How the Dynamics of the Metal-Binding Loop Region Controls the Acid Transition in Cupredoxins

Licia Paltrinieri,[†] Marco Borsari,[†] Gianantonio Battistuzzi,[†] Marco Sola,[‡] Christopher Dennison,[§] Bert L. de Groot,[∥] Stefano Corni,[⊥] and Carlo Augusto Bortolotti*,[‡,⊥]

[†]Department of Chemical and Geological Sciences, University of Modena and Reggio Emilia, via Campi 183, 41125 Modena, Italy
[‡]Department of Life Sciences, University of Modena and Reggio Emilia, via Campi 183, 41125 Modena, Italy
[§]Institute for Cell and Molecular Biosciences, Medical School, Newcastle University, Newcastle upon Tyne NE2 4HH, U.K.
[∥]Computational Biomolecular Dynamics Group, Max-Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany
[⊥]CNR-Nano Institute of Nanoscience, via Campi 213/A, 41125 Modena, Italy

Ⓢ Supporting Information

ABSTRACT: Many reduced cupredoxins undergo a pH-dependent structural rearrangement, triggered by protonation of the His ligand belonging to the C-terminal hydrophobic loop, usually termed the acid transition. At variance with several members of the cupredoxin family, the acid transition is not observed for azurin (AZ). We have addressed this issue by performing molecular dynamics simulations of AZ and four mutants, in which the C-terminal loop has been replaced with those of other cupredoxins or with polyalanine loops. All of the loop mutants undergo the acid transition in the pH range of 4.4−5.5. The main differences between AZ and its loop mutants are the average value of the active site solvent accessible surface area and the extent of its fluctuations with time, together with an altered structure of the water layer around the copper center. Using functional mode analysis, we found that these variations arise from changes in nonbonding interactions in the second coordination sphere of the copper center, resulting from the loop mutation. Our results strengthen the view that the dynamics at the site relevant for function and its surroundings are crucial for protein activity and for metal-containing electron transferases.
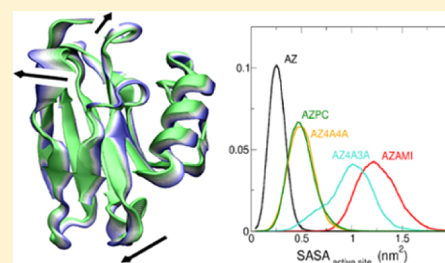
Ionization of amino acid side chains is crucial to structural and functional properties of biomolecules, such as stability, solubility, binding, and enzymatic activity.[1−3] For this reason, the activity of some proteins can be dramatically modulated by intracellular pH changes, which have been shown to be involved in the regulation of physiological processes like apoptosis and proliferation, migration, and transport.[4] In this respect, an interesting case study is provided by cupredoxins, which are copper-containing redox proteins that act as electron carriers in several crucial cellular processes in both plants and bacteria.[5−7] The strong metal equatorial ligands at their T1 copper center are provided by the thiolate sulfur of a Cys and the nitrogen atoms of two His imidazoles. A Met usually acts as a fourth, weaker axial ligand, imposing a distorted tetrahedral geometry.[6] A fifth weaker axial ligand is present in some species, as in the case of azurin (Gly45). Most of the blue copper proteins, such as plastocyanin (PC), amicyanin (AMI), pseudoazurin, and stellacyanin, undergo a pH-dependent structural rearrangement, known as the acid transition, which consists of protonation of the C-terminal His ligand in the Cu(I) protein, with pK_a values ranging from ∼7 to 5 (see ref 6 and references cited therein). As a consequence of these structural changes, the reduction potential ($E°$) of the protein increases dramatically, rendering the species biologically inactive. Several hypotheses have been formulated concerning

the possibility that this pH-induced structural rearrangement serves a physiological purpose.[8−10] In the case of PC, for example, it could downregulate photosynthetic activity, as a consequence of significant lowering of the pH in the thylakoid lumen under extreme exposure to light.[8,10,11] It is possible that the acid transition of cupredoxins endows them with the ability to act as "pH sensors", as proteins whose activities are sensitive to small, physiologically relevant, changes in pH.[4] Therefore, understanding the pH dependence of functional properties of biomolecules is central to the elucidation of their physiological activity. Moreover, the ability of a protein to undergo reversible conformational changes upon an external stimulus, such as a change in pH, with subsequent changes in one or more functional properties like $E°$ could provide a biomolecular switch, with potential applications in biomolecular electronics, biosensing, or molecular machines.[12,13]

The acid transition is not observed for azurin (AZ), possibly the most studied cupredoxin, and a pK_a value of <2 has been estimated.[14] What are the molecular determinants that hinder protonation of the His117 ligand in AZ? As three of the copper-binding residues (Cys112, His117, and Met121, AZ

numbering), including the His whose protonation triggers the acid transition, are situated on the C-terminal hydrophobic loop, previous work has been devoted to the elucidation of how the structural features of this secondary structure element could affect the pK_a of the His ligand. To this end, four AZ mutants, in which this C-terminal loop has been replaced with those of other cupredoxins (AMI and PC, thus generating the AZAMI and AZPC loop mutants, respectively) or with non-native polyalanine loops (AZ4A3A and AZ4A4A), were designed and produced.[15−17] These loop mutants were then characterized by determining their crystallographic structures and functionally via spectroscopic and electrochemical studies.[15−19]

The loop length, which was found to correlate with the pK_a of His protonation, was suggested as one of the factors affecting the thermodynamics of the acid transition, together with second-coordination sphere effects such as hydrogen bonding and π interactions[15−18,22] and differences in the solvent accessibility of the C-terminal His ligand.[17,23,24] Nevertheless, it has been difficult to state which specific molecular features relate the length of the loop and/or its composition to changes in functionality. For example, the hydrophobic loop of AZ is composed of the same number of residues as that of AZ4A3A, but only the latter species undergoes the acid transition. Therefore, additional investigations are required to elucidate the determinants of the thermodynamics of the acid transition in cupredoxins. Keeping in mind that the functionality of a protein is ultimately controlled by its dynamics,[25−27] we decided to go beyond the structure−function relationship and add the time dimension to the investigation of AZ and of its loop mutants. We therefore performed molecular dynamics simulations of AZ, AZAMI, AZPC, AZ4A3A, and AZ4A4A (whose active site structures are displayed in Figure 1), looking for dynamic aspects of cupredoxins that could play a key role in influencing the acid transition.

