids and propionates and is absent near Lys and Val. Acidic residues are drawn in red, basic in blue, hydrophilic in yellow, and polar in maroon; hemes are green.



#### Structural differences between redox states

The dynamical model seems to be much more sensitive to the redox state of the molecule than the crystal structures. This can be seen clearly in Fig. 9, which displays the atomic position deviations between oxidized and reduced states as a function of the atomic number in the sequence in both

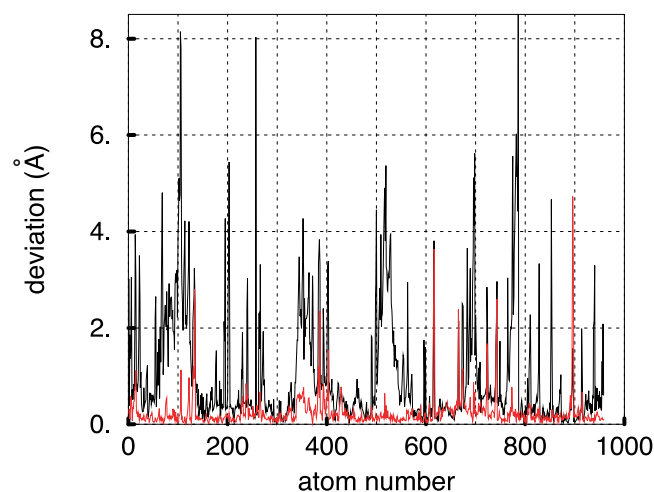


FIGURE 9 Deviation between oxidized and reduced structures for the model mean structures obtained at the end of the dynamics (black) and for the pH 5.7 crystal structures (red). See Fig. 3 a caption for details.

cases. This effect concerns the side-chain reorientations as well as the main-chain shifts. That these large differences found in the simulations are due to the change of redox state and do not correspond simply to the final states of two different pathways of two independent dynamics cannot be formally excluded but is very unlikely, precisely because the initial states (i.e., the crystal structures) are very similar when compared with the final quasi-equilibrium states reached by the dynamical models.

Many side chains, mostly charged ones, show large conformational differences and bond coordination changes (see below). For instance, most of the carboxyl groups of the acidic patch surrounding the solvent-exposed surface of heme I (Glu2, Asp3, Asp25, Glu26, Glu29, Glu34, and Glu50) have their orientation changed. In addition, several extended regions show large deviations between redox states. They are the same regions as those presenting high deviations with respect to the crystal structures described above.

As far as the hemes are concerned, each one shows in the dynamic models different structural modifications associated with the redox change, which will not all be detailed here. A general observation is that many more changes in the conformation of carboxylates of A propionates are observed in the models than in the crystal structures. In the crystal structures, the only case of a conformational change

concerns the propionate A arm of heme III (see Appendix). Another difference between model and crystal structures concerns the interactions of the propionates of heme III with their environment. In the crystal, the propionates interact directly only with water molecules. In the models, these propionates are found interacting with two lysines. In the oxidized form, O1A and O2D interact with  $N_\epsilon$  of Lys30. In the reduced model, Lys30 is shifted away from O2D and replaced by a coordination of  $N_\epsilon$  of Lys91 to O2D. Also in contrast to the crystal structure, the solvent between propionates A and D does not show much structured hydration. In the reduced state, as a consequence of the interaction with Lys91, the hydration of the propionate D seems to be reduced. The discrepancy between model and crystal structures concerning the propionate structures and interactions is possibly related to the question of the protonation state of the carboxylates in the models. It might result from our hypothesis that they are deprotonated in both redox states.

One interesting point is the fact that the dynamics reproduces rather well the rotation of the side chain of Gln81 that is observed in the crystal forms at both pHs, thus confirming the significance of this result. The region where Gln81 is located is a region structurally invariant with respect to the pH in the crystal structures and fairly well conserved in the models with respect to the crystal structures. This amino acid is located near the center of the cytochrome, in the vicinity of propionate A of heme II; of Cys39, ligand of heme I; of His41, ligand of heme II; and of Cys82, ligand of heme III. This region is occupied also by a structured water molecule network, although the water molecule positions in the models and the crystal structures do not coincide exactly. In the oxidized form, the O $\epsilon$ 1 atom of the amido group of Gln81's side chain interacts with carbonyl O of Cys39. In the reduced form, the side chain of Gln81 is more rotated toward the solvent and the amido group makes only H-bonds with water molecules. In the models, in addition, and in contrast to the crystal structures, a concomitant rotation of propionate A of heme II toward the solvent by 90° around its CAA-CBA bond is observed on reduction. These side-chain reorientations are accompanied, in the models as well as in the crystal structures, by a complete reorganization of the water network in this region near heme II. A mutagenesis study of Salgueiro et al. (2001) indeed suggests that the H-bond network around the hemes can affect the redox reaction.

