Theor Chem Acc (1999) 101:223–227
DOI 10.1007/s002149800m75

Theoretical
Chemistry Accounts

Regular article

Information about the biologically relevant properties of Clostridium pasteurianum rubredoxin obtained from modeling and dynamics simulations of molecular variants*

J.-M. Moulis

CEA, Département de Biologie Moléculaire et Structurale, Laboratoire Métalloprotéines, 17 rue des Martyrs, F-38054 Grenoble Cedex 9, France; e-mail: jean-marc.moulis@cea.fr, Tel.: +33-476-885623, Fax: +33-476-885872

Received: 16 May 1998 / Accepted: 4 August 1998 / Published online: 2 November 1998

Abstract. Rubredoxins are small electron transfer proteins containing one iron atom at their active site. The rubredoxin from the anaerobic bacterium *Clostridium pasteurianum* has been subjected to molecular dynamics studies starting from the minimized solvated structure. The results of the simulations have been compared with identical ones carried out with selected mutated forms of the protein obtained by molecular modeling. Surface residues, which are highly conserved among rubredoxins and close to the cysteine ligands, can be replaced by glutamates, i.e. long chain carboxylates. The main structural consequence is a shift of the protein backbone bearing conserved aromatic residues. Reciprocally, substitution of the aromatic residue closest to the iron atom shifts the cysteine-containing peptide fragments. These observations have been related to the changes in electron transfer and redox properties previously measured for this set of rubredoxin molecular variants.

Key words: Rubredoxin – Iron – Sulfur – Electron transfer – Reduction potential

1 Introduction

The non-heme iron proteins rubredoxins have been found in a variety of anaerobic bacteria [1]. In the few cases in which their function is known, they are part of electron transfer chains [2], with their single iron atom exchanging one electron between the ferrous and ferric states.

Several structures of rubredoxins have been solved, some of them with exceptional accuracy for proteins [1]. The iron is coordinated by four cysteine residues to the short polypeptide chain of between 45 and 55 amino acids. A slightly different situation is found in *Pseudomonas oleovorans*, where the rubredoxin mediating electron transfer between NADPH and the enzyme catalyzing the hydroxylation of alkanes is longer and contains two iron atoms [3]. For short rubredoxins, such as the one isolated from *Clostridium pasteurianum*, the polypeptide chain folds mainly as β-sheets connected by 3_{10} turns [4; Fig. 1]. The iron atom of the active site can be replaced by a variety of other metals such as Co, Ni, Cd [5–7], or Zn [8]. In the latter case the near two-fold symmetry exhibited by the iron-substituted site [4] defines a folding motif which is strikingly reminiscent of part of that found in cysteine-containing Zn-finger proteins, a large family of cellular components involved in protein-protein or protein-DNA interactions [9].

Despite this wealth of information, the structural features responsible for the functional properties of these proteins remain unclear. Recently, site-directed mutagenesis experiments based on the nucleotide sequence of the gene from *C. pasteurianum* [10] have revealed significant shifts of either the values of the reduction potential [11–13] or the electron self-exchange rate constant [13]. In an effort to obtain further insight into the structural consequences of substituting some of the amino acids close to the iron atom, molecular modeling and dynamics studies have been performed with some of the recently generated molecular variants [13] and are reported herein.

2 Methods

The starting model of *C. pasteurianum* rubredoxin used in the present calculations is the one derived from 1.1 Å-resolution X-ray crystallographic data [4; Protein Data Base code 1IRO]. For disordered residues, only the side chain conformations of highest occupancies were considered. The C-terminal residue of the sequence was not included in the calculations as it is undefined in the crystal structure. This residue is not expected to be of major importance in the behavior of the active site, as it lies more than 15 Å away from the iron.