■ METHODS

All MD simulations were performed in water using GROMACS[28] and the AMBER99 force field.[29] As in previous work,[30] we used as the starting point of the simulations the crystal structures of AZ and loop mutants AZAMI, AZPC,
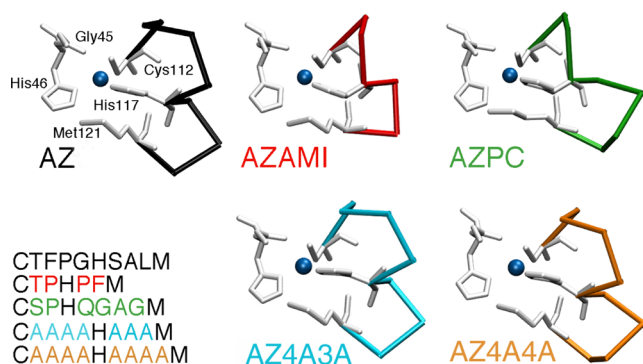
Figure 1. Active site structures of the five proteins investigated in this work. The five metal-coordinating residues are displayed as white licorice and, for the sake of clarity, are labeled only in AZ. The copper atom is displayed as a blue sphere. The hydrophobic loop is displayed as a C_α trace and colored differently for AZ and its mutants; the color code is mantained throughout the paper. At the bottom left, the amino acid sequences of the different loops are also shown. The image was prepared with VMD[20] using the crystal structures of PDB entries 4AZU,[21] 2FTA,[15] 2HX7,[16] 3FSW,[18] and 3FSZ.[18]

AZ4A3A, and AZ4A4A (PDB entries 4AZU, 2FTA, 2HX7, 3FSW, and 3FSZ, respectively). Water was represented with the TIP3P model.[31] For each starting structure, one protein molecule was solvated in a periodic rhombic dodecahedral box. The simulations were performed in the *NVT* ensemble at 300 K. The temperature was kept constant by the isokinetic temperature coupling.[32]

Because the acid transition is observed only for the reduced species,[9] simulations were performed exclusively in the reduced ensemble. Molecular dynamics does not take into account variations of the bonds between copper and its ligands, nor does it allow observation of the breaking or formation of chemical bonds. Therefore, our work did not address the reasons for the different behavior of the two oxidation states of cupredoxins with respect to the acid transition. The atomic charges of the active site, calculated at the DFT level of the theory, were taken from the literature.[33] The distances and the angles between Cu and the five coordinating atoms were constrained to the values in the corresponding crystal structures, to avoid unrealistic distortion of the copper geometry that could be caused by the treatment of the Cu−ligand interactions at a classical level of theory.[33] It is reasonable not to simulate the whole mechanism of His protonation and dissociation, because our goal is to determine whether the characteristics of the system before the transition takes place affect the propensity to undergo this process.

Periodic boundary conditions were applied to the simulation box, and long-range electrostatic interactions were treated with the particle mesh Ewald method[34] using a grid spacing of 0.12 nm combined with a fourth-order B-spline interpolation to compute the potential and forces. The real space cutoff distance and the van der Waals cutoff distance were set to 0.9 nm. A time step of 2 fs was used. Bond lenghts were constrained with the LINCS algorithm. The N- and C-termini of the proteins were modeled as NH_3^+ and CO_2^-, respectively. Positive counterions were added (Na^+) to make the simulation box neutral.

The starting structures were subjected to energy minimization in vacuum (3000 steps) using the steepest descent method, followed by a two-step minimization protocol in solvent, using the conjugate gradient method. The first minimization was performed with the coordinates of the protein held fixed, allowing only the water and the counterions to move, and the second was performed on both the protein and the solvent molecules. The temperature of the system was then increased form 50 to 300 K in 100 ps MD before the production simulations were performed. Three independent production runs (replicas) were conducted for each protein, to improve the conformational sampling.[35] The length of the first simulation was 100 ns, and the length of each additional replica was 60 ns. SASA, radial distribution functions, and average structure values were calculated using the g_sas, g_rdf, and g_rmsf tools, respectively, implemented in GROMACS version 4.5.3. Values and errors were obtained averaging over all replica simulations. Functional mode analysis was implemented as previously described.[36,37] Before analysis, trajectories in the GROMACS XTC format were least-squares fit to a reference frame to filter out overall translation and rotation.[37] The first half of the trajectory was used for model building, while the remainder served for the cross validation of the model. The number of PLS components to be used was evaluated by calculating the correlation coefficients for the model training subset (R_m) and the cross-validation subset (R_c) as a function of the number of

components, and by choosing the smallest number of latent vectors that would ensure convergence of the R_c value. Controls to assess the significance of the FMA-derived modes were conducted as follows: the trajectories for two different proteins (generically labeled as A and B) were merged into one, saving only the coordinates of a common subset of atoms (residues 1−112). The merged trajectory was least-squares fit to a reference frame to filter out overall translation and rotation. The time dependence of the SASA of the active site for proteins A and B was also merged into a single file. Again, the first half of the trajectory, belonging to protein A, was used for model building, while the second part, obtained by MD of protein B, served for the cross validation of the mode. A very low R_c value is expected if the differences between A and B are indeed significant. The Y72A AZ, P114F-AZAMI, and P114F-AZPC mutants were prepared *in silico* by homology modeling with SWISS-MODEL version 8.05[38−42] using as input the FASTA sequence of the mutants, and as the template the PDB structure of AZ (4AZU, chain A[21]), AZAMI (2FTA, chain A[15]), and AZPC (2HX7, chain A[16]), respectively. The simulation setup and the analysis tools are the same as those used for other proteins.

