X-ray Absorption Studies and Homology Modeling Define the Structural Features That Specify the Nature of the Copper Site in Rusticyanin

J. Günter Grossmann,[‡] W. John Ingledew,[§] Ian Harvey,[||] Richard W. Strange,[‡] and S. Samar Hasnain*[,‡]

Molecular Biophysics Group, Daresbury Laboratory, Warrington, Cheshire WA4 4AD, U.K., Department of Biological and Medical Sciences, University of St. Andrews, St. Andrews, KY116 9AL, U.K., and School of Applied Sciences, De Montfort University, The Gateway, Leicester LE1 9BH, U.K.

Received November 7, 1994; Revised Manuscript Received February 13, 1995[⊗]

ABSTRACT: Rusticyanin, a blue copper protein, possessing the highest redox potential among this class of proteins and a high stability at acidic pH reveals homology with the C-terminal end of the other single copper containing blue proteins and an interesting homology to parts of the blue copper domain of the multi-copper proteins such as the nitrite reductases. Extended X-ray absorption fine structure (EXAFS) data at pH 2.0 reveal that Cu is ligated to two His and a Cys in the inner coordination sphere, similar to other blue copper centers. Modeling studies suggest that His85 is the ligating histidine from the N-terminal end. Its neighboring residue is a serine rather than the asparagine found in all known blue Cu proteins. The high stability of the copper site may arise in part due to this substitution. The Cu binding site is surrounded by aromatic residues which may provide further protection for the metal in an acidic environment. In addition, the high number of solvent-exposed uncompensated lysine residues is likely to be of functional relevance under low pH conditions. EXAFS data show a very small change (relative to azurin) in the copper site upon reduction, consistent with a more constrained copper center in rusticyanin compared to azurin and a higher redox potential.

Rusticyanin is a small type-1 blue copper protein and is thought to be a principal component in the iron respiratory electron transport chain of *Thiobacillus ferrooxidans* (Ronk et al., 1991; Ingledew et al., 1977). The protein possesses the highest redox potential of the single blue copper proteins, 680 mV versus a more typical redox potential of ~300 mV. In addition, the high stability of this protein at pH 2 is also intriguing.

The crystallographically characterized blue proteins have shown a basic framework of a copper site where copper ligation is provided by two His, one Cys, and one Met. The lack of methionine in stellacyanin has been thought to be responsible for the lowest redox potential, 184 mV, in the family and has been extensively investigated by site-directed mutagenesis (Karlsson et al., 1991; Murphy et al., 1993; Romero et al., 1993) of azurins from *Pseudomonas aeruginosa* and *Alcaligenes denitrificans*.

Amino acid sequences of rusticyanin from three different strains of *T. ferrooxidans* have become available recently (Yano et al., 1991; Ronk et al., 1991; Nunzi et al., 1993) showing very high degree of homology (>90%) among themselves but a very low invariance (<20%) with azurin, plastocyanin, and pseudoazurin. These data have helped in identifying the three Cu ligands from the C-terminus, a His, Cys, and a Met. However, the nature of the fourth ligand, normally a His from the N-terminal end, is under some dispute and different candidates (His39, 57, or 85, a Met, and Asp73) have been proposed. Blake and co-workers (Ronk et al., 1991) constructed a structural model using sequence alignment with other single blue Cu proteins and favored the Asp73 as a ligand and suggested that such a ligation is responsible for the high redox potential of the protein. Since then, the H46D mutant of azurin has become available (Germanas et al., 1993). This mutant exhibits a redox potential of 297 mV and spectroscopic properties that are not similar to those of rusticyanin thus casting doubts about the proposed model of Ronk et al. (1991) for rusticyanin. Recently, the preparation and characterization of a Co(II)-substituted rusticyanin was reported (Strong et al., 1994). The comparison of UV−vis and EPR spectra with other cobalt-substituted blue copper proteins suggests a coordination sphere in rusticyanin with two histidines, a methionine, and a cysteine.

We provide EXAFS data which strongly support a coordination of two histidines and a cysteine in the inner coordination sphere. Our data are in general agreement with an earlier EXAFS study carried out under different conditions by Holt et al. 1990. A detailed comparison of the amino acid sequence of rusticyanin based on the crystallographic structures of small type-1 copper proteins as well as the multi-copper proteins nitrite reductase and ascorbate oxidase shows that the second histidine ligand is most likely to be the His 85, as proposed by Nunzi et al. (1993). Sequence homology has been translated into structural homology, and a three-dimensional model is proposed. This is discussed in terms of the high redox potential, high acid stability of the protein, and the nature of the copper site.

MATERIALS AND METHODS

Sample Preparation. Rusticyanin from *T. ferrooxidans* was isolated according to the procedure of Lewis et al. (1984). Initial protein solutions containing 50 μM copper were made up in 50 mM β-alanine/sulfate buffer (pH 3.5).

* To whom correspondence should be addressed.
[‡] Daresbury Laboratory.
[§] University of St. Andrews.
[||] De Montfort University.
[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1995.

0006-2960/95/0434-8406$09.00/0

Samples for EXAFS measurements were prepared by dialysis in 10% H_2SO_4 to give a final pH of 2.0. The copper content in the samples was increased to approximately 1 mM by further dialysis using aquacide. Reduced rusticyanin samples were prepared by adding ascorbic acid and observing the bleaching of the intense blue color of the oxidized samples. Samples were placed in EXAFS cells with mylar windows and frozen in liquid nitrogen. The integrity of the oxidized sample was checked before and after exposure to X-rays from the synchrotron by measuring its electron paramagnetic resonance (EPR) spectrum. This was accomplished using a sample cell which permitted direct transfer of the frozen protein solution between the EPR spectrometer and the EXAFS instrument, thus avoiding thawing of the sample.

EXAFS Data Collection. X-ray absorption spectra at the Cu K-edge were measured in fluorescence mode with a 13-element Ge solid-state detector on wiggler EXAFS station 9.2 at the Synchrotron Radiation Source, Daresbury Laboratory. The station was equipped with a silicon 220, order sorting, double crystal monochromator which was set up to perform harmonic rejection at 50%. The electron beam energy was 2 GeV, and the average current was 170 mA. During data collection, a liquid nitrogen cryostat was employed to maintain a sample temperature of approximately 80 K. Data were collected in k-space using a k^3 weighted regime for the counting time, with a maximum counting time of 80 s. A total of eight scans each were recorded on oxidized and reduced rusticyanin at pH 2.0. X-ray damage and changes in the oxidation state of the protein were monitored by comparison of the data sets during data collection; no changes were observed.

EXAFS Data Analysis. Data reduction and analysis were performed on a Convex C220 computer. After summing and background removal, analysis of the raw EXAFS data was accomplished via the full curved wave multiple scattering approach using the Daresbury Laboratory program EXCURV92 (Binsted et al., 1992). Constrained and restrained refinement methods were used to incorporate geometrical information in the analysis and to reduce the ratio of refinable parameters to observations (Hasnain & Strange, 1990; Binsted et al., 1992; Strange et al., 1995). Simulations were carried out on raw EXAFS data weighted by k^3 from $k = 3.5-11.0$ Å^{-1}.

