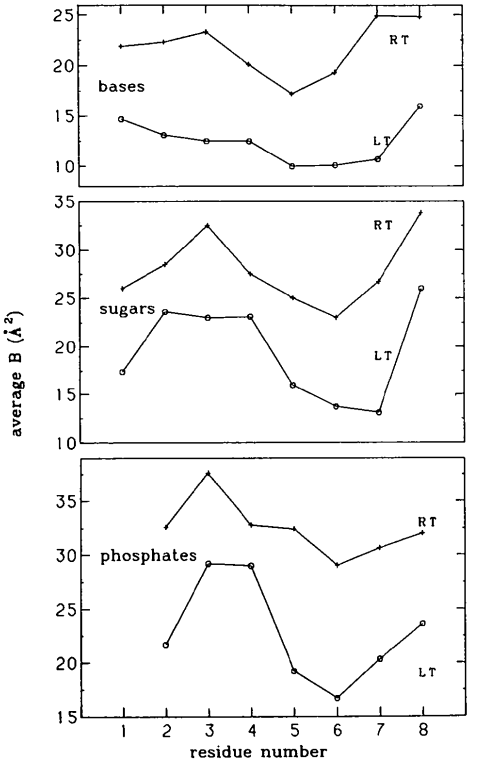

Fig. 3. The distribution of average B's for bases, sugars and phosphates.

the first group (13·5 Å^2) and larger for the second and third groups, 16·8 and 18·1 Å^2, respectively.

Concluding remarks

The successful cooling of several different DNA crystals by different techniques, as demonstrated here and elsewhere (Drew *et al.*, 1982), suggests that cryotechniques may be applicable to a wide variety of DNA crystals. The major advantage of data collection at cryotemperature is the apparent stability of the crystal. Since the preparation of suitable crystals is often a very difficult task, the need for fewer crystals may also be of considerable significance. The shock-cooling technique is experimentally very convenient compared to slow cooling over a period of days and, moreover, it eliminates all complications associated with the use of capillaries. However, this study and the study of d(GGGTACCC) (space group $P6_1$; Eisenstein, Frolow, Shakked & Rabinovich, in preparation) show that no improvement in resolution is achieved on applying cryotechniques to crystals of DNA fragments. Unlike crambin and BPTI, the damage caused to DNA crystals, which contain more than 50% solvent, outweighs the benefits of lowering the B values.

We thank the United States/Israel Binational Science Foundation (BSF) Jerusalem, Israel (grant No. 84-00264), the Fund for Basic Research administrated by the Israel Academy of Science and Humanities, and the Israel Cancer Association for financial support.

References

DICKERSON, R. E., KOPKA, M. L. & PJURA, P. (1985). *Biological Macromolecules and Assemblies*, Vol. 2, edited by A. MCPHERSON & F. JURNAK, pp. 38–126. New York: John Wiley.

DREW, H. R., SAMSON, S. & DICKERSON, R. E. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 4040–4044.

FROLOW, F. (1988). Unpublished results.

HENDRICKSON, W. A. & KONNERT, J. H. (1981). *Biomolecular Structure, Conformation, Function and Evolution*, Vol. 1, edited by R. SRINIVASAN, pp. 43–47. Oxford: Pergamon Press.

HOPE, H. (1988). *Acta Cryst.* B**44**, 22–26.

RABINOVICH, D., HARAN, T., EISENSTEIN, M. & SHAKKED, Z. (1988). *J. Mol. Biol.* **200**, 151–161.

SAENGER, W. (1984). *Principles of Nucleic Acid Structure*. New York: Springer.

SHAKKED, Z. & RABINOVICH, D. (1986). *Prog. Biophys. Mol. Biol.* **47**, 159–195.

Acta Cryst. (1988). B**44**, 628–636

Refinement of the Structure of Pseudoazurin from *Alcaligenes faecalis* S-6 at 1.55 Å Resolution

BY K. PETRATOS, Z. DAUTER AND K. S. WILSON

European Molecular Biology Laboratory (EMBL), c/o Deutsches Elektronen Synchroton (DESY), Notkestrasse 85, *D*-2000 *Hamburg* 52, *Federal Republic of Germany*

(*Received* 26 *April* 1988; *accepted* 13 *July* 1988)

Abstract

The crystal structure of the redox protein pseudoazurin (123 amino acid residues; molecular weight 13 000 daltons) from *Alcaligenes faecalis* has been refined by fast Fourier restrained least-squares minimization. Cycles of rebuilding were carried out to escape from local minima. Individual isotropic temperature factor

0108-7681/88/060628-09$03.00

values were refined separately for all atoms. The R factor was reduced from 0·400 (for 2647 reflections in the 6·0–2·8 Å resolution range) to 0·180 (for all 19 770 reflections in the 9·0–1·55 Å resolution range) with a final estimated accuracy in atomic positions of 0·15 Å. The final model comprises 917 protein atoms and 93 solvent molecules. The root-mean-square shift of the main-chain atoms between the final and the initial model is 0·94 Å (maximum shift 1·8 Å). Most of the larger shifts were the result of rebuilding on the graphics system. The average atomic temperature factor, B, is 23·0 Å^2 for all atoms. Side-chain atoms with high B values were omitted, and their positions checked from difference maps. The three carboxy-terminal residues were omitted from the final model as no single conformation could be assigned from the observed electron density. All other protein atoms were included.

Introduction

Pseudoazurin is a blue copper protein from the potent denitrifying bacterium *Alcaligenes faecalis* strain S-6. It was first isolated and studied by Kakutani, Watanabe, Arima & Beppu (1981*a*). The molecule of pseudoazurin consists of 123 amino acid residues with one tightly bound copper ion.