■ RESULTS AND DISCUSSION

The proton triggering the acid transition must be provided by the solvent, and we have therefore focused our efforts on the interplay between the protein, and in particular the region around its copper active site, and surrounding water molecules. The first quantity that we calculated was the SASA of the active site, where the active site is defined as the metal atom and its five ligands (Gly45, His46, Cys112, His117, and Met121, AZ numbering). In fact, changes in the SASA of the active site can quantitatively describe the extent of the protein−solution interface that can affect, through electrostatic and dipolar interactions, the apparent dielectric constant and (de)stabilize the protonated His117.[24,43] The calculated SASA values, averaged over three independent simulations (replicas hereafter), are listed in Table 1, and the plot of the experimental

Table 1. Experimental pK_a Values for Protonation of the C-Terminal His Ligand[a]

protein	experimental pK_a[b,c]	calculated ⟨SASA⟩ (nm^2)[b]	SASA from X-ray (nm^2)[b]
AZ	<2	0.242 (0.003)	0.109 (0.022)
AZAMI	5.5 (0.2)	1.024 (0.057)	0.979 (0.249)
AZPC	4.4 (0.2)	0.504 (0.023)	0.443 (0.044)
AZ4A3A	4.5 (0.1)	0.882 (0.071)	0.260 (0.018)
AZ4A4A	4.4 (0.1)	0.506 (0.050)	0.241 (0.035)

[a]Calculated SASA values of the protein active site, averaged over all the replicas performed in the reduced ensemble, for the five proteins under investigation. Mean of the active site SASA for the different chains of the crystal structure. [b]The associated standard error is given in parentheses. [c]Taken from refs 15, 16, and 18.
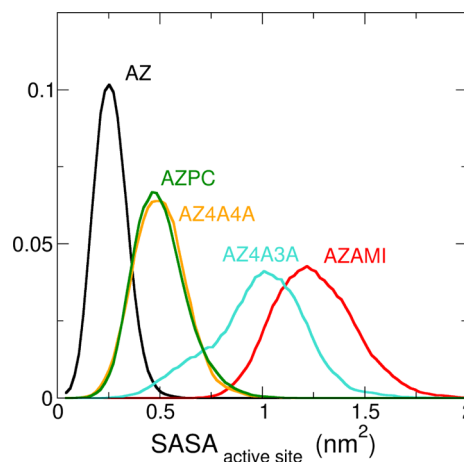
pK_a values versus the calculated active site SASAs is displayed in Figure 2. It can be readily seen that the protein featuring the highest pK_a value, AZAMI, is also the one for which the highest SASA was obtained. On the other hand, the accessibility of the active site of AZ to the solvent is much smaller than that of the other species, in accordance with its lack of an acid transition at pH >2. AZPC, AZ4A3A, and AZ4A4A all feature intermediate pK_a values, and their SASA values lie between the two extremes
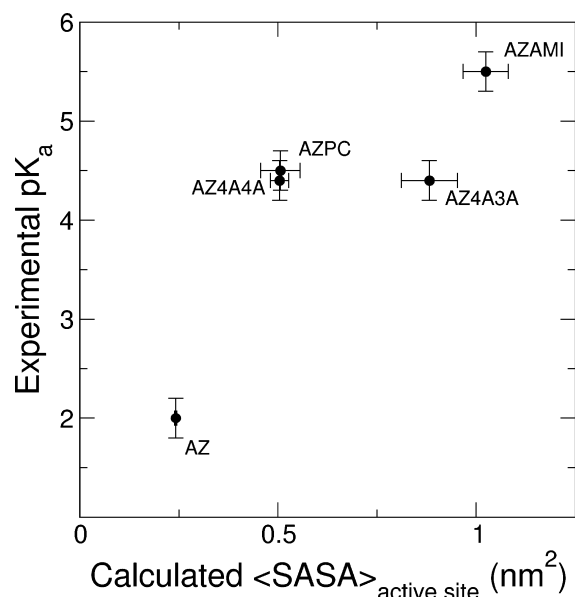
Figure 2. Experimental pK_a values for the protonation of the C-terminal His ligand vs the calculated SASA of the active site. The pK_a value for AZ could not be determined experimentally, and its corresponding data point on the plot was drawn assuming a pK_a value of 2.

represented by AZAMI and AZ. Therefore, a general trend can be observed: the more the protein active site is exposed to the surrounding solvent, the higher the corresponding pK_a of His protonation.

The SASA values displayed in Figure 2 are average quantities resulting from the whole sampling provided by our simulations and are, therefore, indicated as ⟨SASA⟩. It is also interesting to take into account the fluctuations of the SASA values with time. Thus, we calculated the normalized probability distributions of the active site SASA along the 100 ns MD simulations, and they are plotted in Figure 3. The distribution for AZ is much narrower than that of the loop mutants and never exceeds 0.5 nm^2. All the other proteins span a significantly wider range of SASA values. This finding is related to different dynamics of the five proteins along the simulation time. The conformations accessed by AZ along our simulations all feature very close

Figure 3. Normalized distribution of the calculated SASA values for the active sites of the five investigated proteins.

active site SASA values, while the fluctuations of the other species, and especially those of AZAMI and AZ4A3A, lead to a structural sampling displaying a set of conformations characterized by significant changes in the solvent accessibility of the active site. Therefore, the average active site ⟨SASA⟩ for AZ not only is the smallest but also is the one that features significantly smaller fluctuations with respect to its four mutants. The differences concerning the active site SASA among the five proteins are not simply structural but stem from the different dynamics of the molecules, as suggested by the better correlation between experimental pK_a and calculated SASA, with respect to that with SASA values obtained from crystal structures (see Table 1). Differences between the calculated SASA and the corresponding values from crystal structures are most likely due to the low temperature at which X-ray diffraction data were collected and the close packing of molecules in the crystal. In general, proteins in solution and under physiological conditions can display a structural heterogeneity larger than that shown in a crystal structure.[44,45] Nevertheless, as previously suggested,[17] a correlation can also be observed when plotting pK_a versus SASA from crystal structures (see Figure S1 of the Supporting Information), thus supporting our MD results.

