Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

# Crystallographic and physicochemical studies on metal and anion substituted human lactoferrin

A thesis presented in partial fulfilment of the requirements
for the degree of Doctor of Philosophy in Chemistry
at Massey University

Clyde Alexander Smith

1992

## Abstract

Lactoferrin, isolated from human colostrum, is an 80 kDa glycoprotein capable of binding two ferric ions concomitantly with two carbonate anions. The protein consists of a single polypeptide chain folded into two similar halves, each of which bind one iron and one carbonate. It is a member of the transferrin family of proteins, which includes serum transferrin, an iron transport protein found in the blood, ovotransferrin from avian and reptilian egg-white and melanotransferrin expressed by human melanoma cells. These proteins have important roles in the transport of iron around the body and the iron-binding function of lactoferrin has implications for the health of breast-fed infants.

Two lactoferrin complexes with copper substituted for both iron atoms and the other substituted with copper and an oxalate anion for one of the carbonate anions, have been prepared and crystallised, and the structures of both determined by X-ray crystallography to 2.1 Å resolution. Whereas in diferric lactoferrin, both anions are coordinated in a bidentate fashion to the two iron atoms, in the dicupric complex the carbonate in the N-terminal site is monodentate, resulting in a 5-coordinate copper geometry with an elongated apical ligand (≈2.7 Å). In the second complex, oxalate replaces carbonate in the C-terminal half only. Analysis of the structure indicates that the extent of closure of the lobes has an important role in determining the stereochemistry at the metal binding site.

In addition to the structural studies, the binding of a variety of other metal ions, including other transition metals, aluminium, the lanthanides and thorium have been investigated spectroscopically. All of these metal ions form 2:1 complexes with lactoferrin but with varying degrees of stability. Some binding constants for a number of the lanthanides have been estimated, firstly to provide a comparison with results obtained for serum transferrin and secondly to investigate the relationship between the size of the metal ion and the strength of binding. In the course of the investigation of the binding of the lanthanide ions, it was established that cerium interacts with lactoferrin to give a complex which slowly developes a brown colour after standing for a few days. This can be attributed to the slow oxidation of Ce(III) to Ce(IV).

These results, when compared with earlier structural analyses on lactoferrin, show that the protein is very flexible and can accommodate other metals without altering it's overall structure to any great extent. The predicted binding of other metal ions, including vanadium and the lanthanides, is described based on a combination of the crystallographic and physicochemical studies.

## Acknowledgements

I wish to thank my supervisors, Ted Baker and Andrew Brodie for their advise, assistance and encouragement throughout the course of this work. I would also like to thank Bryan Anderson for his help in data collection and with the computing, Heather Baker for her help with protein purification and some of the spectroscopic work, and the other members of the Crystallography Group (past and present). Ted and Heather must also be thanked for teaching us how to put up wallpaper.

I should also like to thank Eric Ainscough and Gill Norris for their valued contributions and helpful discussions, and Catherine Day for critically reading some of this work.

Thanks are also due to the following:

Dr. Peter Lindley for supplying the coordinates of diferric rabbit serum transferrin and the iron-saturated N-lobe half-molecule of rabbit serum transferrin. I would also like to thank him for arranging a visit to his laboratory in 1991.

Drs. Derek Knighton and Jeff Plowman at the Dairy Research Institute for allowing me to use the CD instrument, and for their generous help with running the spectra.

All the mothers who kindly donated breastmilk for use in this study. Thanks also to the staff at the Palmerston North, Wanganui, Otaki and Dannevirke hospitals for arranging for the collection of the samples.

Massey University and the Chemistry and Biochemistry Department for the award of a Graduate Assistantship and the position of Assistant Lecturer.

And last, but certainly not least my family for putting up with me and my work for the past four years. Special thanks to Claire for her support and encouragement, and for helping with the final stages of this thesis.

# Table of contents

## List of figures

## List of tables

## List of Plates

## Abbreviations

| | | | |
|---|---|---|---|
| Lf | lactoferrin | Tf | serum transferrin |
| OTf | ovotransferrin | MTf | melanotransferrin |
| bLf | bovine lactoferrin | apoLf | metal-free human lactoferrin |
| mLf | mouse lactoferrin | cOTf | chicken ovotransferrin |
| dOTf | duck ovotransferrin | hTf | human transferrin |
| bTf | bovine transferrin | rTf | rabbit transferrin |
| msTf | Tobacco Hornworm transferrin | XTf | Xenopus transferrin |
| rat Tf | rat serum transferrin | $rTf_N$ | N-lobe of rabbit transferrin |
| $Lf_N$ | isolated N-lobe of human lactoferrin (from recombinant DNA) | | |
| apoTf | metal-free human serum transferrin | $Fe_2Lf$ | diferric human lactoferrin |
| apobTf | metal-free bovine transferrin | apoOTf | metal-free hen ovotransferrin |
| $Fe_2Tf$ | diferric human transferrin | $Fe_2OTf$ | diferric hen ovotransferrin |
| $Fe_2bTf$ | diferric bovine transferrin | $Fe_2rTf$ | diferric rabbit transferrin |
| $Cu_2Tf$ | dicupric serum transferrin | $Cu_2Lf$ | dicupric lactoferrin |
| $Cu_2oxLf$ | dicupric-(oxalato-carbonato) lactoferrin | | |
| $Mn_2Lf$ | dimanganese(III) lactoferrin | $Co_2Lf$ | dicobalt(III) lactoferrin |
| $Ce(III)_2Lf$ | dicerium(III) lactoferrin | $Ce(IV)_2Lf$ | dicerium(IV) lactoferrin |
| EDTA | ethylenediaminetetraacetic acid | EDDA | ethylenediaminediacetic acid |
| NTA | nitrilotriacetic acid | β-me | β-mercaptoethanol |
| d.d. water | distilled deionised water | SDS | sodium dodecylsulphate |
| UV/Vis | electronic absorption spectroscopy | ESR | electron spin resonance |
| XANES | X-ray-absorption near-edge-structure | CD | circular dichroism |
| UV | Ultraviolet | | |
| EXAFS | extended-X-ray-absorption-fine-structure | | |
| ESEEM | electron spin echo envelope modulation | | |

# CHAPTER 1

## A review of the literature

### 1.1 Historical and functional aspects of the Transferrins§

The first, albeit inadvertent, discovery of a member of the transferrin family occurred in 1900 when Osborne and Campbell, while investigating the composition of egg white, found an unknown protein associated with ovalbumin. They proceeded to name it "conalbumin". It was partially purified by ammonium sulphate precipitation but no other properties were reported.

In 1944, while engaged in research into the production of a bacteriophage effective against various forms of dysentery, Schade and Caroline came across the protein again. They observed that raw egg-white had antimicrobial properties. It was also noted that the addition of iron as either $Fe^{2+}$ or $Fe^{3+}$ inhibited this activity and produced a brown-red colouration.

Two years later Atherton and co-workers established that this "bacteria-inhibiting, iron-binding" protein was Osborne and Campbell's conalbumin. It is now called ovotransferrin (OTf), although occasionally it is still referred to by the original name given over 90 years ago.

Between 1927 and 1937, Barkan and colleagues found that the iron in human serum was both non-dialysable and non-ultrafiltrable and therefore bound to a protein. This iron-binding protein of serum was purified by Schade and Caroline in 1946 and the physical properties investigated by Cohn in 1947 and Schade in 1949. Between them they found that the molecular mass was around 90000 Da; that it bound iron, copper and zinc; that the maximum absorption of the iron-saturated form was at about 460 nm and that two iron atoms bound per protein molecule, with the concomitant binding of two carbonate ions. The name serum transferrin (Tf) was suggested by Holmberg and Laurell in 1947.

In 1939, Sorenson and Sorenson noted a red-coloured protein in a preparation from bovine milk. The protein was isolated and characterised by Groves in 1960 and simply called the "red protein". The corresponding protein in human milk was first purified in 1958 by

---

§ Using Bezkorovainy (1980), Brock (1985), Schade (1985), Aisen & Harris (1989) and Aisen (1989) as secondary references unless otherwise stated.

Johansson as a by-product of the α-lactalbumin purification. He reported that it had a maximum absorption near 460 nm and that it bound iron very tightly. Montreuil gave it the name "lactotransferrin" to show the similarity to transferrin, although it is now generally shortened to lactoferrin (Lf). Since 1960, lactoferrin has been identified in the milk of other mammalian species and in varying quantities in other bodily fluids (Table 1.1).

Table 1.1: The occurrence of lactoferrin in various human bodily secretions [a]

| secretion | concentration (mg/L) |
|---|---|
| milk | 1550 - 5000 [b] |
| serum | 0.8 - 3.5 |
| saliva | 8 |
| tears | 1400 |
| bile | 1600 |
| seminal fluid | 550 |
| urine | 0.3 |

[a] Derived from Bezkorovainy (1980). [b] Human colostrum contains a very high level of lactoferrin which typically falls to 1500 - 2000 mg/L about 10 days postpartum.

The most recently-discovered member of the transferrin family is a membrane-bound protein isolated from the surface of melanoma cells (Rose *et al.*, 1986). Originally called p97, it is commonly known as melanotransferrin (MTf). It displays similar characteristics to the other transferrins, although it has recently been shown that it can only bind one iron atom (Baker *et al.*, 1992).

With the possible exception of melanotransferrin, whose role is not clearly established, these proteins all function as controls of free iron levels in physiological fluids. Under the conditions present in these fluids, Fe(III) is the thermodynamically stable oxidation state, but at physiological pH $Fe^{3+}$ ions readily undergo hydrolysis to form insoluble hydroxides. The solubility product of $\approx 10^{-38}$ (Sillen *et al.*, 1971) for $Fe(OH)_3$ results in a free $Fe^{3+}$ concentration at pH 7.4 of around $4 \times 10^{-18}$ mol/L. It is clear that there is a need for an iron

chelating agent to prevent precipitation of $Fe(OH)_3$ and to increase the available iron concentration to a level capable of supporting hemoglobin synthesis. The transferrins in higher animals fill this role.

The ability of the transferrins to hold iron strongly (binding constants $\approx 10^{21}$, Harris and Pecoraro, 1983) may explain their bacteriostatic properties (Brock, 1985; Schade, 1985) and their role in the suppression of iron-catalysed free-radical formation (Baldwin *et al.*, 1984). Serum transferrin also has a role in iron transport, particularly to immature red cells. Iron-loaded transferrin interacts with specific receptors on the cell surface, the transferrin-receptor complex is internalised and the iron is released in an acidic intracellular vesicle (Octave *et al.*, 1983), although more recent studies suggest that low molecular weight chelators could facilitate iron release inside the cell (Pollack *et al.*, 1991). Transportation to protoporphyrin IX (a heme precursor) via ferritin then occurs. It has been noted that diferric transferrin interacts more effectively with the receptors than apotransferrin, due to conformational differences between the two forms of the protein. The diferric form is also a more efficient iron donor than monoferric transferrin (Bezkorovainy, 1980).

Serum transferrin, along with the ovotransferrins, the lactoferrins and melanotransferrin, are generally thought of as belonging to a family of evolutionarily-related proteins. However, Huebers and Finch (1987) state that because the lactoferrins cannot supply iron to reticulocytes (Cox *et al.*, 1979) it seems inappropriate to class them as transferrins. Based upon all the spectroscopic and structural similarities between the proteins, it is clearly evident that they should be classified together and that they are likely to have originated from a common ancestor (see later in this section).

The majority of the transferrins are glycoproteins, in that they have either one or more attached carbohydrate chains (Bezkorovainy, 1980), making up between 6 and 8% of the molecular mass (Aisen & Listowsky, 1980). They are capable of binding two $Fe^{3+}$ ions, along with two (bi)carbonate anions, with the subsequent release of six protons, three per iron atom (Gelb & Harris, 1980). The ternary complexes formed all show a maximum absorption in the visible spectrum at about 460-470 nm which gives rise to their characteristic salmon-pink colour. In addition, the transferrins will also bind a wide variety of other metal

ions in oxidation states ranging from +2 to +5 and a number of anions can substitute for the carbonate and still produce ternary complexes with spectroscopic properties similar to the native complex (see Chapter 2).

## 1.2 Structural characteristics of the transferrins

The determination of the molecular mass of the transferrins was for a long time the domain of the physical biochemists, who estimated these values from measurement of the hydrodynamic parameters of the proteins. The values obtained were all around 80000 Da, although the host of different methods used (viscosity, diffusion and sedimentation studies, gel filtration, gel electrophoresis and ultracentrifugation) led to considerable variation in the estimates for the same protein. For example, values ranging from 73000 to 90000 were given for human serum transferrin (Bezkorovainy, 1980). The elucidation of the amino acid and carbohydrate sequences for a number of the transferrins has now settled the uncertainty. The most recent estimates of the molecular masses for a number of transferrins are given in Table 1.2.

Table 1.2: The molecular masses of some members of the transferrin family of proteins

| Protein | Molecular mass (Daltons) | Number of amino acid residues | Number of glycan chains |
|---|---|---|---|
| human serum transferrin | 79570 [a] | 679 [a] | 2 [b] |
| hen ovotransferrin | 77700 [c] | 686 [c] | 1 [d] |
| human lactoferrin | 80000 [e] | 691 [e] | 2 [f] |
| bovine lactoferrin | 86100 [g] | 689 [g] | 4 [g] |
| mouse lactoferrin | 75000 [i] | 688 [i] | 1 [i] |
| melanotransferrin | 97000 [j] | 719 [j] | 1 or 2 [j] |

[a] MacGillivray *et al.*, (1983). [b] Spik *et al.* (1975) [c] Williams *et al.* (1982a). [d] Graham & Williams (1975). [e] Anderson *et al.* (1989). [f] Spik *et al.* (1982) [g] Pierce *et al.* (1991). [i] Pentecost & Teng (1987). [j] Rose *et al.* (1986). This protein consists of two lobes totalling 694 residues, with a 25 residue piece attached to the C-terminus.