The phase shifts employed in the data analysis were calculated *ab initio* using the program MUFPOT. The Cu, N, C, and O phase shifts have previously been used to simulate a number of Cu−histidine model compounds (Strange et al., 1987) as well as the Cu site of superoxide dismutase (Blackburn et al., 1987). Full curved wave multiple scattering analysis (Gurman et al., 1986) was included in the above cases. The S phase shift was obtained from EXCURV92, which uses a simplified version of the MUFPOT program. The central atom (copper) phaseshifts employed here were calculated assuming a neutral $Z + 1$ 1s core hole excited state. This approximation and the S phase shifts were also used to analyze the XAFS of other type-1 copper proteins, including a number of methionine-121 azurin mutants (Murphy et. al., 1993) and stellacyanin (Strange et al., 1995), where errors in fitting the inner shell distances to about 2.6 Å were found to be ±0.02 Å.

Sequence Alignment and Modeling. The amino acid sequence of rusticyanin from *T. ferrooxidans* characterized by Nunzi et al. (1993) was used in this work. The sequence shows rather low (≤21%) overall similarity to that of other type-1 copper binding proteins (see Table 1). Sequence identity is quite high in the C-terminal region where three conserved copper binding residues (Cys138, His143, and Met148, rusticyanin numbering scheme) can be located (Ronk et al., 1990; Nunzi et al., 1993; Figure 1). A sequence identity of less than 25% is considered rather low for predicting a three-dimensional structure of a protein (Blundell et al., 1987). However, even at low overall sequence homology a good model of the active site can be attained by comparison (Chothia & Lesk, 1986). Thus, we have concentrated on modeling the chain segments composing the hydrophobic core and the copper center, located in a confined area near the protein surface. We have not attempted to model side chain conformations in the core of the protein accurately due to the low sequence identity.

Table 1: Amino Acid Sequence Identity Matrix for the Structural Alignment of Blue Copper Proteins[a]

	NiR1	AAN	AZA	CBP	PAZ	PLC	AO3	NiR2	AO1	AO2	RTf3
NiR1	-	20.0	15.6	15.8	20.4	16.2	12.1	10.7	25.4	6.1	**19.4**
AAN	21	-	14.1	13.7	27.9	20.2	12.4	14.3	13.3	6.7	**20.0**
AZA	20	18	-	13.7	20.4	15.2	11.7	14.8	12.5	10.2	**14.8**
CBP	15	13	13	-	12.9	15.8	10.5	5.3	4.2	6.3	**13.7**
PAZ	19	26	19	12	-	26.9	12.9	14.0	14.0	6.5	**18.3**
PLC	16	20	15	15	25	-	13.1	9.1	8.1	7.1	**16.2**
AO3	19	13	15	10	12	13	-	14.0	18.9	4.5	**10.3**
NiR2	16	15	19	5	13	9	21	-	17.2	7.6	**8.7**
AO1	31	14	16	4	13	8	23	21	-	6.6	**14.8**
AO2	8	7	13	6	6	7	6	10	8	-	**4.5**
RTf3	**30**	**21**	**19**	**13**	**17**	**16**	**16**	**13**	**18**	**6**	-

[a] The upper diagonal of the matrix shows the percentage of identical residues in identical positions. The lower diagonal gives the raw number of identities. Values for rusticyanin [RTf3, sequence version 3 according to Nunzi et al. (1993)] are the result of our sequence alignment (see also Figure 1). Abbreviations: NiR1, nitrite reductase domain 1; AAN, amicyanin; AZA, azurin; CBP, cucumber basic protein; PAZ, pseudoazurin; PLC, plastocyanin; AO3, ascorbate oxidase domain 3; NiR2, nitrite reductase domain 2; AO1 and AO2, domains 1 and 2 of ascorbate oxidase. The latter three domains (denoted by shaded boxes) do not contain a type 1 copper binding site but have been included as they are thought to be structurally related and have resulted from gene multiplication.

We have used X-ray coordinates of five members of the family of small blue copper proteins from the Brookhaven Protein Data Bank (PDB; Abola et al., 1987; Bernstein et al., 1977) as a basis for modeling rusticyanin. These were amicyanin from *Paracoccus denitrificans* (1AAN, Durley et al., 1993), azurin from *A. denitrificans* (2AZA, Baker 1988), cucumber basic protein (1CBP, Guss et al., 1988), pseudoazurin from *Alcaligenes faecalis* S-6 (2PAZ, Adman et al., 1989), and plastocyanin from *Lombardi poplar* (1PLC, Guss et al., 1992). In addition, the multidomain structures of nitrite reductase from *Achromobacter cycloclastes* (1NRD, Godden et al., 1991) and ascorbate oxidase from *Zucchini* (1AOZ, Messerschmidt et al., 1992) allowed an interesting approach for evaluating the homology with rusticyanin.

The Cα traces of the five small blue copper proteins as well as the two and three domains of nitrite reductase (NiR) and ascorbate oxidase (AOZ), respectively, were superimposed on a Silicon Graphics Indigo workstation running InsightII (Biosym Technologies, Inc., San Diego, CA). Amino acid residues in topologically equivalent positions were recorded in order to derive a structurally based multiple sequence alignment (Figure 1). Automatic sequence alignment methods such as those implemented in the University

