Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Auracyanin B structure in space group $P6_5$

Mihwa Lee, Megan J. Maher, Hans C. Freeman and J. Mitchell Guss*

School of Molecular and Microbial Biosciences, University of Sydney, NSW 2006, Australia

Correspondence e-mail: m.guss@mmb.usyd.edu.au

Received 6 April 2003
Accepted 23 June 2003

PDB Reference: auracyanin B, 1ov8, r1ov8sf.

The structure of auracyanin B, a 'blue' copper protein produced by *Chloroflexus aurantiacus*, has previously been solved and refined in the hexagonal space group $P6_422$ with a single molecule in the asymmetric unit. The protein has now been crystallized in space group $P6_5$, with unit-cell parameters $a = b = 115.9$, $c = 108.2$ Å. In the new crystal form, the asymmetric unit contains four protein molecules. The structure has been solved by molecular replacement and refined at 1.9 Å resolution. The final residuals are $R = 19.2\%$ and $R_{\rm free} = 21.9\%$. In relation to the earlier crystal structure, the doubling of the unit-cell volume and the lower symmetry are explained by small rotations of the molecules with respect to one another.

1. Introduction

Chloroflexus aurantiacus is a thermophilic green photosynthetic bacterium which produces two closely related 'blue' copper proteins, auracyanin A and auracyanin B (Bond *et al.*, 2001; McManus *et al.*, 1992; Van Driessche *et al.*, 1999). Both proteins are likely to be involved in photosynthetic electron transfer. The genes code for 162- and 235-residue peptides, respectively. Each peptide comprises a soluble functional domain of 140 residues preceded by an N-terminal tail (GenBank accession Nos. AF494277, AAB38318; Van Driessche *et al.*, 1999). In auracyanin A, the 22-residue N-terminal tail is a signal peptide normally lost on export of the protein to the periplasm (Van Driessche *et al.*, 1999). The N-terminus of the soluble domain is a modified cysteine, probably acetyl-*N*-cysteine-*S*-glycerol, which is thought to act as a membrane anchor *via* esterified fatty-acid chains. In contrast, auracyanin B has a 95-residue N-terminal tail which displays a high degree of sequence identity with membrane-tethering peptides identified in other proteins (Bond *et al.*, 2001). The sequences of the soluble functional domains of auracyanins A and B are 38% identical and the proteins have identical redox properties (Rooney *et al.*, 2003). The reasons why the organism produces two such closely similar proteins are not yet clear.

The structure of the 140-residue soluble domain of auracyanin B was solved previously in space group $P6_422$ (unit-cell parameters $a = b = 115.7$, $c = 54.6$ Å), with one protein molecule per asymmetric unit (Bond *et al.*, 2001; PDB code 1qhq). The structure was solved by MAD and refined at 1.55 Å resolution. The molecule has a typical cupredoxin fold closely resembling that of the bacterial cupredoxin azurin.

We here report a new crystal form of auracyanin B. The unit cell is hexagonal, as in the previously characterized form, but the *c* axis is approximately doubled and the number of molecules in the asymmetric unit is increased from one to four.

We have successfully refined this new form of auracyanin B in space group $P6_5$ (unit-cell parameters $a = b = 115.9$, $c = 108.2$ Å). The molecular and crystal structures are closely similar to those previously refined in space group $P6_422$ with a single molecule in the asymmetric unit. Small differences in packing account for the lower symmetry.

If the only change from the $P6_422$ unit cell were a doubling of the c axis, then the new structure would be in space group $P6_522$ with two molecules in the asymmetric unit. The diffraction data could indeed be scaled and merged with good statistics in space group $P6_522$ and two molecules per asymmetric unit could be located by molecular replacement. However, the hypothesis that the space group is $P6_522$ was rejected when the structure failed to refine.

Table 1
Statistics of the data and refinement.

Values in parentheses refer to the highest resolution bin.