1.2.1 *Primary structure:*

The primary structures of a number of transferrins have been determined, including hTf (MacGillivray *et al.*, 1983), cOTf (Jeltsch & Chambon, 1982; Williams *et al.*, 1982), Lf (Metz-Boutigue *et al.*, 1984), bLf (Pierce *et al.*, 1991), MTf (Rose *et al.*, 1986), mLf (Pentecost & Teng, 1987), pTf (Baldwin & Weinstock, 1988) and the serum transferrins from *Xenopus*, XTf (Moskaitis *et al.*, 1990) and the Tobacco Hornworm, *Manduca sexta*, msTf (Bartfeld & Law, 1990). All show a remarkable similarity in their sequences, as shown in Table 1.3. Figure 1.1 shows the primary sequences of Lf, Tf and OTf aligned to emphasise the homology between the three proteins.

Table 1.3: Primary structure relationships between some members of the transferrin family of proteins

| | estimated sequence identity (%) [a] | | | |
|---|---|---|---|---|
| | Tf | OTf | MTf | mLf |
| Lf | 55 | 51 | 40 | 71 |
| Tf | — | 54 | 42 | 57 |
| OTf | — | — | 39 | 48 |
| MTf | — | — | — | 40 |

[a] The amino acid sequences were aligned with respect to the conserved metal and anion binding residues and the X-ray structures (where known), as shown in Figure 1.1.

The primary structures also indicate a two-fold internal sequence homology in the proteins, consistent with a gene-duplication event (see later). A proposal that there may be a 6-fold homology (Mazurier *et al.*, 1983), based on very tenuous sequence similarities, has since been shown to be wrong by X-ray crystallographic studies. In fact, it has been known since 1979 that diferric rabbit serum transferrin ($Fe_2$rTf) consisted of two lobes (Gorinsky *et al.*, 1979) and that a two-fold sequence homology would be a distinct possibility for all members of the transferrin family.

Human serum transferrin was found to contain 679 amino acid residues, with a 42% identity between residues 1-336 and 337-679 when aligned (MacGillivray *et al.*, 1983). In human lactoferrin, when the sequences of the two halves are aligned, there is a 40% identity,

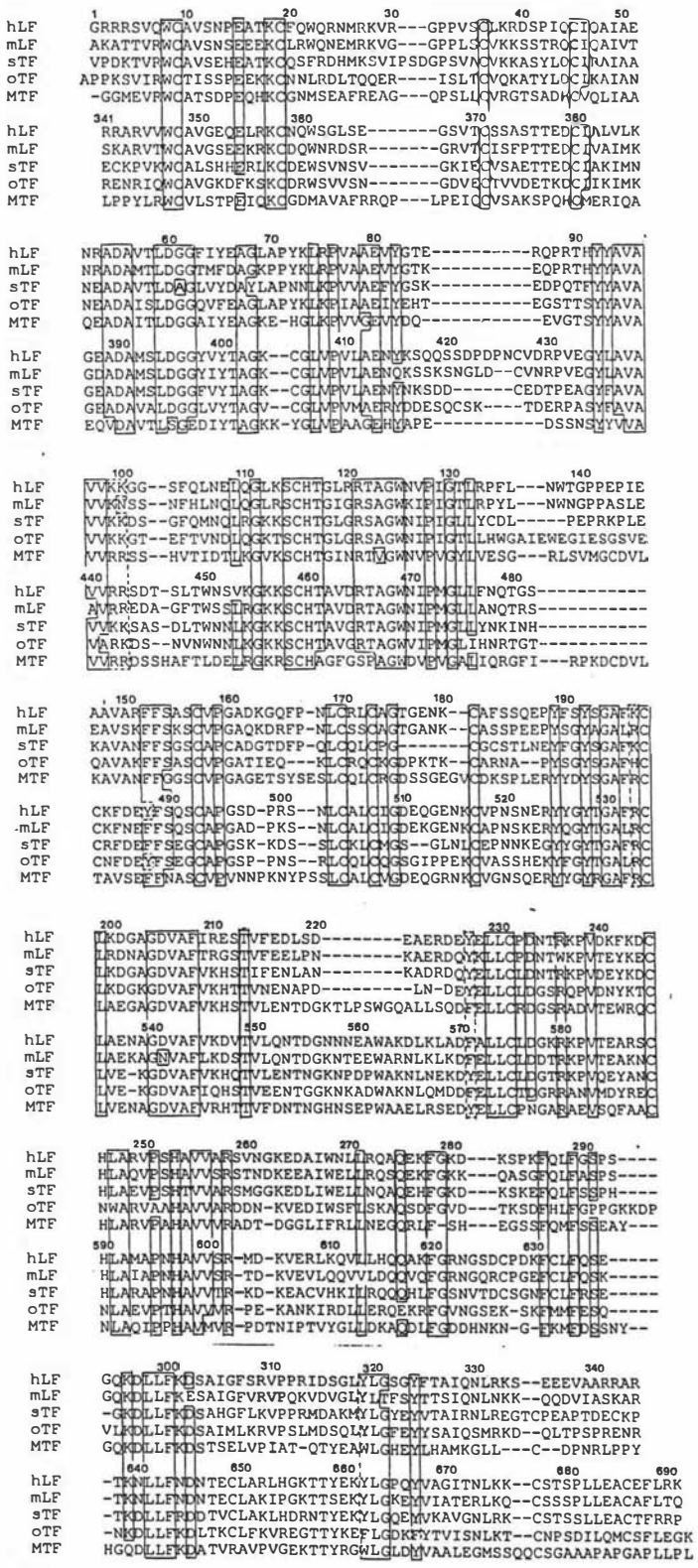


Figure 1.1: Aligned amino acid sequences of human lactoferrin, mouse lactoferrin, human serum transferrin, chicken ovotransferrin and melanotransferrin.

with a connecting peptide of 12 residues joining residues 1-333 and residues 345-691 (Anderson *et al.*, 1989). The sequence of chicken ovotransferrin (cOTf) was derived from the cDNA nucleotide sequence of transcribed cOTf mRNA (Jeltsch & Chambon, 1982) and also by conventional amino acid sequencing (Williams *et al.*, 1982). It consists of 686 amino acids with a 37% internal sequence identity. Alignment with the sequence of Lf suggests that the two halves of the molecule are joined by a 12-residue piece of polypeptide chain.

### 1.2.2 *Fragmentation studies:*

Further evidence for a bilobal structure came with the preparation and purification of iron-binding fragments from a variety of the transferrins. It was established that most members of the family could be cleaved into two fragments of roughly similar size, both of which had the ability to bind iron. Table 1.4 summarises some of the results.

Table 1.4: Selected metal binding fragments of the transferrins

| Protein | Fragment | Method | Reference |
|---|---|---|---|
| cOTf | N | chymotrypsin - 30% Fe-loaded | Williams (1974) |
| cOTf | C | subtilisin - 30% Fe-loaded | Williams (1975) |
| cOTf | N & C | immobilized subtilisin - $Fe_2$OTf | Keung *et al.* (1982) |
| bTf | N & C | trypsin - $Fe_2$bTf | Brock *et al.* (1976) |
| hTf | N | proteolysis - Tf + FeNTA | Evans & Williams (1978) |
| hTf | C | proteolysis - Tf + Fe(III)-citrate, $FeCl_3$, Fe(II)-ascorbate | Evans & Williams (1978) |
| rTf | N | subtilisin - $Fe_2$rTf | Heaphy & Williams (1982) |
| rTf | N | immobilised subtilisin | Sarra & Lindley (1986) |
| rTf | C | chymotrypsin - 20% saturated with FeNTA | Heaphy & Williams (1982) |
| Lf | N & C | chymotrypsin or trypsin - 30% saturated with FeNTA | Bluard-Deconinck *et al* (1978) |

Perhaps the most interesting fragments reported so far are the 18000 Da fragment isolated from duck ovotransferrin (dOTf) (Evans & Madden, 1984) and the 18500 Da fragment from human lactoferrin (Legrand *et al.*, 1984). When diferric dOTf was treated with trypsin, three major iron-binding fragments were isolated. Two were found to have molecular masses of 36000 - 40000 Da, while the third fragment was found to have a molecular mass of 15000 - 17000 Da. It had a $\lambda_{max}$ at 425 nm and was yellow in colour. The heavier-weight fragments bound one Fe atom/molecule whereas the light fragment only contained 0.57 Fe atoms/molecule. Lindley and colleagues have been successful in crystallising this latter fragment (Jhoti *et al.*, 1988). It is suggested that this fragment corresponds to domain 2 of the N-terminal lobe of dOTf and therefore cannot provide all of the iron-binding ligands available to the metal in the intact protein. The vacant coordination sites around the iron atom could possibly be occupied by solvent (P.F. Lindley, personal communication).

The 18500 Da fragment from human lactoferrin also corresponds to a portion of the N-terminal half and it was reported that this quarter-molecule contained carbohydrate and bound up to 0.4 $Fe^{3+}$/mol of fragment (Legrand *et al.*, 1984). The iron-saturated fragment was reported to show a shift in visible absorption maximum from 465 to 430 nm, although no spectra were presented.

### 1.2.3 *Ligands at the metal-binding sites:*

The wealth of spectroscopic work available on the transferrins led to the proposal of a number of models for the metal-binding sites. Tyrosine was the first residue to be implicated as a metal binding ligand, based on a variety of complementary techniques. It was shown that three $H^+$ were released per ferric ion bound to ovotransferrin and that groups with a $pK_a$ > 10 were involved (Warner & Weber, 1953), indicative of either tyrosine or lysine. It was also observed that the binding of $Fe^{3+}$ and $Cu^{2+}$ to ovotransferrin resulted in an increased absorption at 245 and 290 nm (Warner & Weber, 1953), due to the deprotonation of tyrosine and a resulting intense $\pi \rightarrow \pi^*$ electronic transition (Wetlaufer, 1962). These observations, along with the red colour of diferric ovotransferrin and the yellow colour of dicupric

ovotransferrin, characteristic of metal complexes of phenolate ligands (Patch & Carrano, 1981), led to the proposal that either two or three tyrosines were bound to each metal ion.

Raymond and co-workers followed metal binding to transferrin and a phenolate model compound, EHPG, using ultraviolet difference spectroscopy. From the extinction coefficients of the metal-transferrin and the corresponding metal-EHPG complexes, they calculated that only two tyrosines are involved in the binding of the majority of metal ions studied (Pecoraro *et al.*, 1981). There was also some suggestion that larger metal ions such as the lanthanides could possibly bind to three tyrosines.

The other amino acid firmly established as a ligand in metal-binding was histidine. ESR studies initially suggested that two histidine residues per metal ion were involved when iron and copper are bound to transferrin, ovotransferrin and lactoferrin (Aasa *et al.*, 1963, Windle *et al.*, 1963), although Zweier and Aisen (1977) found that specifically-bound $Cu^{2+}$ cations interacted with only one imidazole nitrogen ligand each. The latter conclusions were based upon the observed triplet superhyperfine splitting of the low field $^{65}Cu$ hyperfine line. The authors pointed out that the interaction of only one imidazole nitrogen could be due to the propensity of $Cu^{2+}$ ions to adopt a 4-coordinate (possibly 5-coordinate) geometry and that in the case of $Fe^{3+}$, which occupies a rhombic environment in the transferrins (Aasa, 1970; Aisen *et al.*, 1974), a second histidyl residue could be involved.

No other amino acids were implicated in the first coordination sphere, although fluorescence (Lehrer, 1969) and circular dichroism (CD) spectra (Tomimatsu and Donovan, 1976) initially suggested that tryptophan residues could be involved, either directly bound to the metal or near enough to the binding sites to be substantially perturbed by conformational changes in the protein upon metal binding. This latter possibility was favoured by Brown & Parry (1974) based on their CD study of bovine lactoferrin (bLf) and was later reiterated by Meares and co-workers in their fluorescence studies on transferrin (Meares & Ledbetter, 1977, O'Hara *et al.*, 1981). In addition, it was established that arginine, while not involved in direct binding to the metal, was probably associated with the binding of the synergistic anion (Rogers *et al.*, 1978, Zweier *et al.*, 1981).

### 1.2.4 *Structural studies:*

Early work on the 3-dimensional structures of the transferrins resulted in preliminary crystallographic data being available for a number of the proteins, including rTf (Al-Hilal *et al.*, 1976), Lf (Baker & Rumball, 1977), hTf (DeLucas *et al.*, 1978) and cOTf (Abola *et al.*, 1982). The work of Gorinsky *et al.* (1979) produced a 6 Å resolution electron density map of $Fe_2$rTf which showed a well defined volume of density and indicated that the molecule was composed of two lobes as predicted. They also suggested that each lobe might contain one iron binding site, but the resolution of the map was not sufficient to locate the metal atoms. A low resolution Fourier map of cOTf also indicated a bilobal structure for this protein (Abola *et al.*, 1982).