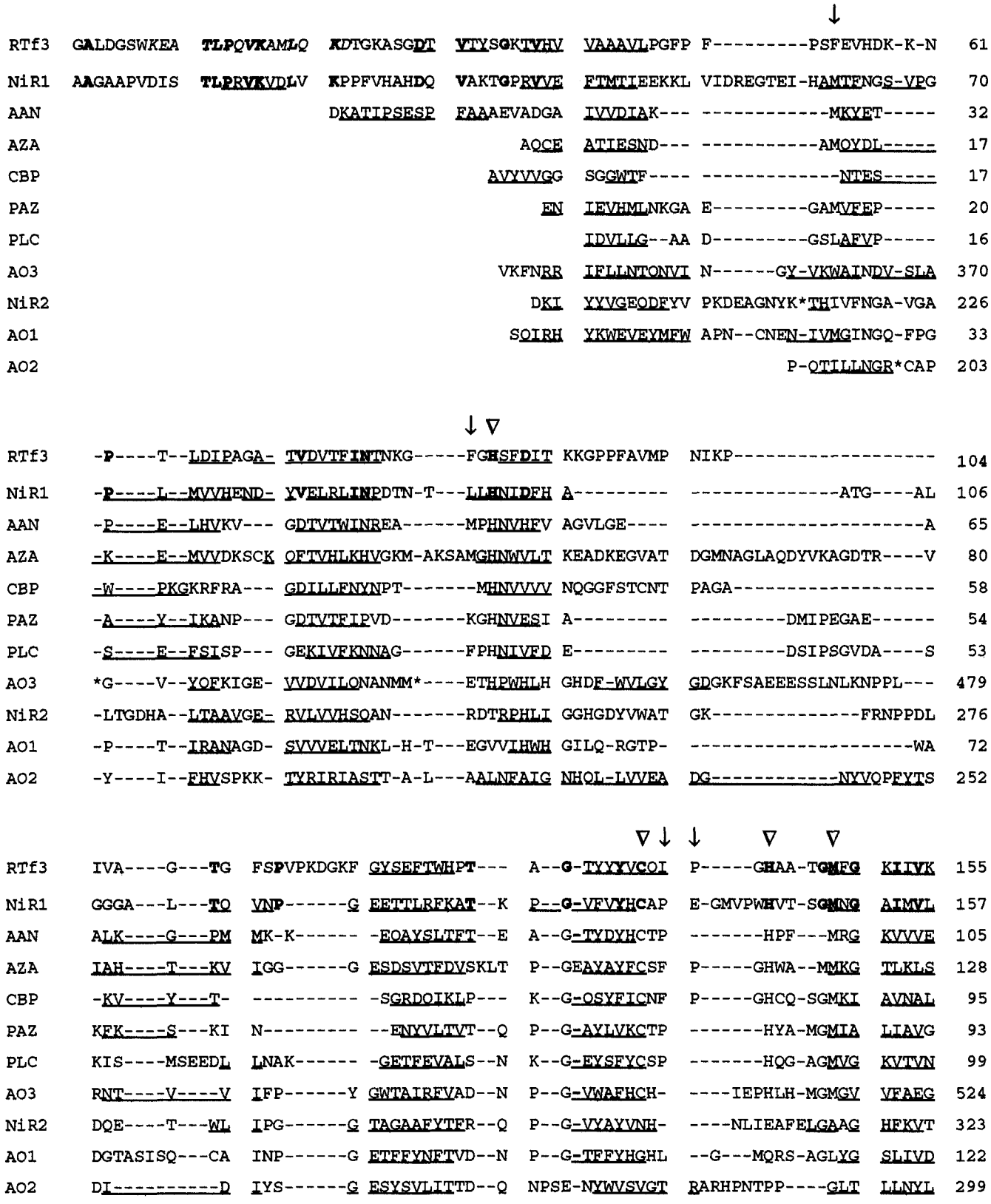

FIGURE 1: Alignment of rusticyanin with blue copper proteins of known structure. Abbreviations of proteins are the same as for Table 1. Underlined residues correspond to β-strands, as determined by X-ray crystallography and, in the case of rusticyanin, by the prediction methods of Chou–Fasman and Garnier–Osguthorpe–Robson. A predicted α-helix at the N-terminal of rusticyanin is printed in italics. Conserved type-1 copper ligands and hydrophobic residues near the copper site are marked by triangles (∇) and arrows ($\downarrow$), respectively. Identical residues between rusticyanin and NiR domain 1 are highlighted. Asterisks (*) denote omitted loop extensions in NiR2, AO2, and AO3.

of Wisconsin Genetics Computer Group software (Devereux et al., 1984) produced a convincing alignment between the sequences of rusticyanin and the domain 1 of NiR (Godden et al., 1991). The N- and C-terminal regions show relatively high sequence identity and the characteristic type 1 Cu ligands (2xHis, 1xCys, and 1xMet) were matched in both sequences. These features as well as secondary structure assignments for rusticyanin from hydrophobic cluster analysis

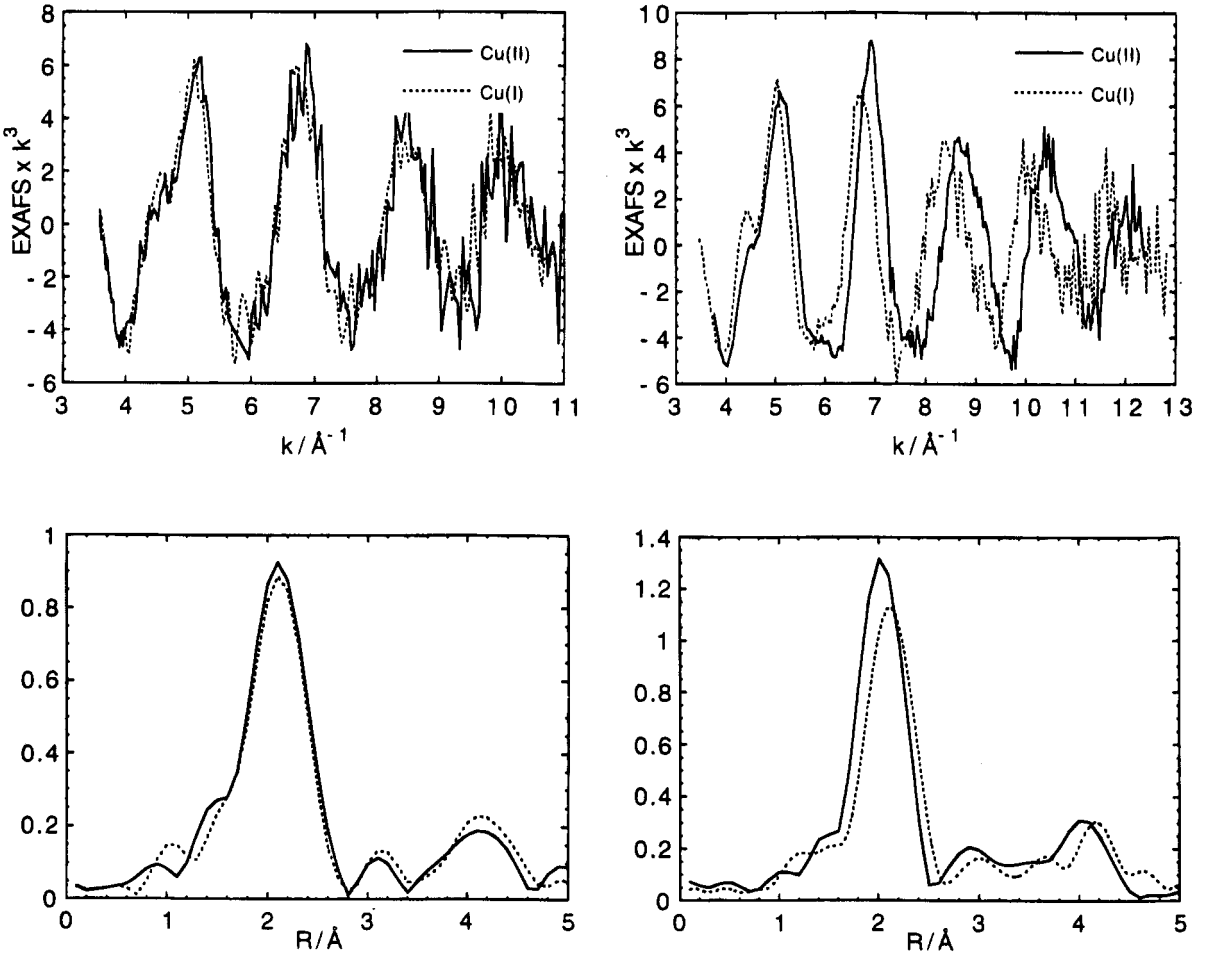

FIGURE 2: k^3-weighted raw EXAFS and Fourier transforms measured at the Cu K-edge of (a, left) oxidized (solid line) and reduced (dashed line) rusticyanin at pH 2 and (b, right) oxidized (solid line) and reduced (dashed line) azurin from *P. aeruginosa* at pH 8.0. The data for Cu(II) azurin are taken from Murphy et al. (1993). Reduced azurin data were prepared in a manner similar to reduced rusticyanin (unpublished results).