The biological function of the protein as deduced from *in vitro* experiments (Kakutani, Watanabe, Arima & Beppu, 1981*b*) is the electron transfer to nitrite reductase which is the last member of the electron transport chain carrying out the anaerobic respiration of the bacterium. Pseudoazurin was crystallized independently in two laboratories (Adman, Beppu & Watanabe, 1984; Petratos, 1984). We reported the crystal structure of the protein at 2·9 Å resolution (Petratos, Banner, Beppu, Wilson & Tsernoglou, 1987). The structure is predominantly extended β-sheet (55%), comprising 30% turns and 15% α-helices. It can be described as a β-barrel consisting of two β-sheets, I and II according to the notation of Chothia & Lesk (1982), which comprise eight β-strands, and an α-helical hairpin which is formed by two α-helices at the C-terminus.

The crystal structures of three other blue copper proteins: azurin from *Pseudomonas aeruginosa* (Adman & Jensen, 1981), poplar plastocyanin (Guss & Freeman, 1983) and azurin from *Alcaligenes denitrificans* (Norris, Anderson & Baker, 1986) have been determined. The structures of plastocyanin and azurin from *Alcaligenes denitrificans* have been refined to high resolution. The folding of these proteins is similar to pseudoazurin.

Protein structures were refined by least-squares minimization firstly in reciprocal space with conventional calculation of the structure factors and gradients (Watenpaugh, Sieker, Herriott & Jensen, 1973). Greatly improved behaviour in refinement was obtained by the introduction of restraints on the atomic parameters (Hendrickson & Konnert, 1980). For macromolecules the restraints are necessary because the ratio of observations to parameters is too low, even at the higher resolutions, to refine coordinates independently. In recent years, the time requirements for the least-squares program have been dramatically reduced with the implementation of fast Fourier alogrithms for the computation of both the structure factors and the gradients (Agarwal, 1978). Such a program was used in the present work (Baker & Dodson, 1980).

In order to obtain an accurate structure, especially around the Cu site which is important for the function of the protein redox system, we have carried out the crystallographic refinement of the oxidized state of pseudoazurin at pH 6·8 to a nominal resolution of 1·55 Å. It is this refinement that we describe here. The structure of the reduced state will be studied later.

Data collection and processing

Pseudoazurin was crystallized as described earlier (Petratos, 1984). Crystals are hexagonal, space group $P6_5$, cell dimensions $a = b = 50·0$ (1), $c = 98·5$ (3) Å. They are truncated hexagonal bipyramids. One molecule of protein per asymmetric unit gives a crystal volume to protein mass ratio of 2·70 Å^3 dalton^{-1}, which is in the range of values stated by Matthews (1968) for crystals of proteins. Large crystals (0·6 × 0·6 × 0·8 mm) of pseudoazurin when freshly prepared diffract X-rays to 1·3 Å resolution. However, after about one hour of exposure, the diffraction degrades to a nominal resolution of 1·55 Å owing to radiation damage.

Diffraction data were collected on CEA Reflex-25 film at the EMBL instrument X31 of the DORIS storage ring at DESY, Hamburg. The data were recorded in July 1986 when DORIS was operated in single-bunch mode at 5·3 GeV and about 20–30 mA with approximately 1 h between injections. The size of the apertures of the collimator was 0·4 × 0·4 mm. The crystal-to-film distance was 75 mm. A single crystal of Si, channel-cut along (111), provided monochromated X-rays of $\lambda = 1·009$ Å with $\Delta\lambda/\lambda = 0·0015$. On 62 three-film packs interleaved with 25 μm thick steel foils, we recorded 108 356 Bragg reflections which correspond to 19 770 unique data to 1·55 Å resolution. A total of three crystals was used. Two were mounted in the capillary tube with the unique sixfold symmetry axis parallel to the direction of the spindle axis of the Arndt–Wonacott camera. This way one makes use of the inherent symmetry of the crystals, so that only 60° of rotation are required in order to collect a complete data set. A total of 72° was collected from these two crystals, the overlapping data giving better scaling between the crystals. The rotation per film pack was 1·5°. Each exposure required an average of 20 min.

The third crystal was mounted in the glass capillary so that its c axis was approximately at right angles to the spindle and 21° of data were recorded in order to fill the 'blind region' of the resolution sphere uncollected in the previous crystal setting. The films were digitized on an Optronics Photoscan P-1000 microdensitometer using a raster window of 100 × 100 μm, appropriate for an average spot size of 0·8 mm. The values of the optical density in the range of 0–3 OD units were output as integers in the range of 0–255. Integration and reduction of the intensities were subsequently carried out on a VAX 11/750 computer using the *MOSCO* program system (Machin, Wonacott & Moss, 1983). Profile fitting (Rossmann, 1979) was used in order to improve the integration and minimize the background residuals.

On the first film of each pack there were about 1600 fully and 700 partially recorded reflections. On the third film of each pack about 25 strong reflections were 'overloaded' and rejected. In order to limit the loss of important strong reflections of the innermost shells of resolution, a data set extending to 2·5 Å resolution collected on a Nonius CAD-4F diffractometer was merged with the synchrotron data.

The data set represents 97·3% of the expected number of reflections to 1·55 Å resolution. For the entire data set 18 269 reflections had intensities greater than σ_I, 17 111 reflections had intensities greater than $2\sigma_I$ and 15 978 reflections had intensities greater than $3\sigma_I$, the standard deviation, σ_I, being derived from statistical calculations in the *MOSCO* system (Machin, Wonacott & Moss, 1983). The merging R_m factor, $R_m = \sum_{\mathbf{h}} |I_i - I| / \sum_{\mathbf{h}} I$ (I_i is the intensity of an individual measurement, I the mean value for that reflection and the summations are over all measurements), was 0·079 for all data to 1·55 Å resolution. Only 45 reflections were rejected as having $|I_i - I| > 4\sigma_I$. The fraction of the number of the observed reflections with intensities $I > 3\sigma_I$ as a function of resolution is given in Fig. 1. The Wilson (1949) distribution is illustrated in Fig. 2 for the total of 19 770 reflections. The fit to the theoretical straight line is satisfactory for a protein data set with the expected deviations in the range 4·0–10·0 Å. The overall temperature factor obtained from the slope of the least-squares line, 15·5 Å^2, is somewhat less than the B value obtained later from the refinement (23·0 Å^2).

