Acta Cryst. (1994). **D50**, 263–270

Metal Substitution in a Blue-Copper Protein: the Crystal Structure of Cadmium-Azurin at 1.8 Å Resolution

BY KERRY A. BLACKWELL, BRYAN F. ANDERSON AND EDWARD N. BAKER*

Department of Chemistry and Biochemistry, Massey University, Palmerston North, New Zealand

(*Received 1 October 1993; accepted 21 December 1993*)

Abstract

Crystals of cadmium-substituted azurin have been prepared by diffusing Cd^{II} into crystals of apo-azurin grown previously and their structure has been determined at high resolution by X-ray crystallography. Data to 1.8 Å resolution were collected by Weissenberg photography (with image plates) using synchrotron radiation. These data were combined with a 2.2 Å diffractometer data set to give 90% coverage to 1.8 Å. An initial model was derived from the isomorphous Cu^{II}-azurin structure, and the cadmium and ligand positions added from 'omit' maps. Refinement was by restrained least squares (program *PROLSQ*), to a final R value of 0.168 for all data in the range 10.0–1.8 Å (23349 reflections). The final model of 1954 protein atoms, two Cd^{II} ions (occupancy 0.75), four SO_4^{2-} ions and 239 water molecules has r.m.s. deviations of 0.015, 0.045 and 0.013 Å from standard bond lengths, angle distances and planar groups. The protein structure is essentially the same as that of Cu^{II}-azurin, with an r.m.s. deviation of 0.18 Å for 97% of main-chain atoms after superposition of the two structures. The Cd atom is within 0.2 Å of the equivalent copper position, displaced slightly away from the axial Met ligand towards the carbonyl O atom of Gly45. The latter has also moved slightly towards the metal, by a rotation of the peptide unit, to give a Cd—O bond of 2.76 Å. The Cd—S(Cys) bond is lengthened to 2.39 Å. The coordination geometry is slightly more tetrahedral than for Cu^{II}, and the cadmium–oxygen interaction is consistent with the presence of an oxygen ligand in the coordination sphere of stellacyanin.

Introduction

Metal substitution has traditionally provided a powerful means of investigating the structural and electronic states of metal sites in proteins. A classic example is the substitution of ions such as Co^{II} for the spectroscopically silent Zn^{II} ions of zinc enzymes (Vallee & Williams, 1968) and zinc proteins (Roy, Brader, Lee, Kaarsholm, Hansen & Dunn, 1989). The approach has also been used to great advantage for metalloproteins whose spectroscopic properties were sufficiently unusual that few or no model complexes were available for comparison.

The blue-copper proteins include a group of low molecular weight, single-copper, electron-transfer proteins collectively known as cupredoxins (Adman, 1991). The best characterized of these are plastocyanin ($M_r \simeq 10$ kDa), from plants, and azurin ($M_r \simeq 14.6$ kDa), from denitrifying bacteria, but many others have also been investigated (for reviews see Gray & Solomon, 1981; Adman, 1985). These proteins have highly distinctive visible and EPR spectra and a low-symmetry metal site. Substitution of Co^{II} for the native Cu^{II} gave some of the strongest early evidence for Cu-cysteinate coordination (McMillin, Rosenberg & Gray, 1974), and metal substitution studies have also employed species such as Mn^{II}, Ni^{II}, Hg^{II} and Cd^{II}. Extrapolation from the metal-substituted protein to the native protein, however, relies on assumptions that the different metal ions occupy the same, or very similar, positions. We have undertaken the present study to explore this question.

The copper site in azurin is best described as trigonal planar or trigonal bipyramidal, with strong bonds of 2.0–2.1 Å to two His residues (46 and 117) and one Cys (112), and longer axial approaches to the copper of 3.0–3.1 Å from a peptide carbonyl O atom (from Gly45) and the thioether S atom of Met121 (see Fig. 1). These details have been established by high-resolution crystallographic studies of two azurins, from *Alcaligenes denitrificans* (Baker, 1988) and *Pseudomonas aeruginosa* (Nar, Messerschmidt, Huber, van de Kamp & Canters, 1991). Crystal structure analyses of the two apo-azurins further show that the ligand arrangement is stable even in the absence of the bound metal (Nar, Messerschmidt, Huber, van de Kamp & Canters, 1992; Shepard, Kingston, Anderson & Baker, 1993).

Cadmium(II) substitution is of interest because of the use of ^{113}Cd NMR spectroscopy as a probe of the metal sites of azurin, plastocyanin and stellacyanin

* To whom correspondence should be addressed (by air mail).

