Control of Metalloprotein Reduction Potential: The Role of Electrostatic and Solvation Effects Probed on Plastocyanin Mutants[†]

Gianantonio Battistuzzi, Marco Borsari, Lodovica Loschi, Maria Cristina Menziani, Francesca De Rienzo, and Marco Sola*

Department of Chemistry, University of Modena and Reggio Emilia, Via Campi 183, 41100 Modena, Italy

Received November 7, 2000

ABSTRACT: The changes in the thermodynamics of Cu(II) reduction for spinach plastocyanin induced by point mutations altering the electrostatic potential in proximity of the copper center were determined through variable temperature direct electrochemistry experiments. In particular, the functionally important surface residues Leu12 and Gln88 were replaced with charged and polar residues, and Asn38 was substituted with Asp. The mutational variations of the reduction enthalpy and entropy were analyzed with a QSPR (quantitative structure−property relationships) approach, employing global and local theoretical descriptors defined and computed on the three-dimensional protein structure. The correlations found are informative on how electrostatic and solvation effects control the $E^{\circ\prime}$ values in this species through the combined effects on the reduction enthalpy and entropy. The changes in reduction enthalpy can be justified with electrostatic considerations. Most notably, enthalpy−entropy compensation phenomena play a significant role: the entropic effects due to the insertion of charged residues determine $E^{\circ\prime}$ changes that are invariably opposite to those induced by the concomitant enthalpic effects. Therefore, the resulting $E^{\circ\prime}$ changes are small or even opposite to those expected on simple electrostatic grounds. The mutational variation in the reduction entropy appears to be linked to the hydrogen bonding donor/acceptor character of the northern part of the protein, above the metal site, and to the electrostatic potential distribution around the copper site. Both properties influence the reduction-induced reorganization of the water molecules on the protein surface in the same region.

The question of how and to what extent the protein matrix and the solvent concur to the selective stabilization of the two oxidation states of the metal center in redox metalloproteins is an important and largely debated issue of biological redox chemistry (*1*−*6*). Comprehension of these effects would be the key for establishing relationships between the observed reduction potential ($E^{\circ\prime}$) and protein structure/sequence features which, in turn, would allow predictions to be made regarding the functional properties of mutants and newly engineered proteins. Extensive theoretical and experimental work on electron transport metalloproteins (cytochromes, blue copper, and iron−sulfur proteins) has shown that the protein controls the reduction potential of the prosthetic center essentially through metal−ligand electronic interactions and electrostatic effects of permanent and induced protein dipoles, net protein charges, and solvent dipoles (*1*−*6*). In addition, the $E^{\circ\prime}$ values are sensitive to the ionic composition of the medium (*7*, *8*). However, the relative importance of these factors may change in different protein families (*2*, *9*).

Surface net charges have been established to contribute to the $E^{\circ\prime}$ value only by approximately 10−30 mV per charge (*1*−*4*, *6*) depending on their location with respect to the metal center. This moderate effect is attributed mainly to the screening of the Coulombic interactions by the dielectric effect of water. The $E^{\circ\prime}$ variations due to proton uptake or release by protein residues are in general modest as well, unless the acid−base equilibrium involves a metal-binding residue (*10*, *11*). In general, no correlations emerged between the total protein charge and the $E^{\circ\prime}$ values within the same protein family, HiPIPs being the only exception most likely as a result of the constancy or self-compensation of the other effectors of $E^{\circ\prime}$ within the series (*2*, *3*). Interestingly, the changes in $E^{\circ\prime}$ induced by point mutations involving addition/suppression of surface charges in proximity of the metal center are not always those expected from simple electrostatic considerations (*2*, *12*). Hence, further investigation is needed to gain a deeper understanding of these electrostatic effects.

Here, we analyze the role played by net and fractional charges as modulators of $E^{\circ\prime}$ by making use of a combined experimental−computational approach. In particular, we have measured electrochemically the variations of the reduction thermodynamics of spinach plastocyanin (pc)[1] (*13*) determined by point mutations involving changes in the net surface charges in proximity of the metal center and derived quantitative models for the rationalization and interpretation of the observed behaviors. We replaced the surface residues Leu12 and Gln88 with positively charged (Arg) and nega-

[†] This work was supported by the Ministero dell'Università e della Ricerca Scientifica e Tecnologica of Italy (Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale, PRIN 1998) and by CNR.

* Correspondence should be addressed to this author at the Department of Chemistry, University of Modena and Reggio Emilia, Via Campi 183, 41100 Modena, Italy. Tel: ++39-059 205 5037, Fax: ++39 059 373 543, E-mail: sola@unimo.it, URL: www.chimica.unimo.it/websola/bioichemgroup.html.

10.1021/bi002565d CCC: $20.00
Published on Web 05/04/2001

tively charged (Glu) residues. In addition, we probed for the effects of changing polarity (with the Leu12His and Leu12Gln mutants) and solvent accessibility of the site (Leu12Gly). Leu12 contributes to the hydrophobic patch at the northern site of the protein and seems to be involved in the interaction with photosystem 1 (PS1) (*14*). The side chain of residue Gln88 belongs to the acidic patch at the eastern site of the protein which is involved in the interaction with cytochrome *f* (*15*). Moreover, we have monitored the effect of substituting Asn38 with Asp. Asn38 does not contribute to the protein surface, but it has been shown to occupy a strategic position for the stabilization of the architecture of the copper site (*16*).

Our starting consideration is that addition/suppression of a charge not only exerts a Coulombic effect on the selective stabilization of the two redox states, which is an exquisitely enthalpic effect, but also can in principle influence the entropy change of the electron exchange process. In fact, perturbation of the charge distribution can modify the mobility of more or less localized protein regions and the hydrogen bonding network within the protein hydration shell to a different extent in the two redox states. Therefore, entropy changes are likely to contribute to the measured $\Delta E^{\circ\prime}$ due to the modification of surface charges. It follows that interpretations/predictions of the effects of charge changes on $E^{\circ\prime}$ based simply on electrostatic considerations hold only for one of the contributions to the free energy change accompanying protein reduction, and may thus be misleading. We show in this paper that deeper insights can be obtained by means of a quantitative analysis of the variations in the reduction enthalpy and entropy due to charge changes, employing global and local theoretical descriptors defined and computed on the three-dimensional (3D) protein structures.

The techniques of computational chemistry are being applied increasingly in the generation of theoretical indexes able to describe biologically active molecules in quantitative structure−property relationship (QSPR) studies (*17*). In fact, the ability to understand the mutual influence of structure and property and to obtain meaningful quantitative rationalization of the variation of the property of interest depends primarily on the availability of descriptors able to decipher the encrypted code of the molecular structure. The theoretical QSPR methodology has assumed an immense practical importance in the development of medicinally active compounds (small molecule), whereas applications of QSPR techniques to peptide and proteins has been less developed. Current studies rely essentially on amino acid side chain scores derived by the application of the principal component analysis (PCA) to descriptor matrixes of different nature, such as empirical scales, 3D descriptors, interaction properties descriptors, etc. (*18*). Sound statistical models with good predictive ability have been reported in the literature, but the approach suffers two severe drawbacks: the descriptors are computed on isolated amino acids or on very simple model systems, and the physical interpretation of the correlations obtained is often difficult.

