RESEARCH PAPERS

Acta Cryst. (1997). D53, 493–506

Primary Sequence and Refined Tertiary Structure of *Pseudomonas fluorescens* Holo Azurin at 2.05 Å

XAVIER LEE,[a]* TANYA DAHMS,[b] HOA TON-THAT,[a] DAO-WEI ZHU,[a]† JOHN BIESTERFELDT,[a] PATRICIA H. LANTHIER,[c] MAKOTO YAGUCHI[c] AND ARTHUR G. SZABO[d]

[a]*Department of Cancer Biology, Research Institute, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, Ohio 44195, USA,* [b]*Department of Biological Sciences, 1392 Lilly Hall of Life Sciences, Purdue University, West Lafayette, IN 47906–1392, USA,* [c]*Institute for Biological Sciences, National Research Council of Canada, Ottawa, 100 Sussex Drive, Ontario K1A 0R6, Canada, and* [d]*Department of Chemistry and Biochemistry, University of Windsor, 401 Sunset Avenue, Windsor, Ontario N9B 3P4, Canada. E-mail: lee@xtal.ri.ccf.org*

(*Received 23 April 1996; accepted 28 January 1997*)

Abstract

This paper reports the primary sequence and refined crystal structure of *Pseudomonas fluorescens* holoazurin. The crystal structure has been determined by molecular replacement on the basis of the molecular model of azurin from *Alcaligenes denitrificans*, and refined by the method of molecular dynamics simulation and energy-restrained least-squares methods. *P. fluorescens* was crystallized in the orthorhombic space group $P2_12_12_1$ with unit-cell dimensions $a = 31.95$, $b = 43.78$, $c = 78.81$ Å. The asymmetric unit is composed of only one molecule. The final R value is 16.7% for 6691 reflections to a resolution of 2.05 Å. This azurin structure shows some interesting features at His35 and His83. Part of the main chain of strand 3 including His35 O are involved in the contact between two symmetrically related molecules. *P. fluorescens* is also compared with the other azurin structures in terms of primary sequence, crystal packing, solvent structure and Cu-site geometry. The difference in fluorescence decay behavior of two holoazurins from *P. fluorescens* and *P. aeruginosa* and the correlation between the fluorescence quenching and electron transfer are discussed.

1. Introduction

Azurin is classified as a member of the blue copper proteins, so called because of their intense band in the visible part of the spectrum ($\lambda_{max} = 600$–625 nm). These proteins also have distinctive electron paramagnetic resonance and redox potentials (for reviews, see Adman, 1985, 1991). Azurin and plastocyanin are the best characterized of all the blue copper proteins. Plastocyanin is involved in the photosynthetic systems of plants and some algae and azurin in the redox systems of certain bacteria (Ryden, 1984). For a variety of reasons extensive chemical, spectroscopic and kinetic studies of these two proteins have been carried out (Adman, 1985), and amino-acid sequences are known for at least 12 azurins and 30 plastocyanins (Ryden, 1984). X-ray crystallographic structural analysis have provided the three-dimensional structures of plastocyanin from at least three different sources (Colman *et al.*, 1978; Guss & Freeman, 1983), and at least three azurins, from *Pseudomonas aeruginosa* (Adman, Stenkamp, Sieker & Jensen, 1978; Adman & Jensen, 1981), *Alcaligenes denitrificans* (Norris, Anderson & Baker, 1983; Baker, 1988) and *Pseudomonas denitrificans* (Korszun, 1987), Recent high-resolution studies of mutants of *P. aeruginosa* and at two different pH values offered a detailed picture of the copper site, the environment of His35, and also ruled out the possible role of His35 in the electron-transfer mechanism (Nar, Messerschmidt, Huber, van de Kamp & Canters, 1991*a*; Romero *et al.*, 1993).

Most of the interest in the blue proteins has centered on elucidating details of the copper site. The structural analysis of plastocyanin showed that the Cu atom (Cu^{II}) forms three strong ligand bonds (2.0–2.1 Å), with one cysteine (S^-) and two histidine (N) ligands, and a considerably weaker bond (2.9 Å) to a methionine (S) ligand. The geometry is best described as distorted tetrahedral. The high-resolution crystal structures of amicyanins from *Paracoccus denitrificans* (Durley & Mathews, 1994) and from *Thiobacillus versutus* (Romero *et al.*, 1993) showed the same features for the Cu site. The azurins were found to share the same four ligands, in a similar geometry (Adman *et al.*, 1978; Adman & Jensen, 1981), and an additional fifth interaction to a main-chain carbonyl O atom (Baker, 1988; Nar *et al.*, 1991*b*) based on the high-resolution refined structure.

The blue absorption spectrum, the high-energy fluorescence spectrum and the ESR spectral behavior of azurin has stimulated enormous interest in its spectroscopic properties. Mutation of residues surrounding Trp48 causes substantial changes in the fluorescence

† Current address: MRC group in Molecular Endocrinology, CHUL Research Center and Laval University, 2705 Boulevard Laurier, Ste-Foy, Québec, G1V 4G2, Canada.

lifetimes and in the distribution of the lifetimes of the mutants (Gilardi, Mei, Rosato, Canters & Finazzi, 1994). The fluorescence spectra of the single tryptophan residue Trp48 in both the holo and apo forms of azurin from *P. aeruginosa* and *P. fluorescens* have a spectral maximum at 308 nm. This is very unusual for tryptophan in a protein (Grinvald, Schlessinger, Pecht & Steiberg,. 1975) as the emission usually takes place at 340–350 nm. Time-resolved fluorescence studies from our laboratory indicated that there is a conformational heterogeneity of the tryptophan in the holo form of these two azurins (Szabo, Stepanik, Wagner & Young, 1983; Hutnik & Szabo, 1989). However, the relative proportions of the fluorescence components were very different, indicating that the relative amounts of the conformers in these two azurins are different. Recent studies further showed that the fluorescence quenching behaviors are very different between holo and apo azurins and between azurin in solution and in the crystalline state. (Dahms & Szabo, 1997). In an effort to understand the structural difference between azurins of different sources and the difference in quenching we undertook a detailed crystallographic study of the *P. fluorescens* azurin.

In this paper we present the primary sequence and the refined tertiary structure of *P. fluorescens* and its comparison with the existing structures of the other azurins.

2. Materials and methods

2.1. *Purification*

Azurin from *P. fluorescens* (ATCC 13525) biotype A was prepared according to a previous report (Hutnik & Szabo, 1989). The copper-to-protein stoichiometry was determined by atomic absorption and amino-acid analysis to be 1:1. The spectral ratio (A_{625}/A_{280}) was 0.55, indicating that there is no significant amount of apoazurin in the sample. There was only one single band at 14.4K in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) and one band at pH 6.55 in an isoelectric focusing experiment (Hutnik & Szabo, 1989).

2.2. *Sequencing*

Twelve azurin amino-acid sequences have been deposited in the Protein Data Bank (Swiss-Prot, Vol. 28), all of which exhibit a high degree of sequence similarity. The previously unknown sequence of *P. fluorescens* (ATCC 13525) biotype A (Stanier, 1947) was determined using a sequencing strategy modified from that of Dell & Morris (1977). The protein was digested by trypsin to generate tryptic peptides. The peptide mixture was purified by reverse-phase high-pressure liquid chromatography (HPLC) and the amino-acid sequence of the purified tryptic peptides was determined by gas-phase microsequencing. The assignment of amino acids to each peptide was supported by electrospray mass spectrometry. The tryptic peptide sequences were then aligned by homology with other known *P. fluorescens* azurins, thus eliminating the necessity for obtaining overlap information. The proposed sequence was found to be consistent with electrospray mass spectrometric analysis and amino-acid composition.

2.2.1. *Reduction and alkylation.* A 500 µg portion of azurin was dissolved in 500 µl of 0.1 *M* Tris–HCl buffer pH 8.5 containing 6 *M* guanidium chloride. This sample was reduced by the addition of 0.0694 mg (450 nmol) of fresh dithiothreitol (Calbiochem, La Jolla, CA, USA) under argon at room temperature for 1 h. *S*-carboxymethylation was then effected by the dropwise addition of 1.256 mg (10 equivalents per mole of thiol) of fresh iodoacetic acid (Sigma Chemical Co., St Louis, MO, USA) followed by vortexing. The solution was flushed with argon, capped and incubated for 60 min in the dark at room temperature. The sample was dialyzed extensively against water at 277 K for 24 h and then lyophilized.

2.2.2. *Tryptic cleavage.* Trypsin (bovine pancreas, sequencing grade) was obtained from Boehringer Mannheim. A 500 µg portion of reduced and alkylated protein was dissolved in 500 µl of 0.1 *M* ammonium bicarbonate buffer pH 8.1 containing 0.1 m*M* $CaCl_2$. Trypsin was added in a 1:100(*w*/*w*) ratio and the mixture was incubated for 1 h at 310 K. Again the addition of trypsin was repeated, followed by a further 30 min incubation at 310 K. The digestion was terminated by lyophilization of the mixture.

The resultant tryptic digest peptides were fractionated by elution with an acetonitrile gradient (1% min^{-1}) with 0.1%(*v*/*v*) trifluoroacetic acid at a flow rate of 1 ml min^{-1} from a Hamilton PRP-3 column (4.1 × 150 mm) previously equilibrated in 0.1% trifluoroacetic acid pH 2.0. The effluent was monitored at 220 nm and individual peptides were collected and lyophilized.

2.2.3. *C-terminal analysis.* Carboxypeptidase P (*Penicillium janthinellum*, sequencing grade) was obtained from Boehringer Mannheim. A 250 µg portion of reduced and alkylated protein was dissolved in 500 µl of 10 m*M* sodium acetate buffer pH 4.5 containing 0.05% Brij-35.

