ELSEVIER

Biophysical Chemistry 125 (2007) 532–539

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Structural, dynamical and functional aspects of the inner motions in the blue copper protein azurin

Bruno Rizzuti [a], Luigi Sportelli [b], Rita Guzzi [b,*]

[a] *Laboratorio Licryl CNR-INFM, Dipartimento di Fisica, Università della Calabria, Ponte P. Bucci, 87036 Rende, Italy*
[b] *Laboratorio di Biofisica Molecolare, Dipartimento di Fisica, Università della Calabria, Ponte P. Bucci, 87036 Rende, Italy*

Received 4 August 2006; received in revised form 15 November 2006; accepted 16 November 2006
Available online 23 November 2006

Abstract

Molecular dynamics was applied to dissect out the internal motions of azurin, a copper protein performing electron transfer. Simulations of 16.5 ns were analyzed in search of coordinated displacements of amino acid residues that are important for the protein function. A region with high conformational instability was found in the 'southern' end of the molecule, far away from the copper site and the binding sites for the redox partners of azurin. By excluding the 'southern' region from the subsequent analysis, correlated motions were identified in the hydrophobic patch that surrounds the protein active site. The simulation results are in excellent agreement with recent NMR data on azurin in solution [A. V. Zhuravleva, D. M. Korzhnev, E. Kupce, A. S. Arseniev, M. Billeter, V. Y. Orekhov, Gated electron transfers and electron pathways in azurin: a NMR dynamic study at multiple fields and temperatures, J. Mol. Biol. 342 (2004) 1599–1611] and suggest a rationale for cooperative displacements of protein residues that are thought to be critical for the electron transfer process. A number of other structural and dynamic features of azurin are discussed in the context of the blue copper protein family and an explanation is proposed to account for the variability/conservation of some regions in the cupredoxins.

Keywords: Azurin; Blue copper proteins; Molecular dynamics; Essential dynamics; Cooperative motions

1. Introduction

Azurin is a small (N=128 residues) copper protein, acting as an electron transfer shuttle in *Pseudomonas aeruginosa* and other bacteria. The presence of the copper ion gives this protein a number of features, including an intense blue color, a high reduction potential and a small parallel hyperfine coupling in the electron spin resonance spectrum [1]. In addition, azurin is fairly resistant to thermal denaturation: the temperature of melting is T_m=86 °C and the conformation stability at 20 °C and pH 7 is ΔG=60 kJ·mol^{-1} [2,3], both values unusually high for a (non-thermophilic) protein of this size. The robustness of azurin is likely related to the presence of the copper ion and necessary to perform its biological role [4]. In fact, the folding of azurin is common with the other proteins of the same family, the cupredoxins, and consists in a two-sheet β-sandwich with eight β-strands arranged in a Greek-key topology. The high β-sheet content provides stiffness for large protein regions and ensures that the coordination of the copper ion in the active site is similar in the reduced and oxidized state, in between the distorted tetrahedral geometry preferred by Cu(I) and the tetragonal geometry preferred by Cu(II). Thus, no significant conformational changes accompany the redox reaction, leading to a small reorganization energy and a consequently high rate of electron transfer.

Because of their peculiar spectroscopic and structural properties, azurin and other blue copper proteins have been intensively studied in the last two decades by using a number of experimental techniques [5], as well as employing computational methods [6]. In particular, Molecular Dynamics (MD) simulation has been used as a tool to investigate the dynamics of azurin with the aim of unraveling the microscopic details of its function [7,8]. Anharmonic inner motions involving single residues that may be important in the electron transfer process were suggested [7]. The basic features of the coordination of the copper ion with the protein matrix, as the latter undergoes large scale motions in the presence of the solvent, were clarified [9].

* Corresponding author. Tel.: +39 0984 496077; fax: +39 0984 494401.
E-mail address: guzzi@fis.unical.it (R. Guzzi).

0301-4622/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2006.11.003

Moreover, the comparison of the MD results obtained for azurin and plastocyanin [8] strongly suggests that similarities in the structure, dynamics and function are ultimately related within the entire blue copper protein family.

In this work, the inner motions of azurin are dissected out in connection with the conditions assessing their equilibration. The results show that a few nanoseconds of simulations are sufficient to determine fluctuations of the protein backbone, but not global collective motions. Convergence of sampling for coordinate displacements is obtained by identifying and excluding from the subsequent analysis a protein region with high conformational instability. This region is small and poorly structured and constitutes the 'southern' end of the protein [1], remote from the active site and the binding sites for the redox partners of azurin. After excluding the 'southern' end, the analysis of the inner motions in the 'northern' region around the active site suggests a rationale for the collective mobility evidenced in recent NMR experiments [10,11] that is believed to contribute to the efficiency of the electron transfer process in azurin. Furthermore, from a structural point of view, an explanation is proposed to account for two seemingly-conflicting features of the blue copper protein family: (a) high variability in the relative position of the two β-sheets and (b) strong conservation for the region that is part of the two sheets and extends from the active site to the hydrophobic core.