To our knowledge this is the first time in which a QSPR approach employing descriptors originally derived to rationalize the reactivity determinants of small molecules such as drugs is extended to a series of protein variants in order to pinpoint the causes which determine variations in a molecular property crucial for the function, such as the $E^{\circ\prime}$ value for electron transport proteins.

EXPERIMENTAL PROCEDURES

Proteins. Recombinant wild-type spinach plastocyanin and the L12G, L12E, L12K, L12H, L12Q, Q88K, Q88E, and N38D mutants were isolated from JM105 *E. coli* cells transformed with the expression system pTrc99A containing the construct of the peptide signal of *P. aeruginosa* azurin and the plastocyanin gene, following the procedure described elsewhere (*19*, *20*). The transformed cells were a gift from Prof. Klaus Bernauer and Prof. Peter Schurmann of the Université de Neuchâtel, Switzerland. To generate the mutants Leu12Glu, Leu12Lys, Leu12His, Leu12Gln, Gln88Glu, Gln88Lys, and Asn38Asp, the wild-type gene was mutated following the procedure of the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and using the mutagenic oligonucleotides L12Ef, L12Er, L12Kf, L12Kr, L12Hf, L12Hr, L12Qf, L12Qr, Q88Ef, Q88Er, Q88Kf, Q88Kr, N38Df, and N38Dr. The presence of the mutation was confirmed by standard sequencing methods. JM105 *E. coli* competent cells were transformed with the constructed plasmids and plated on LB plates containing 50 μg/mL ampicillin. Cells harboring the plasmids were selected by their ability to grow on LB plates containing 50 μg/mL ampicillin. The proteins were purified following the procedure described elsewhere (*19*, *20*); purified proteins gave a single band on PAGE electrophoresis (Bio-Rad Mini Protean) around 10 000 Da. All chemicals were reagent grade and were used without further purification. Nanopure water was used throughout.

Electrochemical Measurements. Cyclic voltammetry (CV) experiments were performed with a Potentiostat/Galvanostat PAR model 273A. A 1 mm diameter pyrolitic graphite disk (PGE) was used as working electrode, and a saturated calomel electrode and a 5 mm diameter Pt as a reference and counter-electrode, respectively. Potentials were calibrated against the MV^{2+}/MV^{+} couple (MV = methyl viologen) (*21*). All the reduction potentials reported in this paper are referred to the standard hydrogen electrode. The electric contact between the reference electrode and the working solution is obtained with a Vycor set. All measurements were carried out under argon using a cell for small volume samples ($V = 0.5$ mL) under thermostatic control. Scan rates varied from 0.02 to 0.5 V s^{-1}. The cleaning procedure of the working electrode is crucial to the voltammetric response. The PGE was first treated with anhydrous ethanol for 5 min, and then polished with alumina (BDH, particle size of about 0.015 μm) water slurry on cotton wool for 5 min; finally the electrode was treated in an ultrasonic pool for about 5 min. Modification of the electrode surface was performed by dipping the polished electrode into a 1 mM solution of polylysine and morpholine (1 mM) for 30 s and then rinsing it with Nanopure water. Protein samples were freshly prepared before use, and their concentration, in general about 0.1 mM, was checked spectrophotometrically. A single voltammetric wave was observed for all species which was either reversible or quasi-reversible. Peak separation in CV

[1] Abbreviations: QSPR, quantitative structure−property relationships; pc, spinach plastocyanin; PCA, principal component analysis; CV, cyclic voltammetry; PGE, pyrolytic graphite electrode; SCE, saturated calomel electrode; SHE, standard hydrogen electrode; SI, Hodgkin similarity index; BIC, bonding information content; DPSA, difference in the charged partial surface areas of the protein; *wt*, wild type.

experiments varied from 60 to 90 mV for scan rates in the range 0.02−0.2 V s^{-1}. Anodic and cathodic peak currents were almost identical, and both were proportional to protein concentration and $\upsilon^{1/2}$ (υ = scan rate), indicating a diffusion-controlled electrochemical process. Given the reversibility or quasi-reversibility of the electrochemical process, the symmetrical shape of the voltammograms, and the almost negligible influence of the scan rate on the half-wave potentials, the $E_{1/2}$ values (taken as the average of the cathodic and anodic peak potentials) can be confidently assumed as the $E^{\circ\prime}$ values. The temperature dependence of the reduction potential was determined with a "nonisothermal" cell (*22*, *23*) in which the reference electrode is kept at constant temperature, while the half-cell containing the working electrode and the Vycor junction to the reference electrode is under thermostatic control with a water bath. The temperature was varied from 5 to 40 °C. With this experimental configuration, the reaction entropy for reduction of the oxidized plastocyanin ($\Delta S^{\circ\prime}_{rc}$) is given by (*22*−*24*)

$$\Delta S^{\circ\prime}_{rc} = S^{\circ\prime}_{red} - S^{\circ\prime}_{ox} = nF(dE^{\circ\prime}/dT)$$

Thus, $\Delta S^{\circ\prime}_{rc}$ was determined from the slope of the plot of $E^{\circ\prime}$ versus temperature which turns out to be linear under the hypothesis that $\Delta S^{\circ\prime}_{rc}$ is constant over the limited temperature range investigated. With the same assumption, the enthalpy change ($\Delta H^{\circ\prime}_{rc}$) was obtained from the Gibbs−Helmholtz equation, namely, from the slope of the $E^{\circ\prime}/T$ versus $1/T$ plot.

The nonisothermal behavior of the cell was carefully checked by determining the $\Delta H^{\circ\prime}_{rc}$ and $\Delta S^{\circ\prime}_{rc}$ values of the ferricyanide/ferrocyanide couple (*22*−*25*). For each protein, the experiments were performed at least 2 times, and the reduction potentials were found to be reproducible within ±2 mV.

Computational Procedures. The crystallographic structure of spinach plastocyanin (*26*) was obtained from the Brookhaven Protein DataBank (PDB entry 1ag6). Mutants were obtained by substituting the appropriate residue in the wild-type structure using the Biopolymer module of the InsightII software package (MSI, San Diego, CA, 1995).

Refinement and Analysis of the Protein 3D Structures. Molecular mechanics and molecular dynamics calculations were performed using the program CHARMM (*27*). The minimization procedure consisted of 50 steps of steepest descent, followed by a conjugate gradient minimization until the rms gradient of the potential energy was less than 0.001 kcal/(mol·Å). The united atom force-field parameters, a 12 Å nonbonded cutoff, and a dielectric constant of 80 were used. The minimized coordinates were used as the starting point for dynamics. During dynamics, the lengths of the bonds involving hydrogen atoms were constrained according to the SHAKE algorithm (*28*), allowing an integration time step of 0.001 ps. The β-sheet structure was maintained by applying harmonic constraints (*26*) between opposite strands. The coordination geometry of the Cu site was maintained by applying weak harmonic constraints to the interatomic distances Cu−ND1 of the two His ligands, Cu−SD of the Met ligand, and Cu−SG of the Cys ligand. The plastocyanin and its mutants were considered in their oxidized state. The CHARMM default parameters were used for Cu (i.e., formal charge of +2 was given to the Cu atom).