(Nunzi et al., 1993), Chou–Fasman (Chou & Fasman, 1978), and Robson–Garnier (Garnier et al., 1978) predictions were then used to align the sequence of rusticyanin manually to the multiple alignment. Interestingly, rusticyanin and domain 1 of NiR possess almost identical chain lengths. Two sections of the polypeptide are difficult to align; in particular, residues 50–64, for which the corresponding region in domain 1 of NiR is longer as it provides interactions with domain 2 and residues 92–119, a segment which shows a high structural divergence among the blue copper proteins. These portions of the polypeptide chain are not likely to interfere with the copper binding site nor to disturb the β-sandwich framework. Some confidence can be obtained in the multiple sequence comparison in our final alignment (Figure 1) due to the buried nature of hydrophobic residues, hydrogen-bonding patterns of β-sheets, and surface polarity in the predicted structure. At the N-terminal end the sequence comparison with domain 1 of NiR lends credence to the suggestion that rusticyanin is extended by approximately 35 residues at the N-terminus compared to other small blue copper proteins (except for amicyanin). The high similarity between the N-termini of rusticyanin and domain 1 of NiR suggests a structural model of rusticyanin based on NiR is likely to be appropriate. However, the assignment of an α-helix for rusticyanin from residue 8 to 22 using secondary structure predictions (see Figure 1) disagrees with the β-strand present in NiR. Thus, the model of the N-terminal of rusticyanin should be regarded as tentative. However, in view of its remoteness to the Cu site, it is unlikely to affect the immediate structure around the metal center.

The model of rusticyanin was constructed by replacing side chains of domain 1 of NiR with side chains of rusticyanin according to the alignment in Figure 1. The domain 1 of NiR was chosen because it is the closest in length (157 residues) to rusticyanin (155 residues) and shows, along with amicyanin, the highest sequence identity (19.4%) to rusticyanin among the other blue copper proteins. For two small portions, segments of pseudoazurin (residues 49–53, rusticyanin numbering) and azurin (residues 139–142, rusticyanin numbering) had to be incorporated in order to compensate for different length of the loops. In addition, two segments (residues 92–104 and 114–119, rusticyanin numbering) could not be aligned and were therefore modeled to connect the β-strands using peptide fragments from the protein database. Thus, the proposed structure for these sections is only suggested as plausible conformations. Residues 114–119 are distant from the Cu site. Even though residues 92–104 may stabilize part of the polypeptide comprising the active site triad residues (Cys138, His143, and Met148), a substantial influence on the overall geometry of the Cu site is not expected.

Energy minimization, using the Powell optimization procedure in X-PLOR (Brünger, 1990), was carried out to eliminate improper atomic overlaps and poor geometry. We used the empirical force-field and the all-hydrogen parameter

file of X-PLOR. The initial 100 steps of minimization concentrated on the protein alone. Protein was placed in a 5 Å shell of water, and 2000 cycles of energy minimization were carried out. Although the side chains of the active site ligands are highly conserved in the type-1 copper proteins, the copper−ligand distances found in our EXAFS experiments were used as constraints during energy minimization in order to optimize the model for the copper site. Bond angles at the copper site were assigned according to the crystal structure of *A. denitrificans* azurin (Baker, 1988). Improper-angle restraints were used to keep the Cu atom coplanar with respect to the two imidazole rings. The two histidine ligands were not distinguished and averaged Cu−histidine parameters were used. Molecular mechanics calculations were performed on a Convex C220 computer. The resulting molecular model of rusticyanin has rms deviations from ideal bond lengths of 0.012 Å and from ideal bond angles of 3.2°.

RESULTS

EXAFS. A comparison of the raw EXAFS data for the oxidized and reduced forms of the protein shows that only small changes occur at the copper site upon reduction (Figure 2a). By contrast, a much larger change is observed in the copper site for oxidized and reduced *P. aeruginosa* azurin (Figure 2b). An examination of the inner coordination shell showed that the data for rusticyanin required the presence of at least two low-*Z* atoms and an S back-scatterer, presumably from the cysteine ligand. From this preliminary analysis the presence of another back-scatterer at ~2.5 Å from Cu was evident. Simulations were carried out with models containing either a second S, from the methionine ligand, or a low-*Z* ligand. Evidence for histidine ligation in both oxidized and reduced protein is clear from Figure 2a, where the peaks in the Fourier transform at ~3 and ~4 Å are characteristic of imidazole and arise from scattering of the photoelectron by the N and C atoms of the imidazole ring (Strange et al., 1987). The number of coordinated histidines was investigated by carrying out simulations with either one or two histidines and with histidines coordinated at the same or at different distances from the copper atom. All atoms belonging to the imidazole ring of histidine were included in the calculations, allowing a distinction between one or two histidines to be made with more confidence than is possible with single scattering calculations.

Simulations with One Histidine. One of the inner shell low-*Z* atoms was assumed to be a nitrogen from histidine. In the oxidized data this ligand refined to 1.85 Å from the Cu, with the S(Cys) at 2.16 Å and the second low-*Z* atom (an oxygen) at 1.96 Å. The composition of the inner shell was completed by including the S(Met) at 2.59 Å. The fit parameters are given in full in Table 2. The simulation has a fit index of 1.03 and is shown in Figure 3a. A similar analysis of the reduced protein gives a copper site with 1N(His) at 1.84 Å, S(Cys) at 2.15 Å, 1O at 1.98 Å, and the S(Met) at 2.60 Å. This simulation has a fit index of 0.89. These simulations account for the amplitude of the major peak (~2−2.6 Å) but are clearly inadequate in reproducing the amplitude of either the second or third shell features of the Fourier transform, indicating a higher number of histidine coordination.

Simulations with Two Histidines. When two histidines are assumed to be coordinated at the same distance, the back-scattering contribution from the imidazole ring is doubled. The amplitude of the outer shell peaks is accounted much better by this model and the fit to the EXAFS is significantly improved, with the fit index 15% smaller in the oxidized data (Figure 3b) and 18% smaller in the reduced data. In both cases the two N(His) refined to 1.94 Å with the S(Cys) at 2.17 Å and the S(Met) at 2.58 Å. There were also changes in the Debye−Waller terms compared to the simulation using only one histidine, and these details are given in full in Table 2. One important observation is that the Debye−Waller term for the two N(His) is large, at 0.015 Å^2, by comparison with EXAFS results obtained for other type 1 copper proteins using the same analysis procedures [e.g., Murphy et al. (1993)]. We have shown elsewhere (Strange et al., 1995) that a large value for the Debye−Waller term of these atoms implies that the histidines are coordinated at different distances.

