49. T. G. M. Katada *et al.*, *J. Biol. Chem.* **259**, 3586 (1984); A. G. Gilman, *Cell* **36**, 577 (1984); *Annu. Rev. Biochem.* **56**, 615 (1987).
50. A. Levitzki, *Physiol. Rev.* **66**, 819 (1986).
51. A. M. Tolkovsky, S. Braun, A. Levitzki, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 213 (1982); A. M. Tolkovsky and A. Levitzki, *J. Cyclic Nucleotide Res.* **7**, 139 (1981).
52. H. Arad, J. Rosenbusch, A. Levitzki, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6579 (1984).
53. I. Marbach, A. Bar-Sinai, A. Levitzki, manuscript in preparation.
54. A. Levitzki, *J. Recept. Res.* **4**, 399 (1984); *FEBS Lett.* **211**, 113 (1987).
55. M. D. Smigel, *J. Biol. Chem.* **261**, 1976 (1986).
56. The complex between GPPNHP-activated G_s and C was purified beyond the stage described in (*52*) with high-performance liquid chromatography (HPLC) purification steps. The amount of the β subunit of the G_s protein was examined quantitatively with specific antibodies to the β subunit of G_s. Assuming that the turnover number of the maximally activated adenylate cyclase is ~1100 min^{-1} we find that the β subunit to C ratio varies between 1.0 and 2.5 at all purification stages. The β subunit accompanies the GPPNHP-activated enzyme but is absent from adjacent fractions in the chromatographic step.
57. I. Marbach, J. Schiloach, A. Levitzki, *Eur. J. Biochem.* **172**, 239 (1988).
58. A. K. Keenan, A. Gal, A. Levitzki, *Biochem. Biophys. Res. Commun.* **105**, 615 (1982).
59. J. Kirilovsky, S. Eimerl, S. Steiner-Mordoch, M. Schramm, *Eur. J. Biochem.* **166**, 221 (1987).
60. R. J. Scarore and J. B. Abrams, *J. Pharmacol. Exp. Ther.* **223**, 327 (1982).
61. Y. F. Su, L. Cubeddu-Ximenez, J. P. Perkins, *J. Cyclic. Nucleotide Res.* **2**, 257 (1976); Y. F. Su *et al.*, *ibid.*, p. 271.
62. T. K. Harden *et al.*, *Science* **210**, 441 (1980); G. L. Waldo, J. K. Northrup, J. P. Perkins, T. K. Harden, *J. Biol. Chem.* **258**, 13900 (1983); J. M. Stadel *ibid.*, p. 3032.
63. C. Hertel, M. Staehelin, J. P. Perkins, *J. Cyclic Nucleotide Res.* **9**, 119 (1983); M. L. Toews, G. L. Waldo, T. K. Harden, J. P. Perkins, *J. Biol. Chem.* **259**, 11844 (1984); C. Hertel and J. P. Perkins, *Mol. Cell. Endocrinol.* **37**, 245 (1984).
64. Y. F. Su, T. K. Harden, J. P. Perkins, *J. Biol. Chem.* **255**, 7410 (1980).
65. B. Strulovici, J. M. Stadel, R. J. Lefkowitz, *ibid.* **258**, 6410 (1983).
66. W. B. Anderson and C. Jaworski, *ibid.* **254**, 4596 (1979); A. Levitzki and D. Atlas, *Life Sci.* **28**, 661 (1980).
67. P. A. Insel, *J. Biol. Chem.* **258**, 13597 (1981).
68. D. A. Green and R. B. Clark, *ibid.* **256**, 2105 (1981).
69. C. Strader *et al.*, *Cell* **49**, 855 (1987).
70. R. L. Doss, J. P. Perkins, T. K. Harden, *J. Biol. Chem.* **256**, 12281 (1981).
71. J. L. Benovic, R. H. Strasser, M. G. Caron, R. L. Lefkowitz, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2797 (1986).
72. R. H. Strasser, J. L. Benovic, M. G. Caron, R. L. Lefkowitz, *ibid.*, p. 6362; J. L. Benovic *et al.*, *J. Biol. Chem.* **262**, 17251 (1987).
73. H. Shidri and R. L. Somers, *J. Biol. Chem.* **253**, 7040 (1978).
74. R. Zuckerman *et al.*, *Biophys. J.* **47**, 37a (1985); U. Wilden, S. W. Hall, H. Kuhn, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1174 (1986).
75. J. Benovic *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8879 (1987).
76. R. Clark *et al.*, *ibid.* **85**, 1442 (1988).
77. A. O. Davies and R. J. Lefkowitz, *J. Clin. Invest.* **71**, 565 (1983); P. J. Scarpace, L. A. Baresi, D. A. Sanford, I. B. Abrass, *Mol. Pharmacol.* **28**, 495 (1985); R. O. Salonen, *Acta Pharmacol. Toxicol.* **57**, 147 (1985).
78. Supported by NIH grant GM 37110. The author wishes to thank Dr. A. Bar-Sinai for critically reading the manuscript.

Research Articles

Phase Determination by Multiple-Wavelength X-ray Diffraction: Crystal Structure of a Basic "Blue" Copper Protein from Cucumbers

J. MITCHELL GUSS, ETHAN A. MERRITT,* R. PAUL PHIZACKERLEY, BRITT HEDMAN, MITSUO MURATA,† KEITH O. HODGSON, HANS C. FREEMAN

A novel x-ray diffraction technique, multiple-wavelength anomalous dispersion (MAD) phasing, has been applied to the de novo determination of an unknown protein structure, that of the "blue" copper protein isolated from cucumber seedlings. This method makes use of crystallographic phases determined from measurements made at several wavelengths and has recently been made technically feasible through the use of intense, polychromatic synchrotron radiation together with accurate data collection from multiwire electronic area detectors. In contrast with all of the conventional methods of solving protein structures, which require either multiple isomorphous derivatives or coordinates of a similar structure for molecular replacement, this technique allows direct solution of the classical "phase problem" in x-ray crystallography. MAD phase assignment should be particularly useful for determining structures of small to medium-sized metalloproteins for which isomorphous derivatives are difficult or impossible to make. The structure of this particular protein provides new insights into the spectroscopic and redox properties of blue copper proteins, an important class of metalloproteins widely distributed in nature.