Data collection	
Space group	$P6_5$
Unit-cell parameters (Å)	$a = b = 115.9$, $c = 108.2$
Wavelength (Å)	1.08
Resolution (Å)	1.90
Mosaicity (°)	0.32
No. of observations	458718
No. of unique reflections	63349
Completeness (%)	97.7 (96.8)
Redundancy	2.7 (2.7)
$R_{\rm merge}$ (%)	5.9 (46.9)
Average $I/\sigma(I)$	16.9 (2.5)
Refinement	
R (%)	19.2 (27.5)
$R_{\rm free}$ (%)	21.9 (29.6)
No. of reflections in test set	3227 (5.1%)
Protein atoms (including Cu)	4148
Heteroatoms†	27
Water molecules	559
R.m.s.d. bond length‡ (Å)	0.01
R.m.s.d. bond angle‡ (°)	1.58
Mean protein B factor, all non-H atoms (Å^2)	11.8
Mean water B factor (Å^2)	40.6
Estimated standard uncertainties‡	
Coordinates, based on residual R (Å)	0.13
Coordinates, based on $R_{\rm free}$ (Å)	0.12
Ramachandran plot, residues in§	
Most favoured regions (%)	86.1
Additional allowed regions (%)	13.9
Disallowed regions (%)	0.0

† Five SO_4^{2-} and two Cl^-. ‡ Diffraction-component precision index (Cruickshank, 1999) calculated using *REFMAC*5 (Murshudov *et al.*, 1999). § Calculated using *PROCHECK* (Laskowski *et al.*, 1993).

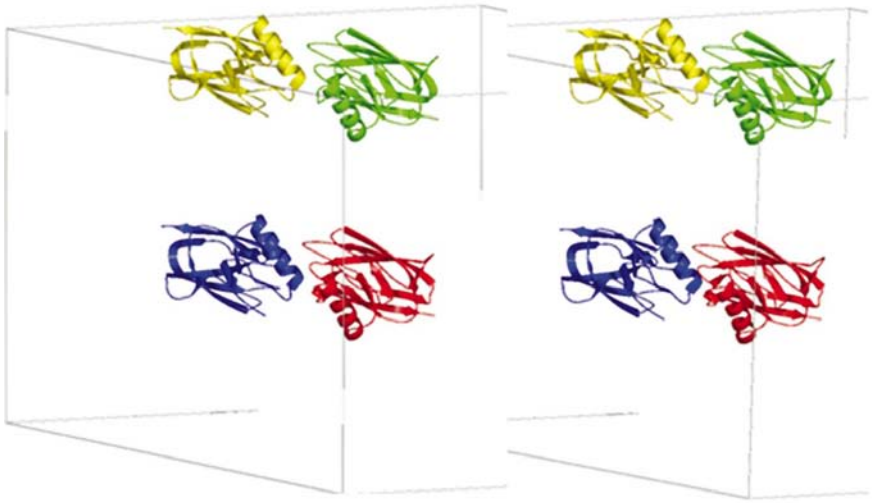

Figure 1
Stereoview of the molecular packing in crystals of auracyanin B in space group $P6_5$. The molecules designated A, B, C and D in the text and in Table 2 are coloured yellow, green, red and blue, respectively. The more extensive intermolecular contacts involving the hydrophobic surface and chloride ion occur between A (yellow) and B (green) and between C (red) and D (blue). If molecule A is equated with the single molecule in the asymmetric unit of the structure of auracyanin B in space group $P6_422$, then the approximate positions of the other three molecules can be generated by applying the crystallographic twofold axes and cell translations of the higher-symmetry space group (x, $x - y$, $2/3 - z$; x, $x - y$, $-1/3 - z$; x, y, $-1 + z$).

2. Materials and methods

2.1. Crystallization

Auracyanin B was isolated from *C. aurantiacus* and purified as described previously (Bond *et al.*, 2001). Crystals were grown by the hanging-drop vapour-diffusion method using 2 µl of protein solution (9 mg ml^{-1}) mixed with 2 µl of reservoir solution (0.1 *M* HEPES pH 7.5, 2.25 *M* Li_2SO_4) at 277 K. Prior to data collection, the crystals were successively transferred to cryoprotectant solutions consisting of the mother-liquor solution containing 5, 10 and 12.5%(*v*/*v*) glycerol, respectively, and were flash-cooled in a stream of nitrogen gas at 100 K.