A simulation with the two histidines at different distances from the copper atom was therefore carried out, and it was made using restrained rather than constrained refinement (Binsted et al., 1992). In this case also, the refinement remained overdetermined. The resulting fit is shown in Figure 3c. In the oxidized protein the histidine ligands were split by 0.19 Å, refining at distances of 1.86 and 2.05 Å. The Debye−Waller terms for the N(His) are now 0.001 and

Table 2: Parameters Used To Simulate the Cu K-Edge EXAFS of Oxidized Rusticyanin at pH 2 Using Alternative Models for the Copper Coordination[a]

shell	atoms	R (Å)	$2\sigma^2$ (Å^2)
	(a) Parameters Used in Simulation with One Histidine		
1	1N(His)	1.85	0.003
2	1C(His)	2.77	0.006
3	1C(His)	2.94	0.006
4	1N(His)	3.94	0.009
5	1C(His)	4.04	0.009
6	1S(Cys)	2.16	0.000
7	1O	1.96	0.010
8	1S(Met)	2.59	0.023
	FI = 1.03, R = 34.7%, E_0 = 21.3		
	(b) Parameters Used in Simulation with Two Histidines Coordinated at the Same Distance		
1	2N(His)	1.94	0.015
2	2C(His)	2.87	0.017
3	2C(His)	3.02	0.017
4	2N(His)	4.04	0.023
5	2C(His)	4.13	0.023
6	1S(Cys)	2.17	0.002
7	1S(Met)	2.58	0.027
	FI = 0.88, R = 31.6%, E_0 = 19.1		
	(c) Parameters Used in Simulation with Two Histidines Coordinated at Different Distances		
1	1N(His)	1.86	0.001
2	1C(His)	2.76	0.004
3	1C(His)	2.98	0.004
4	1N(His)	3.93	0.007
5	1C(His)	4.06	0.007
6	1N(His)	2.05	0.002
7	1C(His)	2.96	0.005
8	1C(His)	3.10	0.005
9	1N(His)	4.14	0.008
10	1C(His)	4.17	0.008
11	1S(Cys)	2.15	0.001
12	1S(Met)	2.60	0.017
	FI = 0.19, R(exafs) = 28.3%, R(distance) = 0.6%, E_0 = 21.3		

[a] R denotes the Cu−ligand distance, and $2\sigma^2$ is the Debye−Waller term for each shell.

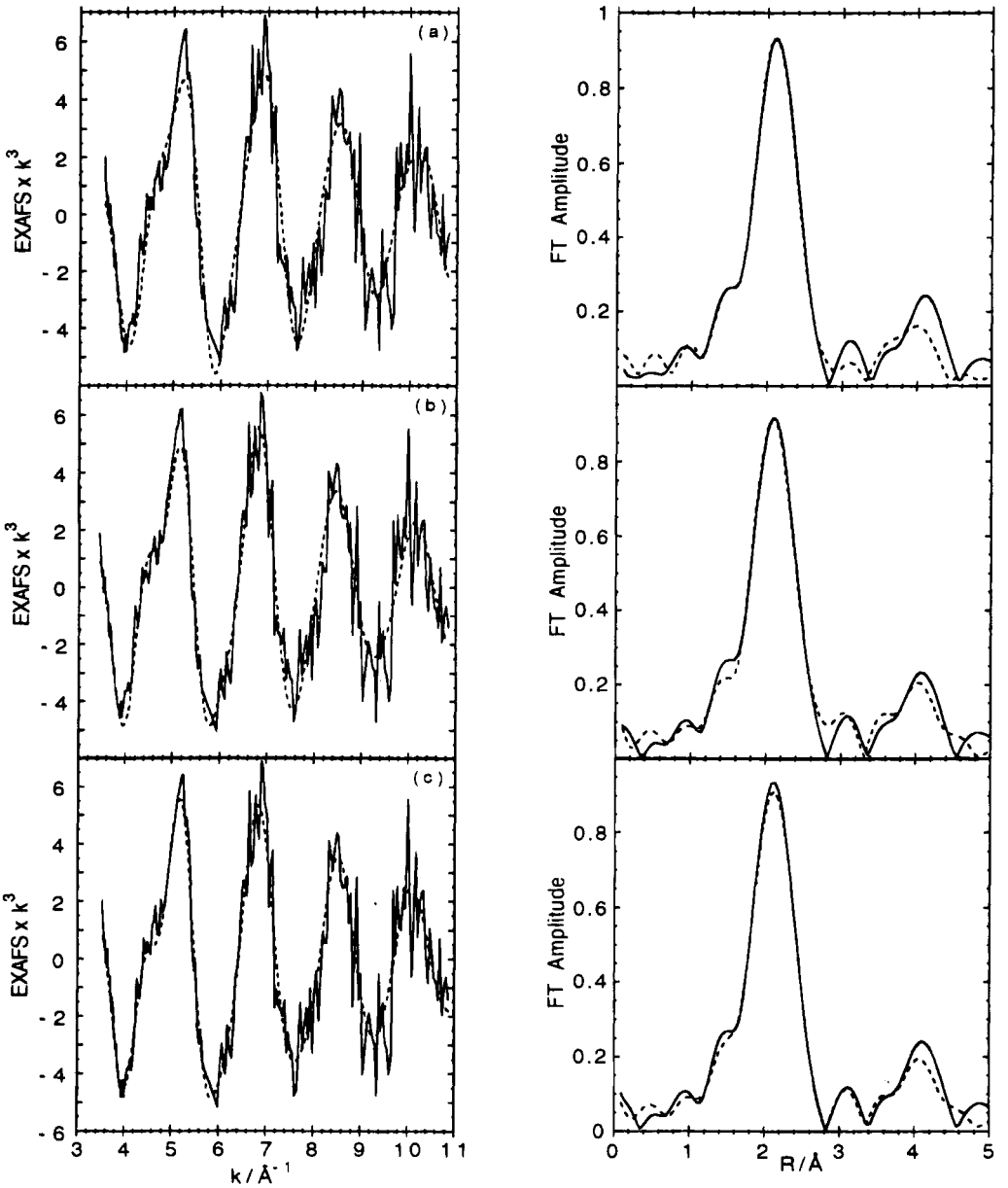

FIGURE 3: Simulation of the k^3-weighted EXAFS of oxidized rusticyanin using the parameters given in Table 2. The copper site is modeled with (a) one histidine ligand, (b) two histidine ligands at the same Cu−N distance, and (c) two histidine ligands coordinated at different distances. In each model 1S(Cys) and 1S(Met) are coordinated to the copper atom.

0.002 Å^2, respectively. The S(Cys) refined to 2.15 Å and the S(Met) to 2.60 Å. The fit index was 0.19. In the reduced protein, the histidines refined to 1.89 and 2.09 Å, with the S(Cys) at 2.16 Å and S(Met) at 2.63 Å, giving a fit index of 0.13. Note that the much smaller values for the fit index in these simulations compared to the previous values arises in part from the weighting scheme used in the restrained refinement (Binsted et al., 1992). The *R* factors given in Table 2 are independent of the weighting scheme.

