Available online at www.sciencedirect.com
SCIENCE DIRECT
Computational Biology and Chemistry 29 (2005) 345–353
Computational Biology and Chemistry
www.elsevier.com/locate/compbiolchem

Folding degrees of azurins and pseudoazurins Implications for structure and function

Ernesto Estrada [a,*], Eugenio Uriarte [b]

[a] *Complex Systems Research Group, X-Ray Unit, Edificio CACTUS, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain*
[b] *Department of Organic Chemistry, Faculty of Pharmacy, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain*

Accepted 23 June 2005

Abstract

A quantitative measure of the degree of folding of azurins and pseudoazurins has been made. We have found that the reduction potential of azurins and pseudoazurins is a function of the contribution to the degree of folding of His117, a key amino acid in electron transfer which is directly bonded to copper in these proteins. The folding degree of His117 explains 95% of the variance in the experimental values of the reduction potential of azurins and pseudoazurins. The change in the folding degree of this amino acid influences several geometric parameters of the main backbones of these proteins. Among them, the angle formed between N(His117)···Cu···S(Cys112), which plays an important role in electron transport, but not the N(His117)···Cu distance, shows some non-linear correlation with the reduction potential of azurins and pseudoazurins. However, it is only able to explain less than 75% in the variance of the reduction potential of these proteins instead of the 95% explained by the folding degree of His117.

Keywords: Folding degree; Blue copper proteins; Structure–function relation; Copper binding site; Protein geometry

1. Introduction

One of the principal goals of the proteome research is the study of protein structure and the understanding of the role it plays in the biochemistry, physiology and pathology of the cell (Domou and Broder, 2004). An understanding of these processes permits the successful search of disease markers as well as of possible targets for virtual drug screening. An important characteristic of the 3D structure of proteins is the degree of folding of the protein chain (Randić and Krilov, 1997, 1999; Estrada, 2004a,b). Because proteins fold to optimize the conformational preferences of amino acids subject to local and global constraints, it is expected that folding degree is related to the structure, evolution and function of proteins.

Families of proteins are comprised of members who share the same three-dimensional structure, perhaps as a consequence of evolution. One of these families is formed by proteins that have a structurally conserved 90-to 150-amino acid sequence module known as the "blue copper binding" domain (Nersissian and Shipp, 2002). The blue copper proteins (BCPs) are a subclass of the copper proteins that catalyze redox reactions in several types of organisms (Lappin, 1991), in which the copper changes from a diamagnetic, Cu(I), to a paramagnetic, Cu(II), oxidation state. Cu(II) is strongly coordinated to the sulfur atom of a cysteine producing an anomalously small parallel hyperfine coupling constant and a very strong absorption band at ca. 600 nm, giving the typical blue color to these proteins (Solomon et al., 1992). This coordination site, the so-called type-1, is completed with two imidazole nitrogens (histidines) in equatorial positions and a sulfur thioether (methionine) in an axial position (Katz et al., 2003). The type-1 site does not show the familiar tetrahedral or square planar/octahedral structures of small Cu(I) and Cu(II) coordination compounds (Broman et al., 1962; Shepard et al., 1990). In fact, it has been described as a strained structure intermediate between those of Cu(I) and Cu(II). Several authors have assigned to this distinctive

* Corresponding author. Tel.: +34 981 563100x16262; fax: +34 981 547077.
E-mail address: estrada66@yahoo.com (E. Estrada).

1476-9271/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiolchem.2005.06.005

structure the significance of being ideal for the shuttling of electrons (Vallee and Williams, 1968; Williams, 1971). It has also been argued that this structure is responsible for the high redox potential of these proteins compared to the Cu(I)/Cu(II) couple in water (James and Williams, 1961; Gray et al., 2000). However, recent calculations fail to support the suggestion that strain plays a significant role in the function of these proteins (Ryde et al., 2000).

The 3D structures of a large number of BCPs, such as azurins, pseudoazurins, stellacyanin, plastocyanins and ceruloplasmin, have been solved (Katz et al., 2003). Azurins are the most prominent members of the BCP family. They show the unique structural feature of a second axial group, the carbonyl oxygen of a glycine. Although the biological redox partner of azurins remains unknown, it is known that they are involved in anaerobic nitrite respiration and have been shown to donate electrons to nitrite reductase, a function thought to involve another BCP, pseudoazurin (De Rienzo et al., 2000). Azurins have been found to induce apoptotic death in macrophages. They form a complex with the tumor suppressor protein p53 stabilizing it and raising its intracellular level (Punj et al., 2003). On the other hand, pseudoazurins are components of the respiratory electron transfer chain that sequentially reduces NO_3^-/NO_2^- to molecular nitrogen and currently, pseudoazurins are the only known in vivo reaction partners of nitrite reductase (Libeu et al., 1997).

Here we study the role played by the folding of azurins and pseudoazurins on their structure and function. The degree of folding quantitatively measures how folded a protein backbone is. We have studied the global and local folding degree for a series of azurins and pseudoazurins and have found that their reduction potential is a function of the degree of local folding of one of the amino acids at the binding site. The explanations and possible implications of these findings are also discussed.

1.1. Degree of protein folding: global and local

The protein folding degree index I_3 transforms the qualitative criterion of folding to a quantitative scale, in which two proteins can be differentiated by measuring the degree of folding of their backbone chains (Estrada, 2002, 2004a). The protein folding degree index is based on the torsion angles of the protein backbone chain, i.e., the so-called ϕ, ψ and ω torsion angles. These are all equal to 180° for a fully extended polypeptide chain, which corresponds to the least folded structure that a chain can hypothetically adopt. The angle ω_i defines the rotation about the C_i–N_{i+1} peptide bond. ϕ_i describes the rotation about N_i–$C\alpha_i$ bond and ψ_i describes the rotation about the $C\alpha_i$–C_i bond. Fig. 1 shows a portion of a protein backbone indicating the these torsion angles. The folding degree index is defined from a graph whose nodes represent ϕ, ψ and ω torsion angles and two nodes are linked if, and only if, the corresponding angles are contiguous in the backbone chain of the protein as indicated in Fig. 1.