The water molecules located in the crystallographic model were removed from the structure and the protein was solvated by a 8 Å thick shell of water molecules using the TIP3P model. Standard

* Contribution to the Proceedings of Computational Chemistry and the Living World, April 20–24, 1998, Chambery, France

parameters defined in CHARMm22 were implemented to describe the interactions involving the atoms of the system and missing hydrogen atoms were built with HBUILD [14]. The active site unit comprising the iron and the four coordinating cysteines was assigned average geometric equilibrium values derived from the most accurate structure available [4], and other relevant parameters from resonance Raman [15] or theoretical calculations [16]. These values were previously reported [13] and apply to a ferric ion (oxidized protein). The overall charges of the molecules were not neutralized but were smoothed over carbon and nonpolar hydrogen atoms. For each of the molecular variants investigated, the substituted residue was replaced on the crystal structure via the MUTATE option of the Quanta package (Molecular Simulations) before starting any calculation. All calculations were carried out in an identical way for all protein forms, smoothing the long-range interactions with a switching function of cut-off distances 11 Å and 14 Å.

A first set of minimization using 50 steps of steepest descents [17] was applied to the solvated crystal structures. Adopted-basis set Newton-Raphton cycles [14] were then run until a root mean square (rms) force of less than 0.01 was reached. Update of coordinates occurred every five steps. The resulting minimized structures were used as the starting coordinates for the simulations.

The equilibrium conditions of bonds involving hydrogens were kept with the SHAKE algorithm [18]. Nonpolar hydrogens were part of the carbon atoms to which they were bound. Molecular dynamics simulations [14] were carried out at 300 K: the system was first heated to this temperature for 3 ps with a time step of 1 fs and equilibrated for 30 ps. Data were then collected every 0.5 ps for the following 100 ps.

Fig. 1. Schematic view of *Clostridium pasteurianum* rubredoxin. A *ribbon* is drawn through the protein backbone (Protein Data Bank 1IRO, [4]). Side chains discussed in the text are drawn in *ball and stick*. The $FeCys_4$ unit (*center top*) is in *gray*, with Cys-9 (right) and Cys-42 (left) on top, and the sites of substitution (Val-8, Gly-10, and Tyr-11) in *lighter gray*. The side chains of Trp-37 and Phe-49 are labeled in the center of the view

3 Results and discussion

The substitutions of amino acids studied in this work are depicted in Fig. 1.They all lie close to the iron atom and have been shown to induce significant differences in either the self-exchange rate constant for V8E (glutamate replacing valine 8) and G10E or the value of the reduction potential for G10E and Y11N [13]. It was thus of interest to check to what structural changes, if any, these substitutions may be related.

Table 1. Comparison of the minimized structures of rubredoxin molecular variants

	Native	Y11 N	G10 E	V8 E
Energy (kcal/mol)	− 10196	−10266	−10363	−10310
Fe—SCys6 (Å)	2.24	2.25	2.26	2.24
Fe—SCys9 (Å)	2.26	2.27	2.26	2.24
Fe—SCys39 (Å)	2.26	2.25	2.28	2.22
Fe—SCys42 (Å)	2.25	2.26	2.25	2.25

3.1 Comparison of minimized structures

The total energy of the minimized native protein is slightly less negative than that of each modified molecule, when the minimization procedures converge (Table 1). This indicates that no major destabilizing effects are contributed by these amino acid substitutions. However, significant localized differences between the four minimized structures can be noticed.

The prosthetic group of the protein, corresponding to the $Fe(SCys)_4$ structural unit, does not exactly display the same geometric values, although the overall structure remains the same. In particular, the lengths of the coordinating Fe—S bonds vary (Table 1). In the case of Y11N a lengthening of the bonds linking the iron to the two sulfur atoms of the most exposed cysteines (numbered 9 and 42 in the sequence, Fig. 1) by ca. 0.01 Å occurs. In contrast, a lengthening of ca. 0.02 Å holds for G10E on the other two bonds involving cysteines 6 and 39. In both cases these changes concern the Fe—S bonds the most remote from the site of the substitution. For V8E, a general shortening of these bonds (0.015 Å on average) results from the substitution of this external residue. The bond and dihedral angles involving the iron atom also vary, but it is difficult to deduce a general trend from the differences.

