Acta Cryst. (1995). D**51**, 1052–1064

Structure of a New Azurin from the Denitrifying Bacterium *Alcaligenes xylosoxidans* at High Resolution

BY FRASER E. DODD AND S. SAMAR HASNAIN*

Molecular Biophysics Group, Daresbury Laboratory, Warrington, WA4 4AD, England, and De Montfort University, Leicester, LE1 9BH, England

AND ZELDA H. L. ABRAHAM, ROBERT R. EADY AND BARRY E. SMITH

Nitrogen Fixation Laboratory, University of Sussex, Brighton BN1 9RQ, England

(*Received* 10 *March* 1995; *accepted* 8 *June* 1995)

Abstract

It has been reported previously that *Alcaligenes xylosoxidans* (NCIMB 11015) grown under denitrifying conditions produces two azurins instead of the single previously identified azurin [Dodd, Hasnain, Hunter, Abraham, Debenham, Kanzler, Eldridge, Eady, Ambler & Smith (1995). *Biochemistry*. In the press]. The new azurin, called azurin II, has been crystallized as blue elongated rectangular prisms with the tetragonal space group $P4_122$ and unit-cell parameters $a = b = 52.65$, $c = 100.63$ Å. X-ray crystallographic data extending to 1.9 Å resolution were collected by the Weissenberg method using 200 × 400 mm image plates and synchrotron X-rays of wavelength 0.97 Å. The three-dimensional structure of azurin II has been solved by the molecular-replacement method using the structure of azurin from *Alcaligenes denitrificans* NCTC 8582 with which this new azurin shows a close homology. The quality of the initial map was sufficient to predict a number of sequence differences. The model is currently refined to an R factor of 18.8% with X-ray data between 8.5 and 1.9 Å. The final model of 961 protein atoms, one Cu atom and 50 water molecules has r.m.s. deviations from ideality of 0.009 Å for bond lengths and 1.7° for bond angles. The overall structure is similar to that of the azurin from *A. denitrificans* NCTC 8582. It has a β-barrel structure with the Cu atom located near the top end of the molecule. The Cu atom is coordinated to Nδ of His46 and His117 at 2.02 Å and to Sγ of Cys112 at 2.12 Å, while the carbonyl O atom of Gly45 and Sδ atom of Met121 provide the additional interactions at 2.75 and 3.26 Å, respectively.

Introduction

In the denitrifying bacterium *Alcaligenes xylosoxidans* (formerly known as *Alcaligenes sp.* NCIMB 11015, *Achromobacter xylosoxidans* or *Pseudomonas denitrificans*) the reduction of nitrite to nitric oxide is accomplished by a blue-copper-containing nitrite reductase (Abraham, Lowe & Smith, 1993, Grossmann *et al.*, 1993, Strange *et al.*, 1995). Two classes of blue-copper proteins, pseudoazurin and azurin have so far been implicated in donating electrons to nitrite reductase: pseudoazurin to the green nitrite reductases of *Alcaligenes faecalis* S-6 and *Achromobacter cycloclastes* (Kakutani, Watanabe, Arima & Beppu, 1981; Liu, Liu, Payne & LeGall, 1986) and azurin to the blue enzyme of *Pseudomonas aureofaciens* (Zumft, Gotzmann & Kroneck, 1987). However, a 'blue-copper protein' (probably azurin I) isolated from *A. xylosoxidans* was shown not to function as an electron donor to the blue nitrite reductase isolated from *A. xylosoxidans*, but cytochrome c_{552} from this organism did function in a coupled assay using yeast lactate dehydrogenase as a primary source of electrons (Miyata & Mori, 1969).

The X-ray crystal structures of two bacterial azurins have been reported, one from *P. aeruginosa* (AzP) (Adman & Jensen, 1981, Nar, Messerschmidt, Huber, van de Kamp & Canters, 1991*b*) and the other from *A. denitrificans* (AzAD) (Norris, Anderson & Baker, 1983; Baker, 1988). The availability of site-directed mutants of both of these azurins has generated new interest in these model electron-transfer proteins and has led to a number of structural studies (Nar, Messerschmidt, Huber, van de Kamp & Canters, 1991*a*; Farver *et al.*, 1993; Murphy, Hasnain, Strange, Harvey & Ingledew, 1993; Romero *et al.*, 1993). An azurin from *A. xylosoxidans* has been crystallized (Strahs, 1969) and was in fact the first copper protein to be crystallographically characterized (cell dimensions and space groups were determined). Despite several attempts, crystallization (Norris, Anderson, Baker & Rumball, 1979) and structure determination of this protein has remained problematical; only the tertiary structure at 3.0 Å resolution has appeared from a solution based on anomalous dispersion (Korzun, 1987). No further details of this work have emerged. We have recently shown that *A. xylosoxidans* is capable of synthesising two biochemically distinct azurins, and that both of these proteins can donate electrons to purified nitrite reductase (Dodd *et al.*, 1995). One of these is the same

* To whom correspondence should be addressed.

azurin as that purified and sequenced by Ambler (1971) while the other, which we have called azurin II, shows a closer similarity to the azurin from *A. denitrificans*. There are precedents for the synthesis of two azurins in a single bacterial species. For example *Methylomonas J* synthesizes two distinct azurins, only one of which is able to function as the primary acceptor to methylamine dehydrogenase (Ambler & Tobari, 1989). In the present case, the crude extracts of *A. xylosoxidans* have copper added to them to activate nitrite reductase (Abraham *et al.*, 1993). However, we also observe both azurin I and azurin II when crude extracts have not been activated by copper. It is not clear why only a single species of azurin in *A. xylosoxidans* has been observed before (Suzuki & Iwasaki, 1962). Both azurins are able to donate electrons to nitrite reductase and have identical redox potentials (Dodd *et al.*, 1995). The amino-acid sequence of azurin II differs from that of azurin I in 43 out of the 129 residues and thus has an identity of 67% to azurin I.

The crystallographic structure of azurin II from *A. xylosoxidans* is presented here. We note that the space group and cell dimensions of azurin II ($a = b = 52.65$, $c = 100.63$ Å) are similar to those published by Strahs (1969) ($a = b = 53.2$, $c = 101$ Å) and Korzun (1987) ($a = b = 52.3$, $c = 99.3$ Å). During the preparation of this manuscript, a preliminary report has appeared on the 2.5 Å structure of an azurin whose N-terminal sequence is the same as azurin II (Inoue *et al.*, 1994). The high-resolution crystallographic refinement of azurin II reported here with the correct amino-acid sequence provides a basis for detailed comparison with the high-resolution structures of AzP and AzAD, the only other two azurins for which high-resolution refined structures are available.

Experimental

Crystallization and preliminary characterization

Two azurins, azurin I and azurin II, were extracted in a yield ratio of 1:2 from *A. xylosoxidans* as described previously (Dodd *et al.*, 1995). Crystals of azurin II were grown using the hanging-drop technique. In a typical experiment, 3 µl of buffered protein solution (10 mg ml^{-1}) was mixed with 3 µl of reservoir buffer over 1 ml reservoirs. Azurin II crystals were grown from 0.1 *M* KH_2PO_4 pH 6.0 over a reservoir containing 620 µl saturated ammonium sulfate (aqueous) and 380 µl 0.1 *M* KH_2PO_4 pH 6.0. At 277 K single crystals were obtained within 7 d; these form elongated rectangular prisms. The largest crystals were 0.1 × 0.1 × 0.5 mm. These were shown to have the unit-cell parameters as $a = b = 52.65$, $c = 100.63$ Å, and the space group was considered to be $P4_122$ or $P4_322$ based upon the systematic absences (Dodd *et al.*, 1995).

