# Metal substitution in transferrins: The crystal structure of human copper-lactoferrin at 2.1-Å resolution

**ARTICLE** *in* BIOCHEMISTRY · JUNE 1992

Impact Factor: 3.02 · Source: PubMed



**4 AUTHORS**, INCLUDING:

Clyde A Smith
Stanford University
**92** PUBLICATIONS **3,219** CITATIONS
SEE PROFILE

Bruce M. Anderson
Virginia Polytechnic Institute and State Unive…
**140** PUBLICATIONS **3,481** CITATIONS
SEE PROFILE

Edward N Baker
University of Auckland
**154** PUBLICATIONS **8,254** CITATIONS
SEE PROFILE

# Metal Substitution in Transferrins: The Crystal Structure of Human Copper-Lactoferrin at 2.1-Å Resolution†,‡

Clyde A. Smith, Bryan F. Anderson, Heather M. Baker, and Edward N. Baker*

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

*Received October 30, 1991; Revised Manuscript Received February 10, 1992*

ABSTRACT: The structural consequences of binding a metal other than iron to a transferrin have been examined by crystallographic analysis of human copper-lactoferrin, $Cu_2Lf$. X-ray diffraction data were collected from crystals of $Cu_2Lf$, using a diffractometer, to 2.6-Å resolution, and oscillation photography on a synchrotron source, to 2.1-Å resolution. The structure was refined crystallographically, by restrained least-squares methods, starting with a model based on the isomorphous diferric structure from which the ligands, metal ions, anions, and solvent molecules had been deleted. The final model, comprising 5321 protein atoms (691 residues), 2 $Cu^{2+}$ ions, 2 (bi)carbonate ions, and 308 solvent molecules has good stereochemistry (rms deviation of bond lengths from standard values of 0.018 Å) and gives a crystallographic $R$ value of 0.196 for 43 525 reflections in the range 7.5–2.1-Å resolution. The copper coordination is different in the two binding sites. In the N-terminal site, the geometry is square pyramidal, with equatorial bonds to Asp 60, Tyr 192, His 253, and a monodentate anion and a longer apical bond to Tyr 92. In the C-terminal site, the geometry is distorted octahedral, with bonds to Asp 395, Tyr 435, Tyr 528, and His 597 and an asymmetrically bidentate anion. The protein structure is the same as for the diferric protein, $Fe_2Lf$, demonstrating that the closure of the protein domains over the metal is the same in each case irrespective of whether $Fe^{3+}$ or $Cu^{2+}$ is bound and that copper could be transported and delivered to cells equally well as iron. The differences in metal coordination are achieved by small movements of the metal ion and anion within each binding site, which do *not* affect the protein structure.

Proteins of the transferrin family, which include serum transferrin, ovotransferrin, and lactoferrin, have a characteristic ability to bind, tightly but reversibly, two $Fe^{3+}$ ions, together with two $CO_3^{2-}$ anions (Brock, 1985; Aisen & Harris, 1989). Crystallographic studies on human lactoferrin (Anderson et al., 1987, 1989) and rabbit serum transferrin (Bailey et al., 1988) have demonstrated that the iron coordination is the same for both proteins, with the metal bound to four protein ligands (two Tyr, one Asp, and one His) and a bidentate $CO_3^{2-}$ anion, and with the two specific sites in each protein being essentially equivalent. Further crystallographic studies on human apolactoferrin have shown that metal binding is associated with a major conformational change in which two protein domains close over the bound metal ion (Anderson et al., 1990).

Binding studies have, however, shown that many other metal ions can be bound by transferrins in place of $Fe^{3+}$. Some, such as $Ga^{3+}$ and $Al^{3+}$, bind with an affinity close to that of $Fe^{3+}$, but others, such as lanthanide ions, bind much more weakly (Aisen & Harris, 1989; Aisen, 1989). This raises a number of interesting questions. Is the metal coordination determined solely by the demands of the protein, irrespective of the size or charge of the metal ion, as is envisaged by the idea of "rack-induced bonding" (Gray & Malmström, 1983), or does the binding site respond to the requirements of different metal ions by differences in ligation and/or differences in the metal-induced conformational change? If the closure of the domains over the metal is different for different metal ions, this could give discrimination in the delivery (or not) of these metal ions to cells. On the other hand, if the protein structure is the same for different bound metals, this implies that these metals could be delivered to cells equally well as $Fe^{3+}$, since receptor binding should be the same.

We have chosen to investigate the structural effects of different metals through crystallographic studies of metal substitution in human lactoferrin. This will enable us to address some of the above questions, as well as provide a structural basis for interpreting some of the many chemical and physical studies previously carried out on metal-substituted transferrins (Brock, 1985; Aisen & Harris, 1989). Here we report a high-resolution crystallographic analysis of human lactoferrin in which $Cu^{2+}$, a metal ion of similar size but different charge, has been substituted for $Fe^{3+}$ in both the specific metal-binding sites of the protein.

## Experimental Procedures

*Purification, Crystallization, and Data Collection.* The purification of human apolactoferrin from fresh colostrum and the preparation and crystallization of dicupric lactoferrin, $Cu_2Lf$,[1] have been described (Norris et al., 1989; Smith et al., 1991). The crystals are orthorhombic, $a$ = 155.9, $b$ = 97.0, and $c$ = 56.0 Å, space group $P2_12_12_1$, and were shown to be isomorphous with those of the diferric protein, $Fe_2Lf$ ($a$ = 156.2, $b$ = 97.4, and $c$ = 55.85 Å).

Two sets of X-ray diffraction data were collected from this complex, one to 2.6-Å resolution (24 387 reflections, 21 318 with $I > 2\sigma_I$, see Table Ia) with a CAD4 diffractometer using Cu K$\alpha$ radiation (Smith et al., 1991) and a second to 2.1-Å resolution collected at the SRS facility at Daresbury, U.K., by the oscillation method (Arndt & Wonacott, 1977) using

† Supported by the U.S. National Institutes of Health (HD-20859), the Health Research Council of New Zealand, and the New Zealand Dairy Research Institute.