Course of model building and refinement

(*a*) *Model building*

The first model of the pseudoazurin had been constructed on an Evans and Sutherland Picture System 2 graphics display system with the interactive computer graphics program *FRODO* (Jones, 1978) as implemented and modified at EMBL–Heidelberg by H. Bosshard, using the diffractometer data extending to 2·9 Å resolution phased with MIR methods (Petratos, 1984). In spite of the incomplete chemical sequence (only 70% was known), 109 residues were fitted in the electron density map. The complete primary structure was determined later (Hormel, Adman, Walsh, Beppu & Titani, 1986) and the model was improved in the same MIR map with nearly all residues fitted. Before the first cycle of least-squares minimization the R factor, $R = \sum_{\mathbf{h}} ||F_o| - |F_c|| / \sum_{\mathbf{h}} |F_o|$ ($|F_o|$, $|F_c|$ denote the observed and calculated structure-factor amplitudes, respectively, and the summations are over all measurements), was 0·400 for 2647 reflections in the resolution range 6·0–2·8 Å.

After the first 30 cycles of least-squares refinement and extending the data in steps to 9·0–1·55 Å resolution the R factor decreased to 0·294. At this stage Fourier maps were calculated with coefficients $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ and phases α_c derived from the

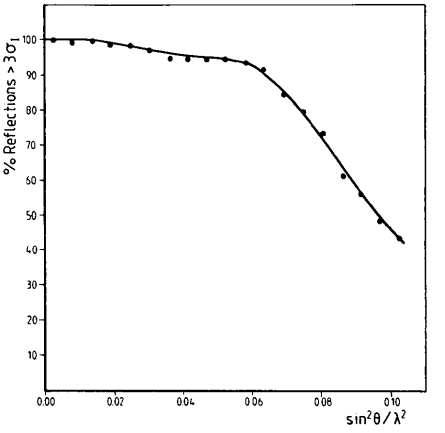

Fig. 1. Fraction of the number of reflections with intensities $I > 3\sigma_I$ as a function of resolution.

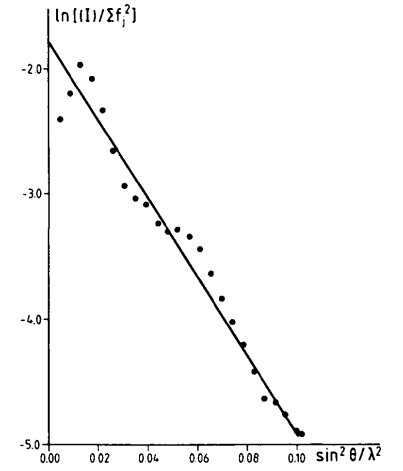

Fig. 2. The Wilson plot for the film data extending to 1·55 Å resolution.

model. The $2|F_o| - |F_c|$ map was contoured at 1·0, 1·4, 1·8 and 6·0 e Å^{-3}. The $|F_o| - |F_c|$ map was contoured at −0·3, −0·2, 0·2 and 0·3 e Å^{-3}.

Intermediate model-building interventions were carried out at the end of each step of the refinement as described in detail in the following section. The electron density maps were contoured as for the first rebuilding stated above except for the last two steps. The contouring levels for the second-last rebuilding were 0·5, 0·9, 1·3 and 6·0 e Å^{-3} and for the last rebuilding 0·4, 0·8, 1·2 and 6·0 e Å^{-3}. The $|F_o| - |F_c|$ maps were contoured at all steps at −0·3, −0·2, 0·2 and 0·3 e Å^{-3}.

All but three C-terminal residues (121–123) were included in the last refinement cycle (No. 141) as well as 93 solvent molecules. The reliability index was 0·180 for all reflections between 9·0 and 1·55 Å resolution. In the $|F_o| - |F_c|$ map there were a few peaks at the level of 0·3 e Å^{-3} which were attributed to the presence of some of the disordered atoms at the C-terminus. The atomic temperature factors, B, were refined independently of the positional parameters in single cycle runs. Occupancies of atoms were not refined, as they are highly correlated to the temperature factors. The rebuilding of the models was carried out on an Evans and Sutherland PS350 graphics system with *FRODO*.

(*b*) *Refinement procedure*

The stereochemistry was restrained according to the weighting scheme shown in Table 1. Structure-factor amplitudes were not given a resolution-dependent weight but were assigned unit weights. The refinement was carried out in 13 steps comprising 141 least-squares cycles. These were interspersed with eight rebuilds of the structure which involved significant movement of certain residues into the correct density and inclusion (or removal) of solvent molecules. The residues that were shifted most lie either at flexible turns of the structure or in the vicinity of the carboxy terminus. A summary of the progress of the refinement is given in Table 2 and the conventional R-factor plot against cycle number is drawn in Fig. 3.

All refinement runs were carried out on a MicroVAX II with a typical cycle for 116 643 reflections in space group $P1$ (equivalent to 19 770 independent observations) and 3032 variables requiring about 40 min of CPU time.

Step 1. Cycles 1–7. The initial model of pseudoazurin, consisting of 942 nonhydrogen atoms with atomic B values set to 12·0 Å^2, was refined with 2647 Bragg reflections in the 6·0–2·8 Å resolution range. The R factor was reduced from 0·400 to 0·219. At no stage of the refinement were any restraints imposed on the known hydrogen bonds between main-chain atoms belonging to β-strands or α-helices.

Step 2. Cycles 8–18. Reflections in the range 6·0–2·4 Å were included (4473). At the start the R factor increased to 0·299. After convergence it decreased to 0·218.