Table 1.5: Preliminary crystallographic data for a number of transferrins and associated fragments

| protein [a] | crystal type | space group | Preliminary cell dimensions (Å) |
|---|---|---|---|
| $Fe_2$Tf | tetragonal | $P4_12_12$ or $P4_32_12$ | a=b=126.9, c=145.2 |
| $Fe_2$Lf | orthorhombic | $P2_12_12_1$ | a=155.5, b=97.3, c=55.5 |
| $Fe_2$Tf | orthorhombic | $P2_12_12_1$ | a=78, b=94, c=112 |
| $Fe_2$cOTf | monoclinic | $P2_1$ | a=89:5, b=60.5, c=73.4, β=95.9° |
| $FerTf_N$ | trigonal | $P3_121$ or $P3_221$ | a=b=66.8, c=137.5 |
| $Fe_2$bLf | orthorhombic | $P2_12_12_1$ | a=138.4, b=87.1, c=73.6 |
| $Fe_2$dOTf | orthorhombic | $P2_12_12_1$ | a=49.6, b=85.6, c=178.7 |
| apodOTf | orthorhombic | $P2_12_12_1$ | a=77.6, b=98.8, c=127.0 |
| $FedOTf_N$ | orthorhombic | $P2_12_12_1$ | a=47.0, b=90.2, c=76.2 |
| FedOTf (18000 Da) | trigonal | $P3_1$ | a=b=41.3, c=81.2 |
| apoLf (native) | orthorhombic | $P2_12_12_1$ | a=222.0, b=115.6, c=77.8 |
| apoLf (deglyc. A) | orthorhombic | $P2_12_12_1$ | a=152.1, b=94.6, c=55.8 |
| apoLf (deglyc. B) | tetragonal | $I_4$ | a=b=189.4, c=55.1 |
| $FeLf_N$ | monoclinic | C2 | a=132.7, b=58.2,c=58.3, β=114.8° |
| $apoFeLf_N$ | tetragonal | $P4_12_12$ | a=58.4, b=58.4, c=217.18 |

[a] Unless otherwise stated, all of these proteins are iron-saturated

A number of other preliminary reports of structural studies on transferrins followed, including the N-terminal fragment of rTf (Sarra & Lindley, 1986), bovine lactoferrin (bLf)

(Norris *et al.*, 1986), diferric and apo dOTf (Rawas *et al.*, 1988), an 18000 Da fragment of dOTf (Jhoti *et al.*, 1988), both native and deglycosylated human apolactoferrin (apoLf) (Norris *et al.*, 1989), the iron saturated N-terminal fragment of dOTf ($FedOTf_N$) (Mikami & Hirose, 1990) and both the iron-saturated ($FeLf_N$) and apo ($apoFeLf_N$) forms of the recombinant N-lobe half of human lactoferrin (C. L. Day, personal communication). Preliminary crystallographic data are given in Table 1.5.

### 1.2.5 *The overall structure of lactoferrin:*

The answers to questions ranging from the nature of the ligands at the binding sites, the significance of the synergistic anion, the conformation of the protein in the vicinity of the binding sites and the inter-site distance came with the solution by X-ray crystallography of the structure of diferric human lactoferrin, $Fe_2Lf$ (Anderson *et al.*, 1987) at 3.2 Å resolution, followed closely by that of diferric rabbit serum transferrin, $Fe_2rTf$ (Bailey *et al.*, 1988) at 3.3 Å resolution. The structure of $Fe_2Lf$ has since been refined at 2.8 Å resolution (Anderson *et al.*, 1989) and more recently the structure of the N-terminal half of rTf has been reported to 2.3 Å resolution (Sarra *et al.*, 1990).

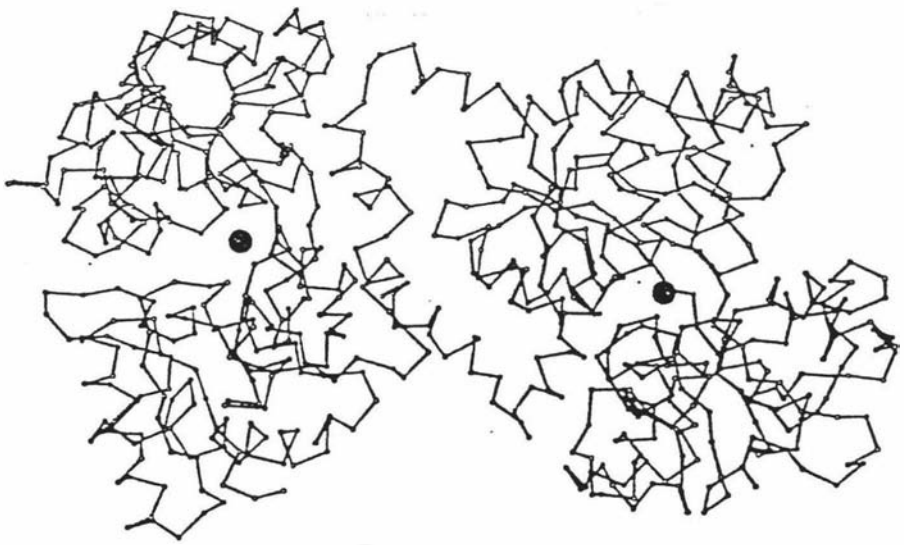


Figure 1.2: $C_\alpha$ plot of diferric lactoferrin, showing the bilobal nature of the molecule (N-lobe on the left and C-lobe on the right). The positions of the two $Fe^{3+}$ ions are indicated by filled circles . The helix connecting the two domains is labelled H.

Both Lf and rTf consist of two virtually identical parts, corresponding to the N- and C-terminal halves of the polypeptide chain, with about 40% sequence identity between the two lobes. Each of the lobes is formed by the folding of a continuous piece of protein chain and contains one $Fe^{3+}$ binding site. The overall structure of the diferric lactoferrin molecule is shown in the form of a $C_\alpha$ trace of the polypeptide chain (Figure 1.2).

Each lobe is further subdivided into two equal-sized yet dissimilar domains, with the metal-binding sites located in the inter-domain cleft. Figure 1.3 gives a schematic representation of the arrangement of the four domains in $Fe_2$Lf (Baker *et al.*, 1987). The two metal-binding sites appear to be deeply buried in the protein, although still accessible to the solvent via water molecules bound in the inter-domain cleft (Anderson *et al.*, 1989).

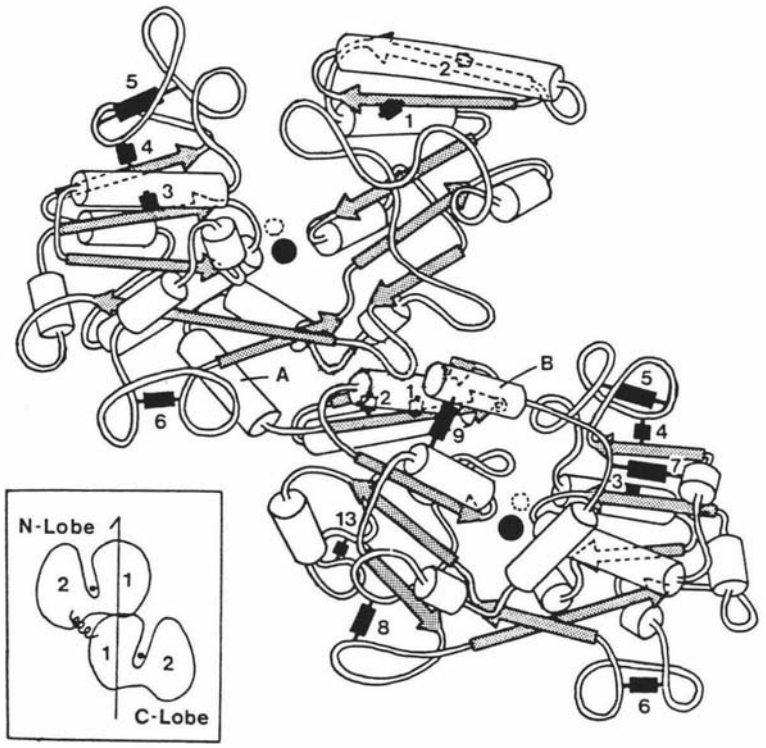



Figure 1.3: Schematic representation of the polypeptide chain in diferric lactoferrin, indicating the major secondary structure elements (helices as cylinders and β-sheets as arrows), with the disulphide bridges shown in black. Helix A connects the two lobes and helix B makes up the C-terminus of the polypeptide chain . The inset shows the relative orientations of the two lobes, which are related by a two-fold screw axis in the direction of the arrow (from Baker *et al.*, 1987).

The N-terminal lobe consists of residues 1 to 333, while the C-terminal half is made up of residues 345 to 691 (Anderson *et al.*, 1989). A 3-turn α-helix (334-344) serves as the link

between the two lobes. However, the two domains in each lobe do not comprise a single piece of the chain. Part of domain 1 is folded first (1-90; 345-433)§ , then the chain crosses behind the binding site to form the whole of domain 2 (91-251; 434-595). It then crosses back to form the remainder of domain 1 (252-320; 596-675). An α-helix heading across the interface towards domain 2 essentially completes the folding. A piece of helix similar to the one linking the two lobes serves as the C-terminus of the protein chain (676-691).

As Baker and colleagues have pointed out (Anderson *et al.*, 1987, Baker *et al.*, 1987), each lobe of lactoferrin shows a remarkable resemblance to a some of the members of a group of bacterial periplasmic binding proteins which are involved in the binding and transport of anions such as sulphate (Pflugrath & Quiocho, 1988) and phosphate (Luecke & Quiocho, 1990), sugars such as L-arabinose (Gilliland & Quiocho, 1981) and L-galactose (Vyas *et al.*, 1983) and the amino acids leucine, isoleucine and valine (Saper & Quiocho, 1983). These single-lobed proteins typically comprise about 300 - 350 residues and have a two-domain structure almost identical to the domain structure of lactoferrin.

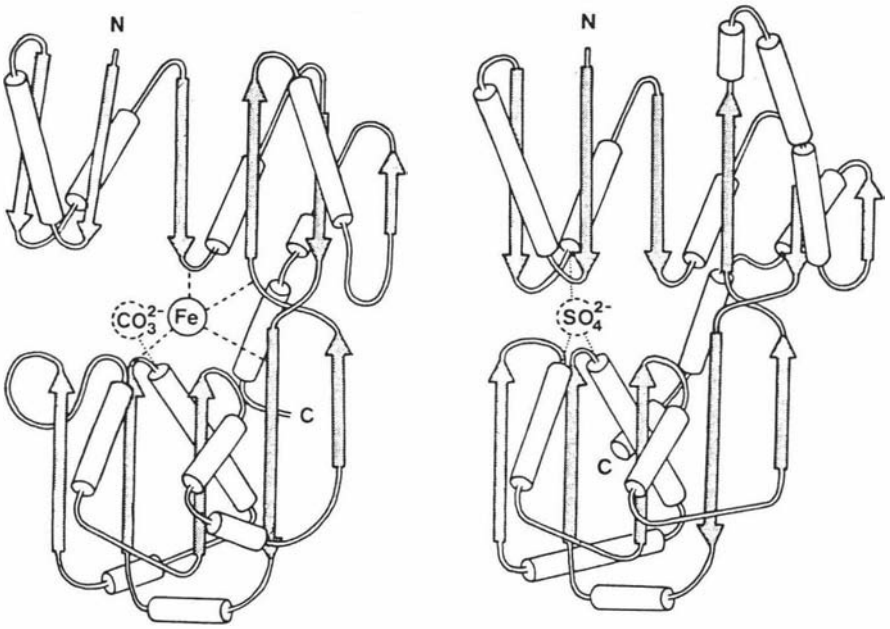



Figure 1.4: Comparison of the folding patterns of (a) the N-lobe of diferric lactoferrin and (b) the sulphate binding protein (SBP) from *Salmonella typhimurium*. N designates the N-terminus and C the C-terminus (Baker *et al.*, 1987).

§ The residue ranges given refer to the N- and C-lobes respectively.

Although there is little amino acid similarity between lactoferrin and these bacterial proteins (the sulphate-binding protein, SBP, for example), virtually all of their secondary structure elements and even the strands linking the two domains match (Figure 1.4), and this has led to the proposal (Baker *et al.*, 1987) that their structural similarities could be due to either divergent evolution from a ion-binding ancestor or the dictates of folding and functional requirements (section 1.2.6).

*Metal and anion binding ligands:* Baker and co-workers have identified the protein ligands responsible for the binding of the metal in Lf as two tyrosines, one histidine and one aspartic acid per iron. The presence of the tyrosines and the histidine were as predicted (section 1.2.3) and although the ligation of an aspartic acid was unexpected it is consistent with lactoferrins strong preference for binding high-spin Fe(III) (Anderson *et al.*, 1987). In addition to the four protein-derived ligands, each iron in $Fe_2Lf$ is also coordinated to the synergistic carbonate anion which binds in a bidentate fashion (the N-lobe site is shown in Figure 1.5). This is in turn hydrogen-bonded to the protein: namely Arg121 (465 in the C-lobe), Thr117 (461) and the N-terminus of an α-helix (helix 5, Anderson *et al.*, 1989).

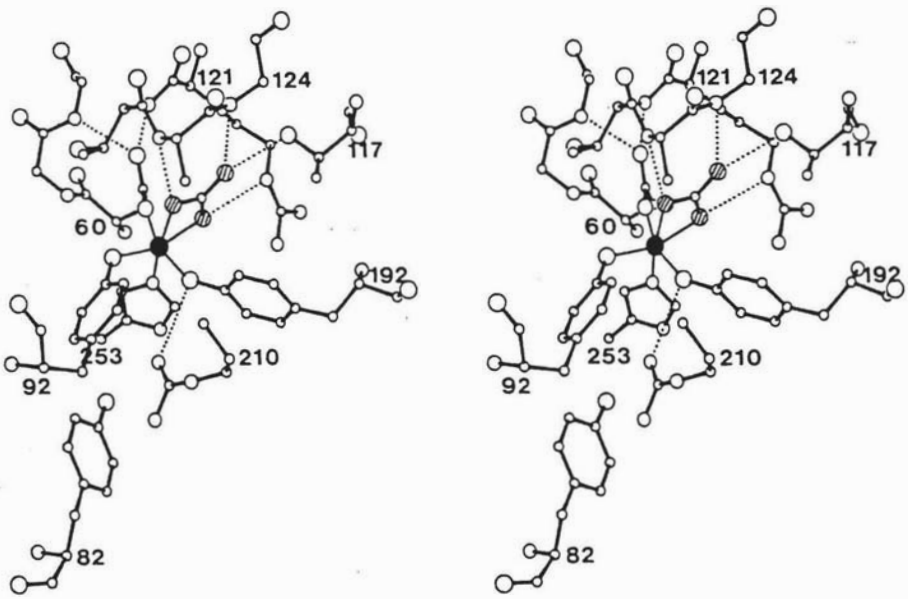



Figure 1.5: A stereo view of the N-lobe iron and carbonate binding sites in $Fe_2Lf$. The iron atom is represented by ● and the carbonate anion by ⊘. Note the bidentate coordination of the anion. The coordination geometry in the C-lobe binding site is almost identical (from Anderson *et al.*, 1989).