© 1994 International Union of Crystallography
Printed in Great Britain – all rights reserved

(Engeseth, McMillin & Otvos, 1984). This showed that the spectra of stellacyanin, the three-dimensional structure of which has not been experimentally determined, and azurin were very similar, but that of plastocyanin was rather different. Substitution of ^{111}Cd has also been used for studies employing perturbed angular correlation of γ-rays (PAC) spectroscopy (Andersson, Bauer & Demeter, 1982). We have, therefore, prepared crystals of cadmium-substituted azurin, from *Alcaligenes denitrificans*, and carried out a high-resolution crystallographic analysis. This shows the details of the cadmium coordination, and the way in which the protein structure responds to incorporation of a non-native metal. It also allows comparison with zinc-azurin, from *Pseudomonas aeruginosa*, the only other metal-substituted azurin to have been analysed crystallographically (Nar, Huber, Messerschmidt, Fillipou, Barth, Jaquinod, van de Kamp & Canters, 1992).

Experimental

Crystal data

Azurin was extracted and purified from *Alcaligenes denitrificans* (NCTC 8582) as previously described (Norris, Anderson, Baker & Rumball, 1979) and crystals of the apo-protein were prepared by vapour diffusion, from solutions of apo-azurin (Shepard, Kingston, Anderson & Baker, 1993). These were then soaked in standard mother liquor (0.1 *M* potassium phosphate buffer, pH 6.0, 75% saturated with ammonium sulfate) containing 0.1 *M* cadmium nitrate, for 3 d.

The cadmium-substituted crystals were isomorphous with both apo-azurin and holo-azurin crystals, having space group $C222_1$, with two molecules in the asymmetric unit and unit-cell dimensions $a = 75.15$, $b = 74.35$, $c = 99.35$ Å (*cf.* apo-azurin $a = 75.1$, $b = 74.1$, $c = 99.5$ Å and Cu^{II}-azurin $a = 75.0$, $b = 74.2$, $c = 99.6$ Å).

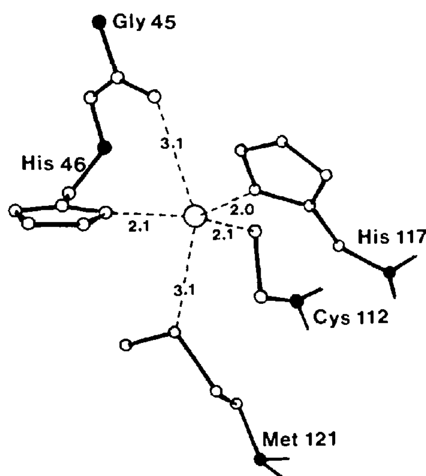

Fig. 1. Copper site in the oxidized form of azurin from *Alcaligenes denitrificans*.

Data collection

Two X-ray data sets were collected. First a 2.2 Å diffractometer data set was collected, using an Enraf–Nonius CAD-4 diffractometer equipped with a helium-filled diffracted-beam path. With the use of rapid data-collection techniques, including a limited-step ω-scan (Hanson, Watenpaugh, Sieker & Jensen, 1979), profile fitting and background averaging (Baker, 1988), the complete data set to 2.2 Å resolution ($\theta = 20.5°$) was obtained from one cadmium-azurin crystal. Data were corrected for radiation damage (intensity fall-off 17% over 3 d, estimated from five standard reflections monitored every 2 h), and absorption (North, Phillips & Mathews, 1968). The 13487 measured reflections represented approximately 98% of the possible data to 2.2 Å resolution. Of these 85% had intensities $I > 2\sigma_I$.

A second set of cadmium-azurin data, to 1.8 Å resolution, was collected by screenless Weissenberg photography using the Weissenberg camera equipped with imaging plates developed by Dr N. Sakabe at the Photon Factory synchrotron radiation source, Tsukuba, Japan (Sakabe, 1991). Data were collected from one cadmium-azurin crystal, mounted first about its *c* axis and then about *a*. In each case a rotation range exceeding 90°, and including both non-rotation axes, was covered; each image covered a rotation of 14°, with an overlap of 0.5° between successive images, giving a total range of 108.5° for eight exposures. The exposure time for the recording of each image was approximately 2 min, giving a total exposure in the X-ray beam of 16 min for each of the two axial rotations. With a camera radius of 430 mm and X-ray wavelength of 1.0 Å, the resolution limit was 1.8 Å. Diffraction images were processed with the program *WEIS* (Higashi, 1989). From a total of 108765 measured reflections to 1.8 Å resolution, 22591 unique reflections were obtained with a merging *R* value ($R = \sum|I - \bar{I}|/\sum\bar{I}$) of 0.060.

Finally, the diffractometer and synchrotron data sets were scaled and merged to give an overall data set of 23366 reflections to 1.8 Å resolution, representing 93% coverage of possible data. Of these 90% had intensities $I > 2\sigma_I$ (74% with $I > 2\sigma_I$ in the highest resolution shell, 1.8–1.9 Å). The merging *R* value for the diffractometer and synchrotron data sets was 0.062.