2. Computational methods

2.1. Molecular dynamics

MD was performed using the GROMACS package [12] with the GROMOS force field 43A1 [13]. The crystal structure of *P. aeruginosa* azurin [14], 4AZU entry in the Protein Data Bank, was used as the starting structure. Only one of the four crystal monomers was considered, together with 80 related crystallization water molecules. Protonation state of the protein residues was adapted to mimic neutral pH. Active site was modeled as previously described [9] using a set of atomic partial charges that takes into account the presence of the Cu ion and employing constraints to fix the bond lengths between the metal and its five native ligands. The protein, together with crystallization waters, was solvated in a rhombic dodecahedron with a nearest image distance of 0.8 nm. The total number of water molecules was 5441 and the simple point charge (SPC) model was used [15]. Periodic boundary conditions were applied.

Residual strain in the system was relieved by 80 steps of steepest descent minimization, using a force constant of 10^3 kJ mol^{-1} nm^{-2} to restrain the position of the protein atoms. Simulations were performed in the NPT ensemble using a time step of 2 fs for integrating the equations of motion. A Berendsen thermostat and barostat were applied [16] with reference temperature of 300 K and pressure of 10^5 Pa and with coupling times of 0.1 and 1 ps, respectively. A twin range cutoff of 0.8 and 1.4 nm was used for nonbonded interactions, updating the neighbour pair list every 5 steps. The SHAKE algorithm [17] was used to constrain the bond lengths of protein and water molecules. Initial atomic velocities were assigned from a Maxwellian distribution at the starting temperature of 250 K. Simulated annealing was performed for 50 ps to gradually increase the temperature up to 300 K. Trajectory data were saved every 0.2 ps during the production run, performed for a total time of 16.5 ns. In addition, three control runs of the same length were performed under slightly different conditions of simulation, by removing the constraint between the Cu ion and, respectively, the His46, His117 or Met121 ligand residues, as previously described [9].

2.2. Data analysis

Correlated atomic motions can be estimated in a straightforward way by using dynamical cross correlations (DCC) [18]. DCC between the position $r_i(t)$ and $r_j(t)$ of the ith and jth atom, respectively, is calculated as:

$$\gamma_{ij} = \frac{\langle r_i r_j\rangle - \langle r_i\rangle\langle r_j\rangle}{[(\langle r_i^2\rangle - \langle r_i\rangle^2)(\langle r_j^2\rangle - \langle r_j\rangle^2)]^{1/2}} \quad (1)$$

where the brackets <..> represent the time average over the simulation trajectory. Correlation can be computed for each protein atom with any other; in our case, only C^α atoms were considered. Thus, DCC directly indicate whether two residues move in the same or opposite directions, though they are not sensitive to perpendicular displacements and to the magnitude of the motion.

In addition, the essential dynamics (ED) technique [19] was also employed. The covariance matrix is calculated from a high-dimensional set of coordinates $x(t)$ of a number of atoms whose collective position determines the protein structure (in our case, the $N = 128$ C^α atoms of azurin were considered). This allows to represent each structure in the simulation ensemble as a point in a $3N$-dimensional space, the conformation space. The covariance of coordinate i and j is defined as:

$$C_{ij} = \langle (x_i - \langle x_i\rangle)(x_j - \langle x_j\rangle)\rangle \quad (2)$$

where the brackets <..> indicate the average over the data points in the trajectory. The corresponding, symmetric matrix **C** can be diagonalized with an orthonormal transformation matrix **R**:

$$\mathbf{C} = \mathbf{R}\ \mathit{\Lambda}\ \mathbf{R}^T \quad (3)$$

where the superscript T denotes the transposed matrix. The matrix **R** contains as columns the eigenvectors, i.e. the principal components, whereas the resulting diagonal matrix **Λ** contains the corresponding eigenvalues, i.e. the mean square positional fluctuations. The eigenvectors can be sorted to decreasing eigenvalues, so that the first principal components correspond to the less constrained degrees of freedom of the protein. The analysis was performed after least squares fitting on the protein C^α atoms of a reference structure to remove the overall molecule rotation. Convergence of the sampled space can be measured in terms of overlap of the fluctuations. The covariance matrix **C**, relative to a given time interval, can be compared with the covariance matrix, indicated as **M**, relative to a different

time interval. The overlap of the two matrices was calculated as [20]:

$$s(\mathbf{C},\mathbf{M}) = 1-\left[\sum_{i=1}^{3N}(\lambda_i^{\mathbf{C}}+\lambda_i^{\mathbf{M}})-2\sum_{i=1}^{3N}\sum_{j=1}^{3N}(\lambda_i^{\mathbf{C}}+\lambda_j^{\mathbf{M}})^{1/2}(R_i^{\mathbf{C}}\cdot R_j^{\mathbf{M}})^2\right]^{1/2} (\mathrm{tr}\mathbf{C}+\mathrm{tr}\mathbf{M})^{-1/2} \quad (4)$$

where λ_i and λ_j are the eigenvalues, R_i and R_j are the eigenvectors, i.e. the columns of the rotation matrix **R**, and tr indicates the trace of the matrices.