‡ The coordinates have been deposited with the Brookhaven Protein Data Bank.

* To whom correspondence should be addressed.

[1] Abbreviations: $Cu_2Lf$, dicupric human lactoferrin; $Fe_2Lf$, diferric human lactoferrin; NAG, *N*-acetylglucosamine; FUC, fucose.

0006-2960/92/0431-4527$03.00/0 

Table I

| | |
|---|---|
| (a) Data Collection (Diffractometer) | |
| number of crystals used | 5 |
| radiation, wavelength (Å) | Cu $K_\alpha$, 1.5418 |
| scan type | step scan in $\omega$ |
| absorption correction | empirical (azimuthal scan) |
| total reflections measured | 30 242 |
| independent reflections | 24 387 |
| observed reflections[a] | 21 318 (87%) |
| $R_{merge}$[b] | 0.032 |
| cell dimensions[c] | |
| $a$ (Å) | 155.90 |
| $b$ (Å) | 96.98 |
| $c$ (Å) | 56.04 |
| (b) Data Collection (Oscillation Photography) | |
| wavelength of radiation (Å) | 0.88 |
| number of film packs | 57 |
| number of films per pack | 3 |
| average $R_{sca}$[d] | 0.073 |
| average $R_{sym}$[d] | 0.068 |
| $R_{merge}$[d] | 0.096 |
| total reflections measured | 228 893 |
| independent reflections (nonzero) | 44 766 |
| estimated mosaic spread (deg) | 0.30 |
| (c) Final Data Set (Diffractometer + Photographic) | |
| $R_{merge}$[e] | 0.089 |
| total unique reflections | 46 683 |
| total unique reflections to 2.1 Å | 44 715 |
| unique observed reflections[f] | 31 818 |

[a] With $I > 2\sigma_I$. [b] $R = \sum|I - \bar{I}|/\sum I$. [c] Average values for the five crystals used in data collection. [d] $R = \sum|I - \bar{I}|/\sum I$; $R_{sca}$ measures the agreement between reflections recorded on successive films of a given film pack, $R_{sym}$ measures the agreement between symmetry-related reflections on the same film, and $R_{merge}$ gives the overall agreement between intensities measured on different films. [e] $R_{merge}$ gives the agreement between intensities measured by the diffractometer and by oscillation photography. [f] With $I > 2\sigma_I$.

radiation of wavelength 0.88 Å produced by a 5-T Wiggler magnet. Two crystals, each mounted about the $a^*$ axis, were used to collect the 90° rotation range required for a full data set. This involved a total of 57 film packs of 3 films each, with the oscillation angle being varied between 1.0° and 2.0°, depending upon the spindle setting (2.0° for $\phi \leq 48°$, 1.5° for $48° \leq \phi < 72°$, and 1.0° for $72° \leq \phi \leq 90°$). The crystal to film distance was 117.9 mm.

The films were digitized with an Optronics P1000 rotating drum microdensitometer and processed by a package of programs developed by Wonacott (1980). The missetting angles, cell parameters, and X-ray beam divergence were refined against the fully recorded reflections in order to improve the estimated fractions of the partial reflections using POSTREF (Winkler et al., 1979). In this way, matching partial reflections on successive films could be added together to give an overall intensity for that reflection. Of the 228 893 reflections measured, 117 529 were fully recorded and 64 087 partially recorded, the remainder being rejected. The fully and partially recorded reflections were scaled and merged into one data set of 44 766 unique, nonzero reflections with an $R_{merge}$ of 9.6% on intensities ($R_{merge} = \sum|I - \bar{I}|/\sum\bar{I}$). X-ray data collection statistics can be found in Table Ia,b.

The final data set was derived from the merging of the film and diffractometer data (Table Ic), giving a total of 46 683 unique reflections ($R_{merge}$ = 8.9% on intensities). This data set is 93% complete to 2.1 Å [47 900 possible reflections, 44 715 measured, 31 818 (71%) with $I > 2\sigma_I$]. The number of observed data ($I > 2\sigma_I$) in the highest resolution shell (2.4–2.1 Å) is 4081 (33% of the total measured).

*Structure Solution and Refinement.* Given the isomorphism of the $Cu_2Lf$ crystals with those of $Fe_2Lf$, the starting model

Table II: Refinement Statistics[a]

| | | |
|---|---|---|
| resolution limits (Å) | 7.5–2.1 | |
| initial $R$ factor | 0.360 | |
| final $R$ factor | 0.196 | |
| number of reflections used[b] | 43 525 | |
| number of protein atoms | 5321 | |
| number of solvent molecules | 308 | |
| other ions | 2 $Cu^{2+}$, 2 $CO_3^{2-}$ | |
| number of sugar residues | 3 NAG and 2 FUC | |
| average $B$ value (protein atoms) (Å$^2$) | 47 | |
| average $B$ value (all atoms) (Å$^2$) | 48 | |
| rms shift in position (final cycle) (Å) | 0.005 | |
| rms shift in atomic $B$ (final cycle) (Å$^2$) | 0.37 | |
| geometrical constraints | target $\sigma$ | rms deviation |
| bond length (Å) | 0.020 | 0.018 |
| angle distance (Å) | 0.040 | 0.064 |
| planarity (Å) | 0.020 | 0.014 |
| chiral volumne (Å$^3$) | 0.15 | 0.194 |
| nonbonded contact (Å) | 0.50 | 0.240 |