Carboxypeptidase P was added in a 1:100(*w*/*w*) ratio and the mixture was incubated at 310 K. Aliquots (50 µl) of reaction mixture were removed at timed intervals (0, 5, 10, 15, 30 and 60 min, and 2.5, 4 and 5 h) and immediately placed in clean eppendorf tubes containing trifluoroacetic acid (10% of final volume) to stop the reaction. The samples were evaporated using a Speed Vac (Savant).

The free amino-acid levels were determined by amino-acid composition analysis and the resulting data were plotted as time *versus* picomoles of amino acid released.

2.2.4. *Amino-acid composition analysis.* Amino-acid composition analyses of the intact protein and of purified tryptic peptides were performed with the Applied Biosystems Inc. (Foster City, CA, USA) model 420A amino-acid analysis system with 0.5–1.0 μg quantities of protein or peptide. Protein samples (2–4 μg) underwent manual vapor-phase hydrolysis using 6 *N* HCl and incubated at 423 K for 60 min under an argon atmosphere. The combined cystine and cysteine content of azurin was determined after oxidation to cysteic acid (Hirs, 1967) and vapor-phase hydrolysis as previously described. After hydrolysis, the samples were dissolved in HPLC-grade water containing 0.025% K_3 EDTA (Applied Biosystems) and an aliquot (5–10 μl) was applied to a sample slide.

2.2.4. *Automated gas-phase sequence analysis.* Automated gas-phase sequence analysis was performed on an Applied Biosystems Inc. (Foster City, CA, USA) model 475A protein sequencing system with 0.1–0.5 nmol quantities of protein and purified peptides. The samples were dissolved in 0.1% trifluoroacetic acid and applied to a glass fiber containing 0.75 mg of polybrene (Applied Biosystems).

2.2.5. *Electrospray mass spectrometry.* A Sciex API III triple quadrupole mass spectrometer (Thornhill, Ontario, Canada) fitted with a pneumatically assisted electrospray interface, and a mass range of 2400 amu e^{-1} was used to analyze samples in 5% acetic acid solutions (0.1–0.2 mg ml^{-1}).

2.2.6. *Amino-acid sequence alignments.* Alignment of amino-acid sequences and the evolutionary relationships of azurins were determined with *GeneWorks* version 2.3.0 software by Intelligenetics Inc, Mountain View, CA, USA.

2.3. *Crystallization*

The crystallization of the *P. fluorescens* holo azurin has been reported (Zhu, Dahms, Willis, Szabo & Lee, 1994). The space group is $P2_12_12_1$ with one molecule in the asymmetric unit and a packing density parameter $V_m = 1.97$ Å^3 Da^{-1} (Matthews, 1968). The cell parameters are $a = 31.95$, $b = 43.78$, $c = 78.81$ Å. The crystal diffracted to 2.05 Å.

2.4. *Data collection*

A complete set of data was collected with only one crystal on the FAST system. The reflection data were evaluated on-line using the program *MADNES* and corrected for absorption, merged and loaded using the program package *CCP*4 from the Daresbury Laboratory (Collaborative Computational Project, Number 4, 1994). There were 7076 unique reflections for the total 27 760 measurements. The $R_{\rm merge}$ $[\sum|I(K) - \langle I\rangle I|/\sum I(K)]$ was 5.0%. The completeness from ∞ to 2.05 is 95.7 and 80% between 2.05 and 2.15 Å.

2.5. *Molecular replacement*

The primary sequence of *P. fluorescens* was unknown at the time the data collection took place. The coordinates of the first molecule of *A. denitrificans* were taken from the Protein Data Bank to serve as a search model. The identity of the amino-acid composition between these two proteins was known to be 80% (Hutnik & Szabo, 1989). The program *MERLOT* in the *CCP*4 package was used to determine the location of the molecule in the unit cell. A rotation search was performed with various resolution ranges of data and different lengths of vectors. A peak appeared constantly at $\alpha = 89.00$, $\beta = 78.50$, $\gamma = 191.50°$. It was about 6σ higher than the background and 2σ higher than the second highest peak. With the same rotation program but finer grids this peak was refined to be $\alpha = 88.71$, $\beta = 78.52$, $\gamma = 195.81°$ with the data between 3 and 8 Å. The translation search was carried out with this orientation of the molecule and with the same range of data. A unique solution was found at the fractions of $x = 27.8/40$, $y = 32.1/56$ and $z = 60.4/100$. This peak was 5.72σ higher than the background and 1.6σ higher than the second highest peak. The correlation factor was 0.368. Graphics display showed that the symmetrically related molecules could be accommodated reasonably well without any overlap. The R factor $(\sum||F_o| - |F_c||/|F_o|)$ for this molecular replacement solution was 0.484 in the resolution range 3.0–8.0 Å. An electron-density map was calculated and it was interpretable.

2.6. *Refinement*

The graphics system used was a Silicon Graphics 4D-35 Personal Workstation equipped with the program *CHAIN* kindly supplied by Professor Quiochio (Baylor College of Medicine, Houston). The program used for the refinement was *X-PLOR* from Dr Brünger at Yale University (Brünger, 1992). The model used was the solution of the molecular replacement with the following modifications in the sequence. Only the sequence of the first 25 residues was determined in our laboratory before the refinement started. Substitution of the remaining sequence was based on the fact that the primary sequence of this particular azurin was likely to be very similar to those of the *P. fluorescens* biotypes B, C and D whose sequences were known (Fig. 1). Apart from the first 25 residues the remainders common to the four azurins (*A. denitrificans* and the *P. fluorescens* biotypes B, C and D) were retained. All the residues common to the biotypes but different from the *A. denitrificans* were changed to those of the biotypes. Any residues not homologous for all the biotypes were changed to alanine. During the whole process special attention was paid to the copper-site geometry. In the early six rounds of refinement only weak potentials were imposed on the active site and subsequently the model was subjected to extensive rounds of refinement without any geometrical

restraints.* Only reflections with $F > 2\sigma$ were used. The course of refinement was initiated with the data between 8.0 and 2.5 Å resolution (Table 1). Between cycles manual correction was performed using *CHAIN*. The R factor ($R = \sum|F_o - F_c|/|F_o|$) dropped from 0.484 to 0.263 after three cycles of refinement with an overall B factor of 15 Å^2. The full sequence of this azurin then became available in our laboratory. With the correct primary sequence and individual B-factor refinement, the R factor was reduced further to 0.227 after only one cycle. One intermediate cycle of refinement was carried out with the data between 8.0 and 2.3 Å before the highest resolution data were included. The R factor at that stage was 0.237. The subsequent refinements were carried out with the data between 6.0 and 2.05 Å resolution and water molecules were added after the R factor reached lower 20. The final R factor is 16.7 with 114 water molecules. The R factor between 2.05 and 2.09 Å is 0.256. During the refinement procedure the free R value was constantly monitored such that no overfitting of the diffraction data occurred. It went from 0.530 at the very beginning to 0.241 at the final cycle. After each cycle of refinement $F_o - F_c$, $2F_o - F_c$ and omit maps were used for the correction of the model. The omit maps were made with the omission of every ten residues from 1 to 129.

Table 1. *Course of refinement of azurin*

Cycles	Resolution (Å)	No. of atoms	R value	Procedure*
1	8.0–2.5	926	0.484–0.321	*A*
2	8.0–2.5	926	0.321–0.281	*A*
3	8.0–2.5	926	0.281–0.263	*A*
4	8.0–2.5	952	0.263–0.227	*B*
5	8.0–2.3	952	0.248–0.237	*B*
6	6.0–2.05	952	0.253–0.240	*B*
7	6.0–2.05	952 + 80 H_2O	0.240–0.211	*B*
8	6.0–2.05	952 + 102 H_2O	0.211–0.205	*C*
9	6.0–2.05	952 + 114 H_2O	0.205–0.195	*D*
10	6.0–2.05	952 + 114 H_2O	0.195–0.185	*D*
11	6.0–2.05	952 + 114 H_2O	0.185–0.176	*D*
12	6.0–2.05	952 + 114 H_2O	0.176–0.167	*D*

* Procedure *A* consisted of the prestage and slowcool from 3000 K. Procedure *B* consisted of the prestage, the slowcool from 3000 K and individual *B*-factor refinement. Procedure *C* consisted of the prestage, the slowcool from 2000 K and the individual *B*-factor refinement. Procedure *D* consisted of the positional and the individual *B*-factor refinement.