Table 1. *Summary of the quality and completeness of data as a function of resolution*

Experimental	
X-ray source	BL6A-2 photon factory, KEK
Camera	Sakabe's Weissenberg camera
Detector	Image plate (200 × 400 mm), and BA100 system
XTF	429.7 mm
Collimator	0.1 mm
Wavelength (Å)	0.97
Coupling constant (° mm^{-1})	2.0
Rotation axis	*c*
No. of crystals	1
Oscillation per sheet (°)	6
Total oscillation angle (°)	94
Total exposure time (min)	45

Data reduction

Resolution range (Å)	No. of reflections observed	No. of independent reflections	Expected No. of reflections	R_{sym} (%)	Completeness (%)
∞ >6.53	4611	352	357	3.8	99.7
6.54–4.64	4477	551	566	3.2	99.1
4.64–3.79	5251	660	690	3.1	95.6
3.79–3.29	5315	669	810	3.5	84.1
3.29–2.94	5228	667	885	4.6	75.4
2.94–2.69	4412	653	968	6.2	67.4
2.69–2.49	3197	643	1049	6.9	61.3
2.49–2.33	2759	648	1121	7.8	57.8
2.33–2.19	2474	600	1181	7.9	50.8
2.19–2.08	2338	624	1258	10.3	49.6
2.08–1.98	1862	588	1245	10.9	44.8
1.98–1.90	1613	555	1357	14.8	40.9

Total No. of measured reflections	43545
Total No. of independent reflections (includes anomolous)	13178
R_{merge} (%)	4.0
Completeness (%)	61.8

X-ray data collection

A high-resolution diffraction data set extending to 1.9 Å was collected using the Weissenberg camera (Sakabe, 1991) at beamline 6A-2 at the Photon Factory, using an X-ray wavelength of 0.97 Å with a crystal to image plate radius of 429.7 mm. Data were collected at 287 K with crystal mounted with the spindle about its *c* axis. Rotation angles of 6° were used with 0.5° overlap and coupling constant of 2° mm^{-1}. A collimator of 0.1 mm was used to define the X-ray beam. The crystal was translated after five to six exposures and thus the data was collected on a fresh part of the crystal before any significant deterioration in diffraction was observed. The higher speed of data collection offered by the Weissenberg method coupled with an off-line image-plate scanner system made it possible to obtain the whole data set from one crystal. 90° of data were collected in 90 min. The storage ring beam current was 353 and 342 mA at the start and end of the data collection, respectively.

The data were reduced using *Weis* (Higashi, 1989) then scaled and merged using the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994), giving a merging *R* factor of 4.0% to 1.9 Å for 43 645 reflections with 13 178 independent reflections. These data have a completeness of 62% to 1.9 Å and 79% to 2.5 Å.

The quality and completeness of data as a function of resolution are summarized in Table 1. A single image plate was used which was offset to obtain higher resolution data. This coupled with the smaller width of the image plate resulted in a lower completeness of the data. Furthermore, as the crystal could only be mounted along the *c* axis due to its morphology, the completeness of higher resolution data is further reduced. Despite these limitations, we note that in the 2.19–2.08 Å shell, approximately 50% of the data are recorded with good statistics ($R_{sym} \simeq 10\%$). A Wilson plot (Wilson, 1949) is shown in Fig. 1, from which an estimate of 13.3 Å^2 for the overall temperature factor was obtained.

Structure determination

Molecular replacement

The structure of azurin II was solved by molecular replacement. Briefly, the coordinates of the refined model of azurin of *A. denitrificans* (Baker, 1988), obtained from the Protein Data Bank at Brookhaven (Bernstein *et al.* 1977; accession number 2AZA), were used as a search model to obtain a molecular-replacement solution. Initial attempts of molecular replacement were made with the *MERLOT* package (Fitzgerald, 1988). Cross-rotation functions at various resolution ranges gave the unambiguous orientation of the molecule (highest peak was selected). The translational searches were attempted in both possible space groups, $P4_122$ and $P4_322$. Both space groups gave equally good solutions, which were very distinct. Rigid-body refinement for one of the solutions in space group $P4_122$ resulted in an *R* factor of 35.8% with X-ray data from 8.0 to 3.0 Å, a value one might expect from the correct solution. However, due to the difficulty in not being able to distinguish between the two space groups, which should have been straightforward at the translation stage, it was decided to approach the molecular-replacement solution using the *AMoRe* (Navaza, 1994) package. The application of this method resolved this ambiguity very clearly and identified $P4_122$ as the correct space group. In this space group, the solution had an *R* factor of 31.5% while in $P4_322$ the top solution had an *R* factor of 44.9%. This is clearly evident from Table 2 and the Harker sections shown in Fig. 2.

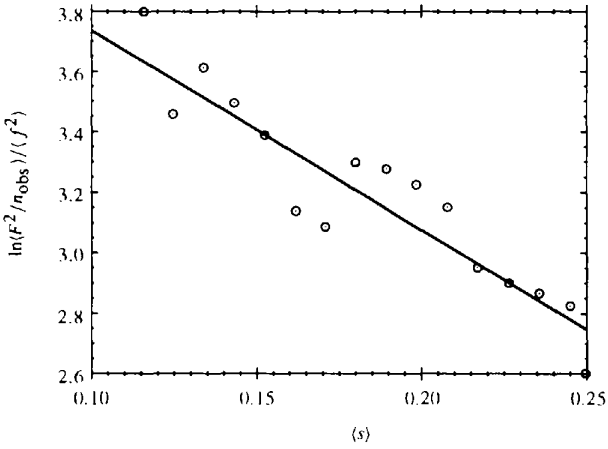


Fig. 1. A Wilson plot (Wilson, 1949) for azurin II based upon 3772 reflections in the range 3.0–2.0 Å. The best straight-line fit gives a *B* value of 13.3 Å^2.

Table 2. *Molecular-replacement solution details*

Rotation-function solution

Peak height = 19.3
(only one solution above 50% of this height)
α = 49.34
β = 66.34
γ = 211.32

Translation-function solution (top three solutions)

$P4_122$

	tx	ty	tz	*R* factor (%)	Correlation
1	0.200	0.270	0.184	31.9	75.2
2	0.200	0.296	0.403	36.9	43.9
3	0.200	0.269	0.151	46.7	43.6

$P4_322$

	tx	ty	tz	*R* factor (%)	Correlation
1	0.200	0.271	0.433	45.4	44.9
2	0.198	0.270	0.183	46.3	42.7
3	0.203	0.763	0.181	48.5	36.4

Rigid-body refinement

$P4_122$

α	β	γ	tx	ty	tz	*R* factor (%)	Correlation
49.02	65.44	211.67	0.200	0.271	0.184	31.5	75.9

$P4_122$

α	β	γ	tx	ty	tz	*R* factor (%)	Correlation
51.58	66.23	210.64	0.202	0.270	0.185	44.5	49.8

A $2F_o - F_c$ map was generated using the *CCP*4 suite of programs to 2.5 Å and an initial model was built. The quality of this initial electron-density map was sufficiently good to recognise the amino-acid differences relative to azurin from *A. denitrificans* at positions 7, 20, 48, 103 and 104 before the information from amino-acid sequence analysis became available (Dodd *et al.*, 1995). The model was built using the program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991) on an IRIS 4000. The density for some of the side chains, particularly lysines, was fragmented, therefore these were replaced by alanines or serines until the density improved.

Refinement

The refinement of the azurin II structure was carried out in several steps by a combination of positional refinement and simulated annealing with the program *X-PLOR*, version 3.1 (Brünger, 1992*a*,*b*) implemented on a Silicon Graphics Challenge machine. Standard topology and parameter files were employed in the refinement of the protein. In the final stages of positional refinement, the Cambridge library was used as has been suggested by Engh & Huber (1991). A charge of +2 was used for Cu with energy constraints of 80 kcal mol Å^{-2} (335 kJ mol Å^{-2}) for the three strong

Cu ligand (Nδ1 His46, His117 and Sγ Cys112) and 30 kcal mol Å^{-2} (125 kJ mol Å^{-2}) for the Cu–Sδ Met121 interaction. No other restraints were applied to these ligands. Cu was included as the anomolous scatterer with f'' = 2.168 at 0.97 Å. Table 3 summarizes the course of the refinement. R free (Brünger, 1992*b*, 1993) was calculated throughout the refinement using 10% of the reflections.

The starting model for the refinement was built using *O* based on the Cα positions of AzAD. The model was built using the *LEGO* commands using the full azurin II sequence. The model was then fitted to the density. Where density was not visible side chains were truncated, this was the case for 12 out of the 129 residues, of which half were lysines. An overall B factor of 15 Å^2 was used with all atoms in the model having a full occupancy.

The refinement was carried out by successive cycles of simulated annealing and finally by positional refinement. A cycle of simulated annealing (SA) consisted of 40 cycles of Powell minimization (Powell, 1977) followed by a slow-cooling run from 4000 to 300 K in steps of 25 K, each of which had 50 time steps of 0.5 fs, the total time being 3.7 ps. This was followed by 120 cycles of Powell minimization.

Step 1. Two models were used in the initial refinement; one built using maps with data to 3.0 Å and the other with data to 2.5 Å. The models built using these maps were subject to two cycles of SA using data in the ranges 7.0–3.0 Å and 7.0–2.5 Å, respectively. With each of the refined models, the data were extended to 2.0 Å and two cycles of SA carried out. Both models converged to essentially the same refinement with very similar R factors. The Cu–ligand distances were essentially the same in the two refinements.