The structures were thermalized to 300 K with 5° rise per 6000 steps by randomly assigning individual velocities from the Gaussian distribution. After heating, the systems were allowed to equilibrate until the potential energy versus time was approximately stable (34 ps). Velocities were scaled by a single factor. An additional 10 ps period of equilibration with no external perturbation was run. Data were collected every 0.5 ps over the last 50 ps of each simulation. The time-averaged structures were subjected to minimization and data analysis for the determination of the descriptor used in the QSPR modeling.

Molecular Electrostatic Potential Calculations. The University of Houston Brownian Dynamics (UHBD) program (*29*) was used to compute the electrostatic potentials and the dipole moments of the proteins. The N- and C-terminal residues were treated as charged. The ionizable residues were considered in their usual protonation states at pH 7. Both the neutral and the protonated forms were considered for the histidine residue in the L12H mutant. The OPLS (*30*) nonbonded parameter sets for atomic charges and radii were assigned to the protein residues. The charge parameters were assigned to the copper site (Cu and its ligands) using a simple distribution algorithm (*31*), developed on the basis of that proposed by Libeu (*32*). A grid dimension of 110 × 110 × 110 $Å^3$ was assigned together with a 1.0 Å grid spacing for computing the electrostatic potential by finite difference solution of the linearized Poisson−Boltzmann equation. The grid was centered on the global center of mass of the superimposed structures. The dielectric constants of the solvent and the protein were set to 78 and 2, respectively. The dielectric boundary was determined from the van der Waals surface of the protein, and dielectric boundary smoothing (*33*) was implemented. The molecular electrostatic potential grids were computed at an ionic strength of 100 mM.

Calculation of the Similarity Indices (SIs). The minimized average structures were superimposed with the Quanta program (MSI, San Diego, CA) using a standard least-squares fitting algorithm. The fit was optimized on all Cα atoms. For each pair of proteins *a* and *b*, the molecular potentials, $\phi_a(i,j,k)$ and $\phi_b(i,j,k)$, computed on a three-dimensional grid, were compared by calculating the Hodgkin similarity index (SI) (*34*). The SIs for comparison of two proteins are calculated for grid points within the intersection of regions, called "skins", which are chosen to be at a distance σ from the van der Waals surface of each protein and to have a thickness δ. The parameters of σ = 3 Å and δ = 4 Å were chosen so that the region where the potentials are compared is not so close to the protein surfaces as to be highly sensitive to small changes in protein structure and not so far that only the major components of the potentials (e.g., electrostatic monopoles) can be detected. The SI values lie in the range −1 to 1, with values near 1 implying that the proteins have highly similar potentials on the overlapping skins and values near −1 implying that the potentials have inverted distributions. The details of the method have been described by Blomberg et al. (*35*).

Definition of the Descriptors Used in the Selected QSAR Models. (A) Dz. Dz is the component of the dipole moment of the proteins along the *z* axis, computed by means of the UHBD program (*29*). The *z* axis coincides with the main

axis of the elongated protein, with its origin in the geometric center and pointing toward the copper site.

(B) BIC. BIC is the bonding information content defined on the basis of the Shannon information theory (*36*). The statistical treatment of chemical structures by means of information theory relies on the premise that each structure contains a limit set of elements (n) that could be decomposed into disjoint subsets n_i ($i = 1, 2, ..., k$) by means of selected equivalence relations defined on the set: $IC = -\Sigma_i p_i \log_2 p_i$ bit, where $p_i = n_i/n$ is the probability that a randomly selected element will be in the ith subset. The logarithm is taken at a basis 2 in order to measure the information in bits. Therefore, BIC = IC/$\log_2$ (number of bonds counting bond orders) is independent of the three-dimensional structure of the protein and can easily be computed a priori from the secondary structure. According to its definition, this index encodes the branching ratio, unsaturation, and constitutional diversity of a molecule.

(C) SI. SI is the electrostatic potential similarity index computed in the "northern" site region of plastocyanin and its mutants. A vector was chosen having its origin at the center of mass of the superimposed structures and pointing toward the copper site. Only the intersection of those parts of the skin inside a conical region centered on the vector and with a 30° angular extent was considered. This index provides a measure of the similarity between the magnitudes and distributions of the molecular electrostatic potentials of the two proteins compared (*35*).

(D) DPSA. DPSA is the difference in the charged partial surface areas of the protein (DPSA = PPSA − PNSA) (*37*), computed on the minimized average structures. The partial positive and the partial negative surface areas of the proteins are derived as follow: $PPSA = \Sigma(+SA_i)$ and $PNSA = \Sigma(-SA_i)$, where $(+SA_i)$ and $(-SA_i)$ are the surface area contributions of the ith positive and negative atom, respectively. The amino acids included in the calculations are S11−F14, G34−V39, M57−L63, Y83, C84, S85−V93, and Cu100. They contribute to construct the solvent-accessible surface of the proteins in the region of interest, where the Cu atom and the mutated residues L12, N38, and Q88 are located.

(E) IE_{HB}. IE_{HB} is the hydrogen bonding contribution of the residue subjected to mutation to the total energy of the protein. This was computed on the minimized average structures of the proteins by using a dielectric constant $\epsilon = 4r$, with the CHARMM program (*27*).

RESULTS

Reduction Thermodynamics for Wild-Type and Mutated Plastocyanins. The $E^{\circ\prime}$ vs T and $E^{\circ\prime}/T$ vs $1/T$ profiles for *wt* pc and the variants L12E, L12K, Q88E, Q88K, L12H, L12Q, L12G, and N38D are shown in Figures 1 and 2, respectively. The reduction potentials and the thermodynamic parameters for the various species are listed in Table 1. The mutation-induced variations in these parameters are reported in Table 2. It is apparent that insertion of a negative charge, such as in the L12E and Q88E mutants, increases $\Delta H^{\circ\prime}_{rc}$ as compared to the *wt* species (i.e., induces a shift toward less negative values). On the contrary, insertion of a positive charge, as in L12K and Q88K, exerts a less pronounced effect in the opposite direction, which is significant only for the former species, given the experimental error in $\Delta H^{\circ\prime}_{rc}$ of ±2 kJ mol^{-1}. The His residue inserted in place of Leu12 in L12H

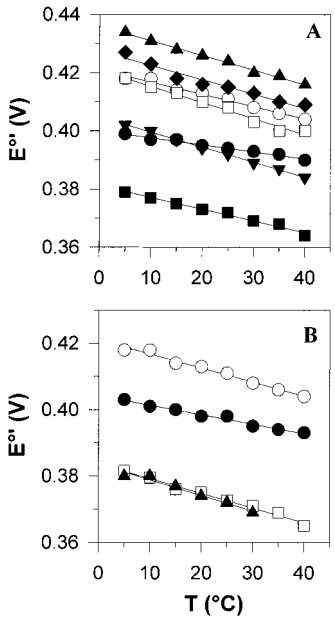

FIGURE 1: $E^{\circ\prime}$ vs T plots for wild-type and mutated spinach (*Spinacea oleracea*) plastocyanins at pH 7 (unless otherwise specified). The slope of the plots yields the $\Delta S^{\circ\prime}_{rc}$ values. (A) Charged substituting residues: L12E (●), L12K (▲), Q88E (■), Q88K (▼), L12H (pH 6.2) (□), N38D (◆), wild-type pc (○). (B) Neutral substituting residues: L12H (pH 7.4) (□), L12Q (▲), L12G (●), wild-type pc (○). Protein concentration, 0.1 mM; base electrolyte, 0.1 M phosphate, pH 7. Solid lines are least-squares fits to the data points. Error bars have the same dimensions of the symbols.