* The copper force field used in the program was defined by Dr Neena Summers, Monsanto Corporation, and kindly given by Dr Rosemary Durley of the Washington University School of Medicine. Dr Summers verified the parameters with seven type 1 copper proteins retrieved from the Brookhaven Protein Data Bank. The force field maintained the ligation state of the metal, in the test structures, during energy minimization and simulated annealing without significantly altering the resulting geometry.

```
AZU2_METJ   ASCETTVTSG DIMTYSTRSI SVPASCAEPT VNFEMKGHMP KTGMGHNWVL  50
AZUR_ALCFA  A-CDVSIEGN DSMQFNTKSI VVDKTCKEPT INLKHIGKLP KAAMGHNVVV  49
AZUR_ALCDE  AQCEATIESN DAMQYNLKEM VVDKSCKQPT VHLKHVGKMA KVAMGHNWVL  50
AZU1_METJ   AGCSVDVEAN DAMQYNTKNI DVEKSCKEPT VNLKHTGSLP KNVMGHNLVI  50
AZUR_ALCSP  AECSVDIAGN DQMQFDKKEI TVSKSCKQPT VNLKHPGKLA KNVMGHNWVL  50
AZUR_BORBR  AECSVDIAGT DQMQFDKKAI EVSKSCKQPT VNLKHIGKLP RNVMGHNWVL  50
AZUR_PSES4  AECSVDIQGN DQMQFSTNAI TVDKACKTPT VNLSHPGSLP KNVMGHNWVL  50
AZUR_PSEAE  AECSVDIQGN DQMQFNTNAI TVDKSCKQPT VNLSHPGNLP KNVMGHNWVL  50
AZUR_PSEFB  AECKTTIDST DQMSFNTKAI EIDKACKTPT VELTHSGSLP KNVMGNNLVI  50
AZUR_PSEPU  AECKVTVDST DQMSFNTKDI AIDKSCKTPT VELTHSGSLP KNVMGHNLVI  50
AZUR_PSEFC  AECKVTVDST DQMSFDTKAI EIDKSCKTPT VDLKHSGNLP KNVMGHNWVL  50
AZUR_PSEFA  AECKVTVDST DQMSFNTKAI EIDKSCKTPT VELTHSGSLP KNVMGHNWVL  50
AZUR_PSEFD  AECKVDVDST DQMSFNTKEI TIDKSCKTPT VNLTHSGSPL KNVMGHNWVL  50
Consensus   AEC.V.I.S. DQMQFNTK.I .VDKSCK.PT VNL.H.G.LP KNVMGHNWVL  50

AZU2_METJ   AKSADVGDVA KEGAHAGADN NFVTPGDKRV IAFTPIIGGG EKTSVKFKVS  100
AZUR_ALCFA  SKKSDESAVA TDGMKAGLNN DYVKAGDERV IAHTSVIGGG ETDSVTFDVS  99
AZUR_ALCDE  TKEADKQGVA TDGMNAGLAQ DYVKAGDIRV IAHTKVIGGG ESDSVTFDVS  100
AZU1_METJ   TKTADFKAVH NDGVAAGEAG NEVKAGDIRV VAHTKLVGGG EKDSVKVDVS  100
AZUR_ALCSP  TKQADMQGAV NDGMAAGLDN NYVKKDDARV IAHTKVIGGG ETDSVTFDVS  100
AZUR_BORBR  TKTADMQAVE KDGIAAGLDN QYLKAGDTRV IAHTKVLGGG ESDSVTFDVA  100
AZUR_PSES4  TTAADMQGVV TDGMAAGLDK NYVKDGDIRV IAHTKIIGSG EKDSVTFDVS  100
AZUR_PSEAE  STAADMQGVV TDGMASGLDK DYLKPDDSRV IAHTKLIGSG EKDSVTFDVS  100
AZUR_PSEFB  SKQADMQPIA TDGLSAGIDK NYLKEGDIVH IAHTKVIGAG EKDSLTIDVS  100
AZUR_PSEPU  SKEADMQPIA TDGLSAGIDK QYLKDGDARV IAHTKVIGAG EKDSVTFDVS  100
AZUR_PSEFC  TTQADMQPVA TDGMAAGIDK NYLKEGDIRI IAHTKIIGAG ETDSVTFDVS  100
AZUR_PSEFA  SSAADMPGIA SDGMAAGIDK NYLKEGDIRV IAHTKIIGAG EKDSVTFDVS  100
AZUR_PSEFD  SKSADMAGIA TDGMAAGIDK DYLKPGDSRV IAHTKIIGSG EKDSVTFDVS  100
Consensus   .K.ADMQ.VA TDGMAAG.DK NYLK.GD.RV IAHTK.IG.G EKDSVTFDVS  100

AZU2_METJ   ALSKDEAYTY FCSYPGHFSM MRGTLKLEE                          129
AZUR_ALCFA  KLKEGEDYAF FCSFPGHWSI MKGTIELGS                          128
AZUR_ALCDE  KLTPGEAYAY FCSFPGHWAM MKGTLKLSN                          129
AZU1_METJ   KLAAGEKYTF FCSFPGHATM MRGTVTVK-                          128
AZUR_ALCSP  KLAAGEDYAY FCSFPGHFAL MKGVLKLVD                          129
AZUR_BORBR  KLAAGDDYTF FCSFPGHGAL MKGTLKLVD                          129
AZUR_PSES4  KLKAGDAYAF FCSFPGHSAM MKGTLTLK-                          128
AZUR_PSEAE  KLKEGEQYMF FCTFPGHSAL MKGTLTLK-                          128
AZUR_PSEFB  KLNAAEKYGF FCSFPGHISM MKGTVTLK-                          128
AZUR_PSEPU  KLAAGEKYGF FCSFPGHISM MKGTVTLK-                          128
AZUR_PSEFC  KLKADGKYMF FCSFPGHIAM MKGTVTLK-                          128
AZUR_PSEFA  KLAAGTDYAF FCSFPGHISM MKGTVTVK-                          128
AZUR_PSEFD  KLTAGESYEF FCSFPGHNSM MKGAVVLK-                          128
Consensus   KL.AGE.Y.F FCSFPGH..M MKGT.TLK-                          129
```

Fig. 1. Alignment of *P. fluorescens* biotype A (ATCC 13525) holoazurin with other azurins. Precursors of *P. aeruginosa* and *A. denitrificans* were omitted to facilitate the alignment *P. fluorescens* is highlighted. PSEFA, *P. fluorescens* biotype A (ATCC 13525); PSEFB, *P. fluorescens* biotype B; PSEFC, *P. fliorescens* biotype C; PSEFD, *P. fluorescens* biotype D (also *P. chlororaphis* ATCC 17414); PSEAE, *P. aeruginosa*; PSES4, *Pseudomonas Sp.* (also *P. denitrificans* ATCC 13867); PSEPU, *P. putida* (NCIB 9869); ALCFA, *A. faecalis*; ALCDE, *A. denitrificans*; ALCSP, *Alcaligenes Sp.*; BORBR, *B. bronchiseptica*; METJ1, *Methylomonas J* iso-1; METJ2, *Methylomonas J* iso-2.

The final model consists of 952 protein atoms and 114 solvent atoms. r.m.s. deviations from standard values for bond lengths and angles are 0.010 Å and 2.59°. R.m.s. values for dihedrals and impropers are 25.8 and 0.91°.

3. Results

3.1. *Primary structure*

3.1.1. *The primary sequence and the amino-acid composition.* The primary sequence of the *P. fluorescens* azurin is given in Fig. 2. The amino-acid composition of the same molecule reported earlier (Hutnik & Szabo, 1989) was revised after sequence determination (Table 2). The isolated peptide sequences and the peptide designations are shown in Fig. 2. The peptide sequences were determined by automated sequence analysis and with the methods described above. The N-terminal 30 residues were determined originally by N-terminal sequence analysis. C-terminal analysis with carboxypeptidase showed that peptide T-5 corresponded to the C-terminus of the protein.

3.1.2. *Molecular weight.* The mass spectrometric analysis of the intact azurin revealed two components with a mass difference of approximately 60 Da. The average mass derived from the major component was 13 686 Da. while the mass from the minor component was 13 626 Da. The mass difference of 60 Da suggested that the major component was the copper-containing holoazurin and that the minor one was the apoenzyme.

Code	Elution Time (minutes)	Sequence
T-2	2.7	EGDTR
T-4	10.2	VIAHTK
T-5	11.5	GTVTVK
T-6	14.0	IIGAGEK
T-7	14.7	NYLK
T-8	16.0	AIEIDK
T-12	20.2	DSVTFDVSK
T-13	21.7	VTVDSTDQMSFNTK
T16(T-6 + T-12)	23.8	IIGAGEKDSVTFDVSK
T-18	25.2	TFTVELTHSGSLPK
T-30	33.8	NVMGHNWVLSSAADMPG1ASDGMAAGIDK
T-31	35.0	LAAGTDYAFFCSFPGHISMMK

(*a*)

```
    5   10   15   20   25   30   35   40   45   50
AECK VTVDSTDQMSFNTK AIEIDK SCK TFTVELTHSGSLPK NVMGHNWVL
        T-13        T-8          T-18          T-30

   55   60   65   70   75   80   85   90   95  100
SSAADMPG1ASDGMAAGIDK NYLK EGDTR VIAHTK IIGAGEK DSVTFDVS
                     T-7  T-2  T-4  T-6  T-12

  105  110  115  120  125
K LAAGTDYAFFCSFPGHISMMK GTVTVK
       T-31          T-5
```

(*b*)

Fig. 2. (*a*) Sequences determined for tryptic peptides derived from cleavage of reduced and carboxymethylated *P. fluorescens* (ATCC 13525) and (*b*) the whole amino-acid sequence.

Table 2. *Amino-acid composition of P. fluorescens (ATCC 13525) azurin*

Residue		Amino-acid composition analysis Found	Predicted*	Published (Hutnik)	Sequence
Asx		13.80	14	15	14
	Asp				10
	Asn				4
Glx		5.72	6	6	6
	Glu				5
	Gln				1
Ser		10.21	11–12	11	12
Gly		10.99	11	11	11
His		3.18	4	4	4
Arg		1.17	1	1	1
Thr		10.52	11–12	12	12
Ala		12.00	12	13	12
Pro		2.82	3	3	3
Tyr		1.61	2	2	2
Val		8.55	9–10	10	10
Met		5.48	6	6	6
Cys		2.88	3	3	3
Ile		6.73	7–8	7	8
Leu		4.82	5	5	5
Phe		5.93	6	6	6
Trp		N/D	1	1	1
Lys		10.88	11–12	12	12
Total			123–128	128	128

* 12 residues of alanine are assumed. The values shown are the average of two determinations following vapour-phase hydrolysis for 60 min at 423 K. Losses due to hydrolysis were not corrected for.