The comparison can be extended to the whole structures. Although many criteria could be used, it may be worth considering the rms differences between the minimized structures with that of the native protein taken as reference. From the plots concerning the main chain atoms (N, Cα, and C) and Fe only, it appears that the protein reorganizes as a whole upon minimization from the crystallographic model with an average deviation of 0.55 Å (Fig. 2, top panel). The mean deviations are similar for the molecular variants, but more specific movements of the main chain are clearly seen (Fig. 2). Although the amplitude of these changes is weaker for Y11N, the main chain of all three modified proteins mainly shifts from its position in minimized native rubredoxin for residues 46–49. Residue 49 is a fully conserved phenylalanine among rubredoxins. In addition, the backbone regions 6–8 and 29–31 in V8E and 36–37 in G10E display the next largest shifts from their

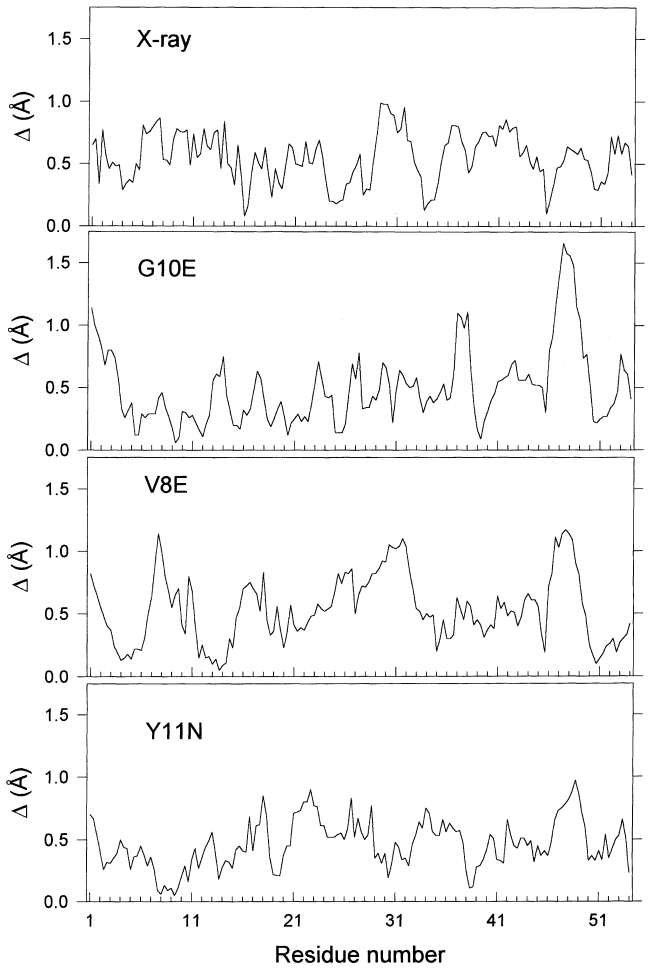

Fig. 2. Structural comparison of the minimized structures. The root mean square differences between the minimized structure of the native protein and the crystallographic model (*top*), minimized G10E (*second from top*), minimized V8E (*third from top*), and minimized Y11N (*bottom*) are plotted for the N, Cα, and C atoms of each residue in sequencing order. The right-most point is for the Fe atom

minimized positions in native protein. The amino acids 6–8 are close to the site of the substitution in V8E, but the other two regions contain also conserved aromatic residues, Phe-30 and Trp-37, respectively. Interestingly, the side-chains of these aromatic residues do not move as much as the backbones in these regions.