Refinement

The starting model for refinement was taken from the isomorphous Cu^{II}-azurin structure (Baker, 1988),

but with the removal of the metal ion and potential ligands (the side chains of His46, His117, Cys112 and Met121, and the carbonyl O atom of Gly45) from both molecules of the asymmetric unit, and all solvent molecules and ions. Refinement was by restrained least squares, using the fast Fourier version of the program *PROLSQ* (Hendrickson & Konnert, 1980). Parameters which were restrained included bond lengths, inter-bond angle distances, planar groups, non-bonded contact distances, chiral volumes and B factors. No restraints were placed on the cadmium bond lengths and angles, however, and no non-crystallographic symmetry restraints were imposed on the two crystallographically independent molecules in the asymmetric unit. The refinement strategy was essentially as described for the structure analysis of apo-azurin (Shepard, Kingston, Anderson & Baker, 1993), with periodic loosening and tightening of restraints and rebuilding from $F_o - F_c$, $2F_o - F_c$ and 'omit' maps, using the interactive graphics program *FRODO* (Jones, 1978).

Initial refinement was based on the 2.2 Å diffractometer data set, using all data in the resolution range 5.0–2.2 Å. The Cd atoms and their ligands were built into the model after the initial round of refinement when the R value was 0.23. The cadmium occupancy was estimated as 0.8 from comparison of the density at each metal site with that of the tightly bound water molecule attached to His117 $N_{\varepsilon 2}$ in each molecule; this had proved a good guide to the level of residual copper density in the apo-azurin structure (Shepard, Kingston, Anderson & Baker, 1993). The occupancies were revised downwards slightly to 0.75 in each site during refinement.

When the higher resolution synchrotron data set became available, the combined (diffractometer + synchrotron) data set was used to complete the refinement. Data in the resolution range 5.0–1.8 Å were used initially, with this being extended to 8.0–1.8 and then 10.0–1.8 Å as the solvent model was added. No σ cut-off of the data was used. Solvent molecules were added to the model where peaks in $F_o - F_c$ maps were greater than three times the r.m.s. deviation of the map, occupied stereochemically reasonable positions (*i.e.* were within hydrogen-bonding distance of other hydrogen-bonding groups) and were not close to protein groups whose conformations were in any doubt. All were regarded as water molecules, except for four which were recognized as SO_4^{2-} ions from the size and shape of their electron-density peaks and from their environments; SO_4^{2-} ions were found in the same sites in the apo-azurin structure (Shepard, Kingston, Anderson & Baker, 1993). Protein groups and solvent molecules with high B factors (>50 $Å^2$) were omitted from the model on several occasions for rebuilding or checking in 'omit' maps.

Table 1. *Refinement details*

(*a*) Course of the refinement

Round	Cycles	R value	Atoms	Comments
1	7	0.288→0.230	1916	5.0 2.2 Å CAD-4 data. Ligands and metal ions removed from model. No solvent
2	17	0.242→0.170	1956	5.0–2.2 Å CAD-4 data. Ligands and Cd^{2+} ions included. No solvent
3	16	0.206→0.155	2080	7.0–2.2 Å CAD-4 data. 124 water molecules added
4	10	0.283→0.266	1916	5.0 1.8 Å synchrotron/CAD-4 data. Ligands and metal ions removed No solvent
5	14	0.236→0.213	1956	5.0 1.8 Å synchrotron/CAD-4 data. Ligands and Cd^{2+} ions added. No solvent
6	15	0.230→0.187	2080	5.0–1.8 Å synchrotron/CAD-4 data. 124 water molecules
7	14	0.204→0.173	2191	8.0–1.8 Å synchrotron/CAD-4 data 215 water molecules, 4 SO_4^{2-} ions
8	15	0.178→0.168	2215	10.0–1.8 Å synchrotron/CAD-4 data. 239 water molecules, 4 SO_4^{2-} ions

(*b*) Refinement parameters

Parameter	Value
No. of reflections (no σ cut-off)	23349
No. of parameters refined*	8862
R.m.s. deviations from ideal values	
Bond (1–2) distances (Å)	0.015
Angle (1–3) distances (Å)	0.045
Planar groups (Å)	0.013
Chiral volumes ($Å^3$)	0.188
Mean B factor (all atoms) ($Å^2$)	27.7

* Overall scale and temperature factors and individual x, y, z, B values for 2215 atoms.

Some details of the refinement are given in Table 1. The final model comprised 1954 protein atoms (from two azurin molecules), two Cd^{2+} ions, four SO_4^{2-} ions and 239 water molecules. The protein structure was restrained close to standard geometry, with r.m.s. deviations from standard bond lengths, angle distances and planar groups of 0.015, 0.045 and 0.013 Å, respectively. The final crystallographic R value was 0.168 for all data in the range 10.0–1.8 Å (23 349 reflections). The atomic coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank.*

Results

The quality of the final model has been estimated from the agreement with the observed X-ray data, from comparisons of the two independently refined molecules of the asymmetric unit, and from comparisons with other protein structures refined at a similar resolution. A plot of the R value as a function of

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1AIZ, R1AIZSF). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Supplementary Publication No. SUP 37012). A list of deposited data is given at the end of this issue.