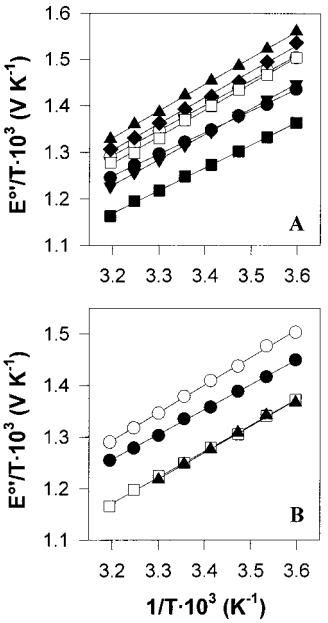

FIGURE 2: $E^{\circ\prime}/T$ vs $1/T$ plots for wild-type and mutated spinach (*Spinacea oleracea*) platocyanins at pH 7 (unless otherwise specified). The slope of the plots yields the $-\Delta H^{\circ\prime}_{rc}$ values. (A) Charged substituting residues: L12E (●), L12K (▲), Q88E (■), Q88K (▼), L12H (pH 6.2) (□), N38D (◆), wild-type pc (○). (B) Neutral substituting residues: L12H (pH 7.4) (□), L12Q (▲), L12G (●), wild-type pc (○). Protein concentration, 0.1 mM; base electrolyte, 0.1 M phosphate, pH 7. Solid lines are least-squares fits to the data points. Error bars have the same dimensions of the symbols.

Table 1: Thermodynamics of Protein Reduction for Wild-Type and Mutated Spinach Plastocyanin[a]

species	$E^{\circ\prime}$ (mV)	$\Delta H^{\circ\prime}_{rc}$ (kJ mol^{-1})	$\Delta S^{\circ\prime}_{rc}$ (J mol^{-1} K^{-1})	$-\Delta H^{\circ\prime}_{rc}/F$ (mV)	$T\Delta S^{\circ\prime}_{rc}/F$ (mV)
wt	+411	−52	−41	+539	−127
L12E	+394	−45	−23	+466	−71
L12K	+424	−55	−49	+570	−151
Q88E	+372	−47	−39	+487	−120
Q88K	+392	−53	−50	+549	−154
L12H (pH 6.2)	+408	−55	−54	+570	−167
L12H (pH 7.4)	+372	−48	−39	+497	−120
L12Q	+372	−49	−45	+508	−139
L12G	+398	−46	−28	+477	−86
N38D	+415	−54	−48	+560	−148

[a] Values refer to pH 7.0, unless otherwise specified. $T = 25$ °C. All measurements were done in 0.1 M phosphate buffer. Average standard errors on $\Delta H^{\circ\prime}_{rc}$ and $\Delta S^{\circ\prime}_{rc}$ are ±2 kJ mol^{-1} and ±6 J mol^{-1} K^{-1}, respectively. The sum $-\Delta H^{\circ\prime}_{rc}/F + T\Delta S^{\circ\prime}_{rc}/F$ often does not exactly correspond to $E^{\circ\prime}$ since, because of the experimental error, the $\Delta H^{\circ\prime}_{rc}$ and $\Delta S^{\circ\prime}_{rc}$ values are rounded to the closest integer.

Table 2: Mutation-Induced Changes in the Enthalpic and Entropic Contributions to $E^{\circ\prime}$ for Spinach Plastocyanin[a]

mutant	$\Delta E^{\circ\prime}$ ($=E^{\circ\prime}_{mut} - E^{\circ\prime}_{wt}$) (mV)	$-\Delta\Delta H^{\circ\prime}_{rc}/F$ (mV)	$T\Delta\Delta S^{\circ\prime}_{rc}/F$ (mV)
L12E	−17	−73	+56
L12K	+13	+31	−24
Q88E	−39	−52	+7
Q88K	−19	+10	−27
L12H (pH 6.2)	−3	+31	−40
L12H (pH 7.4)	−39	−42	+7
L12Q	−39	−31	−12
L12G	−13	−62	+41
N38D	+4	+21	−21

[a] Values refer to pH 7.0, unless otherwise specified. $T = 25$ °C.

shows a pK_a value of 6.9 (*38*), and therefore is mostly protonated at pH 6.2. Measurements were not performed at lower pH because of the incipient protonation and detachment from the metal of the ligand His87 (*10*, *38*, *39*). Analogously to above, the presence of an additional positive charge induces a decrease in the reduction enthalpy (Tables 1 and 2).

The variation in reduction entropy also differs for the two kinds of mutations. In fact, insertion of Glu and Lys results in positive and negative $\Delta\Delta S^{\circ\prime}_{rc}$ values, respectively (Table 2). It is noteworthy that in all cases the changes in $E^{\circ\prime}$ due to the enthalpic and entropic terms for the same mutant have opposite signs (Table 2), indicating that enthalpy−entropy compensation phenomena are operative. These effects are not observed for L12Q, for which the small variations in the enthalpic and entropic contributions to $E^{\circ\prime}$ as compared to *wt* pc sum up to yield a significant $E^{\circ\prime}$ change (Table 2).

Quantitative Description of Mutational Changes in Redox Thermodynamics. Among the great variety of molecular descriptors tested, we have selected those which furnish the best statistical indexes in the correlations with the reduction thermodynamic parameters of the studied plastocyanin mutants. The data values of global and local molecular descriptors selected are listed in Table 3. All these descriptors are computed on the minimized molecular dynamics average structure of the proteins, as described under Experimental Procedures.

The variation of the reduction enthalphy for the series of mutants considered is well explained by making use of the *z* component of the protein dipole moment (D*z*) (Figure 3a):

$$\Delta H^{\circ\prime}_{rc} = -0.0686\mathrm{D}z - 51.795,\ n = 9,\ R^2 = 0.858,\ s^2 = 2.38,\ F = 42.45,\ \text{omitted: N38D} \quad (1)$$

where n is the number of compounds, R is the correlation coefficient, s is the standard deviation, and F is the value of the Fisher ratio. The positive end of the dipole moment vector in the wild-type protein makes an angle of about −80° with the *z* axis, which extends from the protein center of mass toward the copper site. The regression shows that less negative $\Delta H^{\circ\prime}_{rc}$ values are associated with decreasing values of D*z*, namely, with a mutational shift of the center of gravity of the positive charge of the protein which moves away from the copper site, toward the opposite end of the oblong protein. The variant N38D is a clear outlier.