Table 1. *Weighting scheme and standard deviations after the last cycle of the refinement*

		σ*	Standard deviation (Δ)	No. of parameters
Distances (Å)				
Bond lengths	(1–2 neighbours)	0·020	0·010	933
Bond angles	(1–3 neighbours)	0·040	0·034	1263
Dihedral angles	(1–4 neighbours)	0·040	0·034	303
Planar groups		0·020	0·011	155
Chiral volumes (Å^3)		0·120	0·110	146
Non-bonded contacts (Å)				
Single torsion contacts		0·500	0·191	324
Multiple torsion contacts		0·500	0·294	389
Possible hydrogen-bonding contacts		0·500	0·128	78
Torsion angles (°)				
Peptide plane	(ω)	3·0	2·2	120
Staggered	(χ_1)	15·0	16·5	171
Orthonormal	(χ_2)	20·0	34·1	10

* The weight in each class of restraints corresponds to $1/\sigma^2$.

Step 3. Cycles 19–22. More reflections were included (4710) by extending the resolution to the range 9·0–2·4 Å. Initially the R factor increased to 0·234 but then dropped to 0·222.

Step 4. Cycles 23–30. The data included were extended to 1·55 Å leading to a large increase of the R factor to 0·360. Convergence was achieved at an R factor of 0·294. After cycle No. 30 the first major rebuild of the model was carried out employing the $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ electron density maps. Searching the difference map 44 solvent molecules were included in the atom list with B values of 20·0 Å^2.

Step 5. Cycles 31–45. The atomic temperature factors were refined for the first time in cycles 34, 35, 41 and 42. After refinement the average B value was at 23·7 Å^2 and the R factor 0·207.

Step 6. Cycles 46–53. Only 33 of the solvent peaks were retained. The rest were deleted from the list of atoms as having refined to B values greater than 50·0 Å^2. The 25 atoms belonging to the last four residues of the molecule were assigned zero occupancy as having refined to high B values ($>60{\cdot}0$ Å^2). In addition, two wrongly assigned residues in the low-resolution structure, Lys 52 and Gln 106, were corrected to Gly 52 and Lys 106, respectively, which agrees with the chemical sequence. The errors in the sequence were due to misleading features of the medium-resolution MIR map. The atomic B values were refined in cycles 49 and 50. The R factor decreased to 0·219. This was the only step in which we attempted to use data in the range 7·0–1·55 Å in order to examine the effect of omitting the low-resolution terms from the refinement.

Step 7. Cycles 54–67. At the beginning of this step all atomic B values were reset to 12·0 Å^2 except for the

Table 2. *Description of the course of the refinement*

Step	1	2	3	4	5	6	7	8	9	10	11	12	13
No. of cycles	7	11	4	8	15	8	14	16	21	11	8	11	7
No. of parameters varied	2828	2828	2828	2828	2960	2837	2864	2864	2906	2870	2855	3035	3032
No. of reflections	2647	4473	4710	19770	19770	19770	19770	19770	19770	19770	19770	19770	19770
No. of atoms	942	942	942	942	986	945	954	954	968	956	951	1011	1010
No. of solvent molecules	(0)	(0)	(0)	(0)	44	33	40	40	51	73	70	102	93
Resolution range (Å)	6·0–2·8	6·0–2·4	9·0–2·4	9·0–1·55	9·0–1·55	7·0–1·55	9·0–1·55	9·0–1·55	9·0–1·55	9·0–1·55	9·0–1·55	9·0–1·55	9·0–1·55
Mean temperature factor B (Å^2)	12·0	12·0	12·0	12·0	23·7	25·1	18·8	21·8	21·6	20·3	18·9	21·8	23·0
R factor at start (%)	40·0	29·9	23·4	36·0	32·9	23·8	27·6	28·6	28·1	25·8	18·8	19·4	18·4
R factor at end (%)	21·9	21·8	22·2	29·4	20·7	21·9	19·2	19·0	18·7	18·4	18·2	18·0	18·0
Standard deviations from ideal values													
Bond lengths (1–2 distances) (Å)	0·079	0·108	0·074	0·150	0·029	0·026	0·021	0·026	0·026	0·013	0·011	0·011	0·010
Bond lengths (1–3 distances) (Å)	0·175	0·203	0·163	0·239	0·074	0·068	0·051	0·045	0·044	0·039	0·036	0·035	0·034
Planar groups (Å)	0·043	0·054	0·039	0·077	0·021	0·020	0·016	0·014	0·012	0·012	0·011	0·011	0·011
Chiral volumes (Å^3)	0·458	0·602	0·488	0·804	0·354	0·378	0·288	0·332	0·297	0·131	0·120	0·109	0·110
Torsion angle (ω) (°) of peptide plane	6·4	23·6	5·8	9·7	3·9	3·5	2·8	2·7	2·3	2·2	2·2	2·1	2·2

40 solvent molecules which were reset to 20·0 Å^2. The B values were refined in cycles 54, 55, 61, 62 and 63. Residue Ala 120 was reintroduced and the carboxyl group of Glu 1, as not having electron density at 0·5 e Å^{-3}, was removed from the model. The R factor decreased to 0·192.

Step 8. Cycles 68–83. Before this step a major rebuild was carried out with most attention paid to the orientation of the side chains and the backbone of the loops which connect the β-strands. The only class of structural parameters that had an unusually high root-mean-square deviation from ideality at this stage was the chiral volume (r.m.s.d. 0·332 Å^3). The atomic B values were refined in cycles 68, 69, 75 and 76 and the R factor was 0·190.

Step 9. Cycles 84–104. At the start the B values of the 917 protein atoms were reset to 12·0 Å^2 and the corresponding ones of 51 solvent molecules to 20·0 Å^2. The carboxyl group of Glu 1 removed in step 7 was reintroduced in the model. The atomic temperature factors were refined in cycles 84, 85, 91 and 97. The final average B value was 21·6 Å^2 and the R factor 0·187.

Step 10. Cycles 105–115. At the start of this step a total of 34 side-chain atoms belonging to residues Lys 10, Lys 24, Lys 38, Lys 46, Lys 55, Lys 59, Lys 109, Lys 117, Glu 51, Glu 54 and Glu 62 were assigned an occupancy of zero as having refined to high B values. The 883 protein atoms in the model were given a B value of 21·0 Å^2 (the average value from the previous step) and the 73 solvent molecules were assigned a B value of 35·0 Å^2. The thermal parameters were refined in cycles 105, 106 and 113. At the end of this step the R factor was 0·184 and the stereochemistry had improved considerably (see Table 2).