As expected from the similar energy values for all molecular forms, the side chains of the substituting residues are relatively easily included in the structures. Glutamates in positions 8 or 10 point toward the solvent, as predicted for these external segments of the molecule (Fig. 1). The replacement of the bulky Tyr-11 by the smaller asparagine occurs without major structural rearrangement. This indicates that the polar amide group of Asn-11 can be accomodated in the interior of the protein. In the minimized structure of Y11N, a water molecule is located close to the position of the OH of Tyr-11 of the minimized native protein and is a hydrogen bond acceptor from the NH_2 group of Asn-11. The same water molecule bridges Asn-11 to Asp-19 through hydrogen bonds, in an arrangement reminiscent of the H-bonding between OH of Tyr-11 and one of the carboxylic oxygens of Asp-19 in the native structure.

These observations are in qualitative agreement with the results of our previous electron self-exchange rate constant measurements [13]. The latter have been interpreted as arising from the reactivity of the bimolecular complex built around the two β-turns containing the ligands and easily accessible carboxylic groups, like Glu-8 or Glu-10, are likely to perturb the interaction. In contrast, the buried Asn-11 is not expected to significantly interfere in this reaction, as is observed [13].

3.2 Comparison of molecular dynamics simulations

The simulations of these slightly different molecules have been carried out for a relatively short time (100 ps) only in order to evaluate the stability over time of structural differences and their functional consequences, if any.

Indeed, these simulations reveal some changes. The electrostatic potential around the iron atom is contributed by neighboring polar groups belonging to the protein or the solvent. Concerning the latter, the exposed gap between the turns holding the ligands is a potential zone of solvent access to the active site (Fig. 1). A parameter related to the gap between these two ligand-containing segments is the distance between the carbon (carbonyl) atoms of residues 8 and 41 located at the rims of the β-turns. As shown in Fig. 3, this distance varies for the considered molecules with G10E > native > V8E > Y11N on average. In line with this observation, the number of water molecules neighboring the iron atom increases on average and the distance of closest approach decreases following the same ordering, as already reported [13], i.e. as a function of the gap between

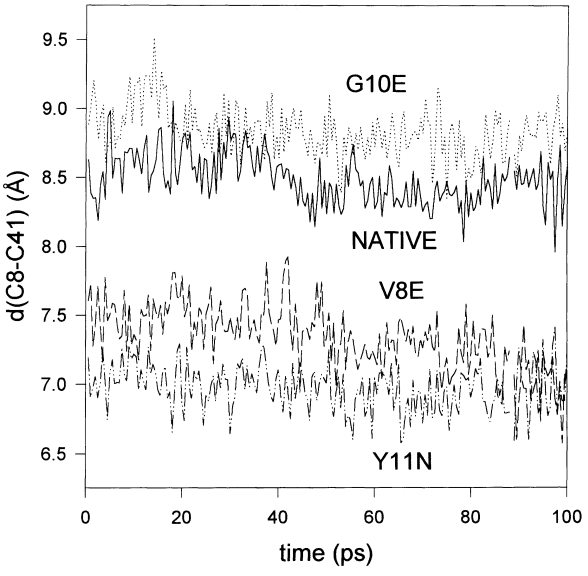

Fig. 3. Time evolution of the distance between the ligand-bearing segments during the simulations. The distance between the main chains of residues 8 and 41 is a good indicator of the gap between the β-turns containing the cysteines. The variation of this distance during the simulations is plotted for the native protein (*full line*), Y11N (*dotted and broken line*), G10E (*dotted line*), and V8E (*broken line*)

these β-turns. This order exactly matches the ranking found for the values of the reduction potential at −117, −95, −75, and −22 mV for G10E, native, V8E, and Y11N, respectively [13].

Other polar groups surrounding the redox site in rubredoxins are the peptide amide groups involved in the six H-bonds with the sulfurs of coordinating cysteines [4]. Some significant changes are apparent during the simulations, with a general increase (weakening) of these bonds for G10E, variable differences for V8E, and almost no variation for Y11N. As an example, the evolution of the distance between the sulfur atom of Cys-6 and the peptide nitrogen of residue 8 is shown in Fig. 4 for all molecular variants. The predicted H-bond between these atoms is clearly lost or very weak during the simulation for G10E with a mean distance of 4.26 Å, whereas it is nearly identical for the other variants.