Additional insights are obtained from the bonding information content (BIC), defined on the basis of the Shannon information theory (*36*). The BIC descriptor explains the observed variation in the reduction enthalpies by capturing the salient structural characteristics of the protein, expressed as branching ratio, unsaturation, and constitutional diversity of the mutated residues. These molecular characteristics regulate the protein potential for nonspecific interactions, i.e., dispersion forces among nonpolar moieties (Figure 3b):

$$\Delta H^{\circ\prime}_{rc} = -3.71\mathrm{BIC} + 266.318,\ n = 9,\ R^2 = 0.629,\ s^2 = 6.245,\ F = 11.89,\ \text{omitted: N38D} \quad (2)$$

Therefore, it appears that an increased amount of dispersion interactions within the protein induces a stabilization of the

Table 3: Experimental Thermodynamic Parameters and Theoretical Molecular Descriptors Computed on the Minimized Average Structure of Plastocyanin and Its Mutants and Used for QSPR Modelling

mutant	$E^{\circ\prime}$ (mV)	$\Delta H^{\circ\prime}_{rc}$ (kJ/mol)	$\Delta S^{\circ\prime}_{rc}$ [J/(mol·K)]	D_z (D)	BIC	SI	DPSA (Å^2)	IE_{HB} (kcal/mol)
wt	411	−52	−41	28.00	85.06	1.000	6498.1	−2.23
L12E	394	−45	−23	−91.15	84.89	−0.170	6314.6	−2.19
L12K	424	−55	−49	36.85	86.51	0.270	6772.4	−5.61
Q88E	372	−47	−39	−44.54	84.42		6507.4	
Q88K	392	−53	−50	28.82	86.09		6847.4	
L12H (pH 7.4)	372	−48	−39	−54.19	85.40	0.160	6648.4	0.00
L12H (pH 6.2)	408	−55	−54	4.12	85.70	0.430	6795.4	−3.03
L12Q	372	−49	−45	−47.45	85.51	0.128	6691.6	0.00
L12G	398	−46	−28	−96.03	83.85	0.014	6440.6	−3.29
N38D	415	−54	−48	−124.93	84.42		6618.7	−3.86

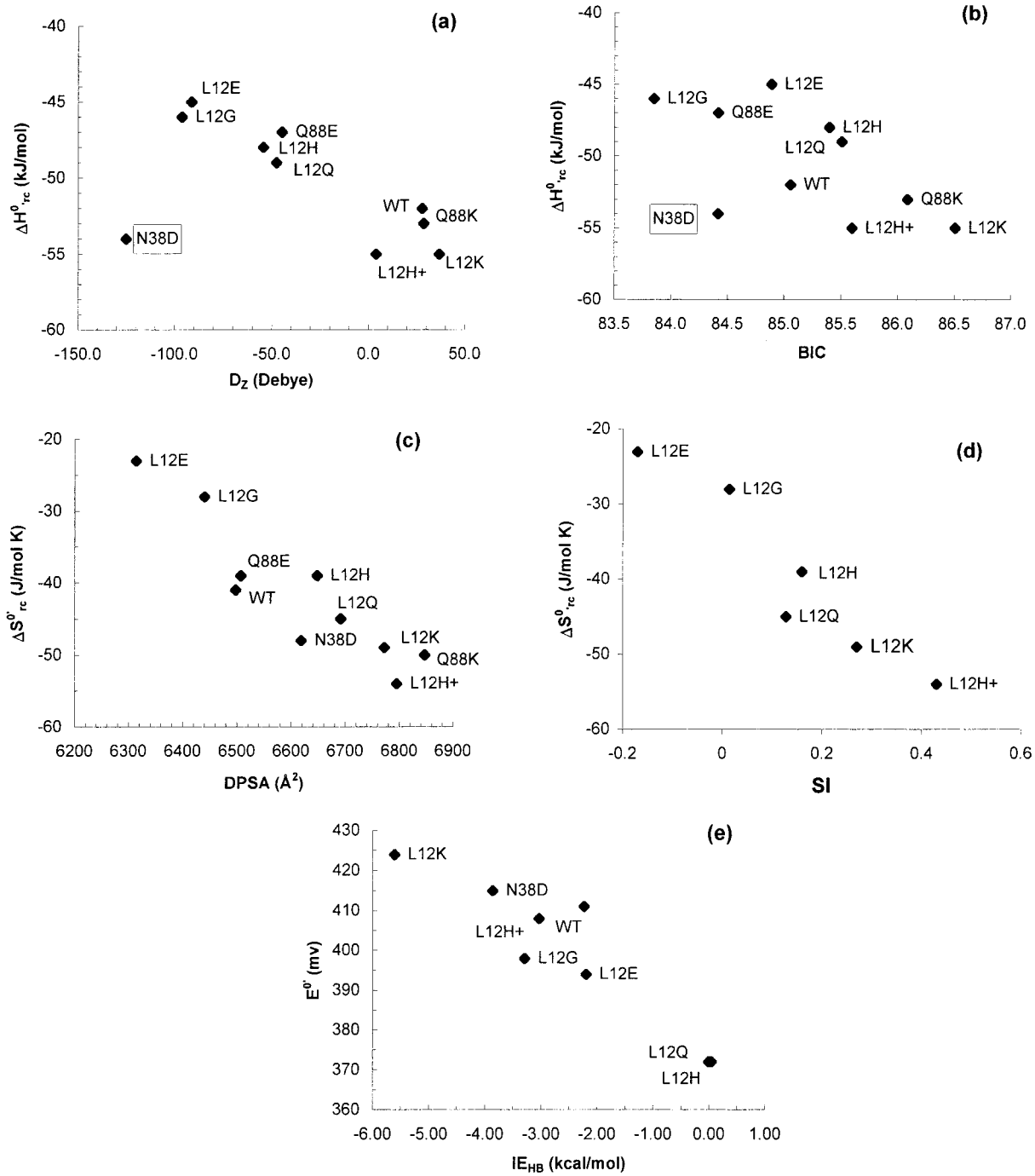

FIGURE 3: Correlations between protein structure descriptors and experimental thermodynamic parameters. Outliers are boxed.

reduced state, leading to more negative $\Delta H^{\circ\prime}_{rc}$ values. Again the variant N38D is an outlier.

The variation of the reduction entropy turns out to correlate with charged partial surface area (CPSA) descriptors (*37*), which model polar intermolecular interactions. In particular, the difference in partial positive and negative surface area of the proteins (DPSA), a descriptor that might be related to the protein capability to establish intermolecular hydrogen bonding interactions, yields a good linear relationship with $\Delta S^{\circ\prime}_{rc}$ (Figure 3c):

$$\Delta S^{\circ\prime}_{rc} = -0.053\text{DPSA} - 309.18,\ n = 10,\ R^2 = 0.853,\ s^2 = 18.14,\ F = 40.77 \quad (3)$$

An inspection of the role of each component of this descriptor reveals that an increase in the PNSA in the northern part of the protein is responsible for less negative $\Delta S^{\circ\prime}_{rc}$ data values.