It is interesting to note that the residues that comprise the anion site come from a small region of the polypeptide chain, located on one domain (N2 and C2) only. Thus the anion site is likely to be unchanged by any domain movements (proposed by Baker *et al.* (1987) as the likely accompaniment of metal release). On the other hand, the 4 protein ligands are widely spaced along the chain, Asp60 (395) located on domain 1, Tyr192 (528) on domain 2 and the other two on each of the two backbone strands running behind the binding sites. It was suggested that this arrangement would be advantageous if a large scale conformational change took place upon metal binding or release (Anderson *et al.*, 1989).

Comparison of the primary structures of the transferrins show that in hTf, rTf, cOTf, bLf, mLf and Lf, the eight protein ligands are all conserved, as are the anion binding residues, one arginine and one threonine and the residues at the N-terminus of α-helix 5 (in each lobe) (Table 1.6). However in two transferrins, namely melanotransferrin (MTf) and the serum transferrin from *Manduca sexta* (msTf), there are significant differences in the residues involved in metal binding which could possibly explain why it has been observed that both of these proteins only bind one iron atom (Baker *et al.*, 1992; Bartfeld & Law, 1990).

Table 1.6: Conserved residues possibly involved in metal binding in some of the transferrins (based on the eight known ligands in diferric lactoferrin)

| | Lf | hTf | cOTf | rTf [a] | bLf | mLf | MTf | msTf |
|---|---|---|---|---|---|---|---|---|
| | Asp60 | Asp | Asp | Asp | Asp | Asp | Asp | Asp |
| | Tyr92 | Tyr | Tyr | Tyr | Tyr | Tyr | Tyr | Phe |
| | Tyr192 | Tyr | Tyr | Tyr | Tyr | Tyr | Tyr | Tyr |
| N-lobe [b] | His253 | His | His | His | His | His | His | Gln |
| | Arg121 | Arg | Arg | Arg | Arg | Arg | Arg | Arg |
| | Thr117 | Thr | Thr | Thr | Thr | Thr | Thr | Thr |
| | | | | | | | | |
| | Asp395 | Asp | Asp | Asp | Asp | Asp | Ser | Asp |
| | Tyr435 | Tyr | Tyr | Tyr | Tyr | Tyr | Tyr | Asn |
| | Tyr528 | Tyr | Tyr | Tyr | Tyr | Tyr | Tyr | Asp |
| C-lobe [c] | His597 | His | His | His | His | His | His | Arg |
| | Arg465 | Arg | Arg | Arg | Arg | Arg | Ser | Thr |
| | Thr461 | Thr | Thr | Thr | Thr | Thr | Ala | Ser |

[a] Personal communication from Dr. P.F. Lindley. [b] Residues making up the N-terminus of helix 5 are essentially conserved between Lf, hTf, cOTf, RTf, bLf and mLf. [c] Helix 5 is again conserved in the first six proteins, although Asn470 is changed to a Val in cOTf.

In MTf, the four N-lobe ligands are conserved, as are Arg121 and Thr117 (lactoferrin numbering) The only difference is at the N-terminus of helix 5, where Ala123 is changed to Val. This change is unlikely to lead to any differences in hydrogen bonding involving the anion. The C-lobe has four significant sequence changes with respect to the other transferrins (except msTf), namely Asp395 to Ser; Thr461 to Ala; Arg465 to Ser and Thr466 to Pro. It is conceivable that this latter difference could greatly affect the stability of the closed structure, as in all the other transferrins there is are interdomain hydrogen bonds between Asp395 and Thr466. The two changes to the anion-binding residues (461 and 465), coupled with the change from Asp395 to Ser may preclude binding of iron and carbonate in this site.

The *Manduca sexta* transferrin also has an essentially intact N-lobe binding site, although Tyr92 has been replaced by Phe. It has been suggested that the next residue in the sequence, a tyrosine, could take the place of Tyr92 in metal binding (Bartfeld & Law, 1990). In addition, Trp125 and Asn126 at the N-terminus of helix 5 are changed to Tyr and Lys respectively. In the C-lobe, only Asp395 is conserved. The N-terminus of helix 5 has also seen extensive changes, namely residues Ala467, Trp469 and Asn470 being changed to Ser, Leu and His respectively. In both lobes, these changes at the helix N-terminus should have little or no effect on anionbinding, although it is very unlikely that the C-lobe metal binding site would be able to coordinate a metal ion.

It is interesting to note that water has not been identified as being present in the first coordination sphere of either $Fe_2$Lf or $Fe_2$rTf. EXAFS studies on diferric ovotransferrin (Hasnain *et al.*, 1987) showed that the coordination number of the $Fe^{3+}$ ion decreased from six to five when the protein was freeze-dried and it was initially suggested that there was a water molecule bound to the iron which was removed upon freeze-drying. Rather than being indicative of a water molecule coordinated to the iron as originally suggested, these observations could be explained in terms of second sphere water molecules hydrogen-bonded to one or more of the metal ligands being lost on freeze-drying, thus altering the coordination geometry of the metal (Sarra *et al.*, 1990). More recent EXAFS results showed that in the case of copper coordination to cOTf, the average coordination number for the two copper

atoms is 4.5 (Garratt *et al.*, 1992), which was reduced upon freeze-drying. This was interpreted in terms of the coordination of three protein ligands the presence of both a monodentate and bidentate carbonate, and the subsequent change of the bidentate anion to monodentate following the removal of a second sphere water. The presence of second sphere water molecules could also account for the proton exchange from water molecules in the vicinity of the metal, as observed by Gaber *et al.* (1970) and the ESR line broadening in copper-substituted transferrin (Zweier & Aisen, 1977). In fact, electron density peaks probably representing water molecules have been observed in both iron binding sites of $Fe_2$Lf within 2.8 Å of the $N_{\delta 1}$ atoms of the His253 (597) and the $O_{\eta}$ of Tyr92 (435). Tyr528 is also hydrogen bonded to a solvent molecule although the corresponding residue in the N-lobe interacts with Arg210 rather than a water molecule (Anderson *et al.*, 1989).

*The inter-domain cleft:* One other interesting aspect of the structure of the transferrins is the presence of a large cavity in the inter-domain cleft (see Figures 1.2 and 1.3). It is described as being roughly spherical and about 10 Å in diameter and, in $Fe_2$Lf, is filled with a network of water molecules (Anderson *et al.*, 1989). The cavity is lined by the N-termini of six helices, three from each domain, thus creating a substantial positive charge inside the cavity, due to the partial positive charges associated with the N-termini of the helices (Hol *et al.*, 1978). These helices, in conjunction with the positively-charged Arg121 (465) sidechain, would serve to attract anions into the cleft to be subsequently bound in the anion pocket to the arginine.

Several basic amino acid sidechains also project into this cleft (Anderson *et al.*, 1989) which, along with the large size of the interdomain cavity, could account for the binding of larger anions such as oxalate, malonate, NTA and thioglycolate, known synergistic anions for the transferrins (Schlabach & Bates, 1975, Ainscough *et al.*, 1983). It has even been observed that the dye xylenol orange can bind synergistically to transferrin (Harris & Gelb, 1980) and most of this large molecule could possibly be accommodated in this inter-domain cavity.

*Apolactoferrin:* Evidence for the large metal-induced conformational change, indicated by early physical studies on transferrin which showed that the protein becomes more spherical or compact as it binds iron, came with the determination of the structure of apolactoferrin (apoLf) by molecular replacement (Anderson *et al.*, 1990). It was found that in this structure the N-terminal lobe was wide open with no bound metal or anion. In order to superimpose the N-lobe of apoLf on to $Fe_2$Lf, a rotation of about 53° of the N2 domain (Figure 1.6) was required. This type of movement was anticipated by Baker *et al.* (1987), although perhaps not to this degree. The conformational change involved is the bending of two β-strands in a hinge-type movement facilitated by the pivotting of an α-helix (5 in the numbering of Anderson *et al.*, 1989) in the N2 domain on the helix (11) which runs across the domain interface (Figure 1.6).

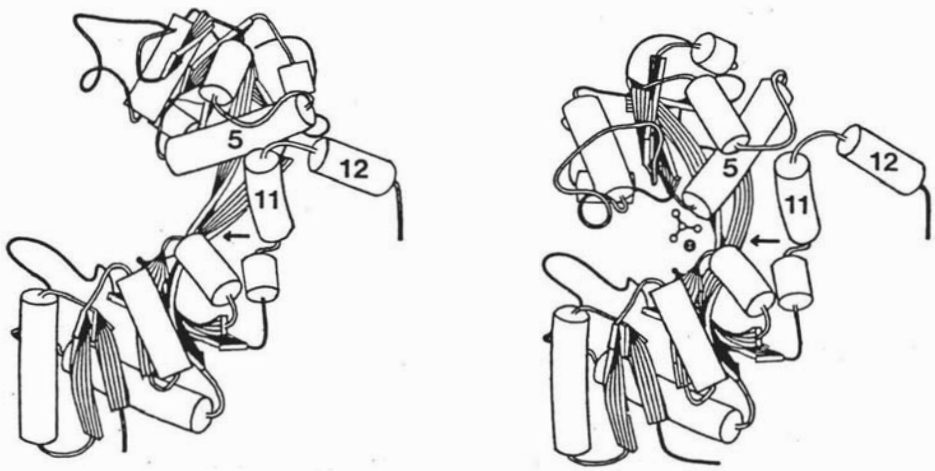



Figure 1.6: Comparison of the secondary structure elements of the N-lobes of (a) diferric lactoferrin and (b) apolactoferrin (from Anderson *et al.*, 1990). The helices involved in the pivotting action are labelled 5 and 11. The arrow points to the β-strands which bend upon domain opening.

A surprising feature of the structure was that the C-lobe remained closed even though there was no evidence for bound iron or bicarbonate. When superimposed on the C-lobe of $Fe_2$Lf, the r.m.s. deviation is only about 0.46 Å. There does not appear to be anything preventing the lobe from opening, although a disulphide bridge which is present only in the C-lobe may create an extra constraint (Anderson *et al.*, 1990). A detailed analysis of the metal and anion binding sites in the C-lobe of apoLf have led to the conclusion that there is a chloride ion in the position normally occupied by the carbonate anion (Baker *et al.*, 1991).

The $Cl^-$ is probably due to the presence of NaCl in the mother liquor from which the crystals were grown. It appears to interact only with the residues in the anion-binding site, which are localised on a small portion of the chain, and hence should not influence the ability of the lobe to open.

Recent studies on one of the bacterial proteins discussed earlier, have shown that, in its unliganded form, the leucine/isoleucine/valine-binding protein LIVBP (Sack *et al.*, 1989) adopts a structure similar to the open N-lobe of apoLf. Other members of this group of proteins show other structural forms related by domain movements: an open structure with the substrate bound to one domain (liganded LIVBP, Sack *et al.*, 1989) and a closed form with the substrate bound in the cleft (SBP, Pflugrath & Quiocho, 1988). The C-lobe of apoLf could be a representative of the remaining form, closed and unliganded (Figure 1.7).

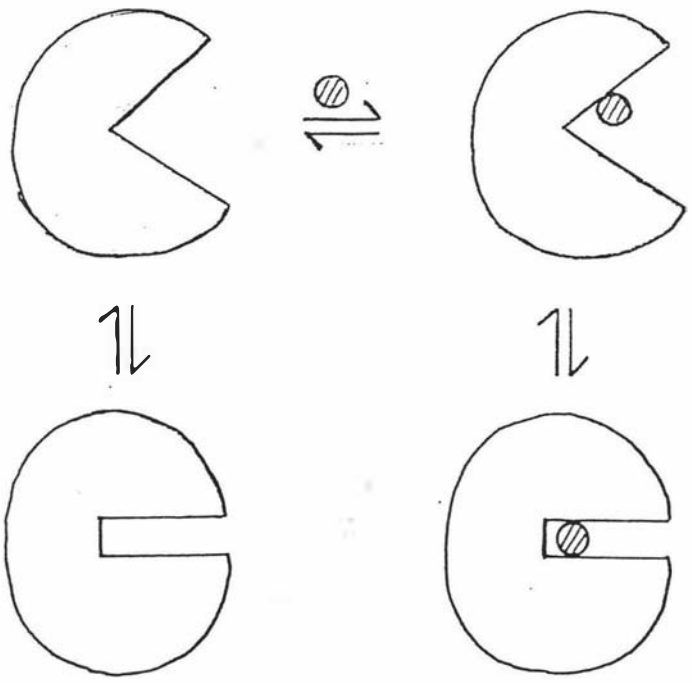


Figure 1.7: The four equilibrium related conformations of a single lobe of lactoferrin.

However, the question still remains; why is the C-lobe closed? One possible explanation could be that an equilibrium between open and closed forms exists. There are four possible

structures which could result from such an equilibrium:

1) Both lobes shut,
2) The N-lobe open and the C-lobe shut,
3) The C-lobe open and the N-lobe shut and
4) Both lobes open

Either the observed structure is selected by crystal-packing (out of the four possible structures in Figure 1.7) or the closed structure is the most stable for the C-lobe in conditions of low ionic strength. An analysis of the structures of the apo protein and the diferric form have not revealed any evidence to suggest that both halves of the molecule cannot be open at the same time (Anderson *et al.*, 1990).

### 1.2.6 *The evolution of the transferrins:*

The evidence for a gene duplication event early on in the evolution of the transferrin family of proteins has been steadily growing since it was first proposed by Greene & Feeney (1968). They determined that hTf, rTf and cOTf were single-chain proteins using ultracentrifugation of the denatured proteins in 8 M urea and 6 M guanidine hydrochloride. This, along with the fact that the transferrins could bind a maximum of two iron atoms per protein molecule, led them to suggest that the proteins consisted of two identical halves.