Further insight into the relationship between the protein active site and the solvent may be obtained by studying the structure of the hydration layers around the metal center. This can be achieved by calculating the radial distribution function $g(r)$ between the copper and the oxygen atoms of the solvent. This is displayed in Figure 4, which provides the probability of
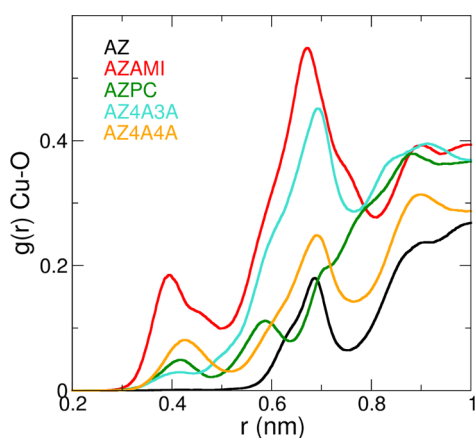

Figure 4. Cu−O radial distribution function $g(r)$. The O atoms taken into account for the calculations are only those belonging to solvent.

finding the oxygen atom of a water molecule a given distance r from the copper atom. The first sharp peak in the Cu−O radial distribution function for AZ falls slightly below 0.7 nm, as previously reported,[46] and corresponds to a solvent molecule hydrogen bonded to His117. On the other hand, all the other proteins, although invariably showing a peak at the same r value, also feature a non-zero probability of finding a water molecule at shorter distances. This is particularly evident in the case of AZAMI, whose first $g(r)$ peak falls around 0.4 nm. This difference in the solvent structure around the copper atom, which likely stems from the substitution of Phe with a much less bulky side chain,[16] as observed for the Phe114Pro mutant of AZ,[47] allows us to hypothesize that water molecules can move significantly closer to the protein active site in the loop

mutants than in AZ, thereby favoring the probability of the acid transition.

All the results collected so far provide a picture of an active site in AZ that is significantly less exposed to the solvent and less prone to undergoing structural fluctuations that could allow water molecules to approach the copper, probably as a consequence of significant differences between the dynamics of AZ and the four mutants. We therefore turned our attention to searching for the structural motions of the loop mutants that are involved in the dynamic changes of the active site SASA over time. This was achieved using functional mode analysis (FMA),[36,37] which allows the identification of the collective atomic modes of a protein that maximally correlate to a functional quantity of interest, in our case the time-dependent SASA of the active site. By applying partial least-squares (PLS)-based FMA[37] to our simulations, the protein dynamics underlying the SASA fluctuations could be detected and compared. The quality of the models calculated by FMA was evaluated by cross validation. We obtained convincing cross validation correlation coefficients (R_c) for the loop mutants (ranging from ≈0.72 to ≈0.85), indicating that a reliable collective model underlying the SASA fluctuations could be built. On the other hand, in the case of AZ, PLS-based FMA could not provide an acceptable model correlating the time dependence of active site SASA to collective protein motions. Indirectly, this again points to the different behavior of AZ with respect to that of the loop mutants. To test whether the differences between the FMA-derived motions were significant, controls were conducted (see Methods) by building an FMA model for one trajectory and cross validation by another. Extremely low R_c values (≈0.15) show that, indeed, the collective dynamics underlying SASA fluctuations differ significantly for the different cases. A three-dimensional representation of the maximally correlated modes contributing to the changes in the active site SASA is depicted in Figure 5.

The motions contributing mainly to SASA fluctuations involve concerted movements of several portions of the protein. Moreover, although the mutants all have the same protein scaffold and the only structural difference lies in the length and composition of the rather short hydrophobic loop, the detected functional modes are significantly different. This finding implies that the C-terminal loop is crucial to overall protein dynamics and, consequently, functional properties, and that modifying a few amino acids in this region can lead to dramatic changes. It is possible to discern which residues contribute most to each SASA-correlated internal mode, by calculating the root-mean-square fluctuation (rmsf) per residue within the collective movement obtained by FMA. In general, loops are the secondary structure elements most involved in the collective movements, and none of the modes describe dramatic structural rearrangements, as expected for an electron transfer protein, which in general can be considered rather rigid systems. In the case of AZAMI, the major contribution to the FMA-derived modes comes from residues belonging to the loops of residues 9−18, 36−44, and 99−107. The collective mode of AZPC is dominated by the motion of the only α-helix, spanning residues 58−71, and by the following loop, but a significant contribution also derives from the loop of residues 36−44. Similar behavior was observed for native plastocyanin, whose loops surrounding the copper atom display the most significant structural rearrangements upon protonation of the C-terminal His92 ligand.[48] For AZ4A3A, the highest rmsf values are displayed again by the residues forming the α-helix
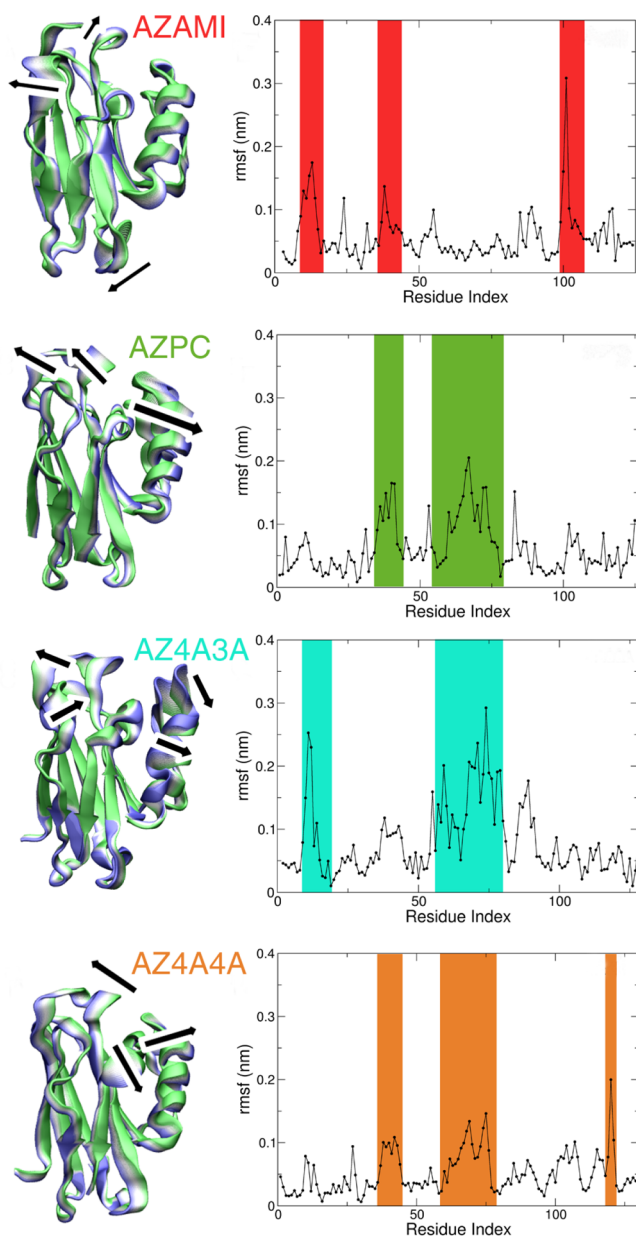
Figure 5. Backbone representation of the collective motions determining the fluctuations in the active site SASA values. No representation could be shown for AZ, as PLS-based FMA could not provide a reliable model.