resolution (Luzzati, 1952) gives an estimated average coordinate error of 0.15 Å, while a σ_a plot (Read, 1986) gives a similar, though slightly higher, value of 0.24 Å. These are shown in Fig. 2. The Fit of the model to the electron density has been quantified by calculating real-space correlation coefficients, using programs from the *CCP*4 protein crystallography suite (SERC Daresbury Laboratory, 1986). For main-chain atoms, 75% of residues have correlation coefficients greater than 0.8, and only five residues out of 258 have values less than 0.5. For side chains 70% have correlation coefficients greater than 0.7 and 20 side chains out of 234 have values less than 0.5. The most poorly defined residues are identified by their high *B* factors and are the same in both molecules. As in other forms of *Alcaligenes denitrificans* azurin they include the N- and C-terminal residues, 1–2 and 128–129, the loop 104–107, and a number of surface side chains (those with $B > 50$ Å^2 are Glu4, Gln14, Lys18, Glu19, Lys27, Gln28, Lys52, Glu53, Gln57, Lys74, Thr78, Lys85, Asp98, Lys101 and Lys126).

A Ramachandran plot of the (φ,ψ) values of both molecules is shown in Fig. 3. All residues are within the allowed regions, with 92% being within the most favoured regions identified by Morris, MacArthur, Hutchinson & Thornton (1992).

Comparisons of the two molecules of the asymmetric unit give a further check on the accuracy of the model, although they do experience very slightly different crystal environments, which undoubtedly leads to some genuine microheterogeneity (Baker, 1988; Shepard, Kingston, Anderson & Baker, 1993). Main-chain atoms for residues 3–127 (500 atoms) can be superimposed with an r.m.s. deviation of 0.27 Å, while atoms of the cadmium site and its immediate surrounds (all atoms of residues 10–12, 15, 35, 44–47, 112–117 and 121; 122 atoms) have an r.m.s. deviation of 0.15 Å. The cadmium–ligand bond distances in the two sites have an r.m.s. difference of 0.07 Å (deviations of up to 0.05 Å from their mean values – see Table 2). The likely error in these bond lengths can be taken as 0.05–0.1 Å. In other respects, including the main-chain torsion angles φ and ψ, and the distribution of *B* factors along the polypeptide chain, the two molecules also correspond very closely; the agreement is essentially as presented in the detailed analyses of Cu^{II}-azurin (Baker, 1988) and apo-azurin (Shepard, Kingston, Anderson & Baker, 1993).

The metal site shows very little disturbance from that of the native Cu^{II} protein. A difference electron-density map calculated with coefficients $|F_{CdAz}| - |F_{CuAz}|$ and phases α_{CuAz} was essentially featureless apart from a large positive peak at the cadmium position, representing the greater electron density of the heavier atom, and a small peak adjacent to the carbonyl O atom of Gly45, representing a movement of this atom towards the cadmium (Fig. 4).

The electron density for the cadmium site in the final $2F_o - F_c$ map is shown in Fig. 5. Analysis of the metal geometry in terms of bond lengths and bond angles (Table 2) shows that it is slightly more tetrahedral than that of Cu^{II}-azurin (Baker, 1988). When compared with the latter, the Cd atom has moved away from the axial S(Met) ligand towards the other axial group, the carbonyl O atom of Gly45. The extent of this movement places the cadmium slightly (0.05 Å) on the carbonyl O-atom side of the plane of the three strongly bound ligands, His46, His117 and

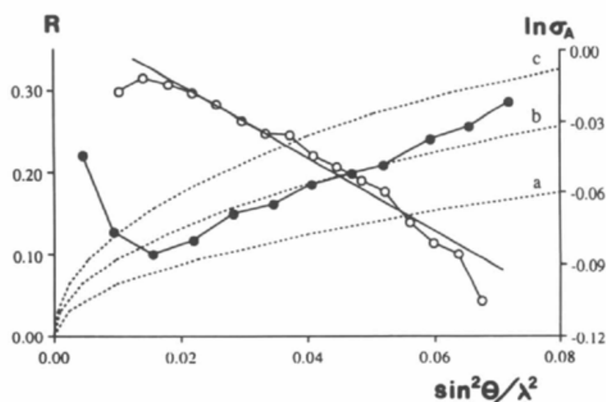

Fig. 2. Error estimates for Cd-azurin. The filled circles represent a Luzzati plot of *R* as a function of resolution with error levels shown as broken lines ($a = 0.10$, $b = 0.15$, $c = 0.20$ Å). The open circles represent a σ_A plot (Read, 1986).

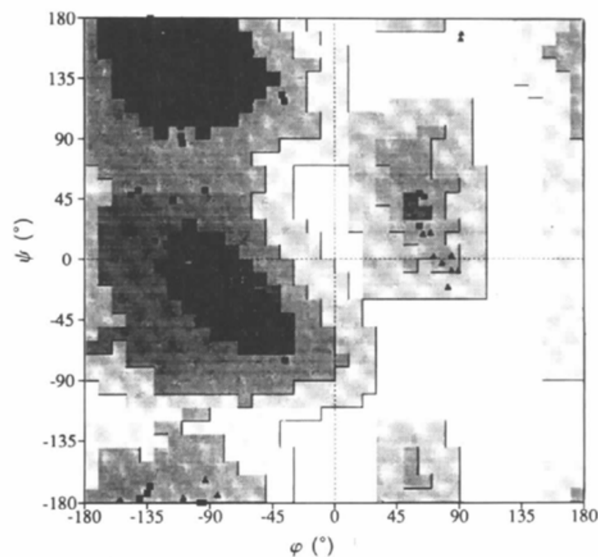

Fig. 3. Ramachandran plot (Ramachandran & Sasisekharan, 1968) for the two independent molecules of Cd-azurin, with glycine residues shown as triangles and all other residues as squares. Plot produced with *PROCHECK* (Morris, MacArthur, Hutchinson & Thornton, 1992).