Noticeable modifications of the hydration network are observed between the two redox forms. Only 38 permanent water sites are common to both dynamical structures, whereas there are 110 water molecules common to the oxidized and reduced crystal structures at pH 5.6. Locally, the hydration network follows the shifts or reorientations of residue side chains (especially carboxyl or amido groups) whose conformation is changed. An example is shown in Fig. 10. But more global changes of the distribution of the branches of the hydration network are also observed. The

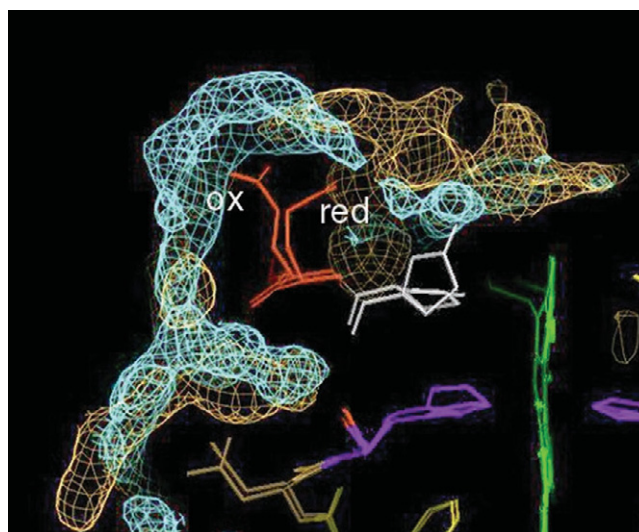


FIGURE 10 Shift of the hydration network in the model structures, accompanying the reorientation of the carboxyl group of Glu2 by reduction. On the picture can be seen heme I and its ligand His24. The color coding is the same as in Fig. 8. The hydration network is represented by the 6 $\sigma$  contour of the water density map, the light-blue contours and yellow contours corresponding to oxidized and reduced forms, respectively. The flipping of the cyclized N-terminus (white) by reduction can be seen, the cycle being perpendicular or parallel to His24 in the oxidized and reduced forms, respectively.

organization of the water molecules is deeply modified around the molecule by reduction, in a complex way, which is difficult to describe in detail. This modification of the ordering of the water molecules reflects probably the redistribution of the electrical field around the cytochrome, induced by the change of redox state. The average degree of water molecule ordering seems to remain about the same in both forms.

### Energetical aspects

More understanding of the difference between the two redox states can be gained by analyzing different aspects of the interaction energies in the system calculated by CHARMM. It should be emphasized that all the interaction energies presented here are potential energies and so cannot be directly equated to the free energies of the different redox states of the protein. The order of magnitude of different energies as calculated by CHARMM is given in Table 3 in kcal mol<sup>-1</sup>. As can be seen, the energy fluctuations of the whole system are dominated by the fluctuations of the solvent self-energy, and the solvent-protein interaction energy constitutes an important contribution to the total energy. One can also see that the self-energy of the cytochrome is lower in the oxidized state, which means that it is more stable in vacuum than the reduced one.



**TABLE 3** Evolution of energy terms during the dynamics

	Stages of dynamics	Beginning	Middle	End
System total energy	ox	$-36734 \pm 427$	$-37268 \pm 410$	$-37611 \pm 422$
	rd	$-36878 \pm 403$	$-37607 \pm 392$	$-37653 \pm 404$
Solvent self-energy	ox	$-34475 \pm 432$	$-35180 \pm 415$	$-35198 \pm 428$
	rd	$-34657 \pm 407$	$-34972 \pm 398$	$-35272 \pm 411$
$c_3$ Self-energy	ox	$1137 \pm 35$	$1147 \pm 33$	$1144 \pm 36$
	rd	$1248 \pm 33$	$1268 \pm 33$	$1251 \pm 34$
$c_3$ -Solvent interaction energy	ox	$-3396 \pm 55$	$-3443 \pm 55$	$-3557 \pm 63$
	rd	$-3469 \pm 51$	$-3695 \pm 58$	$-3632 \pm 65$

Each energy value ( $\text{kcal mol}^{-1}$ ) is the mean value of the corresponding energy calculated by CHARMM, averaged over  $\sim 500$  ps at the beginning, the middle, and the end of the dynamics, given with its corresponding root mean square value. ox, oxidized; rd, reduced.

### Transient regime

By looking at the variation of the different interaction energies during the transient regime, it was found that there is an evolution of the potential energies of the system and that the energies that show the main variations with time are the solvent self-energy and the cytochrome-solvent interaction energy. This is shown in Table 3, which gives the variation of the different energies between the beginning and the end of the simulation. From Table 3 it can be deduced that the fluctuations of the structure of the cytochrome, which are characteristic of the transient regime described above, are related to the evolution of the solvent and its interaction with the cytochrome and that they are not due to the minimization of the cytochrome bond energy by relaxation of residual internal stresses (the  $c_3$  self-energy remaining almost constant during the dynamics). Three factors can contribute to the effect, namely, 1) the mutual local structural adaptation of the water molecules and the cytochrome, 2) the slight increase of the cytochrome SAS by the change of the configuration of solvent-exposed regions of the cytochrome (see Table 4), and 3) the increase of the number of water molecules at the solvent-cytochrome boundary by progressive hydration of that initially depleted region.