and by the following loop, together with the loop of residues 9−18. The ensemble-weighted maximally correlated mode for AZ4A4A arises from the concerted movements of the α-helix, the loop of residues 36−44, and the final part of the introduced polyalanine loop. Therefore, with the exception of the involvement of the loop of residues 99−107 for AZAMI, the collective modes that cause fluctuations of the active site SASA values involve the concerted movements of the loops surrounding the copper atom and the α-helix. The mutated hydrophobic loop does not move significantly, except in AZ4A4A, as it contains three of the copper ligands whose distance to the metal atom is kept constant throughout the whole simulation time.

So far, we have shown that the ability of AZ and the loop mutants to undergo the acid transition most likely depends on the extent and fluctuations in the exposure of the active site to

solvent due to collective protein motions. The final question that needs to be addressed is why the region around the copper site in AZ is so significantly less flexible, thus hindering His protonation and detachment even at low pH values. In AZ, and, to a lesser extent, in the loop mutants, the copper atom is shielded from direct contact with the solvent by the presence of a large hydrophobic area, formed by the close packing among the loops of residues 9−18, 36−44, and 72−79 as well as by the ligand-containing loop of residues 112−121. The relative position of these loop regions is kept constant by a complex network of hydrogen bonds, all falling within the so-called "second coordination sphere".[47,49,50] The importance of some key residues that do not bind copper to the spectroscopic and functional properties of AZ (and other cupredoxins) is well-documented.[22,47,49−51]

The hydrogen bonds accepted by the thiolate group of Cys112 from the backbone amides of Asn47 and Phe114 are among the most important to the functionality of AZ.[22,47,49−51] The former is present in all the investigated species, while the Cys112−Phe114 interaction is absent in both AZAMI and AZPC.[15,16,18] This is because both variants feature a Pro in place of the Phe.[16] In both AZ4A3A and AZ4A4A, the backbone NH group of Ala114 does act as a donor of a hydrogen bond to the Cys112 thiolate, but fluctuations in this bond length in the two mutants are much higher than in AZ, often exceeding 3.5 Å, particularly for AZ4A4A (data not shown).

Another potentially important second-coordination sphere interaction in AZ is the hydrogen bond between the imidazole ring of His46 and the backbone carbonyl of Asn10, which is maintained in both AZPC and AZ4A4A. This bond is also present in the AZAMI and AZ4A3A crystal structures but is quickly and permanently lost during our simulations for both species. As a consequence, the loop spanning residues 9−18 is more free to move away from the active site, thus significantly contributing to the collective motions underlying SASA fluctuations for AZAMI and AZ4A3A (see Figure 5).

The enhanced freedom of movement of the α-helix and the loop that follows it in the four loop mutants is more difficult to justify. One tentative explanation could lie in the absence of Phe114 in the mutants with respect to AZ. The phenyl ring of this residue is located very close to the aromatic side chain of Tyr72, a residue that belongs to the loop located after the helix, and a T-shaped π−π interaction could exist between Phe114 and Tyr72 (see Figure 6A). To test this hypothesis, we monitored both the distance r between the centers of mass of the two aromatic rings and the angle θ between the normals to the ring planes along our simulations of AZ, yielding the following values: $r = 0.50 \pm 0.03$ nm, and $\theta = 63.0 \pm 0.5°$. Therefore, the relative orientation of the two lateral chains is rather constant, and the distance between the two rings does not change markedly. Moreover, these values are in agreement with those for the most stable T-shaped structure calculated *in vacuo* between tyrosine and phenylalanine side chains ($r = 0.51$ nm, and $\theta = 74.2°$).[52] Although the stacked conformation is known to be the most stable for a Tyr−Phe π−π interaction, it has also been stated that distal Phe−Tyr interactions are prevalently T-shaped,[52] as in our case, where the distance between the C_α atoms of the two residues is 0.59 ± 0.02 nm. To determine whether the Tyr72−Phe114 interaction is important for the structural stability of AZ, we created *in silico* and simulated the Y72A mutant of azurin, which cannot have a π−π interaction with Phe114. Removal of the aromatic side
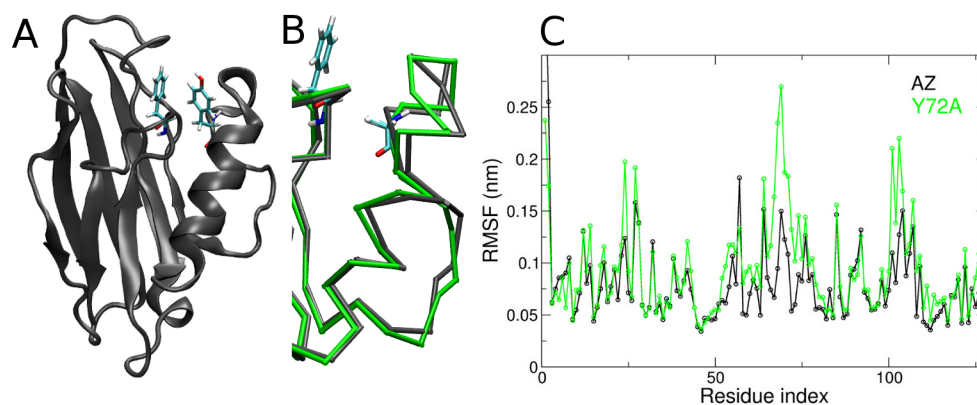
Figure 6. (A) Cartoon representation of AZ. Tyr72 and Phe114 are represented as licorice sticks. (B) Superimposition of the average structure along the MD simulation obtained for WT azurin (black) and the Y72A mutant (green) with Phe114 and the mutated Ala72 displayed as licorice sticks. (C) Comparison of the rmsf per residue for AZ (black) and Y72A (green).