Step 11. Cycles 116–123. At the start three of the solvent molecules were removed from the atom list as having refined to B values greater than 50·0 Å^2. Moreover, the side-chain atoms Cγ of residues Glu 51 and Glu 62 were given zero occupancies as not showing any density in the difference maps. The R factor decreased slightly (0·182) and the average B after refinement in cycle 118 was 18·9 Å^2.

Step 12. Cycles 124–134. Before this step the $2|F_o| - |F_c|$ map was contoured at a lower level (0·5 e Å^{-3}) allowing the 28 side-chain atoms omitted at

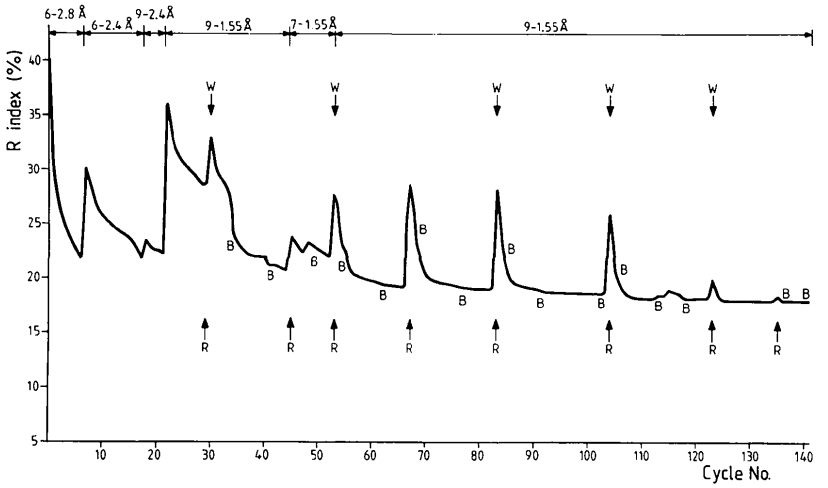

Fig. 3. Reliability factor R plotted against cycle number of the refinement. Rebuilds of the model, inclusion of solvent molecules and refinement of temperature factors are indicated, respectively, with the symbols: R, W and B. At the top of the diagram is shown the resolution range of the data included in each refinement step.

step 10 or 11 of the residues Lys 24, Lys 38, Lys 46, Lys 59, Lys 117, Glu 51, Glu 54 and Glu 62 to be reintroduced to the model. Only eight protein atoms from the side chains of Lys 10, Lys 55 and Lys 109 were now not included as well as the three C-terminal residues (Ser 121-Ala 122-Lys 123) which showed only 'fragmented' electron density in the difference Fourier maps. The solvent structure was extended to 102 molecules. The R factor was 0·180.

Step 13. Cycles 135–141. Before the start of this step the $2|F_o| - |F_c|$ map was contoured at an even lower level (0·40 e Å^{-3}) close to the r.m.s. value of the map (0·36 e Å^{-3}). As a result the eight side-chain atoms omitted previously from the refinement were reintroduced. In addition nine of the solvent molecules had to be removed as having refined to positions close to the newly included side chains. The B values were refined in cycles 135 and 141. The R factor increased initially to 0·184 but decreased to a final 0·180. The quality of the refined structure is clear from the sections of the final $2|F_o| - |F_c|$ electron density map illustrated in Fig. 4. A comparison of the superimposed initial and final refined models of the structure is shown in Fig. 5.

Refinement statistics

The standard deviations of the stereochemical parameters are summarized in Table 1. They were obtained from the correlation matrix after the last cycle of the refinement. Only 1·7% of all distances between next, second-next and third-next neighbours deviated by more than $2\sigma_D$ from the corresponding ideal values and two distances exceeded a deviation of $4\sigma_D$. Only 11 non-bonded atom pairs that are related *via* a single torsion formed contacts that are shorter than the ideal values by more than 0·4 Å. The final difference Fourier synthesis was practically featureless. There were eight positions where the residual density had values above

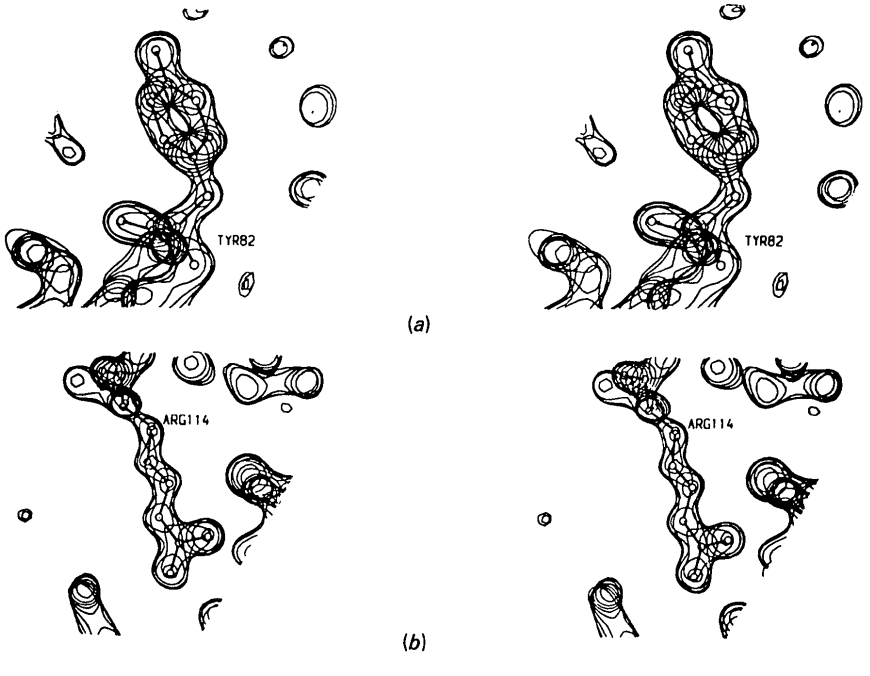

Fig. 4. Stereo pairs of two parts of the final electron density map calculated at 1·55 Å resolution showing: (*a*) Tyr 82 (scale = 4·80 mm Å^{-1}) and (*b*) Arg 114 (scale = 4·10 mm Å^{-1}) with a contour level of 0·9 e Å^{-3} for both. A superposition is shown of 20 sections along the Z axis.