The mass of 13 626 Da corresponded (within 1 Da) to the molecular weight of the apoenzyme of 13 626.49 calculated from the amino-acid sequence of 128 residues (13 624.5, assuming Cys3–Cys26 disulfide). A second batch of purified azurin was analyzed again by mass spectrometry. In this case, a single component consisting of ten spectral lines revealed an average mass of 13 685 Da. These findings suggested that the first

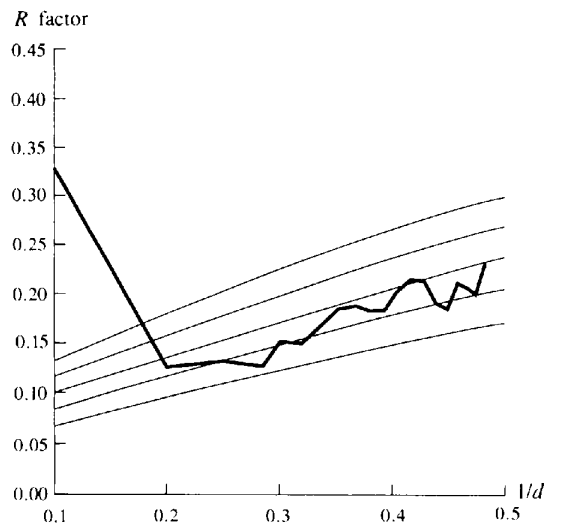

Fig. 3. Luzzati plot of the R values as a function of resolution. Error lines (a = 0.20, c = 0.30 Å) indicate a maximum mean coordinate error of ~0.28 Å.

batch of purified azurin contained a small amount of contaminating apoazurin and that the second purification contained only the holoenzyme. The predicted M_r including one Cu atom was 13 689.04 Da (13 386 if Cys3–Cys26 is a disulfide and two protons are lost when binding Cu). The mass spectrometric analysis of the peptides were within 1 Da of the M_r calculated from their sequence.

3.1.3. *Amino-acid sequence alignments.* The *P. fluorescens* (ATCC 13525) sequence was compared with the 12 azurin sequences found in the Protein Data Bank (Swiss-Prot, Vol. 28) and was found to have the greatest homology (84%) with *P. fluorescens* biotype C (ACTT 17400). It retains, in common with the other azurins, the metal-binding residues His46, Cys112, His117 and Met121 (Adman & Jensen, 1981; Norris, Anderson & Baker, 1986). It also contains Cys3 and Cys26, which are known to form an intramolecular disulfide bridge. This azurin was found to have the single Trp48 residue and the His35 residue as previously reported (Hutnik & Szabo, 1989). The C-terminal end of the protein, -G-T-V-T-V-K is identical to that of METJ1. However, this azurin has four residues, Ser52, Pro57, Ser61 and Thr106, which are not found in any other azurin. Two of these residues, Pro57 and Ser61, are conservative substitutions.

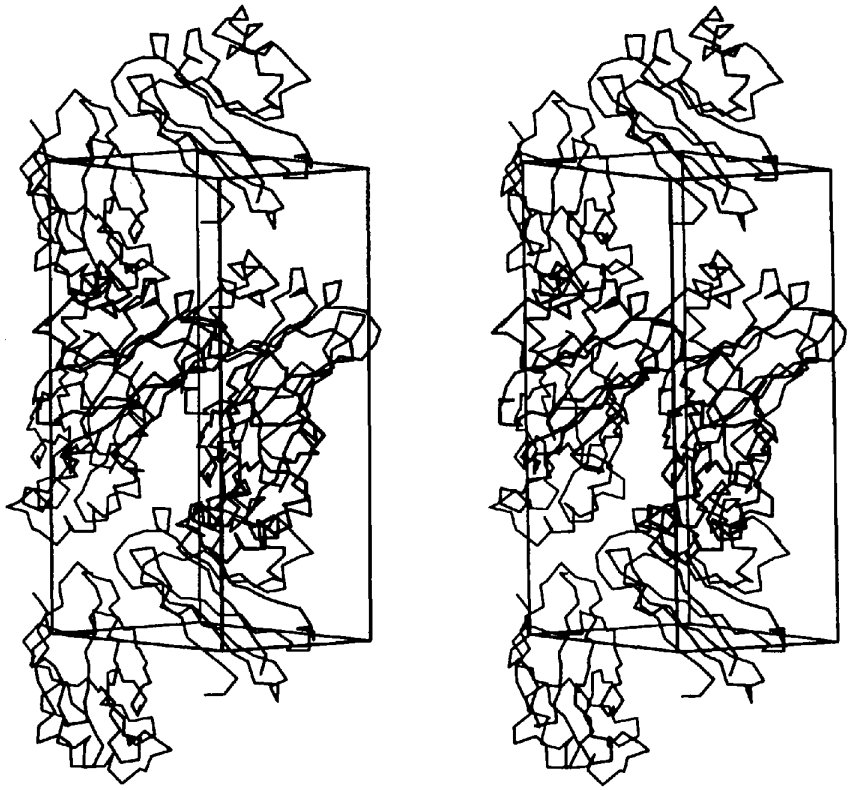

Fig. 4. Packing of the symmetrically related molecules in the unit cell.

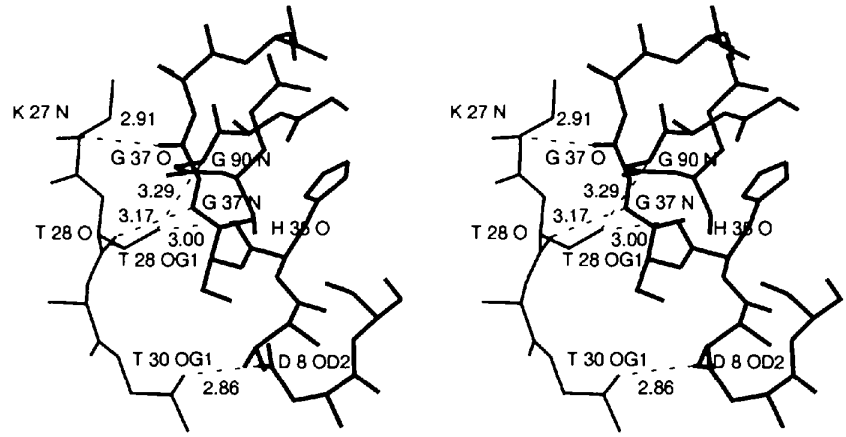

Fig. 5. The major contacts between two symmetrically related molecules. The molecule is in bold while the symmetrically related molecule is drawn with lighter lines.

3.2. *Tertiary structure*

3.2.1. *The accuracy of the final model.* The final model consists of 952 protein atoms and 114 water molecules. Fig. 3 shows a Luzzati plot (Luzzati, 1952) for the refined azurin structure. Except at the lowest resolution and at the highest resolution the *R*-value distribution falls between the first and third lines. This indicates that the expected coordinate error is between 0.25 and 0.30 Å. We estimate that for well defined parts of the structure, the copper site and all internal side chains, the positional error is much less due to low *B* factors (4–12 Å^2). Other regions have higher *B* factors and lower accuracy. They include the chain termini, residues 1–2 and 127–128, some connecting loops between the β-strands, and especially the residues 102–106 (see *Description of the structure*). Long side chains are less accurately defined, with *B* factors between 30 and 45 Å^2. These include those of Lys4, Gln12, Lys18, Glu21, Lys24, Met64, Glu75, Lys85 and Met120.

3.2.2. *Crystal packing. P. fluorescens* molecules are tightly packed in the crystal with a solvent content of 36%. The crystal packing is shown in Fig. 4. The regions involved in direct contact are residues 8, 35–37 and 90–94 in one molecule, and residues 27–30, in the symmetrically related molecule. All levels of interaction, *i.e.* main-chain to main-chain, main-chain to side-chain and side-chain to side-chain participate in the contact. The major bridges are between residues 35–37 of one molecule and 27–30 of the symmetrically related molecule (Fig. 5). Gly37 N hydrogen bonds with the Thr28 O#* (3.17 Å) while Gly37 O with Lys27 N# (2.91 Å). They are reinforced by the side-chain to main-chain interactions, between Thr28 OG1# and His35 O (3.00 Å) and Thr28 OG1# and G90 N (3.29 Å). Side-chain to side-chain interaction again strengthens these contacts through the hydrogen bond between Thr30 OG1# and Asn8 OD2 (2.86 Å), which is located very close to Ser36. In addition, ND2 of Asn16, another residue in the neighborhood forms one hydrogen bond with the Ser94 OG#. The presence of water molecules close to this interface further strengthens these contacts. Unlike *A. denitrificans*, members of the hydrophobic patch (13, 44, 114–120) are not involved in direct contact with the neighboring molecule nor is the C terminus. The N terminus does contact the neighboring molecule through a hydrogen bond between the Ala1 N and the Gln12 OE2# (2.53 Å). There are other minor contacts through the side chains such as the Asn42 ND2 with Ser119 O# (2.83 Å) and through the medium solvent. The important and interesting feature is the direct involvement of the His35 in the contact with the symmetrically related molecule.

3.2.3. *B factors.* Average *B* factors for main-chain and side-chain atoms of the azurin molecules are plotted in Fig. 6. They show the familiar pattern of globular proteins. The copper ligands and the copper-binding loops have very low *B* factors (Gly45, His46, His117, Met112). For the loop formed by the residues 101–106 the values are much higher as this loop is situated on the surface of the molecule and the density is also much poorer than the rest of the molecule. The average *B* factors are 11.0 and 14.3 Å^2 for the main chain and for the side chains, respectively. Compared with other azurin structures the *B* factors of this azurin are lower. A Wilson plot indicated that the average *B* factor is 14.1 Å^2.

* # stands for the symmetrically related molecule.

4. Description of the tertiary structure

4.1. *General organization*

The structure of azurin is as described by Adman *et al.* (1978) and Norris *et al.* (1983), resembling a β-sandwich structure (Chothia & Lesk, 1982). The molecule is shown schematically in Fig. 7. The Cu site is in a

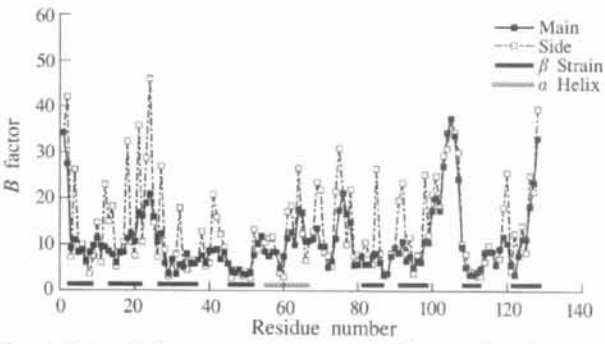

Fig. 6. Plot of *B* factors as a function of residue number. For each residue the *B* factor is the mean of main- or the side-chain atoms. Residues involved in β-strands and α-helices are indicated.