chain of Tyr72 leads to a higher degree of freedom of movement of the α-helix region, as witnessed by the comparison between the rmsf per residue of WT AZ and Y72A. In fact, while for AZ the region encompassing residues 50−80 does not undergo significant fluctuation throughout the whole MD sampling, the same protein region for the Y72A variant displays the highest rmsf values (Figure 6).

Further insights into the role played by Phe114 were provided by the construction and modeling *in silico* of the P114F mutants of AZAMI and AZPC. In the case of P114F-AZAMI, reverting Pro to Phe leads to a dramatic decrease in the active site SASA, whose value for P114F-AZAMI is 0.684 ± 0.019 nm^2, significantly smaller than that for AZAMI. The $g(r)$ plot for P114F-AZAMI is also remarkably different from that of AZAMI, because of the removal of the 0.4 nm peak (see Figure S2A of the Supporting Information). The differences between AZAMI and its mutant can be ascribed to restoration of the π−π interaction between the aromatic rings of Phe114 and Tyr72, as witnessed by the r and θ values for the P114F-AZAMI variant (see Figure S2B of the Supporting Information). On the other hand, the P114F mutation in AZPC does not affect the active site SASA and the changes in the $g(r)$ plot are opposite to those observed for P114F-AZAMI. It is important to stress that it in AZAMI and AZPC the newly introduced Phe is adjacent to the His ligand, whereas in AZ, there are two intervening residues: therefore, care should be taken when directly comparing the P114F mutants to AZ, and only indirect inferences can be obtained by these additional simulations. However, these results suggest that Phe114 is one of the structural attributes contributing to the differences between the species but is not the only factor affecting the protein−solvent interplay, which is determined by a number of structural and dynamical features.

Overall, our results are in agreement with previous works that investigated the internal mobility of cupredoxins, which demonstrate that these proteins, even under native conditions, can access multiple conformations involving weakly populated states that are essential for their functionality.[53,54] Similar findings were also obtained for other systems[44] and, in particular, cytochrome *c*,[27,55] another ET protein that is usually considered to be rigid.

CONCLUSIONS

We have investigated the molecular effectors of the acid transition in cupredoxins by comparing the dynamic behavior

of WT AZ and four variants in which the C-terminal hydrophobic loop was mutated. We find that the SASA of the protein active site largely affects the pK_a of the C-terminal His ligand on this loop, which triggers the whole structural transition. Notably, the average SASA value seems to be crucial to the pH-dependent rearrangement along with the extent of its flucutations over time. Indeed, the effect of loop mutations is not only structural but also dynamic. Another significant difference between AZ and its loop mutants concerns the structure of the water layer surrounding the protein active site, as witnessed by the copper−oxygen radial distribution function $g(r)$. The absence of peaks at distances of <0.7 nm for AZ, in contrast with all the other investigated species, suggests that for the loop mutants solvent molecules can approach the active site, thus enhancing the probability of the acid transition. To identify the protein regions mostly contributing to the SASA fluctuations, we performed FMA-based investigations. These allowed us to ascribe the different behavior of the loop mutants to the enhanced freedom of movement of the loops surrounding the active site with respect to AZ, which is a consequence of the loss of some crucial second-coordination sphere interactions. Our results converge to strengthen the growing awareness that the functional properties of ET proteins are ultimately determined by the dynamic interplay between the protein scaffold and its environment. Similar conclusions have already been drawn for several biomolecules, very different from ET species in both structure and function, including membrane voltage channels.[56,57]

ASSOCIATED CONTENT

S Supporting Information

Details of the comparison between experimental pK_a and XRD active site SASA for all the investigated mutants. Details of the interaction between Phe114 and Tyr72 for P114F-AZAMI and P114F-AZPC mutants. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +39 059 2055096. E-mail: carloaugusto.bortolotti@ unimore.it.

interaction on biological electron transfer. C.A.B. thanks Carabel s.r.l. for financial support.

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

AZ, azurin; MD, molecular dynamics; PC, plastocyanin; AMI, amicyanin; VMD, Visual Molecular Dynamics; SASA, solvent accessible surface area; FMA, functional mode analysis; PLS, partial least-squares; rmsf, root-mean-square fluctuation; WT, wild type; ET, electron transfer; PDB, Protein Data Bank.

REFERENCES

(1) Di Russo, N., Estrin, D., Marti, M., and Roitberg, A. (2012) pH-Dependent Conformational Changes in Proteins and Their Effect on Experimental pK$_a$s: The Case of Nitrophorin 4. *PLoS Comput. Biol. 8*, e1002761.

(2) Grimsley, G. R., Scholtz, J. M., and Pace, C. N. (2009) A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci. 18*, 247−251.

(3) Narayanan, A., and Jacobson, M. P. (2009) Computational studies of protein regulation by post-translational phosphorylation. *Curr. Opin. Struct. Biol. 19*, 156−163.