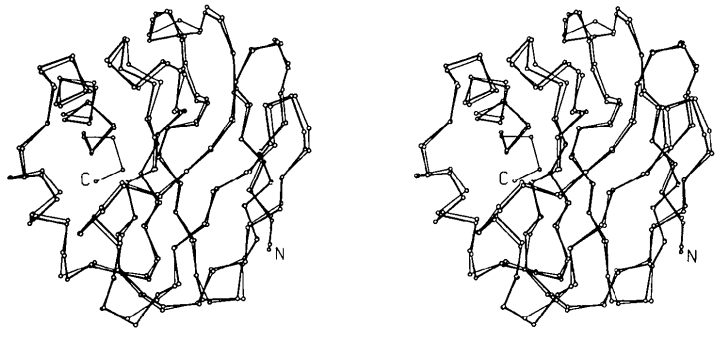

Fig. 5. A stereo Cα plot of the superimposed initial (thin lines) and final refined (thick lines) model of the structure of pseudoazurin. The amino and carboxy termini of the molecule are denoted by N and C respectively. The last three Cα atoms of the final model are not shown as they were not included in the last cycles of the refinement.

0·30 e Å^{-3}. These positions lay in the region of the map which corresponds to the C-terminal residues, but were not interpretable as a single or as a combination of different conformations of this. The map showed no electron density below −0·30 e Å^{-3}.

The quality of the refined structure is also indicated in the Ramanchandran plot, Fig. 6, in which the values of the main-chain torsion angles φ and ψ are plotted pairwise. The sterically allowed regions for all amino acids except glycyl residues are indicated by dashed lines (Ramakrishnan & Ramachandran, 1965). Almost all residues of the protein lie in the allowable region except for the eight glycines and Asn 61 (indicated by an arrow in the plot). Asn 61 is the third residue of a β-reverse turn and hydrogen-bonded to His 40. Residues Met 16, Asp 47 and Ser 95 appear in the left-handed α_L region of the plot. All three residues lie in turns of the structure. Met 16 is held in position by hydrogen bonds to Asn 9 and Arg 114, Asp 47 is the third residue of a β-reverse turn (type II) and Ser 95 is involved in intermolecular contacts in the crystal lattice.

The final average B value of all protein and solvent molecules was 23·0 Å^2. The main-chain atoms and carbonyl oxygens refined to an average B value of 15·6 Å^2, whereas the side-chain atoms refined to an average B value of 27·1 Å^2.

An upper limit for the average coordinate error can be assessed according to Luzzatti (1952) by plotting the R factor as a function of resolution (Fig. 7). Except for the low-resolution range the data points follow satisfactorily the line drawn for a mean coordinate error of 0·15 Å. However, the Luzzatti plot overestimates the error in the model because it assumes that the disagreement between F_o and F_c is only due to the coordinate errors. The r.m.s. difference of the coordinates of the main-chain atoms between the start and the end of the refinement was 0·94 Å. The corresponding r.m.s. deviations for the C_α atoms of the models obtained at the end of the steps 1–4 and the final refined model were 0·78, 0·73, 0·72 and 0·70 Å, respectively. This indicates that although the R factor was low after each of these early steps, the structure was still some distance from the final model.

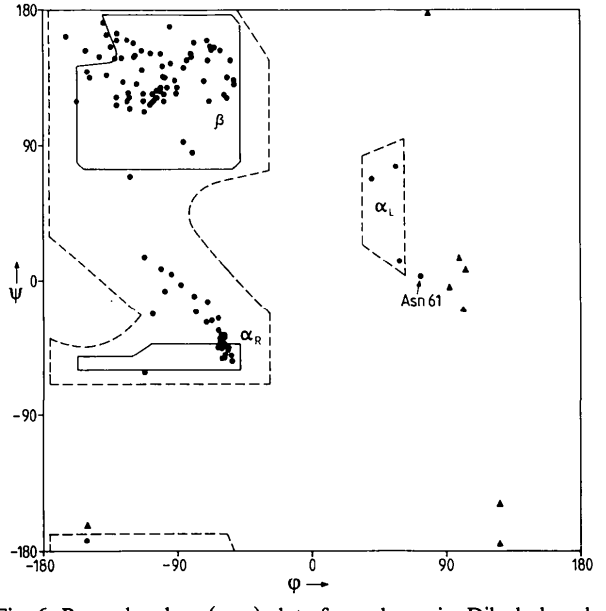

Fig. 6. Ramachandran (φ, ψ) plot of pseudoazurin. Dihedral angle regions for α-helices, β-pleated sheets and left-handed α-helices are indicated, respectively, by α_R, β and α_L. Glycines are indicated by triangles, all other residues by circles.

The geometry of the copper site

The copper ion (Cu^{2+}) and its ligand residues (His 40, Cys 78, His 81 and Met 86) are shown in the computer drawing of Fig. 8. The geometrical parameters of the Cu site are given in Table 3. The estimated error in bond lengths is 0·05 Å and in bond angles 3°. The ligands of the metal, 40Nδ_1, 78Sγ, 81Nδ_1 and 86Sδ, form a distorted tetrahedron. The Cu^{2+} ion lies 0·43 Å out of the plane defined by the three 'strong' ligands (40Nδ_1, 78Sγ, 81Nδ_1) towards the fourth, weaker ligand 86Sδ. The cysteine S atom, 78Sγ, is a hydrogen-bond acceptor of 41N, whereas the imidazole N atoms 40Nε_2 and 81Nε_2 are hydrogen-bond donors to 9Oδ_1 of

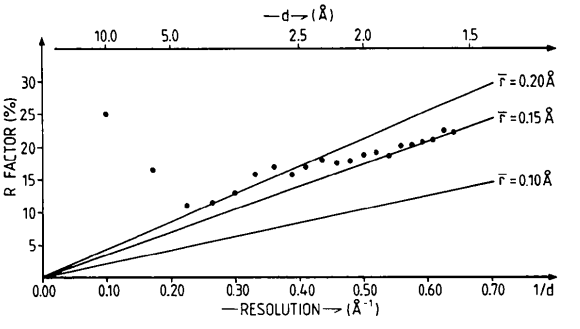

Fig. 7. Plot of R factor *vs* resolution after Luzzatti (1952). All reflections with $F_o > 3\sigma_F$ were used.