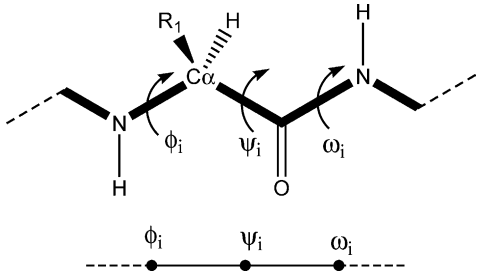


Fig. 1. Illustration of a protein backbone section representing the ϕ, ψ and ω torsion angles and the projections of these angles as the nodes of a chain graph used for defining the folding degree index (see text for explanations).

Let **A** be the adjacency matrix of this chain representing the adjacency of torsion angles in the protein backbone and let **T** be a diagonal matrix of the cosines of ϕ_i, ψ_i and ω_i angles (Estrada, 2002, 2004a). Define $\mathbf{B} = \mathbf{A} + \mathbf{T}$ as a matrix representing the protein backbone, which has the spectrum $\sigma = (\lambda_1, \lambda_2, \ldots, \lambda_t)$, where N is the number of atoms in the protein backbone and $t = N - 3$ is the number of ϕ, ψ and ω torsion angles. Then, the folding degree index I_3 is defined as (Estrada, 2000, 2002, 2004a):

$$I_3 = \frac{1}{N-3}\sum_{j=1}^{N-3} e^{\lambda_j} \tag{1}$$

The index I_3 represents a global characterization of protein folding degree through the sum of contributions coming from individual torsion angles, pairs, triples, quadruples, etc. of contiguous torsion angles, in a way in which larger sequences of contiguous torsion angles receive lower weights than shorter ones. This interpretation follows directly from the fact that I_3 can be expressed as the infinite sum of spectral moments μ_k of **B** divided by $k!$, where the spectral moment μ_k represents the sum of all closed walks of length k in the chain (Estrada, 2000, 2002, 2004a):

$$I_3 = \frac{1}{N-3}\sum_{k=0}^{\infty} \frac{\mu_k}{k!} \tag{2}$$

Each closed walk is associated with a sequence of torsion angles in the chain. For instance, closed walks of length zero represent the nodes of the chain, and corresponds to the single torsion angles, closed walks of length two stand for the edges of the chain, and symbolizes the pairs of adjacent torsion angles in the chain, etc.

It has been previously shown that Eq. (2) can be expressed simply as a sum of the local spectral moments of **B** (Estrada, 2004b):

$$I_3 = \frac{1}{N-3}\sum_{k=0}^{\infty}\sum_{t=1}^{N-3} \frac{\mu_k(t)}{k!} \tag{3}$$

where the local spectral moment $\mu_k(t)$ consists simply of the diagonal entries of the kth power of **B** corresponding to the torsion angle t (ϕ, ψ or ω) in the chain. Let i denote an amino acid of the protein with torsion angles ϕ_i, ψ_i and ω_i. Then, the

contribution of this amino acid to the global protein degree of folding has been defined as the infinite sum of the local spectral moments corresponding to the ϕ_i, ω_i torsion angles (note that ω_i corresponds to the peptide bond which is shared by amino acids i and $i+1$) (Estrada, 2004b):

$$I_3(i) = \sum_{k=0}^{\infty} \frac{\mu_k(\phi_i) + \mu_k(\psi_i)}{k!} \quad (4)$$

Using graph spectral theory we prove that the local contribution of the ith amino acid to the global protein folding can be expressed as follows (see Appendix A):

$$I_3(i) = \sum_{j=1}^{N} e^{\lambda_j} \{[\nu_j(\Psi_i)]^2 + [\nu_j(\phi_i)]^2\} \quad (5)$$

where $\nu_1, \nu_2, \ldots, \nu_N$ is an orthonormal basis of R^N composed by eigenvectors of **B** associated to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$. $\nu_j(\psi_i)$ and $\nu_j(\phi_i)$ denotes the component of ν_j corresponding to the torsion angles ψ and ϕ of the ith amino acid. Eq. (5) was given by Estrada (2004b) in a different formulation but expressing exactly the same. Here we provide a more condensed formula for these contributions using some recent results (Estrada and Rodríguez-Velázquez, 2005).