The analysis of the similarity index data values (SI) shows that, in general, dissimilarity with respect to the wild type, which has been chosen as reference compound, leads to less negative $\Delta S^{\circ\prime}_{rc}$ values, the most dissimilar mutant being L12E. This descriptor, computed on a small portion of the protein around the copper site, explains 90% of the variation in the entropic contribution to the reduction potential of the complete L12 series of mutants (Figure 3d).

$$\Delta S^{\circ\prime}_{rc} = -55.87\text{SI} - 31.92,\ n = 5,\ R^2 = 0.906,\ s^2 = 16.22,\ F = 38.74 \quad (4)$$

Moreover, this index is very useful in numerically synthesiz-

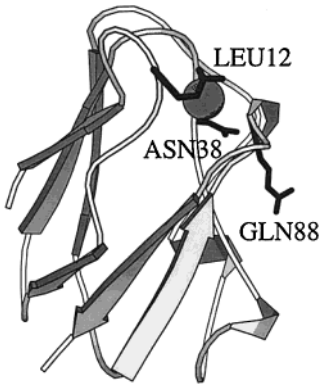

FIGURE 4: Structure of spinach plastocyanin (*26*) showing the residues subjected to mutation in this work (Leu12, Gln88, and Asn38).

ing the great amount of information obtained by the visual inspection/comparison of the molecular properties, such as electrostatic potentials and shapes.

Finally, a good correlation is obtained between the reduction potential and the contribution of the hydrogen bonding energy (IE_{HB}) of the residue in position 12 to the total energy of the protein (Figure 3e). This residue, being the mutated amino acid closest to the Cu site, has been chosen as a probe to test the influence of metal site rigidity conferred by hydrogen bonding networks on the reduction potential. The mutant N38D has also been included in the correlation, because, as already pointed out, Asn38 plays an important role in the copper site architecture, and the consequences of its mutation are reflected also in L12 behavior. On the contrary, Q88 mutants are located too far away to be effective.

$$E^{\circ\prime} = -9.44\mathrm{IE}_{\mathrm{HB}} + 375.41,\ n = 8,\ R^2 = 0.864,\ s^2 = 59.26,\ F = 37.69 \quad (5)$$

DISCUSSION

The spatial position of Leu12, Gln88, and Asn38, which are the residues subjected to mutation in this study, is shown in Figure 4. The selection of these residues was dictated by their crucial location in the protein structure. Leu12 belongs to the 'northern' hydrophobic patch of spinach plastocyanin, facing the Cu ligand His87 which is suggested to be important in the electron-transfer reaction of plastocyanin with the redox partners (*14*), and is located exactly above the metal site. Mutations of this residue resulted in significant changes in the kinetics of the reaction with photosystem I, as well as in the reduction potential and spectroscopic properties (*40*). In particular, substitution of Leu12 with alanine and asparagine was found to induce a change in the chemical shift of the Cu ligand His37 in the reduced protein, alterations of the EPR parameters of the cupric form, and a decrease in $E^{\circ\prime}$ as compared to the wild-type species, which were interpreted as indicative of a slight distortion of the coordination geometry and modification of the electronic structure of the Cu site (*14*). Moreover, L12E and L12K mutations have been shown to affect the rate constants for the photoinduced electron-transfer reaction between cytochrome *c* and plastocyanin (*41*).

The side chain of Gln88 lies in the acidic patch of the eastern site of the protein, close to Tyr83, a residue which has been suggested to be involved in the ET reaction with cytochrome *f* (*15*, *42*). It has been reported that the Q88E mutation causes a 20% increase in the overall second-order rate constant for reduction by cytochrome *f* (*43*), while both the Q88N (*44*) and Q88K (*45*) mutations influence the plastocyanin−PSI interaction.

Insertion of charged side chains in the solvent-exposed positions 12 and 88 should not alter appreciably protein folding and the coordination features of the copper center. Consistently, no appreciable mutation-induced changes in the electronic spectra were detected, which indicates that these mutations leave the geometric and electronic properties of the Cu site essentially unchanged, as noted previously (*14*). Different considerations dictated the choice of residue Asn38 as a target for point mutation. Asn38 is a residue buried in the protein environment and rigorously conserved among the plastocyanin family: mutations to Gln, Thr, and Leu showed its critical importance in establishing an extended hydrogen bonding network adjacent to the copper, which helps to preserve the integrity of the type 1 copper site (*16*).

The Enthalpy Change. Simple electrostatic considerations predict that insertion of a negative charge in the surroundings of the metal would selectively stabilize the more positively charged oxidized state, inducing an increase in reduction enthalpy (less negative $\Delta H^{\circ\prime}_{rc}$ values) and a consequent decrease in reduction potential. Opposite arguments hold for insertion of a positive charge. Moreover, the magnitude of the enthalpic effect is expected to be inversely proportional to the distance of the charge from the metal. Experimental mutation-induced enthalpy changes, interpreted in light of the 3D structure of the protein variants, invariably bear out these predictions (Tables 2 and 3) (the charges of Glu and Lys in position 12 turn out to be closer to the metal as compared to position 88).

Appropriate molecular descriptors computed on the minimized average structures allowed the establishment of satisfactorily quantitative models for the reduction thermodynamics parameters with good interpretative and predictive power. Notably, 86% of the variation in the reduction enthalpy is explained by the mutation-induced change in the component of the total dipole moment of the molecule along the *z* axis (D*z*) (Figure 3a, eq 1), due to the shift of the center of gravity of the negative and positive protein charge with respect to the metal center, in accord with electrostatic expectations. In this context, it is noteworthy that an important determinant for the stabilization of the reduced state due to the insertion of positively charged residues seems to be the increased capability of these mutants to establish dispersion interactions with the rest of the protein and, therefore, to delocalize the positive charge introduced, as expressed by the bonding information content (BIC) descriptor (*36*) (Figure 3b, eq 2).

The variant N38D is an outlier in both of the above correlations. In fact, the selective stabilization of the oxidized state expected from both descriptors is not detected experimentally. This may be explained by considering the strategic position of Asn38 in the protein structure. The functional role of this largely conserved residue seems to be connected to its participation to an extended hydrogen bonding network with residues Cys84 and Ser85, which electronically and conformationally stabilizes the Cu site (*16*). Accordingly, the Cys→Cu charge-transfer band in N38D is blue-shifted