Table 2. *Metal-site geometry compared with* Cu^{II}*-azurin*

Mean values are of two independent molecules (deviations from the mean are in parentheses).

Bond lengths (Å)	Cd^{II}-azurin	Cu^{II}-azurin
M—O(45)	2.76 (4)	3.13 (4)
M—$N_{\delta1}$(46)	2.25 (2)	2.08 (1)
M—$N_{\delta1}$(117)	2.21 (1)	2.00 (1)
M—S_{γ}(112)	2.39 (5)	2.15 (3)
M—S_{δ}(121)	3.23 (3)	3.11 (1)
Bond angles (°)		
O(45)—M—$N_{\delta1}$(46)	80 (3)	74 (2)
O(45)—M—$N_{\delta1}$(117)	86 (3)	80 (2)
O(45)—M—S_{γ}(112)	105 (4)	104 (1)
O(45)—M—S_{δ}(121)	150 (2)	147 (1)
$N_{\delta1}$(46)—M—$N_{\delta1}$(117)	106 (2)	105 (4)
$N_{\delta1}$(46)—M—S_{γ}(112)	132 (1)	135 (0)
$N_{\delta1}$(46)—M—S_{δ}(121)	73 (2)	77 (2)
$N_{\delta1}$(117)—M—S_{γ}(112)	121 (1)	119 (3)
$N_{\delta1}$(117)—M—S_{δ}(121)	90 (0)	96 (2)
S_{γ}(121)—M—S_{δ}(121)	102 (1)	107 (2)
Other distances (Å)		
O(45)···S_{δ}(121)	5.78 (4)	5.99 (3)
M···N_2S plane*	0.05 (1)	−0.12 (3)
O(45)···N_2S plane*	2.71 (1)	2.86 (2)
S_{δ}(121)···N_2S plane*	−3.06 (1)	−3.10 (5)
N(47)···S_{δ}(112)	3.45 (2)	3.48 (3)
N(114)···S_{γ}(112)	3.45 (2)	3.58 (8)
O(10)···$N_{\varepsilon2}$(46)	2.58 (7)	2.72 (9)
OW(170)···$N_{\varepsilon2}$(117)	2.69 (2)	2.72 (7)

* Plane defined by $N_{\delta1}$(46), $N_{\delta1}$(117), S_{γ}(112).

Cys112 [compared with 0.1 Å on the other side, towards S(Met), in the Cu^{II} structure]. At the same time the carbonyl O atom of Gly45 has moved approximately 0.15 Å towards cadmium and the N_2S plane by rotation of the Gly45-His46 peptide unit. The M—O distance is thus markedly shortened (0.37 Å), whereas all the other metal–ligand distances are slightly lengthened (by 0.1–0.3 Å), consistent with the larger radius of Cd^{II}, compared with Cu^{II}. The greatest change in the latter is the 0.24 Å increase in the M—S(Cys) distance. These movements can be seen in Fig. 6, where the two structures are superimposed.

The rest of the protein structure is hardly altered from the Cu^{II}- and apo-azurin structures. The ligand side chains make the same hydrogen bonds (His46 $N_{\varepsilon2}$···O(10) 2.58 Å; His117 $N_{\varepsilon2}$···OW(170) 2.69 Å; Cys112 S_{γ}···N(47) 3.45 Å; Cys112 S_{γ}···N(114) 3.45 Å) but all are shortened slightly to allow the expansion of the metal site. The radius of the cavity between the equatorial ligands, His46 $N_{\delta1}$, His117 $N_{\delta1}$ and Cys112 S_{γ}, is expanded from 1.16 Å in apo-azurin and 1.24 Å in Cu^{II}-azurin to 1.35 Å in Cd^{II}-azurin. There appears to be a small associated movement of the loop between Cys112 and His117 (especially Phe114 and Pro115), which can be seen in Fig. 6. The size of the movement is only marginally significantly (r.m.s. deviation 0.15 Å for the main-chain atoms of this loop) but the fact that the same phenomenon is seen for both molecules suggests that it may be a real effect. In general, however, superposition of the polypeptide chain shows very close correspondence throughout the molecule. Taking all main-chain atoms (N, C_{α}, C, O) but omitting the poorly ordered N- and C-terminal residues 1–2 and 128–129 gives an r.m.s. deviation of 0.13 Å between Cd^{II} and Cu^{II} structures and 0.14 Å between Cd^{II} and apo structures.