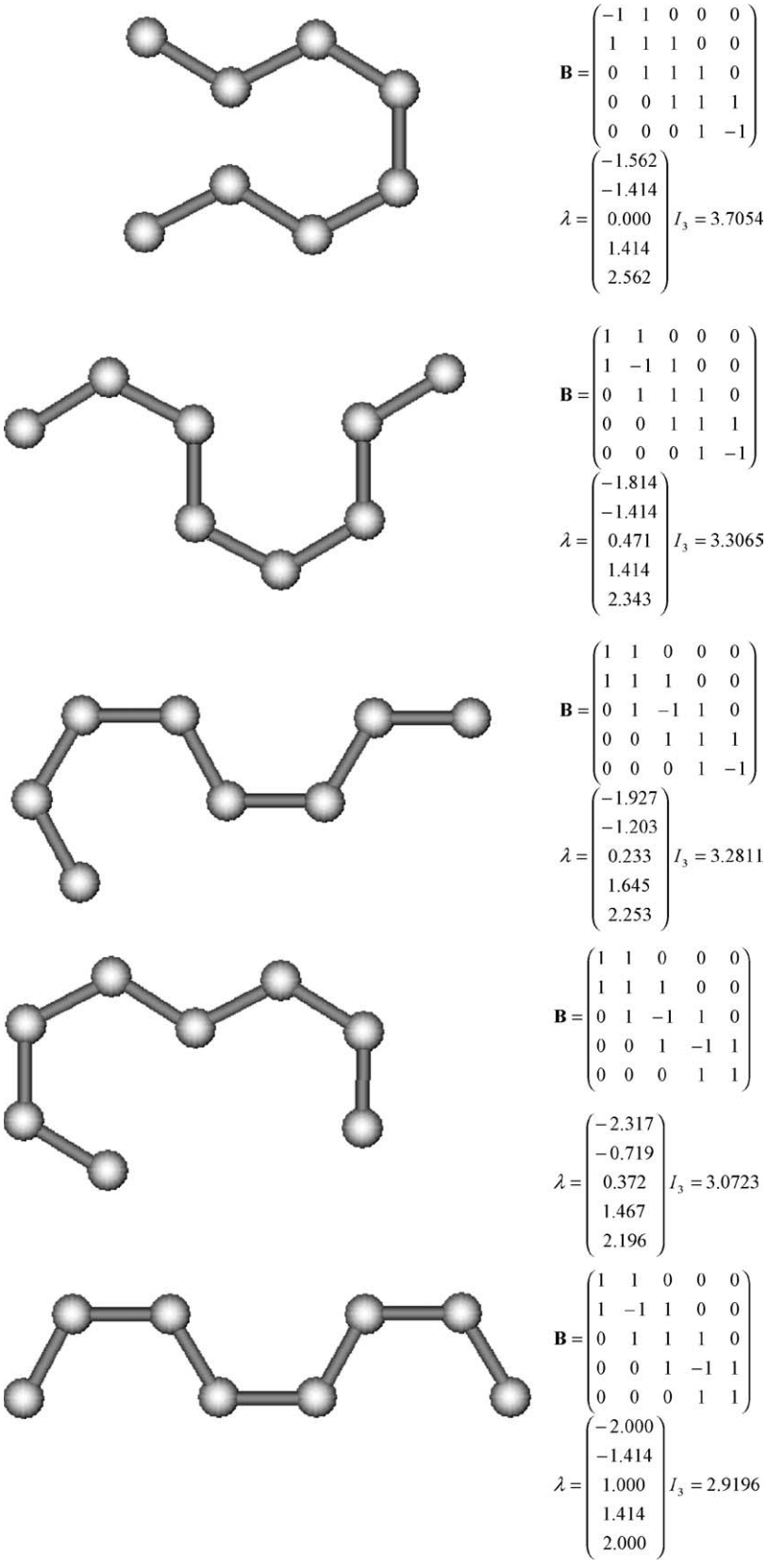
We have previously shown that I_3 and $I_3(i)$ describe very well the folding degree of protein chains by studying more than 150 proteins of different sizes and types. We have also analysed the influence of protein size, crystal packing, temperature, pH and secondary structure on the values of global and local folding degree of proteins as well as the applications of these indices to the classification of protein domains, protein similarity and structure–function relationships (Estrada, 2000, 2002, 2003, 2004a,b). However, we still explain here some of the most important advantages of using I_3 instead of other descriptors to account for the folding degree of a (protein) chain. The first argument against the use of this folding degree index comes from the application of the Occam razor or principle of parsimony which admonishes us to choose from a set of otherwise equivalent models of a given phenomenon the simplest one. Then, the question is Why to use the more complicated I_3 index instead of the simplest sum of dihedral angles of the backbone chain? We answer this question by using Fig. 2 where we illustrate several chains with different folding but having the same sum of dihedral angles. These chains have only eight vertices, which produce these five different conformations, but as the length of the chain increases, the number of such conformations with the same sum of dihedral angles and different folding increases dramatically. Thus, the conclusion is simple and straightforward: the simple sum of dihedral angles (or a function of them, such as cosines) is not enough to account for the degree of folding of a chain. It is also important to consider the distribution of such folded regions in the chain. In Fig. 2, we can see that when these folded regions are more to the centre of the chain the larger the folding degree index is. The second question that can arise in the context of protein folding degree is immediately answered by this same example, i.e., why not to use the simplest percentage of helix and strands instead of the more complicated I_3 index? As a simple example we will consider the five chains represented in Fig. 2. Let call extended (E) the region formed by four vertices in these chains which form a dihedral angle of 180° and twisted (T) similar regions forming a dihedral angle of 0°. Using these designations, we will see that the "secondary" structure of the first chain is represented by the sequence ETTTE. The other chains have sequences: TETTE, TTETE, TTEET and TETET, respectively. This means that all these chains have 40% of E and 60% of T conformations. The only thing changing is the distribution of these regions along the chain. This situation is also typical in protein chains. For instance, the following three proteins have the same percentage of helix (38%) and strand (15%) but they have different values of the folding degree index (in parenthesis): 1OYB (3.0629); 1APS (3.1012) and 1TCA (3.1226). Other examples are, for instance, 2HBG and 1CPC which have 78% of helix and 0% of strand and I_3 indices of 3.3718 and 3.3381, respectively as well as 1LEN and 2BBK with 3% of helix and 50% of strand and I_3 values of 2.6750 and 2.7165, respectively.

1.2. Proteins dataset

The dataset of azurins and pseudoazurins studied in this work is comprised of the following 15 proteins, for which PDB codes and crystallographic resolution (in parentheses) are also given (see Ryde et al., 2000 for coordination of Cu(II) and references for crystallographic data): *A. xylosoxidans* azurin I (pH 8.0), 1RKR (2.45 Å); *P. putida* azurin (pH ∼ 7), 1NWP (1.60 Å); *A. denitrificans* (M121H) azurin (pH 3.5), 1A4C (2.45 Å); *A. denitrificans* (M121Q) azurin, 1URI (1.94 Å); *P. aeruginosa* azurin (pH 5.5), 4AZU (1.9 Å); *P. aeruginosa* (M121A) azurin (pH 5.1), 2TSA (2.2 Å); *P. fluorescens* azurin, 1JOI (2.05 Å); *A. xylosoxidans* azurin II (pH 6.5), 1DYZ (1.75 Å); *P. aeruginosa* (M121E) azurin (pH 6.0), 1ETJ (2.3 Å); *A. denitrificans* azurin (pH 5.0), 2AZA (1.8 Å); *P. aeruginosa* azurin (pH 9.0), 5AZU (1.9 Å); *A. faecalis* pseudoazurin (pH 6.8), 1PAZ (1.55 Å); *A. faecalis* pseudoazurin (pH 7), 8PAZ (1.6 Å); *M. extorquens* pseudoazurin, 1PMY (1.5 Å); *A. cycloclastes* pseudoazurin, 1ZIA (1.54 Å).