THE CLASSIC PHASE PROBLEM IN X-RAY CRYSTALLOGRAPHY can be solved with the use of anomalous scattering effects. As the energy of an incident x-ray beam is varied across the absorption edge of an element, there may be substantial changes in the real and imaginary components (f' and f'') of the x-ray scattering. In crystal structures that contain atoms with large "anomalous scattering" effects, the net observed intensity of each Bragg reflection will then be energy dependent. In such cases, the differences between the Bragg intensities measured from a single crystal at several x-ray energies may be used to directly derive crystallographic phases and hence to determine the crystal structure. Multiple-wavelength anomalous dispersion (MAD) phase assignment is potentially applicable to any macromolecular crystal structure that contains one or more anomalous scatterers (*1*). Metallopro-

J. M. Guss, M. Murata, and H. C. Freeman are in the Department of Inorganic Chemistry, University of Sydney, Sydney, New South Wales, 2006, Australia. E. A. Merritt, R. P. Phizackerley, and B. Hedman are in the Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, CA 94309. K. O. Hodgson is in the Department of Chemistry, Stanford University, Stanford, CA 94305.

*Present address: Department of Biological Structure, University of Washington, Seattle, WA 98105.
†Present address: Department of Biochemistry, University of Georgia, Athens, GA 30606.

Table 1. Anomalous dispersion terms and R_{sym} values at the four x-ray wavelengths used in data collection.

$$R_{sym} = \{\sum_{hkl}\sum_{i}|I_i - \bar{I}_{hkl}|\} / \{\sum_{hkl}\sum_{i} I_i\}$$

where the second summation is over all redundant and all space-group symmetry-equivalent measurements at a given *hkl*. The R_{sym} values under (1) were obtained when the data were processed with a conventional model for coincidence loss as a function of detector count rate. The values under (2) resulted from the empirical scaling procedure described in (*12*). All R_{sym} values are for the data to 2.5 Å resolution.

X-ray energy (wavelength)	f' (electrons)	f'' (electrons)	R_{sym} (1)	R_{sym} (2)	Maximum count rate (kHz)	Unique reflections
10.0301 keV (1.2359 Å)	−1.61	3.27	0.120	0.046	68	14109
9.0022 keV (1.3771 Å)	−6.17	4.17	0.107	0.045	66	10830
8.9900 keV (1.3790 Å)	−8.11	2.54	0.097	0.045	66	10781
8.0414 keV (1.5416 Å)	−2.55	0.60	0.055	0.043	57	7746

teins are obvious candidates for the technique; even proteins without metal atoms in their native state may be made amenable to MAD phase assignment by chemical modification or by co-crystallization with an anomalous scatterer (*2*). Many of the difficulties inherent in isomorphous replacement methods are thus bypassed: data are collected from a single crystal form, a laborious search for derivatives is unnecessary, and the question of imperfect isomorphism does not arise. The phasing power of the MAD technique actually increases for higher resolution data, since the magnitude of the anomalous dispersion scattering does not decrease with scattering angle. The application of MAD phasing has been made technically feasible through the use of intense polychromatic synchrotron radiation together with accurate data collection from multiwire electronic area detectors.

The MAD phasing technique appeared particularly well suited to solving a difficult and long-standing problem. In 1971 and 1974, two groups of investigators independently reported the occurrence of a basic copper-containing protein in cucumbers (*3, 4*). In view of the occurrence of the protein in several plant sources (*5, 6*), the names "cusacyanin" and "plantacyanin" were proposed. As the protein has spectroscopic and redox properties that show that it belongs to the class of blue copper proteins, we refer to it merely as CBP, "cucumber basic blue protein." We crystallized CBP in 1976, and preliminary crystallographic data were recorded (*7*). Only one heavy-atom isomorphous derivative was successfully prepared (with mercuric acetate), and then only from crystals of the native protein cross-linked with glutaraldehyde. A map calculated by single isomorphous replacement techniques defied interpretation. Our attempts to solve the structure by molecular replacement with models based on the known structure of another blue copper protein, plastocyanin, also failed. However, the structure was readily solved with MAD phasing.

Because MAD phasing for protein structure analysis is so new, too few experiments have been completed to determine how large an anomalous dispersion signal is required to solve a protein structure of a given size [although we have studied this question theoretically (*1*)]. The phasing power of the MAD technique is greater when the signal is large, as is the case at the L absorption edges of the lanthanides (*8*). The large signal at the Tb L_{III} edge (f' ~28 electrons, and f'' ~20 electrons) was exploited by Kahn *et al.* in the determination of the *Opsanus tau* parvalbumin structure (*9*). The substitution of Tb^{3+} at the two Ca^{2+} binding sites in this protein (molecular weight, M_r 10,100) introduced a large anomalous dispersion signal. In contrast, the magnitudes of f' and f'' are typically less than 10 electrons at the K absorption edges of the transition elements. Thus it is significant that in the present work the signal from a single Cu atom in the native CBP (M_r 10,100) was sufficient for structure determination with MAD phasing (Table 1) (*10*).

Experimental. Crystals of CBP were grown by hanging-drop vapor diffusion against 40 percent polyethylene glycol–6000 in 0.1*M* phosphate buffer (*p*H 6.0). The x-ray energies for data collection were chosen after characterization of the energy dependence of the anomalous dispersion terms f' and f'' exhibited by the sample crystals in the x-ray region that spans the CuK absorption edge. For this purpose, the x-ray fluorescence from a single, oriented crystal of CBP was measured as a function of the incident x-ray energy with a scintillation counter positioned in the horizontal plane and within 2 cm of the sample crystal at 90 degrees to the 95 percent horizontally polarized incident beam. Figure 1 shows the variations in f' and f'' observed near the CuK absorption edge. Two of the x-ray energies used for the data collection were chosen to lie at the absorption edge: one at the point of maximum f'', and one at the point of maximum negative f'. The remaining two energies were chosen approximately 1 keV above and below the edge (the latter specifically at the CuK_α line). Bragg intensities were measured from two crystals of CBP with dimensions 0.37 mm by 0.37 mm by 0.13 mm and 0.37 mm by 0.37 mm by 0.08 to 0.12 mm, respectively, by using the area detector facility built specifically for exploiting the MAD phasing technique at the Stanford Synchrotron Radiation Laboratory (SSRL) (*11, 12*). To the extent possible, the diffraction geometry was chosen so that Bijvoet pairs of reflections (F^+ and F^-) were measured simultaneously on different portions of the detector (*13*). The 85,374 integrated Bragg intensities were partitioned into 140 bins, each bin corresponding to a rotation of the sample crystal by about 8 degrees at a single energy. A linear scale factor was assigned to each bin to minimize the overall R_{sym}, and the redundant and symmetry-equivalent observations were averaged to yield a consensus value of F^+ and F^- for each reflection at each energy (Table 1).

