CMLS, Cell. Mol. Life Sci. 61 (2004) 1123–1142
1420-682X/04/101123-20
DOI 10.1007/s00018-004-3181-5

CMLS Cellular and Molecular Life Sciences

Review

Computational approaches to structural and functional analysis of plastocyanin and other blue copper proteins

F. De Rienzo[a, b, *], R. R. Gabdoulline[c], R. C. Wade[c], M. Sola[a, b] and M. C. Menziani[a]

[a] Chemistry Department, University of Modena and Reggio Emilia, Via Campi 183, 41100 Modena (Italy), e-mail: derienzo.francesca@unimore.it
[b] INFM National Center on nanoStructures and bioSystems at Surfaces (S3), Modena (Italy)
[c] EML Research gGmbH, Villa Bosch, Schloss-Wolfsbrunnenweg 33, 69118 Heidelberg (Germany)

Received 9 May 2003; received after revision 24 November 2003; accepted 28 November 2003

Abstract. Computational techniques are becoming increasingly important in structural and functional biology, in particular as tools to aid the interpretation of experimental results and the design of new systems. This review reports on recent studies employing a variety of computational approaches to unravel the microscopic details of the structure-function relationships in plastocyanin and other proteins belonging to the blue copper superfamily. Aspects covered include protein recognition, electron transfer and protein-solvent interaction properties of the blue copper protein family. The relevance of integrating diverse computational approaches to address the analysis of a complex protein system, such as a cupredoxin metalloprotein, is emphasized.

Key words. Plastocyanin; blue copper proteins; electron transfer; protein-protein association; structure-function relationship; QSAR; computational analysis.

Introduction

Interest in theoretical approaches to biological sciences has increased dramatically during the last decade. This has been spurred by the huge amount of experimental information now available on many relevant biological systems due to the development of molecular biology and genetic engineering techniques. Advances in algorithms and computer hardware have increased the potential for computational approaches to aid the interpretation of experimental results and the design of further experiments, as well as to provide detailed insights into biological processes that are unattainable experimentally. Computational modeling and simulation techniques permit the study of the structure-dynamics-function relationships of a single protein molecule on timescales from femto- to milliseconds. The effects of structural modifications of molecules on the biological reactions in which they are involved can be analyzed for individual molecules and translated to effects on observable macroscopic biochemical parameters, such as reaction rates, association constants, rate constants for electron transfer (ET), redox potentials etc.

In this review, we will discuss recent computational studies directed towards the structural and functional characterization of the cupredoxins, a superfamily of redox proteins, with particular attention to plastocyanin (Pc). The overall aim is to understand whether and how the function of these proteins is controlled by the protein structure and physicochemical properties of the solvent environment. The structural and functional characterization, which is preliminary to this analysis, is also featured. The papers related to these topics and published in the period from 2000 to early 2003 are analyzed in detail, but previous works are also briefly reported where relevant. This review is complementary to several recent reviews which deal with electron transfer and association in redox pro-

* Corresponding author.

teins, either by discussing the experimental and theoretical perspectives [1–3], or by pointing out recent computational developments [4, 5].

Redox reactions are amongst the simplest chemical reactions and play a primary role in biology. They are involved in fundamental processes for life, such as respiration, photosynthesis, metabolism and catabolism. Thus, understanding the mechanisms of these reactions will improve our basic knowledge of biological phenomena. For instance, reproducing the cascade of physicochemical reactions that transform light energy into chemical energy during photosynthesis is an intriguing challenge for the research community in view of the fact that it can provide virtually all the energy required for all living organisms on our planet to survive. Besides this, the investigation of electron transfer in nature also impacts on electronic technology aimed at exploiting biomolecules, and in particular ET proteins, for implementing new nanodevices [6, 7].

The cupredoxins, also known as blue copper (BC) proteins, are small (10–20 kDa) soluble copper proteins that function as ET agents shuttling electrons, one by one, from a donor to an acceptor protein in the respiratory and photosynthetic chains of many bacteria and plants. The barrel-like structure of cupredoxins is represented in figure 1 (this picture was drawing using MOLSCRIPT [8]). ET between a cupredoxin and one of its partners is mediated by formation of a dynamic and transient complex which is hard to study due to its short life and energetic instability [1, 3]. The difficulty of obtaining structural information about such complexes reduces the possibilities to relate structural data to experimental findings and, therefore, protein structure to function. In addition, the complexes formed seem to undergo reorganization prior to electron transfer [1, 3, 4], thus further complicating the resulting picture. The overall ET process can be written as a sequence of reaction steps which include (i) formation of the complex between the two redox partners, which is necessary to bring the proteins into contact (eq. 1); (ii) structural reorganization of the complex to bring the proteins into the right geometry for electron transfer (eq. 2); (iii) the electron transfer itself (eq. 3); and (iv) dissociation of the complex after the redox reaction has taken place (eq. 4):

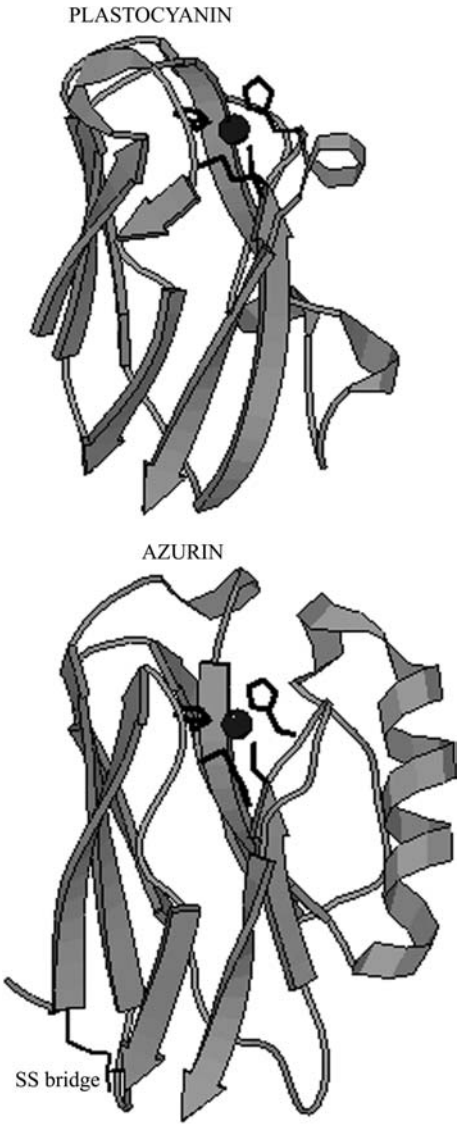

Figure 1. Representation of the crystal structures of two cupredoxins: Pc from spinach (pdb code: 1ag6) *(top)* and Az from *P. aeruginosa* (pdb code: 4azu) *(bottom)*. The Cu ion is in CPK and the Cu ligands are in stick. The secondary structure elements are highlighted: arrows represent β strands and cylinders represent α helices. The picture was drawn using MOLSCRIPT [8].

$$A_{red} + B_{ox} \underset{k_{off}}{\overset{k_{on}}{\longleftrightarrow}} (A_{red}/B_{ox})' \quad (1)$$

$$(A_{red}/B_{ox})' \underset{k_{der}}{\overset{k_{reorg}}{\longleftrightarrow}} (A_{red}/B_{ox})'' \quad (2)$$

$$(A_{red}/B_{ox})'' \underset{k_{rET}}{\overset{k_{ET}}{\longleftrightarrow}} (A_{ox}/B_{red}) \quad (3)$$

$$(A_{ox}/B_{red}) \underset{k_{on'}}{\overset{k_{off'}}{\longleftrightarrow}} A_{ox} + B_{red} \quad (4)$$

where k_{on} and k_{off} are the rate constants for the association process, k_{reorg} and k_{der} are the rate constants for the direct and reverse reorganization reaction, k_{ET} and k_{rET} are the rate constants for the direct and reverse ET reaction, and $k_{off'}$ and $k_{on'}$ are the rate constants for the complex dissociation reactions (which are not necessarily the same as k_{off} and k_{on}). Now, it is evident that the overall process is influenced by both the association and ET steps, which depend respectively on the interaction and redox properties of the proteins, which, in turn, depend on the structural features of the protein. In fact, the redox properties of the cupredoxins are due to the copper atom and to the

spatial disposition of the amino acid copper ligands in the protein. On the other hand, the ability of the cupredoxins to complex their protein partners is due to their overall protein conformation and dynamics in solution.

Therefore, great theoretical and computational effort has been devoted by several research laboratories to (i) describe the cupredoxin structure, encompassing both the copper site geometry and the overall protein structure and dynamics in relation to their environment; (ii) unravel the molecular mechanisms underlying cupredoxin function, especially the association and electron transfer processes; and (iii) understand and rationalize the relationship between cupredoxin structure and function.

The cupredoxin structure

The copper site

The metal site of the BC proteins has long been studied by researchers interested in the unusual physicochemical properties of its oxidized form, i.e. the bright blue color, narrow hyperfine splitting in its electron spin resonance spectra and high reduction potentials with respect to copper aqueous complexes. Many theoretical computations have been performed to explain these peculiarities and to reproduce the electronic and spectroscopic features [9–11]. In this section, we will focus our attention on the geometric structure of the Cu site and comment only on the results and conclusions reported most recently.

In the cupredoxin-like structure, the copper (fig. 2) is bound to the protein matrix by three strong ligands (the distance between the Cu atom and the ligand atom is $d_{Cu\text{-}L} \approx 2$ Å) arranged in an almost trigonal plane. These ligands are the N^{δ} atoms of two histidines and the S^{γ} atom of a cysteine thiolate. In addition, the copper is coordinated by one or two weaker axial ligands ($d_{Cu\text{-}L} \approx 3$ Å), typically the S^{γ} atom of a methionine and the carbonyl O atom of a glycine (in azurin) [12] (fig. 2) or the amide O atom of a glutamine (in stellacyanin) [13]. The Cu site has a distorted trigonal geometry that is intermediate between that preferred by Cu(I) complexes (tetrahedral) and that preferred by Cu(II) complexes (tetragonal). Thus, the change in geometry with redox is limited, and the energy required for this structural rearrangement is small. These observations led to the induced rack theory [14] and the entatic state theory [15], based on the hypotheses that the protein matrix forces the Cu(II) site into a catalytically poised geometry similar to the Cu(I) geometry. These theories, which trace the peculiar properties of the cupredoxins and their electron transfer capability back to the strained geometry of their active site, have been challenged recently [9, 10], and the issue has been addressed through different theoretical methods.

Ryde et al. [9, 10] performed quantum-mechanical (QM) optimization in vacuo of the geometry of both the oxidized and reduced Cu sites, using the B3LYP hybrid density functional (DFT) method [10, 16]. As a QM treatment of the whole protein is not tractable, they constructed models of increasing complexity, including only the copper and its amino acid ligands. The best representative minimum model for the prototypical Cu site (i.e. the Pc site with two His, a Met and a Cys in the coordination sphere of copper) was found to be that with imida-

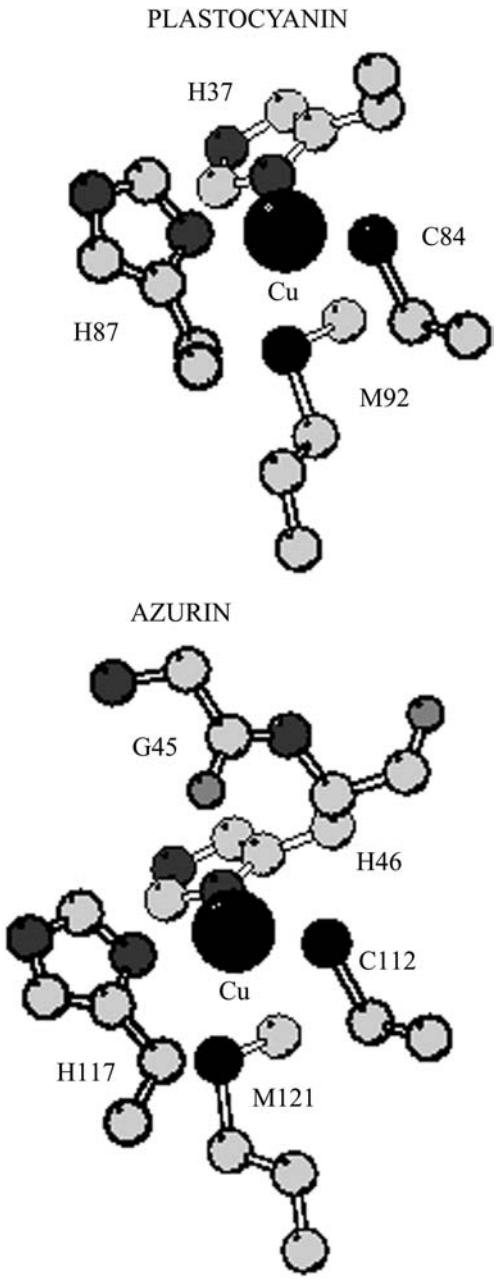

Figure 2. Representation of the Cu-site structures of Pc from spinach *(top)* and Az from *P. aeruginosa (bottom)*. The Cu ion is in CPK and the Cu ligands are in ball and stick. Color code: black, copper and sulfur atoms; dark gray, nitrogen atoms; gray, oxygen atoms; light gray, carbon atoms. The picture was drawn using MOLSCRIPT [8].

zoles (Im), methylthiolate (CH_3S^-), and dimethyl sulfur [$(CH_3)_2S$], modeling the histidines, the cysteine, and the methionine, respectively. Surprisingly, the results of the calculations on the copper site were in clear contrast with the rack and entatic state hypotheses. The in vacuo optimized geometry of the isolated Cu(II) site was highly similar to the experimental geometry of the site in the protein structure. This suggested that the geometry of the Cu(II) site in cupredoxins is not strained by the protein matrix. The main discrepancies with the experimental complexes (determined in aqueous solution) were found in the Cu-S_{Cys} distance, which was slightly (0.1 – 0.2 Å) too long, and in the Cu-S_{Met} distance, which was slightly (0.2 – 0.5 Å) too short.