2. Results and discussion

The values of the degree of global folding index for the 15 proteins studied are given in Table 1. The average folding degree for these proteins is 2.7675, a low value compared to that found in other protein families. For instance, the average value of I_3 for 152 proteins studied by Estrada (2004a) is 2.9767. The average value of I_3 for azurins and pseudoazurins agrees with the fact that blue copper proteins form mainly-β sandwich structures. Azurins show a slightly higher values of folding compared to pseudoazurins. This order of



$$\mathbf{B} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\lambda = \begin{pmatrix} -1.562 \\ -1.414 \\ 0.000 \\ 1.414 \\ 2.562 \end{pmatrix} \quad I_3 = 3.7054$$

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\lambda = \begin{pmatrix} -1.814 \\ -1.414 \\ 0.471 \\ 1.414 \\ 2.343 \end{pmatrix} \quad I_3 = 3.3065$$

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\lambda = \begin{pmatrix} -1.927 \\ -1.203 \\ 0.233 \\ 1.645 \\ 2.253 \end{pmatrix} \quad I_3 = 3.2811$$

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\lambda = \begin{pmatrix} -2.317 \\ -0.719 \\ 0.372 \\ 1.467 \\ 2.196 \end{pmatrix} \quad I_3 = 3.0723$$

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\lambda = \begin{pmatrix} -2.000 \\ -1.414 \\ 1.000 \\ 1.414 \\ 2.000 \end{pmatrix} \quad I_3 = 2.9196$$

Fig. 2. Representation of a chain of eight vertices with different folding conformations but having the same sum of dihedral angles. The matrices representing the adjacency of torsion angles in the protein backbone, their eigenvalues and the I_3 indices of each conformation are also given.

Table 1
Global folding degree indices (I_3) and other structural parameters of the azurins and pseudoazurins studied and their reduction potentials measured vs. a normal hydrogen electrode (NHE)

Protein	PDB	I_3	SDA	R_G	%Helix	%Strand	d_{N-Cu}[a]	N–Cu–S[b]	I_3 (H117)[c]	E^0
A. xylosoxidans azurin I (pH 8.0)	1RKR	2.7699	0.1600	13.53	21	35	1.906	124.89	8.0171	305
P. putida azurin (pH ∼ 7)	1NWP	2.7738	0.1615	13.52	19	36	1.944	119.53	7.8008	295
A. denitrificans (M121H) azurin (pH 3.5)	1A4C	2.7748	0.1611	13.48	18	36	1.948	122.30	8.5988	350
A. denitrificans (M121Q) azurin	1URI	2.7750	0.1620	13.54	17	36	2.064	116.97	7.5856	263
P. aeruginosa azurin (pH 5.5)	4AZU	2.7759	0.1628	13.52	16	34	2.111	123.62	8.2833	310
P. aeruginosa (M121A) azurin (pH 5.1)	2TSA	2.7777	0.1619	13.44	15	34	2.390	127.72	8.7991	373
P. fluorescens azurin	1JOI	2.7781	0.1636	13.64	21	34	2.013	125.25	7.5508	–
A. xylosoxidans azurin II (pH 6.5)	1DYZ	2.7807	0.1624	13.51	19	34	1.989	121.05	7.9916	305
P. aeruginosa (M121E) azurin (pH 6.0)	1ETJ	2.7819	0.1639	13.52	19	36	2.020	125.15	7.8623	220
A. denitrificans azurin (pH 5.0)	2AZA	2.7820	0.1640	13.54	16	36	2.008	121.64	7.5784	285
P. aeruginosa azurin (pH 9.0)	5AZU	2.7926	0.1676	13.50	17	34	2.056	122.74	7.9672	293
A. faecalis pseudoazurin (pH 6.8)	1PAZ	2.7300	0.1512	12.78	17	38	2.125	111.59	7.5416	269
A. faecalis pseudoazurin (pH 7)	8PAZ	2.7302	0.1511	12.86	17	37	2.013	113.50	7.5052	269
M. extorquens pseudoazurin	1PMY	2.7448	0.1550	12.98	14	36	1.968	110.08	7.3612	260
A. cycloclastes pseudoazurin	1ZIA	2.7458	0.1551	12.80	20	35	1.989	113.49	7.4820	260

Protein databank (PDB) codes are also given.
[a] Distance between N(His117)· · ·Cu.
[b] Angle between N(His117)· · ·Cu· · ·S(Cys112).
[c] Contribution of His117 to I_3 index.

folding degree follows the general trend of the percentages of secondary structures for these proteins (Hooft et al., 1996). For instance, azurins have about 19% helix and 34% strand on average, while pseudoazurins have 17% helix and 36.5% strand. The differences between I_3 index and the sum of dihedral angles as well as of the percentage of secondary structure of the proteins studied here is well illustrated in Fig. 3. Despite there is a general trend indicating a relationship between I_3 index and the sum of dihedral angles this relation does not exist for some particular series of proteins as those shown in Fig. 3. It is also remarkable that despite a linear correlation is observed between I_3 index and the percentage of secondary structure of several families of proteins this is not the case for specific families as the one studied here (see Fig. 3). These differences arises by the incomplete accounting of the folding degree by the sum of dihedral angles and the percentage of secondary structure, which consider the whole (backbone chain) as the sum of their parts (angles or amino acids) and do not consider the way in which such parts are distributed along the whole. This characteristic is typical of all complex system and it is simply stated as "the parts cannot contain the whole".