Fig. 7. Schematic Richardson-type diagram of the folding of the azurin molecule, β-strands shown as arrows. Copper site at the top (northern) end of the β-barrel. The program *MOLSCRIPT* (Kraulis, 1991) was used to plot this diagram.

cavity formed between several of the β-strands and their connecting loops. A 28-residue segment (residues 53–80 linking strands 4 and 5 appears like an extra 'flap' on the outside of the β-barrel. This flap contains the only α-helix in the molecule. An extensive hydrophobic patch is at one end of the molecule, above the copper site, and a large number of charged side chains distributes over the rest of the surface. The refinement has revealed 114 bound water molecules but no SO_4^{2-}.

4.2. *Conformational angles*

The torsion angles (φ, ψ) for the polypeptide backbone are shown in the familiar Ramachandran plot (Fig. 8). The clustering round about (–120, 135°) corresponds to the predominance of β-character in the structure; the other cluster around about (–60 , –30°) results from the residues in turns and the single α-helix. All non-glycine residues are within allowed regions. 102 residues are in the most favorable regions. Non-glycine residues in the left-handed α region are Met13, Asn71 and Tyr72. The first one is in a type I turn, while the latter two are type III turns. One glycine residue, Gly76, is also in the left-handed α region. Five glycine residues are outside the allowed regions, Gly37, Gly67, Gly90, Gly105, Gly116. Except for Gly105 the other four are all invariant in azurins sequenced to date, confirming their indispensable role in configuration.

4.3. *Hydrogen bonding*

The organization of the β-strands and the turns in *P. fluorescens* are identical to that of *A. denitrificans*. Therefore, the hydrogen bonding involved in the main-chain contacts are very similar for both *P. fluorescens* and *A. denitrificans*. The replacement of Thr61 ($\varphi = -68$, $\psi = -48°$) by Ser61 ($\varphi = -61$, $\psi = -51°$) in the α-helix has very little effect on the conformational angles. As in *A. denitrificans* the C═O of residue 62 hydrogen bonds only to a water molecule while the carbonyl groups of residues 63 and 64 make 1–4 hydrogen bonds, to 66 NH and 67 NH although the latter is quite loose (4.09 Å) compared with that in *A. denitrificans* (3.39 Å). This indicates that one end of the helix is slightly looser in *P. fluorescens*. β-strand 2 also has a 'kink' in the middle, at residue 17, which transfers its hydrogen bonding from strand 1 to strand 8. The kink seems less obvious with $\varphi = -150$ and $\psi = 152°$ instead of $\varphi = -117$ and $\psi = 135°$ in the case of *A. denitrificans*. The torsion angles of residues 16 and 18 are very similar in these two cases.

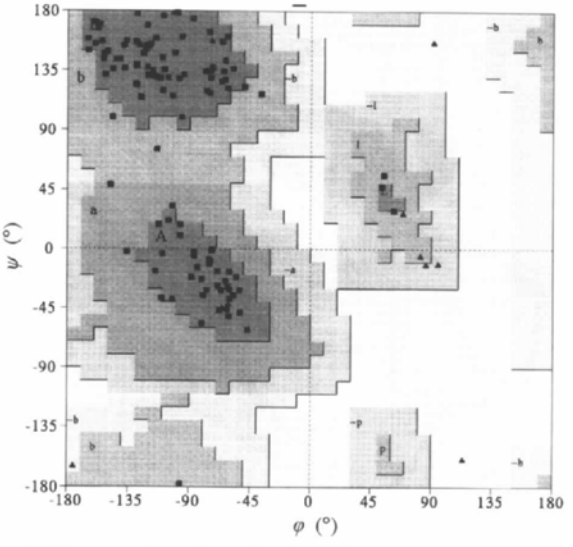

Fig. 8. Ramachandran plot of main-chain dihedral angles. Glycine residues are indicated as triangles. 103 residues are in the most favored regions [A, B, L] and nine residues are in the additional allowed regions [a,b,l,p]. There are 11 glycine and three proline residues.

Table 3. *Side-chain to main-chain hydrogen bonds*

Local	Long range (crosslinking)
Asn16 OD···HN 16	Ser9 OG···O 34
Asp23 OD1···HN 25	Thr10 OG1···OG 14
His35 ND1···O 36	Asp11 OD1···HN 39
Ser42 OG···HN 42	Asp11 OD2···HN 38
Thr51 OG1···O 52	Asp16 OD2···HN 8
Ser61 OG···HN 57	His35 NE2···O 44
Asp77 OD1···HN 79	His46 NE2···O 10
Thr84 OG1···O 85	Asn47 OD1···HN 113
Glu91 OE1···HN 88	Asn47 ND2···O 71
Asp98 OD2···HN 100	Asp71 OD1···HN 86
Cys112 SG···HN 114	Thr84 OG1···HN 48
	Asp93 OD···HN 85
	Tyr108 O···HN 103
	Ser113 OG···O 72

Sequence differences in *P. fluorescens* results in more differences in side-chain to main-chain hydrogen bonding (Table 3). In *A. denitrificans* there are 15 pairs of local side-chain to main-chain hydrogen bonds and 16 pairs of long range (crosslinking) bonds. Out of the 15 local hydrogen bonds only ten pairs are strictly retained. The substitutions of Thr10 for Asn10 and Ser14 for Gln14 changes the side-chain to main-chain interaction to side-chain to side-chain interaction. Thr10 OG1 does not form hydrogen bonds with either NH12 or NH14. It hydrogen bonds with the OG of Ser14 (2.95 Å). This in fact has the same effect as the side-chain to main-chain interactions to secure the turning of the main chain. The substitutions of Asp16 by Asn16 and Ser61 by Thr61 do not change the interaction. Two bound water molecules around the imidazole of His83 replace its hydrogen bond between ND1 and O84.

His35 ND1 hydrogen bonds with Ser36 O (2.72 Å). The torsion angles of the following two residues, Ser36 ($\psi = 164°$) and Gly37 ($\varphi = 119°$) make it possible. This is similar to the flip of the peptide bond observed in *P. aeruginosa* at pH 5.5 (Nar *et al.*, 1991*b*) with similar torsion angles of Pro36 ($\psi = 152°$) and Gly37 ($\varphi = 119°$) (Fig. 9). In *P. aeruginosa* the flip was shown as a pH effect whereas in *P. fluorescens* this is a stable conformation. It was not observed in *A. denitrificans* with

completely different torsion angles of Val36 ($\psi = 7°$) and Gly37 ($\varphi = -88°$) (Baker, 1988). This conformation of the peptide bond in *P. fluorescens* facilitates the major contacts between Gly37 N and Thr28 O# of the symmetrically related molecule. The His35 NE2 retains a hydrogen bond with Met44 O (2.81 Å) as in both *A. denitrificans* and *P. aeroginosa*. Since Ser9, Ser36 and Met44 belong to β-strands 1, 3 and 4 this certainly strengthens their interaction.

From temperature-jump and stopped-flow experiments (Antonini *et al.*, 1970; Pecht & Rosen, 1973; Brunori, Greenwood & Wilson, 1974; Wilson, Greenwood, Brunori & Antonini, 1975; Rosen & Pecht, 1976) it was inferred that the fast bimolecular process of electron transfer is coupled to a slow monomolecular reaction. This slow process was ascribed to a conformational equilibrium between two forms of reduced azurin, only one of which participates in the redox equilibrium (Wilson *et al.*, 1975; Rosen & Pecht, 1976). H^1 NMR studies demonstrated that the rate of protonation of His35 was shown to be comparable to the slow relaxation rate obtained by temperature-jump experiments (Hill & Smith, 1976). From these and the other experiments it was concluded that the azurin conformational equilibrium plays a regulatory role in its electron-transfer reactivity (Silvestrini, Brunori, Wilson & Darley-Usmar, 1981; Farver & Pecht, 1981) and that a surface patch of azurin located in the vicinity of His35 side chain is involved in the azurin–cytochrome electron transfer (Farver & Pecht, 1981; Mitra & Bersohn, 1982). However, recent experiments on the azurin–cytochrome redox reaction with site-specific mutants of azurin (van de Kamp, Floris, Hali & Canters, 1990; van de Kamp, Silvestrini *et al.*, 1990) have shown that fast electron transfer to cytochrome c_{551} does not require His35 and the hydrophobic patch on the azurin surface around copper ligand His117, rather than the His35 patch, play a key role in its interaction with cytochrome.