[a] PROLSQ refinement against 2.1-Å resolution film + diffractometer data. [b] Between 7.5- and 2.1-Å resolution and with $I > 0.5\sigma_I$

for refinement was derived from that of $Fe_2Lf$, refined at 2.2-Å resolution (M. Haridas, B. F. Anderson, and E. N. Baker, manuscript in preparation). From this were removed the metal-binding ligands, the metal ions, the carbonate anions, and the solvent molecules. The initial $R$ factor for data in the resolution range 5.0–2.1 Å (40 785 reflections with $I > 0.5\sigma_I$) was 0.360. The structure was refined by restrained least-squares methods using the FFT version of the program PROLSQ (Hendrickson & Konnert, 1980). Restraints were placed on bond lengths, angle (1–3) distances, planar groups, chiral volumes, temperature factors, torsion angles, and van der Waals contacts. Target values and standard deviations were as in Table II. For the first phase of refinement (26 cycles), the geometry was tightly restrained, and an overall $B$ value of 25 Å$^2$ was applied. At this stage ($R$ = 0.301), $2F_o - F_c$ density maps were examined on an Evans and Sutherland PS 330 interactive graphics system and the metal ions and ligands added into the model using the program FRODO (Jones, 1978). The carbonate anions were similarly built into the appropriate $2F_o - F_c$ and $F_o - F_c$ density after a further 27 cycles ($R$ = 0.276).

Further refinement proceeded in a series of phases, each involving a number of cycles of least-squares refinement (typically 20–30) of atomic positions and individual isotropic temperature factors, followed by manual rebuilding of the model on the graphics system. During each least-squares refinement phase, the strategy consisted of loosening the weights on the geometrical constraints and then slowly tightening them again over 12–15 cycles. The $R$ factor generally dropped by up to 4% upon relaxing the geometry and increased again by about 1% when the geometrical constraints were subsequently tightened. Restraints were applied to the $B$ values in order to prevent excessive shifts during refinement (±3 Å$^2$) and to maintain the $B$ values of adjacent atoms to within 4 Å$^2$ of each other.

Solvent molecules were fitted into $F_o - F_c$ density (contoured at 2 times the rms deviation of the map), provided that they made hydrogen bonds of reasonable geometry to neighboring groups and were not close to any part of the structure whose conformation was in doubt. At intervals, solvent molecules with high $B$ values or too-close contacts, as well as a number of the less well defined loops and side chains on the molecular surface, were checked by excluding them from the refinement and calculating $2F_o - F_c$ and $F_o - F_c$ omit maps to establish their correct position or conformation. The carbonate anions

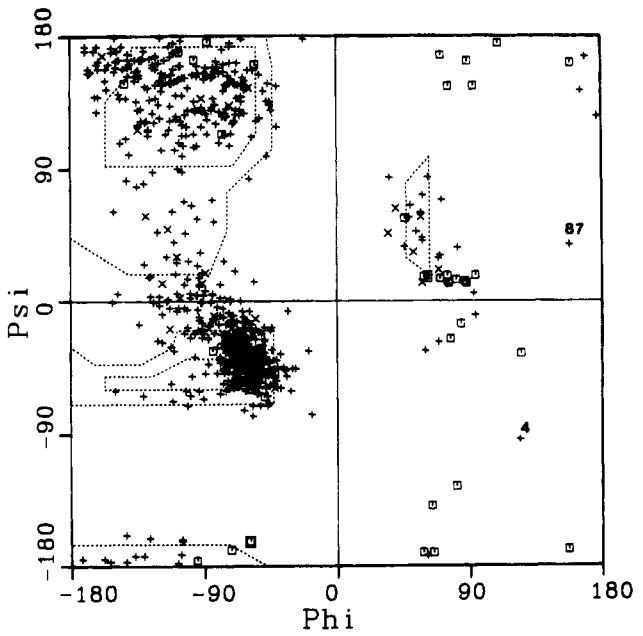


FIGURE 1: Ramachandran plot of the main-chain torsion angles for $Cu_2Lf$, showing the energetically allowed regions (Ramakrishnan & Ramachandran, 1965). Glycine residues are indicated by squares; other residues are indicated by crosses.

were also treated in this way several times during the course of the refinement.

Late in the refinement it became possible to locate some of the sugar residues of the glycan chains, from $F_o - F_c$ maps. One *N*-acetylglucosamine (NAG) moiety was found attached to each of the *N*-glycosylated residues, Asn 137 and Asn 478. On the basis of the published carbohydrate sequence (Spik et al., 1982), each glycan chain has a fucose and another NAG attached to the asparagine-linked NAG, at the O6 and O4 positions, respectively. Both of these residues could be located in the N-lobe, but only the fucose in the C-lobe. These were also constrained to standard geometry during subsequent refinement.

Special attention was paid to the copper atoms during refinement. At the resolution of this analysis, the metal density was not resolved from the ligands and it was necessary to impose bond length constraints to prevent large oscillations in position. Loose constraints ($\sigma$ = 0.05 Å) were imposed to keep the metal–ligand bond lengths close to their initial values. During refinement the shifts in the bond lengths were monitored and target values were occasionally updated to take into account the direction of these shifts. The long Cu–Tyr bond was not restrained at any stage, however.

After about 350 cycles of refinement (15 phases of refinement and model building on a Evans and Sutherland PS300 graphics system), the final value of the crystallographic *R* factor was 0.196 for 43 525 reflections between 7.5- and 2.1-Å resolution. Relevant refinement statistics are given in Table II.

## Results

*The Refined Model.* The final model consists of 5321 protein atoms (691 amino acid residues), 308 solvent molecules, assumed to be water, 62 carbohydrate atoms (three *N*-acetylglucosamine and two fucose residues), two $Cu^{2+}$ ions, and two $CO_3^{2-}$ ions. The protein and carbohydrate structures have geometry close to ideal, with rms deviations of 0.018 and 0.064 Å from standard values of bond lengths and angle (1–3) distances, respectively. The average maximum error is estimated, from a Luzzati plot (Luzzati, 1952), at 0.2–0.3 Å; the error is probably somewhat less than this in well-defined regions (the binding site for example) but substantially more in some poorly defined external parts of the molecule.