3. Results and discussion

To assess the stability of the simulation a set of protein properties was monitored for the reference and control runs. The overall findings show that azurin maintains its native topology fold. The radius of gyration is 1.37±0.01 nm (uncertainty is the standard deviation) and it slightly fluctuates around the starting value, 1.370 nm. The solvent accessible surface area (SASA) in simulation is 29±1 nm^2, consistent with the value of 28.8 nm^2 derived from the X-ray data [14]. The minimum value of SASA, 0.1 nm^2, is found for the whole fragment 45–51, that is centered around the unique tryptophan, Trp48, along the primary structure. Thus, azurin maintains its hydrophobic core in the simulation.

Only 4 out of the 65 main chain–main chain hydrogen bonds are lost in all of the MD runs with respect to the crystal structure. The secondary and tertiary structure of the protein is not altered. The only bond lost between two β-strands, the one between O···N of Gln8 and Asn16, is replaced by the N···$O^{\delta 1}$ bond between the same residues. The hydrogen bond lost between Lys41–Gly45 is weak already in the crystal configuration [14], with an interaction energy $E=-7$ kJ·mol^{-1} (calculated with the DSSP program of Kabsh and Sender [21]), about half the value for the Pro40–Met44 bond that is conserved in the same helical turn. The bond between Gly63–Ser66 is replaced by two hydrogen bonds between Asp62–Ser66 and Gly63–Gly67, determining for the terminal part of the helix of azurin an interconversion between the two forms 3_{10} and α, as also observed in NMR spectroscopy [22,23]. Finally, the rupture of the Lys74–Asp77 hydrogen bond contributes to explain the fast internal motion of the corresponding protein turn, as found with NMR [24].

A number of newly-formed hydrogen bonds is observed in the simulation runs, but none of them alters the protein secondary structure. In particular, no main chain–main chain bond is observed between the two β-sheets, in agreement with MD simulations of other cupredoxins [8,25,26]. In contrast, several permanent and transient hydrogen bonds that involve side chains, i.e. main chain–side chain or side chain–side chain bonds, are registered. The resulting dynamic network of hydrogen bonds contribute to maintain the coordination between the two β-sheets that conjunctly form the protein β-sandwich. This allows the sheets, if needed, to accommodate their reciprocal position by moving slightly with respect to each other without destroying the protein scaffold or exposing to the solvent the hydrophobic core.

Fig. 1 shows the root mean square deviations (RMSD) of the position of the C^{α} atoms from the starting structure for the reference MD run, as a function of the simulation time. Deviations in the first 2.5 ns of simulation are very low (below 0.20 nm) but a second rough plateau is obtained in the curve after 3.5 ns (black line). The mean value of the C^{α} deviations of this second plateau remains lower than 0.25 nm, indicating that the inner mobility of the protein main chain is moderate. The jump observed in the RMSD curve in the time interval 2.5–3.5 ns is due to the positional rearrangements of a few mobile residues in an otherwise fairly rigid backbone. In fact, it is sufficient to exclude the first two residues in the protein N-terminus (Fig. 1, gray line) to decrease significantly the RMSD values in both the plateaus, whereas no modification is observed in the 2.5–3.5-ns interval, when the main chain N-terminus deviates less from its starting position. Arbitrarily excluding a few other residues close to the N-terminus and with correlated dynamics, as will be discussed later on, would both smooth the jump between the two plateaus and further decrease the values of the RMSD curve of a quantity ranging between 0.04–0.1 nm.

To investigate why azurin requires several nanoseconds of simulation to equilibrate in spite of low deviations of the atomic positions, the structure–dynamics relationship has to be considered in detail. In a first approximation, any protein can be simplified in terms of elements such as helices, sheets, hydrophobic cores etc [27]. The main structural units of azurin are two β-sheets and a 13-residue α-helical insertion. In the 'northern' region of the molecule there is the copper site surrounded by a hydrophobic patch [14]. In the 'southern' region the two β-sheets are connected by the two turns 23–26 and 103–106 and by the unique disulfide bridge; the two main chain termini are also present. A hypothesis to explain the dynamics of azurin is that its structural elements have more internal rigidity than the protein as a whole, but several nanoseconds are required for all these sub-units to achieve a collective coordinated motion in simulation. This hypothesis was tested in

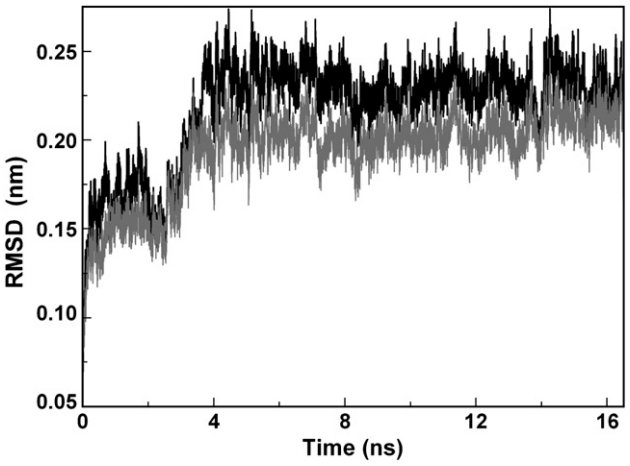

Fig. 1. Positional root mean square deviations from the crystal structure of all the C^{α} protein atoms (black line) and excluding the first two residues (gray line), as a function of the simulation time.