The data used for the MAD phase assignment comprised 3550 independent reflections (99 percent of the accessible data) measured

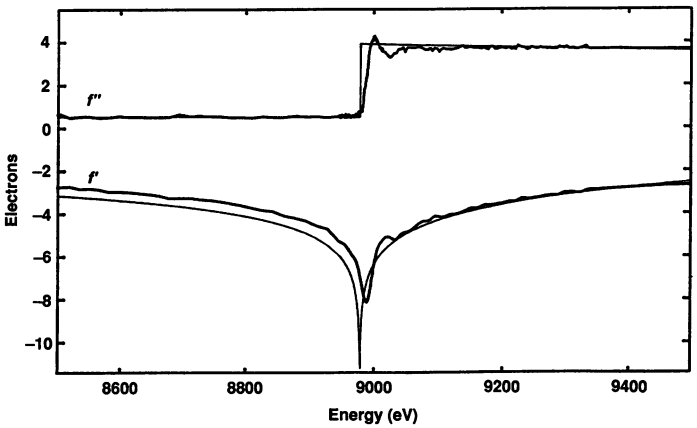


Fig. 1. Energy dependence of the anomalous dispersion terms f'' and f' in the region of the CuK absorption edge. Values of f'' and f' are in electrons. Experimental values for f'' (heavy line) were obtained from x-ray fluorescence from a single crystal of CBP; ideal f'' values (thin line) for atomic Cu are from (*58*). Experimental values for f' are derived by numerical integration from the f'' spectrum with the Kramers-Kroenig relation; ideal f' values (thin line) are from Hönl theory (*59*). Derivation of the experimental f'' and f' values was performed with an in-house program DISCO (*60*).

to a 2.5 Å resolution; 2095 were represented by all eight possible observations (F^+ and F^- at four energies) and 3430 were represented by four or more observations. From the multiple observations for each unique reflection we derived an estimate of the partial scattering contribution F_{Cu} of the Cu atom to that reflection with the algorithm suggested by Karle (*14*) and implemented by Hendrickson (*15*) in his program MADLSQ. A Patterson map that used coefficients F^2_{Cu} from the data arbitrarily limited to 3.5 Å resolution revealed the Cu atom location (subject to a sign ambiguity in one coordinate). The Cu coordinates thus obtained were refined by least squares against the estimated F_{Cu} for the entire data set. At this point, we could calculate the scattering contributions F_{Cu} and ϕ_{Cu} of the refined Cu atom partial structure factor at each energy. From these and the measured intensities we derived a crystallographic amplitude F_P and phase ϕ_P for the normal (nonanomalous) scattering component of each reflection with a procedure analogous to that used for the assignment of multiple isomorphous replacement (MIR) phases, but with the additional need to estimate the "native protein" amplitude F_P. The phases were assigned as follows: for possible values of ϕ_P taken at 10-degree intervals, the estimate of F_P was refined to minimize the lack-of-closure residual expressing the disparity between the observed amplitudes Fo_i and the predicted amplitudes Fc_i:

$$\text{Lack of closure} = \Sigma\,(Fo_i^2 - Fc_i^2)^2 \quad (1)$$

where

$$|Fc_i|^2 = [F_{Cu}\cos\phi_{Cu} + F_P\cos\phi_P]^2 + [F_{Cu}\sin\phi_{Cu} + F_P\sin\phi_P]^2 \quad (2)$$

By analogy with the Blow-Crick formulation for MIR phases (*16*), the likelihood associated with the phase angle ϕ_P was taken to be:

$$P(\phi_P) = \exp[-\Sigma(Fo_i - Fc_i)^2/2nE^2] \quad (3)$$

The $2n$ observations Fo_i are not independent in the sense that the Fo from different derivative crystals are independent in the Blow-Crick formulation. The F_P term (which in the MIR case is simply the Fo for the native crystal) is a refined quantity rather than a constant. For both of these reasons, the "error" term E^2 is not strictly equivalent to that in the Blow-Crick formulation; here E^2 was treated as an empirically determined constant. As in MIR phasing, given the likelihood distribution one may choose either the most probable phase or a "best" centroid phase estimate and a figure of merit.

Two electron density maps were calculated at this point, one for each of the possible signs of the z coordinate of the Cu atom. Both used data to 3.0 Å resolution and figure-of-merit–weighted centroid phases. Only one of these maps was clearly interpretable as a protein structure with a well-defined molecular boundary, thus resolving the ambiguity in the z coordinate of the Cu atom. Prior to fitting a model, we reduced the noise in this map by solvent flattening (Fig. 2). The polypeptide backbone corresponding to 90 of the 96 residues in the known sequence (*17*) could be traced in a minimap at this point. The electron density of only six residues (12 to 14 and 23 to 25) was sufficiently weak or discontinuous to cause uncertainties in interpretation. Further model-building and optimization were performed on an Evans and Sutherland PS300 display system, with the program FRODO (*18*). The remaining six residues were identified; further refinement by means of the program PROLSQ should provide additional information (*19*). The present residual R is 0.22 for the 7167 reflections recorded at λ = 1.2359 Å in the range 1.8 Å $\leq d \leq$ 6.0 Å; at 1.8 Å resolution the data set is 84 percent complete at this wavelength, whereas at 3.0 Å resolution the data set is 98 percent complete.