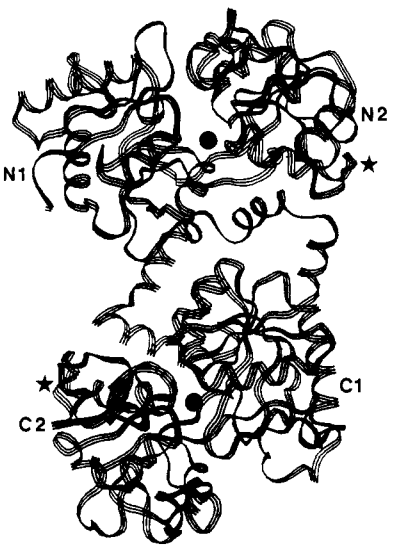


FIGURE 2: Ribbon diagram of dicupric lactoferrin, showing the bilobal nature of the $Cu_2Lf$ molecule (N-lobe above and C-lobe below) and the nomenclature of the four domains. The positions of the two $Cu^{2+}$ ions are indicated by filled circles (●). The positions at which the carbohydrate chains are attached are indicated by an asterisk (*).

A Ramachandran plot of the main-chain conformational angles, $\phi$ and $\varphi$ (Figure 1), shows that virtually all lie within the allowed and partially allowed regions of conformational space, as expected. The exceptions are mostly in the ill-defined surface regions of the structure, namely, external loops at residues 85–88, 218–224, 281–284, and 416–426 and the N-terminus, residues 1–5, where the poor quality of the electron density maps makes interpretation of the correct folding of the polypeptide chain very difficult. Plots of residue number against average main-chain or side-chain atomic temperature factors (not shown) indicate quite clearly the location of these portions of the structure, and although every effort was made to determine their correct conformation with the use of omit maps, they remain poorly defined in relation to the rest of the structure.

The 308 solvent molecules were located independently of the $Fe_2Lf$ structure (i.e., no prior knowledge of the solvent positions in $Fe_2Lf$ was assumed). Superposition of the $Cu_2Lf$ and $Fe_2Lf$ structures (based on least-squares fitting of their $C_\alpha$ positions) followed by analysis of the superimposed solvent structure showed that 128 out of 308 $Cu_2Lf$ water molecules were in conserved positions (within 1.2 Å).

The first 2–3 residues in the two glycan chains are included in the model, but their density is weak. As in the diferric lactoferrin structure, no interpretable density was visible beyond them, and the carbohydrate structure is clearly flexible and poorly ordered. In the N-lobe the first NAG moiety is covalently bonded through $C_1$ to $N_{\delta 2}$ of Asn 137 and makes favorable hydrogen bonds to $O_{\epsilon 1}$ of Glu 110 and a solvent molecule. The FUC residue is $\alpha$-linked through its $C_1$ atom to $O_6$ (NAG), and the second NAG residue is attached via a $\beta$ linkage to $O_4$ of the first NAG. The C-lobe carbohydrate chain, attached to Asn 478, does not, however, come within hydrogen-bonding distance of any part of the protein molecule.

*Polypeptide Chain Folding.* Figure 2 shows a schematic representation of the polypeptide chain conformation of the $Cu_2Lf$ molecule. The overall folding is essentially identical to that observed in the diferric lactoferrin structure, the details of which have been given previously (Anderson et al., 1989).

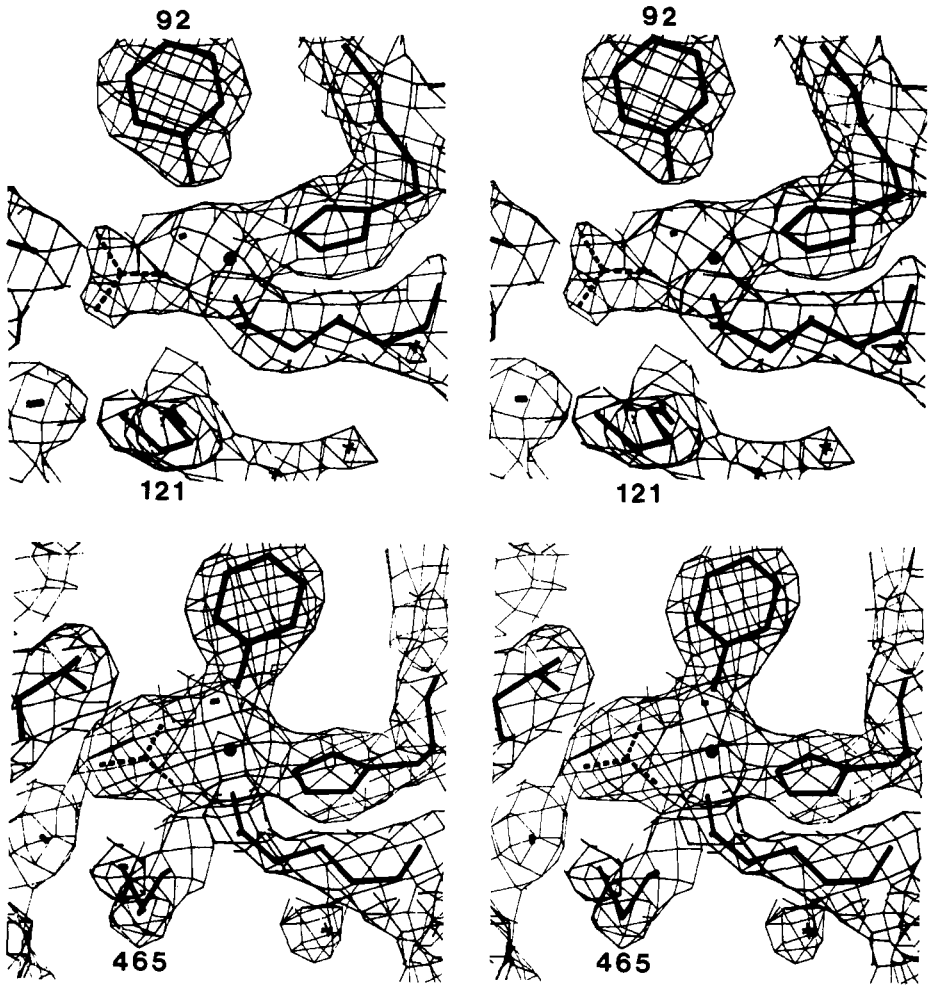


FIGURE 3: Stereoviews of the electron density in a $2F_o - F_c$ map calculated prior to the inclusion of the anions into the model ($R$ = 0.276). (a, top) The N-lobe site, showing the branched nature of the anion density and the separation between the density of Tyr 92 (top) and the copper atom. (b, bottom) The C-lobe site, with its bidentate anion. Note the closer approach of the side chain of Arg 465 to the anion, compared with the corresponding residue in the N-lobe. In each case, the anion is shown with dashed lines and the copper atom as a filled circle (●).