In contrast to other blue proteins, where the Cu−S(Met) distance is ~3 Å [for example, see discussion in Murphy et al. (1993)], the inclusion of the S(Met) contribution in the EXAFS of rusticyanin is clearly significant. For example, in the final model for the oxidized protein, the fit index was 0.26 (36% larger) when the S(Met) was excluded from the fits. In earlier studies of related structures, such as amicyanin (Lommen et al., 1991) and plastocyanin (Scott et al., 1982; Penner-Hahn et al., 1989), the Cu−S(Met) contribution does not make a significant EXAFS contribution. In *P. aeruginosa* azurin (Murphy et al., 1993) the presence of this sulfur in the EXAFS simulation improves the fit index by only 1%. This observation may be attributed to the much shorter Cu−S(Met) distance (2.6 Å) and thus stronger photoelectron backscattering in rusticyanin compared to these other proteins.

The copper site of rusticyanin as determined by EXAFS is therefore similar to that of other known type 1 Cu proteins, such as azurin and plastocyanin, in that there are two histidine ligands and two sulfur ligands present. However, the copper site in rusticyanin is quite distinct in two respects. First, the methionine ligand is much shorter in the oxidized protein than in any other type 1 Cu protein including pseudoazurin from *A. faecalis* S-6, which has a Cu−S(Met) distance of ~2.7 Å (Adman et al., 1989; Petratos et al., 1988). A shorter Cu−S(Met) bond (~2.4 Å) has been reported in the low pH reduced form of *A. faecalis* S-6 (Vakoufari et al., 1994). Secondly, there is only a minimal change in the Cu site coordination upon reduction to Cu(I) (see also Figure 2). The Cu−ligand distances are increased by ≤0.04 Å upon reduction (Table 2). The changes are much smaller than those observed by EXAFS between the oxidized and reduced

states of azurin (Figure 2b), plastocyanin, and stellacyanin (manuscript in preparation). These small changes are of direct relevance to the high redox potential of rusticyanin.

An earlier EXAFS study of rusticyanin, using ammonium sulfate precipitation at room temperature, also showed the presence of two histidines at the Cu site in both oxidation states (Holt et al., 1990). There are, however, significant differences seen in the present study. These may arise from a number of factors, including an improved signal to noise ratio in the present work due to the use of an energy-discriminating solid-state Ge detector, solution versus paste samples, room temperature versus 80 K, and the more rigorous analysis procedures adopted here. For example, full account of multiple scattering has been taken in the present analysis.

Model Evaluation. Although the blue copper proteins used in our modeling studies show only a low degree of sequence homology (<25% sequence identity; see Table 1), their core structures as determined by X-ray crystallography are very similar and commonly characterized by the so-called β-barrel or β-sandwich topology. On this basis, a model for rusticyanin from *T. ferrooxidans* was considered to consist of a very similar β-sheet topology. This is supported by circular dichroism spectroscopy of rusticyanin (Ronk et al., 1991) which revealed mainly β-structural elements. Correct sequence alignment is the most crucial part for any structure predicted on the grounds of its similarity to molecules of known three-dimensional structure. In the case of the model for rusticyanin proposed here, the distribution of hydrophobic and hydrophilic residues which make up the core of the β-sandwich together with the polarity of surface residues and the spatial location of the copper ligands gives confidence to the sequence alignment.

Figure 4 shows the predicted structure for rusticyanin. The core structure is formed by two face-to-face β-sheets (β-sandwich) comprised of four and five strands of polypeptide backbone, respectively. β-Strands that face the solvent on one side and the protein interior on the other side have hydrophilic and hydrophobic residues in alternating order. Thus, the interior of the molecule is mainly hydrophobic while the surface shows a large number of polar/charged residues (particularly lysine residues). This is consistent with the typical distribution of side chains on a protein surface. There are a number of uncompensated lysine residues which point away from the protein surface and essentially do not cluster with negatively charged residues. This aspect may be responsible for rusticyanin's acid stability and its ability to function at low pH. In this connection it is interesting to note that the recently isolated and characterized blue copper protein halocyanin (Scharf & Engelhard, 1993), which requires a high pH ($\geq$9) environment, is composed of an extraordinary high amount of negatively charged amino acid residues; no lysines and only five arginine residues could be detected. The multiple sequence alignment predicts an N-terminal extension of the β-sandwich of more than 30 residues compared to the small blue copper proteins such as azurin or plastocyanin.

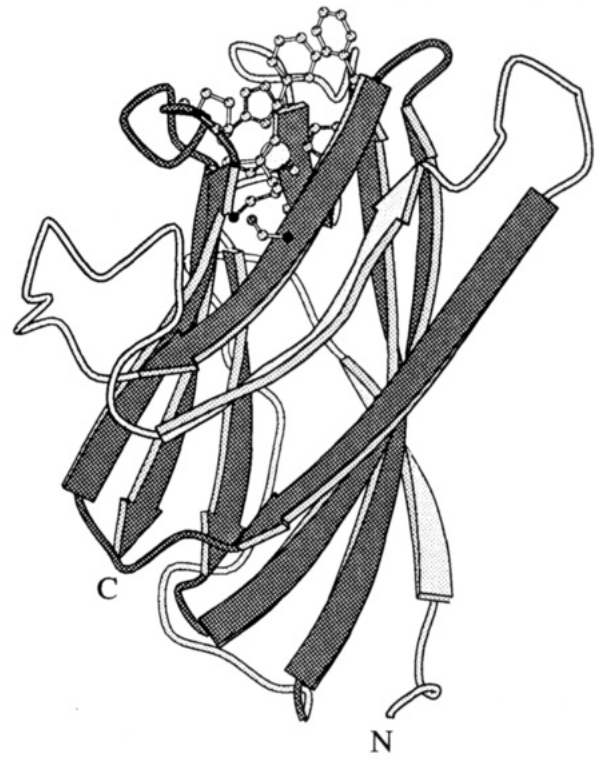

FIGURE 4: Schematic view of the energy-minimized model for rusticyanin based on the overall fold of domain 1 of nitrite reductase. Light gray has been used to highlight main differences from the NiR polypeptide backbone as well as speculative segments. The latter include both the N-terminal residues and extended loop sections (residues 92−104 and 114−119). Important amino acids in the vicinity of the copper site as well as the copper ligands (His85, Cys139, His143, and Met148) are represented by atomic models. The copper atom is shown as a large sphere. The picture was produced with the program Molscript (Kraulis, 1991).

DISCUSSION

Our primary interest is centered around the Cu site which lies at one end of the barrel between several of the β-strands and their connecting loops. As is known from previous reports (Ronk et al., 1991; Yano et al., 1991; Nunzi et al., 1993), three copper ligands of rusticyanin can be clearly identified in the C-terminal region which contains the conserved triad of residues among blue copper proteins Cys138, His143, and Met148. Since our EXAFS results are in favor of a second histidine ligand, the most likely fourth copper ligand appears to be His85. This ligand has not only been suggested in previous sequence comparisons of small blue copper proteins (Yano et al., 1991; Nunzi et al., 1993) but is strongly suggested by the sequence alignment with the type-1 copper binding domains of nitrite reductase and ascorbate oxidase. However, the high redox potential and the intriguing stability of the copper site of rusticyanin at low pH suggest distinct aspects of the metal site in rusticyanin.