Step 2. The model obtained from the 2.0 Å refinement of the initial 2.5 Å model (cycles 1*b* and 1*d*) was used with data to 1.9 Å which provided us with

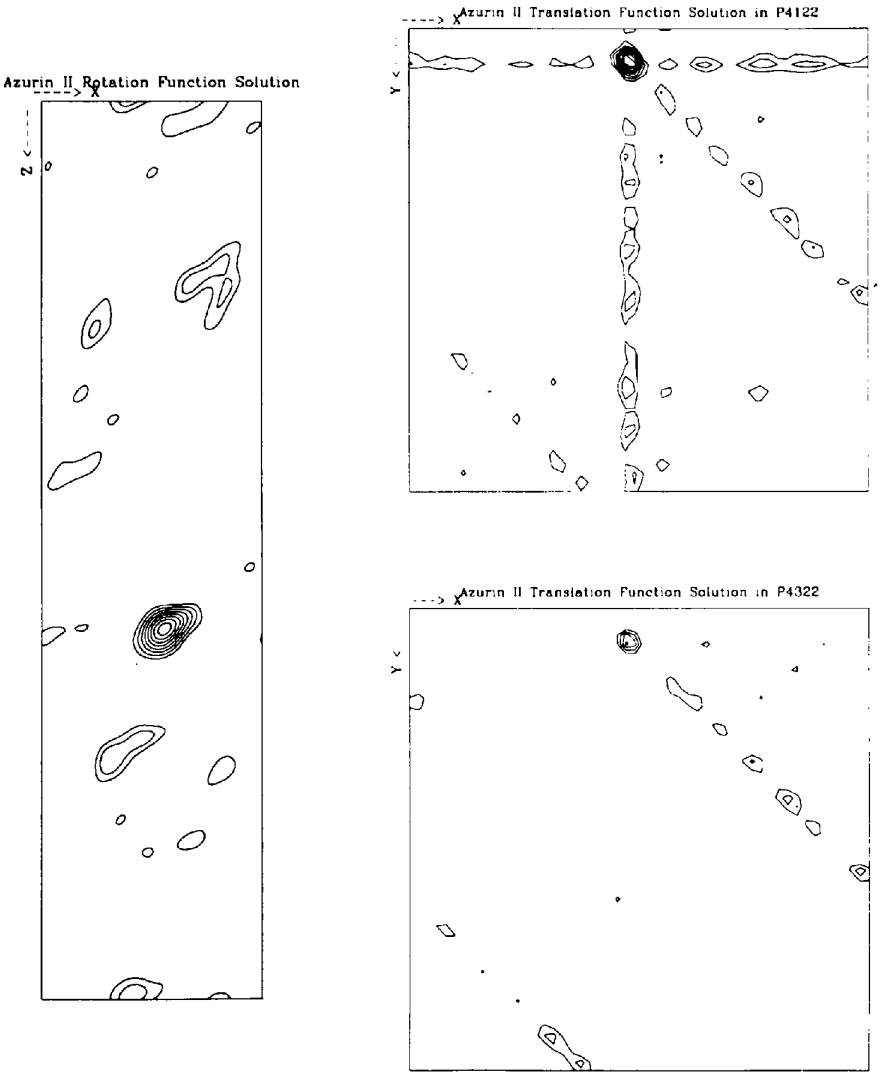


Fig. 2. The rotation- and translation-function solution for azurin II. The translation-function maps are contoured at the same levels. These clearly show that $P4_122$ is the correct space group.

Table 3. *Course of refinement*

Cycle	*R* value	*R* free value	Atoms	Resolution range (Å)	Comments
1*a*	32.1 → 20.1	32.1 → 29.3	918	7.0–3.0	
1*b*	33.7 → 23.0	33.4 → 29.0	918	7.0–2.5	
1*c*	26.0 → 24.8	29.3 → 31.2	918	7.0–2.0	From cycle 1*a*
1*d*	26.0 → 24.7	29.0 → 30.5	918	7.0–2.0	From cycle 1*b*
2	35.5 → 25.6	35.6 → 30.4	918	7.0–1.9	From cycle 1*d*
3	26.6 → 24.9	31.4 → 29.9	923	7.0–1.9	Some side chains added
4*a*	26.2 → 24.8	30.4 → 29.8	938	7.0–1.9	
4*b*	24.8 → 23.4	29.8 → 27.8	962	7.0–1.9	Weak data excluded
5	24.9 → 21.9	27.8 → 27.3	1014	7.0–1.9	52 waters included
6	21.3 → 19.2	27.9 → 23.7	1010	8.5–1.9	4 waters removed, *B* refinement
7	20.9 → 17.5	21.3 → 20.7	1037	8.5–1.9	27 waters added, data extended
8	20.1 → 18.8	21.1 → 20.7	1012	8.5–1.9	23 waters removed

Refinement parameters	
No. of reflections	12162
No. of parameters	4048
R.m.s. deviations	
Bond distances (Å)	0.009
Bond angles (°)	1.7
Mean *B* factors	
Main chain (Å^2)	15.9
Side chains (Å^2)	18.9
Waters (Å^2)	25.5

12 667 reflections and (918 × 3) parameters. An SA run yielded an *R* factor of 25.6% with *R* free of 30.4%. This procedure was repeated which resulted in a small improvement. The quality of the model was assessed using *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993) at this stage before proceeding with further refinement.The Ramachandran plot (Ramachandran, Ramakrishnan & Sasisekharan, 1963; Ramachandran & Sasisekharan, 1968) showed 87% of the residues in the most favoured regions with only Gln2 and Val17 in the generously allowed regions.

Step 3. The model was carefully examined using *O* at this stage. Truncated residues were replaced in those cases where density had improved. This was followed by a cycle of SA. After another examination with *O*, density for all of the remaining truncated residues could be seen except for Lys27, Gln28, Lys34 and Lys56. A cycle of SA followed. The Ramachandran plot now showed that 89% of the residues were found in the most favoured regions, with the remainder in the additionally allowed regions. The *R* factor was 24.9%.

Step 4. A further cycle of rebuilding of suspect side chains and SA was carried out. At this stage, it was noted that some 200 reflections which corresponded to the weakest data (lowest two bins) had an R_{merge} an order of magnitude worse than the rest. These were removed. 40 cycles of positional refinement led to a significant reduction in the *R* factor to 23.4%.

Step 5. At this stage water molecules were chosen using the $F_o - F_c$ map with the *CCP*4 *Peakmax* and *Watpeak* programs (Collaborative Computational Project, Number 4, 1994). A total of 200 potential waters in the range of 1.4–5 Å from the protein atoms were examined using *O* and a total of 52 water molecules which corresponded to well defined $F_o - F_c$ density were selected for their inclusion in the model. A cycle of SA was carried out.

Step 6. The model was examined using *O*. Four water molecules which had moved out of density were deleted from the model. The remaining side chains were built. Positional refinement was carried out. This was followed by a grouped *B*-factor refinement. The groups were the main chain, side chains, water and the Cu. The resulting grouped *B* values were 13.8, 16.1, 23.3 and 13.0 Å^2. A cycle of SA was carried out followed by individual atom restrained *B* refinement resulting in an *R* factor of 19.2%.

Step 7. The model was examined in detail at this stage. Three water molecules were deleted, but 27 new water molecules were added using the $2F_o - F_c$ map. Two different refinements were carried out; one with the original data range and the other with data extended to 8.5 Å. Both of the refinements led to essentially similar *R* factors. Further refinement proceeded with data from 8.5 to 1.9 Å with 12 162 reflections. A cycle of SA followed. A minimization with the Cambridge database library as suggested by Engh & Huber (1991) resulted in further improvement in the *R* factor and subsequent refinements were carried out using this library.

Step 8. A detailed inspection was undertaken of residues which showed significant departure from ideal geometry The previous cycle of SA had moved 23 water molecules to distances >5 Å from the protein, these were removed. A cycle of positional refinement was carried out with overall *B* values for protein atoms and water molecules set to 20 and 30 Å^2, respectively. This was followed by a cycle of *B* refinement and positional refinement. The exclusion of 23 water molecules did not result in an increase in *R* free. The final model consisted of 961 protein atoms and 50 water molecules and one Cu atom. The final *R* factor for 12 162 observations

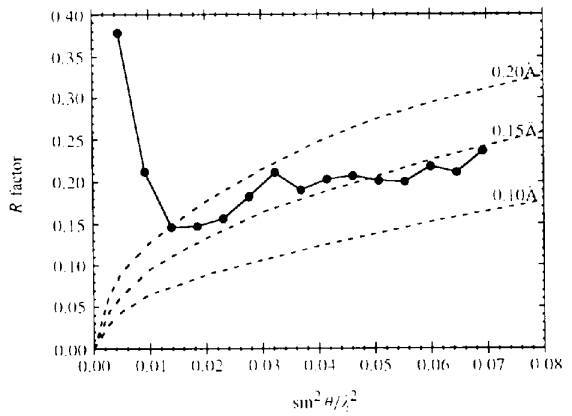
between 8.5 and 1.9 Å was 18.8% with an R free of 20.7%.