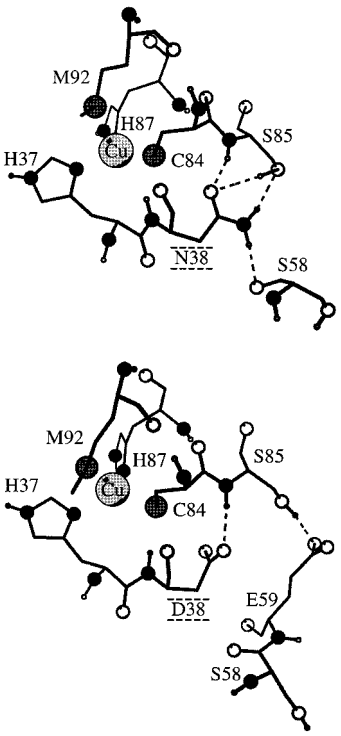

FIGURE 5: Hydrogen bonding network established by Asn38 (top) and Asp38 (bottom). Nitrogen atoms are represented by black circles, sulfur atoms by gray circles, and oxygens by white circles. Only polar hydrogens are shown. The hydrogen bonding interaction energy (IE_{HB}) between Asn28 and Ser85, as computed by the CHARMM program, is −8.34 kcal/mol; moreover, an additional hydrogen bond is achieved with Ser58 ($IE_{HB} = -2.92$ kcal/mol). In the Asp38 mutant, only one hydrogen bond with Ser85 is conserved ($IE_{HB} = -3.61$ kcal/mol).

by approximately 10 nm as compared to the *wt* species, and the mutant tends to lose the metal with time. The molecular dynamics calculations confirm this hypothesis. In fact, as shown by Figure 5, the hydrogen bonding network of the native protein undergoes significant modifications upon N38D mutation, and, on average, the hydrogen bonding intramolecular interaction energy of the Asn38 residue ($IE_{HB} = -11.26$ kcal/mol) is remarkably higher than that of an aspartate in the same position ($IE_{HB} = -3.86$ kcal/mol).

The Entropy Change. The entropic effects due to the mutational insertion of charged residues determine $E^{\circ\prime}$ changes that are invariably opposite to those induced by the concomitant enthalpic effects (Table 2). These enthalpy−entropy compensation phenomena are most likely the result of mutation-induced changes in the solvation properties of the molecule. Indeed, insertion/suppression of a net charge on the protein surface might induce a more or less localized modification of the H-bonding network in the hydration shell of the protein. A number of models for the hydration of nonpolar solutes and biopolymers recognize that these solvent reorganization effects induce largely compensating enthalpy and entropy changes (*46*−*56*). In the case of the Q88K mutant, the entropic effects prevail on the enthalpic effects, thus making the resulting $E^{\circ\prime}$ change opposite to that expected on electrostatic grounds (Table 2). Therefore, it is apparent that water plays a dual role in attenuating the effect of introduction/suppression of a surface net charge on the reduction potential of a metalloprotein: one is related to the high dielectric constant that weakens the Coulombic interaction between the charge and the metal, and the other is due to enthalpy−entropy compensation phenomena arising from solvent reorganization effects. This observation may provide further explanation for the inconsistencies noted sometime between the observed variations of $E^{\circ\prime}$ due to addition/suppression of surface charges and those predicted on the basis of simple Coulombic arguments (*12*).

The QSPR model that highlights the correlation between the variation in the reduction entropy and the charged partial surface area descriptors (*37*) (Figure 3c, eq 3) nicely supports this view. In fact, it appears that an extended negative surface area localized on the northern part of the protein, able to accept hydrogen bonding, turns out to favor entropically the reduced state. A partial positive surface area with hydrogen bonding donor propensity instead favors the oxidized state. Both these effects are opposite to the expected electrostatic (enthalpic) effects of the same charges on the $E^{\circ\prime}$ of the metal center.

The variation in $\Delta S^{\circ\prime}_{rc}$ along the series of mutants is well depicted by considering simultaneously the descriptors used. In fact, mutation of the wild-type protein with negatively charged or neutral residues induces modifications of the electrostatic potential distribution around the copper site (Figure 3d, eq 4) that enforces the dipolar characteristics of the protein variants. Moreover, the positive end of the dipole moves toward the main axis of the oblong molecule (*z*) away from the copper site (Figure 3a, eq 1). This determines the alignment of the water molecule dipole moment approximately head to tail on the hydrogen bonding acceptor surface area in proximity of the copper site. The opposite holds for mutation of the wild-type protein with positively charged residues. In this case, the modifications induced in the electrostatic potential distribution around the copper site (SI) produce a shift of the dipole moment in the opposite direction, and therefore a reorganization of the water molecules on the protein surface near the copper site is required, probably influenced also by the hydrophobicity of the environment (*13*) (Figure 3b, eq 2). Therefore, these findings support the hypothesis (*6*) that the electrostatic field experienced by the solvent around the protein is an important determinant for the reduction potential in metalloproteins.

The importance of the capability of the mutated residue to establish a strong hydrogen bonding network in close proximity to the copper site and its consequence on the $E^{\circ\prime}$ variation is shown in Figure 3e (eq 5). This trend indicates that a rigidity of the copper site induced by an extended hydrogen bonding network established by residues in positions 12 and 38 favors the reduced species.

Finally, a comparative analysis of the dynamic behavior of wild-type plastocyanin and its mutants is particularly informative for some aspects. For example, regarding solvation effects, mutation of Leu12 to Gly was thought to increase the accessibility of the copper site to the solvent. In this case, the enthalpic and entropic stabilization of the oxidized and reduced state, respectively, of the L12G mutant as compared to wild-type pc (Tables 1 and 2) could be justified by the increased interaction with solvent dipoles (*13*). However, molecular dynamics simulations show that this mutation increases the conformational flexibility of the loop formed by residues 7−12, which, on average, results in a shift of this polypeptide segment toward the copper site. Therefore, the solvent-accessible surface of His87 (the only

Cu ligand that protrudes to the protein surface) remarkably decreases in the minimized average structure of the L12G mutant (18 Å^2) as compared to *wt* pc (38 Å^2). However, we note that the hydrogen bonding acceptor surface area of the northern half of the L12G protein turns out to be more extended with respect to the same region of wild-type pc. Therefore, the major contribution to the stabilization of the oxidized state and to the increase in reduction entropy in this mutant appears to be related to the change in the electrostatic field around the copper site due to the conformational change and to the effect of bound and mobile water molecules lying in an extended region around the copper site, respectively, instead of water penetration into the copper cavity.

ACKNOWLEDGMENT

M.C.M. and F.D.R. are especially grateful to CICAIA (Centro Interdipartimentale Calcolo Automatico e Informatica Applicata) for the InsightII program.