Similar calculations on the azurin (Az) site, with a fifth additional carbonyl ligand covalently bound to one of the Cu-ligating histidines through the backbone ($ImCH_2CH_2NHCOCH_3$), and on the stellacyanin site, with CH_3CONH_2 to model glutamine instead of methionine, further confirmed the previous results. To test their results further, Ryde et al. [10] also took into account (i) the effect of the solvent on the copper site geometry, by performing geometry optimization of the Cu(I) prototype site in an implicit solvent model with varying dielectric constant (ε) [17]; and (ii) the effect of the protein surroundings on the active site structure, by applying a combined quantum-mechanical and molecular mechanics (QM/MM) approach to the whole protein structure [18]. This allowed treatment of the interesting part of the system (i.e. the copper site) with the DFT-B3LYP functional [16] and of the rest of the protein with the AMBER molecular mechanics force field [19, 20]. The results showed that the experimental structures were better reproduced by complete treatment of the site environment, and the resulting Cu-site geometry improved, becoming more similar to that found in the cupredoxin crystal structures. The energetic strain (between 31 and 50 kJ/mol), computed as the difference between the isolated in vacuo site geometry and the QM/MM optimized site geometry, was evaluated to be in the normal range for incorporation of a metal ion into a protein. Ryde et al. [10] concluded therefore that the Cu(II) site in BC proteins is not mechanically strained by the protein matrix and that the peculiar properties of BC proteins are rather due to the physicochemical effects of the copper ligands, especially of the coordinating Cys. More precisely, they argued that the peculiar electronic structure of the Cu-S_{Cys} bond gives reasons for both the tetrahedral coordination structure and the physicochemical properties of the site.

Comba et al. [21], although applying a similar approach, came to rather different conclusions. They used a mixed QM/MM method to analyze the oxidized structures of two other cupredoxins, amicyanin (Ami) and rusticyanin (Rus). The QM region was defined by the minimum model Cu(Im)[$S(CH_3)_2$]SH, which is smaller than the one used by Ryde et al. [10], and the calculations were performed with approximate DFT-B3LYP functional [16]. The MM part, which was calculated with the MM3 force field [22], also differs from the calculations of Ryde et al. [10], which considered the whole protein structure. Here, the MM part included only the C-terminal loop containing the downstream Cu-ligands (residues 92 – 98 for Ami and 138 – 148 for Rus), and the backbone of the upstream Cu-coordinating histidine (H53 for Ami and H85 for Rus) (fig. 3). The atoms at the border between the QM/MM optimized fragment and the protein backbone were frozen in their initial positions, to account for the effects of the protein matrix excluded from the calculations. The comparison of the QM/MM structure and the DFT optimized minimum model allowed the authors to evaluate the influence of the backbone on the active site geometry. The results showed that in both Ami and Rus, the most significant differences in the QM/MM system are related to the angles and distances involving the S_{Met} (for instance, the Cu-S_{Met} bond is elongated by 0.2 – 0.4 Å). The strained energy was computed to be 74 $kJmol^{-1}$ for Ami and 82 $kJmol^{-1}$ for Rus. These energy values are slightly higher than those determined by Ryde et al. [10], supporting the

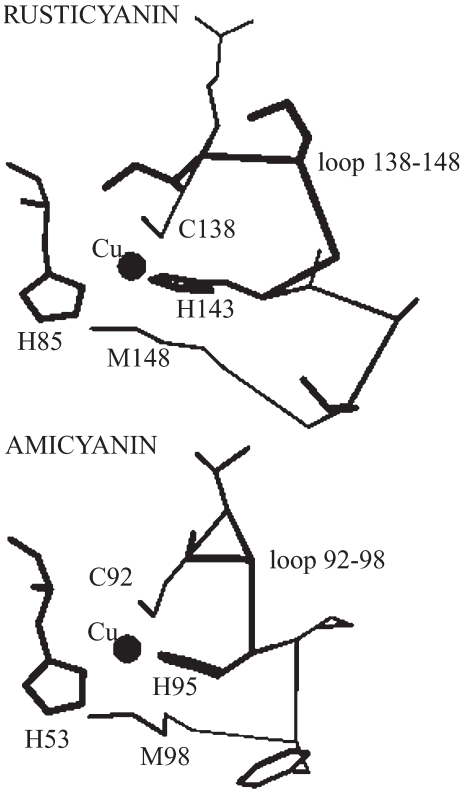

Figure 3. Representation of the Cu-site structures of rusticyanin from *Thiobacillus ferrooxidans* (pdb code: 1rcy) *(top)* and amicyanin from *P. denitrificans* (pdb code: 1aac) *(bottom)*, with the Cu-coordinating residues and the downstream loop used in the QM/MM calculations by Comba et al. [21]. The picture was drawn using MOLSCRIPT [8].

idea of some strain being present in the structure. Thus, the authors concluded that the protein matrix has some entatic influence on the Cu site geometry and that this is connected with the enforcement of the Cu-S_{Met} distance.

Protein structure and dynamics

Small changes in the active site geometry of BC proteins can tune their redox properties. Similarly, conformational fluctuations of the protein structure can affect not only protein recognition and binding but also electron transfer properties. Small movements of atoms in protein regions far from the copper site can induce fluctuations of residues near the Cu site and/or in regions involved in protein binding. These internal dynamic processes, which are essential to protein functionality, depend on protein structural features and might be influenced by mutation of the protein sequence.

Recently, Arcangeli et al. [23] performed molecular dynamics (MD) simulations and essential dynamics analyses of BC proteins in order to identify large concerted protein motions that can be functional to recognition and ET processes. They analyzed and compared the dynamics of the fully hydrated structures of Pc from poplar and Az from *Pseudomonas aeruginosa*. These two proteins differ both in their active site architecture and in their three-dimensional (3D) structure (figs 1 and 2). Az, which has a longer sequence than Pc, is characterized by a five Cu-ligand site (instead of the four residue Pc Cu site) (fig. 2), an extra α helix protruding outside the barrel-like fold and a surface disulfide bridge located at the opposite end of the protein structure with respect to the Cu site (fig. 1).

The results of the essential dynamics analyses showed that the predominant internal motions of the proteins are restricted to a subspace of a few degrees of freedom. In fact, the anharmonic essential motions crucial for the proteins to perform protein binding and electron transfer could be described by three or four eigenvectors. The analysis of protein motions along these eigenvectors highlighted the presence of correlated movements between functional regions (likely involved in recognition or electron transfer) and structural regions far from them. Concerted motions of Pc mainly involve turns, in particular the loop containing the downstream Cu ligands (fig. 1), the hydrophobic loop in the northern end of the protein, containing the highly conserved residue L12, and the two acidic patches highlighted in figure 4. In Az, the region opposite to the Cu site is relatively rigid, due to the presence of the disulfide bridge; thus, concerted motions involve mainly flexible hydrophobic regions near the Cu site which are considered relevant for mediating protein binding.

The MD and essential dynamics studies of a Pc mutant were also performed to evaluate the effects of structural mutations on protein dynamics [24]. The poplar Pc sequence was mutated to obtain the insertion of a surface disulfide bridge opposite to the Cu site end. The MD simulation highlighted that similarly to the wild-type (wt) Pc, the mutant maintains both its globular shape and secondary structure during simulation, and the structure and geometry of the Cu site are almost unaffected by introduction of the disulfide bridge. However, analysis of the H-bond network revealed that some H-bond interactions involving residues near the Cu site in the wt Pc (and thought to be relevant for Pc function) are lost or modified in the mutant protein. This causes higher flexibility and larger fluctuations in the regions around Cu in the mutant with respect to the wt Pc which can be correlated to differences in the dynamics of the ET process mediated by the two proteins. The introduction of the SS bridge also influences the essential dynamics of the mutated protein, where residues in the northern site, both belonging to the hydrophobic patch and to the downstream Cu-ligand loop (fig. 1), move in a concerted fashion with the large negative patch (fig. 4), while the region of the SS bridge is not involved. The differences in the functional concerted fluctuations observed for the SS mutant and the wt Pc might be related to differences in partner recognition and in the dynamic process of binding.

The overall picture emerging from the work by Arcangeli et al. [23, 24] is that internal protein motions operating in

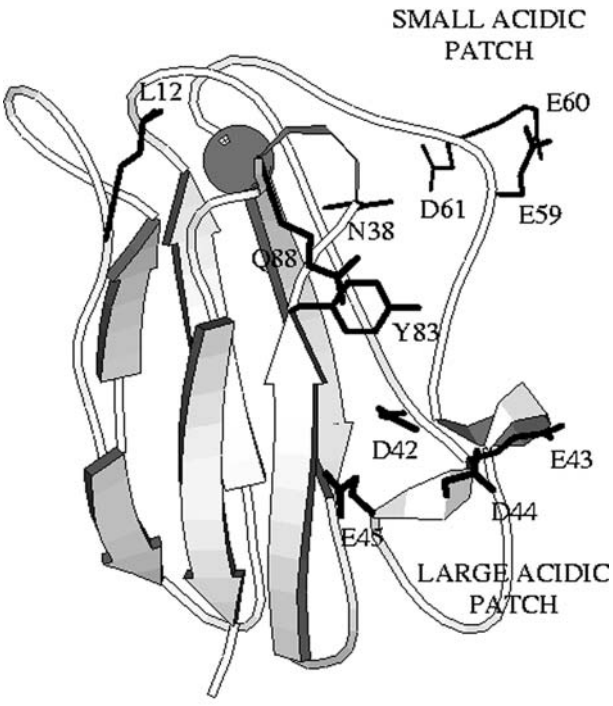

Figure 4. Mapping on the 3D structure of spinach Pc of the relevant residues L12, N38 and Q88 in the northern hydrophobic site and of the two acidic regions (residues 59–61 and 42–45) in the eastern site. Y83, which lies in the middle of the eastern site is also shown. The hydrogen atoms are not shown. The picture was drawn using MOLSCRIPT [8].

a concerted fashion and on a large scale can be fundamental for protein functionality by indirectly influencing the structure and conformation of restricted protein regions which are directly involved in BC protein function.

Protein structure and solvent

The protein environment, which is often aqueous, affects the structure and folding dynamics, and, therefore, the functionality of globular proteins. In fact, solvent-protein interactions, together with the interactions between residues in the protein matrix, facilitate the folding process and establishment of intermolecular interactions with other complex systems. Furthermore, to be properly folded and fully functional, a protein requires a minimum level of hydration. Of course, the contrary is also true, i.e. as far as water affects protein structure, protein can modify the structural and dynamic behavior of water.

The mutual influence of water molecules and BC proteins has been addressed with MD simulations and other computational studies in recent papers with the aim of improving understanding of solvent-protein interaction mechanisms in general as well as understanding their influence on the redox properties of BC proteins [25–27].

Bizzari and Cannistraro [25] recently reviewed several different computational methods and simulation techniques to investigate the local and global organization and dynamic properties of solvent water around BC and other globular proteins, pointing out the role of the solvent in controlling protein functionality and comparing experimental information. They discuss the meaning and usefulness of theoretical quantities introduced to characterize water surface characteristics, on the one hand, and to address problems such as protein folding, molecular recognition and structural transitions, on the other. For example, the solvent accessible surface area can be used to quantify the surface area of the protein in contact with the solvent. The analysis of the solvent accessible surface area for a large number of proteins in different solvent media suggested that exposure to the solvent depends both on the size of the protein and on the hydrophilic or hydrophobic nature of the protein side chains. The water residence time, i.e. the period during which a single water molecule can be found within a given distance (generally the first coordination shell) of a protein surface atom, can help to elucidate the structural and dynamic behavior of water molecules at the protein surface.