### Detailed interaction energy analysis

The analysis of the interaction energy between hemes and residues, detailed residue by residue and heme by heme,

**TABLE 4** SAS of crystallographic and dynamical model structures of *Dafri* cytochrome  $c_3$ , in  $\text{\AA}^2$ , calculated with CCP4 (1994) program IMOLAREA

	Crystal structure	Dynamic model			
		1	2	3	4
SAS of oxidized form	6246	6429	6584	6644	$6600 \pm 80$
SAS of reduced form	6246	6501	6677	6665	$6695 \pm 55$

Columns 1–4 correspond to different stages of the dynamical model structure: column 1, after end of preparation runs; column 2, average during first 500 ps of dynamics; column 3, average at end of dynamics; and column 4, overall average and root mean square of fluctuations during the dynamics.

sheds an interesting light on the cytochrome's structural properties. These interaction energies are represented in Fig. 11 *A*, which shows that the regions constituting the kernel of the molecule described above are those that interact directly with the hemes: the regions 4–6, 21–42, and 54–64 interact each with two hemes, and the region 71–86 interacts with the four hemes and can therefore be considered as the center of the molecule, especially the very central region Tyr78-Cys82. It shows also that the less stable regions 7–20 and 87–103 interact strongly with heme IV, in contrast to the less stable regions 43–53 and 65–70, which do not interact with any heme. In Fig. 11 *B*, it can be seen that the residues that exhibit a noticeable variation of interaction energy with the hemes on changing the redox state may be classified in four groups: 1) the Fe ligand histidines (His24, His27, His40, His41, His63; His79, His86, and His100), which exhibit relatively large and uniform interaction energy variations by reduction; 2) a few residues showing large variations because of change of coordination ((i) Lys30 and Lys91, which change their coordination with heme III propionate carboxylates A and D, (ii) Tyr78, which forms, in the reduced form, a H-bond with the heme I carboxylate D, and (iii) Arg17, which in the oxidized form interacts with Glu16 (interaction between Arg17-N $\zeta$ 1 and Glu16-O $\epsilon$ 2) and in the reduced form is shifted closer to heme I (near methyl group CMB) while Glu16's side chain is fully immersed in the solvent); 3) the region 71–86, where a series of residues from the kernel, i.e., close to the hemes, show small energy variations; and 4) the region 6–21, where the residues show also small energy variations, together with a rather large position shift, apparently driven by the attraction of Arg17 by the reduced heme I.

Two main conclusions can be drawn from these results. The first one is that the residues whose interaction with the hemes is varying strongly with the redox state are mainly basic residues like Lys and Arg. This result may have some meaning in the case of the numerous basic cytochromes  $c_3$ , where a patch of lysines are found in the vicinity of heme IV. A recently published example of such a Lys configuration change by reduction in the heme vicinity has been given by Brennan et al. (2000) in the case of cytochrome  $c_3$  of *Desulfovibrio gigas*. The second one is that the numerous