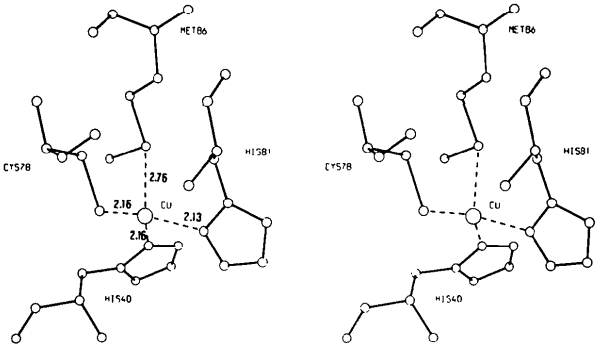

Fig. 8. Stereo computer drawing of the Cu site in pseudoazurin. The metal-to-ligand distances are given in Å.

Table 3. *The geometry of the Cu site*

Distances		Distance of Cu ion from planes		Angles	
Cu–40Nδ_1	2·16 Å	Imidazole ring of His40	0·10 Å	40Nδ_1–Cu–78Sγ	136°
Cu–78Sγ	2·16	Imidazole ring of His81	0·04	40Nδ_1–Cu–81Nδ_1	100
Cu–81Nδ_1	2·13	40Nδ_1, 78Sγ, 81Nδ_1	0·43	40Nδ_1–Cu–86Sδ	87
Cu–86Sδ	2·76	40Nδ_1, 78Sγ, 86Sδ	0·66	78Sγ–Cu–81Nδ_1	112
Cu–39O	3·79	40Nδ_1, 81Nδ_1, 86Sδ	1·08	78Sγ–Cu–86Sδ	108
78Sγ–41N	3·61	81Nδ_1, 78Sγ, 86Sδ	0·73	81Nδ_1–Cu–86Sδ	112
40Nε_2–9Oδ_1	2·70			Cu–78Sγ–78Cβ	105
81Nε_2–O*w*	2·76				

Asn 9 and to a solvent molecule, which is assumed to be water, respectively.

The geometry of the Cu site is clearly similar, but not identical, to those of azurin and plastocyanin. The Cu^{2+} ion, as in the other proteins, is tightly bound in a hydrophobic environment near the surface of the molecule. In azurin the ligands form a trigonal bipyramid (Norris, Anderson & Baker, 1986), with the carbonyl oxygen of Gly 45 as the weak fifth ligand, 3·1 Å distant from the copper. In pseudoazurin the comparable Cu–O39 distance is 3·8 Å, too long for the oxygen to be considered as a ligand. The copper site is much more like in plastocyanin (Guss & Freeman, 1983) but slightly more regular tetrahedral as judged by the values of the angles between the metal and its ligands, 78Sγ–Cu–81Nδ_1 and 81Nδ_1–Cu–86Sδ, which are 112° in pseudoazurin but 123 and 103° in plastocyanin. The angle between the copper, 78Sγ and 78Cβ is 105°, which is very close to the value reported for plastocyanin (107°). This angle is implicated in the multiple mode of binding of the thiolate S and the Cu^{2+} ion in order to produce the characteristic optical and EPR spectra of the protein (Penfield, Gewirth & Solomon, 1985). The extensive hydrogen bonding at the vicinity of the metal will be described in a separate communication.

Concluding remarks

The crystal structure of pseudoazurin from *Alcaligenes faecalis* strain S-6 was refined to a nominal resolution of 1·55 Å thanks to the availability of synchroton X-rays from the storage ring DORIS. The use of fast Fourier methods as applied in the restrained least-squares minimization allowed a rapid and straightforward refinement. The overall folding of the structure has not changed from that reported at 2·9 Å resolution. Significant shifts include residues lying at turns on the surface of the structure and at the carboxy terminus. The geometry of the redox centre has changed slightly and more accurate orientations of most of the side chains were determined. There were no 'breaks' in the electron density map which becomes 'weaker' only at the site of turns and the carboxy terminus of the molecule as expected for the flexible parts of a structure. The first shell of hydration of the protein in this crystal form has been determined.

Adman and co-workers at the University of Washington (Seattle, USA) have carried out a refinement of the same structure to 2·0 Å resolution using diffractometer data (Adman, 1987, private communication). The refined atomic coordinates from both laboratories are similar with an r.m.s. difference of 0·09 Å, but those determined in the present study at 1·55 Å allow a more accurate definition of the atomic parameters. The smaller discrepancy of 0·09 Å between these two sets of coordinates derived from independently measured data and refinement is perhaps a better measure of the error in the coordinates than that derived from the Luzzati plot.