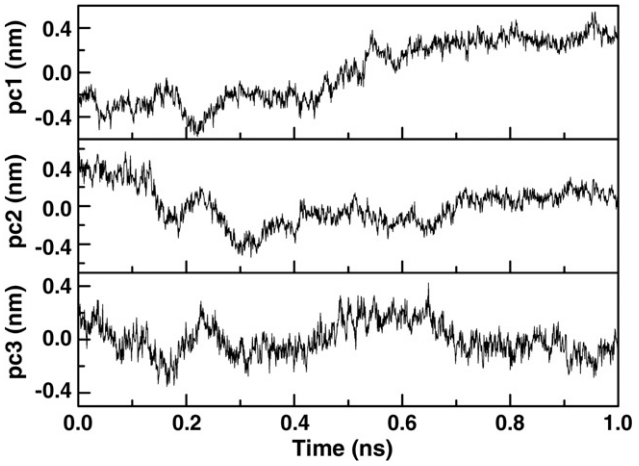

Fig. 2. First three principal components calculated in the first nanosecond of simulation.

terms of cross-correlated motions of protein residues. DCC were calculated [18] between the position of the *i*th and *j*th C^{α} atom ($i<j$) for the part of the MD trajectory corresponding to the first plateau in the atomic deviation curve (Fig. 1), i.e. 0.5–3 ns, and compared with the values obtained for five different 2.5-ns portions of the same trajectory (4–6.5, 6.5–9, 9–11.5, 11.5–14 and 14–16.5 ns).

The results show that, choosing a threshold value of 0.3 for the modulus of DCC, 460 residue–residue correlated motions were found for the 0.5–3-ns interval, whereas at least 855 were found for each of the other five intervals. This indicates that, in simulation, collective dynamics is enhanced for the structural elements of azurin on a timescale beyond the nanosecond. In addition, negative DCC were found between the protein N-terminus and the turn 23–26 in the first part of the MD trajectory. On the contrary, the correlation becomes positive after a few nanoseconds of simulation, as expected because of the presence of the disulfide bridge between the two cysteine residues 3 and 26. The disulfide bridge contributes to the coordination of the two β-sheets of azurin both from a structural [28] and dynamical [25] point of view. Thus, the observation that the dynamics of the –SS– bridge changes at different timescales indicates that it takes time for the two β-sheets of azurin to set up a synchronous motion after the simulated annealing.

The mechanism here highlighted for azurin, with the two β-sheets coordinated but with slightly independent dynamics, complements the previous observation of a dynamic network of hydrogen bonds maintaining the mutual arrangement between the sheets (as also found in other cupredoxins [26]). This could contribute to explain why, in the blue copper protein family, there has been little structural constraint for the whole sheets to move relative to each other during the course of evolution, while there have been very strong constraints holding the strands in their positions in the sheets [29]. Furthermore, the same mechanism can explain why nanosecond simulations are, in some cases, insufficient to obtain a proper sampling of protein inner motions even in the presence of a well-defined, highly-cooperative structure. Undersampling of conformational space in the simulation of β-barrels on the 10-ns scale is not uncommon [30], depending in most cases on the presence of sub-domains lacking secondary structure, or on the details of the protein architecture and environment.

The ED technique [19] was used to extract the essential degrees of freedom of azurin and further investigate the protein inner motions. Unlike DCC, ED is sensitive to twist and mutually perpendicular displacements, as well as to the magnitude of the motion [19]. The analysis was performed on the C^{α} atoms and the crystal structure was used as reference structure to remove the rotation of the molecule. The first three principal components in the first nanosecond of simulation are shown in Fig. 2. When the sampling along an eigenvector is incomplete, principal components resemble cosines with the number of periods equal to half the eigenvalue index [31]. Thus, the curves obtained for azurin confirm that the protein structure must gradually equilibrate before any conclusion on its coordinated inner motions can be drawn.

A more thorough ED analysis was performed in the time interval 3.5–16.5 ns, corresponding to the second plateau in the atomic RMSD curve (Fig. 1). The reference structure for the least squares fit was the one that deviated the least from the average structure in the same time interval. The results indicate that the distribution of the principal components is close to a Gaussian, confirming that the protein moves around only one structure. As shown in Fig. 3, the first eigenvalues of the covariance matrix depend on the principal component, whereas the difference between two consecutive eigenvalues rapidly levels down to zero as the eigenvector index increases. The first 30 and 60 principal components are required to account for, respectively, 80% and 90% of the conformational fluctuations (Fig. 3, inset). This result is considerably different compared to other proteins, because usually 10–20 principal components correspond to 90% of the fluctuations [19]. The dissimilarity can be explained in terms of the rigid body-like behavior of the structure of azurin, with few preferential and many almost-equivalent collective degrees of freedom available. The same behavior, with 30 principal components responsible for 80% of the fluctuations, is reported in a 10-ns MD simulation of