(4) Srivastava, J., Barber, D. L., and Jacobson, M. P. (2007) Intracellular pH sensors: Design principles and functional significance. *Physiology 22*, 30−39.

(5) Gray, H. B., Malmström, B. G., and Williams, R. (2000) Copper Coordination in Blue Proteins. *JBIC, J. Biol. Inorg. Chem 5*, 551−559.

(6) Dennison, C. (2005) Investigating the Structure and Function of Cupredoxins. *Coord. Chem. Rev. 249*, 3025−3054.

(7) Savelieff, M. G., Wilson, T. D., Elias, Y., Nilges, M. J., Garner, D. K., and Lu, Y. (2008) Experimental evidence for a link among cupredoxins: Red, blue, and purple copper transformations in nitrous oxide reductase. *Proc. Natl. Acad. Sci. U.S.A. 105*, 7919−7924.

(8) Guss, J., Harrowell, P., Murata, M., Norris, V., and Freeman, H. (1986) Crystal structure analyses of reduced (CuI) poplar plastocyanin at six pH values. *J. Mol. Biol. 192*, 361−387.

(9) Battistuzzi, G., Borsari, M., Canters, G., deWaal, E., Leonardi, A., and Sola, M. (2002) Thermodynamics of the acid transition in blue copper proteins. *Biochemistry 41*, 14293−14298.

(10) Hass, M. A. S., Christensen, H. E. M., Zhang, J., and Led, J. J. (2007) Kinetics and mechanism of the acid transition of the active site in plastocyanin. *Biochemistry 46*, 14619−14628.

(11) Sato, K., Kohzuma, T., and Dennison, C. (2003) Active-Site Structure and Electron-Transfer Reactivity of Plastocyanins. *J. Am. Chem. Soc. 125*, 2101−2112.

(12) Bortolotti, C. A., Paltrinieri, L., Monari, S., Ranieri, A., Borsari, M., Battistuzzi, G., and Sola, M. (2012) A surface-immobilized cytochrome c variant provides a pH-controlled molecular switch. *Chem. Sci. 3*, 807−810.

(13) Boyle, M. M., Smaldone, R. A., Whalley, A. C., Ambrogio, M. W., Botros, Y. Y., and Stoddart, J. F. (2011) Mechanised materials. *Chem. Sci. 2*, 204−210.

(14) Jeuken, L. J. C., van Vliet, P., Verbeet, M. P., Camba, R., McEvoy, J. P., Armstrong, F. A., and Canters, G. W. (2000) Role of the Surface-Exposed and Copper-Coordinating Histidine in Blue Copper Proteins: The Electron-Transfer and Redox-Coupled Ligand Binding Properties of His117Gly Azurin. *J. Am. Chem. Soc. 122*, 12186−12194.

(15) Li, C., Yanagisawa, S., Martins, B. M., Messerschmidt, A., Banfield, M. J., and Dennison, C. (2006) Basic Requirements for a Metal-Binding Site in a Protein: The Influence of Loop Shortening on the Cupredoxin Azurin. *Proc. Natl. Acad. Sci. U.S.A. 103*, 7258−7263.

(16) Li, C., Banfield, M. J., and Dennison, C. (2007) Engineering Copper Sites in Proteins: Loops Confer Native Structures and Properties to Chimeric Cupredoxins. *J. Am. Chem. Soc. 129*, 709−718.

(17) Li, C., Sato, K., Monari, S., Salard, I., Sola, M., Banfield, M. J., and Dennison, C. (2011) Metal-Binding Loop Length Is a Determinant of the pK$_a$ of a Histidine Ligand at a Type 1 Copper Site. *Inorg. Chem. 50*, 482−488.

(18) Sato, K., Li, C., Salard, I., Thompson, A. J., Banfield, M. J., and Dennison, C. (2009) Metal-Binding Loop Length and not Sequence Dictates Structure. *Proc. Natl. Acad. Sci. U.S.A. 106*, 5616−5621.

(19) Battistuzzi, G., Borsari, M., Dennison, C., Li, C., Ranieri, A., Sola, M., and Yanagisawa, S. (2009) Active site loop dictates the thermodynamics of reduction and ligand protonation in cupredoxins. *Biochim. Biophys. Acta 1794*, 995−1000.

(20) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual Molecular Dynamics. *J. Mol. Graphics 14*, 33−38.

(21) Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G. (1991) 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*, 765−772.

(22) Yanagisawa, S., Crowley, P. B., Firbank, S. J., Lawler, A. T., Hunter, D. M., McFarlane, W., Li, C., Kohzuma, T., Banfield, M. J., and Dennison, C. (2008) π-Interaction Tuning of the Active Site Properties of Metalloproteins. *J. Am. Chem. Soc. 130*, 15420−15428.

(23) Battistuzzi, G., Borsari, M., Loschi, L., and Sola, M. (1997) Redox thermodynamics, acid-base equilibria and salt-induced effects for the cucumber basic protein. General implications for blue-copper proteins. *JBIC, J. Biol. Inorg. Chem. 2*, 350−359.

(24) Hunter, D. M., McFarlane, W., Sykes, A. G., and Dennison, C. (2001) Effect of pH on the Self-Exchange Reactivity of the Plant Plastocyanin from Parsley. *Inorg. Chem. 40*, 354−360.

(25) Karplus, M., and Kuriyan, J. (2005) Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U.S.A. 102*, 6679−6685.

(26) Henzler-Wildman, K., and Kern, D. (2007) Dynamic Personalities of Proteins. *Nature 450*, 964−972.

(27) Bortolotti, C. A., Amadei, A., Aschi, M., Borsari, M., Corni, S., Sola, M., and Daidone, I. (2012) The Reversible Opening of Water Channels in Cytochrome c Modulates the Heme Iron Reduction Potential. *J. Am. Chem. Soc. 134*, 13670−13678.

(28) Berendsen, H., van der Spoel, D., and van Drunen, R. (1995) GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun. 91*, 43−56.

(29) Wang, J., Cieplak, P., and Kollman, P. A. (2000) How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem. 21*, 1049−1074.