Structure of CBP. The structure of CBP is shown as a Cα plot in Fig. 3. The backbone consists of eight strands of polypeptide. Part of strand 1 and all of strand 2 have irregular conformations. The NH_2-terminal region of strand 1 and substantial portions of the remaining strands have β conformations. Only five of the strands—1, 3, 6, 7, and 8—form a β sandwich. Strand 2 covers one side of the sandwich. Strands 4 and 5 are bent and twisted so that their directions are roughly perpendicular to the other polypeptide strands. Near the beginning of strand 4 lies His^{39}, one of the Cu-binding residues. The other Cu ligands are Cys^{79}, His^{84}, and Met^{89}. These three residues are located on a double loop linking strands 7 and 8. A second Cys residue, Cys^{85}, also lies on this double loop, but

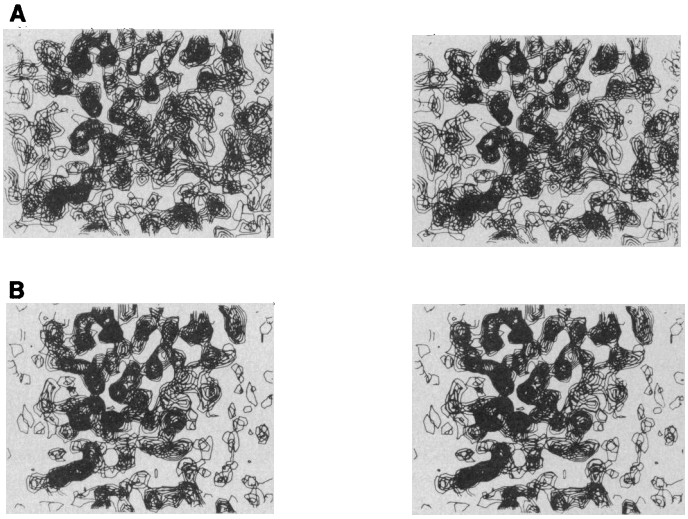
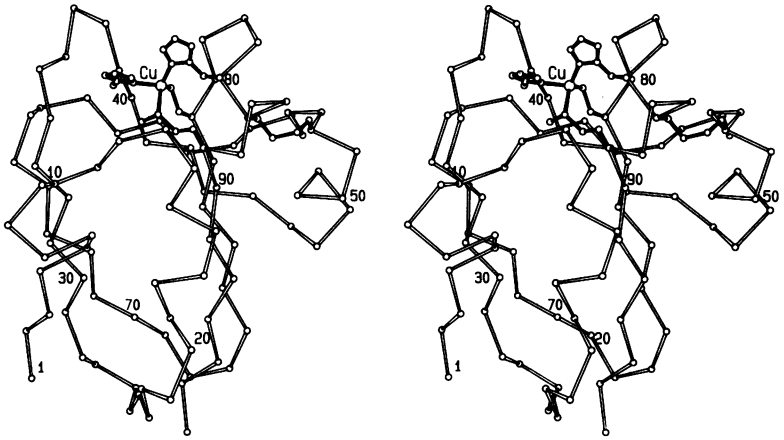


Fig. 2. Stereo sections from electron density maps at 3.0 Å resolution, (**A**) before and (**B**) after solvent flattening. Seven successive sections separated by intervals of 1.1 Å along z are shown. The contour intervals are 1σ, beginning at the 1σ level (σ being the estimated standard deviation of the electron density). To produce map B, MAD phase likelihood distributions were input in the form of Hendrickson-Lattman coefficients (*61*) to the molecular envelope and phase recombination stages of Wang's ISAS program package (*62*). The data to 3.0 Å resolution were used to generate a molecular envelope corresponding to 35 percent solvent (the theoretical solvent content being 47 percent). The map was calculated after three cycles of phase recombination.

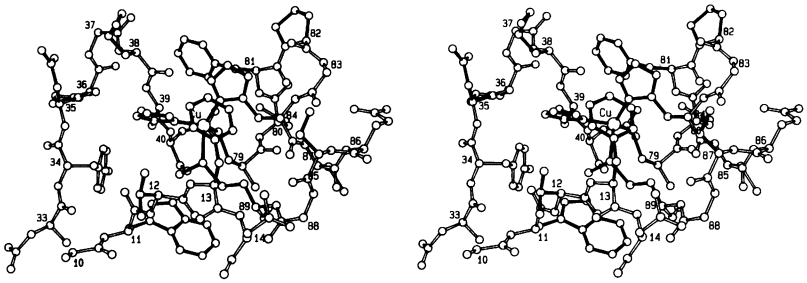
Table 2. Comparison between cucumber basic blue protein (CBP) (*3, 5, 17*) and stellacyanin (Sc) (*32–34, 39, 57*).

Parameter	CBP	Sc
Molecular weight M_r	10,100	20,000
Cu atoms per molecule	1	1
Cys residues per molecule	3	3
Met residues per molecule	2	0
pI	~10.5	9.9
E^0 (mV)	317*	184†
Electronic absorption bands		
λ_{max} (nm), ϵ_{max} (M^{-1} cm^{-1})	443, 2030	450, 942
	597, 3400‡	617, 3549§
	750, 1800	789, 341
X-band EPR parameters		
g_x	2.02¶	2.025**
g_y	2.08	2.077
g_z	2.207	2.287
A_x (cm^{-1})	0.006	0.0057
A_y (cm^{-1})	0.001	0.0029
A_z (cm^{-1})	0.0055	0.0035

*Other values: 270 mV (*30*), 340 mV (*54*). †A recent redetermination: 191 mV (*55*). ‡Other values: 593 nm (900M^{-1} cm^{-1}) (*4*), 593 nm (2900M^{-1} cm^{-1}) (*31*), and 595 nm (2000M^{-1} cm^{-1}) (*54*). §From (*39*). Other values: 604 nm (3820M^{-1} cm^{-1}) (*33*), and 605 nm (4050M^{-1} cm^{-1}) (*56*). ¶From (*3, 5*). Similar values are reported in (*30*). **From (*57*).



Cucumber basic protein



Cucumber basic protein

Fig. 3 (top). Stereoview of the CBP molecule, showing the Cα atoms of the polypeptide backbone, the side chains of the Cu-binding residues (His^{39}, Cys^{79}, His^{84}, and Met^{89}), and the cystine disulfide bridge between Cys^{52} and Cys^{85}. **Fig. 4 (bottom).** Stereoview of the Cu site in CBP. The exposed imidazole ring edge of His^{84} is surrounded by the side chains of Phe^{13}, Met^{38}, Phe^{81}, Pro^{82}, and Ser^{87}; that of His^{39} by the side chains of Thr^{12}, Phe^{13}, Asn^{35}, and Met^{38}. The side chain of Trp^{11} is seen below that of Met^{89}.

is not coordinated to the Cu atom. Between the end of strand 4 and the beginning of strand 5 are two turns of helix. The second turn of helix finishes at a third Cys residue, Cys^{52}. A disulfide bridge joins Cys^{52} to Cys^{85}. With respect to the Cu site, the disulfide bridge lies on the distal side of the double loop in the polypeptide backbone: neither of the S atoms is within bonding distance of the Cu atom.