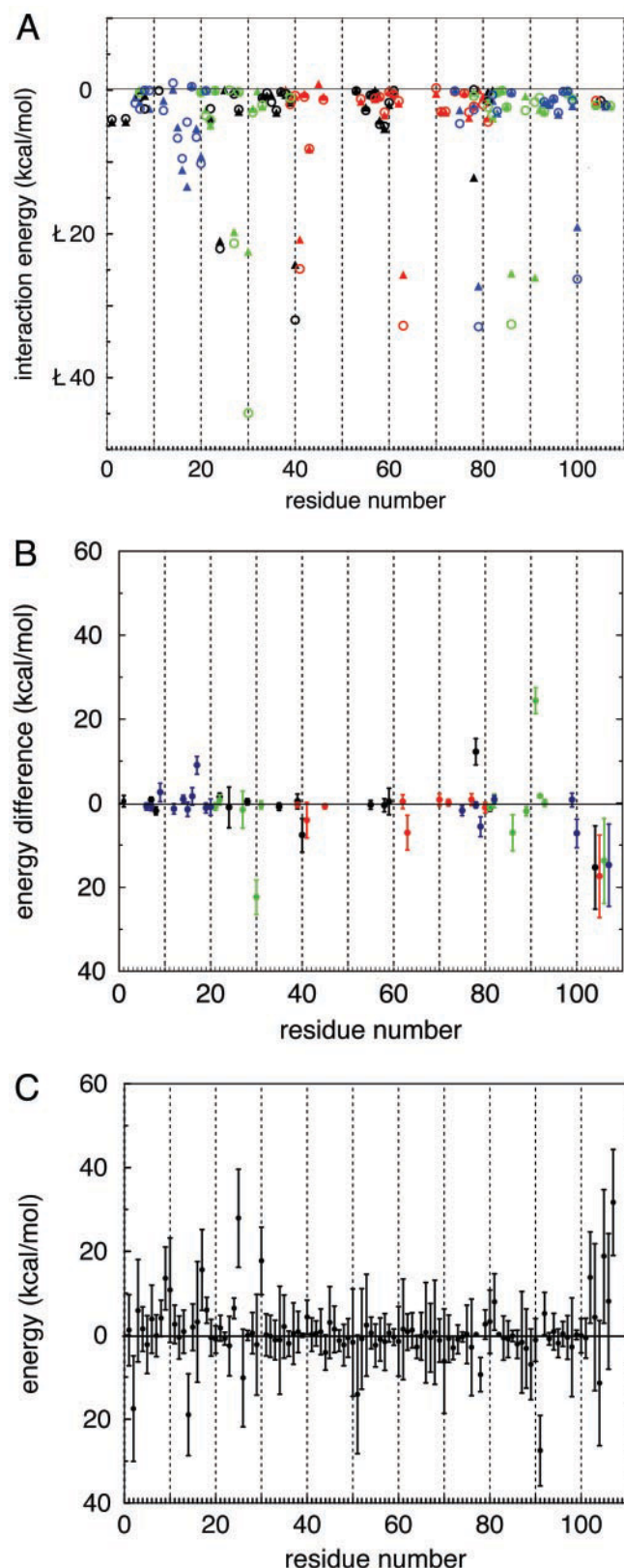


FIGURE 11 (A) Diagram, residue by residue, of the time-averaged and fluctuations of the heme-residue interaction energies (unit = 1 kcal mol<sup>-1</sup>) calculated by averaging over ~0.5 ns at the end of the dynamic simulations. Circles and triangles correspond to the interactions in the oxidized

acidic residues that are characteristic of this structure do not participate directly in these interactions.

The difference of the residue-solvent interaction energy for each residue between the two redox states is given in Fig. 11 C. In contrast to the case of the heme-residue interaction, the acidic amino acids are the most concerned by this interaction, in addition to a few basic or polar amino acids. Many of the side-chain conformational changes mentioned in the section discussing the structural differences between redox states in the models do involve large residue-solvent interaction energy variations. The sign of the energy variations that are observed is variable and seems not to depend on the charge of the residues. After close examination of the related structural changes, these energy variations seem to be primarily related to: 1) a change in the H-bond coordination of the side chain, 2) a change of the water ordering in the vicinity of the side chain, and 3) a covering/uncovering of part of an amide or a carboxylate group.

In case 1, the rule is that the value of the residue-solvent interaction energy is decreased when the change of redox state induces the breaking of one or more residue-residue side-chain H-bonds existing in one state, replaced by residue-solvent interactions in the new state. This case applies to residues Lys14, Glu16, Arg17, His24, Asp25, Lys30, Tyr78, and Lys91 and to the propionate D molecules of heme I and heme III. This change induces probably also an increase of configurational entropy, by local increase of the disorder of the system.

In case 2, the value is decreased in the redox state where the solvent shows higher ordering in the vicinity of the residue side chains, which can be inferred from a higher number of permanent water sites or from the structure of the surrounding hydration network density map. In effect, more stable H-bonds should correspond to lower time-averaged interaction energies. This seems to be the case for Glu2, Pro18, His40, Asp51, and heme II.

Case 3 is invoked for Gln81 and propionate D of heme IV, where the structural change, which can explain the energy variation, is the uncovering, in the reduced form, of the oxygens Oe1 of the amide group of Gln81 and O1D of the carboxyl group of propionate D of heme IV, allowing

and reduced forms, respectively. The color coding corresponds to interactions with heme I (black), heme II (red), heme III (green), and heme IV (blue). The points at the right of the picture correspond to hemes I, II, III, and IV (residue numbers 104, 105, 106, and 107, respectively). Not shown are the residues that have vanishing interaction energy with the hemes. (B) Difference between oxidized and reduced forms of the time-averaged interaction energies between hemes and residues shown in *a* (unit = 1 kcal mol<sup>-1</sup>), with estimated RMSD. Comments and color coding are the same as for *a*. (C) Difference, residue by residue, of the time-averaged residue-solvent interaction energy between oxidized and reduced states (unit = 1 kcal mol<sup>-1</sup>). The error bars represent the estimated root mean square time fluctuations of these energy differences.

**TABLE 5** Heme energy (in kcal mol<sup>-1</sup>) in both oxidized (ox) and reduced (rd) forms calculated at the end of the dynamics by CHARMM and averaged over 500 ps

Heme energy:		Self-energy	Self-energy + interaction with residues	Self-energy + interaction with residues and solvent
Heme I	ox	66.43	-35.0	-198.2
	rd	81.74	-24.3	-176.0
	ox-rd	-15.31	-10.7	-22.2
Heme II	ox	65.12	-34.9	-204.8
	rd	82.53	-7.0	-195.8
	ox-rd	-17.41	-27.9	-9.0
Heme III	ox	68.31	-70.6	-198.8
	rd	82.04	-48.6	-184.2
	ox-rd	-13.73	-22.0	-14.6
Heme IV	ox	68.83	-58.4	-151.2
	rd	83.59	-39.0	-163.6
	ox-rd	-14.76	-19.4	+12.4

For discussion, see text.

the creation of new H-bonds with water molecules. Indeed, Gln81 could be considered as almost belonging to the first case because, in the oxidized state, the distance O $\epsilon$ 1(Gln81)-O(Cys39) is 3.45 Å and could correspond to a H-bond.