In the crystal structure of *A. denitrificans* azurin, His35 is buried, explaining its failure to titrate in the pH range 4.5–9.0, and is not protonated as expected at pH 5.0 (Baker, 1988). This is due to the fact that a hydrogen bond formed between Val36 NH and Gly9 O (2.75 Å) reduces the opening of the small cleft in which the His35 imidazole is located, thus preventing the side chain of His35 from gaining access to solvent. Under such circumstances the imidazole of His35 forms two hydrogen bonds, NE2H···O=C(44) and ND1···HN(37). On the other hand, Nar *et al.* have detected a pH-induced conformational transition involving residues 36–37 in *P. aeruginosa* azurin (Nar *et al.*, 1991*b*). The peptide unit between residues Pro36 and Gly37 flips over to adopt the conformation Pro36 ($\psi = 152°$), Gly37 ($\varphi = 119°$) at pH 5.5 (Nar *et al.*, 1991*b*). At this pH, instead of forming ND1···HN(37) the protonated His35 ND1 H atom is involved in a strong hydrogen bond to O36 as in *P. fluorescens*. Pro36 in *P. aeruginosa* precludes a hydrogen bond of NH to Gly9 O and that permits the access of the His35 side chain to solvent (Fig. 10). In *P. fluorescens*, the His35 ND1 is protonated at pH 7.0 despite the bond between Ser36 NH and Ser9 O (3.33 Å) as in *A. denitrificans*. This weak hydrogen bond leaves a wider opening in the cleft and allows the His35 side chain have access to solvent. This results in a conformation of residues 35–37 similar to that in *P. aeruginosa* at pH 5.5. The distance between the Ser9 O and the Ser36 N is 3.33 Å in *P. fluorescens* and approximately the same in *P. aeruginosa* without this

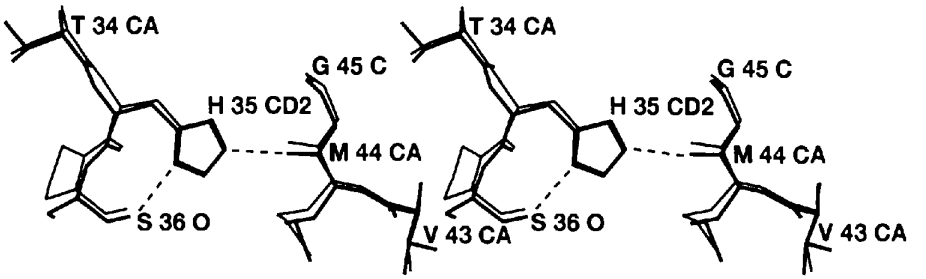

Fig. 9. The conformational similarity of the peptide stretch Thr34-His35-Ser36 in *P. fluorescens* (bold line) and the Ser34-His35-Pro36 in *P. aeruginosa* at pH 5.5 (lighter line).

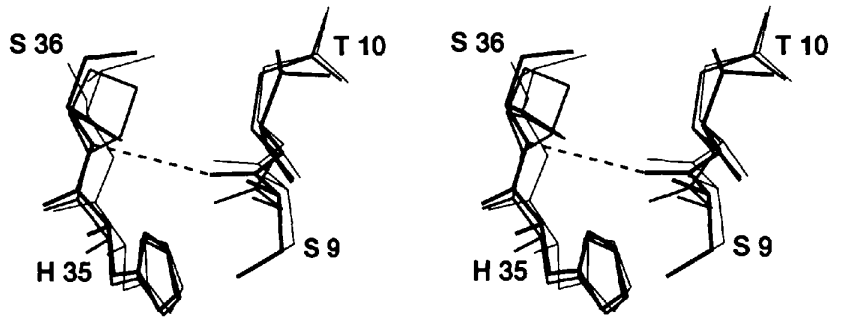

Fig. 10. Comparison of the opening of the cleft where the imidazole of the His35 is located. *P. fluorescens* is shown in bold *P. aeruginosa* in a medium line and *A. denitrificans* in light line.

particular hydrogen bond, compared with the 2.75 Å formed between the Ser9 O and the Val36 N in *A. denitrificans*. This opening in *P. fluorescens* favors the protonation of His35 at pH 7.0. Conformational changes of the residue segment 35–37 due to pH are unlikely to happen in *P. fluorescens* crystals as this segment participates in the crystal contacts with the symmetrically related molecule. Therefore, the conformation of residues 36–37 is a stable one in the crystal structure.

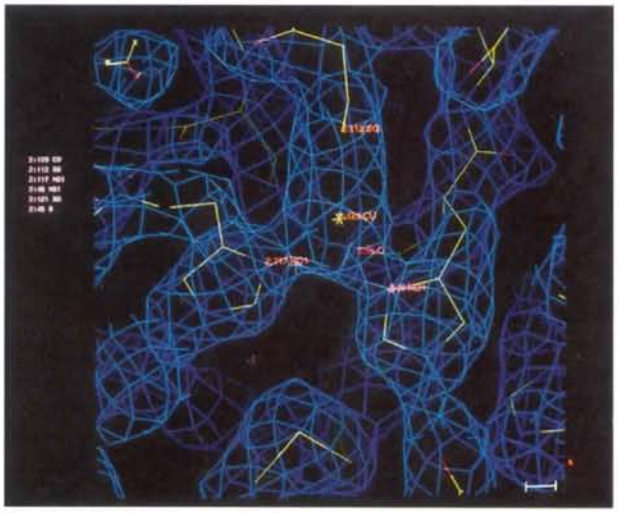

Fig. 11. $2Fo - Fc$ electron-density map at the at the Cu site. Contoured at 1.4σ.

Out of the 16 long-range side-chain to main-chain hydrogen bonds in *A. denitrificans* only 13 remain as the Tyr15 and Tyr110 are changed to Phe in *P. fluorescens*. No hydrogen bonds can be formed after these substitutions although the spaces of the side chains are retained. Also the distance between Cys112 SG and Asn47 N becomes too long to be considered to be a hydrogen bond (3.52 Å). The remaining 13 pairs are very similar to those of the *A. dentrificans*. Most of them are at the northern end of the molecule, involving residues towards the ends of the β-strands and in the connecting loops at the top of the molecule. These undoubtedly contribute to the stability of the structure surrounding the copper site. A new pair is formed on account of the substitution of Thr10 for Asn10.

4.4. *Copper site*

Fig. 11 shows the electron density of the Cu site. Details of the copper geometry are given in Table 4. The Cu atom makes three strong bonds, with the thiolate S atom of Cys112 (mean Cu—S = 2.13 Å), and the imidazole N atoms of His46 and His117 (mean Cu—N = 2.10 and

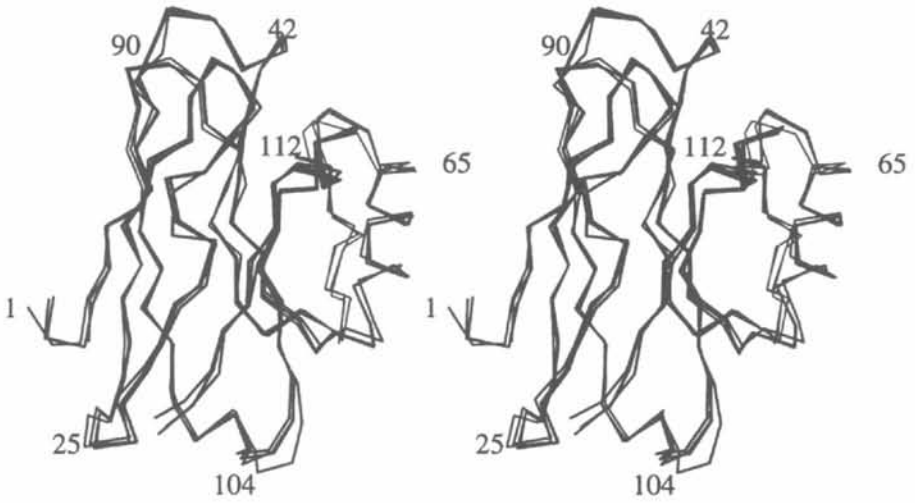

Fig. 12. The environment of the His83 side chain.

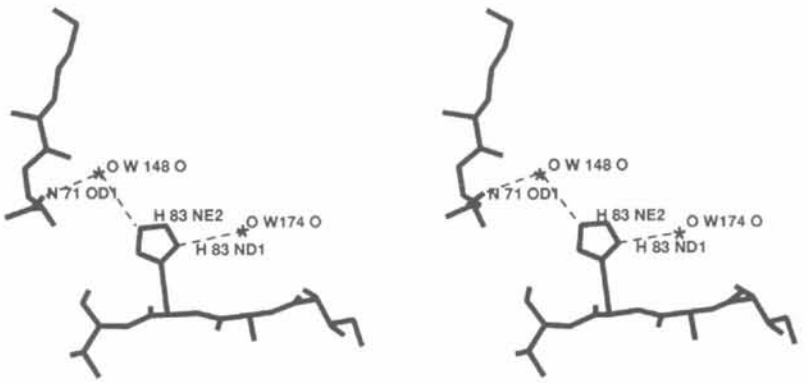

Fig. 13. Overlaps of the C^{α} atoms of five azurins.

Table 4. *Copper-site geometry*

	P. fluorescens	*A. denitrificans* Mean	*P. aeruginosa* pH 5.5	*P. aeruginosa* pH 9.0
(*a*) Cu—ligand bond lengths (Å)				
Cu—ND1(45)	2.10	2.08	1.99	1.98
Cu—ND1(117)	2.06	2.01	2.11	2.06
Cu—SG(112)	2.10	2.12	2.27	2.28
Cu—O(45)	2.99	3.16	2.84	2.83
Cu—SD(121)	3.23	3.12	3.18	3.15
(*b*) Ligand—Cu—ligand bond angles (°)				
O(45)—Cu—ND1(46)	72	74	73	77
O(45)—Cu—SG(112)	105	103.5	98	98
O(45)—Cu—ND1(117)	85	79.5	89	88
O(45)—Cu—SD(121)	143	147	149	149
ND1(46)—Cu—SG(112)	136	135	133	134
ND1(46)—Cu—ND1(117)	98	104.5	103	103
ND1(46)—Cu—SD(121)	73	77.5	78	74
SG(112)—Cu—ND1(117)	126	119.5	123	122
SG(112)—Cu—SD(121)	109	107.0	110	110
ND1(117)—Cu—SD(121)	87	96.0	87	88
(*c*) Angles at ligand atoms (°)				
Gly45 C—O—Cu	132	128	132	131
His46 CG—ND—Cu	126	123	130	130
His46 CE1—ND—Cu	127	129	125	125
Cys112 CB—SG—Cu	114	107	109	108
His117 CG—ND1—Cu	124	122	130	130
His117 CE1—ND1—Cu	128	127	124	124
Met121 CG—SD—Cu	139	141.5	140	140
Met121 CE—SD—Cu	83	101.5	101	103
(*d*) Other distances (Å)				
O(45)···SD(121)	5.89	5.96	5.64	5.74
(*e*) Deviation from plane through ND1(46), SG(112), ND1(117) (Å)				
Cu	0.04	0.13	0.03	0.03
O(45)	−2.81	−2.86	−2.74	−2.74
SD(121)	3.09	3.10	3.00	3.03

2.06 Å, respectively). These three ligands are arranged in a distorted trigonal planar configuration around the Cu atoms, with bond angles in the trigonal plane of 129, 95 and 135°. Two much weaker bonds form between the Cu and C=O of Gly45 (2.96 Å) and the thioether S atom of Met121 (3.23 Å). The Cu atom lies very close to the N^2S plane through the three strongly bound ligands, displaced only very slightly (0.04 Å) toward Met121. The evidence here argues for a trigonal bipyramidal (N_2S_2O) geometry.