Another structural peculiarity of rubredoxins is the set of six conserved aromatic residues (Tyr-4, -11, and -13, Phe-30 and -49, and Trp-37 in the *C. pasteurianum* protein), the side chains of which build the major part of the inner core of the molecule. Replacing one of them by a nonaromatic residue is likely to disturb the structure. Excepting Tyr-11, the aromatic residues closest to the iron atom (and to Tyr-11) are Phe-49 and Trp-37 (Fig. 1). When plotting the distance separating the closest atoms of these residues over the time of the simulation (Fig. 5), Y11N displays a significantly longer value (mean 4.92 Å) compared to the other forms. Some comparatively minor differences can also be noticed among the latter (Fig. 5). Interestingly, the shorter this distance (Fig. 5), the wider the gap between the β-turns holding the ligands (Fig. 3).

Other geometric relationships among the structures obtained during these simulations could be derived, but an extensive comparison is beyond the scope of this report. The few examples given above nevertheless sustain some conclusions pertaining to the relevance of the present results to the biochemical properties previously observed [13].

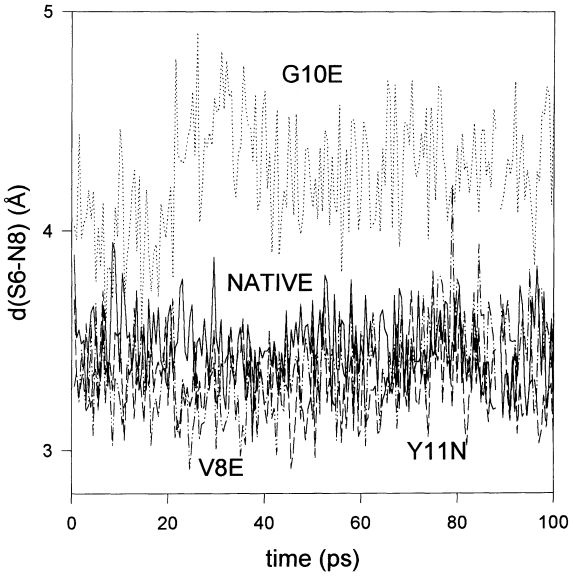

Fig. 4. Time evolution of the NH8-S6 potential hydrogen bond during the simulations. The symbols are as in Fig. 3

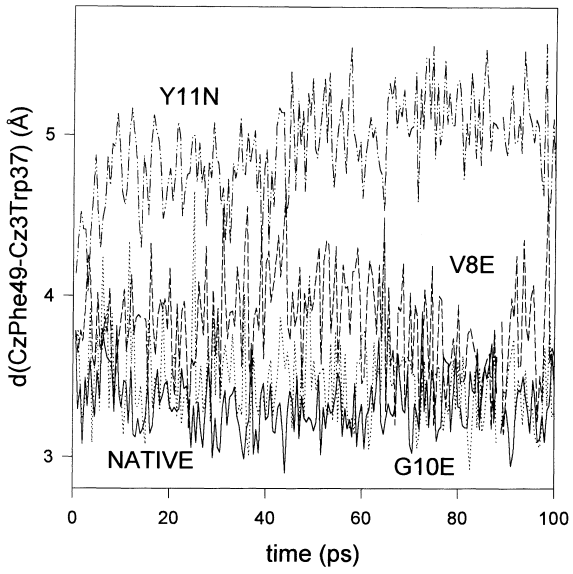

Fig. 5. Time evolution of the aromatic core compactness during the simulations. The index used is the distance between the outermost atoms of residues 37 and 49 (see Fig. 1). The symbols are as in Fig. 3

4 Conclusions

The large difference in the values of the electron-self exchange rate constants measured for V8E and G10E, on the one hand, and the native protein, on the other, can be readily explained by the repulsion between molecules owing to the presence of exposed carboxylates at the interaction site in the former cases [13]. This conclusion is borne out by the present results which do not show any sign of intramolecular interactions of these carboxylates with other protein groups in the solvated molecule. In contrast, the strong electrostatic repulsion suggested between molecules of V8E or G10E is far less intense in the cases of the native or Y11N proteins.