Luise et al. [26] presented a detailed MD study of how the residence time of water molecules on the surface of the BC protein Az depends on the solvent accessibility of the protein surface and on the specific solvent-exposed atom types. The simulation was performed at a high hydration level (5.0 g H_2O/g protein). They found that the main factor affecting water residence time is the local solvent accessible surface area, i.e. the water residence time is shorter in regions with large exposed surface and longer in crevices or grooves. In addition, the intrinsic chemical properties of specific protein sites seem to influence the water residence time only in regions with reduced accessibility, where water molecules make favorable interactions with polar or charged groups and unfavorable interactions with apolar groups. In other words, protein conformation and strength of the interaction between the water molecules and the protein influence the rate of exchange between bulk water and protein bound water. These findings by Luise et al. [26] are in good agreement with the results of a different analysis performed by Kuhn et al. [28] on a large series of X-ray protein structures, where the water residence time is related to protein topography.

Similar results were obtained by analysis of solvent diffusive properties at the protein surface [25]. Interestingly, while water at a large distance from the protein surface behaves like bulk water, the mobility of water at the surface is variable, going from trapped to extremely mobile molecules, as a probable consequence of local effects due to the interaction of the intrinsic electric dipole of water molecules with the protein surface atoms.

Another relevant descriptor of water-protein system behavior [25] is the structural and dynamical reorganization of the H-bond network that is established, at sufficient hydration level, between the solvent molecules and the surface protein sites and that can influence the network of intramolecular H -bonds, thus affecting the functionality of the protein. An example of the role played by hydration water was reported by Buning and Comba [27], who investigated the acidic structural transition of two cupredoxins, Ami and Pc, performing MD studies. The protonation state of these proteins is pH dependent due to the presence of a partially exposed histidine in the coordination sphere of the Cu atom (downstream histidine in figs 1 and 2). This His, which becomes protonated and flips away from the copper at low pH value, is considered to be responsible for the switching off/on of the ET activity through its protonation/deprotonation mechanism. The simulations were performed with the AMBER force field [19, 20] and a charge distribution for the Cu site obtained by previous GAUSSIAN-B3LYP calculations [16]. The solvent was implicitly defined through a distance-dependent dielectric constant ε, and a large number of conformers of the Ami and Pc MD structures were obtained by varying the torsional angle of the C-terminal His and clustered in families depending on the torsional values of their C-terminal loop, where the Cu ligand His is found. The results showed three different conformers for Pc corresponding to (i) the unprotonated Cu-ligating His form at pH 7, (ii) the protonated unbound form at pH 3.8 and (iii) an intermediate between the other two. For Ami, very similar His conformers to Pc were observed. On the basis of these calculations and previous experimental informa-

tion, Buning and Comba [27] proposed a model mechanism in which the stabilization of the low pH form is preceded by the decomplexation of His, the isomerization of the protein to a solvent accessible His form and protonation of the flipped His. In this picture, where the role of the solvent molecules is fundamental for protonating the flipped His and stabilizing the unbound form, the extent of solvent accessibility to the C-terminal His determines the possibility of the protein to attain a stable protonated state and, therefore, to lose its specific functionality.

Cupredoxin function

Protein association

Correct and timely interaction of specific molecules is crucial to the proper functioning of cells and fundamental to ET processes. Specific association depends on the intrinsic and specific structural properties of the protein systems involved, such as their overall shape, local geometry and surface charge distribution. Therefore, protein-protein interfaces should be geometrically and chemically complementary. However, for a complex to be formed, it is first necessary that the two systems come into contact. This may occur by active transport or through diffusion. During diffusion, the proteins undergo random Brownian motion and are primarily subject to long-range electrostatic forces, which may guide the molecules to orient in a way favorable to association. If the Brownian motions alone were taken into account, the diffusional collision rate of two proteins would be on the order of $10^9 - 10^{10}$ M^{-1} s^{-1} [29]. Orientational restraints, requiring the correct alignment of proteins for binding, decrease this rate to 10^6 M^{-1} s^{-1} [30], and electrostatic interactions can increase this rate up to $10^9 - 10^{10}$ M^{-1} s^{-1} [31, 32]. The environment, through pH, ionic strength, viscosity etc., can modulate the association process [32, 33]. At the end of the diffusional association, the two proteins form a so-called encounter complex, and from this point the proteins can rearrange into a bound complex. ET complexes are transient complexes that commonly have short lifetimes favoring a high throughput of electrons. As a consequence, obtaining experimental information on the conformation of the docked complex is a formidable task, and simulation techniques have become essential to address this issue.

For BC proteins, the most extensive studies concern Pc. This protein shuttles electrons from cytochrome b_6/f (Cyt b6/f) to photosystem I (PSI). Both these proteins are large (200–300 kDa) and membrane-embedded complexes. In contrast, Pc is a small (about 10 kDa) and soluble protein able to diffuse in the lumen and exchange electrons with its two membrane-constrained partners. Most modeling studies have been done on Pc and its physiological electron donor partner, Cyt f. This is a consequence of the availability of a large number of crystal and solution structures of Pc from different species [34], and of the availability of several Cyt f structures [35]. The crystal structure of PSI was only published recently [36, 37], and therefore it has not yet been the subject of extensive studies. Recently, the structure of two Pc/Cyt f complexes have become available: one structure is from higher plant (spinach Pc/turnip Cyt f) [38] and the other one is from a cyanobacterial source (*P. laminosum*) [39].

Pc has an eight-stranded flattened Greek-key β-barrel fold (fig. 1), containing a type-I copper atom coordinated by two histidines, a cysteine and a methionine (fig. 2) [12, 34, 40–41]. The surface structural features of Pc depend on the biological source they come from [42]. A common feature is the presence of a hydrophobic area, the 'northern' site, around the solvent exposed Cu ligand H87 (spinach Pc numbering). Higher-plant Pcs are strongly acidic proteins, with two ridges of negatively charged residues located at positions 59–61 and 42–45, respectively (fig. 4). These patches define the so-called 'eastern' or 'negative' site, where Y83 (spinach Pc numbering), one of the putative entry/exit points for electrons, is also situated. A peculiar case is the Pc from the fern *D. ryopteris crassirhizoma*, where the acidic patch almost reaches the northern site. Pcs from algae are still acidic, though the two acidic patches at the 'eastern' site are smaller and topologically different than in higher-plant proteins. Finally, in bacterial Pc, the acidic site is nearly completely suppressed. A representation of the electrostatic features of the different Pc species is provided in figure 5. The northern and eastern site regions are believed to be involved in either the association or the ET processes.

Cyt f [35, 43] is a c-type heme protein constituent of the Cyt b6/f complex which is anchored to the membrane by a single C-terminal transmembrane helix. The N-terminal part of Cyt f (fig. 6), which protrudes from the membrane into the lumen, has two β-sheet domains arranged to form an elongated structure; the heme is bound to the larger domain with the N-terminal α-amino group as a ligand to the heme iron. The structure from the eukaryotic species from turnip has been known since 1994 [43]. This protein, although slightly negative as a whole, has a ridge of basic residues that could recognize the negative patch of Pc (fig. 6). The commonly accepted hypothesis of electrostatic steering is therefore supported for eukaryotic Cyt f and Pc by significant local electrostatic surface complementarity (fig. 6), at least for the proteins in vitro [1, 3, 44–45]. The more recent structure of bacterial Cyt f [35, 46], on the contrary, does not show any basic ridges in the heme region and, as a whole, is far more negative than the eukaryotic species (net charge of cyanobacterial Cyt f: –12e vs. eukaryotic Cyt f: –1e), complementarily to its redox partner (fig. 6). Recent reviews [1–3] report on computational attempts to define the structure of the

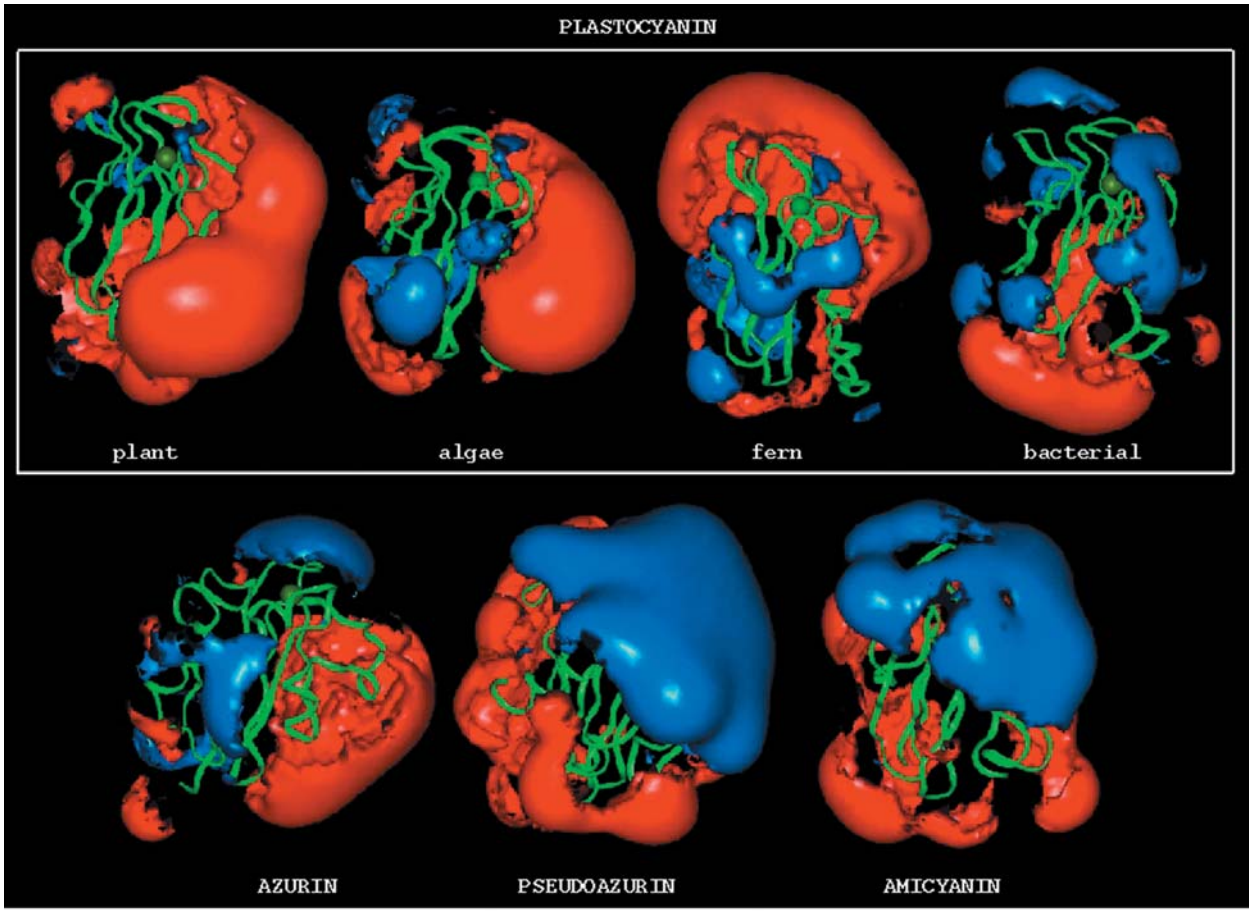

Figure 5. Electrostatic potentials (MEPs) of the most representative cupredoxin subfamilies, computed at 100 mM ionic strength. The isopotential contours are at –0.5 (red) and +0.5 kcal/mol/e (blue).

eukaryotic Pc/Cyt f complex. Here we outline the most relevant ones.

Pearson et al. [47] and Soriano et al. [48] used a combination of geometric and electrostatic criteria to propose a few possible eukaryotic Pc/Cyt f docked structures, in which the distance between the Cu atom in Pc and the Fe atom in Cyt f is short enough to allow the transfer of electrons. Pearson et al. [49] also carried out an electrostatically steered Brownian dynamics (BD) simulation of the poplar Pc/Cyt f complex with a continuum solvent model. Ullmann et al. [50] used a combination of electrostatically based Monte Carlo and MD techniques to define an ensemble of possible low-energy Pc/Cyt f complexes. However, although most of the computational papers report electrostatically driven docking of Pc and Cyt f from higher plants, experimental studies [51–52] support the idea that the initial electrostatic complex is not ideal for electron transfer, which would instead be effective only after complex rearrangement.