The only one distinct difference would be the much smaller angle of Met121 CE—SD—Cu. It is only 83° in *P. fluorescens* compared with the 101–103° in *A. denitrificans* and *P. aeruginosa* (Baker, 1988; Nar *et al.*, 1991*b*). The density map clearly indicates the smaller angle in our case.

4.5. *Solvent structure*

In this structure of *P. fluorescens* no density peak corresponding to sulfate ion has been found. Therefore, all the peaks have been treated as water molecules. A total of 114 water molecules are included in the final model of the *P. fluorescens* azurin molecule. The water molecules have been given serial numbers after the protein atoms and in order of increasing *B* factors. By hydrogen-bond contacts we mean those O and N atoms which are within hydrogen-bond distance, *i.e.* 2.4 to 3.3 Å (Roe & Teeter, 1993).

The overall pattern of the distribution of the water molecules follows the same picture given by Baker (1988) for *A. denitrificans*. *P. fluorescens* contains no truly internal water molecules and 53% of the water molecules are in the shallow depressions created by the side chains of the molecule and loops of the polypeptide chains. The rest are in the deeper pockets between the flap and its neighboring loops or in the interfaces between the symmetrically related molecules. However,

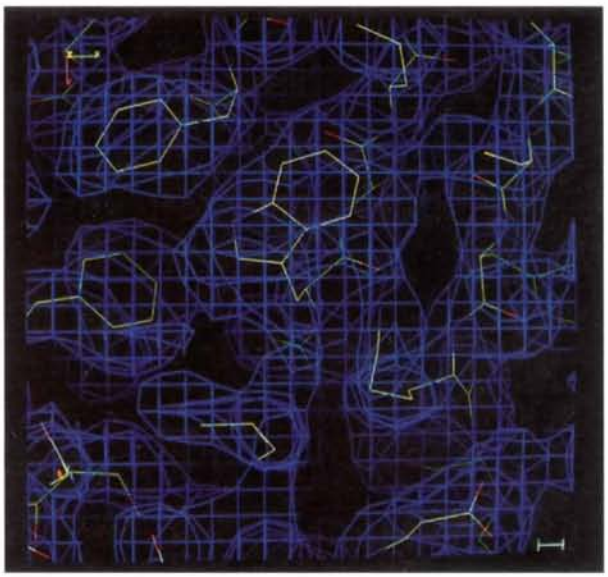

Fig. 14. Electron density shows the Phe side chains adjacent to Trp48.

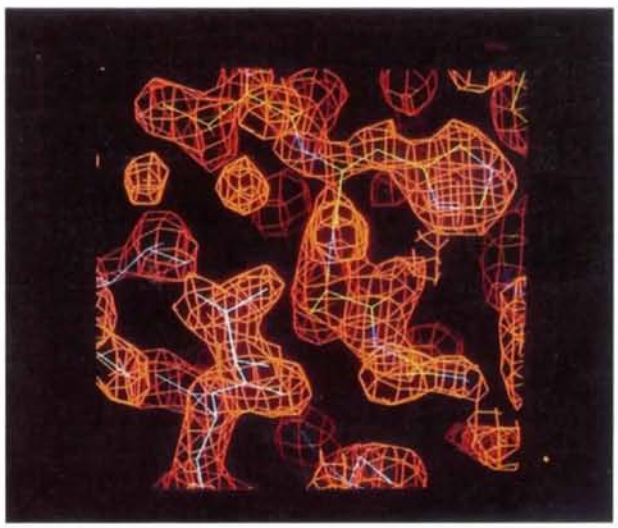

Fig. 15. $2F_o - F_c$ electron-density map showing the direct contact between the His35 O and Thr28 OG1# of the neighboring molecule. Contoured at 1.4σ.

Table 5. *Water molecule environment*

Number of hydrogen bonds to protein atoms	Total number of hydrogen bonds						
	0	1	2	3	4	5	Total
0	4	8	6	0	0	0	18
1	0	12	17	14	1	0	44
2	0	0	8	15	11	0	33
3	0	0	0	6	9	0	15
4	0	0	0	0	1	2	3
Total	4	20	31	35	22	2	114
Average B factor (Å^2)	46	39	36	30	34	20	

there are still some differences between the solvent structure of these two azurin molecules probably due to the difference in the crystal packing. First of all there are fewer water molecules per one molecule of *P. fluorescens* azurin. We have only 114 water molecules in the final model instead of 140 for *A. denitrificans*. This is consistent with the tight packing of the crystal matrix with only 36% of solvent content. The overall B factor of water molecules in *P. fluorescens* is 27 Å^2. The number of hydrogen bonds made by water molecules is summarized in Table 5. Of the total 114 water molecules there are only four without hydrogen bonding within 3.3 Å.

4.6. *Comparison of different azurin molecules*

Fig. 12 shows the overlaps of the C^α atoms of five azurin molecules. They are *A. denitrificans*, *P. aeruginosa* at pH 5.5, *P. aeruginosa* at pH 9.0, *P. denitrificans* and *P. fluorescens* (Bernstein *et al.*, 1977). Except for the surface loops they are well overlapped, especially the fragments involved in the ligands with Cu. The lowest r.m.s. deviation is 0.61 Å between *P. aeruginosa* at pH 5.5 and *P. fluorescens* while the highest is 0.78 Å between *P. fluorescens* and *A. denitrificans*. However, the molecules pack very differently in the crystals for these four azurins due to different space groups and the different numbers of the molecules in an asymmetric unit. The crystal packing of the *P. fluorescens* azurin is the simplest one with only one molecule in an asymmetric unit. The geometry of the copper site is not much different from that of the other azurins. However, there are quite a few local differences due to the substitutions of residues. The most distinctive ones would be the differences in His35 and His83. The structural difference in His35 between *P. fluorescens* and the other azurin is discussed in the section *Hydrogen bonding*. The His83 imidazole ring flips almost 180° as the torsion angle between CB and CG is –93° in *P. fluorescens* instead of 91° in *A. denitrficans*. Therefore, no hydrogen bond can be formed between His83 ND1 and 84 C=O and no protonation of His83 can occur at NE2. Instead water molecule W174 forms hydrogen bond with ND1 (2.86 Å). W148 forms one hydrogen bond with NE2 of the imidazole ring (3.30 Å) and another with the carbonyl O atom of K70 (Fig. 13). This can be explained as the pH for crystallization was 7.0 instead of 5.0 in the case of *A. denitrificans* (Ugurbil, Norton, Allerhand & Bersohn, 1977).

5. Discussion

Comparison of the X-ray crystallographic and time-resolved fluorescence data allows for the determination of the most probable Trp quenching mechanism for Trp48 in *P. fluorescens*. The time-resolved fluorescence decay is dominated by a very short lifetime component both in solution (106 ps, 94% of the total fluorescence) and in the crystalline state (54 ps, 80–90% of the total fluorescence). This short fluorescence decay time has been attributed to quenching of the Trp fluorescence through the Cu^{II}–ligand complex (Hutnik & Szabo, 1989). Inspection of the protein model from the X-ray crystallographic data reveals a distance of under 10 Å between the Trp indole ring and the center of the copper (Cu^{II}—CG Trp = 10, Cu^{II}—CDI Trp = 9.8 Å). Such a distance is consistent with the Dexter-type electron-exchange mechanism and provides further support for such mechanism in protein systems (Birks, 1970). The electron-exchange mechanism becomes less favorable at greater distances, suggesting that the longer decay time observed for Trp in *P. fluorescens* azurin must arise from a conformation in which the Trp indole ring is further from the Cu. Another potential contributor to Trp quenching may include a local cysteinyl (Cys) sulfhydryl group (Cowgill, 1970) where the C^α of the Cys is within 7.1 Å of the Trp CG.

Alternatively there are several Phe side chains which surround the Trp side chain, Phe15, Phe110 and Phe97 (Fig. 14). Based upon the absorption spectral overlap for Phe and Trp, normally one might expect that resonance energy transfer might occur from Phe to Trp (Teale & Weber, 1957). It is possible that the Phe ring could act as a wire for exchange of electrons between Cu and Trp, increasing the efficiency of an electron-exchange quenching process and accounting for the very short Trp decay time. The fluorescence decay behavior of two holoazurins from *P. fluorescens* and *P. aeruginosa* was best fit by three decay components whose lifetime values were similar in these two proteins (Hutnik & Szabo, 1989). The fluorescence decay parameters are given in Table 6. One decay time was found to be ~5 ns, another between 0.5 and 0.4 ns, and a third at 0.10 ns. These three decay components are suggested to originate from either three different conformations of the Trp side chain or different protein conformations of the Cu^{II}–ligand complex. The pre-exponential terms, C_i, are equivalent to the relative concentrations of these conformational states (Hutnik & Szabo, 1989). The relative concentration of the long decay component was 0.089 for *P. fluorescens* but only 0.021 for *P. aeruginosa*. That of the middle decay time is almost the same (0.067–0.046) while the fractional concentration for the third component are significantly

Table 6. *Fluorescence decay parameters*

The fluorescence decay parameters are the best-fit values which are obtained from the exponential decay model. Each value is the mean of at least two determinations. The precision of the data per exponential decay was typically as follows: τ, ±0.01; τ_2, ±0.01; τ_3, ±0.01. Standard errors were typically as follows: τ_1, ±0.03; τ_2, ±0.03; τ_3, ±0.002. τ_i is the symbol used for excited singlet state lifetime. C_i represents the fractional fluorescence for each respective component.