Results

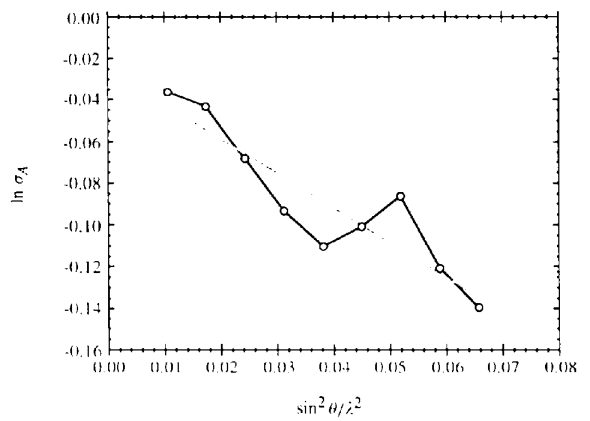
Quality of the structure

A number of criteria have been applied here to assess the quality of the final model. The structure of azurin II is determined and refined with data up to 1.9 Å resolution with a small number (50) of water molecules to a moderately good R factor of 18.8%. Furthermore, only a slightly larger value of 20.7% for the 'free R factor', the so-called R free or an unbiased R factor (Brünger 1992*b*, 1993) calculated by excluding 10% of the reflections from refinement, suggests the reliability of the overall structure. The R factor remained the same when these reflections were included in the final refinement.

A global estimate for the accuracy of a structure can be obtained in several ways. A plot, Fig. 3(*a*), of R factor against resolution calculated by the method of Luzzati (1952), provides an average error of 0.15 Å in atomic position. A σ_A plot (Read 1986) of $\ln(\sigma_A)$ as a function of $\sin^2\theta/\lambda^2$ is shown in Fig. 3(*b*). The slope of the σ_A plot indicates an average error of 0.25 Å. From this, we suggest that the average error in atomic positions for the structure is between 0.15 and 0.25 Å, with well ordered regions *i.e.* with low B, tending to have errors no more than 0.15 Å and in some cases significantly better than



(*a*)



(*b*)

Fig. 3. (*a*) Plot of the variation of R factor with resolution. The theoretical variations (Luzzati, 1952) of 0.10, 0.15 and 0.20 Å coordinate error with R factor are shown in dashed lines. (*b*) Plot of the variation of $\ln(\sigma_A)$ with resolution. The slope of the line of best fit gives an estimate of the r.m.s. coordinate error.

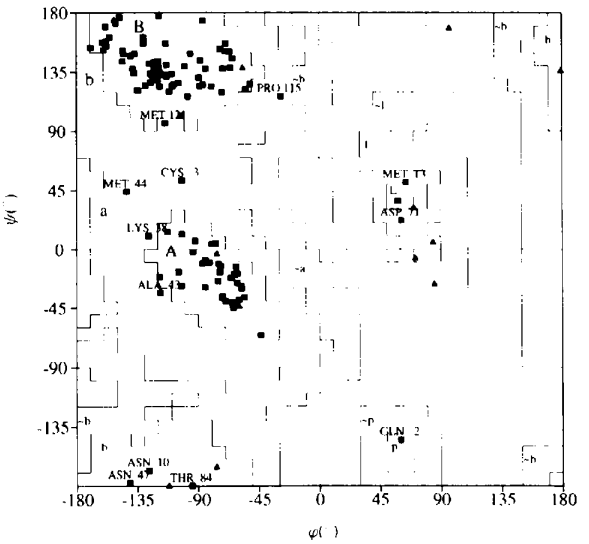


Fig. 4. A Ramachandran plot (Ramachandran, Ramakrishnan, & Sasisekharan, 1963; Ramachandran & Sasisekharan, 1968) of the main-chain torsion angles for the final refined model of azurin II. The plot was calculated using *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993). Glycines are shown as triangles and the other residues as squares. Of the 108 non-glycine and non-proline residues, 97 (90%) were in the core regions, and ten (9%) are in the additionally allowed regions. A single residue, Gln2, is in a generously allowed region. All residues outside the core regions are labelled.

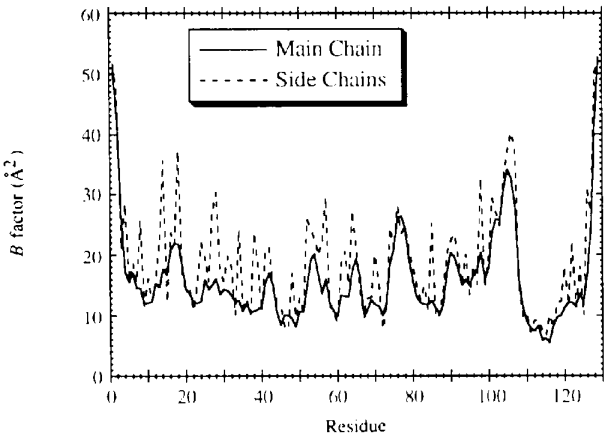


Fig. 5. A B-factor plot showing the averaged isotropic temperature factors of the main chain (solid line) and side chains (dashed line) of the final refined model of azurin II. The two highest non-termini peaks in the main-chain plot correspond to the outside of the flap region (76–78) and to a loop region (102–105). These regions are shown in green in Fig. 6.

that. The higher *B*-value regions are likely to have larger errors.

In addition to the assessment of the structure through an overall quality factor, it is important to assess the accuracy and reliability of different parts of the structure. The geometry of the main chain and side chains have been analysed using the program *PROCHECK* (Laskowski *et al.*, 1993). Ramachandran plots (Ramachandran *et al.*, 1963; Ramachandran & Sasisekharan, 1968) of the dihedral angles φ/ψ are given in Fig. 4. The majority of non-glycine residues, 91%, are in the 'most favoured' regions, defined as the core regions, with none of the residues in the disallowed regions. The main-chain and side-chain properties of the structure, as detailed in Laskowski *et al.* (1993), fall either in the 'Inside' or 'BETTER' categories indicating that these are in good agreement with other well refined structures. A further guide to the quality of the structure

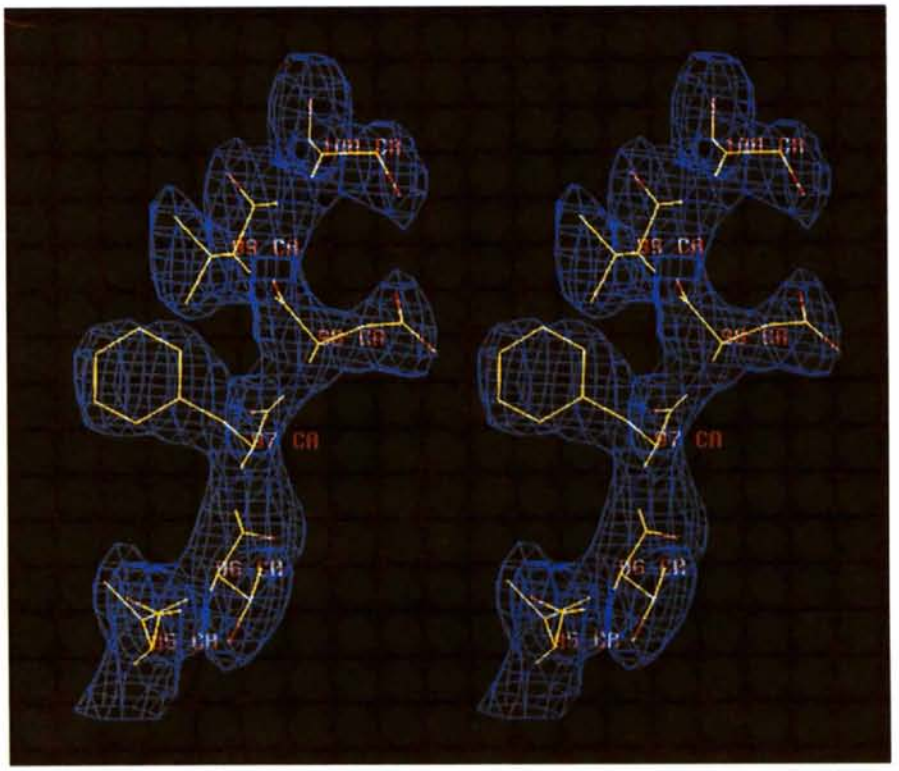


Fig. 6. Electron density for peptide region 95–100.