The amino acid sequences and the studies of the iron-binding fragments obtained from various transferrins by proteolysis supported the gene duplication theory by showing that two virtually identical halves existed, which were able to bind iron quite independently of each other. Greene & Feeney (1968) suggested that the ancestral protein from which the transferrins had evolved was probably of molecular mass 38000 Da, containing one iron-binding site. Williams (1982), backed by the amino acid sequences of cOTf and hTf, suggested that the precursor protein resembled the N-terminal lobe of cOTf, in terms of the numbers and positions of the disulphide bridges, with a molecular mass of about 40000 Da.

Some recent discoveries of transferrin-like iron-binding proteins in some invertebrate species such as the tarantula spider, *Dugesiella hentzi*, the crab, *Cancer magister*, the tobacco hornworm, *Manduca sexta*, the amphibia *Xenopus laevis* and *Xenopus borealis* and, perhaps most notably, the ascidian, *Pyura stolonifera*, may lend support to this theory also (see below). At the same time, however, their existence raises a number of questions.

The tarantula "transferrin" was found to be similar in most respects to mammalian and avian transferrins, with particular regard to spectroscopic characteristics and the ability to bind two $Fe^{3+}$ ions per protein molecule (Lee *et al.*, 1978). The molecular mass, reported as 100000 Da, confuses the issue somewhat, as this is significantly greater than the 80000 Da of the more well-known transferrins. The higher molecular mass could be the result of incorrect measurement, as was the case with hagfish transferrin, initially reported as having a mass of 44000 Da (Palmour & Sutton, 1971) but later amended to 75000 - 80000 Da by Aisen *et al.* (1972). If the molecular mass is correct, then the extra 20000 Da could be due either to insertions of extra amino acids or to the presence of extra glycosylation sites on the surface of the protein. The other possibility of course, is that this is not a member of the transferrin family, in that it has a totally different primary and tertiary structure. However, for nature to have arrived, along completely separate evolutionary lines, at an iron transport protein with identical iron-binding properties to the transferrins but with a totally different size and structure, seems unlikely.

The iron-binding protein found in *Cancer magister* has iron-binding and spectroscopic properties closely resembling serum transferrin. Two iron atoms bind per protein molecule, resulting in a visible absorption maximum near 465 nm (Huebers *et al.*, 1982). Difference spectral results show a maximum at 300 nm, similar to that found in other transferrins (Tan & Woodworth, 1969, Teuwissen *et al.*, 1972, Pecoraro *et al.*, 1981) and these results suggest the presence of at least one tyrosine residue in each binding site. Bicarbonate is required for optimal colour formation, again similar to the mammalian and avian transferrins. Surprisingly this transferrin-like protein has a molecular mass near 150000 Da and consists of a single polypeptide chain.

The invertebrate *Pyura stolonifera*, which originated about 200 million years ago (Huebers & Finch, 1987) also has an iron-binding protein analogous to transferrin. It is a protein of molecular mass 41000 Da containing one iron-binding site, with a $\lambda_{max}$ = 430 nm (Martin *et al.*, 1984).

Neither the crab nor the Pyura "transferrins" lose their iron in the presence of apoTf, although it was observed that these proteins had the ability to associate with mammalian reticulocyte receptors for transferrin, where the iron is released and incorporated into heme (Huebers *et al.*, 1982, Martin *et al.*, 1984). The tarantula "transferrin" does not donate iron to rat reticulocytes (Lee *et al.*, 1978). These results suggest that delivery of iron via a circulating protein has ancient origins and is not solely confined to vertebrates. Little seems to have altered concerning iron transport, even though the proteins responsible have undergone some major changes. Huebers *et al.* (1982) suggest that all iron transport proteins which can provide a selective distribution of iron to tissues should be classified as transferrins and that even though the tarantula and the crab proteins may have a different molecular mass, they can still be thought of as members of this family.

The observed interaction of the crab and Pyura "transferrins" with mammalian reticulocytes is somewhat surprising, in that lactoferrin is unable to bind to specific transferrin receptors. Baker *et al.* (1991) suggested that relative lobe orientations could be a factor in determining whether or not receptor binding takes place. Superposition of the C-lobe of $Fe_2$Lf on to the N-lobe requires a 180° rotation, compared with 167° in the case of $Fe_2$rTf. This implies that Lf is essentially "linear" while rTf is slightly "bent". If both lobes of the incoming iron-loaded transferrin are required for interaction with the receptor, then any differences in the lobe orientations would inhibit specific binding. Clearly there must also be specific residues on the molecular surface acting as a receptor recognition site. The differences in the ability of Lf and Tf to bind to Tf receptors could be a combination of the differences in structure and in the location of these receptor sites. The crab transferrin could possess a similar "patch" of residues to enable it to bind to rat reticulocytes (Huebers *et al.*, 1982) although the arthropods, being a much older species than the mammals and birds, would not normally be expected to have an iron binding protein similar to the mammalian and

avian transferrins. However, this does not preclude the possibility that there was an ancient single-sited iron-binding protein from which both the invertebrate and vertebrate transferrins evolved.

If the ancestor of the modern-day transferrins was a protein of molecular mass ≈40000 Da capable of binding one iron atom, then the protein isolated from *Pyura stolonifera* may well be representative of the ancient transferrin precursors. Williams (1982) suggested that the ancestral protein may have been anchored in a cell membrane where it served as an iron receptor. This supposition was based on the observation that injection of the isolated half-molecule fragments of cOTf into mice resulted in the rapid excretion of the proteins in the urine (Williams *et al.*, 1982a). The 40000 Da fragments were too small to exist free in the blood stream. The recent discovery of melanotransferrin (Rose *et al.*, 1986), a double-sited transferrin with over 40% sequence homology with human lactoferrin (Baker *et al.*, 1987), lends support to this proposal. Melanotransferrin could be considered to represent a stage in the evolution of the transferrins where gene duplication had taken place but separation from the cell membrane had not yet occurred (Williams, 1982).

The similarity in folding between Lf and several bacterial binding proteins (Section 1.2.5), extends also to the nature and location of their anion sites (Baker *et al.*, 1987; 1990). This has led to the suggestion that perhaps these two proteins diverged from a common single-sited anion-binding precursor, and that since then the transferrins have evolved an additional metal-binding site and undergone gene duplication (Baker *et al.*, 1990). If this were the case, then this would place the divergence event well before the proposed gene duplication which Williams (1982) puts at approximately 500 million ago. The ancestor to *P. stolonifera* must have appeared between the addition of the metal site and the gene duplication, branching off on an evolutionary path quite distinct from the vertebrates and some of the other invertebrates.

The existence of the two-sited tarantula and hornworm transferrins suggests that gene duplication may have taken place prior to the appearance of these species, meaning that the figure of 500 million years given by Williams (1982) may well be an underestimate. The other possibility is that the arthropod and insect transferrins evolved along a different path

which may account for the low sequence identity (26 - 28%) between msTf and the other transferrins (Bartfeld & Law, 1990). Figure 1.8 shows one possible course of events based on the evidence outlined here and that given by Williams (1982).

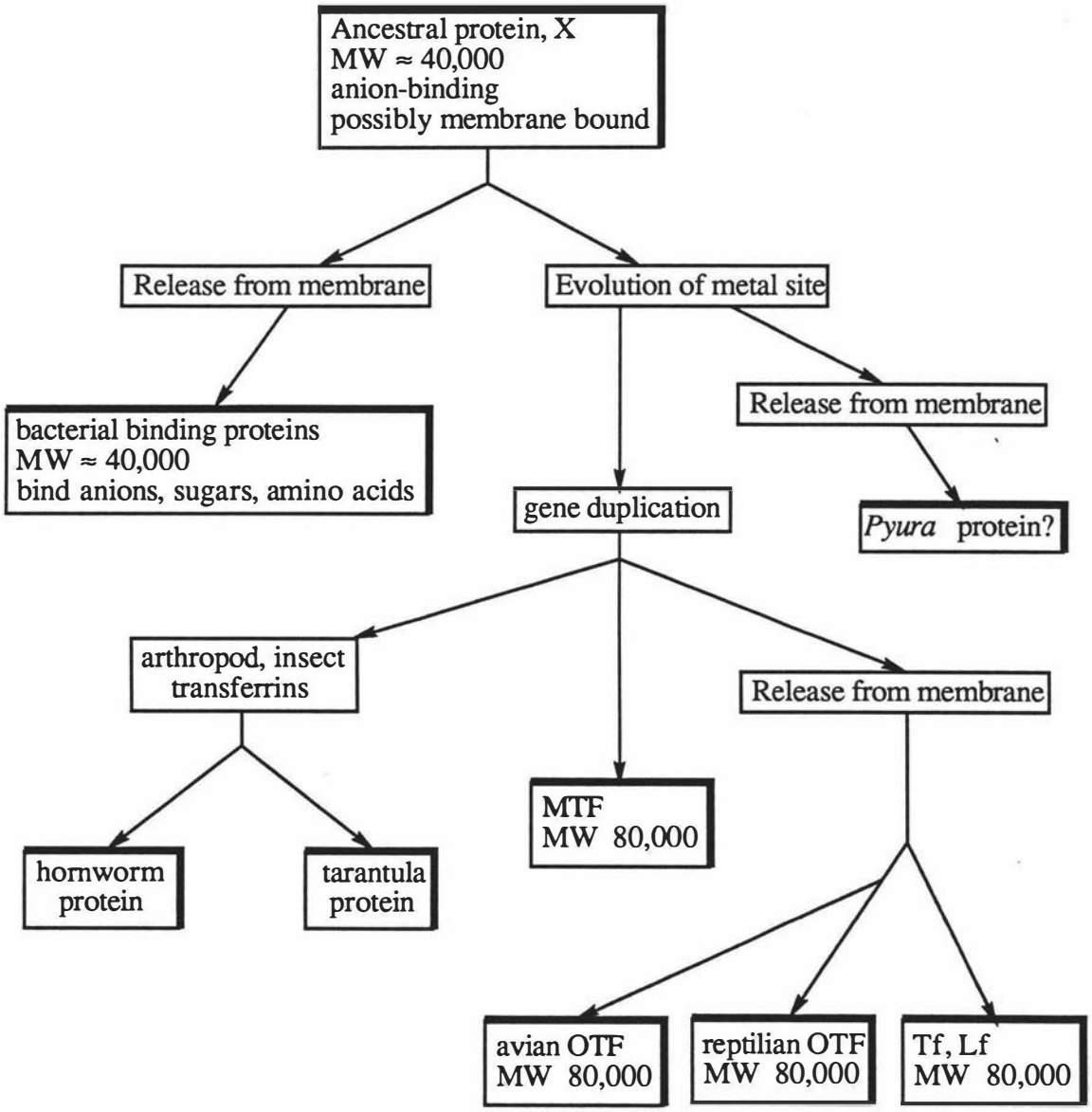



Figure 1.8: A possible evolutionary scheme for the modern-day transferrins, using a single-sited anion binding protein as the common precursor and incorporating the ideas of Williams (1982) and Baker *et al.* (1990).

Several observations resulting from the structural analysis of human lactoferrin and rabbit serum transferrin support the theory that the transferrin precursor may have been an anion-binding protein which evolved a metal binding site prior to gene duplication. These are outlined below.

(i) The majority of metalloproteins and metalloenzymes do not require the concommitant binding of an anion in order to coordinate metal ions. Azurin, for example, forms a simple binary complex with copper (Baker, 1988) or cadmium (E. N. Baker, personal communication), and the zinc in carboxypeptidase A (Rees *et al.*, 1983) and carbonic anhydrase (Eriksson *et al.*, 1988) is coordinated to protein residues and solvent molecules only. If the transferrin ancestor was assumed to be an anion-binding protein, then the requirement for a suitable anion in the binding of metal ions could be simply an artifact of evolution.

(ii) The close similarity in secondary structure and folding of the sulphate binding protein (SBP) (Pflugrath & Quiocho, 1988), despite a less than 10% sequence homology between SBP and lactoferrin (Baker *et al.*, 1987). The sulphate binding site lies at the end of a helix which is analogous to helix 5 (121 - 136; 465 - 481) in lactoferrin, although the hydrogen bonding interactions involving the sulphate are somewhat different than those observed in the transferrins. As in the transferrins, the interdomain cleft in SBP is lined with the N-termini of six helices, inducing a substantial positive charge. This positive charge is important for the initial attraction of the sulphate and the subsequent binding of the anion.

(iii) The anion binds to the apo protein before the metal ion (Kojima & Bates, 1981) and the resultant binary complex is relatively stable in solution. It is not known whether the binding of the anion alone can induce the conformational change required to close the domain, although both liganded "open" and liganded "closed" forms of the leucine-isoleucine-valine binding protein, LIVBP (Sack *et al.*, 1989) and SBP (Pflugrath & Quiocho, 1985; 1988) respectively.

(iv) The anion binding site in the transferrins (those residues which interact via hydrogen bonds with the carbonate) all come from a small region of the polypeptide chain (117 - 124 in the N-lobe and 461 - 468 in the C-lobe, lactoferrin numbering), while the ligands which bind to the iron come from four widely spaced parts of the chain. It has been suggested that this could have advantages for metal binding and release associated with a large conformational change (Anderson *et al.*, 1989), although it is unusual, in that the binding sites or active sites in most metalloproteins and metalloenzymes tend to have the metal binding ligands

concentrated in one or two regions of the polypeptide chain. In the 2Fe-2S ferredoxin from *Anabaena* 7120 (Rypniewski *et al.*, 1991), for example, the sulphur atoms from cysteines 41, 46, 49 and 79 are involved in the iron-sulphur cluster, while the two inorganic sulphurs hydrogen bond with mainchain nitrogen atoms of residues 40, 42, 44, 45 and 46.