2.2. Data collection and processing

Synchrotron data to 1.9 Å resolution were recorded from a single crystal on beamline 7-1 at the Stanford Synchrotron Radiation Laboratory. The data were recorded on a MAR345 imaging plate (X-ray Research, Hamburg). A total of 84 successive frames were collected, with a 0.5° oscillation for each frame. The diffraction data were processed and scaled with the *HKL* suite of programs: *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The data could be processed in either space group $P6$ or $P622$ with similar results ($R_{\rm merge}$ = 5.9% in $P6$, 6.4% in $P622$). Statistics for the data in $P6$ are presented in Table 1. Insufficient data along the reciprocal axis (0, 0, l) were recorded to identify reliably the nature of the screw axis.

The unit-cell parameters of the new and previously observed crystal forms are almost identical, with the exception of a doubling of the c axis from 54.6 to 108.2 Å. Since the two crystal forms are so closely related, the hypothesis that the crystals were twinned was tested by a statistical analysis using *TRUNCATE* (Collaborative Computational Project, Number 4, 1994). The Stanley factor $\langle I^2\rangle/\langle I\rangle^2$ (Stanley, 1972; Yeates, 1997) was found to be 2.4,

suggesting that there was no twinning. Standard values of the $\langle I^2\rangle/\langle I\rangle^2$ ratio are ∼2.0 for untwinned data and ∼1.5 for twinned data (Stanley, 1972; Yeates, 1997).

2.3. Structure solution and refinement

The crystal structure was solved by molecular replacement using *AMoRe* (Navaza, 2001). The auracyanin B structure in space group $P6_422$ (Bond *et al.*, 2001; PDB code 1qhq) was used as the search model. The Cu atom, other ions and solvent molecules were deleted. If the complexities that arose during the structure solution, described below, had been apparent at the time, reflection data at the Cu edge would have been recorded to enable a MAD solution of the structure.

Initial attempts to solve the structure by molecular replacement were carried out in the high-symmetry space groups $P622$, $P6_122$, $P6_222$, $P6_322$, $P6_422$ and $P6_522$. In each case, two molecules could be located in the asymmetric unit. The best agreement factors were obtained in space groups $P6_222$ and $P6_522$. The models were refined in stages using rigid-body, simulated-annealing, torsion and Cartesian parameterization using *CNS* (Brünger *et al.*, 1998). The use of torsion angles as variables is known to increase the radius of convergence of the refinement procedure and is useful to escape local minima (Rice & Brünger, 1994). In all cases,

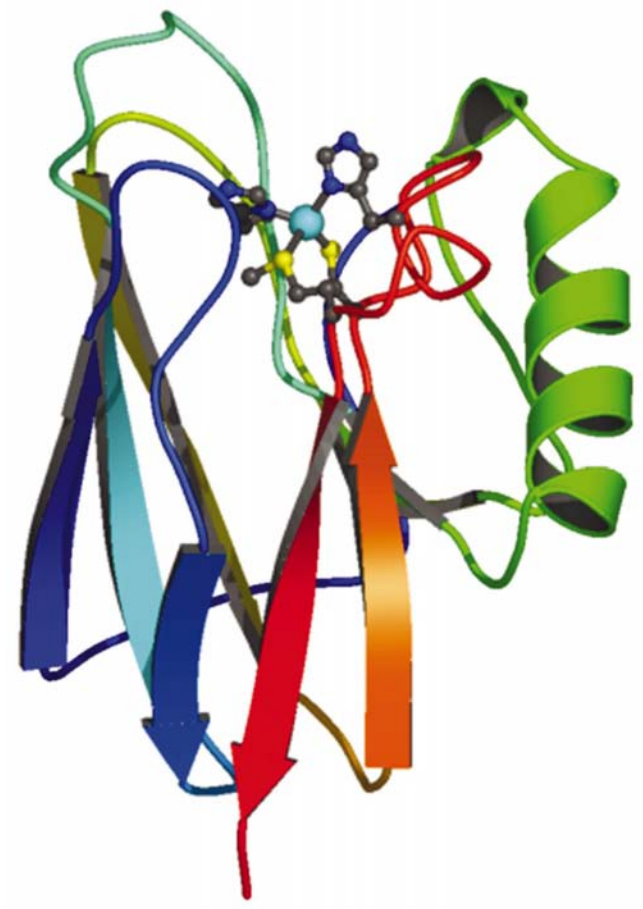

Figure 2
The molecule of auracyanin B. The polypeptide chain is coloured from blue at the N-terminus to red at the C-terminus. The Cu atom is represented by a cyan sphere. The Cu-binding residues are shown in a ball-and-stick representation.