Solvent molecules in the cadmium-azurin crystal structure were located independently of those found for the Cu^{II} and apo-azurin structures. Comparison of the three solvent models showed that 140 water molecules (60%) were common to all three using a cut-off of 1.0 Å for equivalence.

Discussion

The structure of cadmium-azurin seen in the present study extends the picture of the blue-copper site that has been developed through comparative high-resolution crystallographic studies of both azurin and plastocyanin. Removal of copper from either protein leaves the binding site essentially unchanged (Garrett, Clingeleffer, Guss, Rogers & Freeman, 1984; Shepard, Kingston, Anderson & Baker, 1993), although in both cases an alternative configuration of one of the His ligands (His117 in azurin) is accessible to permit metal removal or insertion (Garrett, Clingeleffer, Guss, Rogers & Freeman, 1984; Nar, Messerschmidt, Huber, van de Kamp & Canters, 1992). The implication is that the binding site is primarily determined by the constraints of the protein structure. This has functional advantages in allowing redox changes to occur with minimal structural change (Shepard, Anderson, Lewandoski, Norris & Baker, 1990).

The present study, however, together with a similar analysis of zinc-substituted azurin (Nar, Huber, Messerschmidt, Fillipou, Barth, Jaquinod, van de Kamp & Canters, 1992) shows that a limited flexibility does exist when a non-native metal is substituted. Thus in Zn^{II}-azurin the metal atom is displaced substantially towards the carbonyl O atom of Gly45, which in turn moves towards the zinc through a twist of the Gly45-His46 peptide; the result is a much shorter M—O bond (2.32 Å for Zn, compared with 2.96 Å for Cu in *P. aeruginosa* azurin), and a much more tetrahedral geometry.

A similar effect is seen in cadmium-azurin (this work), though much less pronounced. The movements of the metal and Gly45 O are of a similar nature but the resulting M—O distance is longer, at 2.76 Å (although still substantially shorter than the Cu—O distance of 3.13 Å in *A. denitrificans* azurin). The lesser movement may be attributable to the larger size of Cd^{II} compared with Zn^{II} (which allows

a significant Cd—O interaction with a longer bond) or to different coordination preferences. The larger metal in turn is accommodated by the protein structure with little disturbance because the longer bonds to the equatorial His and Cys ligands are taken up through shorter hydrogen bonds with the surrounding protein structure.

The Cd—O distance of 2.76 Å implies a weak to moderate bond, when compared with 'normal' Cd—O distances in small molecule complexes of 2.2–2.3 Å (Greaney, Raston & White, 1975). The movement of the Cd atom away from both sulfur donors, with the second largest change being the movement of 0.24 Å away from Cys112, is consistent with a preference of Cd^{II} for oxygen and nitrogen ligands rather than sulfur.

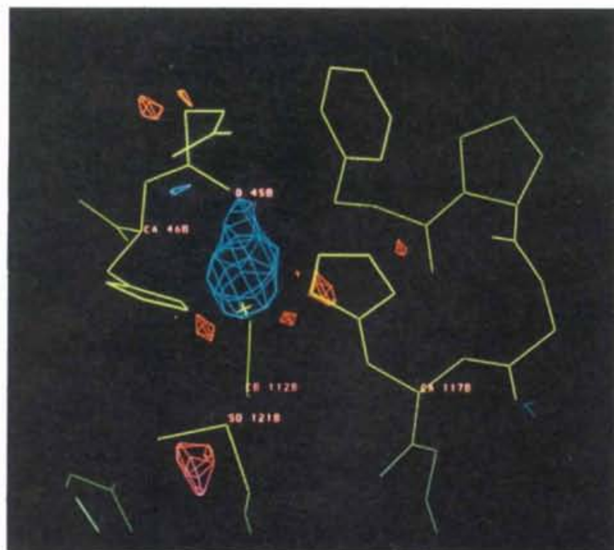

Fig. 4. Difference electron density map calculated with coefficients $|F_{CdAz}| - |F_{CuAz}|$ and phases α_{CuAz}. Positive density (blue) contoured at 5σ, negative density (red) at -5σ. Model for Cu-azurin is superimposed. The large positive peak results from small movements of the metal ion and the Gly45 carbonyl O atom towards each other when Cd^{II} is substituted for Cu^{II}.

Changes in the metal site on metal substitution appear more pronounced for azurin than for plastocyanin. In Hg^{II}-substituted plastocyanin, for example, the metal geometry stays essentially the same, while all of the metal–ligand bond distances are increased slightly (Church, Guss, Potter & Freeman, 1986). One major difference between the metal environments of plastocyanin and azurin is the much closer approach of the carbonyl O atom of Gly45 to the coordination sphere; the equivalent O atom in Cu^{II}-plastocyanin is 0.7 Å further away from the metal than in Cu^{II}-azurin. Part of this difference (0.2 Å) is attributable to the greater displacement of copper from the N_2S plane in plastocyanin than in azurin (Baker, 1988) but this still makes the O atom 0.5 Å closer to the N_2S plane in azurin. This fundamental difference adds an extra possibility for geometrical flexibility in the azurin site, as the metal can find potential ligands on either side of the N_2S plane.