Briefly, the molecule is divided into two lobes, representing its N-terminal and C-terminal halves. Each lobe is further divided into two domains, N1 and N2 in the N-lobe and C1 and C2 in the C-lobe (Figure 2). The two copper sites, one in each lobe, are each at the interface between the two domains, approximately 12–15 Å from the protein surface, at the end of a deep solvent-filled cleft. As in the diferric molecule, the two lobes have very similar structures; least-squares superposition shows that 82% of the $C_\alpha$ atoms of the N-lobe can be matched with corresponding atoms in the C-lobe with an rms deviation of 1.2 Å.

*The Metal and Anion Sites.* The metal ions, metal-binding ligands, and carbonate anions were not included in the starting model to prevent bias of the $Cu_2Lf$ electron density. The first $2F_o - F_c$ electron density map clearly showed the positions of the ligands. A region of very high density (10–12 times the rms deviation of the map) could be identified in each binding cleft as the position occupied by the copper atom; two such atoms were subsequently added. Inclusion of the two copper atoms still left a large wedge of residual density at each site, between the copper atoms and a pocket formed by the side chains of Arg 121 (465 in the C-lobe) and Thr 117 (461) and the N-terminus of helix 5. As in the $Fe_2Lf$ structure (Anderson et al., 1989), this was interpreted as the site of the bound $CO_3^{2-}$ (or $HCO_3^-$) ion. In contrast to $Fe_2Lf$, however, there were clear differences between the N- and C-terminal sites, both in the shape and orientation of the anion density and in the ligand arrangement around each copper. In the C-terminal site, the anion density was triangular with one apex pointing toward Thr 461 and helix 5, consistent with an anion bound in a bidentate fashion to the metal (Figure 3b). In the N-terminal site, however, the density was bifurcated in a manner suggesting monodentate coordination (Figure 3a). Moreover, in the N-terminal site there was a clear break between the density of one ligand (Tyr 92) and the copper density, suggesting a longer $Cu-O_\eta$ (Tyr 92) bond.

In the C-terminal site, the carbonate was initially fitted in a symmetrical bidentate fashion (Cu–O distances of 2.0 Å) that gave four favorable hydrogen-bonded contacts, one each with Arg 465, Thr 461, and the main-chain peptide nitrogens of residues 467 and 468 (at the N-terminus of helix 5). The copper was placed such that distances to the side chains of Asp 395, Tyr 435, Tyr 528, and His 597 were all ≈2.0 Å. In the N-terminal site the anion was fitted so that it was bound in a monodentate fashion (Cu–O distance of 2.0 Å). This not only gave a good fit to the electron density but also resulted in the formation of five hydrogen bonds, two to Arg 121 and one each to Thr 117 and the peptide nitrogens of residues 123 and 124 at the N-terminus of helix 5. A bidentate carbonate would not adequately fit the available density or interact favorably with the anion-binding residues. The copper atom in this site was placed such that the distances to Asp 60, Tyr 192, and His 253 were all about 2.0 Å and that to Tyr 92 was ≈2.6 Å.

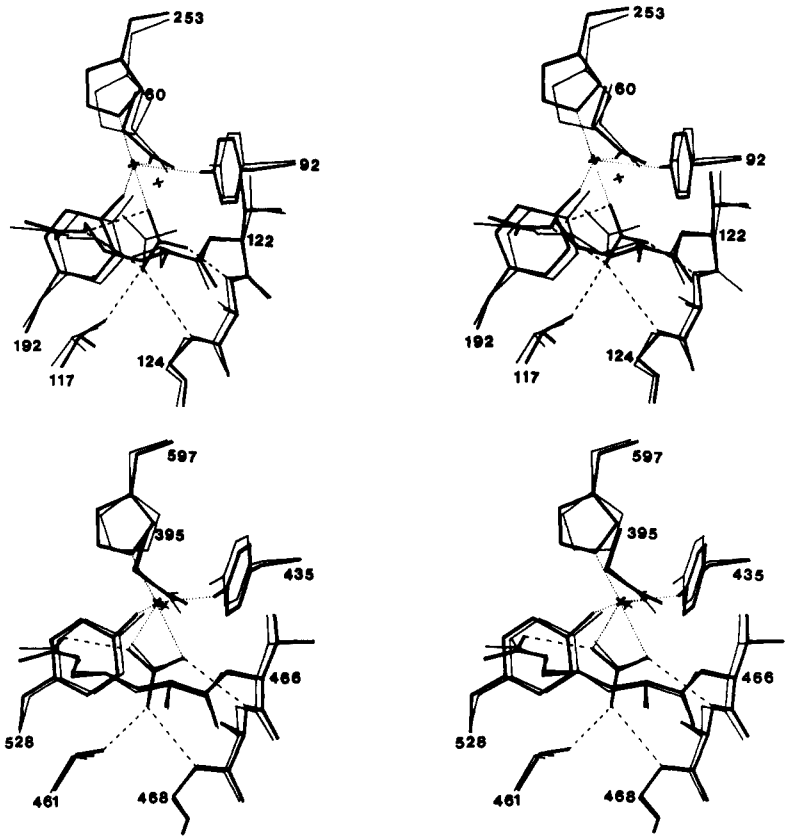


FIGURE 4: Stereoviews of (a, top) the N-lobe and (b, bottom) the C-lobe metal-binding sites of dicupric lactoferrin (heavy lines) superimposed upon diferric lactoferrin (thin lines). The hydrogen bonds associated with each anion (dashed lines) and the metal–ligand bonds (dotted lines) are indicated for the dicupric structure only.