In order to clarify the roles of individual residues in the function of azurins and pseudoazurins we have studied the influence of amino acid contributions to the global folding on the values of the reduction potential of these proteins. First, the contribution of all residues to the global folding degree of the BCPs studied were calculated. We have been able to obtain a very good model relating both properties for these proteins, showing that the reduction potential of these proteins is explained to a large extent by the folding degree of the His117 (His81 in pseudoazurins) residue according to the following linear regression model, which has a correlation coefficient of 0.974 (see Fig. 4):

$$E^0 = 75.2(\pm 5.2) I_3(\text{His117}) - 297.8(\pm 41.3)$$

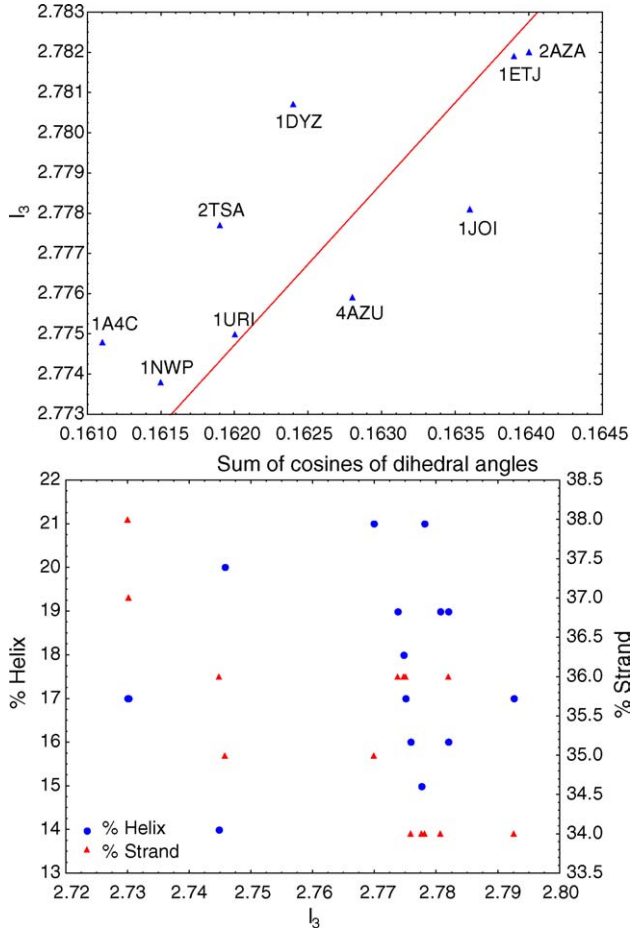


Fig. 3. Relationships between the folding degree index of azurins and pseudoazurins and the sum of cosines of ihedral angles of the main backbone as well as the percentage of secondary structure.

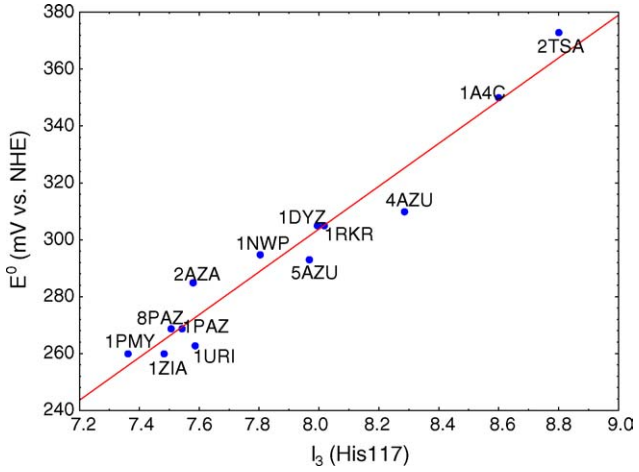


Fig. 4. Plot of the reduction potential (E^0 mV vs. NHE) as a function of the contribution of histidine 117 (azurins) or histidine 81 (pseudoazurins). The straight line corresponds to the linear regression model between both variables.

This histidine residue serves as a conduit between the Cu ion and the protein surface and is solvent accessible, which makes it an important element for the function of these proteins (see Fig. 5). The His46 (His40 in pseudoazurins) residue is further away from the protein surface and is not accessible to solvent (Pozdnyakova et al., 2001).

According to this model, an increase of the folding degree of His117 (His81 in pseudoazurins) translates into an increment of the reduction potential of azurins and pseudoazurins. As we have previously remarked, this residue is important for the function of BCPs. In particular, it has been demonstrated (Jeuken et al., 2000) that the imidazole ring of His117 provides an excellent pathway for the electron to transfer to external partners, and its loss destroys the electronic coupling between the copper atom and the redox partner. It has been previously observed that replacing His117 by a glycine results in little change in the copper site geometry (Jeuken et al., 2000). However, the differences found in the reduction potential of *wild type* azurin and His117Gly azurin are probably due to slight changes in the structure of the protein close to His117 (Jeuken et al., 2000). This observation coincides with our findings that a change in the folding degree of His117 has a significant effect on the reduction potential of azurins and pseudoazurins. This difference in the folding degree of His117 of azurins and pseudoazurins can be reflected in different structural parameters of both the binding site and the protein backbone in the neighborhood of this residue. We have investigated two of these structural parameters: the distance between copper and the nitrogen atom of His117 (His81 in pseudoazurins) and the angle formed between the sulfur of Cys112 (Cys78 in pseudoazurins), the copper atom and the nitrogen of His117 (His81 in pseudoazurins). Our supposition is guided by the proposed scheme of electron flow in BCPs, for which (Cys112)···Cu···His117···H_2O is an important path (see Fig. 5). In Fig. 6, we illustrate the enlargement of the N···Cu distance and widening of the S···Cu···N angle for 2TSA ($E^0 = 373$ mV) in comparison with the *A. denitrificans* azurin (PDB 1URI), which has a reduction potential of only 263 mV. We have found no relationship between the N···Cu distance and the