In this respect, a particularly interesting study was performed by Ubbink et al. [38] and Crowley et al. [39], who employed a combined nuclear magnetic resonance (NMR)/MD technique to determine, respectively, the structure of the eukaryotic [38] and bacterial [39] complexes formed by Pc and the soluble part of Cyt f. They collected an extensive set of experimental NMR chemical shift data during Pc/Cyt f complex formation that were then used, together with other experimental information [39, 53], to perform a restrained MD calculation. The eukaryotic complex structures in the NMR ensemble (pdb code: 2pcf) are similar to one of the structures identified previously by Ullmann et al. [50], although they have a shorter average Cu-Fe distance ($d_{Cu\text{-}Fe} \approx 10.9$ vs. 14 Å), with Cu site region of Pc facing the heme region of Cyt f. The eastern site is also part of the complex interface and interacts with a region rich in basic residues on the surface of the Cyt f small domain. The results of the study by Ubbink et al. [38] suggested a two-step mechanism for complex assembly: (i) formation of a relatively unspecific, dynamic encounter complex followed by (ii) reorganization into a specific complex, suitable for electron transfer, under the influence of short-range interactions. The accepted scheme of the overall reaction is that reported in equations 1–4, and in the case of Pc and Cyt f can be written as follows:

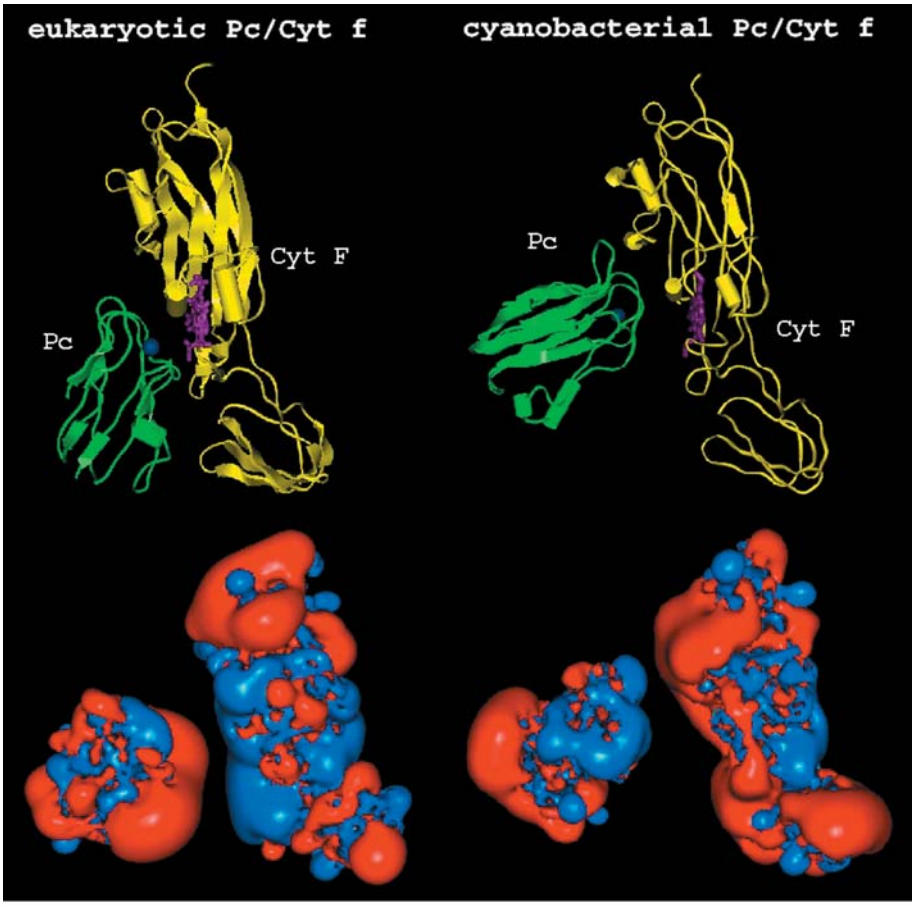

Figure 6. (*Top)* Representation of the structures of the eukaryotic Pc/Cyt f complex (pdb code: 2pcf) [38] *(left)* and the cyanobacterial Pc/Cyt f complex from *P. laminosum* [39] *(right)*. The secondary structure elements are highlighted (see fig. 1). The Cu ion and the heme are represented in CPK and sticks, respectively. Color code: green, Pc structure; blue, copper atom; yellow, Cyt f structure; pink, heme; (*bottom)* MEPs (at ionic strength of 100 mM) of the Pc and Cyt f structures from the eukaryotic protein complex (2pcf) *(left)* and of the cyanobacterial Pc and Cyt f species from *P. laminosum* (pdb code: 1baw and 1ci3, respectively) *(right)*. Isopotential contours as in figure 5.

$$Pc_{ox} + Cytf_{red} \underset{k_{off}}{\overset{k_{on}}{\longleftrightarrow}} (Pc_{ox}/Cytf_{red})' \underset{k_{der}}{\overset{k_{rearr}}{\longleftrightarrow}} \quad (5)$$

$$(Pc_{ox}/Cytf_{red})'' \underset{k_{rET}}{\overset{k_{ET}}{\longleftrightarrow}} (Pc_{red}/Cytf_{ox}) \underset{k'_{on}}{\overset{k'_{off}}{\longleftrightarrow}} Pc_{red} + Cytf_{ox}$$

where the rate constants have the same meanings as for equations 1–4; and the ratio of k_{on} to k_{off} corresponds to the association constant K_A.

The cyanobacterial Pc/Cyt f complex structures in the NMR ensemble [39] are quite different from the angiosperm structures, in that Pc interacts with Cyt f in a 'head on' manner, with its main axis almost perpendicular to the plane of the Cyt f surface (fig. 6). The northern hydrophobic site is primarily involved in complex formation, while the eastern site (which has no acidic features) is not part of the interface. The relative involvement of electrostatic interactions in bacterial complex assembly is further supported by the negligible dependence of complex structure and association on the increasing ionic strength and variation of pH values [35, 54–55].

The publication of these works [38–39] spurred new experimental and computational studies [49, 56–59], since the structure of the Pc/Cyt f complex can be used as a starting point for proposing new experimental assays and further computation and statistical analyses of the available experimental data.

From this perspective, De Rienzo et al. [59] exploited the known structural, mutation [38] and kinetic data [53] on the Pc/Cyt f ET reaction to gain new insights into the role of molecular recognition in the overall ET process. BD calculations were performed to simulate the association of Pc and Cyt f. This allowed the evaluation of the effects of mutations on the Pc/Cyt f association rate, and the comparison of the results obtained with the experimental second-order kinetic constants (k_2) for the overall ET process measured by Kannt et al. [53] for the in vitro ET reaction between wt and mutant Pcs from spinach and the soluble part of Cyt f from turnip:

$$Pc_{ox} + Cytf_{red} \underset{k_{-2}}{\overset{k_2}{\longleftrightarrow}} Pc_{red} + Cytf_{ox} \quad (6)$$

The k_2 values do not provide direct information about the ET or association mechanisms; however, they are useful in comparative studies for estimating the effects of mutations in Pc and/or Cyt f on the overall ET rate [1]. The mutations reported by Kannt et al. [53] are seven single and multiple point mutations (D42N, E43N, E43K, E43Q/D44N, E59K/E60Q, E59K/E60Q/E43N and Q88E) involving residues located in the eastern site (fig. 4). The analysis of the experimental results for the interaction of these mutants with the wt form of Cyt f suggested that at least in this set of mutants, the overall reaction is modulated by the association step, while the ET step itself appears to be almost unaffected by mutations [45, 53].

On this basis, it was possible to assume that association is the limiting step of the reaction in equation 6 and, therefore, to compare the experimental k_2 values with the bimolecular diffusional association rate constants k_{on} (hereafter referred to as k_c), computed by simulating protein-protein encounters by electrostatically steered BD [60] and to examine their dependence on protein mutations and on the nature of the physical environment [33, 60–61], with particular attention to ionic strength. Although good agreement was found between calculated and experimental values, the k_c values always overestimated the experimental k_2 values. Overestimation might be due to the intrinsic limits of the BD simulations, in which protein conformational flexibility and hydrodynamic interactions, as well as phenomena such as postdiffusional association rearrangement or conformational gating effects (which might affect real system dynamics) are not modeled [32, 60].

Besides the influence of ionic strength, another crucial aspect of protein association is dependence on pH. This effect is fundamental when protein association involves proton release or uptake. Crnogorac et al. [62] developed a method (the Modified Proton Linkage Model) for studying the effects of pH on noncovalent association of proteins. Differently from the traditional Proton-Linkage Model [63], which allows determination of the number of protons exchanged during association, the Modified Proton Linkage Model is also able to determine the pKa values of those residues at the complex interface which are responsible for the pH dependence of the association constant. The method is based on the definition and identification of groups of isoacidic residues, i.e. residues that, due to their similar pKas, are treated together as a group with a composite pKa, thus reducing the number of variables taken into account. To test their model, the authors chose a pair of well-characterized globular proteins, spinach Pc and horse-heart cytochrome c (Cyt c), that can bind together through their complementary charged patches: the acidic area of Pc abuts the basic area of Cyt c. The pH dependence observed for the formation of the wt complex changed as a consequence of mutations of one or both wt protein sequences and of the specific mutation introduced. To study this trend, Crnogorac and collaborators [62] compared the variation of association constants for the wt complex and three other complexes formed by wt Cyt c and different Pc mutants (D42N, E59Q, E60Q) within a pH interval ($5.4 < pH < 9.0$) compatible with the stability and optimal functioning of the proteins. The results of the calculations were in good agreement with the experimental data and provided some interesting hints for elucidating microscopic details that cannot be obtained experimentally. In particular, they revealed that upon association, the ionic forms of both the basic and acidic residues are stabilized with respect to the neutral forms. This is due to a few Pc acidic residues with anomalously high pKa values (about 6.3) and a few Cyt c basic residues with anomalously low pKa values (about 7.0), which change their pKas during complex formation and become, respectively, deprotonated and protonated. The authors were also able to identify two acidic residues (namely E59 and E61, see fig. 4) of Pcs with a high probability of releasing a proton upon binding. This work shows the feasibility of identifying unusual pKa values for potentially biologically relevant ridges of residues.

Very recently, Gross and Pearson [64] carried out a BD study (using the Macrodox program [65]) of an algal (*Chlamydomonas*) Pc/Cyt f complex formation. Algal Pc and Cyt f largely resemble their respective relatives from higher plants, as far as 3D structures and the interaction properties are concerned. Consistently, they found for the *Chlamydomonas* Pc/Cyt f complex a structure that is highly similar to those previously predicted for eukaryotic complexes by computer simulation [47–50, 59] and coupled NMR/MD simulation techniques [38]. In particular, both the eukaryotic and the algal complexes show (i) important hydrophobic interactions at the complex interface between nonpolar residues in the heme region of Cyt f and in the Cu region of Pc; and (ii) electrostatic interactions between residues forming a positive patch and a negative patch on the Cyt f and Pc surfaces, respectively. In addition, all the complexes obtained share the same essential conformation and orientation, whose randomness is increased by increasing the ionic strength and/or introducing mutation at key interface residues. All these findings support the fundamental role played by the electrostatics in complex formation.

Gross and Pearson [64] also compared the algal Pc/Cyt f structure with the results obtained by an analogous study performed on the complex formed by the Cyt f and cytochrome c6 (Cyt c6) from *Chlamydomonas*. Interestingly, they found that the Cyt c6/Cyt f complex closely resembles the Pc/Cyt f complex, in hydrophobic and electrostatic interactions. Cyt c6, which is expressed by the organism in particular growth conditions, can substitute Pc in its interaction with Cyt f. In fact, these two proteins, although very different in sequence and structure, share the same interaction properties.

Similarly to De Rienzo et al. [59], the global ET rate constant estimated by Gross and Pearson [64] are overestimated with respect to the experimental values, probably as a consequence of the intrinsic limitations of the BD program used or of the choices and assumptions made on the simulated system.

Recently, the high-resolution crystal structure of the whole PSI complex from the cyanobacterium *Prochlorothrix elongatus* [36, 37] has become available. PSI (which is the electron acceptor of Pc) is a multisubunit transmembrane protein complex that takes part in the reduction of NAD to NADH. Due to the extreme complexity of the PSI structure, only the structures of some subunits were known previously [66]. Taking advantage of the new high-resolution structure of PSI by Jordan et al. [37], Myshkin et al. [67] docked Pc from *Prochlorothrix hollandica* to the PSI lumenal subunits PsaA and PsaB, which are considered to be involved in the binding of Pc from crystallographic and mutagenesis studies. While the structure of *P. hollandica* Pc was available from NMR studies [68], the *P. hollandica* PsaA and PsaB were modeled by homology to the known PSI structure of *P. elongatus*. The Pc structure was then docked to the lumenal loops using different rigid-body algorithms, based on geometric criteria [69, 70], energetic criteria [71] or both [72], for comparison purposes.

The best docking conformation was defined as the one which best satisfied a predefined set of geometric and distance criteria, imposed on the basis of the available experimental data on the complex interface and the ET process.

The free energy of binding (ΔG_{bind}) was evaluated for the complex structure selected by means of the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) approach developed by Kollmann et al. [73, 74]. This free-energy study was performed on the wt complex and the complexes formed with two Pc mutants relevant to the ET process. The computational results obtained by Myshkin et al. [67] were in reasonable agreement with the experimental information [75], thus supporting further applications of this approach to other bacterial Pc/PSI complexes.