This is also observed in the binding site of azurin (Shepard *et al.*, 1990), where the copper atom is coordinated to Cys112, His117 and His46 in a trigonal plane with weak axial interaction with Met121 and Gly45, and in a number of calcium binding proteins, where the calcium binding loop consists of 12 sequential residues (Herzberg & James, 1985).

It is possible that it was the anion binding site which was initially present in the transferrin ancestor and that subsequent closure of the protein brought several residues with oxygen and nitrogen donor atoms into close proximity. There may have been some initial attraction by the open anion-bound complex for any metal ions present (assuming an equilibrium between "open" and "closed" liganded forms - see Figure 1.7), and a more highly specific metal binding site evolved from this.

## 1.3 Concluding remarks

The preceding review of past and present studies on the transferrin family of proteins is by no means complete. In attempting to introduce those areas into which this present thesis fits, several important aspects of transferrin chemistry have not been addressed. Perhaps the most important is the question of the function of these proteins.

Clearly the transferrins have roles not only in the movement of iron around the body but also in the regulation of the levels of this essential (and potentially toxic!) metal ion. Lactoferrin, the subject of this study, is essential to the health of new-born infants in its ability to deprive gastrointestinal bacteria of iron (Bullen *et al.*, 1972; 1974). While perhaps not so important in the transport process, it has been proposed that lactoferrin could be acting as an iron carrier following the dicovery of an intestinal lactoferrin receptor (Cox *et al.*, 1979). Neither transferrin nor ovotransferrin could deliver their iron to intestinal tissues.

It is evident that the function of lactoferrin and the transferrins in general is dependent both on the structure of the proteins and on the strength of binding of the iron and the stability of the ternary iron-carbonate-protein complex.

# CHAPTER 2

## Introduction

### 2.1 The interaction of metals with the transferrins

Iron, the first metal associated with members of the transferrin family (Bezkorovainy, 1980), appears to be the most physiologically significant of the metal ions bound by the serum transferrins, although metals such as aluminium (Trapp, 1983), manganese (Lonnerdal, 1984; Davidsson *et al.*, 1989), zinc (Blakeborough *et al.*, 1983) and some heavy metals (Quarterman, 1983) have been identified as being bound to the transferrins *in vivo.* Clinical medicine has exploited the ability of transferrin to bind other metal ions in the use of $Ga^{3+}$ as an imaging agent for displaying a variety of *in vivo* processes (Harris & Sephton, 1977).

Specific receptors for iron-saturated transferrin have been identified on immature red cells and a scheme for the release of the bound iron has been proposed (Octave *et al.*, 1983; Aisen, 1989). The equilibrium binding constants for the binding of iron (see Table 3, Appendix I) are some 100 - 1000 times greater than metal ions of similar charge and ionic radii (Martin *et al.*, 1987) and the bacteriostatic effect of human milk on bacteria such as *E. coli* can be attributed to the great affinity that lactoferrin has for iron (Bullen *et al.*, 1972; 1974).

#### 2.1.1 *Metal binding - general aspects:*

Among the many other metal ions which can be bound by transferrins (Table 2.1 and Table 1 in Appendix I) are the first row transition metals (Warner & Weber, 1953; Aisen *et al.*, 1969; Tan & Woodworth, 1969; Chasteen *et al.*, 1977 and Ainscough *et al.*, 1979), the group 13 metals (Donovan & Ross, 1975; Bertini *et al.*, 1983; 1988; Harris & Pecoraro, 1983), some of the second and third row transition metals (Tan & Woodworth, 1969; Stjernholm *et al.*, 1978 and Duffield & Taylor, 1987), the lanthanides (Luk, 1971) and the actinides, $Th^{4+}$ (Harris *et al.*, 1981) and $Pu^{4+}$ (Duffield & Taylor, 1987). These metal ions are not bound as tightly as iron (see Section 4.1.2).

Table 2.1: Metal ions which have been reported to bind to the transferrins [a]

| d-block metals | | | group 13 | f-block metals | |
|---|---|---|---|---|---|
| 1st row | 2nd row | 3rd row | metals | lanthanides | actinides |
| $Sc^{3+}$ | $Cd^{2+}$ | $Hf^{4+}$ | $Al^{3+}$ | $La^{3+}$ | $Th^{4+}$ |
| $V^{3+}$, $V^{4+}$, $V^{5+}$ | | $Pt^{3+}$ | $Ga^{3+}$ | $Pr^{3+}$ | $Pu^{4+}$ |
| $Cr^{3+}$ | | | $In^{3+}$ | $Nd^{3+}$ | |
| $Mn^{2+}$, $Mn^{3+}$ | | | $Tl^{3+}$ | $Sm^{3+}$ | |
| $Fe^{2+}$, $Fe^{3+}$ | | | | $Eu^{2+}$, $Eu^{3+}$ | |
| $Co^{2+}$, $Co^{3+}$ | | | | $Gd^{3+}$ | |
| $Ni^{2+}$ | | | | $Tb^{3+}$ | |
| $Cu^{2+}$ | | | | $Ho^{3+}$ | |
| $Zn^{2+}$ | | | | $Er^{3+}$ | |

[a] References can be found at the end of Table 1 in Appendix I.

Aisen (1980) proposed three criteria that should be met when labelling a metal ion as specifically bound:

1) No more than two metal ions are bound
2) Binding to iron-saturated transferrin is not observed
3) One bicarbonate (or other anion acceptable by the protein) is bound for each metal ion bound.

Brock (1985) added one more criterion to this list, saying that the metal in question (non-iron) should be displaced by iron when the latter is added. Most of the metal ions mentioned above seem to obey these criteria and can, therefore, be considered as being specifically bound to the transferrins. There are, however, a number of interesting cases where one and sometimes more of the above criteria are not obeyed, even though the metal ion binds specifically. An example of this is the apparent similarity in the affinity of bovine lactoferrin for $Fe^{3+}$ and $VO^{2+}$, in that a large excess of one metal will displace the other (Carmichael and Vincent, 1979). Ainscough *et al.* (1980) reported an anomalous result concerning the binding of chromium(III) to human lactoferrin, in that only one $Cr^{3+}$ is displaced by iron, impying that the second $Cr^{3+}$ is not labile. Non-specific binding of a number of different

metal ions has also been observed, including $Cu^{2+}$, $Zn^{2+}$ (Ainscough *et al.*, 1980; Brock, 1985) and vanadate (Harris & Carrano, 1984), where, in addition to the two metal ions bound at the specific sites, an undetermined number of others can bind.

It is also interesting that at high pH and with no suitable anion present, transferrin is capable of binding one $Cu^{2+}$ ion in or near one of the specific sites, resulting in a non-specific complex where the $Cu^{2+}$ appears to interact with four equivalent nitrogen ligands, as shown by ESR (Zweier & Aisen, 1977). Vanadium also seems to have this ability, as the vanadyl ion ($VO^{2+}$) binds to both sites of ovotransferrin without an anion (Casey & Chasteen, 1980), while the vanadate ion [$VO_2^+$ or $VO(OH)^{2+}$] is able to bind to transferrin in the absence of carbonate (Harris & Carrano, 1984). It is possible that oxo or hydroxo ligands retained on the two vanadium species could act in a similar way to the synergistic anion, shielding the vanadium from the positively charged anion-binding pocket and satisfying the hydrogen bonding potential of the residues involved.

However, in general, most of the metals which bind to the transferrins behave according to the conditions set down by Aisen (1980) and Brock (1985). They all bind in the two specific binding sites, leading to the associated deprotonation of one (if not both) of the tyrosines known to be available in each site. In most cases a suitable synergistic anion is required, and they are displaced from the protein by $Fe^{3+}$. In addition, lowering of the pH of the ternary metal-anion-protein complex facilitates the release of the bound metal ion.

## 2.2 The interaction of anions with the transferrins

The synergistic anion requirement of the transferrins can be met with anions other than (bi)carbonate, including oxalate, malonate, glycolate, thioglycolate, EDTA and nitrilotriacetate (NTA). Schlabach and Bates (1975) tested the ability of 23 dicarboxylic acids, 6 inorganic anions, several carboxylic acids with proximal aldehyde, ketone, amine and thiolate groups as well as acetate, propionate, NTA and EDTA, in their ability to promote the development of a coloured iron-transferrin complex. Of these anions, only the dicarboxylates, NTA, EDTA and some of the mixed anions such as glyoxylate, acetoacetate,

pyruvate and thioglycolate were found to be synergistic (Table 2.2 and Table 2 in Appendix I). This led to the so-called "interlocking sites" model for the interaction of the anion with the metal and the protein, shown schematically in Figure 2.1a. It was found that in order to display the properties of a synergistic anion, the anion must possess a carboxylate group and a second electron donor group no more than 7 Å away (Schlabach and Bates, 1975; see Figure 2.1b). It was proposed that the carboxylate group was either hydrogen-bonded or electrostatically linked to the protein, while the second functional group coordinated to the metal ion.

Table 2.2: Anions, in addition to carbonate, which have been reported to be capable of forming ternary complexes with iron and transferrin (taken from Schlabach & Bates, 1975)

| di- | Anions containing one carboxylate and one of the following: | | | | | |
|---|---|---|---|---|---|---|
| carboxylate | aldehyde | ketone | alcohol | amine | thiol | chelates |
| oxalate | glyoxylate | pyruvate | glycolate | glycine | thioglycolate | NTA |
| malonate | | acetoacetate | lactate | phenylalanine | | EDTA |
| maleate | | ketomalonate | phenyllactate | | | xylenol |
| | | a-ketoglutarate | malate | | | orange[a] |
| | | | phenylglycolate | | | |
| | | | gluconate | | | |
| | | | salicylate | | | |

[a] Reported by Harris & Gelb (1980).

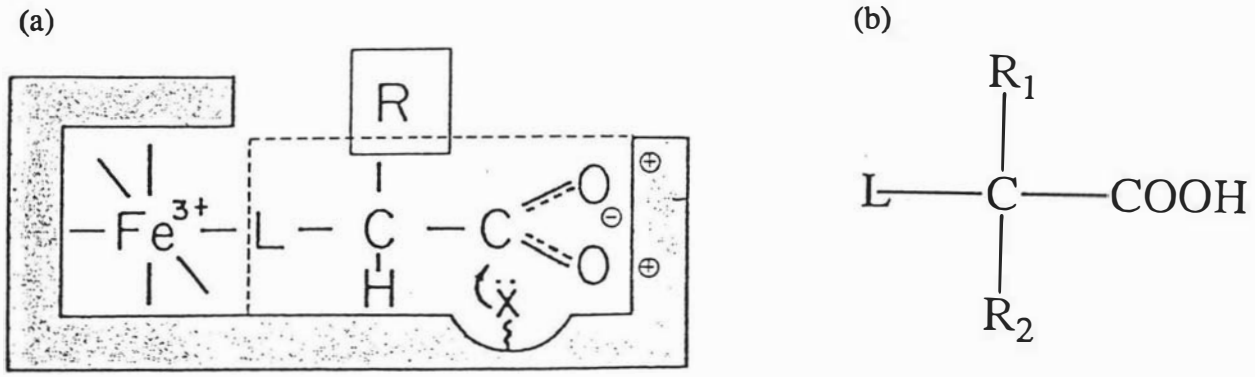



Figure 2.1: a) The "interlocking sites" model of anion binding to the transferrins, where L indicates the second (proximal) functional group and b) the general form of the synergistic anions for the transferrins (from Schlabach & Bates, 1975).

This work was extended by Campbell and Chasteen (1977), who used the vanadyl ion as an ESR probe to establish the synergistic nature of a number of anions. Synergistic anions could be classified as two types (II and III - type I were non-synergistic). Type II anions (mononegative free anions including lactate, salicylate, glycolate and thioglycolate) allowed only one $VO^{2+}$ to bind, whereas type III anions (dinegative free anions such as malonate, oxalate, NTA and carbonate) enabled two $VO^{2+}$ to bind in two different conformations depending on the conditions. It was concluded that these results were in accordance with the "interlocking sites" model and that the anion bridges the metal and anion sites by coordinating to the metal through the second functionality. An EXAFS study of thioglycolate-substituted $Fe_2$OTf (Schneider *et al.*, 1984) also indicated that the sulphur atom of the anion is coordinated to the iron, consistent with the "interlocking sites" model. It was observed that the Fourier transformed EXAFS spectrum of thioglycolate-$Fe_2$OTf exhibited an extra feature at 2.3 Å which was not seen in the spectrum of $Fe_2$OTf containing carbonate as the synergistic anion. Schneider *et al.* (1984) assigned this as an Fe-S interaction, noting that the distance (2.3 Å) was consistent with those observed for ferric thiolate complexes (2.27 - 2.29 Å).

Although these latter results suggested that the anion is directly bound to the metal ion, as the "interlocking sites" model suggests, this question had first been addressed by Harris *et al.* (1974), using $^{13}C$-labelled carbonate. They found that the carbon atom in the anion could be no more than 9 Å from the metal ion in a dicobalt transferrin complex. The ESR experiments on $Fe_2$Tf substituted with nitroxyl spin-labelled oxalate and malonate (Najarian *et al.*, 1978) were also interpreted in terms of direct coordination of the anion to the iron. More recently, electron spin echo envelope modulation (ESEEM) studies by Eaton *et al.* (1989, 1990) have been shown to be consistent with a direct interaction of the metal ion (iron or copper) with the anion (carbonate or oxalate).

As noted above and in Section 1.2.5, the 2.8 Å resolution structure of $Fe_2$Lf reported by Anderson *et al.* (1989) provided the definitive answer, at least in terms of the carbonate anion. It was found that residual electron density adjacent to each iron atom could be best fitted by a carbonate anion bound in a bidentate fashion (see Figure 1.5). Although the

corresponding sites in the structure of $Fe_2$rTf have not been analysed at high resolution, the 2.3 Å resolution structure of the N-terminal half of the rTf molecule (Sarra *et al.* 1990) also shows that the anion is best represented by a bidentate carbonate.