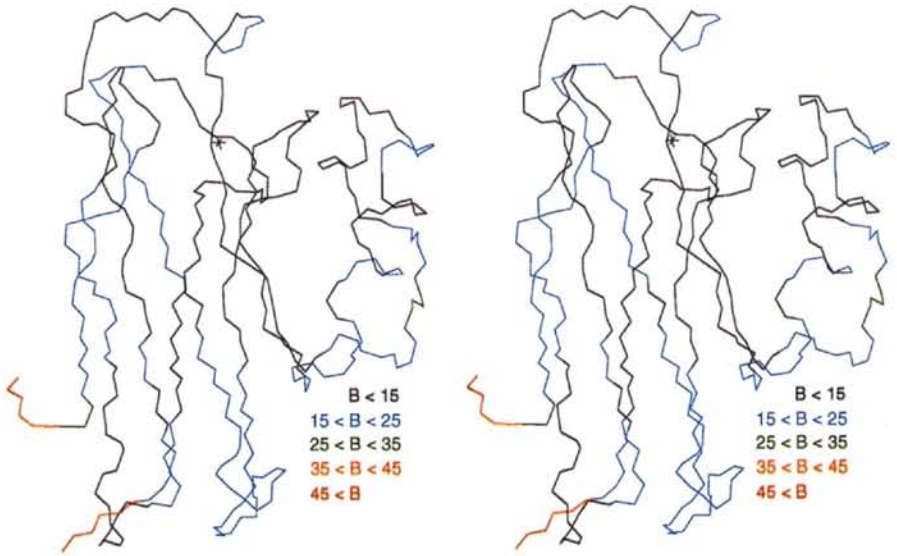


Fig. 7. A stereo main-chain plot of azurin II colour coded with Co *B* factors.

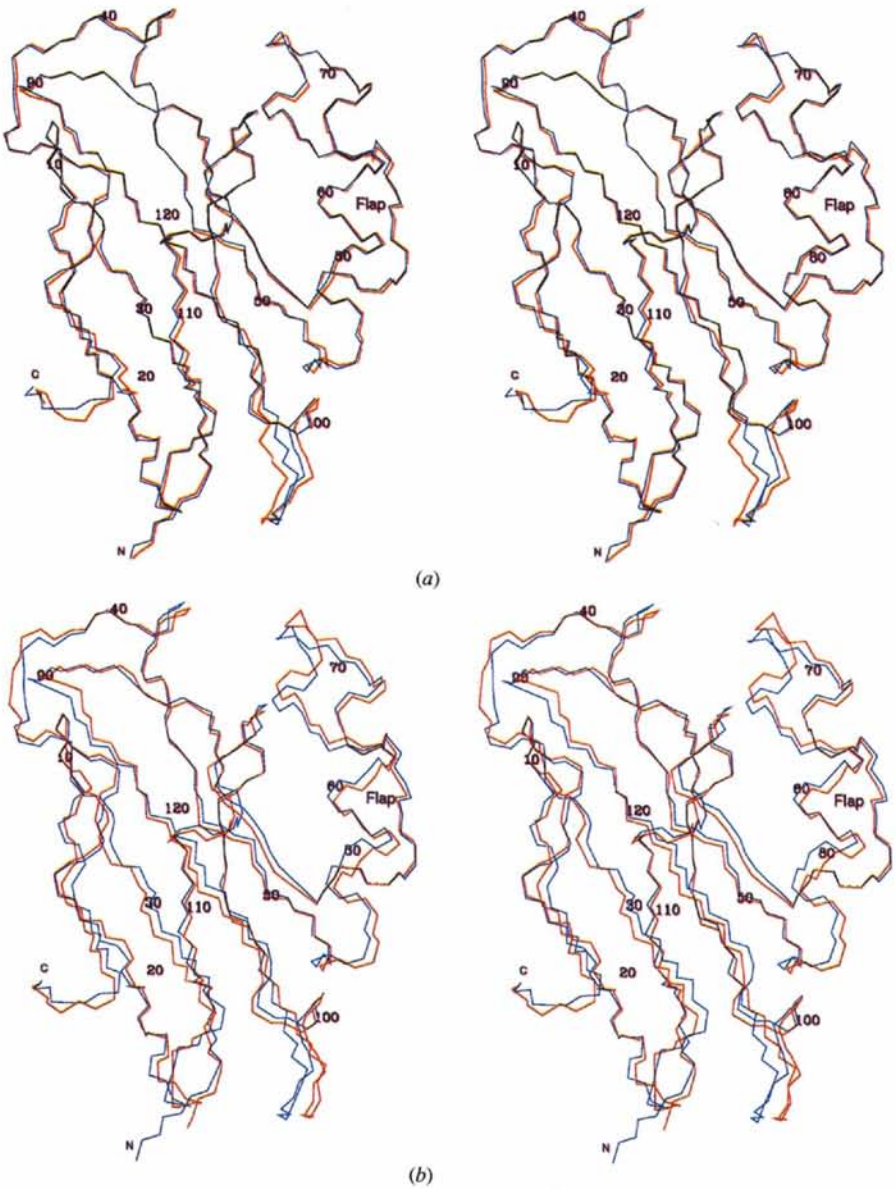


Fig. 8. (*a*) A stereo main-chain plot of azurin II (blue) with AzAD (red) with the Cα's superposed. The only significant departure occurs in the loop region around residues 101–108. (*b*) A stereo main-chain plot of azurin II (blue) with AzP (magenta) with the Cα's superposed.

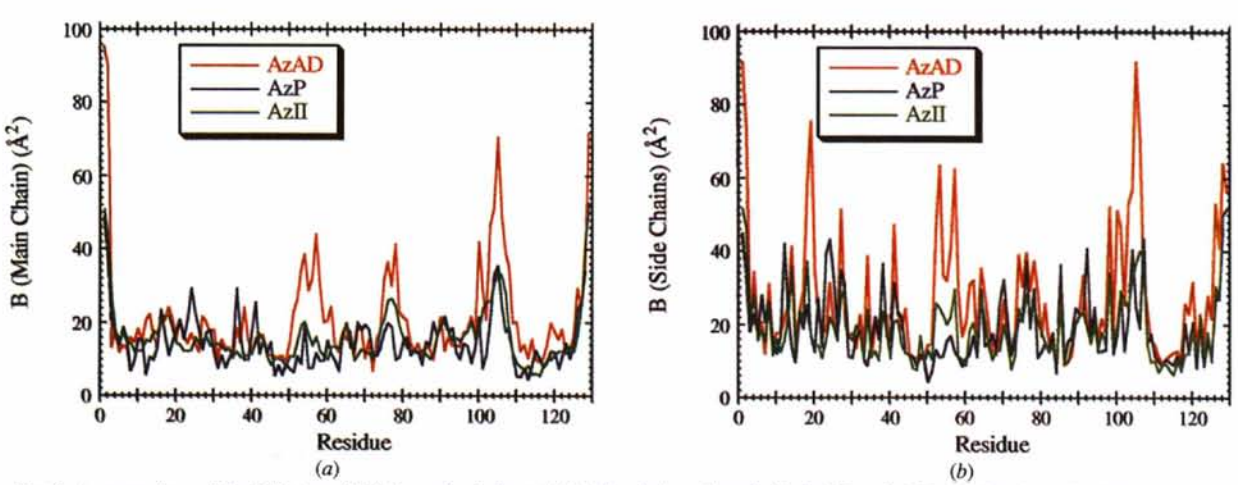


Fig. 9. A comparison of the *B* factors of (*a*) the main chain and (*b*) side chains of azurin II, AzAD and AzP as a function of residue number.

in different regions is provided by an inspection of atomic B values, which encompasses not only the static and dynamic disorder but an incorrectly positioned atom will show a particularly high B value. The most poorly defined residues are often identified with high B values, $B > 50$ Å^2. Fig. 5 shows the B values for the main-chain and side-chain atoms. The two termini residues, Ala1 and Ser129, show B factors of 51 and 53 Å^2, respectively. In addition, as would be expected, the loop 104–107 shows a B value in excess of 30 Å^2, about double that of the mean B value of 15.9 Å^2 for the main-chain atoms. The mean B value for the side chains is 18.9 Å^2 while the mean B value for the solvent is 35 Å^2 with only five waters with B values of around 55 Å^2. The well behaved nature of B values suggest that the low completeness of high-resolution data has not caused serious anisotropy in the atomic coordinates. The quality of the structure and resolution can be further assessed from the quality of the map. Fig. 6 shows the electron density for residues 95–100; the density around the Cu site is shown in the section on Cu site in Fig. 10. The overall quality of the final map is what would be expected from a high-resolution (~2 Å) structure. Indeed, the quality of density at the Cu site is very similar to what has been published for AzAD, AzP and their mutants, all of which have been determined to a resolution better than 2 Å.