In the early stages of the refinement, the Cu–ligand and Cu–anion distances were all restrained to target values near 2.0 Å [with the exception of the Cu–O(Tyr 92) bond, which was not restrained at all]. Small consistent shifts in the bond distances were observed, and the target values were altered in accord with these changes. It was noted that the coordination of the C-terminal anion had a tendency to become more asymmetric during the course of the refinement. Although the resolution of electron density maps was not sufficient to unequivocally distinguish between symmetric and asymmetric coordination, and the hydrogen-bonding patterns were similar for both, the asymmetric geometry resulting from the least-squares refinement was maintained. It was also observed quite early in the refinement that the Cu–O(Tyr 528) bond distance in the C-lobe was consistently shifted toward a longer value. The restraints on this distance were removed to allow the side chain to move to its most favored position. In order to achieve this without allowing the copper to move with Tyr 528, the constraints on the other bond distances were temporarily tightened to 0.03 Å. The restraints were removed from all the bond distances near the end of the refinement, and very little change was observed.

In the final model, the copper geometries in the two sites are different, the N-site being approximately square pyramidal with a long (2.8 Å) apical bond (Tyr 92), while the C-site is best described as distorted octahedral with five bonds of 2.0–2.2 Å and the sixth (Tyr 528) at about 2.4 Å. A list of relevant bond lengths in the two copper sites is given in Table III, and stereoviews of the two sites are shown in Figure 4a,b. The error in the Cu–ligand bond distances is estimated at ≈0.2 Å. It should be noted that the positions of the copper atoms, relative to the ligands, are such that in the N-site atom O(2) of the anion is trans to the His ligand, whereas in the C-site O(1) is trans to the equivalent His (Table III).

Table III: Bond lengths and Selected Bond Angles in the Copper Coordination Sphere

| bond[a] | N-lobe site (Å) | C-lobe site (Å) |
|---|---|---|
| Cu–O60 (395) | 2.0 | 2.0 |
| Cu–O92 (435) | 2.8 | 2.1 |
| Cu–O192 (528) | 2.0 | 2.4 |
| Cu–N253 (597) | 2.0 | 2.1 |
| Cu–$O_1$ (anion) | | 2.2 |
| Cu–$O_2$ (anion) | 1.9 | 2.0 |
| **angle[a]** | **N-lobe site (deg)** | **C-lobe site (deg)** |
| O60 (395)–Cu–O92 (435) | 84 | 82 |
| O60 (395)–Cu–O192 (528) | 158 | 170 |
| O60 (395)–Cu–N253 (597) | 116 | 82 |
| O60 (395)–Cu–$O_1$ (anion) | | 84 |
| O60 (395)–Cu–$O_2$ (anion) | 85 | 80 |
| O92 (435)–Cu–N253 (597) | 101 | 106 |
| N253 (597)–Cu–$O_1$ (anion) | | 166 |
| N253 (597)–Cu–$O_2$ (anion) | 159 | 113 |

[a] Donor atoms indicated are for the N-terminal site, with the equivalent atoms in the C-terminal site given in parentheses.

*Comparison of $Cu_2Lf$ with $Fe_2Lf$.* The similarities in the overall folding of the polypeptide chains for $Cu_2Lf$ and $Fe_2Lf$, and in the relative orientations of the four domains in each structure, were expected from the closely isomorphous nature of the $Cu_2Lf$ and $Fe_2Lf$ crystals. The similarity in the polypeptide fold is shown by superposition of the $C_\alpha$ atoms of $Fe_2Lf$ onto the refined $Cu_2Lf$ model. The rms deviation in $C_\alpha$ positions between the two structures is only 0.30 Å. A series of superpositions of the domains of $Cu_2Lf$ and $Fe_2Lf$, summarized in Table IV, shows little or no relative domain movement resulting from the substitution of copper(II) for iron(III). The lack of structural changes beyond the metal-

Table IV: Domain Relationships between $Cu_2Lf$ and $Fe_2Lf$[a]

| | | $Fe_2Lf$ | | | |
|---|---|---|---|---|---|
| | | N1 | N2 | C1 | C2 |
| $Cu_2Lf$ | N1 | 0.34 Å | 0.55° | 0.24° | 0.14° |
| | N2 | | 0.28 Å | 0.24° | 0.35° |
| | C1 | | | 0.29 Å | 0.10° |
| | C2 | | | | 0.27 Å |

[a] The diagonal elements give the rms deviation in $C_\alpha$ positions when equivalent domains of $Cu_2Lf$ and $Fe_2Lf$ are superimposed. The off-diagonal elements show the relative rotations between pairs of domains.

binding sites was also demonstrated by an $|F_{Fe_2Lf}| - |F_{Cu_2Lf}|$ difference electron density map calculated using phases from the refined diferric structure. This map was essentially featureless, with the only significant peaks being in the vicinity of the two metal-binding sites. A positive and negative difference peak was associated with each of the metal sites, clearly indicating a shift in the position of the copper atoms with respect to the iron. The movement for each metal ion was estimated to be in the order of 0.3–0.8 Å.