These three rules are not exhaustive because they do not explain all observed energy variations, for instance, for Thr9, Asp10, Glu26, Glu29, Gly92, and Lys102.

As far as solvent exposure is concerned, we found no correlation between variations in the SAS of these residues and their interaction energies with solvent. What is involved is not the raw SAS variation but the possibility of the formation or annihilation of H-bonds with water molecules.

Table 5 shows for the four hemes different values of the heme potential energy variation by reduction as calculated by CHARMM, namely 1) the heme self-energy variation, 2) the heme self-energy variation + heme-residue interaction energy variation, and 3) the heme self-energy variation + heme-residue interaction energy + the heme-solvent interaction energy variation. The dispersion of the energy values between hemes increases when adding these interaction contributions. Both types of interactions depend on very contingent (i.e., nonconserved) local structural features such as 1) the occurrence of two lysines in the vicinity of the propionates of heme III or the formation of a H-bond on the OH of Tyr78 in the reduced state and 2) variations of the heme-water interaction. This may explain the great diversity of redox potentials existing among the cytochromes *c*<sub>3</sub>. But in addition to the two direct contributions to heme potential energy variation considered here, our evaluation shows that there are also nonnegligible indirect contributions to the redox potential, especially from the residue-solvent interaction energy variations that have same order of magnitude as the first ones. The case of Gln81 can be given as an example. Gln81 shows a change of configuration that seems

to be characteristic of *Dafri* reduction but that does not produce any variation of direct Gln81-heme interaction energy, although it produces a rather large Gln81-solvent interaction energy variation. The residue-solvent interaction energy variations must be considered indeed as global contributions to the redox potentials, because none of them can be attributed to any particular heme. As a general conclusion, it can be said that even though direct simulation does not permit the evaluation of the hemes' redox potential, it does permit the identification of some of the important interactions and their energy contributions, which have to be taken into account for doing so.

## CONCLUSIONS

Qualitatively, the dynamical modeling provides an informative picture of the structures and the energetics of cytochrome *c*<sub>3</sub> in its fully oxidized and fully reduced states. The main results of the dynamical simulations can be summarized as follows:

First, the dynamical model leads to a stable system, free of ad hoc stabilization constraints. With respect to a preceding similar dynamical simulation by Soares et al. (1998), this achievement is probably related to a more consistent treatment of long-range electrostatic interactions. The simulation requires several nanoseconds to reach a dynamical equilibrium. This rather long stabilization time seems to be a consequence of a combined effect of water-cytochrome interaction and external loop oscillations. The water-cytochrome interaction takes a long time to reach equilibrium. Similar behavior, related to the time required for penetration of the water into the structure coupled to slow structural fluctuations of protein loops, has been noted before (see, for example, Garcia and Hummer, 2000; Sterpone et al., 2001).

Second, the stabilized structures agree reasonably well with the crystal structures. The agreement is rather good for the tetraheme cluster and its neighborhood, but there are also regions with large deviations between both kinds of structures. For instance, the side chains of the amino acids of the acidic patch identified around heme I in *Dafri* by Nørager et al. (1999) all show large deviations. The redox-linked structural changes in the neighborhood of hemes I and II, namely, the reorientation of Gln81 and the modification of the hydration network, is well reproduced by the simulations. This structural result pointed out by Nørager et al. has been confirmed by the NMR study of Pereira et al. (2002), suggesting a role in the redox reaction. Among the permanent water molecule sites predicted by the model, only a few overlap with the crystal water molecules. A very good agreement is obtained between the crystal structure *B*-factors, after subtraction of the mosaicity contribution, and model *B*-factors, except for peripheral side chains.

Third, the model provides a detailed description of the structure of the solvent around the cytochrome, in the form of a hydration network and a number of permanent water



molecule sites. The topography of the hydration network around the molecule is shown to be very sensitive to the redox state of the cytochrome.

Fourth, Nørager et al. (1999) proposed the hypothesis that the remarkable structural stability of the tetraheme cluster in all cytochromes  $c_3$  resulted from a direct specific heme-heme interaction. However, our energetic analysis shows that the direct heme-heme interaction is almost entirely due to van der Waals forces, is not very strong, being of the order of  $-2 \text{ kcal mol}^{-1}$ , and is almost independent of the redox state. In the experimental study of Pereira et al. (2002), they observed no positive redox cooperativity between the hemes for this acidic cytochrome, in contrast to the type I cytochrome  $c_3$ . Even if these results cannot be directly compared, it might explain the weak interaction energies between the hemes observed in the simulations.