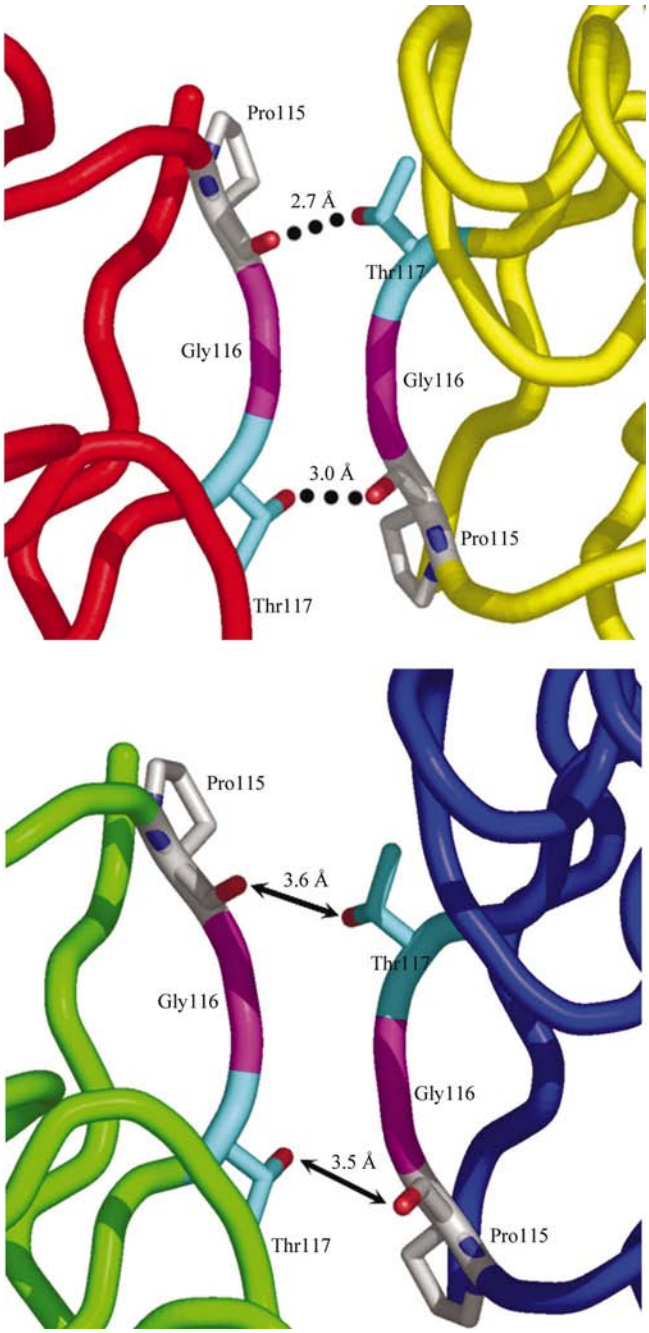

Figure 3
Details of a second contact region between molecules, illustrating the deviation from exact crystallographic symmetry. The top panel shows the contact between molecule *A* (yellow) and *C* (red) transformed by the symmetry operation $(y, 1 + y - x, 7/6 + z)$. The lower panel shows the equivalent contact between molecules *D* (blue) and *B* (green) transformed by the symmetry operation $(y, 1 + y - x, 1/6 + z)$. Since these contacts do not involve molecules related by crystallographic symmetry, the reverse contacts are similar but not identical (Table 4). The right-hand molecule in the lower panel is displaced downwards in comparison with the upper panel, resulting in the loss of the hydrogen bonds (dotted).

Table 2
Comparisons between the four molecules (C^{α} atoms) in the asymmetric unit.

Rotation angles were calculated using *LSQMAN* (Kleywegt *et al.*, 2001).

	B		*C*		*D*	
	R.m.s.d. (Å)	Rotation angle (°)	R.m.s.d. (Å)	Rotation angle (°)	R.m.s.d. (Å)	Rotation angle (°)
A	0.10	179.3	0.05	180.0	0.10	1.71
B			0.10	1.7	0.06	180.0
C					0.10	179.4

Table 3
Refinement of the auracyanin B structure in different space groups.