Electron transfer

The rate constant for electron transfer k_{ET} is given by the Fermi golden rule as [3]:

$$K_{ET} = (2\pi/\hbar)|H_{AD}|^2 FC \quad (7)$$

where $\hbar$ is $h/2\pi$, H_{AD} is the electronic coupling matrix element between a donor (D) and an acceptor (A) site (i.e. the superposition of the wavefunctions of D and A), and FC is the Frank-Condon factor, which is mainly related to the structural and vibrational properties of the protein complex system. According to Marcus interpretation [3]:

$$FC = [1/(4\pi\lambda RT)^{1/2}] \exp[-(\Delta G^o + \lambda)^2/(4\lambda RT)]. \quad (8)$$

Therefore, k_{ET} depends on H_{AD}, on the free energy of the reaction ΔG^o (i.e. the reduction potential) and on the reorganization energy λ. H_{AD}, ΔG^o and λ are often evaluated by theoreticians as separate contributions.

When modeling and simulating ET in large biological systems, it is first necessary to decide the level of approximation to be used to describe the systems. ET between macromolecular systems often covers distances of 20 Å or more, and thus it is necessary to consider models with large numbers of atoms. An ab initio QM treatment, in which the electron distributions are modeled, would be the most correct and complete approach. However, treating a whole protein-protein complex with ab initio or even semiempirical QM techniques is still prohibitive from a computational point of view. Thus, two main alternatives to a full ab initio QM approach are normally envisaged: (i) using QM (ab initio or semiempirical) techniques in a restricted volume of the protein system, or treating the whole protein system using (ii) empirical energy-based methods or (iii) QM/MM approaches.

QM methods

An interesting example of the first approach can be found in the paper by Larsson [76]. Here, the postulate is that the requirements for a redox operative Cu site are a small reorganization energy barrier to the electron transfer and a connection to the outer world through suitable ET pathways.

λ is conventionally considered to be made of a contribution from the structural changes involving the first coordination sphere of the redox site (inner-sphere λ) and a contribution from the surrounding protein matrix and the solvent (outer-sphere λ). In the case of the BC proteins, the inner-sphere λ is due to the four (or five, in the case of Az) residues directly bound to the copper center (fig. 2). Total λ has been evaluated from ET experimental assays [76] to be quite small (0.4–1.2 eV), thus explaining of the fast ET process. Theoretical estimates are limited to inner-sphere λ, preventing direct quantitative comparison with experiments.

Larsson et al. [76–77] evaluated inner-sphere λ arising from differences in the bond lengths and bond angles of Cu ligand atoms observed between oxidized and reduced X-ray Cu-site structures as:

$$\lambda = {}^1/_2 \Sigma_i k_i \delta R_i^2 \quad (9)$$

where δR_i are the changes upon electron transfer in the ith bond length or bond angle, and k_i are the force constants obtained from the vibrational modes.

They obtained $\lambda = 0.5$ eV for the oxidation of the Cu(I) site of Az from *Alcaligenes denitrificans*. The largest contribution (0.34 eV) to this value is due to the change of the Cu-S_{Cys} bond length, as expected since the singly

occupied orbital of the Cu(II) site filled at reduction is a dCu-pS_{Cys} π-antibonding orbital. However, the author discuss that, λ being quadratic in δR, small variations in the geometric values (of the order of the experimental error in crystallographic data) might significantly alter the results.

Another relevant study was performed by Ryde et al. [10, 78], who estimated inner-sphere λ from their calculated model of the Pc site (see discussion above) to be 62 kJ/mol, which corresponds to about 0.64 eV. Similarly to Larsson [76, 77], they found that comparison of theoretical with experimental estimates of λ is not at all straightforward. In fact, the experimental value also contains the contribution from the outer sphere, which includes the protein matrix, the solvent and the redox partner docked to the protein. Thus, Ryde et al. [10, 78] provided an approximate value of the total λ by combining the computed inner-sphere λ with an estimate for outer-sphere contribution (42 kJ/mol = ~0.45 eV) obtained by force-field calculations on the complex formed by Pc and its electron donor Cyt f docked in a few different conformations. The resulting total λ was about 100 kJ/mol (≈1.04 eV), which is comparable to the experimental value obtained for Pc (102 kJ/mol, i.e. ~1.06 eV) [79].

On the basis of their calculations on a model Cu-site structure, the authors then discusses how a low reorganization energy is achieved by BC proteins through an appropriate choice of Cu ligands. According to the authors, being the inner-sphere λ calculated on model systems comparable to that of the Cu site in the protein, neither the reorganization energy nor the site structure seem to be affected by the protein matrix. In open contrast with other researchers [80] who support the idea that a low λ is achieved through a constrained coordination imposed by the protein matrix on the Cu site, Ryde et al. [10, 78] affirm that the low reorganization energy observed is achieved by the site flexibility. This is ascribed to the presence in the first coordination shell of particular Cu-coordinating atoms, such as S_{Met}, which seems to be able (as demonstrated by calculations) to modulate its position with respect to the copper at very low cost.

In addition to the reorganization energy evaluation, Larsson [76] also presented an ab initio QM study of ET pathways in Pc. He modeled a limited region of the Pc structure, obtained by cutting out the protein around the Cu site and subsequent saturation of the dangling bonds with H atoms. This region contains the Cu atom and its four amino acid ligands (H87, H37, M92 and C84) (fig. 2) plus the residues connecting C84 to H87 (E85 and P86) and Y83, directly bound to C84. Y83, located in the eastern site, midway between the small and the large acidic patches (fig. 4), might be involved in ET. Larsson used a semiempirical QM method (INDO/S) [81] to evaluate the possibility that this residue is the entry/exit site for electrons. More precisely, he hypothesized a proton-transfer-assisted ET in which, after association with the partner protein, Y83 loses its hydroxyl proton (which has a relatively low pKa value of ~9), leaving an activated conjugated system that could donate an electron to the partner. An electron coming from the proper Cu site could then 'move' readily [81] to the Y83 to substitute the lost electron. The calculations showed that the Highest Occupied Molecular Orbital (HOMO) of the reduced Cu-site–Y83 system is localized isotropically on the Cu atom and its ligand residues, with a peak on the solvent-exposed H87 and vanishing electron density on Y83. Calculations on the system after removal of the Y83 hydroxyl proton showed instead that HOMO is completely localized on the activated tyrosine. Thus, the hypothesis is that excess electron charge on the deprotonated Y83 is transferred to the other protein (intermolecular ET) and replaced by another electron from the Cu (internal ET). This might be the electron pathway connecting the Cu site of Pc with the outer world, i.e. in this case, the electron acceptor of Pc, PSI.

FF methods

An alternative way to treat the ET and to evaluate the H_{AD} term in equation 7 for BC proteins was presented some years ago by Ullmann et al. [50]. They used Pathways, a semiempirical approach developed by Beratan et al. [82], to identify the different possible ways in which the electron can move from the heme of Cyt f to the Cu site of Pc, taking into account the whole structure of the protein complex.

The Pathways model allows one to evaluate the electron coupling (H_{AD}) between an electron donor site and an acceptor site that are not directly coupled and that are separated by an intervening medium, called the bridge. As shown in equation 7, the k_{ET} is proportional to $|H_{DA}|^2$ through the FC factor. Therefore, by estimating H_{DA}, the method provides an estimate of the rate of the ET process. The quantum-mechanical effects are not explicitly accounted for, and the method is based on a parameterization of electron tunneling through different bond types. The concept of electron tunneling path is introduced and is meant as a combination of covalent, H-bond and van der Waals interactions connecting the donor site to the acceptor site. Each bond type has an associated factor describing the probability that an electron is transferred through that bond (the attenuation of the electronic coupling is lower for ET through covalent bonds and it increases progressively for H-bonds and van der Waals contacts). Many different electron pathways connecting the same donor and acceptor sites can be found, whose electronic couplings differ depending on the nature of the bonds in each path. The method may be deficient in describing situations in which not only the nature of the bridging medium but also the solvent or other local factors largely affect the ET process. Due to the approxima-

tions of the method, only trends in H_{DA} values should be considered, since absolute values have little meaning.

Ullmann et al. [50] applied the Pathways method to six different configurations of the Pc/Cyt f complex that they obtained after docking and MD simulations (see paragraph 2a). The configuration with the strongest electronic coupling was the one with the shortest Cu-Fe distance, not the most energetically stable one. The most energetically stable configuration had the second-best electronic coupling. The best path showed the electron moving from a heme propionate to P86 in the northern site of Pc, through van der Waals interactions, and then directly to the Cu through H87 or the path S85-C84, which are tightly coupled to the Cu (fig. 7). In all cases, the path C84-S85-P86-H87 in the northern site seems to be involved in electron transfer. These electronic pathways can also be envisaged in the NMR/MD ensemble structures determined by Ubbink et al. [38], where, however, an additional putative electron path is also suggested, which connect directly the Fe-ligand Y1 directly to the Cu-ligand H87 (see fig. 7).

Ullmann et al. [50], similarly to Larsson [76], also speculated on the possibility of a different electron pathway, which could involve interaction of Y83 with a charged side chain of Cyt f. Here, Y83 is treated as an electron acceptor, in contrast to Larsson's hypothesis. In the energetically most stable structure of Ullmann et al. [50] a hydrogen bond or a cation-π interaction between Y83 and Cyt f K65 could be established. The cation-π couple (formed between the positive ion of the K65 and the aromatic ring of Y83) could act as an electron acceptor and form a radical intermediate during the electron transfer from Fe to Cu. Even though the formation of this intermediate has never been detected experimentally, it could theoretically form in vivo during the interaction of Pc with the physiological partner. In this case, the transfer of the electron from Fe to Cu would not be direct but mediated by Y83. To analyze this process, Ullmann et al. [50] performed Pathways calculations by splitting the problem in two parts, from Fe to Y83 and from Y83 to Cu, and considering all the conformations recorded during the MD simulation that led to the most stable optimized conformation. The analysis highlighted a pathway, at the beginning of the simulation, which was mainly composed of hydrogen bonds connecting Fe and Cu through K65 of Cyt f and Y83 of Pc (fig. 7). This path was not persistent in the simulation due to disruption of the H-bond network involving Cyt f K65, as a consequence of its diversion towards the Pc acidic patch. In this situation, the ET pathway runs through a different route via backbone atoms connecting K65 of Cyt f and Y83 of Pc. This second path was characterized by a much lower electronic coupling than the first one, suggesting that the structures of Pc and Cyt f need to be in a precise orientation and conformation for the complex to be electron transfer active. The presence of these electron pathways connecting Pc Y83 and Cyt f K65 cannot be observed in the NMR/MD structures by Ubbink et al. [38], since the K65 side chain points towards the small acidic patch residues S58 and E59, far away from Y83.

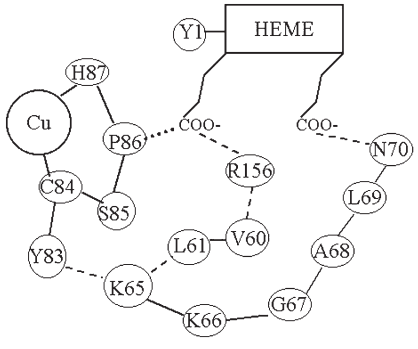

Figure 7. Scheme of the possible ET pathways from the heme in Cyt f to the Cu in Pc proposed by Ullmann et al. [50]. The residues at the Pc/cyt f complex interfaces, the Cu atom and the heme (with its propionates) are represented. Symbols: straight line (—), covalent bonds; broken line (- - -), hydrogen bonds; dotted line (•••), van der Waals interactions.

The Pathways analysis [50] provided a viable explanation for the results of cross-linking experiments [83–84], which demonstrate that ET from Fe to Cu can be precluded by covalently bonding an acidic residue in the eastern site of Pc to a lysine of Cyt f. Ullmann et al. [50] hypothesized that formation of a relatively rigid (or covalent) complex structure might prevent the complex rearrangement that is necessary for ET, and/or disrupt the interaction between Pc Y83 and Cyt f K65, thus completely blocking the ET reaction.

Besides λ and H_{AD}, the third factor to be considered is the variation of the free energy ΔG^o due to the redox reaction, i.e. the redox potential. In De Rienzo et al. [85], the variation of the experimental redox potential E_m (midpoint redox potential) in a Pc mutant series [53] was studied by computing the redox potential differences $(\Delta E_m)_c$ between the Cu(I) and Cu(II) states of the wt Pc and its mutants. Redox potential differences were computed as

$$(\Delta E_m)_c = (E_m Pc)_c - (E_m wt)_c \quad (10)$$

where $(E_m Pc)_c$ is the computed redox potential value of each mutant and $(E_m wt)_c$ is that of the wild type. The values $(E_m wt)_c$ and $(E_m Pc)_c$ were computed considering the electrostatic free energy difference between the reduced and the oxidized form of the each protein [85].