Fifth, analysis of the various interaction and self-energies contributing to the potential energies of the dynamical structures illustrates the sorts of structural change that are energetically important when the redox state of the protein changes. These changes occur in the protein, in the solvent, and in the interaction between them. The analysis suggests thus different ways in which the bi-histidinyll cytochromes  $c_3$ , despite their large variety of sequences and peripheral structures, can modulate their redox potentials to have similar magnitudes. Of course, calculation of the redox potentials directly from simulations of the type we have done here is impractical and requires special techniques, such as those used by Baptista et al. (1999), by Ullmann (2000), and by Louro et al. (2001) in their works on cytochrome  $c_3$ .

Although qualitatively the simulations agree with, and complement, the experimental results on many points, there are some quantitative differences. The main one concerns the greater mobility or flexibility of the simulation compared with the crystal structures, which is revealed by a much larger side chain mobility in the models, much larger structural differences between oxidized and reduced model forms, and a reduced number of permanent water sites inside the kernel of the cytochrome  $c_3$  molecule.

These discrepancies could be due to a number of factors. The first one that can be invoked is that our simulations were performed for (effectively) isolated cytochrome molecules in solution whereas the crystal structures correspond to molecules constrained by the crystal-packing interactions. It is clear from Fig. 6 that the mobility of certain surface residues of the model protein is incompatible with the molecular crystal packing. Even if compatible, these regions, with such a high mobility, would be too disordered to be resolved at atomic resolution by crystallography. It is conceivable that the intermolecular interactions in the very highly concentrated molecular solutions tend to quench loop oscillations, thus permitting the occurrence of crystallization although it is more usually assumed that molecules with disordered regions do not crystallize. Therefore, the discrepancies between experimental and simulated structures result

more likely from imperfections in the force field, the simulation procedures, and the difficulty of simulating proteins directly with variable protonation states for their titratable residues (see Lounnas, 1999, for a discussion of the validity of some of these approximations).

## Supplementary material

The coordinates of *Dafri* crystal structures at pH 8 (files *Dafri.oxi.pH8.pdb* and *Dafri.red.pH8.pdb* for the oxidized and reduced forms, respectively) are deposited in Protein Data Bank format (Berman et al., 2000) as supplementary material.

## APPENDIX

### Crystal structure of *D. africanus* cytochrome $c_3$ at pH 8

As part of the present work, two new crystal structures of *Dafri* were solved at pH 8 to study the influence of pH on the structure of the oxidized and reduced forms (S. Nørager and M. Roth). The crystals used for this determination were grown at pH 5.6 in air, i.e., in the oxidized form, with the technique published by Nørager et al. (1999). They were then soaked in a buffer at pH 8 for several hours. The crystals did not show any visible sign of degradation. The reduced form was prepared subsequently by dithionite reduction of the crystal followed by flash cryocooling at liquid nitrogen temperature, in anaerobic conditions with the method described by Nørager et al. (1999). The crystals were kept cryocooled for x-ray diffraction data collection on the beamline BM30A/FIP at the European Synchrotron Radiation Facility (Grenoble, France). The crystals diffracted to  $\sim 2\text{-}\text{\AA}$  resolution. The oxidized and reduced structures were solved with programs WARP and REFMAC from the CCP4 suite of programs (CCP4, 1994), starting from the known pH 5.6 isomorphous crystal structures after having discarded the water molecules. The structure models were progressively corrected with the graphical program O (Jones et al., 1991), after successive REFMAC refinements, by inspection of the weighted  $2F_o - F_c$  and  $F_o - F_c$  maps calculated by REFMAC. The refined oxidized structure corresponds to  $R = 0.217$  and  $R_{\text{free}} = 0.248$  with  $\sim 180$  water molecules, and the refined reduced structure to  $R = 0.242$  and  $R_{\text{free}} = 0.272$  with  $\sim 160$  water molecules. The final standard  $R$  and  $R_{\text{free}}$  factors were not as good as they were for the pH 5.6 crystal structures. The structural accuracy at pH 8 is thus lower than it was at pH 5.6. This might result from a slight decrease of crystal quality by soaking the crystals in the buffer at pH 8.