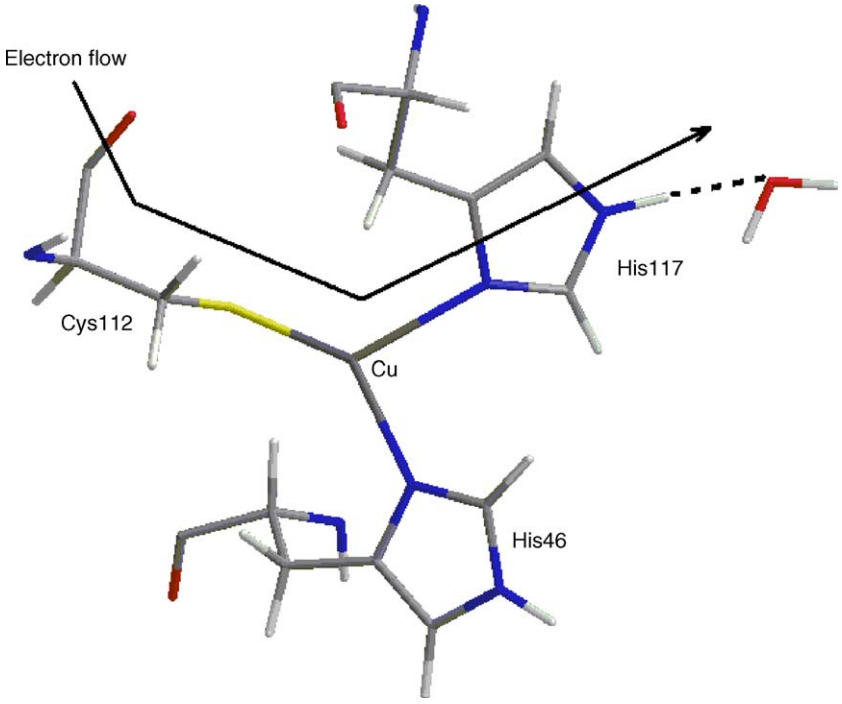


Fig. 5. Schematic diagram of the flow of electrons in blue copper proteins using as an example the crystallographic structure of the binding site of *P. aeruginosa* (M121A) azurin (pH 5.1), PDB code 2TSA.

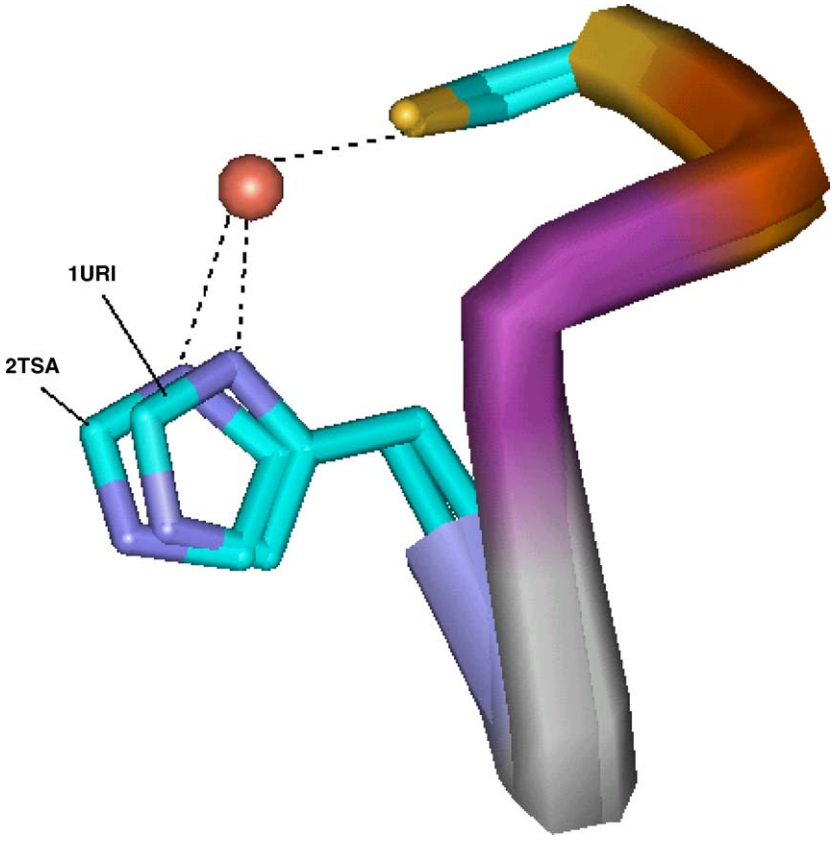


Fig. 6. Superposition of the crystallographic structures of a portion of the binding site of two azurins with different reduction potentials: 2TSA has a reduction potential of 373 mV and 1URI of 263 mV. The first protein has a N(His117)···Cu distance of 239.03 pm and N(His117)···Cu···S(Cys112) angle of 127.72°, while the second has 202.83 pm and 116.98°, respectively. The dotted lines are provided as guides.

reduction potential of azurins and pseudoazurins (graphic not shown).

As a consequence of the changes in the folding degree of His117 the Cys112···Cu···His117 angle appears to be related to the reduction potential of these proteins. This angle can be related to the efficiency in the electron flow (see Fig. 5), showing that an optimal angle, approximately 130°, increases the reduction potential of these BCPs. However, this quadratic correlation, illustrated in Fig. 7, shows a correlation coefficient of only 0.86, which indicates that there are other factors which are accounted by the folding degree contribution of His117 but not by this geometric parameter. Another possibility is that the higher folding degree of His117 makes this residue more accessible by solvent (Jeuken et al., 2000), influencing the efficiency of the electron transfer, a water molecule interacting with the imidazole ring of His117 to assist in the electron flow (see Fig. 5).