Superpositions of the separate N- and C-terminal lobes of $Cu_2Lf$ on to the corresponding lobes of $Fe_2Lf$, using only the secondary structure elements in the least-squares calculation, show the structural movements in the binding sites (Figure 4a,b). In the N-terminal site, the copper atom is 1.0 Å from the equivalent iron position, displaced in a direction away from Tyr 92 and toward the opening of the binding cleft. The anion has rotated in its plane by 20° about one oxygen (Figure 4a), and it is the combination of the cation and anion movements which allows the anion to become monodentate in $Cu_2Lf$ compared with bidentate in $Fe_2Lf$. Small movements of the Asp 60, Tyr 192, and His 253 side chains allow the donor atoms to remain within coordinating distance of the copper. The Tyr 92 side chain, however, appears to be in an almost identical position to its counterpart in $Fe_2Lf$, resulting in the observed elongation of the Cu–$O_\eta$ bond length. In the C-terminal site, the copper has shifted to a lesser extent (0.4 Å), and the four ligand side chains have moved correspondingly to keep in contact with the metal ion. The carbonate anion in this site occupies essentially the same position and orientation in both $Cu_2Lf$ and $Fe_2Lf$.

## Discussion

The copper-lactoferrin structure analysis has demonstrated that the polypeptide chain conformation when lactoferrin binds $Cu^{2+}$ is the same as when $Fe^{3+}$ is bound. That is, when $Cu^{2+}$ binds to apolactoferrin in its "open" configuration (Anderson et al., 1990), it is able to induce the same domain closure as $Fe^{3+}$. The extent of closure is the same within 0.1–0.5°. This closed structure therefore represents a particularly stable structure which is determined, at least in part, by interactions between the protein domains rather than by the metal (although the metal and anion clearly give greatly enhanced stability and rigidity, through their interactions with groups from both domains). This observation is consistent with the closed but metal-free C-lobe of apolactoferrin (Anderson et al., 1990), where again the same closed structure was seen even though no metal ions were bound.

The copper-lactoferrin structure also shows, however, that there is sufficient internal flexibility in each binding cleft to allow some change in metal coordination. In this way the protein can accommodate, to a degree, the different conformational preferences of a different metal ion. This appears to be more true of the N-terminal site than the C-terminal site. In both sites the metal ion is able to move (1.0 Å in the N-lobe and 0.4 Å in the C-lobe), but in the N-lobe site the anion also moves, resulting in a five-coordinate copper in this site, compared with six-coordinate in the C-lobe site.

It is not clear what fundamental difference between the sites gives rise to the two different copper geometries. The anion pocket formed by the side chain of Arg 121 (465) and the N-terminus of helix 5 is larger in the N-lobe than in the C-lobe, perhaps because the Arg side chain is held farther away by a hydrogen bond to Ser 191 (Gly 527 in the C-lobe). A tighter anion pocket could explain the lack of movement of the carbonate ion in the C-lobe, although we note that (i) chicken ovotransferrin has Gly instead of Ser at position 191 but apparently shows the same effect (see below) and (ii) the C-lobe site is nevertheless able to expand to accept an oxalate ion (Smith et al., 1991). Apart from the immobility of the anion, there is no obvious reason why the copper atom cannot move in the C-site as in the N-site, as three of its protein ligands appear free to move, being hydrogen bonded only to water molecules, and the fourth (Asp 395) has essentially the same position and orientation as its equivalent in the other lobe.

We believe that the difference between the two sites probably results mainly from the cumulative effect of a number of small changes in interactions associated with the extent of domain closure in each lobe. As in diferric lactoferrin, the C-lobe is more closed over the metal site than is the N-lobe, the difference corresponding to a domain rotation of about 6° (Anderson et al., 1989). We may expect that this effect may differ from one transferrin to another.

The geometry in the N-lobe is clearly a more favorable one for Cu(II) than that in the C-lobe. A square pyramidal geometry is commonly found for small-molecule copper(II) complexes, and the apical Cu–O bond length of 2.8 Å is typical for such complexes (Hathaway & Billings, 1970; Hathaway, 1987). The octahedral geometry in the C-lobe is less common for Cu(II) since Jahn–Teller distortion usually results in a tetragonal geometry (four short bonds and two longer apical bonds). However, the marked tendency of the anion to become asymmetrically bound during refinement could be taken as reflecting the movement of the copper toward a less symmetric stereochemistry (a movement restricted by the relative immobility of the carbonate).

The monodentate coordination of the anion in the N-lobe site is suggestive of $HCO_3^-$ rather than $CO_3^{2-}$. Potential hydrogen-bonding interactions in the binding site are also consistent with $HCO_3^-$; one noncoordinated oxygen, O(1), of the anion is 2.9 Å from the phenolic oxygen, $O_\eta$, of the apical ligand, Tyr 92, and the angles of this approach suggest hydrogen bonding [C–O(1)···$O_\eta$ = 101° and C$\zeta$–$O_\eta$···O(1) = 111°]. This implies that either the anion or the Tyr side chain is protonated in this site. This may also partially explain proton release titrations (Gelb & Harris, 1980) which indicate that fewer protons are released when $Cu^{2+}$ is bound, compared with $Fe^{3+}$ (two protons per $Cu^{2+}$ and three protons per $Fe^{3+}$).

The presence of both five- and six-coordinate copper centers also emphasizes the way in which the differences between the two binding sites of transferrins become more pronounced when metal ions other than $Fe^{3+}$ are bound. The electronic spectra of the N- and C-substituted monocupric complexes of chicken ovotransferrin, for example, clearly indicate this difference. The N-terminal site has an estimated visible absorption maximum at 450 nm, while that of the C-terminal site is close to 430 nm (Yamamura et al., 1984). Dicupric ovotransferrin has a maximum absorption near 440 nm, an average of the values for the two sites. The structural data obtained from dicupric lactoferrin suggests that these differences observed in ovotransferrin could be the result of different

coordination of the anion in the two sites. Further support for such a conclusion comes from recent EXAFS studies on dicupric chicken ovotransferrin (Garratt et al., 1991), where the average coordination number of the two $Cu^{2+}$ has been estimated to be 5.5. It has been suggested that this could indicate the presence of at least one monodentate (bi)carbonate anion.