The model presented for rusticyanin in this paper provides neat explanations for the copper site being more stable than in other blue copper proteins. One is the hydrophobic patch above the copper site on the otherwise highly hydrophilic surface of the molecule. His143, equivalent to His117 in azurin, is the only Cu ligand exposed to solvent. In azurin, the exposed histidine has been extensively discussed in terms of a route for electron transfer. In the case of rusticyanin His143 lies in the center of hydrophobic, mainly aromatic residues (Figure 5): 2xPhe, 1xPro, and 1xIle. The presence of *two* aromatic residues (Phe54 and 83) has no counterpart in the region of the metal sites for the known type-1 copper binding proteins and is a feature that will make the copper in rusticyanin less accessible to solvent. In other small blue copper proteins the hydrophobic patch is primarily formed by methionine (a rather flexible side chain which can be

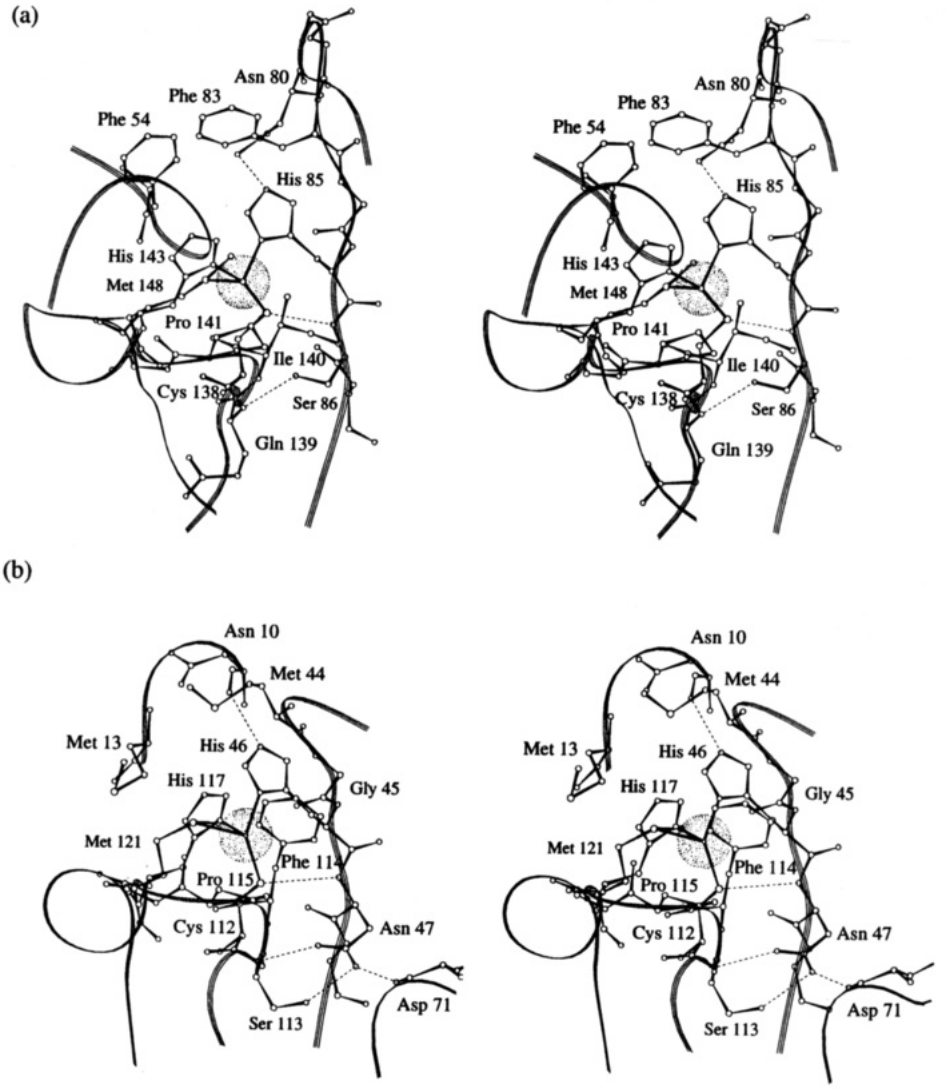

FIGURE 5: Close-up of the copper binding site (a) in the model of rusticyanin and (b) in the crystal structure of azurin (Baker, 1988). Copper ligands, hydrophobic residues, and amino acids forming hydrogen bonds (dotted lines) are shown.

involved in hydrogen bonding through its sulfur atom) and proline residues. In rusticyanin it is likely that the two rigid aromatic rings are able to protect the histidine effectively from becoming protonated at low pH. Consequently, the 4-fold coordination of copper and the concomitant redox activity will be maintained even in an acidic environment.

The increased stability of the metal binding site may also arise from a tightly controlled hydrogen-bonding network comprising both indirect and direct interactions with the copper ligands. In our sequence comparison Ser86 replaces the strictly conserved asparagine residue found in this position for all the small blue copper proteins. This substitution would have a striking effect on the geometry of the copper site since it is adjacent to the first histidine ligand. The shorter side chain of Ser86 is able to form a hydrogen bond (Oγ) with the backbone NH of Glu139 to link the loop containing the triad of copper ligands near the C-terminal end (Figure 5a). This is in contrast to azurin or plastocyanin where the asparagine side chain (Oδ1 and Nδ2) interacts with two different polypeptide segments of the molecule (Figure 5b) forming three hydrogen bonds. It is known from hydrogen-bonding studies that, unlike Ser, the side chain of Asn tends to make rather "long-range" interactions which cross link more distant parts of the polypeptide chain (Baker & Hubbard, 1984). This feature is related to the relatively inflexible side chain of Asn (i.e., the conformational rigidity of the NH_2 group) which does not favor short-range hydrogen bonds. In contrast, the shorter but more flexible side chain of Ser86 may more easily adopt a conformation such that the two copper ligating parts of the polypeptide chain are held together more tightly. In addition, Ser86 is also able to interact directly with a copper ligand via a NH$-$S hydrogen bond between its backbone amino group and Cys138, an interaction also found in azurin, plastocyanin, or pseudoazurin. This bond not only stabilizes the orientation of the cysteine side chain but furthermore supports a compact binding of the two distinct polypeptide segments which provide all of the copper ligands (Figure 5a,b). The orientation of His85 is maintained by a hydrogen bond between the side chain Oδ1 of Asn80 and His85 Nϵ2 (Figure 5a). The latter hydrogen bond is analogous to the Glu49 Oϵ2$-$His53 Nϵ2 interaction in the vicinity of the copper binding pocket of amicyanin.

The protein fold used for the model of rusticyanin is a β-sandwich commonly found in blue copper proteins (Adman, 1991). It is known from early studies on the family of small copper-containing proteins that, despite low sequence identities (<30%), the topology of the protein fold is generally maintained and the residues in the vicinity of the copper center—the site with the minimal amino acid substitutions—form almost identical structures (Chothia & Lesk, 1982). Thus, it is not surprising that the copper site

of rusticyanin is a close structural analogue of other type 1 copper binding proteins. The stability of rusticyanin at acidic pH appears to arise from several stabilizing factors including hydrogen bonding from Ser86, aromatic rings in the hydrophobic patch, and the overall distribution of solvent exposed lysine residues.