Overall structure

The overall structure of azurin II is very similar to the two refined structures of azurins, AzAD (Baker, 1988) and AzP (Nar *et al.*, 1991*a*,*b*). A stereo picture where colour coding of the main-chain B factor has been included is shown in Fig. 7. Like the other azurins, the azurin II molecule consists of eight β-strands, which form a β-barrel structure. Azurin II shows a high degree of homology *i.e.* 89 and 60% with both AzAD and AzP, respectively. The close similarity of the structure of azurin II with AzAD and AzP is displayed in Fig. 8. The r.m.s. fits to main-chain atoms of AzAD and AzP give differences of 0.36 and 0.70 Å, respectively. The only significant differences between azurin II and AzAD are near the termini and in part of strand 7 from residues 101 to 108. The differences with AzP are more complex and occur in several places including the C-terminus and residues 101–108. The relative spacings between different strands differ significantly. The flap region which is distant from the Cu site also differs significantly. The difference between AzP and AzAD in the 101–104 segment was suggested to arise from the lack of a proline at position 104 in AzP, causing the loop to take a different position in relation to neighbouring chains. Azurin II also lacks this proline, and the particular loop takes an intermediate position between that of the AzP and AzAD.

Fig. 9 shows a comparison of B factors for azurin II, AzP and AzAD. Each of the three structures show a similar trend in B values for both the main and side chains. The higher mobility of different segments of the structure is reflected in the higher B values (see Fig. 5 also). Both of the termini are equally well defined unlike AzAD where the N-terminus had a significantly higher B value ($\simeq$ 100 Å^2) compared to the C-terminus ($\simeq$ 60 Å^2). In general, the B values for AzAD are much higher than azurin II for which the values are similar to those reported for AzP. In particular, the B values for residues 53–56, 78–80 and 102–105 are much higher in AzAD compared to those in the other two structures. The latter segment is in the loop region at the 'southern' end while the other two belong to segments going into and out of the flap, respectively. The side-chain B values for a much larger number of residues is high in the case of AzAD compared to AzP and azurin II.

The copper site

The copper site lies in what is known as the 'northern' part (Guss & Freeman, 1983) of the molecule. Like in the case of AzAD and AzP, there are three strong bonds to Cu from Nδ of His46 and His117 and Sγ Cys112. In addition, two groups are found axial to the copper, one being the Sδ of Met121 and the other the carbonyl O atom from Gly45, Fig. 10. Metric details of the Cu site of azurin II from unrestrained refinement are given in Table 4 and Fig. 11, where these are compared with the refined Cu site of AzAD. The three strong ligands are arranged in a distorted trigonal plane around the copper with the Cu atom located in this plane. In the case of AzAD and AzP, the Cu atom is displaced by ~0.1 Å from the plane towards the Cu—SMet bond compared to 0.36 Å in poplar plastocyanin (Guss, Bartunik & Freeman, 1992).

During the refinement of azurin II, special attention was paid to the copper-site geometry. Only weak restraints were applied to the Cu–ligand distances as are detailed in *Refinement* above. These restraints were the same as those used by Nar *et al.* (1991*a*,*b*) for AzP refinement. To remove bias of the starting model, extensive refinement was undertaken in the final stages without any restraints. The unrestrained refinement showed shifts of 0.05 Å for Nδ117 and 0.06 Å for Sγ112; no shifts were observed for the other ligands.

It is clear that the overall geometry of the copper site in azurin II is very similar to the other two refined structures of azurin. Also, the metal–ligand distances are in reasonable agreement not only among the azurins but also with the well refined structure of plastocyanin. However, close inspection of Table 4 reveals some differences which are larger than would be expected from the quoted e.s.d's. For example, the Cu–Sγ distance for AzP (Nar *et al.*, 1991*b*) is about 0.13 Å longer than found for all the other cases. It is also longer compared to the extended

Table 4. *Details of the copper geometry for azurin* II *and comparison with the refined* Cu *sites of AzAD and AzP*

Also included are details of the high-resolution structures of plastocyanin (Guss *et al.*, 1992) where the Cu site is expected to be similar but lacks the fifth interaction. Information from EXAFS studies of single crystal and solutions are given where available (see text for references).

Cu ligand (azurin)	AzII	AzAD	AzP	AzP (EXAFS)	Cu ligand (plastocyanin)	Pc 1.6 Å	Pc 1.33 Å	Pc EXAFS (crystal)	Pc EXAFS (solution)
46 Nδ	2.02	2.08	2.03	1.92	37 Nδ	2.04	1.91	1.98	1.94
117 Nδ	2.02	2.01	2.11	1.92	87 Nδ	2.10	2.06	1.98	2.02
112 Sγ	2.12	2.12	2.25	2.12	84 Sγ	2.13	2.07	2.08	2.09
121 Sδ	3.26	3.12	3.15	3.04	92 Sδ	2.90	2.82		
45 O	2.75	3.16	2.97	2.79					

Bond angles around CU (°)	
Sγ(112)—Cu—Nδ(46)	134.03
Sγ(112)—Cu—Nδ(117)	119.59
Sγ(112)—Cu—Sδ(121)	103.69
Nδ(46)—Cu—Sδ(121)	74.80
Nδ(46)—Cu—Nδ(117)	106.36
Sδ(121)—Cu—Nδ(117)	88.7
O(45)—Cu—Nδ(46)	76.35
O(45)—Cu—Sδ(121)	147.41
O(45)—Cu—Nδ(117)	86.10
O(45)—Cu—Sγ(112)	106.36

Bond angles at ligands (°)		
Gly45	C—O—Cu	135.86
His46	Cγ—Nδ—Cu	130.58
	Cε—Nδ—Cu	119.91
Cys112	Cβ—S—Cu	115.88
His117	Cγ—Nδ—Cu	126.64
	Cε—Nδ—Cu	123.36
Met121	Cγ—Sδ—Cu	142.90
	Cε—Sδ—Cu	101.07
	Cγ—Sδ—Cε	100.55

Cu out of plane distance (Å)	0.015
Cys112 hydrogen bonds	
Sg—HN(Asn47)	2.58
Sγ—HN(Phe114)	2.66

X-ray absorption fine-structure spectroscopy (EXAFS) value of AzP in solution (Strange *et al.*, 1993). We also note that the average Cu–His distances are longer in the case of AzAD (0.11 Å), AzP (0.15 Å) and AzII (0.10 Å) compared to the EXAFS value for AzP. The 1.33 Å refined structure of plastocyanin (Guss *et al.*, 1992) shows an excellent agreement with the metrical information obtained from EXAFS data of solution (Murphy *et al.*, 1990) and single crystal (Scott, Hahn, Doniach, Freeman & Hodgson, 1982) plastocyanin. We note that EXAFS data for the Cu–His and Cu–Cys are expected to be accurate to within ±0.03 Å. The differences discussed above are of similar magnitude as would be expected for an oxidation/reduction step, thus from a chemical as well as functional point of view it is important to establish the accuracy of this metrical information from a crystallographic structure. Guss *et al.* (1992) point out that the systematic shortening of Cu–ligand distances in the 1.33 Å structure of plastocyanin compared to 1.6 Å structure arise not only from the extension of resolution (0.03 Å) but also from the use of a combined SR(film)/diffractometer data (0.06 Å) as against to just the diffractometer data.