	$P6_5$	$P6_522$	$P6_222$
Residual R (%)	19.2	32.9	32.0
Residual R_{free} (%)	22.0	41.6	39.4

Table 4
Intermolecular contact distances (Å) in the current structure ($P6_5$) and in the previously published structure ($P6_422$).

(*a*) Pro115O–Thr117 O^{γ}.

	Pro115*A*–Thr117C^{i}	Pro115*C*–Thr117A^{ii}	Pro115*D*–Thr117B^{iii}	Pro115B–Thr117D^{iv}
$P6_5$	2.7	3.0	3.5	3.6
$P6_422$†		2.6 (3.5)		

(*b*) Gly116 C^{α}–Gly116 C^{α}.

	Gly116*A*–Gly116C^{i}, Gly116*C*–Gly116A^{ii}	Gly116*D*–Gly116B^{iii}, Gly116*B*–Gly116D^{iv}
$P6_5$	3.4	4.7
$P6_422$	4.5	

Symmetry codes: (i) $x, 1 + y - x, -5/6 + z$; (ii) $x - y, x, 11/6 + z$; (iii) $y, 1 + y - x, 1/6 + z$; (iv) $y - x, x, 5/6 + z$. † The side chain of Thr117 in $P6_422$ has two conformations.

highly significant peaks of approximately equal height occurred in electron-density difference maps at the expected positions of the Cu atoms, which had been omitted from the initial models. The best model in $P6_222$ included two Cu atoms and 95 water molecules and refined to residuals R = 36.4% and R_{free} = 39.8%. The model refined in $P6_522$ with two Cu atoms and 100 water molecules had similar residuals R = 37.0% and R_{free} = 41.3%.

The failure to refine the structure successfully in any of the allowed higher-symmetry space groups led us to try refinement in the lower-symmetry space groups $P6$, $P6_1$, $P6_2$, $P6_3$, $P6_4$ and $P6_5$ with four molecules in each asymmetric unit. Using *AMoRe*, the best molecular-replacement solutions were found in $P6_2$ and $P6_5$. The correlation coefficients and R factors for data in the resolution range 30–4.0 Å were 0.755 and 49.0% ($P6_2$), and 0.760 and 48.0% ($P6_5$), respectively. Rigid-body refinement for data in the resolution range 30–4.0 Å gave better statistics in $P6_5$ (correlation coefficient 0.83, residual R = 29.7%) than in $P6_2$ (0.79, 36.8%). Several rounds of positional and temperature-factor refinement in both space groups using *REFMAC*5 (Murshudov *et al.*, 1999) with data in the range 30–2.0 Å resulted in residuals of R = 27.0% and R_{free} = 30.0% in $P6_5$, compared with residuals of 35% and 38.7% in $P6_2$, respectively. It was concluded that the correct space group is $P6_5$. Further refinement was carried out using *CNS* (Brünger *et al.*, 1998) with all data to 1.9 Å resolution using strict fourfold non-crystallographic symmetry (NCS) restraints. The refinement consisted of simulated-annealing and positional refinement with manual model building using the program *O* (Jones *et al.*, 1991). The inclusion of a Cu atom and 118 water molecules for each monomer improved the residuals R and R_{free} to 25.1 and 26.5%. Refinement of the model was completed with *REFMAC*5 with TLS (Winn *et al.*, 2001). Strong initial NCS restraints were progressively relaxed until they were removed entirely in the final cycles. The residuals R and R_{free} converged to 19.2 and 21.9%, respectively. The stereochemical quality of the structure was validated using *PROCHECK* (Laskowski *et al.*, 1993) and *WHATCHECK* (Hooft *et al.*, 1996). All residues fall into the allowed regions of the Ramachandran plot. The refinement statistics are included in Table 1.

3. Results and discussion

3.1. Description of the structure

The final model consists of four molecules of auracyanin B with 556 protein residues, 559 water molecules, four Cu atoms, two chloride ions and five sulfate ions in space group $P6_5$ (Fig. 1). The first N-terminal residue (Ala) was not located in electron density, so that each molecule comprises 139 residues (residues 2–140). One residue (Gln50) in each molecule was modelled in two conformations. The estimated standard uncertainty in the position of an average atom, evaluated as the Cruickshank diffraction-component precision index (DPI; Cruickshank, 1999), is 0.13 Å.