Conclusion. The modeling study presented here suggests that rusticyanin possesses a β-sandwich structure typical of the single blue copper proteins, often associated with their high stability. The greater stability of the Cu site in rusticyanin must partly arise from a more protective nature of the site, as revealed by the molecular modeling. The higher hydrophobicity of this region compared to other blue Cu proteins may further contribute toward this stability. The high number of uncompensated lysine residues located on the surface of the rusticyanin model has direct implications for the functionality of rusticyanin in a highly acidic environment. The presence of a serine next to the ligating histidine instead of an asparagine may result in a tighter Cu cavity through tighter hydrogen bonding. This finding is consistent with relatively smaller structural changes observed in rusticyanin upon reduction compared to azurin and as such is an elegant explanation of the high redox potential of rusticyanin.

NOTE ADDED IN PROOF

A global fold of rusticyanin has been deduced using ^{15}N and ^{1}H NMR (Hunt et al., 1994).

ACKNOWLEDGMENT

We thank S.E.R.C. for funding and the Daresbury Laboratory for provision of the experimental and computational facilities. We would also like to thank Prof. E. Adman for helpful comments and for providing unrefined coordinates of nitrite reductase (NiR) including also its side chains and Dr. L. Murphy for assistance and interest in this project.

REFERENCES

Abola, A., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) in *Crystallographic Databases—Information Content, Software Systems, Scientific Applications* (Allen, F. H., Beregerhoff, G., & Sievers, R., Eds.) pp 107−132, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester.

Adman, E. T. (1991) *Adv. Protein Chem. 42*, 145−197.

Adman, E. T., Turley, S., Bramson, R., Petratos, K., Banner, D., Tsernoglou, D., Beppu, T., & Watanabe, H. (1989) *J. Biol. Chem. 264*, 87−99.

Baker, E. N. (1988) *J. Mol. Biol. 203*, 1071−1095.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, F. M., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Binsted, N., Strange, R. W., & Hasnain, S. S. (1992) *Biochemistry 31*, 12117−12125.

Blackburn, N. J., Strange, R. W., McFadden, L. M., & Hasnain, S. S. (1987) *J. Am. Chem. Soc. 109*, 7162−7170.

Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E., & Thornton, J. M. (1987) *Nature 326*, 347−352.

Brünger, A. T (1990) X-PLOR Manual, Version 2.1, Yale University, New Haven, CT.

Chothia, C., & Lesk, A. M. (1982) *J. Mol. Biol. 160*, 309−323.

Chothia, C., & Lesk, A. M. (1986) *EMBO J. 5*, 823−826.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45−148.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387−395.

Durley, R., Chen, L., Lim, L. W., Mathews, F. S., & Davidson, V. L. (1993) *Protein Sci. 2*, 739−752.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97−120.

Germanas, J. R., Di Bilio, A. J., Gray, H. B. & Richards, J. H. (1993) *Biochemistry 32*, 7698−7702.

Godden, J. W., Turley, S., Teller, D. C., Adman, E. T., Liu, M. Y., Payne, W. J., & LeGall, J. (1991) *Science 253*, 438−442.

Guss, J. M., Merritt, E. A., Phizackerley, R. P., Hedman, B., Murata, M., Hodgson, K. O., & Freeman H. C. (1988) *Science 241*, 806−811.

Guss, J. M., Bartunik, H. D., & Freeman, H. C. (1992) *Acta Crystallogr. B48*, 790−811.

Hasnain, S. S., & Strange, R. W. (1990) *Biophysics and Synchrotron Radiation.* (Hasnain, S. S., Ed.) pp 104−122, Ellis Horwood Ltd., Chichester, U.K.

Holt, S. D., Piggott, B., Ingledew, W. J., Feiters, M. C., & Diakun, G. P. (1990) *FEBS Lett. 269*, 117−121.

Hunt et al. (1994) *J. Mol. Biol. 244*, 370−384.

Ingledew, W. J., Cox, J. C., & Halling, P. J. (1977) *FEMS Microbiol. Lett. 2*, 193−197.

Karlsson, B. G., Nordling, M., Pascher, T., Tsai, L.-C., Sjolin, L., & Lundberg, L. G. (1991) *Protein Eng. 4*, 343−349.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Lewis, C. A., Lappin, A. G., & Ingledew, W. J. (1984) *Biochem. Soc. Trans. 12*, 503.

Lommen, A., Pandya, K. I., Koningsberger, D. C., & Canters, G. W. (1991) *Biochim. Biophys. Acta 1076*, 439.

Messerschmidt, A., Ladenstein, R., Huber, R., Bolognesi, M., Avigliano, L., Petruzzelli, R., Rossi, A., & Finazzi-Agró, A. (1992) *J. Mol. Biol. 224*, 179−205.

Murphy, L. M., Strange, R. W., Karlsson, G., Lundberg, L., Pascher, T., Reinhammar, B., & Hasnain, S. S. (1993) *Biochemistry 32*, 1965−1975.

Nunzi, F., Woudstra, M., Campèse, Bonicel, J., Morin, D., & Bruschi, M. (1993) *Biochim. Biophys. Acta 1162*, 28−34.

Pascher, T., Karlsson, B. G., Nordling, M., Malmström, B. G., & Vänngård, T. (1993) *Eur. J. Biochem. 212*, 289−296.

Pendry, J. B. (1974) *Low Energy Electron Diffraction*, Academic Press, New York.

Penner-Hahn, J. E., Murata, M., Hodgson, K. O., & Freeman, H. (1989) *Inorg. Chem. 28*, 1826.

Petratos K., Dauter, Z., & Wilson, K. S. (1988) *Acta. Crystallogr. B 44*, 628−636.

Romero, A., Hoitink, C. W. G., Nar, H., Huber, R., Messerschmidt, A., & Canters, G. W. (1993) *J. Mol. Biol. 229*, 1007−1021.

Ronk, M., Shively, J. E., Shute, E. A., & Blake, R. C., II (1991) *Biochemistry 30*, 9435−9442.

Scharf, B., & Engelhard, M. (1993) *Biochemistry 32*, 12894−12900.

Scott, R. A., Hahn, J. E., Doniach, S., Freeman, H. C., & Hodgson, K. O. (1982) *J. Am. Chem. Soc. 102*, 168.

Strange, R. W., Blackburn, N. J., Knowles, P. F., & Hasnain, S. S. (1987) *J. Am. Chem. Soc. 109*, 7157−7162.

Strange, R. W., Reinhammar, B., Murphy, L. M., & Hasnain, S. S. (1995) *Biochemistry 34*, 220−231.

Strong, C., Harrison, S. L., & Zeger, W. (1994) *Inorg. Chem. 33*, 606−608.

van de Kamp, M., Silvistrini, M. C., Brunori, M., van Beeumen, J., Hall, F. C., & Canters, G. W. (1990) *Eur. J. Biochem. 194*, 109−118.

Vakoufari, E., Wilson, K. S., & Petratos, K. (1994) *FEBS Lett. 347*, 203−206.

Yano, T., Fukumori, Y., & Yamanaka, T. (1991) *FEBS Lett. 288*, 159−162.

BI942591+