No basic structural difference between pH 5.6 and pH 8 homologous forms were found. The main results are the following. 1) Many water molecule positions and most of the Zn ion positions are identical, but at pH 8 no As atoms are present because of the change of buffer (Tris-HCl instead of cacodylate). 2) At pH 8, in contrast to the pH 5.6 structures, no multiple side-chain configurations were found, but there were a higher number of disordered peripheral side chains, especially in the reduced form. 3) By looking at the four available structures, the following residues have side chains with variable orientation (a structural variability starting usually at the level of C $\gamma$  along the side chain): Glu2, Lys14, Glu16, Arg17, Glu26, Glu29, Lys30, Glu34, Val44, Glu50, Glu87, Lys88, Lys91, Val94, Met95, Glu98, and Lys102. Most of them are solvent-exposed charged amino acids. 4) In the reduced form at pH 8, in the region between Trp83 and Met95, a shift of the backbone of  $\sim 0.5 \text{ \AA}$  is found with respect to the oxidized one. In this region also, Trp83 is flipped by  $\sim 70^\circ$  with respect to its orientation in all other structures, being closer to the BC side of heme IV and parallel to His79, which is one of the ligands of heme IV Fe, instead of being parallel to the heme plane. 5) At pH 5.6, the propionate

A of heme III showed a different orientation in the oxidized and reduced forms, a fact initially attributed to the change in redox state. At pH 8, this propionate has the same orientation in both forms, which is intermediate between those at pH 5.6. This observation tempers somewhat our previous conclusions about the strict relationship between redox state and propionate or amino acid carboxylate orientation (Nørager et al., 1999). The same point of view has also been recently expressed by Louro et al. (2001). 6) At both pH 8 and 5.6, the same change of orientation of Gln81 is observed as a function of the redox state. This change is also found in the molecular dynamics simulation and is discussed in the main text.

By comparing the structures at both pH values, it can be concluded that the heme-protein interactions seem not to be modified by the change of pH, or at least no structural changes are observed which could give this indication; on the contrary, the configuration of the solvent-exposed charged side chains seems to be sensitive to pH.

We thank the Commissariat de l'Énergie Atomique and the Center National pour la Recherche Scientifique for support of this work.

## REFERENCES

- Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1994. Prediction of pH dependent properties of proteins. *J. Mol. Biol.* 238:415–436.
- Baptista, A. M., Martel, P. J., and C. M. Soares. 1999. Simulation of electron-proton coupling with a Monte Carlo method: application to cytochrome  $c_3$  using continuum electrostatics. *Biophys. J.* 76: 2978–2998.
- Becke, A. D. 1993a. A new mixing of Hartree-Fock and local density functional theories. *J. Chem. Phys.* 98:1372–1377.
- Becke, A. D. 1993b. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* 98:5648–5652.
- Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.
- Breneman, C. M., and K. B. Wiberg. 1990. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* 11(3):361–373.
- Brennan, L., D. L. Turner, A. C. Messias, M. L. Teodoro, J. LeGall, H. Santos, and A. V. Xavier. 2000. Structural basis of the network of functional cooperativities in cytochrome  $c_3$  from *Desulfovibrio gigas*: solution structures of the oxidized and reduced states. *J. Mol. Biol.* 298:61–82.
- Bret, C. 2000. Modélisation de la polarisation dans les champs de forces moléculaires et étude des structures redox du cytochrome  $c_3$  de *Desulfovibrio africanus* par dynamique moléculaire. Ph.D. thesis. Université Joseph Fourier, Grenoble, France.
- Brooks, B., R. E. Bruccoleri, B. B. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. A program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* 4:187–217.
- CCP4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Cryst. D.* 50:760–763.
- Chirlian, L. E., and M. M. Francl. 1987. Atomic charges derived from electrostatic potentials: a detailed study. *J. Comput. Chem.* 8:894–905.
- Essmann, U., L. Perera, L. M. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
- Friesner, R. A. 1991. New methods for electronic structure calculations on large molecules. *Annu. Rev. Phys. Chem.* 42:341–367.
- Garcia, A. E., and G. Hummer. 2000. Water penetration and escape in proteins. *Proteins Struct. Funct. Genet.* 38:261–272.
- Gilson, M. K. 1993. Multiple-site titration and molecular modeling: two rapid methods for computing energies and forces for ionizable groups in proteins. *Proteins.* 15:266–282.
- Gordon, E. H. J., A. D. Pike, A. E. Hill, P. M. Cuthbertson, S. K. Chapman, and G. A. Reid. 2000. Identification and characterisation of a novel cytochrome  $c_3$  from *Shewanella frigidimarina* that is involved in Fe(III) respiration. *Biochem. J.* 349:153–158.
- Hay, P. J., and W. R. Wadt. 1985. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* 82:299–310.
- Jones, T. A., J.-Y. Zou, S. W. Cowan, and M. Kleywegt. 1991. Improved methods for building protein models in electron density maps and for the location of errors in these models. *Acta Cryst. A.* 47:110–119.
- Lonsdale, K. 1985. Atomic scattering factors for x-rays: temperature and other modifying factors. In *International Tables of Crystallography*, Vol. 3. The International Union of Crystallography, editor. Reidel Publishing, Dordrecht. 233–236.
- Lounnas, V. 1999. Molecular dynamics simulation of proteins in aqueous environment. In *Hydration Processes in Biology*. M. C. Bellissent-Funel, editor. IOS Press, Amsterdam. 261–290.
- Lounnas, V., B. M. Pettitt, and G. N. Philipps, Jr. 1994. A global model of the protein-solvent interface. *Biophys. J.* 66:601–614.
- Louro, R. O., I. Bento, P. M. Matias, T. Catarino, A. M. Baptista, C. M. Soares, M. A. Carrondo, D. L. Turner, and A. V. Xavier. 2001. Conformational component in the coupled transfer of multiple electrons and protons in a monomeric tetraheme cytochrome. *J. Biol. Chem.* 276: 44044–44051.
- Louro, R. O., T. Catarino, J. LeGall, and A. V. Xavier. 1997. Redox-Bohr effect in electron/proton energy transduction: cytochrome  $c_3$  coupled to hydrogenase works as a 'proton thruster' in *Desulfovibrio vulgaris*. *J. Biol. Inorg. Chem.* 2:488–491.
- MacKerell, A. D., Jr., D. Bashford, M. Bellot, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-MacCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. I. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem.* 102:3586–3616.
- Magro, V., L. Pieulle, N. Forget, B. Guigliarelli, Y. Pétillot, and E. C. Hatchikian. 1997. Further characterization of the two tetraheme cytochromes  $c_3$  from *Desulfovibrio africanus*: nucleotide sequences, EPR spectroscopy and biological activity. *Biochim. Biophys. Acta* 1342: 149–163.
- Martel, P. J., C. M. Soares, A. M. Baptista, M. Fuxreiter, G. Nàray-Szabó, R. O. Louro, and M. A. Carrondo. 1999. Comparative redox and pKa calculations on cytochrome  $c_3$  from several *Desulfovibrio* species using continuum electrostatic methods. *J. Biol. Inorg. Chem.* 4:73–86.
- Nørager, S., P. Legrand, L. Pieulle, C. Hatchikian, and M. Roth. 1999. Crystal structure of the oxidized and reduced acidic cytochrome  $c_3$  from *Desulfovibrio africanus*. *J. Mol. Biol.* 290:881–902.
- Pereira, P. M., I. Pacheco, D. L. Turner, and R. O. Louro. 2002. Structure-function relationship in type II cytochrome  $c_3$  from *Desulfovibrio africanus*: a novel function in a familiar heme core. *J. Biol. Inorg. Chem.* 7:815–822.
- Pessanha, M., L. Brennan, A. V. Xavier, P. M. Cuthbertson, G. A. Reid, S. K. Chapman, D. L. Turner, and C. A. Salgueiro. 2001. NMR structure of the haem core of a novel tetrahaem cytochrome isolated from *Shewanella frigidimarina*: identification of the haem-specific axial ligands and order of oxidation. *FEBS Lett.* 489:8–13.
- Pieulle, L., J. Haladjian, J. Bonicel, and E. C. Hatchikian. 1996. Biochemical studies of the c-type cytochromes of the sulfate reducer *Desulfovibrio africanus*: characterization of two tetraheme cytochromes  $c_3$  with different specificity. *Biochim. Biophys. Acta.* 1273:51–61.
- Pople, J. A., and D. L. Beveridge. 1970. Approximate Molecular Orbital Theory. McGraw-Hill, New-York.
- Salgueiro, C. A., P. N. da Costa, D. L. Turner, A. C. Messias, W. M. A. M. van Dongen, L. M. Saraiva, and A. V. Xavier. 2001. Effect of hydrogen-bond networks in controlling reduction potentials in *Desulfovibrio vulgaris* (Hildenborough) cytochrome  $c_3$  probed by site-specific mutagenesis. *Biochemistry.* 40:9709–9716.
- Soares, C. M., P. J. Martel, J. Mendes, and M. A. Carrondo. 1998.