(30) Paltrinieri, L., Borsari, M., Ranieri, A., Battistuzzi, G., Corni, S., and Bortolotti, C. A. (2013) The Active Site Loop Modulates the Reorganization Energy of Blue Copper Proteins by Controlling the Dynamic Interplay with Solvent. *J. Phys. Chem. Lett. 4*, 710−715.

(31) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys. 79*, 926−936.

(32) Brown, D., and Clarke, J. H. R. (1984) A Comparison of Constant Energy, Constant Temperature, and Constant Pressure Ensembles in Molecular Dynamics Simulations of Atomic Liquids. *Mol. Phys. 51*, 1243−1252.

(33) van den Bosch, M., Swart, M., Snijders, J. G., Berendsen, H. J. C., Mark, A. E., Oostenbrink, C., van Gunsteren, W. F., and Canters, G. W. (2005) Calculation of the Redox Potential of the Protein Azurin and Some Mutants. *ChemBioChem 6*, 738−746.

(34) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.

(35) Papaleo, E., Ranzani, V., Tripodi, F., Vitriolo, A., Cirulli, C., Fantucci, P., Alberghina, L., Vanoni, M., De Gioia, L., and Coccetti, P. (2011) An acidic loop and cognate phosphorylation sites define a molecular switch that modulates ubiquitin charging activity in Cdc34-like enzymes. *PLoS Comput. Biol. 7*, e1002056.

(36) Hub, J. S., and de Groot, B. L. (2009) Detection of functional modes in protein dynamics. *PLoS Comput. Biol. 5*, e1000480.

(37) Krivobokova, T., Briones, R., Hub, J. S., Munk, A., and de Groot, B. L. (2012) Partial least-squares functional mode analysis: Application

to the membrane proteins AQP1, Aqy1, and CLC-ec1. *Biophys. J. 103*, 786−796.

(38) Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics 22*, 195−201.

(39) Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res. 37*, D387−D392.

(40) Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: An automated protein homology-modelling server. *Nucleic Acids Res. 31*, 3381−3385.

(41) Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis 18*, 2714−2723.

(42) Peitsch, M. C. (1995) Protein modeling by E-mail. *Nat. Biotechnol. 13*, 658−660.

(43) Isom, D. G., Castañeda, C. A., Cannon, B. R., and García-Moreno, B. (2011) Large Shifts in pK_a Values of Lysine Residues Buried Inside a Protein. *Proc. Natl. Acad. Sci. U.S.A. 108*, 5260−5265.

(44) Henzler-Wildman, K., and Kern, D. (2007) Dynamic personalities of proteins. *Nature 450*, 964−972.

(45) Blake, P. R., Day, M. W., Hsu, B. T., Joshua-Tor, L., Park, J., Hare, D. R., Adams, M. W. W., Rees, D. C., and Summers, M. F. (1992) Comparison of the X-ray structure of native rubredoxin from *Pyrococcusfuriosus* with the NMR structure of the zinc-substituted protein. *Protein Sci. 1*, 1522−1525.

(46) Cascella, M., Magistrato, A., Tavernelli, I., Carloni, P., and Rothlisberger, U. (2006) Role of Protein Frame and Solvent for the Redox Properties of Azurin from *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A. 103*, 19641−19646.

(47) Yanagisawa, S., Banfield, M. J., and Dennison, C. (2006) The Role of Hydrogen Bonding at the Active Site of a Cupredoxin: The Phe114Pro Azurin Variant. *Biochemistry 45*, 8812−8822.

(48) Hass, M. A. S., Thuesen, M. H., Christensen, H. E. M., and Led, J. J. (2004) Characterization of μs-ms Dynamics of Proteins Using a Combined Analysis of ^{15}N NMR Relaxation and Chemical Shift: Conformational Exchange in Plastocyanin Induced by Histidine Protonations. *J. Am. Chem. Soc. 126*, 753−765.

(49) Warren, J. J., Lancaster, K. M., Richards, J. H., and Gray, H. B. (2012) Inner- and outer-sphere metal coordination in blue copper proteins. *J. Inorg. Biochem. 115*, 119−126.

(50) Wilson, T. D., Yu, Y., and Lu, Y. (2013) Understanding copper-thiolate containing electron transfer centers by incorporation of unnatural amino acids and the CuA center into the type 1 copper protein azurin. *Coord. Chem. Rev. 257*, 260−276.

(51) Marshall, N. M., Garner, D. K., Wilson, T. D., Gao, Y., Robinson, H., Nilges, M. J., and Lu, Y. (2009) Rationally Tuning the Reduction Potential of a Single Cupredoxin Beyond the Natural Range. *Nature 462*, 113−116.

(52) Chelli, R., Gervasio, F. L., Procacci, P., and Schettino, V. (2002) Stacking and T-shape competition in aromatic-aromatic amino acid interactions. *J. Am. Chem. Soc. 124*, 6133−6143.

(53) Korzhnev, D. M., Karlsson, B. G., Orekhov, V. Y., and Billeter, M. (2003) NMR detection of multiple transitions to low-populated states in azurin. *Protein Sci. 12*, 56−65.

(54) Ma, L., Hass, M. A. S., Vierick, N., Kristensen, S. M., Ulstrup, J., and Led, J. J. (2003) Backbone Dynamics of Reduced Plastocyanin from the Cyanobacterium *Anabaenavariabilis*: Regions Involved in Electron Transfer Have Enhanced Mobility. *Biochemistry 42*, 320−330.

(55) Krishna, M. M. G., Lin, Y., Rumbley, J. N., and Englander, S. W. (2003) Cooperative Omega Loops in Cytochrome *c*: Role in Folding and Function. *J. Mol. Biol. 331*, 29−36.

(56) Alessandrini, A., and Facci, P. (2011) Changes in single K channel behavior induced by a lipid phase transition. *Commun. Integr. Biol. 4*, 346−348.

(57) Alessandrini, A., and Facci, P. (2011) Unravelling lipid-protein interaction in model bilayers by AFM. *J. Mol. Recognit. 24*, 387−396.