The Cu atom is located beneath the surface at one end of the molecule (Fig. 4). The donor atoms are Nδ(His^{39}), Sγ(Cys^{79}), Nδ(His^{84}), and Sδ(Met^{89}). The coordination is distorted from a tetrahedral geometry, but further refinement is required before the bond lengths and bond angles at the Cu atom can be stated with confidence. At this stage there is no evidence for a fifth Cu-ligand bond or close Cu-polypeptide contact. Both of the His ligands have their distal (Cδ–Nε) imidazole ring edges exposed to the solvent, the immediate environment of His^{84} being more hydrophobic than that of His^{39} (Fig. 4). The accessibility surfaces of the two imidazole rings appear to be contiguous. On the side of the Cu site remote from the solvent, the side-chain methyl group of Met^{89} lies in contact with the aromatic side-chain group of Trp^{11}.

Structural comparisons with other blue copper proteins. Three blue Cu proteins—plastocyanin, azurin, and pseudoazurin—have previously been characterized crystallographically. In each the Cu atom is coordinated by the Nδ(imidazole) atoms of two His residues, the Sγ(thiolate) atom of a Cys residue, and the Sδ(thioether) atom of a Met residue (*20–25*). Refinements of the structures of plastocyanin and azurin have shown that the Cu–S(Met) bonds are abnormally long (2.9 and 3.1 Å) (*21, 24*) and that the Cu atom in azurin makes an additional close contact (3.1 Å) with a backbone O(peptide) atom (*24*). Although the Cu–S(Met) bonds are obviously weak, they seem to play a crucial role in tuning the reduction potentials of the blue Cu site (*26, 27*).

The present work shows that the distorted tetrahedral NNSS′ coordination in CBP is analogous to that found at the Cu sites of plastocyanin and azurin, lending further support to the hypothesis that the high redox potentials of the proteins (CBP, 317 mV; plastocyanins, from ~360 to 370 mV; azurins, from ~280 to 320 mV) have a common structural origin (*26, 27*). The folds of the polypeptide backbones of the three proteins are, however, distinctly different (Fig. 5). In azurin, strands 4 and 5 of the polypeptide backbone are part of the β sandwich; connecting the ends of these strands, a flap comprising about 30 residues and including three turns of helix hangs off the main body of the molecule (*23*). In plastocyanin, strand 5 is too irregular to be part of the β sandwich (*20, 21*). In CBP, the β-sandwich structure is further depleted by a bend and twist in strands 4 and 5 that place these strands at a large angle from the other strands. These observations support a suggestion by Adman that there are several subcategories of blue Cu-protein structure (*28*). From the viewpoint of crystallographic methodology, the remarkable difference between the tertiary structures of CBP and plastocyanin explains why molecular replacement methods failed for solving the CBP structure when a search model based on plastocyanin was used.

The CBP structure confirms or explains the results of several antecedent spectroscopic studies. Three of the Cu-binding residues, a Met and two His residues, were predicted from ^1H nuclear magnetic resonance (NMR) redox titrations (*29*). The fourth, a Cys, was to be expected from the intense charge-transfer band at ~600 nm (*30, 31*). The locations of the His and Met ligands in the molecule could be inferred from sequence homology with plastocyanin and azurin in the vicinity of the Cu site (*29*). The prediction of the Cu-binding Cys residue in CBP was less certain because of the presence of two additional Cys residues that have no equivalent in the other two proteins. The proximity of Trp^{11} to the Cu site is consistent with the observation that the quantum yield of a 340-nm fluorescence band typical of Trp is much higher in *apo*-CBP than in Cu(I)- and Cu(II)-CBP (*4*). The observed close contact between the side chain of Met^{89} and the aromatic group of Trp^{11} accounts for the large upfield shift of the ε-CH_3 resonance of Met^{89} in the ^1H NMR

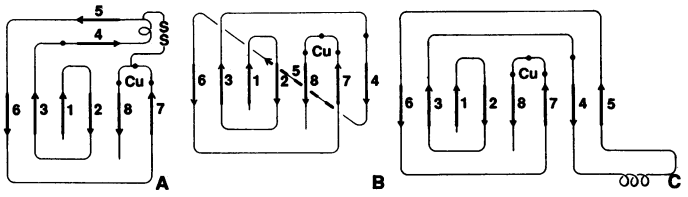


Fig. 5. Schematic representations of the polypeptide backbone folding in (**A**) CBP, (**B**) plastocyanin [adapted from (*20*)], and (**C**) azurin [adapted from (*23*)]. The solid black circles represent the Cu-binding residues. In each case the β sandwich is viewed from the exterior. [(B) and (C) are with permission from *Nature*]

spectrum of Cu(I)-CBP (*29*). In earlier work, this shift was ascribed to a close contact with the phenyl ring of Phe^{13}, it being assumed that Phe^{13} in CBP occupies a position analogous to that of Phe^{14} in plastocyanin and Phe^{15} in azurin (*29*). This incorrect assumption is another casualty of the large difference between the polypeptide backbone folds of CBP and plastocyanin: in CBP, Phe^{13} is not part of the lining of the hydrophobic pocket surrounding the Cu site, but rather is located on the surface of the molecule (Fig. 4).

Structure of a related protein, stellacyanin. The biological function of CBP remains unknown. CBP has nevertheless attracted a great deal of interest because its spectroscopic properties and primary structure closely resemble those of stellacyanin (Sc), an intriguing member of the blue Cu protein family that—being an outlier—may hold the key to an understanding of some of the properties of the blue Cu site. Stellacyanin has the lowest E^0 (184 mV) so far reported for any blue Cu protein (*32*) and—among all of the known blue Cu proteins—no methionine (*33, 34*). Thus the fourth ligand at the Cu site of Sc must be different. An explanation of the spectroscopic and redox properties of Sc awaits a determination of the structure of its Cu site. Unfortunately, Sc has yet to be crystallized, possibly because of the presence of a substantial (40 percent) and heterogeneous polysaccharide component. Strong (44 percent) homology between the primary structures of CBP (96 residues) and Sc (107 residues) has been demonstrated (*17*). Neither this homology, nor the spectroscopic similarities, provides proof of a structural relation between CBP and Sc, but merely renders plausible the hypothesis that such a relation exists (*35*).