## 2.3 Questions to be addressed by this thesis

While the results of spectroscopic studies, and more recently structural studies, have shed light on the possible functions of the transferrins, established the protein ligands responsible for iron binding, shown the structural relationship between the metal and the synergistic anion and indicated some of the similarities and differences between the three main membrs of the transferrin family (Tf, Lf and OTf), they have also raised a number of important questions concerning the binding and release of metals and anions.

The purpose of the present study is not to attempt to answer all the questions related to the transferrin family, but to address a number of the problems with respect to lactoferrin, from both a spectroscopic and crystallographic aspect, and to try to relate some of the results to the transferrin family as a whole. The broad aims of the study are discussed below.

*Metal binding:* A fundamental question relating to the structure, function and evolution of metal binding sites in proteins concerns the extent to which metal binding is determined by the constraints of the protein structure, or by the stereochemical preferences of the metal ion. In the case of a protein where binding induces a conformational change (as is the case with lactoferrin), different metals could induce different kinds of conformational changes. Because the process of metal binding could be dependent on two factors (the protein and the metal ion), two extreme cases can be envisaged:

(i) A totally rigid binding site where the positions of the ligand sidechains are fixed by interactions with the neighbouring residues. The metal ion would be required to fit into a "pre-specified site" and interact with the ligands in a fixed coordination geometry (the idea of *rack-induced bonding* (Gray & Malmström, 1983)).

A good example of this type of interaction is seen in the copper(II) electron transport protein azurin from *Alcaligenes denitrificans* (Shepard *et al.*, 1990; Shepard, 1991), where

the reduction of copper(II) to copper(I) results in minimal alteration to the coordination geometry. The binding site consists of one thiolate sulphur (Cys) and two imidazole nitrogens (His) arranged in a trigonal plane, with two weak interactions in axial positions (a carbonyl oxygen and a thioether sulphur (Met)). It has been suggested that this arrangement is a compromise between a trigonal bipyramidal geometry favoured by copper(II) and a trigonal planar geometry favoured by copper(I) and that the lack of structural change facilitates fast electron transfer (Shepard *et al.*, 1990). In addition, structural studies on cadmium-substituted azurin (E. N. Baker, personal communication) and metal-free azurin (Shepard, 1991) also show that the arrangement of the ligands is retained when the copper is replaced by another metal ions or even when there is no metal ion present.

(ii) The binding site could be totally flexible with the ligands not constrained by the surrounding protein structure, allowing the protein to adapt to give the stereochemistry which is preferred by the incoming metal ion, the *induced-fit* model of binding.

Clearly the *rack-induced bonding* model would result in a significant amount of selectivity in metal binding. The rigidity of the binding site would impose both a size and a geometrical restriction on the possible metal ions which could bind to the protein, in that metal ions with ionic radii larger than the binding cavity may not be accommodated by the protein, and cations which favour geometries other than that provided by the ligands would probably be bound much more weakly than others whose preferred geometry matched the site. The induced-fit model, on the other hand, would allow virtually all metal ions to be bound, irrespective of their geometrical preferences or their size.

The selectivity of metal binding observed in the transferrins suggests that the metal binding sites are not completely flexible, although the extent of this rigidity is not known. A certain amount of size restriction has been observed with respect to the lanthanide ions. Luk (1971) found that for the larger members of the lanthanide group (e.g. $Nd^{3+}$) only one metal ionn was bound by hTf (although, the results of Harris (1986b) who determined binding constants for both sites, throws these earlier observations into doubt), while the smaller ions (e.g. $Tb^{3+}$ and $Eu^{3+}$) resulted in a 2:1 stoichiometry. Similar results were also obtained by Chung (1985). Clearly, the ability of the transferrins to bind 2 equivalents of $Eu^{3+}$, which

has an ionic radius of 0.947 Å[§] (Shannon, 1976) significantly larger than Fe(III) (0.645 Å), indicates that there is some flexibility in the positions of the ligands, although the amount of movement must be limited.

In addition to a restriction in terms of the size of the ions, the geometrical and coordination preferences of the metal also play a major role. The most notable effect of this can be seen in the strength of binding of various metal ions to the transferrins. Iron(III) is bound much more tightly than any other metal ion ($K_{eq} \approx 10^{21}$), even those of similar size. For example, manganese(III), with an ionic radius (0.645 Å) identical to that of iron(III) (Shannon, 1976), cannot remove Fe(III) from $Fe_2Tf$. However, when Fe(III) is added to $Mn_2Tf$ solutions, the iron readily replaces the manganese in the binding site (Brock, 1985).

Metal ions similar to iron with respect to stereochemical preferences, Ga(III) and Al(III) for example, although of similar size to Fe(III) (0.62 and 0.535 Å respectively) and favouring octahedral geometries, have association constants lower than iron ($K_{eq} \approx 10^{19}$ and $10^{13}$ respectively). However, these are considerably greater than the values for metal ions which adopt geometries very much different from that of iron, such as $Zn^{2+}$ which favours tetrahedral coordination and has a $K_{eq}$ of about $10^5$. This suggests once again the inability of the metal binding ligands to alter their positions to any great extent.

It is also possible that factors such as the charge on the metal ion could lead to differences in its interaction with the protein ligands or in the hydrogen bonding interactions of the ligands with surrounding parts of the protein and solvent structure. Crystallographic studies on metal-substituted transferrins are clearly required in order to establish the importance of these various factors.

*Anion binding:* Structural studies on Lf and rTf have shown that carbonate, the physiological synergistic anion, is bound in a bidentate fashion to the iron (discussed in Section 2.2), and bridges between the metal ion and some positively-charged portions of the protein. In the absence of an anion the latter would possibly result in some unfavourable electrostatic interactions with the metal ion. The mode of carbonate binding seen in

§ For a 6-coordinate $Eu^{3+}$. The ionic radius increases with increasing coordination number.

lactoferrin and transferrin appears to be in accord with the *interlocking sites* model of Schlabach & Bates (1975) (Section 2.2), who also suggested a bridging function for the anion. However for other anions it is reasonable to suppose that the carboxylate function might bind preferentially to the metal rather than the proximal ligand if such an opportunity existed, as high spin $Fe^{3+}$ has a preference for oxygen donors such as carboxylate ligands (Cotton & Wilkinson, 1988).

Although it is not possible to give an unequivocal generalised anion binding model for the transferrins based solely on the structural studies with carbonate, three possible models can be envisaged for a synergistic anion of general structure (Figure 2.1b): (i) The anion could be bound with the proximal ligand, L, coordinated to the metal and the carboxylate group interacting with the surrounding protein via hydrogen bonds. This is the Schlabach and Bates *interlocking sites* model; (ii) the anion could be bound with the carboxylate coordinated to the metal, with the proximal ligand interacting with the surrounding protein; (iii) the anion could be coordinated to the metal ion in a bidentate fashion. In this case, the carboxylate group and/or the proximal ligand might also interact with the surrounding protein.

Several more recent studies involving NMR and ESEEM spectroscopy have attempted to establish how the anions interact with metal ions in the transferrins, with conflicting results. For example, NMR studies on the binding of $^{13}$C-oxalate to gallium transferrin were interpreted as 1,1-bidentate coordination of the oxalate (Bertini *et al.*, 1986a), that is one carboxylate was directly attached to the gallium, in a bidentate fashion, while the other end interacted with the protein. In contrast to this, the ESEEM studies on the Cu(II) and $VO^{2+}$ complexes of hTf indicated that oxalate bound in a 1,2-bidentate fashion (Eaton *et al.*, 1989), with both ends attached directly to the metal ion. Definitive evidence of the mode of binding of these anions can only come from high resolution crystallographic studies on anion-substituted transferrin complexes.

*Differences between the transferrins and between the two sites:* Metal and anion substitution studies on the transferrins have raised several questions concerning differences

between the members of the transferrin family. With respect to anion substitution, for example, Ainscough *et al.* (1983) observed that when excess oxalate ions were added to dicupric lactoferrin (prepared with carbonate as the synergistic anion), it appeared that one carbonate was replaced by an oxalate. This resulted in a hybrid complex with carbonate in one site and oxalate in the other, having electronic and ESR spectra different to those of the parent dicarbonato complex. A similar experiment with diferric lactoferrin and excess oxalate has now shown that no anion substitution occurs in this case (Shongwe *et al.*, 1992), indicating that when iron is the associated metal, carbonate is much less susceptible to replacement.

Zweier & Aisen (1977) and Zweier (1980) in their studies on hTf and cOTf respectively, also observed substitution of oxalate for carbonate upon the addition of an excess amount of the former, and it was found that for hTf this substitution occurred in the N-terminal site. It was also established that the two proteins (hTf and cOTf) behaved somewhat differently with respect to anion substitution. For hTf, in the presence of $Cu^{2+}$, only one oxalate was bound to the protein and this was found to be incorporated into the N-terminal site. However, for cOTf, two oxalate ions were bound, one in each site, and the C-terminal site had five times the affinity for oxalate as the N-site.

As the metal and anion sites in the transferrins are closely inter-related, it is likely that factors which cause changes in the metal binding envionment may also affect the coordination of the anion and vice versa. This raises the question of what structural differences between transferrin, ovotransferrin and lactoferrin could account for these observations.

The majority of the high resolution structural work to date has concentrated on human lactoferrin, but most of the spectroscopic studies have been carried out on the other members of the transferrin family. Hence, in order to accurately translate the spectroscopic observations on the transferrins into conclusions about the 3-dimensional structure, it is necessary to be able to relate spectroscopy to structure. The results obtained by spectroscopic and structural methods, although in appearance totally independent, can become powerful tools when combined. It is envisaged that this combined study on the 3-

dimensional structures of dicupric human lactoferrin complex (containing carbonate as the synergistic anion) and a dicupric complex with one carbonate replaced by an oxalate, will, when combined with the results of spectroscopic studies on lactoferrin, lead to a much clearer understanding of the relationship between the structure of the protein in general, the specific structure observed in the metal and anion binding sites, and the nature of the anions and metal ions which bind.

# CHAPTER 3

## Crystallographic studies

As collection and processing of the data for the two lactoferrin complexes was essentially similar in each case, the experimental procedures involved for both will be described together. The preparation of each of the complexes will, however, be described separately for the sake of clarity.

### 3.1 Experimental procedures

3.1.1 *Preparation and crystallisation of the metal-anion-lactoferrin complexes:*

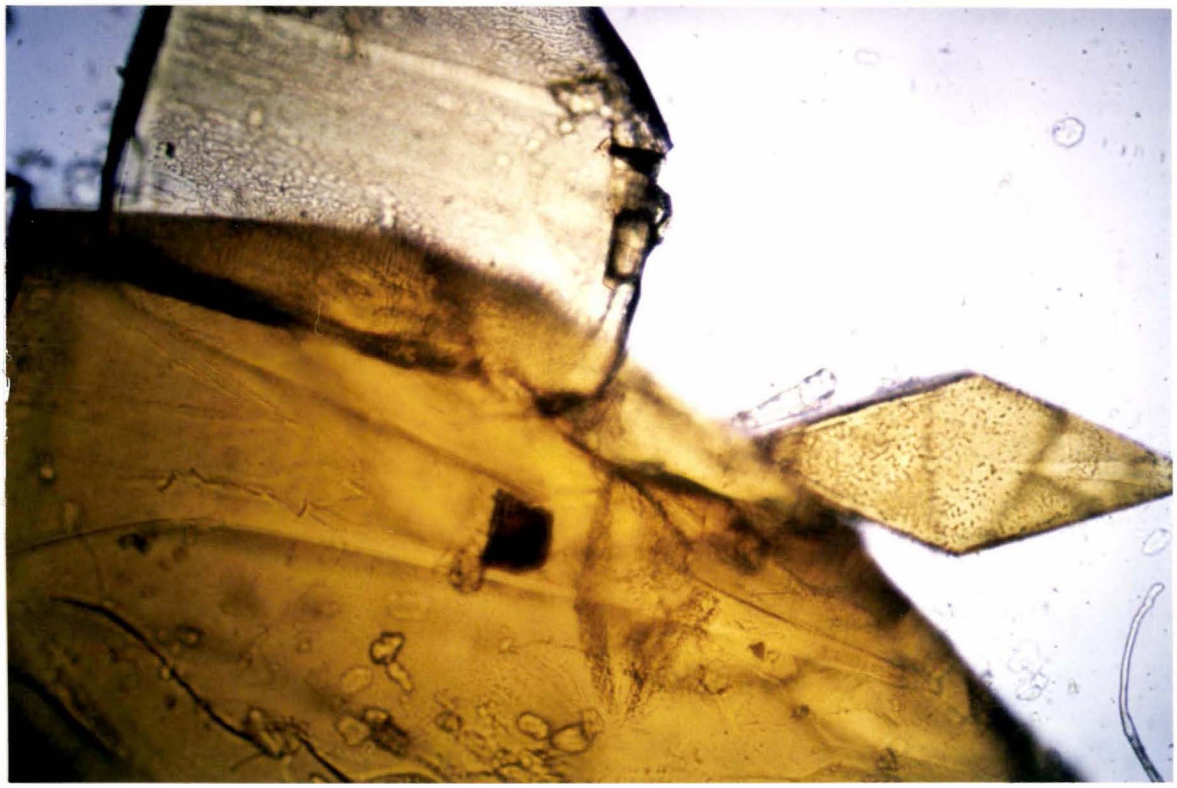
*Dicupric-dicarbonato-lactoferrin ($Cu_2Lf$):* This complex was prepared as described in section 4.2.8. The yellow $Cu_2Lf$ was precipitated by making the solution 70% saturated with ammonium sulphate. The precipitate was dissolved in sufficient 0.025M Tris-HCl/0.1M NaCl, pH 7.8 to give a protein concentration of about 40-50 mg/ml, and loaded into 8 mm dialysis tubing. The solutions were then equilibrated for 7 days in vials containing 18 ml of 0.01M sodium phosphate buffer, pH 7.8, at 4°C and at the end of this time the buffer was made 10% v/v in ethanol. Yellow crystals (Plate 3.1a) of $Cu_2Lf$ grew in two weeks to three months. The mother liquor was colourless at the end of this time, indicating that virtually all of the $Cu_2Lf$ had crystallized.