Briefly, this procedure computes the pKa values of each titratable group in the protein, taking into account its environment and allowing the pK_as of titrating residues to be evaluated from the average protonation states of each site at a given pH. The electrostatic free energy is then calculated by integrating the titration curve.

The calculations were performed for both the reduced and oxidized states of the protein, to evaluate the differ-

ence in the electrostatic free energy of the two states, which, at a given pH, is related to the redox potential by the Nernst equation ($\Delta G^o = - nFE$). If there are no relevant conformational changes of the protein within the range of pH considered for the titration curve, the values of the redox potential obtained at a given pH can be compared with those from experiments carried out at the same pH. For Pc, this is true for pH > 5 [34].

Considering that the available experimental values of the midpoint redox potential are accurate to ±3 mV, the calculated values are in reasonable agreement with the experimental data, though generally underestimated. The underestimation might be a consequence of the lack of conformational flexibility (which could affect the entropy of the system) in the computational procedure.

QM/MM methods

Very recently, Olsson et al. [86] published a paper where they report on the calculations of reduction potentials of Pc and rusticyanin (Rus). These two BC proteins have very different redox potentials (375 vs. 680 mV for Pc vs. Rus, meaning a redox potential difference of 305 mV), while having nearly identical Cu sites. The authors investigated the possible reasons for this difference and compared the results obtained with different simulation techniques, including all-atom classical methods, with or without a polarizable force field, the semiempirical macroscopic protein dipole Langevin dipole method and QM/MM approach.

In the semimacroscopic calculations, a relevant number of configurations for both redox states were produced for averaging the electrostatic energy, which was then used as input in the Linear Response Approximation formulation [87]. Interestingly, this approach gave rise to the best estimates of the reduction potential differences between Pc and Rus: ΔE^o is in fact about 310 mV, in the range of the experimental value (305 mV). Instead, with the long (about 1 ns) all-atoms classical simulations, the reduction potential difference obtained was much larger (500–700 mV). Even repeating the simulations for a large number of protein configurations did not lead to results comparable to the experimental data. This problem was partially overcome using a polarizable force field: the Pc-Rus redox potential difference fell then in the range 269–313 mV, thus being in good agreement with experiments.

The implementation of a QM/MM simulations arose from the need to describe the system to a more accurate level, to represent correctly its physical behavior and to improve the predictive power of simulation and energy calculations. Here, the system was split into three regions: (i) the core, constituted by the copper, its ligand residues and the closest water molecules, which was treated by an ab initio approach using regular full DFT; (ii) the outer region, made of the rest of the protein matrix and the solvent, which was treated with a classical force field; (iii) the region at the interface between the QM and the MM parts, generally constituted by the residues and the water molecules in the second coordination shell of the metal, which was represented by frozen electron density with a frozen density functional (FDFT). This method is faster than classical QM/MM approaches where the interface is represented with point charges, thus it allows one to include in the QM part a larger portion of the whole system and to run the calculations over different configurations in order to obtain an average value of the computed property. The FDFT QM/MM approach, depending on system modeling and implementation, produced an energy difference between the two BC proteins of 180–340 mV. Thus, at the present stage of development, a lower description level is still able to produce better estimates of the reduction potential than an accurate quantum-mechanical treatment of the system.

The authors also studied a model Cu site (with no protein matrix around) in an aqueous environment, to obtain further insight into the origin of the E^o difference between Pc and Rus. They found that the redox potential for this model system ranges from 400 to 600 mV, lying between that of Pc and that of Rus. They deduced that the primary effect of the protein matrix is to tune down the redox potential for Pc and to tune it up for Rus, as a consequence of different orientation and organization of the microscopic protein and water dipoles in the two protein systems.

Olsson et al. [86] pointed out in their work two essential aspects of redox potential calculations: (i) the importance of considering explicitly the influence of the active site environment (both protein and solvent) which contributes to the energy changes related to protein function, and (ii) the need to perform a proper configuration sampling and to consider average property values, as a consequence of the high flexibility and dynamics of the protein systems.

Structure-function relationships

Understanding the structure-function relationship in proteins is fundamental, since this, with the aid of protein engineering and molecular biology techniques, may allow the design of new proteins with desired properties and functionalities.

We have already discussed how theoretical quantities derived from 3D protein structures in water can be useful to describe the different behavior of proteins under different conditions [25, 26]. However, from a different point of view, it is interesting to ask what molecular determinants are intrinsic to protein structure and sequence and whether and how they can be related to specific functions. The idea is to infer from the protein structures indices which contain information about a particular structural property, and to compare variations of these indices

to variations observed experimentally in the behavior of the proteins.

Semi-quantitative structure-function relationships of cupredoxins

Despite the large amount of data now available from mutagenesis studies, kinetic analysis and biochemical assays, a detailed functional characterization of the proteins belonging to the different cupredoxin subfamilies is difficult to achieve experimentally, due to the complexity of the redox cascades in which they are involved. As previously mentioned, all the cupredoxins have similar barrel-like architectures and ET properties [34, 41]; however, their sequence similarity is low, and they can recognize and bind different partner proteins. The redox partners are unknown in the case of many cupredoxins, and to complicate the picture further, the interaction of a cupredoxin with its partner does not seem to be exclusive: in fact, it is possible that more than one donor or acceptor is suitable for interaction with the same cupredoxin and, on the other hand, that the cupredoxin itself might be substituted with other isofunctional proteins [50, 88, 89].

With the aim of gaining a comprehensive, objective and insightful overview of the interaction properties of the cupredoxin superfamily, De Rienzo et al. [90] recently computed indices which describe the interaction properties of the available 3D structures of cupredoxins using the PIPSA (Protein Interaction Property Similarity Analysis) procedure [33, 91]. These indices were used to perform a semi-quantitative comparative analysis of the proteins to identify the differences in the structural features of cupredoxins that can affect their molecular recognition properties. Molecular interaction fields (MIFs), in particular the molecular electrostatic potential (MEP), which is effective at medium and long range, and the hydrophobic interaction field were computed for all the cupredoxins. To obtain descriptors of MIF variability, the electrostatic and hydrophobic properties were then compared quantitatively by means of the pairwise similarity index (SI) [92], which translate the property differences observed between any two proteins into a single number.

The analyses resulted in an extensive comparison of interaction properties for a set of BC proteins with known 3D structure representing the different cupredoxin subfamilies (plastocyanins, azurins, pseudoazurins, amicyanins, rusticyanin, basic copper protein, stellacyanins), providing a classification of the BC proteins with respect to their MIFs.

To verify whether the interaction properties depend strictly on the protein structure, pairwise comparative analyses of the sequence and the structural elements of the proteins were also performed.

The results showed that the hydrophobic interaction properties are strictly related to the protein sequence and local structure. This is probably a consequence of the fact that the hydrophobic interactions are effective at short-range distances from the protein structure and therefore they depend largely on the local structure and sequence variations. Thus, proteins sharing a similar sequence will share similar hydrophobic properties. On the other hand, similarities in the electrostatic properties do not simply mirror the similarities of protein sequences and structures. This was a crucial result, meaning that for understanding the function of a protein, it is not enough to consider simply its local structure or sequence. In fact, the detailed analysis of the electrostatic potential of cupredoxins provided different and complementary information with respect to those related to sequence and structure. For example, it showed that some proteins belonging to different subfamilies (and having very low sequence similarities), such as amicyanins and pseudoazurins, can have similar electrostatic potentials (fig. 5). This result explained the experimental observation that in certain environmental conditions, these two proteins might work as isofunctional proteins. Similarly, some Azs and pseudoazurins, although having different overall binding profiles (fig. 5), showed sufficient local similarity to bind to redox partners belonging to the same family (Cu-NIRs) [90, 93]. On the other hand, it was possible to observe that proteins belonging to the same subfamily, such as eukaryotic, algal, fern and cyanobacterial Pcs, might show different interaction properties (fig. 5), consistent with the differences observed in their binding specificities to their redox partners, as previously discussed.

Besides the different specific applications, what is particularly interesting in this method is the fact that it is rapid, automated and reliable, which renders it particularly suited for large-scale analyses and comparisons of large numbers of experimentally determined and modeled protein structures. The limitations are in the qualitative or semi quantitative nature of the results which, nevertheless, can be used to design experimental assays to obtain further insights into the physiological functions of the proteins.

Quantitative structure-function relationships of cupredoxins

In the field of pharmaceutical chemistry, theoretical methods have long been developed to describe the influence of the structure of drugs (or of potentially active organic molecules) on the function they play in living organisms. Methods to derive quantitative structure-activity (QSAR) and structure-property relationships (QSPR) have been widely used in the literature. They aim is to derive theoretical indices able to decipher the code of variations of the molecular structure and to relate them, in a quantitative way, to experimentally observed variations of specific functional properties (e. g. binding affinity, se-

lectivity, efficacy, activity). Classically, these indices describe the property of a whole molecule and are based on molecular connectivity and the chemical nature of the constituent atoms. Although good predictive models have been obtained for small organic molecules, application of this same technique to large molecules is not yet straightforward due to the complexity of the biological systems, i.e. the huge number of atoms they contain, their conformational flexibility, the difficulty of obtaining experimental information about their functionality etc. However, the overwhelming growth of information about protein function provided in the last years by structural genomic and proteomics studies can help to move the QSAR/QSPR perspective from small organic molecules to macromolecules such as proteins.

De Rienzo et al. have shown [59, 85, 94] that it is possible to use descriptors, typically developed for QSARs of small molecules, to perform QSPR studies of series of protein mutants. This approach proved to be successful in elucidating the causes that determine the variations of molecular properties crucial for protein function. The studies were performed on the Pc/Cyt f system, and the biological functions analyzed were (i) the capability of the two molecular systems to recognize and bind each other, (ii) the electron transfer between them and (iii) the redox properties of Pc. The experimental data, which highlight the trends in these protein functions, are respectively:

(i) the mutation-dependent variation of the association constant K_A (see eq. 5) for the Pc/Cyt f interaction process, which describes the ability of a protein to bind its partner [85];
(ii) the mutation-dependent variation of the overall ET rate constant k_2 (see eq. 6) [59, 85]; and the redox potential E_m [86] of the Pc/Cyt f electron transfer, which relates to the efficiency of the whole process between the two molecules;
(iii) the mutation-dependent variation of the thermodynamic experimental parameters (ΔH^o, ΔS^o, ΔE^o) of the Pc Cu(II) reduction, which describes the redox properties of the Pc variants at an electrode [94].

The K_A, k_2 and E_m data in 1) and 2) were available for a set of spinach Pc mutants with single or multiple mutations in the region of the eastern site involving residues D42, E43, D44, E59, E60 and Q88, which interact with the wt form of Cyt f from turnip (fig. 4) [53]. The thermodynamic data in 3), in contrast, were available for a set of spinach Pc variants obtained by mutating three residues in the region of the Cu site, L12, Q88 and N38 (northern site mutants), which are thought to influence the redox properties of the protein (fig. 4) [94].

The protein descriptors which yield the most significant correlations are related to (i) electrostatic and dipole-dipole interactions, effective at medium- and long-range distances (encoded by indices derived from the dipole moment or the MEP [59, 94]); (ii) polar interactions (encoded by partial charged surface area indices [95]); (iii) hydrogen bond interactions (encoded by descriptors such as the hydrogen bond interaction energy or the surface area accessible to solvent and capable of establishing H-bond interactions) [96]; (iv) dispersion and repulsive interactions (encoded by descriptors which take into account the steric hindrance or the branching ratio [97]).

In addition, estimations of mutation-dependent variation of wt Pc redox potential E_m and overall reaction rate k_2 were obtained by electrostatic free-energy calculations [86] and BD simulations [59], respectively.

The QSPR analysis showed that at least for Pc eastern site mutants, the recognition and binding process is the limiting step, while the intracomplex electron transfer is little affected by mutations [53]. In fact, mutations of the negatively charged residues of the Pc eastern site into positively charged or polar residues strongly affected the electrostatic properties of the protein and especially its dipole moment, reducing the protein polarity and its ability to attract the basic patch of Cyt f and, consequently, the overall reaction rate. Alternatively, it can be said that the more mutations perturb the electrostatic distribution on the wt Pc surface, the more the reaction of the Pc mutant with Cyt f is hampered. This provided a theoretical confirmation to the thesis by Kannt et al. [53] that the overall rate constant for the Pc/Cyt f reaction is mainly affected by electrostatically steered association.

The northern site Pc mutants were used to study the relative variation of the thermodynamic parameters and the redox potential of Pc with respect to an electrode [94]. This was useful to address the issue of how and to what extent the protein matrix and the solvent selectively stabilize the two oxidation states of the Cu atom in spinach Pc and to establish relationships between the redox potential ($E^{\circ\prime}$) and protein structure/sequence features. QSPR analyses were performed to determine how structural factors and electrostatic and solvation effects influence the redox properties of Pc through the combined effects on the reduction enthalpy and entropy.