Many of the reported properties of Sc can be readily explained if this protein does indeed have the same molecular fold as CBP. In the following discussion, several equivalences revealed by a published alignment of the amino acid sequences of CBP and Sc (*17*) assume special significance:

CBP: Trp^{11}...His^{39}...Cys^{52}...Cys^{79}...His^{84}...Cys^{85}...Met^{89}
Sc: Trp^{11}...His^{46}...Cys^{59}...Cys^{87}...His^{92}...Cys^{93}...Gln^{97}

Three of the Cu ligands in Sc can be readily identified. Coordination by the imidazole groups of two His residues is consistent with electron nuclear double resonance (ENDOR) (*36*) and indicated by NMR (*37*) evidence, and the presence of a Cys thiolate group can be inferred from the charge-transfer spectra of Co(II)-Sc as well as Cu(II)-Sc (*38, 39*). According to the above alignment with CBP, the residues involved in these interactions are His^{46}, His^{92}, and Cys^{87}. Only the number and nature of any additional Cu-ligand bonds and close contacts remain to be defined.

An important aspect of the sequence homology is that both CBP and Sc have three Cys residues—two more than are generally found in blue Cu proteins. In the case of Sc, the additional Cys residues have been implicated as Cu-binding residues, either individually (*37, 40*) or in a Cys–Cys disulfide bridge (*41*). By analogy with the disulfide-bridged residues Cys^{52} and Cys^{85} in CBP, the Cys bridge in Sc is now identified as $(Cys^{59})S–S(Cys^{93})$. The distances of the two S atoms of the disulfide bridge from the Cu site (9.4 and 10.8 Å in CBP) eliminate the hypothesis (*41*) that they contribute to the Cu coordination.

If the fourth Cu ligand in Sc is neither a thiolate nor a disulfide group, what can the structure of CBP tell us about it? In the alignment of the sequences of Sc and CBP, the residue in Sc that corresponds to Met^{89} in CBP is Gln^{97} (*17*). A Gln side chain, $–CH_2–CH_2–CONH_2$, has similar dimensions and conformational characteristics to a Met side chain, $–CH_2–CH_2–S–CH_3$. The substitution of an O(amide) for a S(thioether) donor would provide a "more Cu(II)-like" environment for the Cu atom, thus providing a rationalization for the low E^0 of Sc (184 mV) compared with other blue Cu proteins (*27*). It might also (depending on the relative Cu-ligand distances) create an increase in the ligand field at the Met position relative to plastocyanin, as required to account for the rhombic splitting in the electron paramagnetic resonance (EPR) spectrum according to a recent ligand-field analysis (*42*). If the same explanation is applied to CBP, which also exhibits rhombic splitting in the EPR spectrum but has the same combination of ligands as plastocyanin, then the implication is that the Cu–S(Met) distance in CBP is shorter than in plastocyanin. It remains to be seen from the refined structure whether this is the case.

Coordination of the Cu atom in Sc by a side-chain amide group was suggested in a recent conference report of ^1H NMR relaxation measurements on Co(II)-Sc (*43*), and would be compatible with x-ray absorption fine structure (EXAFS) analyses of Sc (*44*) and metal-substituted Sc analogues (*45*). Suggestions that the Cu atom in Sc interacts with an amide group belonging to the polypeptide backbone have been made on the basis of resonance Raman measurements on Cu(II)-Sc (*46*) and ^{113}Cd NMR measurements on Cd(II)-Sc (*47*). We note that the possibility of an additional Cu–O(peptide) contact in Sc is not eliminated by the apparent absence of such a contact in CBP (just as the presence of such a contact in azurin is not prevented by its absence in plastocyanin).

The present description of the disulfide bridge between Cys^{59} and Cys^{93} in Sc is in agreement with recent chemical evidence (*48*). An earlier observation that Sc has a $(Cys^{87})S–S(Cys^{93})$ bridge, implying that the Cu-binding thiolate group belongs to Cys^{59} (*34*), is easily explained. The protein used for the cited experiment (the determination of the amino acid sequence) was necessarily in the apo form: a preliminary computer-graphics simulation with CBP as a model has shown that once the Cu atom was removed, only modest rotations of the Cys^{87} and Cys^{93} side chains about Cβ–Cγ and a <0.2 Å relaxation of the backbone at Cys^{87} are required to permit the disulfide switch

$$(Cys^{87})SH + (Cys^{93})S–S(Cys^{59}) \rightarrow (Cys^{87})S–S(Cys^{93}) + (Cys^{59})SH$$

postulated by Engeseth *et al.* (*48*). The three Cys residues in the model *holo*-protein are so near one another as to suggest that the disulfide switch would be hindered rather than assisted by denaturation of the *apo*-protein.

The structure of CBP provides explanations for two other aspects of the chemistry of Sc. First, the quenching of a Trp fluorescence in Cu(II)-Sc and Co(II)-Sc relative to that in *apo*-Sc (*49*) resembles the fluorescence quenching of CBP (see above) so closely that it would be surprising if the locations of the homologous residue Trp^{11} were not similar in the two proteins. Second, if the surface of Sc over the imidazole ring edges of His^{46} and His^{92} is as exposed as that over His^{39} and His^{84} in CBP (Fig. 4), then the accessibility of the Cu site predicted from the high electron self-exchange rate (*50, 51*), the close adherence to Marcus theory (*52*), the occurrence of redox

reactions at electrodes in the absence of mediators (*52*), and electron spin-echo measurements (*53*), is dramatically confirmed.

The above results definitively demonstrate how MAD phasing can be used to determine protein structures. These experimental phasing results for CBP are consistent with those predicted for MAD phasing based upon one Cs atom in a protein of M_r 12,000 (*1*). For CBP, the magnitudes of the anomalous scattering terms are typically more than a factor of 2 smaller (*1*). With the data collection methodology currently available on synchrotron sources, it should be feasible to obtain phases that are sufficiently accurate for initial structure determination with K edge effects from one anomalously scattering atom in a protein with M_r up to ~25,000. The larger effects from L edges should more than double this M_r range and more accurate data collection should raise these limits even higher.