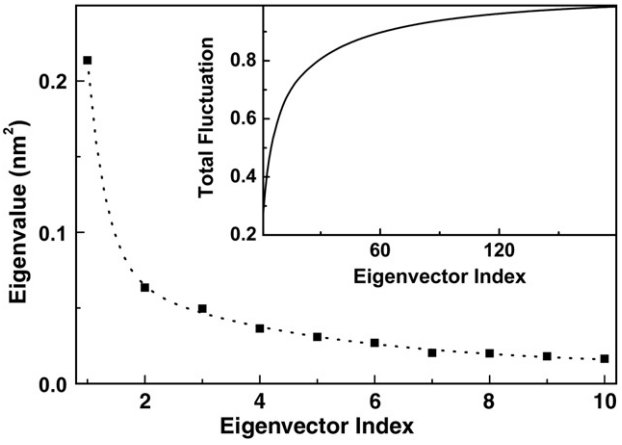

Fig. 3. Eigenvalues as a function of the first 10 eigenvectors, in the time interval 3.5–16.5 ns; points are fitted with a second order exponential (dotted line). Inset: total fluctuation as a function of the eigenvector index, for the first 180 eigenvectors.

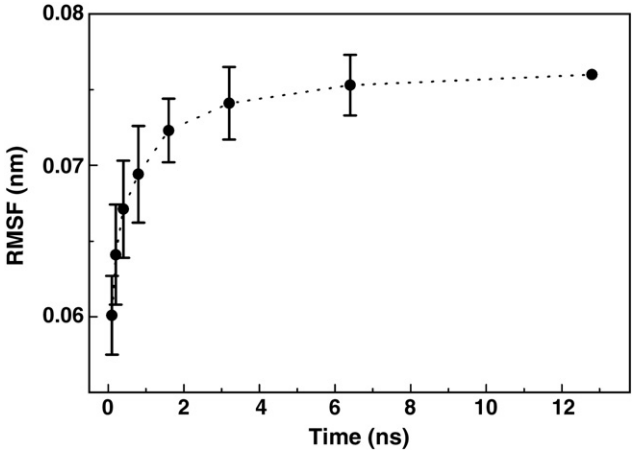

Fig. 4. Root mean square fluctuations of the C^{α} atoms as a function of the length of the subinterval. The averages are calculated over the subintervals in the time interval 3.7–16.5 ns (128 subintervals of 0.1 ns) and the error bars indicate the standard deviations.

plastocyanin [26], a protein of the same family, and may be a feature common to the cupredoxins.

More insight into the dynamics of azurin can be obtained by analyzing the fluctuations of the atomic positions. In the ED analysis, the root mean square fluctuation (RMSF) can be extracted by dividing the square root of the sum of the eigenvalues by the number of atoms [19]. The 3.7–16.5-ns interval of the reference MD run was initially divided in 128 subintervals of 0.1 ns, ED was performed on each subinterval and results were averaged. The procedure was iterated by dividing the 3.7–16.5-ns interval in 64 subintervals, then in 32, 16,... , 1 subinterval. Results are reported in Fig. 4. The fluctuations increase with the simulation time, but convergence is achieved after a few nanoseconds. The analysis of the RMSF relative to the first principal components allows to uncouple the contribution to fluctuations from different regions (data not shown). The highest correlation was found between the fluctuations in the azurin N-terminus and in the turn 103–106, as already hypothesized [24] on the basis of the NMR results. More generally, high correlation was found for the RMSF of all the other residues belonging to the 'southern' pole of azurin, albeit this region is poorly-structured compared to the protein scaffold. Thus, the ED technique reveals that the degrees of freedom associated to this region are collectively involved in chaotic motions.

Furthermore, it can be observed that while the motions of residues in the 'southern' region of azurin appear to be strongly coupled with each other, they do not appear to be coupled to inner motions in the other protein regions. The latter motions do not show at first sight any particular correlation with each other, since the fluctuations of the 'southern' region overpower their contribution in the ED analysis. For this reason, in the subsequent analysis the most mobile residues in the 'southern' site of azurin were isolated and selectively excluded to highlight collective motions in the rest of the protein. The region excluded embraces the protein residues that are not part of the β-structure in the region farthest away from the active site: residues 23–28 and 98–106 plus two residues for each main chain terminus (a total of 19 out of 128 protein residues, corresponding to 15%). High mobility of this region on the nanosecond timescale was also evidenced with NMR [11].

The ED technique was applied again and only the C^{α} atoms of the remaining 109 protein residues were considered in the analysis. To assess the convergence of the sampled space, the overlap of the fluctuations was used in terms of the covariance matrices. The covariance matrix **M** for the time interval 3.5–10 ns was constructed and compared with the covariance matrix **C** relative to the time interval 10–16.5 ns. The overlap of the two matrices is 90%, which indicates good convergence of the