Sample	λ_m (nm)	τ_1	τ_2	τ_3	C_1	C_2	C_3
Pae holo	310	4.89	0.36	0.098	0.021	0.46	0.933
Pfl holo	310	4.91	0.52	0.105	0.089	0.067	0.844

different (0.844–0.933) (Table 6). Statistically six different rotamers of the Trp side chain may possibly be observed in protein structures. In the case of azurin three of them, the C_α–C_β rotamers can be accommodated in the hydrophobic pocket without bumping into the other side-chain atoms. The conformer observed in the refined structure is the most stable conformation with the lowest potential energy ($\chi_1 = -177.30$, $\chi_2 = -89.20°$). The other two $\alpha\beta$ rotamers ($\chi_1 = -177.30$, $\chi_2 = -95.10$; $\chi_1 = -70.40$, $\chi_2 = 100.50°$) are not observed in the final refined structure but could coexist with the most stable one in solution and in crystals. They could only be detected from their different fluorescence decay behavior. The three possible conformations of the Trp48 side chains may account for the heterogeneity shown in the fluorescence experiments (Hutnik & Szabo, 1989). Alternately the Cu^{II}–ligand complex might exist in two different conformations, one a square planar structure and the other a flattened tetrahedral structure. In this latter case the Cu would be further removed from the Trp residue leading to the long decay time. In the crystal structure observed, the Cu–ligand complex is seen to be in a trigonal planar configuration.

We have compared the two Trp residues in the crystal structures of these two proteins based on the refined models and no significant difference in terms of coordinates can be found between the Trp48 side chains. However, differences in the *B* factors of Trp48 in *P. fluorescens* is 4.3 $Å^2$ compared with 9.2 $Å^2$ of the same residue in *P. aeruginosa*. Differences are mainly seen in the atoms CD2, CE2, CD1 and NE1 of the side chain. The magnitudes range from two to five times higher in *P. aeruginosa*. This suggests a much higher level of disorder occurring in Trp48 of *P. aeruginosa*. This might account for the different distribution of the decay components or conformers in *P. aeruginosa* (C_s, 0.089, 0.067, 0.844) compared with that in *P. fluorescens* (C_s, 0.021, 0.046, 0.933) determined from the fluorescence decay measurement. In short, the structural heterogeneity observed in the fluorescence decay behavior and the difference between *P. fluorescens* and *P. aeruginosa* can well be accounted for based on the possible coexistence of three conformers and the difference in temperature factors of the tryptophan residues in these two azurins. The existence of three conformers can only be detected based on the fluorescence decay behavior due to the fact that the fluorescence depends on very local interactions.

The involvement of the His35 patch in close contact with the symmetrically related molecule (Fig. 15) would seem to support that this patch may play a role in electron-transfer function of azurin. Experiments with His35 mutant ruled out this possibility (van de Kamp, Floris *et al.*, 1990). However, this close contact does seem to provide a quenching pathway for Trp in the crystalline state (Fig. 16), as the fluorescence lifetime values are shorter for both holo and apo *P. fluorescens* in the crystalline state (apo = 3.79 ns; holo = 2.769 ns for the long component *versus* apo = 4.68 ns; holo = 4.49 ns in solution) (Dahms & Szabo, 1997).

6. Conclusions

Compared with other azurin structures, the unique features of the *P. fluorescens* are: the stable conformation of the main chain at His35 which is similar to the flip observed in *P. fluorescens* (Nar *et al.*, 1991*b*) and the direct involvement of the His35 patch in a close contact with the symmetrically related molecule. This provides a quenching pathway for the Trp in the crystalline state and explain the shorter fluorescence lifetime values. The refined structure supports the speculation that there is heterogeneity for the Trp48 side chain. The difference in *B* factor between the *P. fluorescens* and the *P. aeruginosa* azurin Trp side chain may account for differences in the Trp rotamer population. Finally it is clear from the refined structure that the primary mechanism for Trp fluorescence quenching is a short-range Dexter-type energy-transfer process. The direct evidence for Dexter type quenching provides another level of understanding to the quenching processes inherent to proteinaceous

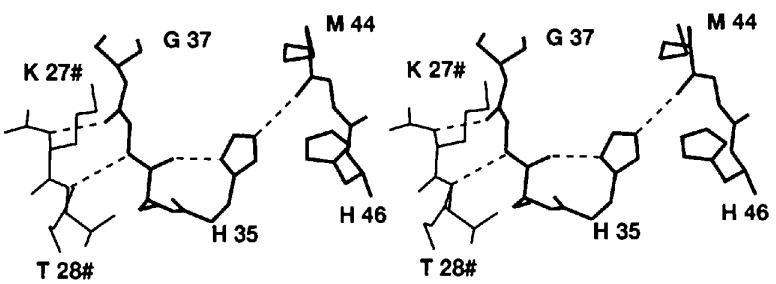

Fig. 16. The possible channel for static quenching.

environments. Taken as a whole this study has provided new ideas for understanding the information which can be obtained from Trp fluorescence and contributes to the growing database which will be used to design novel experiments for protein dynamics and interactions.*

We would like to thank Dr R. Durley of the Washington University School of Medicine for providing us with the parameters for the copper force field used in the refinement and for important discussions. This is NRCC publication number 39555.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1JOI, R1JOISF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: AM0039).

References

Adman, E. T. (1985). *Topics in Molecular and Structural Biology: Metalloproteins*, Vol. 6, edited by P. M. Harrison, pp. 1–42, Weinheim: Verlag Chemie.

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

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

Adman, E. T., Stenkamp, R. E., Sieker, J. C. & Jensen, L. H. (1978). *J. Mol. Biol.* **123**, 35–47.

Antonini, E., Finazzi-Agro, A., Avigliano, L., Guerrieri, P., Rotilio, G. & Mondovi, B. (1970). *J. Biol. Chem.* **245**, 4847–4856.

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

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

Birks, J. B. (1970). *Photophysics of Atomic Molecules*, pp. 569–570. New York: John Wiley.

Brünger, A. T. (1992). *X-PLOR* Version 3.1. Yale University Press, New Haven, CT, USA.

Brunori, M., Greenwood, C. & Wilson, M. T. (1974). *Biochem. J.* **137**, 113–116.

Chothia, C. & Lesk, A. M. (1982). *J. Mol. Biol.* **160**, 309–324.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**51**, 760–763.

Colman, P. M., Freeman, H. C., Guss, J. M., Murata, M., Norris, V. A., Ramshaw, J. A. M. & Venkatappa, M. P. (1978). *Nature (London)*, **272**, 319–324.

Cowgill, R. W. (1970). *Biochem. Biophys. Acta*, **207**, 556–559.

Dahms, T. E. S. & Szabo, A. G. (1997). *Methods Enzymol.* **278**, 202–221.

Dell, A. & Morris, H. R. (1977). *Biochem. Biophys. Res. Commun.* **78**, 3–10.

Durley, R. & Mathews, F. S. (1994). *J. Mol. Biol.* **236**, 1196–1211.

Farver, O. & Pecht, J. (1981). *Isr. J. Chem.* **21**, 13–17.

Gilardi, G., Mei, G., Rosato, N., Canters, G. W. & Finazzi, A. A. (1994). *Biochemistry*, **33**(6), 1425–1432.

Grinvald, A., Schlessinger, J., Pecht, I. & Steiberg, A. (1975). *Biochemistry*, **14**, 1921–1929.

Guss, J. M. & Freeman, H. C. (1983). *J. Mol. Biol.* **69**, 521–563.

Hill, H. A. O. & Smith, B. E. (1976). *Biochem. Biophys. Res. Commun.* **70**, 331–338.

Hirs, C. H. W. (1967). *Methods Enzymol.* **11**, 197–199.

Hutnik, C. M. & Szabo, A. G. (1989). *Biochemistry*, **28**, 3923–3934.

van de Kamp, M., Floris, R., Hali, F. C. & Canters, G. W. (1990). *J. Am. Chem. Soc.* **112**, 907–908.

van de Kamp, M., Silvestrini, M. C., Brunori, M., van Beeumen, J., Hali, F. C. & Canter, G. W. (1990). *Eur. J. Biochem.* **194**. 109–118.

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

Kraulis, P. T. (1991). *J. Appl. Cryst.* **24**, 946–950.

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

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mitra, S. & Bersohn, R. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 6807–6811.

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.

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. (1986). *J. Am. Chem. Soc.* **108**, 2784–2785.

Pecht, I. & Rosen, P. (1973). *Biochem. Biophys. Res. Commun.* **50**, 853–858.

Roe, S. M. & Teeter, M. M. (1993). *J. Mol. Biol.* **229**, 419–421.

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

Ryden, L. (1984). *Copper Proteins and Copper Enzymes*, Vol. I, edited by R. Lontie, pp. 157–182. Boca Raton, Florida: CRC Press.

Rosen, P. & Pecht, I. (1976). *Biochemistry*, **15**, 775–786.

Silvestrini, M. C., Brunori, M., Wilson, M. T. & Darley-Usmar, V. H. (1981). *J. Inorg. Chem.* **14**, 327–338.

Stanier, R. Y. (1947). *J. Bacteriol.* **54**, 191–194.

Szabo, A. G., Stepanik, T. M., Wagner, D. & Young, N. M. (1983). *Biophys. J.* **41**, 233–344.

Teale, F. W. & Weber, G. (1957). *Biochem. J.* **65**, 476–482.

Ugurbil, K., Norton, R. S., Allerhand, A. & Bersohn, R. (1977). *Biochemistry*, **16**, 886–894.

Wilson, M., Greenwood, C., Brunori, M. & Antonini, E. (1975). *Biochem. J.* **145**, 449–457.

Zhu, D. W., Dahms, T., Willis, K., Szabo, A. G. & Lee, X. (1994). *Arch. Biochem. Biophys.* **308**(2), 469–470.