The consequences of the change in the folding degree associated with His117 can be related to several other structural and functional factors of the BCPs, which influence their reduction potential to some degree. This residue is in a protein region characterized by a high degree of folding. This region is formed mainly of a turn, or the combination of a turn and a small 3_{10}-helix (1rkr; 1nwp; 1dyz; 5azu; 1paz; 8paz) or even by only a 3_{10}-helix in 1zia. These characteristic folds make the degree of folding of this region significantly high compared to the rest of protein formed mainly by strands. Consequently, the high folding degree of this region can be

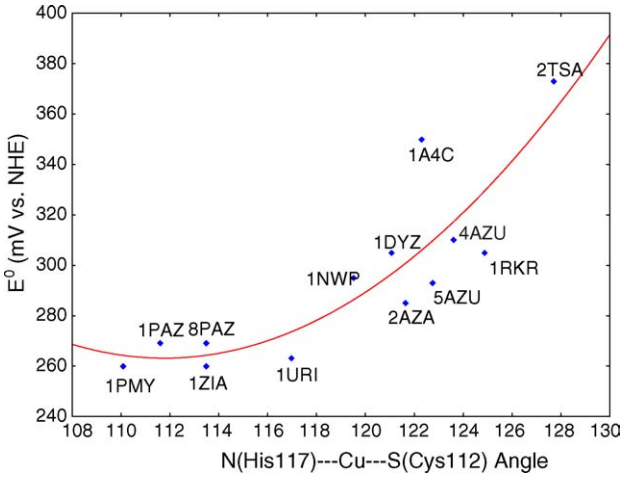


Fig. 7. Plot of the reduction potential of azurins and pseudoazurins as a function of the angle formed between His117 (azurins) or His81 (pseudoazurins), the copper atom and Cys112 (azurins) or Cys78 (pseudoazurins). The line is the quadratic model showing a correlation coefficient of 0.86.

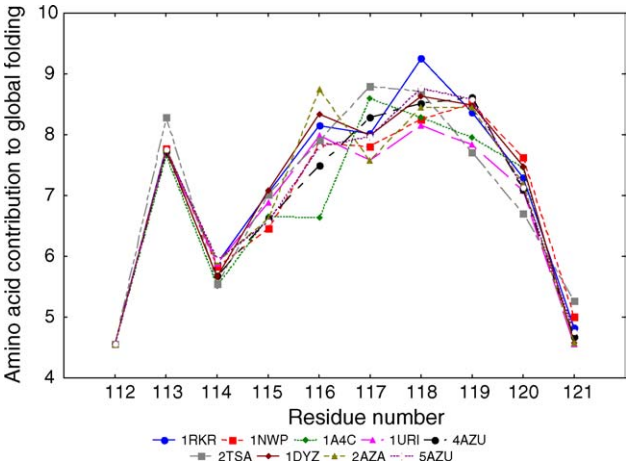


Fig. 8. Folding degree map of residues 112–121 in the azurins studied. Note that Cys112 (bonded to Cu) does not change its folding from one protein to another. However, His117, also bonded to Cu, follows the same trend as the reduction potential of these blue copper proteins.

a necessary prerequisite for the high reduction potential of these proteins. However, only the folding degree contribution of His117 appears to be quantitatively related to the reduction potential of the azurins and pseudoazurins as illustrated in Fig. 8.

3. Conclusions

We have shown that the degree of folding of azurins and pseudoazurins contains important 3D information that has probably been the result of evolutionary pressure over millions of years. The large range of reduction potentials in BCPs is also a result of evolutionary pressure, since it parallels different biological functions of the individual proteins. We have found that the reduction potential of these proteins is a function of the contribution to the folding of His117, a key amino acid in the binding site of these proteins. There are multiple structural consequences of the change in degree of folding of His117 during the evolution of these proteins. We have found that, at a minimum, the N(His117)···Cu···S(Cys112) angle which plays an important role in electron transport, could be involved in the differences in the reduction potential of azurins and pseudoazurins. The effect of the degree of folding of His117 on the efficiency of electron transfer from this residue to H_2O is an interesting problem worthy of further study. We hope that these results shed some new light on the structure–function relationships of this important family of proteins.

Acknowledgement

E.E. thanks "Ramón y Cajal" program, Spain for partial financial support.

Appendix A

Let $I_3(k)$ the contribution of the torsion angle k to the folding degree of the protein, where k is one of the following angles: ϕ_i, ψ_i or ω_i. Then, according to the original definition of folding degree index we have that:

$$I_3(k) = \sum_{l=0}^{\infty} \frac{\mu_l(k)}{l!} \tag{A.1}$$

Let e_i be the orthogonal projection of the unit vector (the ith vector of the canonical base of R^n) on ν_j is:

$$pr_j(e_i) = \frac{\langle e_i, \nu_j \rangle}{||\nu_j||^2} \nu_j = \langle e_i, \nu_j \rangle \nu_j = \nu_j(k) \cdot \nu_j \tag{A.2}$$

where ν_j is the eigenvector associated to the jth eigenvalue of A and $\nu_j(k)$ is its kth element.

Hence, the number of closed walks starting at torsion angle k can be expressed in terms of the spectral properties of the graph as follows:

$$\mu_l(k) = (A^l)_{ii} = \langle A^l e_i, e_i \rangle = \left\langle A^l \sum_{j=1}^{N} pr_j(e_i), \sum_{j=1}^{N} pr_j(e_i) \right\rangle$$

$$= \sum_{j=1}^{N} (\lambda_j)^l [\nu_j(k)]^2 \tag{A.3}$$

Substituting this expression in (A.1) we obtain:

$$I_3(k) = \sum_{l=0}^{\infty} \left(\sum_{j=1}^{N} \frac{(\lambda_j)^l [\nu_j(k)]^2}{l!} \right) \tag{A.4}$$

which can be reordered to obtain the absolutely convergent series:

$$\sum_{j=1}^{N} \left([\nu_j(k)]^2 \sum_{k=0}^{\infty} \frac{(\lambda_j)^l}{l!} \right) = \sum_{j=1}^{N} [\nu_j(k)]^2 \, \mathrm{e}^{\lambda_j} \tag{A.5}$$