REFERENCES

1. Zhou, H.-X. (1997) *JBIC, J. Biol. Inorg. Chem. 2*, 109−113.
2. Bertini, I., Gori-Savellini, G., and Luchinat, C. (1997) *JBIC, J. Biol. Inorg. Chem. 2*, 114−118.
3. Mauk, A. G., and Moore, G. R. (1997) *JBIC, J. Biol. Inorg. Chem. 2*, 119−125.
4. Gunner, M. R., Alexov, E., Torres, E., and Lipovaca, S. (1997) *JBIC, J. Biol. Inorg. Chem. 2*, 126−134.
5. Armstrong, F. A. (1997) *JBIC, J. Biol. Inorg. Chem. 2*, 139−142.
6. Warshel, A., Papazyan, A., and Muegge, I. (1997) *JBIC, J. Biol. Inorg. Chem. 2*, 143−152.
7. Battistuzzi, G., Borsari, M., Loschi, L., and Sola, M. (1997) *JBIC, J. Biol. Inorg. Chem. 2*, 350−359.
8. Battistuzzi, G., Loschi, L., Borsari, M., and Sola, M. (1999) *JBIC, J. Biol. Inorg. Chem. 4*, 601−607.
9. Capozzi, S., Ciurli, S. L., and Luchinat, C. (1998) *Struct. Bonding 90*, 127−160.
10. Sykes, A. G. (1991) *Adv. Inorg. Chem. 36*, 377−408.
11. Battistuzzi, G., Borsari, M., Loschi, L., Martinelli, A., and Sola, M. (1999) *Biochemistry 38*, 7900−7907.
12. Zeng, Q., Smith, E. T., Kurtz, D. M., and Scott, R. A. (1996) *Inorg. Chim. Acta 242*, 245−251.
13. Battistuzzi, G., Borsari, M., Loschi, L., Righi, F., and Sola, M. (1999) *J. Am. Chem. Soc. 121*, 501−506.
14. Sigfridsson, K., Young, S., and Hansson, O. (1996) *Biochemistry 35*, 1249−1257.
15. Ubbink, M., Ejdeback, M., Karlsson, B. G., and Bendall D. S. (1998) *Structure 6*, 323−335.
16. Dong, S., Ybe, J. A., Hecht, M. H., and Spiro, T. G. (1999) *Biochemistry 38*, 3379−3385.
17. Karelson, M., Lobanov, V. S., and Katritzky, A. R. (1996) *Chem. Rev. 96*, 1027−1043.
18. Sandberg, M., Eriksson, L., Jonsson, J., Sjostrom, M., and Wold, S. (1998) *J. Med. Chem. 41*, 2481−2491.
19. Nordling, M., Olausson, T., and Lundberg L. G. (1990) *FEBS Lett. 276*, 98−102.
20. Nusbaumer, C. (1996) Ph.D. Thesis, Université de Neuchâtel, Switzerland.
21. Kuwana, T. (1977) in *Electrochemical studies of biological systems* (Sawyer, D. T., Ed.) *ACS Symp. Ser. No. 38*, American Chemical Society, Washington, DC.
22. Yee, E. L., Cave, R. J., Guyer, K. L., Tyma, P. D., and Weaver, M. J. (1979) *J. Am. Chem. Soc. 101*, 1131−1137.
23. Yee, E. L., and Weaver, M. J. (1980) *Inorg. Chem. 19*, 1077−1079.
24. Taniguchi, V. T., Sailasuta-Scott, N., Anson, F. C., and Gray, H. B. (1980) *Pure Appl. Chem. 52*, 2275−2281.
25. Koller, K. B., and Hawkridge, F. M. (1985) *J. Am. Chem. Soc. 107*, 7412−7417.
26. Xue, Y., Okvist, M., and Young, S. (1998) *Protein Sci. 7*, 2099−2105.
27. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
28. van Gunsteren, W. F., and Berendsen, J. C. (1977) *Mol. Phys. 34*, 1311−1327.
29. Madura, J. D., Briggs, J. M., Wade, R. C., Davis, R., Luty, B. A., Ilin, A., Antonsiewicz, J., Bagheri, M. K., Scott, B., and McCammon, J. A. (1995) *Comput. Phys. Commun. 91*, 57−95.
30. Jorgensen, W. L., and Tirado-Rives, J. (1988) *J. Am. Chem. Soc. 110*, 1657−1666.
31. De Rienzo, F., Gabdoulline, R. R., Menziani, M. C., and Wade, R. C. (2000) *Protein Sci. 9*, 1439−1454.
32. Libeu, C. A. P., Kukimoto, M., Nishiyama, M., Horinouchi, S., and Adman, E. T. (1997) *Biochemistry 36*, 13160−13179.
33. Davis, M. E., and McCammon, J. A. (1991) *J. Comput. Chem. 12*, 909−912.
34. Hodgkin, E. E., and Richards, W. G. (1987) *Int. J. Quantum Chem., Quantum Biol. Symp. No. 14*, 105−110.
35. Blomberg, N., Gabdoulline, R. R., Nilges, M., and Wade, R. C. (1999) *Proteins: Struct., Funct., Genet. 37*, 379−387.
36. Kier, L. B. (1980) *J. Pharm. Sci. 69*, 807−811, and references cited therein.
37. Stanton, D. T., and Jurs (1990) *Anal. Chem. 62*, 2323−2329.
38. Unpublished experiments from our laboratory.
39. McGinnis, J., Sinclair-Day, J. D., Sykes, A. G., Powls, R., Moore, J., and Wright, P. E. (1988) *Inorg. Chem. 27*, 2306−2312.
40. Modi, S., Nordling, M., Lundberg, L. G., Hansson, O., and Bendall, D. S. (1992) *Biochim. Biophys. Acta 1102*, 85−90.
41. Ivkovic-Jensen, M. M., Ullmann, G. M., Young, S., Hansson, O., Crnogorac, M. M., Ejdeback, M., and Kostic, N. M. (1998) *Biochemistry 37*, 9557−9569.
42. Illeraus, J., Altshmied, L., Reichert, J., Zak, E., Herrmann, R. G., and Haelhnel, W. (2000) *J. Biol. Chem. 275*, 17590−17595.
43. Kannt, A., Young, S., and Bendall, D. S. (1996) *Biochim. Biophys. Acta 1277*, 115−126.
44. Sigfridsson, K., Young, S., and Hansson, O. (1997) *Eur. J. Biochem. 245*, 805−812.
45. Olesen, K., Ejdeback, M., Crnogorac, M. M., Kostic, N. M., and Hansson, O. (1999) *Biochemistry 38*, 16695−16705.
46. Blokzijl, W., and Engberts, J. B. N. F. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 1545−1579.
47. Lumry, R., and Rajender, S. (1970) *Biopolymers 9*, 1125−1227.
48. Ben-Naim, A. (1975) *Biopolymers 14*, 1337−1355.
49. Lumry, R., Battistel, E., and Jolicoeur, C. (1982) *Faraday Symp. Chem. Soc. 17*, 93−108.
50. Lee, B. (1985) in *An anatomy of hydrophobicity* (Eisenfeld, J., and Di Lisi, C., Eds) pp 3−11, Elsevier, North-Holland, Amsterdam.
51. Grunwald, E. (1986) *J. Am. Chem. Soc. 108*, 5726−5731.
52. Yu, H.-A., and Karplus, M. (1988) *J. Chem. Phys. 89*, 2366−2379.
53. Matubayasi, N., Reed, L. H., and Levy, R. M. (1994) *J. Phys. Chem. 98*, 10640−10649.
54. Lee, B., and Graziano, G. (1996) *J. Am. Chem. Soc. 118*, 5163−5168.
55. Gallicchio, E., Mogami Kubo, M., and Levy, R. M. (1998) *J. Am. Chem. Soc. 120*, 4526−4527.
56. Canters, G. W., Kolczak, U., Armstrong, F. A., Jeuken, L. J. C., Camba, R., and Sola, M. (2000) *Faraday Discuss. No. 116*, 205−220.

BI002565D