The analysis showed that the changes observed in the reduction enthalpy of Pc can be explained mainly on electrostatic bases, considering the variation of the dipole moment and its components. More precisely, a negative charge introduced in the surroundings of Cu seemed to be capable of stabilizing the Cu(II) state, enhancing the reduction enthalpy with respect to the wt protein (less negative $\Delta H^{o\prime}_{rc}$ values) and decreasing the reduction potential, while opposite considerations hold for insertion of a positive charge. On the other hand, the variation of the reduction entropy seemed to be related to the capability of the protein to donate or accept hydrogen bonds.

Counterintuitively, if an extended negative surface area, which can accept H bonds, is localized on the northern re-

gion of the protein, the Cu(I) state is favored entropically, while if the same region is positive and capable of donating H bonds, then the Cu(II) state is favored. Thus, the insertion of a mutation can modify the charge distribution around the copper site and affect the dipolar features of the wt Pc in such a way as to determine a realignment of the water dipoles at the Cu site, and with consequences on the redox state stability which are opposite to that expected on the basis of electrostatic effects.

The analysis also showed that the presence of a strong hydrogen bond network in the proximity of the Cu can affect the redox potential in the sense that the insertion of residues, which enlarge the H-bond network thus favoring the rigidity of the copper site, tends to stabilize the Cu(I) state.

All these results support the idea that the redox potential of Pc is largely affected by the electrostatic field experienced by the solvent around the protein.

Conclusions

This review summarizes recent research trends in the application of computational approaches to the study of cupredoxins, a superfamily of copper-containing proteins. The interest in these proteins is primarily due to their involvement in complex redox processes fundamental for their host organisms (such as photosynthesis). Cupredoxins are themselves experimentally well characterized, but their interactions with their ET partner proteins are only partially characterized experimentally. To obtain a complete and detailed description of these systems means not only unraveling the microscopic features of their structure and function, but also understanding how the structural modifications and the physicochemical properties of the environment influence the protein functionality.

First, the interest in understanding the influence of the environment on protein local structure is highlighted [10, 21, 23, 24–27]. Experiments and computational studies are aimed at answering questions such as: To what extent are the local properties of the metal site determined by its protein and solvent environment? Which is the influence of protein internal dynamics on structural regions relevant for the protein functionality? How is protein conformation connected to solvent dynamics? Although the answers to these questions are not unique, they all point out the importance of simulating a model system that is as complete and consistent with the real biological system as possible. As a consequence, mixed methods combining computational approaches, such as DFT or quantum-chemical and molecular mechanics methods, are being increasingly applied and preferred to traditional methods due to their versatility and ability to describe the same system at different levels of detail. Thus, the solvent (typically water) can be explicitly introduced in the simulations or described using implicit parameterization. In addition, MD and essential dynamics are extensively used to provide insight into the dynamic behavior of the protein structure in its solvent environment and to relate conformational changes and internal fluctuations to the functional role played by the proteins.

The environment not only affects the protein structure but also protein functionality. In computational studies, it is important to take into account modifications of solvent characteristics such as pH and ionic strength, and to analyze them in the light of the different functions played by the protein [32]. pH and ionic strength are particularly relevant when molecular recognition and protein-protein association are studied [59, 62]; however, other protein functions can also be affected, such as electron transfer, as in the case of cupredoxins [76].

Two fundamental processes in which the BC proteins are involved have been analyzed and studied extensively and from different perspectives: the association of cupredoxins with their partners and their electron transfer. Force-field-based computational techniques (electrostatic potential calculations, MD, docking, BD) are used in combination with experimental assays to understand how proteins bind to each other, to propose hypotheses for the structure of bimolecular protein complexes [1, 38, 67] and to estimate rate constants for association [59]. Deriving clues for understanding the details of ET processes in biology is a far more complicated matter to be approached computationally. Questions such as: What are the requirements for fast electron transfer? How does an electron exit or enter the Cu site? Which residues are directly involved in the ET process? How can we estimate the parameters involved in the ET process, such as reorganization energy, electronic coupling between two redox active sites and reduction potential? These issues are addressed with computational approaches which imply different levels of system description. Ab initio QM methods [76] are used to focus on the local properties of the region around the Cu site such as reorganization energy or change in charge distribution during ET. Force-field-based methods [50, 85] are applied to the analysis of the whole protein system in order to compute the electronic coupling and the ET rate constant between two metalloproteins, or provide estimates of reduction potential. Finally, QM/MM simulations are being implemented with the hope of enhancing the power of computer techniques to predict the redox energy, thanks to a more physical description of the systems.

Recently, a new quantitative approach (QSPR) commonly used in pharmaceutical chemistry has been introduced to rationalize the influence of protein structure on protein function and to quantify their relationship with the aid of regression models [4, 59, 85, 90, 94]. Once validated on different protein systems, the application of QSPR analyses might be exploit in protein engineering.

Acknowledgments. We thank Dr M. Ubbink who kindly provided us with the NMR ensemble coordinates of the Pc/Cyt f complex from the cyanobacterium *P. laminosum*. One of the most frequent orientations of Pc with respect to Cyt f is represented in figure 6, top left.

1 Hope A. B. (2000) Electron transfer amongst cytochrome f, plastocyanin and photosystem I: kinetics and mechanisms. Biochim. Biophys. Acta **1456:** 5–26, and references therein
2 Bendall D. S. (ed.) (1996) Protein Electron Transfer, Bios Scientific Publishers, Oxford, and references therein
3 Bendall D. S., Wagner M. J., Schlarb B. G., Söllick T. R., Ubbink M. and Howe C. J. (1999) Electron transfer between cytochrome f and plastocyanin in *Phormidium Laminosum*. In: The Phototrophic Prokaryotes, pp. 315–328, Peschek et al. (ed.), Kluwer Academic, New York, and references therein.
4 De Rienzo F., Gabdoulline R. R., Wade R. C. and Menziani M. C. (2002) Biologically relevant properties of copper-containing proteins analyzed by means of semiquantitative and quantitative theoretical descriptors. In: Handbook of Copper Toxicology and Pharmacology, pp. 53–80, Massaro Edward J. (ed.), Humana Press, Totowa, NJ, and references therein
5 Banci L. (2003) Molecular dynamics simulations of metalloproteins. Curr. Opin. Chem. Biol. **7:** 143–149, and references therein
6 Rinaldi R., Biasco A., Maruccio G., Cingolani R., Alliata D., Andolfi L. et al. (2002) Solid state molecular rectifier based on self organized metalloproteins. Adv. Mat. **14:** 1453–1457
7 Rinaldi R., Biasco A., Maruccio G., Cingolani R., Alliata D., Andolfi L. et al. (2003) Electronic rectification in protein devices. Appl. Phys. Lett. **82:** 472–474
8 Kraulis P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. **24:** 946–950
9 Ryde U., Olsson M. H. and Pierloot K. (2001) Theoretical biochemistry – processes and properties of biological systems. In: Theoretical and Computational Chemistry, vol. 9, pp. 1–55, Eriksson (ed.), Elsevier Science, and references therein
10 Ryde U., Olsson M. H. M., Roos B. O. and Borin A. C. (2001) A theoretical study of the copper-cysteine bond in blue copper proteins. Theor. Chem Acc. **105:** 452–462, and references therein
11 van Gastel M., Coremans W. A., Sommerdijk H., van Hemert M. C. and Groenen E. J. J. (2002) An ab-initio quantum-chemical study of the blue copper site of azurin. J. Am. Chem. Soc. **124:** 2035–2041, and references therein
12 Adman E. T. (1992) Copper protein structures. Adv. Prot. Chem. **42:** 145–197, and references therein
13 Hart P. J., Nersissian A. M., Herrmann R. G., Nalbandyan R. M., Valentine J. S. and Eisenberg D. (1996) A missing link in cupredoxins: crystal structure of cucumber stellacyanin at 1.6 angstrom resolution. Prot. Sci. **5:** 2175–2183
14 Malmstrom B. G. (1994) Rack-induced bonding in blue-copper proteins. Eur. J. Biochemistry **223:** 711–718
15 Williams R. J. P. (1995) Energised (entatic) states of groups and of secondary structures in proteins and metalloproteins. Eur. J. Biochem. **234:** 363–381
16 Frisch M. J., Trucks G. W., Schlegel A. B., Grill P. M. W., Johnson B. G., Robb M. A. et al. (1994) GAUSSIAN 94, Gaussian, Pittsburgh, PA
17 Tomasi J., and Persico M. (1994) Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. Chem. Rev. **94:** 2027–2094
18 Ryde U. (1996) The coordination of the catalytic zinc in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations. J. Comp.-Aided Mol. Design **10:** 153–164
19 Case D. A., Pearlman D. A., Cadwell J. W., Cheatham III T. E., Ross W. S., Simmering C. L. et al. (1997) Amber 5.0, University of California, San Francisco
20 Weiner S. J., Kollman P. A., Case D. A., Singh U. C., Ghio C., Alagona G. et al. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins J. Am. Chem. Soc. **106:** 765–784
21 Comba P., Lledos A., Maseras F. and Remenyi R. (2001) Hybrid quantum mechanics/molecular mechanics studies of the active site of the blue copper proteins amicyanin and rusticyuanin. Inorg. Chim. Acta **324:** 21–26
22 Allinger N. L. (ed.) (1992) Quantum Chemistry Program Exchange, Bloomington, IN
23 Arcangeli C., Bizzarri A. R. and Cannistraro S. (2001) Concerted motions in copper plastocyanin and azurin: an essential dynamics study. Biophys. Chem. **90:** 45–56
24 Arcangeli C., Bizzarri A. R. and Cannistraro S. (2001) Molecular dynamics simulation and essential dynamics study of mutated plastocyanin: structural, dynamical and functional effects of a disulfide bridge insertion at the protein surface. Biophys. Chem. **92:** 183–199
25 Bizzarri A. R. and Cannistraro S. (2002) Molecular dynamics of water at the protein-solvent interface J. Phys. Chem. B **106:** 6617–6633, and references therein
26 Luise A., Falconi M. and Desideri A. (2000) Molecular dynamics simulation of solvated azurin: correlation between surface solvent accessibility and water residence times. Proteins **39:** 56–67
27 Buning C. and Comba P. (2000) Protonation of the copper(I) form of the blue copper proteins plastocyanin and amicyanin – a molecular dynamics study. Eur. J. Inorg. Chem. **6:** 1267–1273
28 Kuhn L. A., Siani M. A., Pique M. E., Fisher C. L., Getzoff E. D. and Tainer J.A. (1992) The interdependence of protein surface-topography and bound water molecules revealed by surface accessibility and fractal density measures. J. Mol. Biol. **228:** 13–22
29 Smoluchowski M. V. (1917) Versuch einer mathematischen Theorie der Koagulationskinetik Kollider Löschungen. Z. Phys. Chem. **92:** 129–168
30 Northrup S. H. and Erickson H. P. (1992) Kinetics of protein-protein association explained by Brownian dynamics computer simulations. Proc. Natl. Acad. Sci. USA **89:** 3338–3342
31 Schreiber G. and Fersht A. R. (1996) Rapid, electrostatically assisted association of proteins. Nat. Struct. Biol. **3:** 427–431
32 Gabdoulline R. R. and Wade R. C. (2001) Protein-protein association: investigation of factors influencing association rates by Brownian dynamics simulations. J. Mol. Biol. **306:** 1139–1155
33 Gabdoulline R. R. and Wade R. C. (2002) Biomolecular diffusional association. Curr. Opin. Struct. Biol. **12:** 204–213
34 Sykes A. G. (1994) Active-site properties of the blue copper proteins. Adv. Inorg. Chem. **36:** 377–408
35 Carrell C. J., Schlarb B. G., Bendall D. S., Howe C. J., Cramer W. A. and Smith J. L. (1999) Structure of the soluble domain of cytochrome f from the cyanobacterium *Phormidium laminosum*. Biochemistry **38:** 9590–9599
36 Fromme P., Jordan P. and Krauss N. (2001) Structure of photosystem I. Biochim. Biophys. Acta **1507:** 5–31
37 Jordan P., Fromme P., Witt H. T., Klukas O., Saenger W. and Krauss N. (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5 angstrom resolution. Nature **411:** 909–917
38 Ubbink M., Ejdeback M., Karlsson B. G. and Bendall D. S. (1998) The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic NMR and restrained rigid-body molecular dynamics. Structure **6:** 323–335
39 Crowley P. B., Otting G., Schlarb-Ridley B. G., Canters G. W. and Ubbink M. (2001) Hydrophobic interactions in a cyanobac-