The structural observations made here explain the difference in the ^{113}Cd NMR spectra of azurin and plastocyanin following metal substitution (Engeseth, McMillin & Otvos, 1984). In azurin the metal interacts with an oxygen ligand, while in plastocyanin it does not. The similar ^{113}Cd NMR spectra found for azurin and stellacyanin (Engeseth, McMillin & Otvos, 1984), when taken with the observation of a Cd—O bond in cadmium-azurin, further support suggestions that an oxygen ligand is present in the

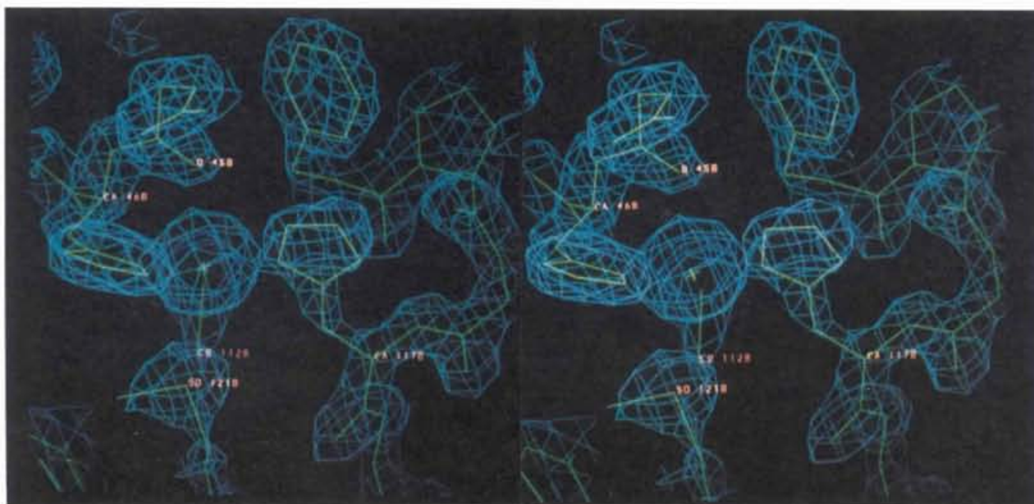

Fig. 5. Stereo diagram of the electron density ($2F_o - F_c$) at the metal site of Cd-azurin (shown for molecule *B*).

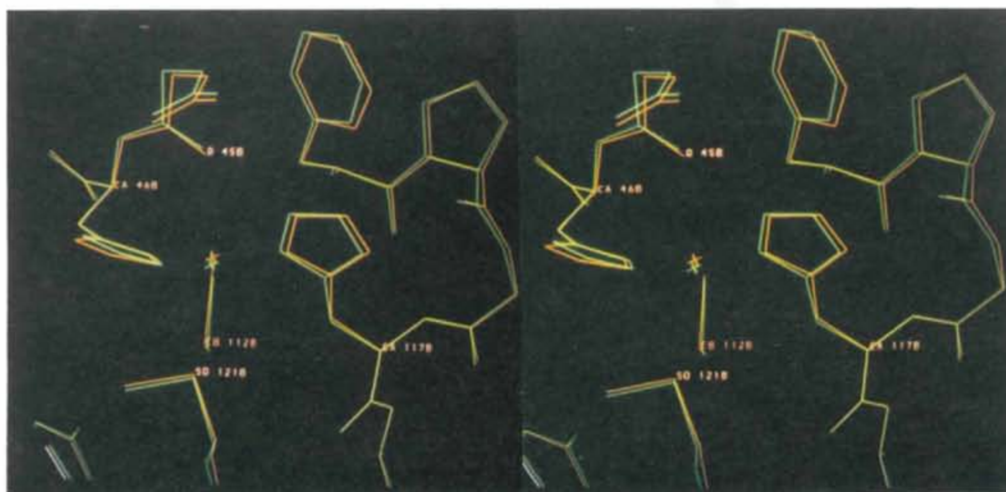

Fig. 6. Comparison of the metal sites for Cd-azurin (orange) and oxidized Cu-azurin (yellow). Superposition based on the main-chain atoms of residues 3–127. Note the small movement of Phe114 and Pro115 accompanying that of the metal ligands.

coordination sphere of stellacyanin. Whether this is the amide O atom of a Gln side chain, as suggested from model-building experiments (Fields, Guss & Freeman, 1991), or a carbonyl O atom analogous to that of Gly45, cannot be ascertained from the present data. A coordinated peptide oxygen could account for the pulsed electron nuclear double resonance (ENDOR) results obtained for stellacyanin (Thomann, Bernardo, Baldwin, Lowery & Solomon, 1991) if a peptide 'flip' occurs at high pH, as proposed.