Since attempts to refine the structure in a high-symmetry space group failed, the relationship between the four molecules in the asymmetric unit is of particular interest. We refer to the four molecules as *A*, *B*, *C* and *D*. Each molecule is represented by its 139 crystallographically located C^{α} atoms. When the molecules are superposed in pairs, the r.m.s. differences ($\leq$0.1 Å) are smaller than the estimated standard uncertainty in the model (DPI = 0.13 Å) and rotations from the orientations that the molecules would have in a higher-symmetry space group are small ($\leq$1.7°) (Table 2). There are no significant differences among the metal-binding sites. There are only four residues in which the positions of corresponding side-chain atoms in molecules *A*, *B*, *C* and *D* vary by more than 0.25 Å (Arg22, Asn40, Gln77 and Asp89). However, these residues have higher atomic temperature factors, $B \geq 30$ Å^2, than the mean temperature factor of the model, $\langle B \rangle$ = 13 Å^2 (Table 1), and are therefore less reliably determined.

In view of the very small deviations from higher symmetry, we cite further evidence that the space group is correctly assigned. The arrangement of molecules in $P6_5$ involves a local

translation of approximately (0, 0, 0.5) relating the pairs of molecules *A*/*D* and *B*/*C*. The consequence of this approximate symmetry is seen in the diffraction data as a systematic effect on intensity with reflection parity ($\langle I(l \text{ even})/I(l \text{ odd})\rangle = 1.3$ for all data to 1.9 Å resolution and 1.7 for data to 10 Å resolution). There are no significant systematic effects according to parity for other reflection classes. Following the completion of the refinement in space group $P6_5$, attempts were made to revert to space groups $P6_222$ and $P6_522$. Molecules *A* and *D* from the $P6_5$ structure were treated as an asymmetric unit of two molecules. The origin was fixed by placing the local twofold axis relating pairs of molecules in $P6_5$ along the crystallographic twofold axis in $P6_522$. This solution was identical to that obtained directly by molecular replacement in these space groups (see above). After ten cycles of rigid-body refinement followed by 20 cycles of positional refinement using *REFMAC*5 (Murshudov *et al.*, 1999), the residuals were $R = 32.0\%$ and $R_{\rm free} = 39.4\%$ in $P6_222$ and $R = 32.9\%$ and $R_{\rm free} = 41.6\%$ in $P6_522$. Further refinement did not improve the R and $R_{\rm free}$ values (Table 3). Even though the deviations from higher symmetry are relatively small (Table 2), all the atoms of the related molecules are affected and result in highly significant differences between the residuals.

3.2. Comparison with the structure in $P6_422$

The molecular structure of auracyanin B in space group $P6_5$ is closely similar to the previously solved structure in space group $P6_422$. The core of the molecule is a sandwich of two β-sheets formed by eight polypeptide strands in a typical cupredoxin fold (Bond *et al.*, 2001) (Fig. 2). The crystals of both forms were grown under the same conditions and the structures contain the same additional SO_4^{2-} and Cl^- ions. There are no significant differences between the main-chain C^α positions in the two structures. However, there are significant differences at a number of side chains. In the $P6_5$ structure, only one residue (Gln50) is modelled with two conformations, compared with eight residues in the $P6_422$ structure (Asn12, Glu13, Thr14, Glu20, Val43, Leu45, Arg111 and Thr117). In both structures, the residues with two side-chain conformations generally have higher temperature factors, $B \geq 30$ Å^2, than the mean value of the model, $\langle B\rangle = 13$ Å^2. Significant differences between side-chain conformations, as indicated by r.m.s. differences ≥ 0.3 Å, occur at eight residues (Arg22, Ser36, Asn40, Asp67, Gln77, Leu83, Asp89 and Thr139), as well as at five of the residues that are modelled with two conformations in one or the other structure (Asn12, Thr14, Glu20, Leu45 and Gln50). Some of the residues with significant differences in side-chain conformation (Asn12, Thr14, Glu20, Thr36 and Thr139) are involved in intermolecular contacts, so that they contribute directly to the existence of the two crystal forms.