However, such changes in the charge distribution at the surface of the molecule are less likely to shift the value of the reduction potential. This property depends on the distribution of dipoles, including solvent molecules, around the redox site [19]. The outcome of the 100 ps simulations is that several sets of such dipoles are perturbed in the molecules bearing amino acid changes. A decreased compactness of the aromatic core, as more clearly evidenced in Y11N (Fig. 5), seems to bring the two cysteine-containing β-turns close together (Fig. 3) and decreases access of the solvent to the redox active ferric ion. This effect may be of major impact in setting the value of the redox potential, as it follows exactly the increase in potential measured among these four molecules. Other kinds of permanent dipoles, such as the peptidic amides involved in H-bonds with the ligands

of the iron (Fig. 4), can also participate in defining the electrostatic potential around the metal atom and probably exert correcting effects to the above phenomenon.

It is the difference between the sums of all these dipolar contributions at both redox levels that defines the value of the reduction potential. This explains why even sophisticated predictive methods to relate structure to the absolute values of proteins' reduction potentials are not yet reliable. Modeling and dynamics simulations of redox proteins, such as rubredoxins [20; this work], are likely to improve these methods. It is hoped that the study of a set of slightly perturbed simple molecules as implemented here may lead the path to a better understanding of this fundamental property of redox proteins.

Acknowledgements. Dr. Serge Crouzy is thanked for enlightening discussions and Dr. Jacques Meyer for careful reading of the manuscript and useful suggestions.

References

1. Sieker LC, Stenkamp RE, LeGall J (1994) Methods Enzymol 243:203
2. LeGall J, Xavier AV (1996) Anaerobe 2:1
3. Lode ET, Coon MJ (1971) J Biol Chem 246:791
4. Dauter Z, Wilson KS, Sieker LC, Moulis J-M, Meyer J (1996) Proc Natl Acad Sci USA 9:8836
5. May SW, Kuo J-Y (1978) Biochemistry 17:3333
6. Kowal AT, Zambrano IC, Moura I, Moura JJG, LeGall J, Johnson MK (1988) Inorg Chem 27:1162
7. Henehan CJ, Pountney DL, Zerbe O, Vašák M (1993) Protein Sci 2:1756
8. Pétillot Y, Forest E, Mathieu I, Meyer J, Moulis J-M (1993) Biochem J 296:657
9. Schwabe JWR, Klug A (1994) Nat Struct Biol 1:345
10. Mathieu I, Meyer J, Moulis J-M (1992) Biochem J 285:255
11. Zeng Q, Smith ET, Kurtz DM Jr, Scott RA (1996) Inorg Chim Acta 242:245
12. Ayhan M, Xiao ZG, Lavery MJ, Hamer AM, Nugent KW, Scrofani SDB, Guss M, Wedd AG (1996) Inorg Chem 35:5902
13. Kümmerle R, Zhuang-Jackson H, Gaillard J, Moulis J-M (1997) Biochemistry 36:15983
14. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comput Chem 4:187
15. Yachandra VK, Hare J, Moura I, Spiro TG (1983) J Am Chem Soc 105:6455
16. Noodleman L, Norman JG Jr, Osborne JH, Aizman A, Case DA (1985) J Am Chem Soc 107:3418
17. Levitt M, Lifson S (1969) J Mol Biol 46:269
18. Rychaert JP, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327
19. Warshel A, Papazyan A, Muegge I (1997) J Biol Inorg Chem 2:143
20. Yelle RB, Park N-S, Ichiye T (1995) Proteins 22:154