As pointed out by Guss *et al.* (1992), it remains problematical to know what precision can be claimed for the atomic position in a correctly determined and competently refined structure. The average coordinate error can be reasonably well estimated from the Luzzati or Read plots. However, these global criteria do not provide an accurate assessment for the accuracy of a particular part of the structure, *e.g.* the metal site in metalloprotein. In a number of cases, the differences in the crystallographically independent molecules in the asymmetric unit have been used to provide an assessment of accuracy of metal–ligand bond distances. For AzAD (Baker, 1988), thus an average 'maximum error' of 0.04 was estimated for Cu–ligand distances while for AzP (Nar *et al.*, 1991*a*,*b*) a precision ranging from 0.01 to 0.11 Å was calculated from the four molecules in the asymmetric unit. Guss *et al.* (1992) pointed out that as the structures of the molecules in the asymmetric unit are not independently determined, the error estimates based on internal self-consistency without allowance for systematic effects are likely to be too low. Guss *et al.* (1992) suggest that metal–ligand distances in a well refined structure of a blue-copper protein are uncertain by 0.04, 0.05 and 0.07 Å at 1.3, 1.6 and at 1.8 Å resolution, respectively. The effect of different refinement protocols on these metal–ligand distances is also not fully understood. Fields *et al.* (1994) have reported two independent refinements for plastocyanin with data collected at 173 K to 1.6 Å resolution. They report a mean-square difference of 0.08 Å for the Cα atoms in the two refinements. For Cu–ligand distances they observe a maximum difference of 0.07 Å. Thus, the variability in the Cu–ligand distances, particularly Cu–His and Cu–Cys, observed in Table 4 between different structures, may only reflect the different qualities of data (R_{merge}, resolution and completeness) and crystal (radiation sensitivity, mosaicity, *etc.*) and differences in refinement protocols used.

The Cu-site in azurin II is well ordered and the atoms participating in copper coordination exhibit low B values. Thus, Cu has a B value of 10.8 Å^2, Nδ of His46 and His117 has B values of 5.1 and 8.0 Å^2, respectively, while O(45), S(112) and S(121) have B values of 11.6, 9.3 and 11.4 Å^2. These compare favourably with the B values of AzP (and AzAD in parentheses), which are 10.7 (12.3), 11.2 (11.6), 8.0 (9.7), 11.6 (11.3), 11.3 (10.0) and 9.5 (13.1) Å^2 for Cu, N(46), N(117), O(45), S(112) and S(121), respectively.

Molecular packing and implications for electron transfer

The hydrophobic patch surrounding His117 dominates the interaction between two azurin II molecules in a

manner very similar to that already noted for AzAD (Baker, 1988) and AzP (Nar *et al.*, 1991*a*,*b*). A comparison of the molecular contacts for azurin II and AzAD is shown in Fig. 12. The importance of the hydrophobic patch and the potential role of His117 as a mediator for electron transfer (Van De Kamp *et al.*, 1990; Chen *et al.*, 1992) is further emphasized by the molecular structure of azurin II. The distance between the Cu atoms of the neighbouring molecules is 14.7 Å with an edge-to-edge His117–His 117 distance of 6.7 Å. A well ordered water molecule (W167) with $B = 10$ Å^2 is found at 2.69 Å from Nε of His117 of molecule *A* and at about 2.64 Å from the carbonyl group of Gly116 of molecule *B*, Fig. 13. The peptide C═O group of residue 116 is among the very few polar groups which protrude in the interface. Thus, the possibility exists for the electron transfer to take place *via* this water molecule through the peptide bond or through space between the two histidines directly. We also note the presence of a number of water molecules in the neighbourhood which could provide the source of protons for efficient cycling of the protein between redox states.

Concluding remarks

The successful refinement of azurin II from *A. xylosoxidans* can be attributed to several factors. Most importantly the realization that the organism has two distinct azurins with only 67% identity between them together with the realization that azurin II has a high sequence identity with azurin from *A. denitrificans* (Dodd *et al.*, 1995). Despite a low completeness, the high quality of the X-ray intensity data ($R_{merge} = 4\%$) obtained from one crystal further facilitated the structure deter-

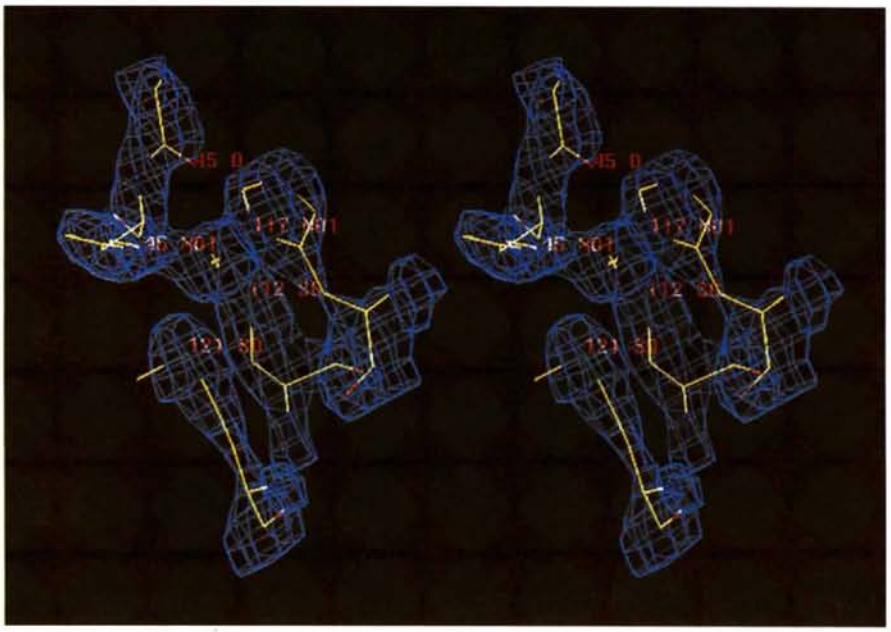


Fig. 10. Electron density around the Cu site.

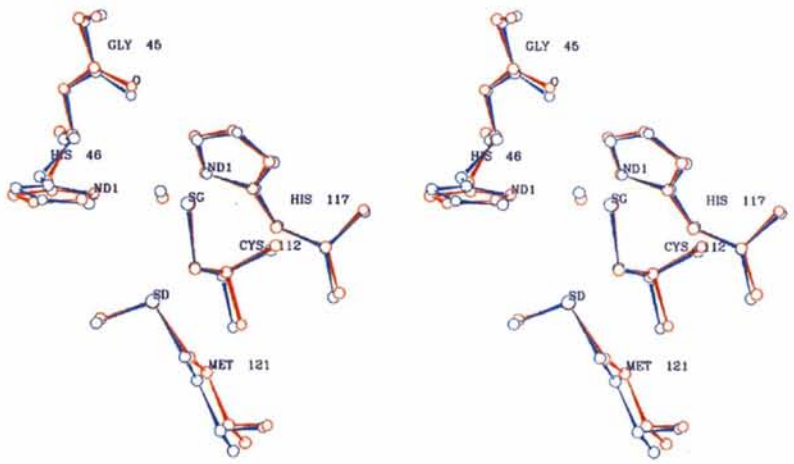


Fig. 11. A superposition of the copper sites of azurin II (blue) with AzAD (red). The ligand positions show a good likeness but the Cu atom can be seen to be differently placed. In azurin II the Cu atoms is in the trigonal plane (displacement is < 0.02 Å) of the strong ligands.

mination and refinement. The structure has provided a basis of comparative study for the three azurins for which relatively high resolution crystallographic structures are now available. The variability of Cu–ligand distances in these metalloproteins raises questions about the atomic precision of well refined and well determined structures. The observation that the *B* values are smaller in azurin II and AzP compared to AzAD may stem from different qualities of data, crystals and also, to some extent, from the exact refinement protocols employed.

The structure of azurin II has provided further support to the idea that the hydrophobic patch around His117 is the crucial region of protein–protein interaction for electron transfer. The presence of a well ordered water and its interaction with the Nε of His117 of one molecule and with the polar carbonyl group of Gly116 provide an interesting alternative where the electron can be transferred either through space between the Nε of His or through the water molecule and the peptide carbonyl. It is tempting to suggest that this region is also involved in

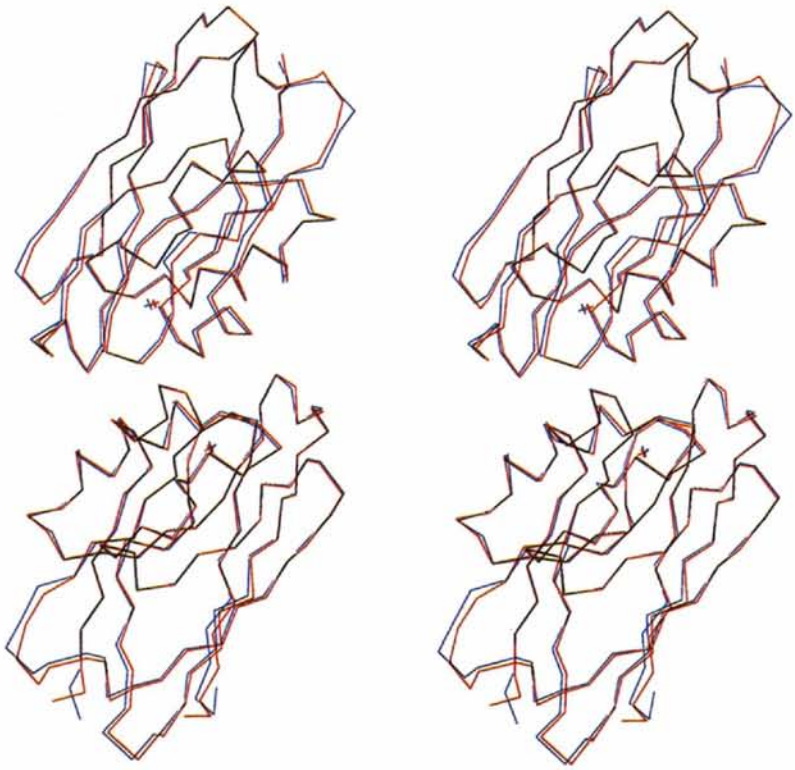


Fig. 12. A stereo superposition of the asymmetric unit on AzAD (red) with two symmetry-related molecules of azurin II (blue).