Finally, we note that significant changes to the metal geometry also occur when different amino acids are substituted for the native ligands (Romero, Hoitink, Nar, Huber, Messerschmidt & Canters, 1993). It can be concluded that the protein structure is designed to provide a particular geometry for the native metal with the native ligands, and that these are optimized for biological function; changes to either metal or ligands can change this delicate relationship.

We thank Dr Sakabe for help with data collection at the Photon Factory synchrotron source and Dr C. A. Smith for help with illustrations. ENB is an International Research Scholar of the Howard Hughes Medical Institute.

References

ADMAN, E. T. (1985). *Metalloproteins*, Vol. 2, edited by P. HARRISON, pp. 1–42. London: Macmillan.

ADMAN, E. T. (1991). *Adv. Protein Chem.* **32**, 145–197.

ANDERSSON, I., BAUER, R. & DEMETER, I. (1982). *Inorg. Chim. Acta*, **67**, 53–59.

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

CHURCH, W. B., GUSS, J. M., POTTER, J. J. & FREEMAN, H. C. (1986). *J. Biol. Chem.* **261**, 234–237.

ENGESETH, H. R., MCMILLIN, D. R. & OTVOS, J. D. (1984). *J. Biol. Chem.* **259**, 4822–4826.

FIELDS, B. A., GUSS, J. M. & FREEMAN, H. C. (1991). *J. Mol. Biol.* **222**, 1053–1065.

GARRETT, T. P. J., CLINGELEFFER, D. J., GUSS, J. M., ROGERS, S. J. & FREEMAN, H. C. (1984). *J. Biol. Chem.* **259**, 2822–2825.

GRAY, H. B. & SOLOMON, E. I. (1981). *Copper Proteins*, edited by T. SPIRO, pp. 1–39. New York: John Wiley.

GREANEY, M. T., RASTON, C. L. & WHITE, A. H. (1975). *J. Chem. Soc. Dalton Trans.* pp. 876–879.

HANSON, J. C., WATENPAUGH, K. D., SIEKER, L. C. & JENSEN, L. H. (1979). *Acta Cryst.* A**35**, 616–621.

HENDRICKSON, W. A. & KONNERT, J. H. (1980). *Biomolecular Structure, Function, Conformation and Evolution*, Vol. 1, edited by R. SRINIVASAN, pp. 43–57. Oxford: Pergamon Press.

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

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

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

MCMILLIN, D. R., ROSENBERG, R. C. & GRAY, H. B. (1974). *Proc. Natl Acad. Sci. USA*, **71**, 4760–4762.

MORRIS, A. L., MACARTHUR, M. W., HUTCHINSON, E. G. & THORNTON, J. M. (1992). *Proteins*, **12**, 345–364.

NAR, H., HUBER, R., MESSERSCHMIDT, A., FILLIPOU, A. C., BARTH, M., JAQUINOD, M., VAN DE KAMP, M. & CANTERS, G. W. (1992). *Eur. J. Biochem.* **205**, 1123–1129.

NAR, H., MESSERSCHMIDT, A., HUBER, R., VAN DE KAMP, M. & CANTERS, G. W. (1991). *J. Mol. Biol.* **221**, 765–772.

NAR, H., MESSERSCHMIDT, A., HUBER, R., VAN DE KAMP, M. & CANTERS, G. W. (1992). *FEBS Lett.* **306**, 119–124.

NORRIS, G. E., ANDERSON, B. F., BAKER, E. N. & RUMBALL, S. V. (1979). *J. Mol. Biol.* **135**, 309–312.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

RAMACHANDRAN, G. N. & SASISEKHARAN, V. (1968). *Adv. Protein Chem.* **23**, 283–437.

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

ROMERO, A., HOITINK, C. W. G., NAR, H., HUBER, R., MESSERSCHMIDT, A. & CANTERS, G. W. (1993). *J. Mol. Biol.* **229**, 1007–1021.

ROY, M., BRADER, M. L., LEE, R. W.-K., KAARSHOLM, N. C., HANSEN, J. F. & DUNN, M. F. (1989). *J. Biol. Chem.* **264**, 19081–19085.

SAKABE, N. (1991). *Nucl. Instrum. Methods Phys. Res.* A**303**, 448–463.

SERC Daresbury Laboratory (1986). *CCP4. A Suite of Programs for Protein Crystallography*. SERC Daresbury Laboratory, Warrington WA4 4WD, England.

SHEPARD, W. E. B., ANDERSON, B. F., LEWANDOSKI, D. A., NORRIS, G. E. & BAKER, E. N. (1990). *J. Am. Chem. Soc.* **112**, 7817–7819.

SHEPARD, W. E. B., KINGSTON, R. L., ANDERSON, B. F. & BAKER, E. N. (1993). *Acta Cryst.* **D49**, 331–343.

THOMANN, H., BERNADO, M., BALDWIN, M. J., LOWERY, M. D. & SOLOMON, E. I. (1991). *J. Am. Chem. Soc.* **113**, 5911–5913.

VALLEE, B. L. & WILLIAMS, R. J. P. (1968). *Proc. Natl Acad. Sci. USA*, **59**, 498–503.