terial plastocyanin-cytochrome f complex. J. Am. Chem. Soc. **123:** 10444–10453
40 Redinbo M. R., Yeates T. O. and Merchant S. (1994) Plastocyanin: structural and functional analysis. J. Bioenerg. Biomembr. **26:** 49–66, and references therein
41 Baker E. N. (1994) Copper proteins with type 1 sites. In: Encyclopedia of Inorganic Chemistry, pp. 883–904, King R. B. (ed.), Wiley Interscience, Chichester, and references therein
42 Sato K., Kohzuma T. and Dennison C. (2003) Active-site structure and electron-transfer reactivity of plastocyanins. J. Am. Chem. Soc. **125:** 2101–2112
43 Martinez S. E., Huang D., Szczepaniak A., Cramer W. A. and Smith J. L. (1994) Crystal structure of chloroplast cytochrome f reveals a novel cytochrome fold and unexpected heme ligation. Structure **2:** 95–105
44 Gong X.-S., Wen J. Q., Fischer N. E., Young S., Howe C. J. et al. (2000) The role of individual lysine residues in the basic patch on turnip cytochrome f for electrostatic interactions with plastocyanin in vitro. Eur. J. Biochem. **267:** 3461–3468
45 Illerhaus J., Altschmidt L., Reichert J., Zak E., Herrmann R. G. and Haehnel W. (2000) Dynamic interaction of plastocyanin with the cytochrome bf complex. J. Biol. Chem. **275:** 17590–17595
46 Wagner M. J., Parker J. C. L., Howe C. J. and Bendall D. S. (1996) Some characteristics of cytochrome f in the cyanobacterium *Phormidium laminosum*: its sequence and charge properties in the reaction with plastocyanin. Biochim. Biophys. Acta **1276:** 246–252
47 Pearson D. C. Jr, Gross E. L. and David E. S. (1996) Electrostatic properties of cytochrome f: implications for docking with plastocyanin. Biophys. J. **71:** 64–76
48 Soriano G. M., Cramer W. A. and Krishtalik L. I. (1997) Electrostatic effects on electron-transfer kinetics in the cytochrome f-plastocyanin complex. Biophys. J. **73:** 3265–3276
49 Pearson D. C. Jr and Gross E. L. (1998) Brownian dynamics study of the interaction between plastocyanin and cytochrome f. Biophys. J. **75:** 2698–2711
50 Ullmann G. M., Knapp E.-W. and Kostic N. M. (1997) Computational simulation and analysis of dynamic association between plastocyanin and cytochrome f. Consequences for the electron-transfer reaction. J. Am. Chem. Soc. **119:** 42–52
51 Qin L. and Kostic N. M. (1993) Importance of protein rearrangement in the electron-transfer reaction between the physiological partners cytochrome f and plastocyanin. Biochemistry **32:** 6073–6080
52 Meyer T. E., Zhao Z. G., Cusanovich M. A. and Tollin G. (1993) Transient kinetics of electron transfer from a variety of c-type cytochromes to plastocyanin. Biochemistry **32:** 4552– 4559
53 Kannt A., Young S. and Bendall D. S. (1996) The role of acidic residues of plastocyanin in its interaction with cytochrome f. Biochim. Biophys. Acta **1277:** 115–126
54 Schlarb-Ridley B. G., Bendall D. S. and Howe C. J. (2003) Relation between interface properties and kinetics of electron transfer in the interaction of cytochrome *f* and plastocyanin from plants and the cyanobacterium *Phormidium laminosum*. Biochemistry **42:** 4057–4063
55 Hart S. E., Schlarb-Ridley B. G., Delon C., Bendall D. S. and Howe C. J. (2003) Role of charges on cytochrome *f* from the cyanobacterium *Phormidium laminosum* in its interaction with plastocyanin. Biochemistry **42:** 4829–4836
56 Schlarb-Ridley B. G., Navarro J. A., Spencer M., Bendall D. S., Hervas M., Howe C. J. and De la Rosa M. A. (2002) Role of electrostatics in the interaction between plastocyanin and photosystem I of the cyanobacterium *Phormidium laminosum*. Eur. J. Biochem. **269:** 5893–5902
57 Cunha C. A., Romao M. J., Sadeghi S. J., Valetti F., Gilardi G. and Soares C. M. (1999) Effects of protein-protein interactions on electron transfer: docking and electron transfer calculations for complexes between flavodoxin and c-type cytochromes. J. Biol. Inorg. Chem. 360–374
58 Soriano G. M., Ponamarev M. V., Piskorowski R. A. and Cramer W. A. (1998) Identification of the basic residues of cytochrome f responsible for electrostatic docking interactions with plastocyanin in vitro: Relevance to the electron transfer reaction in vivo. Biochemistry **35:** 15120–15128
59 De Rienzo F., Gabdoulline R. R., Menziani M. C., De Benedetti P. G. and Wade R. C. (2001) Electrostatic analysis and Brownian dynamics simulation of the association of plastocyanin and cytochrome f. Biophys. J. **81:** 3090–3104
60 Gabdoulline R. R. and Wade R. C. (1998) Brownian dynamics simulation of protein-protein diffusional encounter. Methods **14:** 329–341
61 Elcock A. H., Gabdoulline R. R., Wade R. C. and McCammon J. A. (1999) Computer simulations of protein-protein association kinetics: acetylcholinesterase-fasciculin. J. Mol. Biol. **291:** 149–162
62 Crnogorac M. M., Ullmann G. M. and Kostic N. M. (2001) Effects of pH on protein association: modification of the proton-linkage model and experimental verification of the modified model in the case of cytochrome c and plastocyanin. J. Am. Chem. Soc. **123:** 10789–10798
63 Wyman J. (1964) Linked functions and reciprocal effects in hemoglobin: a second look. Adv. Protein Chem. **19:** 223–286
64 Gross E. L. and Pearson D. C. (2003) Brownian dynamics simulations of the interaction of *Chlamydomonas* cytochrome f with plastocyanin and cytochrome c6. Biophys. J. **85:** 2055–2068
65 Northrup S. H. (1999) Macrodox v. 2.3.1, software for the prediction of the macromolecular interactions, Tennessee Technological University, Cookvilee, TN
66 Falzone C. J., Kao Y.-H., Zhao J., Bryant D. A. and Lecomte J. T. J. (1994) Three-dimensional solution structure of PsaE from the cyanobacterium synechococcus sp. strain PCC 7002, a photosystem I protein that shows structural homology with SH3 domains. Biochemistry **33:** 6052–6062
67 Myshkin E., Leontis N. B. and Bullerjahn G. S. (2002) Computational simulation of the docking of *Prochlorothrix hollandica* plastocyanin to potosystem I: modeling the electron transfer complex. Biophys. J. **82:** 3305–3313, and references therein
68 Babu C. R., Volkman B. F. and Bullerjahn G. S. (1999) NMR solution structure of plastocyanin from the photosynthetic prokaryote, *Prochlorothrix hollandica*. Biochemistry **38:** 4988–4995
69 Vakser I. A. and Alalo C. (1994) Hydrophobic docking: a proposed enhancement to molecular recognition techniques. Proteins **20:** 320–329
70 Gabb H. A., Jackson R. M. and Sternberg M. J. (1997) Modelling protein docking using shape complementarity, electrostatics and biochemical information. J. Mol. Biol. **272:** 106–120
71 Morris G. M., Goodsell D. S., Halliday R. S., Huey R., Hart W. E., Belew R. K. et al. (1998) Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. J. Comput. Chem. **19:** 1639–1662
72 Mandell J. G., Roberts V. A., Pique M. E., Kotlovyi V., Mitchell J. C., Nelson E. et al. (2001) Protein docking using continuum electrostatics and geometric fit. Protein Eng. **14:** 105–113
73 Cheatham T. E., Srinivasan J., Case D. A. and Kollman P. A. (1998) Molecular dynamics and continuum solvent studies of the stability of polyG-polyC and polyA-polyT DNA duplexes in solution. J. Biol. Struct. Dyn. **16:** 265–280.
74 Srinivasan J., Cheatham T. E., Cieplak P., Kollman P. A. and Case D. A. (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. J. Am. Chem. Soc. **120:** 9401–9409
75 Navarro J. A., Myshkin E., De la Rosa M. A., Bullerjahn G. S. and Hervas M. (2001) The unique proline of the *Prochlorothrix hollandica* plastocyanin hydrophobic patch impairs electron transfer to photosystem I. J. Biol. Chem. **276:** 37501–37505

76 Larsson S. (2000) Energy saving electron pathways in proteins. J. Biol. Inorg. Chem. **5:** 560–564
77 Larsson S., Broo A. and Sjölin L. (1995) Connection between structure, electronic spectrum, and electron-transfer properties of blue copper proteins. J. Phys. Chem. **99:** 4860–4865
78 Olsson M. H. M., Ryde U. and Roos B. O. (1998) Quantum chemical calculations of the reorganization energy of blue-copper proteins. Prot. Sci. **7:** 2659–2668
79 Sigfridsson K., Sundahl M., Bjerrum M. J. and Hansson _. (1996) Intraprotein electron transfer in a ruthenium-modified Tyr83-His plastocyanin mutant: Evidence for strong electronic coupling. J. Biol. Inorg. Chem. **1:** 405–414
80 Gray H. B., Malmström B. G. and Williams R. J. P. (2000) Copper coordination in blue proteins. J. Biol. Inorg. Chem. **5:** 551–559
81 Ridley J. and Zerner M. T. (1976) Triplet states via intermediate neglect of differential overlap: benzene, pyridine and the diazines. Theor. Chim. Acta **42:** 223–323
82 Beratan D. N., Betts J. N. and Onuchic J. N. (1991) Protein electron transfer rates set by the bridging secondary and tertiary structure. Science **252:** 1285–1288
83 Gross E. L., Curtiss A., Durell S. R. and White D. (1990) Chemical modification of spinach plastocyanin using 4-chloro-3,5-dinitrobenzoic acid: characterization of four singly-modified forms. Biochim. Biophys. Acta **1016:** 107–114
84 Hibino T., Lee B. H., Yajima T., Odani A., Yamauchi O. and Takabe T. (1996) Kinetic and cross linking studies on the interactions of negative patch mutant plastocyanin from *Silene pratensis* with photosystem I complexes from cyanobacteria, green alga Malmstrom E., and plants. J. Biochem. **120:** 556– 563
85 De Rienzo F., Grant G. H. and Menziani M. C. (2002) Theoretical descriptors for the quantitative rationalization of plastocyanin mutant functional properties. J. Comp.-Aided Mol. Design **16:** 501–509, and references therein
86 Olsson M. H., Hong G. and Warshel A. (2003) Frozen density functional free energy simulations of redox proteins: computational studies of the reduction potential of plastocyanin and rusticyanin. J. Am. Chem. Soc. **125:** 5025–5039, and references therein
87 Lee F. S., Chu Z. T. and Warshel A. (1993) Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the polaris and enzymix programs. J. Comput. Chem. **14:** 161–185
88 Williams P. A., Fülöp V., Leung Y.-C., Chan C., Moir J. W. B., Howlett G. et al. (1995) Pseudospecific docking surfaces on electron transfer proteins as illustrated by pseudoazurin, cytochrome c550 and cytochrome cd1 nitrite reductase. Nat. Struct. Biol. **2:** 975–982
89 Yamanaka T. and Fukumori Y. (1995) Molecular aspects of the electron transfer system which participates in the oxidation of ferrous ion by *Thiobacillus ferrooxidans*. FEMS Microbiol. Rev. **17:** 401–413
90 De Rienzo F., Gabdoulline R. R., Menziani M. C. and Wade R. C. (2000) Blue copper proteins: a comparative analysis of their molecular interaction properties. Prot. Sci. **9:** 1439–1454 and references therein
91 Blomberg N., Gabdoulline R. R., Nilges M. and Wade R. C. (1999) Classification of protein sequences by homology modeling and quantitative analysis of electrostatic similarity. Proteins **37:** 379–387
92 Hodgkin E. E. and Richards W. G. (1987) Molecular similarity based on electrostatic potential and electric field. Int. J. Quant. Chem. Quant. Biol. Symp. **14:** 105–110
93 Dodd F. E., Van Beeumen J., Eady R. R. and Hasnain S. S. (1998) X-ray structure of a blue-copper nitrite reductase in two crystal forms. The nature of the copper sites, mode of substrate binding and recognition by redox partner. J. Mol. Biol. **282:** 369–382
94 Battistuzzi G., Corsari M., Loschi L., Menziani M. C., De Rienzo F. and Sola M. (2001) Control of metalloprotein reduction potential: the role of electrostatic and solvation effects probed on plastocyanin mutants. Biochemistry **40:** 6422– 6430
95 Stanton D. T. and Jurs P. C. (1990) Development and use of charged partial surface area structural descriptors in computer-assisted quantitative structure-property relationship studies. Anal. Chem. **62:** 2323–2329
96 Brooks B. R., Bruccoleri R. E., Olafson B. D., States D. J., Swaminathan S. and Karplus M. (1983) CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J. Comput. Chem. **4:** 187–217
97 Kier L. B. (1980) Use of molecular negentropy to encode structure governing biological activity. J. Pharm. Sci. **69:** 807–811, and references therein