- Molecular dynamics simulation of cytochrome  $c_3$ : studying the reduction processes using free energy calculations. *Biophys. J.* 74:1708–1721.
- Sterpone, F., M. Ceccarelli, and M. Marchi. 2001. Dynamics of hydration in hen egg white lysozyme. *J. Mol. Biol.* 311:409–419.
- Teixeira, V. H., C. M. Soares, and A. M. Baptista. 2002. Studies of the reduction and protonation behavior of tetraheme cytochromes using atomic details. *J. Biol. Inorg. Chem.* 7:200–216.
- Ullmann, G. M. 2000. The coupling of protonation and reduction in proteins with multiple redox centers: theory, computational method and applications to cytochrome  $c_3$ . *J. Phys. Chem. B.* 104:6293–6301.
- Valente, F. M. A., L. M. Saraiva, J. LeGall, A. V. Xavier, M. Teixeira, and I. A. C. Pereira. 2001. A membrane-bound cytochrome  $c_3$ : a type II cytochrome  $c_3$  from *Desulfovibrio vulgaris* Hildenborough. *Chem. Biochem.* 2:895–905.
- van Gunsteren, W. E., and H. J. C. Berendsen. 1990. Computer simulations of molecular dynamics: methodology, applications and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.* 29:992–1023.
- Woods, R. J., M. Khalil, W. Pell, S. H. Moffat, and V. H. Smith, Jr. 1990. Derivation of net atomic charges from molecular electrostatic potentials. *J. Comput. Chem.* 11:297–310.