and this clearly also converges to $I_3(k)$(Estrada and Rodríguez-Velázquez, 2005):

$$I_3(k) = \sum_{j=1}^{N} [\nu_j(k)]^2 \, \mathrm{e}^{\lambda_j} \tag{A.6}$$

Let i be the ith amino acid in the protein chain and let ψ_i and ϕ_i be the two-torsion angle contributing to the folding degree of this amino acid. Then, if we substitute $\nu_j(k)$ by $\nu_j(\psi_i)$ and $\nu_j(\phi_i)$ in (A.6) we obtain the expression for the contribution of ith amino acid to the global degree of folding of the protein:

$$I_3(i) = \sum_{j=1}^{N} \mathrm{e}^{\lambda_j} \{ [\nu_j(\Psi_i)]^2 + [\nu_j(\phi_i)]^2 \} \tag{A.7}$$

References

Broman, L., Malmström, B.G., Aasa, R., Väenngård, T., 1962. Quantitative electron spin resonance studies on native and denatured ceruloplasmin and laccase. J. Mol. Biol. 5, 301–310.

De Rienzo, F., Gabdoullive, R.R., Menziani, M.C., Wade, R.C., 2000. Blue copper proteins: a comparative analysis of their molecular interaction properties. Protein Sci. 9, 1439–1459.

Domou, B., Broder, S., 2004. Implications of new proteomic strategies for biology and medicine. J. Proteome Res. 3, 253–260.

Estrada, E., 2000. Characterization of 3D molecular structure. Chem. Phys. Lett. 319, 713–718.

Estrada, E., 2002. Characterization of the folding degree of proteins. Bioinformatics 18, 697–704.

Estrada, E., 2003. Application of a novel graph-theoretic folding degree index to the study of steroid-DB3 binding affinity. Computat. Biol. Chem. 27, 305–313.

Estrada, E., 2004a. A protein folding degree measure and its dependence on crystal packing, protein size, secondary structure, and domain structural class. J. Chem. Inf. Comput. Sci. 44, 1238–1250.

Estrada, E., 2004b. Characterization of the amino acid contribution to the folding degree of proteins. Proteins Struct. Funct. Bioinf. 54, 727–737.

Estrada, E., Rodríguez-Velázquez, J.A., 2005. Subgraph centrality in complex networks. Phys. Rev. E. 71, 056103.

Gray, H.B., Malmström, B.G., Williams, R.J.P., 2000. Copper coordination in blue proteins. J. Biol. Inorg. Chem. 5, 551–559.

Hooft, R.W.W., Sander, C., Vriend, G., 1996. The PDBFINDER database: a summary of PDB DSSP and HSSP information with added value. CABIOS 12, 525–529.

James, B.R., Williams, R.J.P., 1961. Oxidation-reduction potentials of some copper complexes. J. Chem. Soc., 2007–2019.

Jeuken, L.J.C., van Vliet, P., Verbeet, M. Ph., Camba, R., McEvoy, J.P., Armstrong, F.A., Canters, G.W., 2000. Role of the surface-exposed and copper-coordinating histidine in blue copper proteins: The electron-transfer and redox-coupled ligand binding of His117gly azurin. J. Am. Chem. Soc. 122, 12186–12194.

Katz, A.K., Shimonyi-Livny, L., Navon, O., Bock, C.W., Glusker, J.P., 2003. Copper-binding motifs: structural and theoretical aspects. Helv. Chim. Acta 86, 1320–1338.

Lappin, A.G., 1991. Properties of copper “blue” proteins. In: Sigel, H. (Ed.), Metal Ions in Biological Systems. Marcel Dekker Inc., New York.

Libeu, C.A.P., Kukimoto, M., Nishiyama, M., Horinouchi, S., Adman, E.T., 1997. Site-directed mutants of pseudoazurin: explanation of increased redox potentials from X-ray structures and from calculation of redox potential differences. Biochemistry 36, 12160–13179.

Nersissian, A.M., Shipp, E.L., 2002. Blue copper-binding domains. Adv. Protein Chem. 60, 271–340.

Pozdnyakova, I., Guidry, J., Wittung-Stafshede, P., 2001. Probing copper ligands in denaturated *Pseudomona aeruginosa* azurin: unfolding His117Gly and His46Gly mutants. J. Biol. Inorg. Chem. 6, 182–188.

Punj, V., Das Gupta, T.K., Chakrabarty, A.M., 2003. Bacterial cupredoxin azurin and its interactions with the tumor suppressor protein p53. Biochem. Biophys. Res. Commun. 312, 109–114.

Randić, M., Krilov, G., 1997. Characterization of 3D sequences of proteins. Chem. Phys. Lett. 272, 115–119.

Randić, M., Krilov, G., 1999. On a characterization of the folding of proteins. Int. J. Quantum Chem. 75, 1017–1026.

Ryde, U., Olsson, M.H.M., Roos, B.O., De Kerpel, J.O.A., Pierloot, K., 2000. On the role of strain in blue copper proteins. J. Biol. Inorg. Chem. 5, 565–574.

Shepard, W.E.B., Anderson, B.F., Lewandowski, G.E., Norris, D., Baker, J., 1990. Copper coordination geometry in azurin undergoes minimal change on reduction of copper(II) to copper(I). J. Am. Chem. Soc. 112, 7817–7819.

Solomon, E.I., Baldwin, M.J., Lowery, M.D., 1992. Electronic structures of active sites in copper proteins: contributions to reactivity. Chem. Rev. 92, 521–542.

Vallee, B.L., Williams, R.J.P., 1968. Metalloenzymes: the entatic nature of their active sites. Proc. Natl. Acad. Sci. U.S.A. 59, 498–505.

Williams, R.J.P., 1971. Catalysis by metalloenzymes. Entatic state. Inorg. Chim. Acta Rev. 5, 137–155.