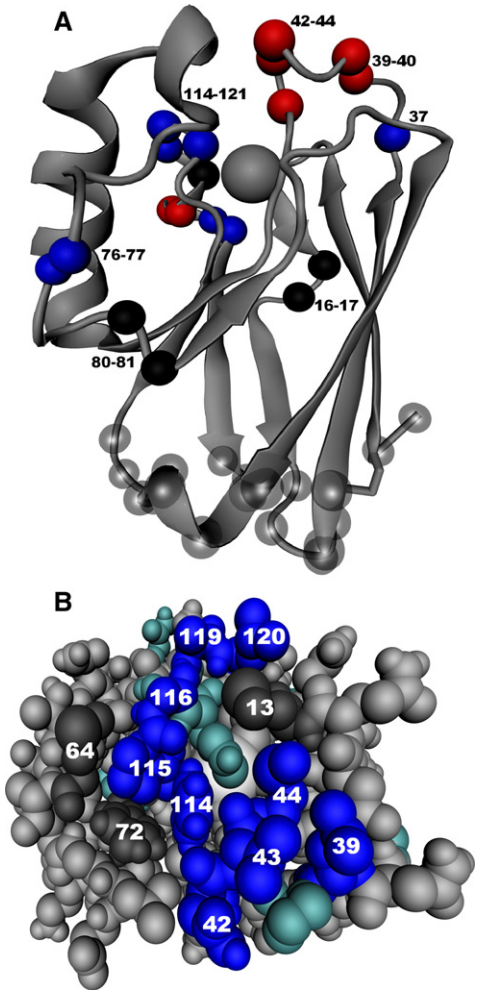

Fig. 5. Mapping of high positional fluctuations (threshold value of 0.08 nm) of the C^{α} atoms onto the structure of azurin. (A) Ribbon model of azurin, with the copper site shown as a large sphere in the 'northern' site of the molecule (top). Transparent spheres indicate the 'southern' region, undergoing chaotic motions. Opaque spheres indicate fluctuations dominated by, respectively, the first (black), second (red) and third (blue) eigenvector. Fluctuations in the helical region are omitted for clarity. (B) Atomic model of the 'northern' site of azurin, perspective top view with the copper site at the center. Residues showing high fluctuation are highlighted depending whether they belong (blue) or not (cyan) to the protein hydrophobic patch. All the residues of the hydrophobic patch are numbered (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

sampled space. The convergence is difficult to obtain even with small proteins [20] and is fundamental to investigate the protein dynamics in search of functionally relevant motions [32]. A detailed analysis of the first principal components was carried out to decouple the contribution to the atomic fluctuations of the different regions of the protein. The first eigenvector mostly describes the movements of the α-helical region, including the structures that connect it to the β-scaffold. The only other region with a significant correlated dynamics is the small loop 16–18, roughly on the opposite side of the molecule, which connects the two portions of the only strand that is common between the two β-sheets. Interestingly, all of these regions show large cooperative motions in high-temperature MD simulations during the early steps of the unfolding process of azurin [33]. Thus, we might speculate that the thermally-induced denaturation of this protein is driven by a selective excitation of specific equilibrium collective motions, on the basis of a general mechanism proposed by Di Nola and coworkers [34].

The fluctuations relative to the successive principal components are progressively weaker, as expected (Fig. 3). Coordinated motions relative to these eigenvectors concern again the α-helical region and the adjacent loops, but more evidently they cluster in the 'northern' end of the protein. When mapped onto the three-dimensional structure of azurin, they show a symmetric distribution centered around the copper site. Fig. 5 shows a ribbon representation of azurin, where residues showing fluctuations higher than a threshold value of 0.08 nm are highlighted as colored spheres according to the eigenvector they belong. Noticeably, nanosecond mobility of residues localized in the region surrounding the active site of azurin was detected in different NMR experiments [10,11]. In detail, both the region including the residues 42–45 and the solvent-exposed residues surrounding the copper ligand His117 show, at room temperature, nanosecond motions in addition to fluctuations at the picosecond timescale [10,11]. Furthermore, at 44 °C the residue Leu39, which makes also part of the hydrophobic patch, shows increased flexibility, too [11]. Nanosecond dynamics is not only limited to the 'northern' side of azurin, but it interests also some other regions lacking of secondary structure. These regions include the turns following the α-helix [10,11,24] and the loop connecting the two portions of the only strand in common between the two β-sheet [11], besides just about the whole 'southern' region [11].

Our simulation results not only reproduce the experimental findings (see Fig. 5), but also explain these observations in the framework of specific, coordinated structural motions. The functional meaning of the cooperative motions in the region surrounding the active site of azurin is not clear but some hypotheses can be formulated, including low energy-cost rearrangement of the copper-binding residues that may contribute to the efficiency of the electron transfer process, or coordinate movements of the protein hydrophobic patch that might help to complex the different protein partners or to trigger the redox reaction.

By focusing the attention on the protein regions that show little mobility, a few other conclusions can be suggested on the basis of our simulation findings. Theoretical studies [35,36] indicate that the electron transfer rates generally depend on the thermal fluctuations of protein conformations on the picosecond timescale (or faster). This picture might suggest, for azurin, an important role for the structural elements that are close to the copper ion and whose dynamics is not strongly influenced by the reciprocal displacements of the two β-sheets. The results obtained in the simulation indicate that such elements should include the fragment 47–49, which forms the hydrophobic core around the Trp48 residue. The fragment 47–49 is connected directly to the His46 copper ligand and indirectly to the Cys112 ligand, through the double hydrogen bond between Val49–Phe111. This hydrogen bond, which is maintained in all of the simulation runs, is also strongly conserved within the blue copper protein family [1]. Moreover, the fragment 47–49 forms the central part of the network of electron transfer pathways proposed for azurin [5] and shows a preserved dynamic stability over the wide range of temperature of activity of this protein [11]. Finally, from a structural point of view, the functional importance of this protein region could also explain why, in all the cupredoxins, the conserved region including the active site extends to a continuous substructure through the center of the molecule [29].