REFERENCES AND NOTES

1. J. C. Phillips *et al.*, *Acta Crystallogr.* **A33**, 445 (1977); J. C. Phillips and K. O. Hodgson, *ibid.* **A36**, 856 (1980).
2. For example, Hendrickson and co-workers are exploring the substitution of Se for S in methionine residues as a means of producing measurable anomalous dispersion effects [W. A. Hendrickson, *Trans. Am. Crystallogr. Assoc.* **21**, 11 (1985)], and the structure of a complex of streptavidin with Se-substituted biotin has been determined from MAD data recorded across the Se K-edge [J. L. Smith, A. Pähler, H. M. K. Murthy, W. A. Hendrickson, *Acta Crystallogr.* **A43**, C-10 (1987); A. Pähler *et al.*, unpublished results].
3. L. E. Vickery, thesis, University of California at Santa Barbara (1971).
4. K. A. Markossian *et al.*, *Biochim. Biophys. Acta* **359**, 47 (1974).
5. V. Ts. Aikazyan and R. M. Nalbandyan, *FEBS Lett.* **55**, 272 (1975); *ibid.* **104**, 127 (1979).
6. ———, *Biochim. Biophys. Acta* **667**, 421 (1981).
7. P. M. Colman *et al.*, *J. Mol. Biol.* **112**, 649 (1977). The space group is $P2_12_12_1$ with a = 30.88 Å, b = 46.41 Å, c = 65.57 Å, V = 93,971 Å^3, and Z = 4.
8. R. C. Lye *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 77, 5884 (1980); L. K. Templeton *et al.*, *Acta Crystallogr.* **A38**, 74 (1982).
9. R. Kahn *et al.*, *FEBS Lett.* **179**, 133 (1985).
10. Two other relevant multiple-wavelength experiments have been reported. S. Harada *et al.* [*J. Appl. Crystallogr.* **19**, 448 (1986)] determined MAD phases for *Rhodospirillum rubrum* cytochrome c′ at 6 Å resolution from diffraction data collected at three wavelengths across the Fe edge and compared them with MIR phases for the same structure. In the structure analysis of *Pseudomonas denitrificans* azurin, the known structure of *Pseudomonas aeruginosa* azurin was used to interpret an electron density map calculated from 1266 coefficients with phases derived from diffraction data at two wavelengths near the CuK edge [Z. R. Korszun, *J. Mol. Biol.* **196**, 413 (1987)].
11. The data were collected at SSRL over a period of 8 days on beam line 1-5AD with an area detector facility [R. P. Phizackerley, C. W. Cork, E. A. Merritt, *Nucl. Instrum. Methods* **A246**, 579 (1986)]. The SPEAR storage ring operated at an energy of 3.0 GeV with a ring current typically falling from 80 to 40 mA over a period of 12 hours before reinjection. The energy setting of the Si(111) two-crystal monochromator was recalibrated against the known absorption edge of a 12.5-μm metallic Cu foil at least once after each reinjection of the storage ring and was found to be stable to within 1 eV. At an x-ray energy of 9 keV, the bandwidth of the two-crystal monochromator was also ~1 eV. Harmonic rejection was achieved by detuning the first of the two monochromator crystals to yield a 10 percent reduction in x-ray flux compared with the fully tuned setting at each collection energy. The beam path from the sample to the detector was 371 mm (including 270 mm in helium). At each of the four x-ray energies, 1095 electronic images were recorded from the first crystal, and 539 images at each energy from the second crystal. Exposure time was controlled by monitoring incident x-ray flux; typical exposure times were between 30 and 60 seconds per image. Each image was recorded sequentially at the four wavelengths. The detector count rate ranged from 25 to 68 kHz, with a corresponding range of coincident event loss from 14 to 65 percent (*12*). Each electronic image resulted from exposure of the sample crystal over a 0.12- or 0.20-degree rotation in steps of either 0.005 or 0.01 degree. The sample crystals were maintained at 293 K during data collection.
12. Analytical methods for the variation of coincidence loss with net count rate proved to be unreliable at the upper end of the range. Thus the electronic images were empirically corrected as follows. On the assumption that the intensity of the solvent ring scattering was relatively constant, each recorded image was rescaled so that the average intensity per pixel within a predefined rectangle on the detector surface was brought to a constant value. The rectangle was chosen to lie within the observed solvent scattering ring, and in any given image the pixels used in the integration of Bragg peaks were excluded from the calculation of the average. This postprocessing of the recorded area detector images yielded R_{sym} values for the reduced data which were 20 to 60 percent lower than the values obtained when the images were processed with a conventional model for coincidence loss (Table 1).
13. E. A. Merritt and R. P. Phizackerley. A description of the general methodology used for accurate collection of multiple energy anomalous dispersion data is in preparation.
14. J. Karle, *Int. J. Quantum Chem. Quantum Biol. Symp.* 7, 357 (1980).
15. W. A. Hendrickson, J. L. Smith, S. Sheriff, *Methods Enzymol.* **115**, 41 (1985).
16. D. M. Blow and F. H. C. Crick, *Acta Crystallogr.* **12**, 794 (1959).
17. M. Murata *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6434 (1982).
18. T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
19. W. A. Hendrickson and J. H. Konnert, in *Computing in Crystallography*, R. Diamond, S. Ramaseshan, K. Venkatesan, Eds. (Indian Academy of Sciences, Bangalore, 1980), pp. 13.01–13.25.