Crystals of the reduced state Cu^{+} of the protein have been obtained after addition of a dilute solution of vitamin C to a droplet containing crystals of the oxidized blue form. The colourless crystals were used for collection of diffraction data on the same synchrotron source to 1·8 Å resolution. We will analyse these data to look for possible changes in the geometry of the redox site.*

We would like to thank Miss Nathalie Pipon for her exceptional assistance in the preparation of the protein samples, crystallizations and data collection on the beam line.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1PAZ, R1PAZSF), and are available in machine-readable form from the Protein Data Bank at Brookhaven or one of the affiliated centres at Melbourne or Osaka. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37025 (as microfiche). Free copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

References

ADMAN, E. T., BEPPU, T. & WATANABE, H. (1984). *Acta Cryst.* A**40**, C-33.
ADMAN, E. T. & JENSEN, L. H. (1981). *Isr. J. Chem.* **21**, 8–12.
AGARWAL, R. C. (1978). *Acta Cryst.* A**34**, 791–809.
BAKER, E. N. & DODSON, E. J. (1980). *Acta Cryst.* A**36**, 559–572.
CHOTHIA, C. & LESK, A. M. (1982). *J. Mol. Biol.* **160**, 309–323.
GUSS, J. M. & FREEMAN, H. C. (1983). *J. Mol. Biol.* **169**, 521–563.
HENDRICKSON, W. A. & KONNERT, J. H. (1980). In *Computing in Crystallography*, edited by R. DIAMOND, S. RAMASESHAN & K. VENKATESAN, pp. 13.01–13.25. Bangalore: Indian Academy of Sciences.
HORMEL, S., ADMAN, E. T., WALSH, K. A., BEPPU, T. & TITANI, K. (1986). *FEBS Lett.* **197**, 301–304.
JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
KAKUTANI, T., WATANABE, H., ARIMA, K. & BEPPU, T. (1981*a*). *J. Biochem.* **89**, 463–472.
KAKUTANI, T., WATANABE, H., ARIMA, K. & BEPPU, T. (1981*b*). *J. Biochem.* **89**, 453–461.
LUZZATTI, V. (1952). *Acta Cryst.* **5**, 802–810.
MACHIN, P. A., WONACOTT, A. J. & MOSS, D. (1983). *Daresbury News*, **10**, 3–9.
MATTHEWS, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
NORRIS, G. E., ANDERSON, B. F. & BAKER, E. N. (1986). *J. Am. Chem. Soc.* **108**, 2784–2785.

PENFIELD, K. W., GEWIRTH, A. A. & SOLOMON, E. I. (1985). *J. Am. Chem. Soc.* **107**, 4519–4529.
PETRATOS, K. (1984). PhD Thesis, Wayne State Univ., Detroit, USA.
PETRATOS, K., BANNER, D. W., BEPPU, T., WILSON, K. S. & TSERNOGLOU, D. (1987). *FEBS Lett.* **218**, 209–214.
RAMAKRISHNAN, C. & RAMACHANDRAN, G. N. (1965). *Biophys. J.* **5**, 909–933.
ROSSMANN, M. G. (1979). *J. Appl. Cryst.* **12**, 225–238.
WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1973). *Acta Cryst.* B**29**, 943–956.
WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.

Acta Cryst. (1988). B**44**, 636–645

Charge-Density Study of Boron Nitrilotriacetate, $C_6H_6BNO_6$, at 100 K: a Comparison of Polar Bonds

BY P. MOECKLI AND D. SCHWARZENBACH

Institut de Cristallographie, University of Lausanne, BSP Dorigny, CH-1015 Lausanne, Switzerland

H.-B. BÜRGI AND J. HAUSER

Laboratorium für chemische Kristallographie, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

AND B. DELLEY

Paul Scherrer Institut für Nuklearforschung, RCA Laboratory, Badenerstrasse 569, CH-8048 Zürich, Switzerland

(*Received* 1 *December* 1987; *accepted* 3 *June* 1988)

Abstract

The charge density in the cage molecule boron nitrilotriacetate has been determined experimentally at 100 K from X-ray (Mo $K\alpha$, $\bar{\lambda} = 0{\cdot}71069$ Å) diffraction data measured to $(\sin\theta/\lambda)_{max} = 1{\cdot}08$ Å^{-1}, and calculated theoretically using an efficient and accurate density functional approach. The structure is noncentrosymmetric, orthorhombic space group $Pn2_1a$, $a = 10{\cdot}468$ (2), $b = 10{\cdot}193$ (7), $c = 6{\cdot}754$ (3) Å, $Z = 4$. The electron density has been refined by least squares to $R_F = 0{\cdot}77\%$. The model includes aspherical rigid atoms represented by a multipole expansion truncated at the hexadecapole level, spherical H atoms, isotropic secondary extinction and the standard structural parameters. Monopolar radial functions were either exponential functions, or free-atom valence-shell functions modified by the κ formalism. The deformation density was constrained to the noncrystallographic molecular symmetry $3m$; partially relaxing this symmetry did not reveal features attributable to intermolecular interactions. Static and dynamic maps with infinite and finite resolutions are presented for both types of monopolar radial functions. The theoretical charge density is in good agreement with the observations. Charge densities in the six bonding regions C—C, B—N, C—N, B—O, C—O and C=O are discussed in terms of peak height and bond polarity: the more electron rich the atoms are, the lower are the peaks; bond peaks tend to be displaced towards the more electronegative atom. In contradiction with chemical expectation, B—N appears to be more polar than B—O, and C—N more polar than C—O. This is due to the representation of the bonding density as ρ(observed) – ρ(promolecule).

0108-7681/88/060636-10$03.00

Introduction

Fig. 1 shows a plot of boron nitrilotriacetate, $C_6H_6BNO_6$ (NTA-B), and the numbering of the atoms. An earlier determination of the crystal and molecular structure of this compound (Müller & Bürgi, 1984) suggested a number of reasons for studying details of the charge-density distribution. Some of these are of a chemical, others of a methodological nature.*

Chemically, NTA-B shows six different types of bonds between B, C, N and O atoms. This fact allows

* For the terminology used in this paper see Coppens (1982).

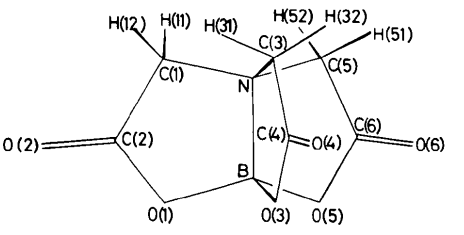

Fig. 1. Numbering of the atoms of the NTA-B molecule.