4. Conclusions

Computational methods represent an important tool to investigate the structure and dynamics of proteins with the aim of better elucidating their functional properties. Indications coming from the simulation could be particularly useful for azurin, since not all the factors that influence the electron transfer process and the mechanism of interaction with the physiological partners of this protein have been totally identified yet [5]. Insight into the properties of azurin might also extend to other cupredoxins, because of the evident similarities within the blue copper protein family.

The results of our simulations show that correlation among the residue motions in azurin is achieved via the coordination of the movements of its different structural elements. Convergence of sampling for collective motions is difficult to obtain in current MD simulations [20]. For azurin, convergence of sampling has been evidenced when a small protein region, the 'southern' end, is excluded on the basis of the indications obtained by analyzing atomic fluctuations, deviations and cross-correlated displacements. It is worth to emphasize that the 'southern' end corresponds to a highly-mobile region revealed also in NMR experiments [11]. Furthermore, this region constitutes the substructure that is the less conserved within the cupredoxin family [29].

The scaffold of azurin is fairly rigid, as indicated by the atomic deviations and fluctuations in the simulations. The principal component analysis suggests that the whole protein has few preferential collective degrees of freedom. In this respect, azurin shows a close similarity with plastocyanin, another member of the cupredoxin family, both from a structural and dynamical point of view. The similarity could extend to other blue copper proteins, sharing the same folding topology. In spite of its stiffness, however, the β-barrel of azurin is able

to dynamically rearrange its structure in the course of the simulation. This feature helps the protein to preserve the hydrophobic core and the conformation of the active site. In both cases, this could have a role in determining the electron transfer properties of azurin.

Collective motions of protein residues affect the hydrophobic patch that surrounds the protein copper site. Nanosecond timescale mobility of amino acid residues in this region was already pointed out in NMR experiments [11]. The simulation additionally suggests the notion that many of those residue displacements are correlated. A sort of self-organization of the hydrophobic patch surface, complementary to the hypothesis of an exogenous conformational rearrangement driven by a partner molecule or by the solvent, could contribute to explain a number of experimental findings, such as the fact that in azurin the redox reaction takes place probably with several partners [5], and the electron transfer rate does not depend significantly on the viscosity of the solvent [37].

Acknowledgment

VMD — Visual Molecular Dynamics [38] was used for the protein displays.