The molecular packing of auracyanin B in the two crystal forms was compared by examining the intermolecular contacts in each structure. In the $P6_422$ structure, there is an intermolecular contact across a crystallographic twofold axis ($x, y, z \rightarrow x, x-y, 2/3-z$) involving a hydrophobic surface and a Cl^- ion (Bond *et al.*, 2001). There are equivalent intermolecular contacts in the $P6_5$ structure owing to local twofold axes of symmetry between the pairs of molecules *A* and *B* and between *C* and *D*. However, another intermolecular contact in the $P6_422$ structure occurs in two slightly different forms in the $P6_5$ structure, accounting for its lower symmetry. This difference can be visualized by looking at the contacts between molecules involving the residues Pro115, Gly116 and Thr117 (Table 4 and Fig. 3). In the $P6_422$ structure this region makes a close contact, including one direct hydrogen bond to a molecule related by a crystallographic twofold axis ($x, y, z \rightarrow y, x, 4/3-z$). The 2.6 Å hydrogen bond is from the carbonyl O atom (Pro115) to O^γ in one of the two conformers of the side chain of Thr117. In the $P6_5$ structure, the four equivalent contacts among molecules *A*, *B*, *C* and *D* fall into two groups: two contacts are the same in the high-symmetry structure (*A*–*C* and *C*–*A*) and the other two contacts (*B*–*D* and *D*–*B*) involve the second conformer of the side chain of Thr117 that was observed in the $P6_422$ structure (Table 4*a*; Fig. 3). The difference between the two sets of contacts is emphasized by the large change in the distances between the C^α atoms of Gly116 (Table 4*b*). The result is that a residue that was observed in two conformations in the $P6_422$ structure has frozen into two different conformations in the $P6_5$ structure, resulting in lower overall symmetry.

4. Conclusions

Crystals of auracyanin B from *C. aurantiacus* in space group $P6_5$, with four protein molecules per asymmetric unit, have a structure similar to that previously reported for crystals in a space group of higher symmetry, $P6_422$. Although the symmetry operation that relates two pairs of molecules in the asymmetric unit is very close to being a crystallographic twofold axis, the structure could only be refined in space group $P6_5$. The only significant differences between the molecular structures of auracyanin B in space groups $P6_5$ and $P6_422$ involve side-chain conformations. Some of these differences help to explain the existence of two crystal forms.

This work was supported by the Australian Research Council (grant DP0208320 to JMG and HCF). MM is an Australian Research Council Postdoctoral Fellow. Access to the facilities of the Stanford Synchrotron Radiation Laboratory (SSRL) was made possible by a travel grant from the Access to Major National Facilities Program administered by the Australian Nuclear Science and Technology Organization. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences. The auracyanin was supplied by Professor R. E. Blankenship, Arizona State University.

References

Bond, C. S., Blankenship, R. E., Freeman, H. C., Guss, J. M., Maher, M. J., Selvaraj, F. M., Wilce, M. C. J. & Willingham, K. M. (2001). *J. Mol. Biol.* **306**, 47–67.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

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

Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.

Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.

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

Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M. & Jones, T. A. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 353–356. Dordrecht: Kluwer Academic Publishers.

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

McManus, J. D., Brune, D. C., Han, J., Sanders-Loehr, J., Meyer, T. E., Cusanovich, M. A., Tollin, G. & Blankenship, R. E. (1992). *J. Biol. Chem.* **267**, 6531–6540.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.

Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rice, L. M. & Brünger, A. T. (1994). *Proteins Struct. Funct. Genet.* **19**, 277–290.

Rooney, M. B., Honeychurch, M. J., Selvaraj, F. M., Blankenship, R. E., Bond, A. M. & Freeman, H. C. (2003). *J. Biol. Inorg. Chem.* **8**, 306–317.

Stanley, E. (1972). *J. Appl. Cryst.* **5**, 191–194.

Van Driessche, G., Hu, W., Van De Werken, G., Selvaraj, F., McManus, J. D., Blankenship, R. E. & Van Beeumen, J. J. (1999). *Protein Sci.* **8**, 947–957.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.

Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.