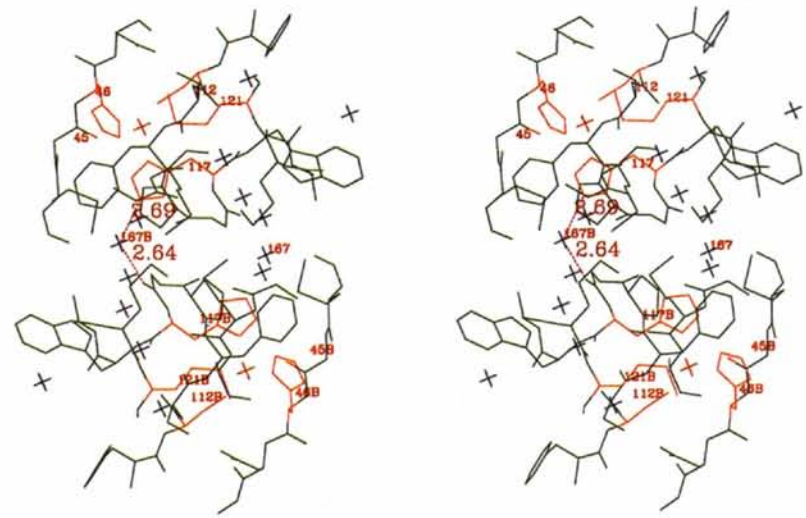


Fig. 13. The interface of the two azurin II molecules shown in Fig. 12. The Cu atoms are 14.7 Å apart. A potential electron pathway between the two copper site *via* His117 Nd, W167, Gly116B O and His117B Nδ is clear (see text for further details).

interaction with nitrite reductase and donates an electron to the type 1 Cu site of the enzyme.

The question why the organism produces two electron-donor proteins with similar redox potentials is interesting and can only be speculated upon at this stage. The close similarity of azurin II to the sole known azurin from NCTC8582 and the close taxonomical relationship between *A. xylosoxidans* 11015 and *A. denitrificans* 8582 does raise the question whether the latter organism also has the other azurin.*

We are grateful to Professor Sakabe for providing experimental facilities at the Photon Factory and for useful discussions. We are also indebted to Dr W. Hunter for his support in the initial stage of the project as well as for his continued interest and encouragement. We also would like to thank Drs Ian Harvey and Richard Strange for their help in data collection and processing. We also thank Professors E. N. Baker and H. Freeman for their comments.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1ARN). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: HE0118).

References

Abraham, Z. H. L., Lowe, D. J. & Smith, B. E. (1993). *Biochem. J.* **295**, 587–593.

Adman, E. T. & Jensen, L. H. (1981). *Isr. J. Chem.* **21**, 8–12.

Ambler, R. P. (1971). In *Recent Developments in the Chemical Study of Protein Structure*, edited by A. Previero, J.-F. Pechere & M. A. Coletti-Previero, pp. 289–305. Paris: INSERM.

Ambler, R. P. & Tobari, J. (1989). *Biochem. J.* **261**, 495–499.

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

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

Brünger, A. T. (1992*a*). *X-PLOR Manual*. Version 3.0, Yale Univ. Press, New Haven, CT, USA.

Brünger, A. T. (1992*b*). *Nature (London)*, **355**, 472–475.

Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.

Chen, L., Durley, R., Poliks, B. J., Hamada, K., Chen, Z., Mathews, F. S., Davidson, V. L., Satow, Y., Huzinga, E., Vellieux, F. M. D. & Hol, W. G. J. (1992). *Biochemistry*, **31**, 4959–4964.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dodd, F. E., Hasnain, S. S., Hunter, W. N., Abraham, W. N., Debenham, M., Kanzler, H., Eldridge, M., Eady, R. R., Ambler, R. P. & Smith, B. E. (1995). *Biochemistry*. **34**, 10180–10186.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Farver, O., Skov, L.K., Pascher, T., Karlsson, B. G., Nordling, M., Lundberg, L. G., Vänngård, T. & Pecht, I. (1993). *Biochemistry*, **32**, 7317–7322.

Fields, B. A., Bartsch, H. H., Bartunik, H. D., Cordes, F., Guss, J. M. & Freeman, H. C. (1994). *Acta Cryst.* D**50**, 709–730.

Fitzgerald, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.

Grossmann, J. G., Abraham, Z. H. L., Adman, E. T., Neu, M., Eady, R. R., Smith, B. E. & Hasnain, S. S. (1993). *Biochemistry*, **32**, 7360–7366.

Guss, J. M., Bartunik, H. D. & Freeman, H. C. (1992). *Acta Cryst.* B**48**, 790–811.

Guss, J. M. & Freeman, H. C. (1983). *J. Mol. Biol.* **169**, 521–563.

Higashi, T. (1989). *J. Appl. Cryst.* **22**, 9–18.

Inoue, T., Shibata, N., Nakanishi, H., Koyama, S., Ishii, H., Kai, Y., Harada, S., Kasai, N., Ohshiro, Y., Suzuki, S., Kohzuma, T., Yamaguchi, K., Shidara, S. & Iwasaki, H. (1994). *J. Biochem.* **116**, 1193–1197.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kakutani, T., Watanabe, H., Arima, K. & Beppu, T. (1981). *J. Biochem.* **89**, 463–472.

Korzun, Z. R. (1987). *J. Mol. Biol.* **196**, 413–419.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291

Liu, M. Y., Liu, M. C., Payne, W. J. & LeGall, J. (1986). *J. Bacteriol.* **166**, 604–608.

Luzzati, P. V. (1952). *Acta Cryst.* **5**, 802–810.

Miyata, M. & Mori, T. (1969). *J. Biochem.* **66**, 463–471.

Murphy, L. M., Hasnain, S. S., Strange, R. W., Harvey, I. & Ingledew, J. W. (1990). *X-ray Absorption Fine Structure*, edited by S. S. Hasnain, pp. 152–154. Chichester: Ellis Horwood.

Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M. & Canters, G. W. (1991*a*). *J. Mol. Biol.* **218**, 427–447.

Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M. & Canters, G. W. (1991*b*). *J. Mol. Biol.* **221**, 765–772.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Norris, G. E., Anderson, B. F. & Baker, E. N. (1983). *J. Mol. Biol.* **165**, 501–521.

Norris, G. E., Anderson, B. F. , Baker, E. N. & Rumball, S. V. (1979). *J. Mol. Biol.* **135**, 309–312.

Powell, M. J. D. (1977). *Math. Program.* **12**, 241–254.

Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 995.

Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 238.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

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

Sakabe, N. (1991). *Nucl. Instrum. Methods A*, **303**, 448–463.

Scott, R. A., Hahn, J. E., Doniach, S., Freeman, H. C. & Hodgson, K. O. (1982). *J. Am. Chem. Soc.* **104**, 5364–5369.

Strahs, G. (1969). *Science*, **165**, 60–61.

Strange, R. W., Dodd, F. E., Abraham, Z. H. L., Grossmann, J. G., Brüser, T., Eady, R. R., Smith, B. E.. & Hasnain, S. S. (1995). *Nature Struct. Biol.* **2**, 287–292.

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

Suzuki, H. & Iwasaki, H. (1962). *J. Biochem.* **52**, 193–199.

Van De Kamp, M, Silverstini, M. C., Brunori, M, Van Beuman, J., Hali, F. C., Canters, G. W. (1990). *Eur. J. Biochem.* **194**, 109–118.

Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318.

Zumft, W. G., Gotzmann, D. J. & Kroneck, P. M. H. (1987). *Eur. J. Biochem.* **168**, 301–307.