References

[1] E.T. Adman, Copper protein structures, Adv. Protein Chem. 42 (1991) 145–197.
[2] C. La Rosa, D. Milardi, D. Grasso, R. Guzzi, L. Sportelli, Thermodynamics of the thermal unfolding of azurin, J. Phys. Chem. 99 (1995) 14864–14870.
[3] R. Guzzi, L. Sportelli, C. La Rosa, D. Milardi, D. Grasso, M.P. Verbeet, G.W. Canters, A spectroscopic and calorimetric investigation on the thermal stability of the Cys3Ala/Cys26Ala azurin mutant, Biophys. J. 77 (1999) 1052–1063.
[4] H.B. Gray, B.G. Malmström, R.J.P. Williams, Copper coordination in blue proteins, J. Biol. Inorg. Chem. 5 (2000) 551–559.
[5] C. Dennison, Investigating the structure and function of cupredoxins, Coord. Chem. Rev. 249 (2005) 3025–3054.
[6] F. De Rienzo, R.R. Gabdoulline, R.C. Wade, M. Sola, M.C. Menziani, Computational approaches to structural and functional analysis of plastocyanin and other blue copper proteins, Cell. Mol. Life Sci. 61 (2004) 1123–1142.
[7] C. Arcangeli, A.R. Bizzarri, S. Cannistraro, Long-term molecular dynamics simulation of copper azurin: structure, dynamics and functionality, Biophys. Chem. 78 (1999) 247–257.
[8] C. Arcangeli, A.R. Bizzarri, S. Cannistraro, Concerted motions in copper plastocyanin and azurin: an essential dynamics study, Biophys. Chem. 90 (2001) 45–56.
[9] B. Rizzuti, M. Swart, L. Sportelli, R. Guzzi, Active site modeling in copper azurin molecular dynamics simulations, J. Mol. Model. 10 (2004) 25–31.
[10] D.M. Korzhnev, B.G. Karlsson, V.Y. Orekhov, M. Billeter, NMR detection of multiple transitions to low-populated states in azurin, Protein Sci. 12 (2003) 56–65.
[11] A.V. Zhuravleva, D.M. Korzhnev, E. Kupce, A.S. Arseniev, M. Billeter, V.Y. Orekhov, Gated electron transfers and electron pathways in azurin: a NMR dynamic study at multiple fields and temperatures, J. Mol. Biol. 342 (2004) 1599–1611.
[12] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, J. Mol. Model. 7 (2001) 306–317.
[13] W.F. van Gunsteren, F.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Biomolecular Simulation: The GROMOS96 Manual and User Guide, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, 1996.
[14] H. Nar, A. Messerschmidt, R. Huber, M. van de Kamp, G.W. Canters, Crystal-structure analysis of oxidized *Pseudomonas aeruginosa* azurin at pH 5.5 and pH 9.0. A pH-induced conformational transition involves a peptide-bond flip, J. Mol. Biol. 221 (1991) 765–772.
[15] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, in: B. Pullman (Ed.), Interaction Models for Water in Relation to Protein Hydration, Intermolecular Forces, Reidel, Dordrecht, 1981, pp. 331–342.
[16] H.J.C. Berendsen, J.P.M. Postma, A. Di Nola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[17] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[18] T. Ichiye, M. Karplus, Collective motions in proteins. A covariance analysis of atomic fluctuations in molecular-dynamics and normal mode simulations, Proteins 11 (1991) 205–217.
[19] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412–425.
[20] B. Hess, Convergence of sampling in protein simulations, Phys. Rev., E 65 (2002) 031910-1–031910-10.
[21] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[22] M. van de Kamp, G.W. Canters, S.S. Wijmenga, A. Lommen, C.W. Hilbers, H. Nar, A. Messerschmidt, R. Huber, Complete sequential ^{1}H and ^{15}N nuclear magnetic resonance assignment and solution secondary structure of the blue copper protein azurin from *Pseudomonas aeruginosa*, Biochemistry—US 31 (1992) 10194–10207.
[23] C.W.G. Hoitink, P.C. Driscoll, H.A.O. Hill, G.W. Canters, ^{1}H and ^{15}N nuclear magnetic resonance assignments, secondary structure in solution, and solvent exchange properties of azurin from *Alcaligenes denitrificans*, Biochemistry—US 33 (1994) 3560–3571.
[24] A.P. Kalverda, M. Ubbink, G. Gilardi, S.S. Wijmenga, A. Crawford, L.J.C. Jeuken, G.W. Canters, Backbone dynamics of azurin in solution: slow conformational change associated with deprotonation of histidine 35, Biochemistry—US 38 (1999) 12690–12697.
[25] B. Rizzuti, L. Sportelli, R. Guzzi, Evidence of reduced flexibility in disulfide bridge-depleted azurin: a molecular dynamics simulation study, Biophys. Chem. 94 (2001) 107–120.
[26] A.R. Bizzarri, G. Costantini, S. Cannistraro, MD simulation of a plastocyanin mutant adsorbed onto a gold surface, Biophys. Chem. 106 (2003) 111–123.
[27] A.R. Fersht, Nucleation mechanisms in protein folding, Curr. Opin. Struct. Biol. 7 (1997) 3–9.
[28] N. Bonander, J. Leckner, H. Guo, B.G. Karlsson, L. Sjölin, Crystal structure of the disulfide bond-deficient azurin mutant C3A/C26A. How important is the S–S bond for folding and stability? Eur. J. Biochem. 267 (2000) 4511–4519.
[29] J. Gough, C. Chothia, The linked conservation of structure and function in a family of high diversity: the monomeric cupredoxins, Structure 12 (2004) 917–925.
[30] J.D. Faraldo-Gómez, L.R. Forrest, M. Baaden, P.J. Bond, C. Domene, G. Patargias, J. Cuthbertson, M.S.P. Sansom, Conformational sampling and dynamics of membrane proteins from 10-ns computer simulations, Proteins 57 (2004) 783–791.
[31] B. Hess, Similarities between principal components of protein dynamics and random diffusion, Phys. Rev., E 62 (2000) 8438–8448.
[32] H.J.C. Berendsen, S. Hayward, Collective protein dynamics in relation to function, Curr. Opin. Struct. Biol. 10 (2000) 165–169.
[33] B. Rizzuti, V. Daggett, R. Guzzi, L. Sportelli, The early steps in the unfolding of azurin, Biochemistry—US 43 (2004) 15604–15609.
[34] D. Roccatano, I. Daidone, M.-A. Ceruso, C. Bossa, A. Di Nola, Selective excitation of native fluctuations during thermal unfolding simulations: horse heart cytochrome *c* as a case study, Biophys. J. 84 (2003) 1876–1883.

[35] I. Daizadeh, A.S. Medvedev, A.A. Stuchebrukhov, Effect of protein dynamics on biological electron transfer, P. Natl. Acad. Sci. USA 94 (1997) 3703–3708.
[36] I.A. Balabin, J.N. Onuchic, Dynamically controlled protein tunneling paths in photosynthetic reaction centers, Science 290 (2000) 114–117.
[37] L.J.C. Jeuken, J.P. McEvoy, F.A. Armstrong, Insights into gated electron-transfer kinetics at the electrode–protein interface: a square wave voltammetry study of the blue copper protein azurin, J. Phys. Chem., B 106 (2002) 2304–2313.
[38] W. Humphrey, A. Dalke, K. Schulten, VMD— Visual Molecular Dynamics, J. Mol. Graph. 14 (1996) 33–38.