The differences between copper and iron coordination, in the N-lobe at least, illustrate the potential danger in extrapolating structural conclusions from one metal to another. Even though both metals are bound to the same ligands, their coordination geometries are different (five-coordinate, square pyramidal, compared with six-coordinate, octahedral), and there is a strong possibility that the protonation state of either the anion or one Tyr ligand is different for copper compared with iron. The difference between the N- and C-lobe sites further emphasizes that spectroscopic signals are likely to be an average of the signals from two rather different metal environments.

A fundamental finding from the work, however, is that, despite the differences in detail in metal coordination, when $Cu_2$Lf is compared with $Fe_2$Lf, the protein structures are the same. We anticipate that the same protein structure should be found when other metals of similar size to $Fe^{3+}$ and $Cu^{2+}$ are bound. This should apply, for example, to $Ga^{3+}$, $Cr^{3+}$, $Mn^{3+}$, $Co^{3+}$, $Ni^{2+}$, $Zn^{2+}$, and possibly $Al^{3+}$. It is likely, however, that larger cations, such as the lanthanide ions, may not be accommodated by the same domain closure, and it is interesting to note that low-angle scattering studies of $Hf^{4+}$-substituted transferrin suggest that it has an "open" structure (Grossman et al., 1991).

These conclusions should be true also of other transferrins, serum transferrin and ovotransferrin, given their strong similarity in binding properties (Brock, 1985; Harris & Aisen, 1989) and the close correspondence of the lactoferrin and transferrin structures (Anderson et al., 1987; Bailey et al., 1988). One important implication is that transferrins carrying metal ions such as $Cu^{2+}$ should bind equally well to receptors as when $Fe^{3+}$ is bound, since the protein structure is unchanged. Presumably, a similar internalization should also occur, though whether the metal is then released will depend on the mechanism of metal release.

## Acknowledgments

We gratefully acknowledge the award of a graduate assistantship (to C.A.S.) by Massey University, and the encouragement and interest shown by Drs. Andrew Brodie and Eric Ainscough, in addition to many useful discussions. Thanks are also due to Dr. Peter Lindley and members of the Birkbeck College transferrin group for their helpful discussions and generous sharing of unpublished data.

**Registry No.** Fe, 7439-89-6; $Cu^{2+}$, 7440-50-8; $CO_3^{2-}$, 3812-32-6.

## References

Aisen, P. (1989) in *Iron Carriers and Iron Proteins* (Loehr, T., Ed.) pp 355–371, VCH Publishers, New York.

Aisen, P., & Harris, D. C. (1989) in *Iron Carriers and Iron Proteins* (Loehr, T., Ed.) pp 241–351, VCH Publishers, New York.

Anderson, B. F., Baker, H. M., Dodson, E. J., Norris, G. E., Rumball, S. V., Waters, J. M., & Baker, E. N. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1769–1773.

Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. W., & Baker, E. N. (1989) *J. Mol. Biol. 209*, 711–734.

Anderson, B. F., Baker, H. M., Norris, G. E., Rumball, S. V., & Baker, E. N. (1990) *Nature 344*, 784–787.

Arndt, U. W., & Wonacott, A. J., Eds. (1977) *The Rotation Method in Crystallography*, North-Holland, Amsterdam.

Bailey, S., Evans, R. W., Garratt, R. C., Gorinsky, B., Hasnain, S. S., Horsburgh, C., Jhoti, H., Lindley, P. F., Mydin, A., Sarra, R., & Watson, J. L. (1988) *Biochemistry 27*, 5804–5812.

Baker, E. N., Anderson, B. F., Baker, H. M., Haridas, M., Jameson, G. B., Norris, G. E., Rumball, S. V., & Smith, C. A. (1991) *Int. J. Biol. Macromol. 13*, 122–129.

Brock, J. H. (1985) in *Metalloproteins Part II, Metal Proteins with Non-Redox Roles* (Harrison, P., Ed.) pp 183–262, Macmillan, London.

Garratt, R. C., Evans, R. W., Hasnain, S. S., Lindley, P. F., & Sarra, R. (1991) *Biochem. J. 280*, 151–155.

Gelb, M. H., & Harris, D. C. (1980) *Arch. Biochem. Biophys. 200*, 93–98.

Gray, H. B., & Malmström, B. G. (1983) *Comments Inorg. Chem. 2*, 203–209.

Grossman, G., Appel, H., Hasnain, S. S., Neu, M., Schwab, F., & Thies, W.-G. (1991) *10th International Conference on Iron and Iron Proteins*, Abstract, Oxford, U.K.

Hathaway, B. J. (1987) in *Comprehensive Coordination Chemistry* (Wilkinson, G., Gillard, R. D., & McCleverty, J. A., Eds.) Vol. 5, pp 533–774, Pergamon Press, Oxford.

Hathaway, B. J., & Billings, D. E. (1970) *Coord. Chem. Rev. 5*, 143–207.

Hendrickson, W. A., & Konnert, J. H. (1980) *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., Ed.) Vol. 1, pp 43–57, Pergamon Press, Oxford.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

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

Norris, G. E., Baker, H. M., & Baker, E. N. (1989) *J. Mol. Biol. 209*, 329–331.

Ramakrishnan, C., & Ramachandran, G. N. (1965) *Biophys. J. 5*, 909–933.

Smith, C. A., Baker, H. M., & Baker, E. N. (1991) *J. Mol. Biol. 219*, 155–159.

Spik, G., Strecker, G., Fournet, B., Bouquelet, S., Montreuil, J., Dorland, L., Van Halbeek, H., & Vliegenthart, J. F. G. (1982) *Eur. J. Biochem. 121*, 413–419.

Winkler, F. K., Schutt, C. E., & Harrison, S. C. (1979) *Acta Crystallogr. A35*, 901–911.

Wonacott, A. J. (1980) *A Suite of Programs for On-Line Evaluation and Analysis of Integrated Intensities on Small Angle Rotation/Oscillation Photographs*, Cambridge, U.K.

Yamamura, T., Hagiwara, S., Nakazato, K., & Satake, K. (1984) *Biochem. Biophys. Res. Commun 119*, 298–304.