20. P. M. Colman *et al.*, *Nature* **272**, 319 (1978).
21. J. M. Guss and H. C. Freeman, *J. Mol. Biol.* **169**, 521 (1983).
22. E. T. Adman and L. H. Jensen, *Isr. J. Chem.* **21**, 8 (1981).
23. G. E. Norris, B. F. Anderson, E. N. Baker, *J. Mol. Biol.* **165**, 501 (1983).
24. ———, *J. Am. Chem. Soc.* **108**, 2784 (1986).
25. K. Petratos *et al.*, *FEBS Lett.* **218**, 209 (1987).
26. H. B. Gray and B. G. Malmström, *Comm. Inorg. Chem.* **2**, 203 (1983).
27. E. W. Ainscough *et al.*, *Biochemistry* **26**, 71 (1987).
28. E. T. Adman, in *Topics in Molecular and Structural Biology. Metalloproteins*, P. M. Harrison, Ed. (Macmillan, London, 1985), vol. 1, pp. 1–42.
29. G. King *et al.*, *FEBS Lett.* **166**, 288 (1984).
30. T. Sakurai, H. Okamoto, K. Kawahara, A. Nakahara, *ibid.* **147**, 220 (1982).
31. T. Sakurai, S. Sawada, A. Nakahara, *Inorg. Chim. Acta* **123**, L21 (1986).
32. B. Reinhammar, *Biochim. Biophys. Acta* **275**, 245 (1972).
33. J. Peisach, W. G. Levine, W. E. Blumberg, *J. Biol. Chem.* **242**, 2847 (1967).
34. C. Bergman, E. Grandvik, P. O. Nyman, L. Strid, *Biochem. Biophys. Res. Commun.* 77, 1052 (1977).
35. The hypothesis that the structures of CBP and Sc are related was originally based on similarities between the ultraviolet-visible and electron paramagnetic resonance (EPR) spectra (Table 2). Both proteins have strong electronic absorptions at ~450 and ~800 nm with molar extinction coefficients $\epsilon_{600} >> \epsilon_{450} > \epsilon_{800}$; the values of the ratio of absorbances A_{280}/A_{600} are similar; the X-band EPR spectra are rhombic and have similar relations between the g and A values. These properties were regarded as sufficiently distinctive to suggest that the Cu sites of CBP and Sc are significantly different from those of plastocyanin and azurin. The resonance Raman spectra, which are sensitive probes of the Cu sites, enhanced the impression that CBP and Sc are more similar to each other than to plastocyanin (*30*).
36. J. E. Roberts, T. G. Brown, B. M. Hoffman, J. Peisach, *J. Am. Chem. Soc.* **102**, 825 (1980).
37. H. A. O. Hill and W. K. Lee, *J. Inorg. Biochem.* **11**, 101 (1979).
38. D. R. McMillin, R. A. Holwerda, H. B. Gray, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1339 (1974).
39. E. I. Solomon *et al.*, *J. Am. Chem. Soc.* **102**, 168 (1980).
40. L. Ryden and J.-O. Lundgren, *Biochimie* **61**, 781 (1979).
41. N. S. Ferris, W. H. Woodruff, D. B. Rorabacher, T. E. Jones, L. A. Ochrymowycz, *J. Am. Chem. Soc.* **100**, 5939 (1978); D. L. Tennent and D. R. McMillin, *ibid.* **101**, 2307 (1979); D. R. McMillin and M. C. Morris, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6567 (1981); O. Farver, A. Licht, I. Pecht, *Biochemistry* **26**, 7317 (1987); O. Farver and I. Pecht, in *Oxidases and Related Redox Systems*, T. E. King, H. S. Mason, M. Morrison, Eds. (Liss, New York, 1988), pp. 270–283.
42. A. A. Gewirth, S. L. Cohen, H. J. Schugar, E. I. Solomon, *Inorg. Chem.* **26**, 1133 (1987).
43. S. Dahlin *et al.*, *Recl. Trav. Chim. Pays-Bas* **106**, 419 (1987).
44. J. Peisach, L. Powers, W. E. Blumberg, B. Chance, *Biophys. J.* **38**, 277 (1982).
45. B. Reinhammar, S. Dahlin, M. C. Feiters, *Recl. Trav. Chim. Pays-Bas* **106**, 360 (1987).
46. O. Siiman, N. M. Young, P. R. Carey, *J. Am. Chem. Soc.* **98**, 744 (1976); D. F. Blair *et al.*, *ibid.* **107**, 5755 (1985).
47. H. R. Engeseth, D. R. McMillin, J. D. Otvos, *J. Biol. Chem.* **259**, 4822 (1984).
48. H. R. Engeseth, M. A. Hermodson, D. R. McMillin, *FEBS Lett.* **171**, 257 (1984).
49. L. Morpurgo, A. Finazzi-Agro, G. Rotilio, B. Mondovi, *Biochim. Biophys. Acta* **271**, 292 (1972).
50. S. Wherland, R. A. Holwerda, R. C. Rosenberg, H. B. Gray, *J. Am. Chem. Soc.* **97**, 5260 (1975).
51. S. Dahlin, B. Reinhammar, M. T. Wilson, *Biochem. J.* **218**, 609 (1984).
52. A. G. Mauk, R. A. Scott, H. B. Gray, *J. Am. Chem. Soc.* **102**, 4360 (1980).
53. W. B. Mims, J. L. Davis, J. Peisach, *Biophys. J.* **45**, 755 (1984).
54. A. M. Nersissian, M. A. Babayan, L.Kh. Sarkissian, E. G. Sarukhanian, R. M. Nalbandyan, *Biochim. Biophys. Acta* **830**, 195 (1985).
55. V. T. Taniguchi, N. Sailasuta-Scott, F. C. Anson, H. B. Gray, *Pure Appl. Chem.* **52**, 2275 (1980).
56. K. E. Falk and B. Reinhammar, *Biochim. Biophys. Acta* **285**, 84 (1972).
57. B. G. Malmström, B. Reinhammar, T. Vänngård, *ibid.* **205**, 48 (1970).
58. W. H. McMaster, N. Kerr Del Grande, J. H. Mallett, J. H. Hubbell, *Publ. UCRL-50174*, Sect. 2, Rev. 1 (University of California, Lawrence Livermore Radiation Laboratory, Livermore, CA, 1969).
59. H. Hönl, *Ann. Phys.* **18**, 625 (1933).
60. K. Eichhorn, personal communication.
61. W. A. Hendrickson, *Acta Crystallogr.* **B27**, 1472 (1971).
62. B. C. Wang, *Methods Enzymol.* **115**, 106 (1985).

63. Supported by grants from the Australian Research Grants Committee (H.C.F.) and from the National Institutes of Health (K.O.H. and R.P.P.). The synchrotron x-ray data were recorded at the Stanford Synchrotron Radiation Laboratory, which is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, and the Division of Research Resources of the National Institutes of Health. The coordinates of the structure determined have been deposited with the Brookhaven Data Bank.

26 April 1988; accepted 13 June 1988