The crystals were then carefully transferred from the ethanol-phosphate mother liquor to a 0.01M phosphate solution containing 2-methyl-2,4-pentanediol (MPD). Anderson *et al.* (1989) had found that the replacement of the more volatile ethanol with MPD greatly improved the quality of crystals of $Fe_2Lf$ and the resulting diffraction pattern. Vials containing phosphate buffer and the following ethanol/MPD mixtures were set up:

10%:0, 8%:2%, 6%:4%, 6%:6%, 4%:8%, 2%:10%, 0:12%, 0:16% and 0:20%

The dialysis bags containing the crystals were transferred from one solution to the other, allowing at least two days to equilibrate in each solution. Careful checks were made on the crystals to ensure that they were not dissolving or cracking.

(a)



(b)

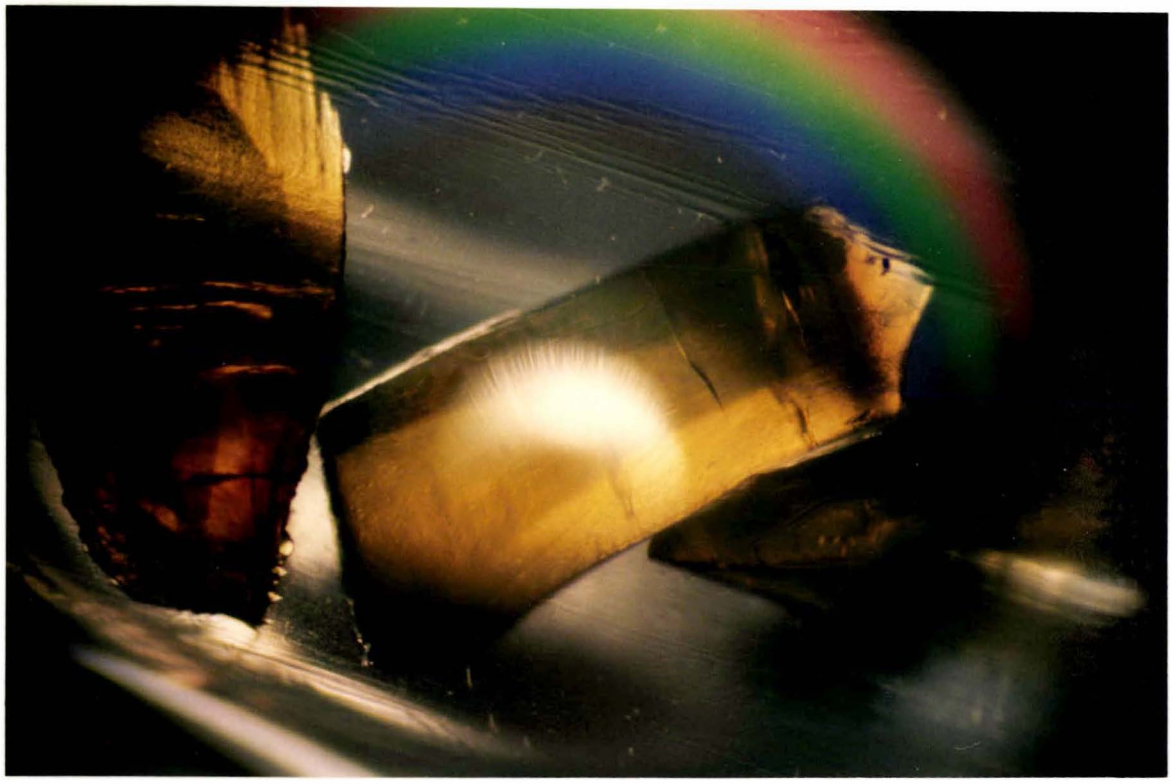


Plate 3.1: Crystals of (a) dicupric dicarbonato lactoferrin ($Cu_2Lf$) and (b) dicupric (carbonato-oxalato) lactoferrin ($Cu_2oxLf$).

*Dicupric-(carbonato-oxalato)-lactoferrin ($Cu_2$oxLf):* The samples of $Cu_2$oxLf used for crystallisation were prepared by the first route described in section 4.2.8 using "EDTA-treated" apolactoferrin as the starting material. One carbonate in $Cu_2$Lf was substituted by oxalate by incubation of $Cu_2$Lf with a 50 molar excess of oxalate. The complex was left to stand and the replacement of the carbonate with oxalate was monitored by UV/visible absorption spectroscopy until formation of the hybrid complex appeared to be complete, typically after 5 days to one week (see section 3.1.9 for details). The solution was then subjected to ultrafiltration (Amicon XM-50 membrane) and concentration using an Amicon microcentricon until the protein concentration was about 60-70 mg/ml (as judged by the decrease in volume). Crystallisation conditions were identical to those described above for $Cu_2$Lf, except that both methanol and ethanol were used as precipitaing agents, whereas only ethanol was used in the case of $Cu_2$Lf. Yellow crystals (Plate 3.1b) appeared within 4 -7 days after the addition of methanol and were allowed to stand in the mother liquor at 4°C for 4 weeks until crystallisation was complete, as indicated by the loss of colour from the mother liquor. These crystals were also transferred to phosphate containing MPD in a similar way to that described above.

Prior to data collection, 9 mg of unused crystalline material was placed in a 1.2 ml quartz cuvette capped by a silicon septum and flushed with argon for 10 minutes. One ml of phosphate buffer at pH 7.8, which had been subjected to repeated evacuation and argon flushing at pH 3.5 to remove all $CO_2$, was injected into the cuvette and the crystals allowed to dissolve. The absorption spectrum was subsequently recorded with an HP 8452A electronic absorption spectrophotometer using 1 ml of fresh degassed phosphate buffer in the reference cell. Figure 3.1 shows the spectrum obtained from the dissolved crystals overlaid on a spectrum of freshly prepared $Cu_2$Lf and $Cu_2$oxLf for comparison. It can be clearly seen that the spectrum from the crystalline hybrid complex is virtually identical to that of fresh $Cu_2$oxLf, in that they both have $\lambda_{max}$ values of ≈424 nm. $Cu_2$Lf has a $\lambda_{max}$ about 10 nm higher (≈434 nm) (see section 4.3.1). The spectrum obtained from the dissolved crystals indicates, at least qualitatively, that little or no replacement of oxalate by carbonate took place during crystallisation.

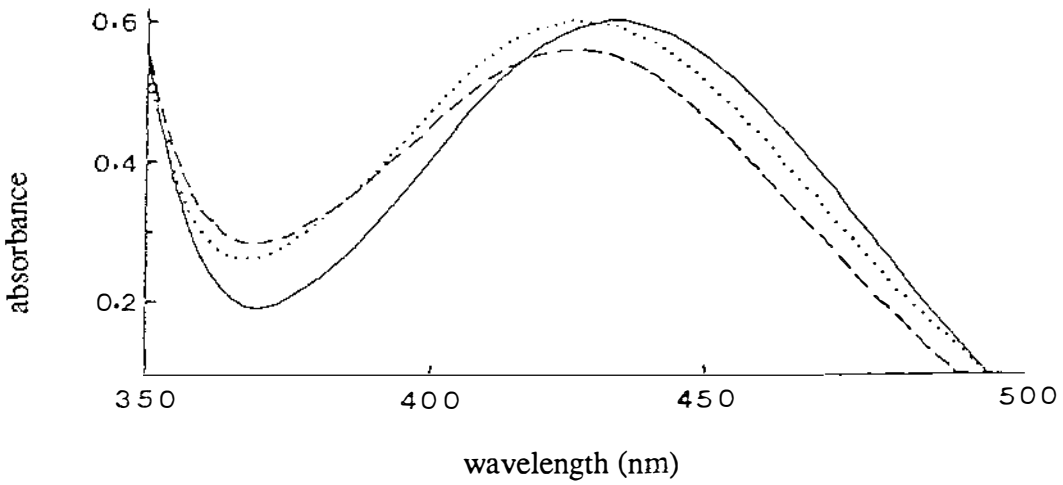



Figure 3.1: Electronic absorption spectrum of a $Cu_2$oxLf (·····) solution prepared by dissolving some crystalline $Cu_2$oxLf in carbonate-free phosphate buffer. A spectrum of freshly prepared $Cu_2$Lf (——) and $Cu_2$oxLf (- - - -) is shown for comparison.

### 3.1.2 *Crystal preparation, mounting and unit cell determination:*

The crystals of $Cu_2$Lf and $Cu_2$oxLf (up to 3 mm long) were carefully cut with a scalpel into smaller fragments about 1 mm in length. These were mounted in thin-walled, non-siliconised glass capillaries along with a little of the mother liquor. The ends of the tubes were sealed with wax. One crystal of each of $Cu_2$Lf and $Cu_2$oxLf were used for preliminary characterisation by precession photography. Two zero-layer precession photographs (90° apart) were taken for each crystal, with the precession angle set at 17° and 15° for $Cu_2$Lf and $Cu_2$oxLf respectively (Plates 3.2a - d). From these, values for the cell dimensions could be estimated (see Table 3.1). It was observed that the $Cu_2$Lf crystal diffracted to about 2.6 Å resolution, where the sp ɔts were quite weak, while the diffraction pattern for the $Cu_2$oxLf crystal, although only recorded to about 3 Å resolution due to the smaller precession angle used, appeared to be stronger and likely to extend beyond 2.5 Å.

Final unit cell dimensions were determined with the aid of a computer-controlled Enraf-Nonius CAD4 four-circle diffractometer. The diffractometer was set to scan through

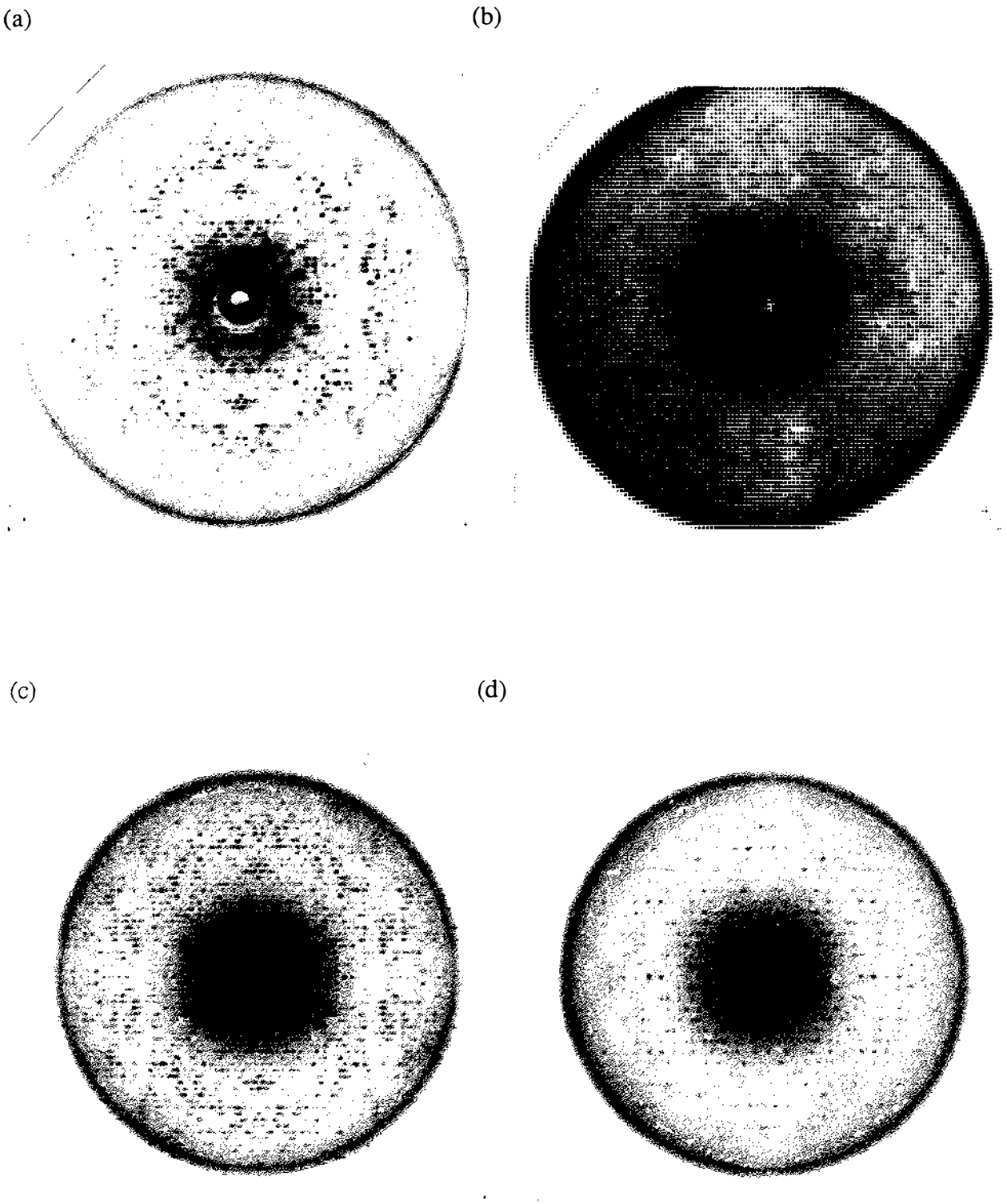



Plate 3.2: Zero-layer precession photographs of the (a) hk0 and (b) 0kl reciprocal lattice nets for $Cu_2Lf$ ($\mu$ = 17°), along with the (c) hk0 and (d) 0kl nets of $Cu_2oxLf$ ($\mu$ = 15°). Observable data can be seen at the edges of the precession circle (2.6 Å and 3.0 Å resolution for $Cu_2Lf$ and $Cu_2oxLf$ respectively).

reciprocal space at low θ angles (near θ = 1.6°, 2.3° and 3.5°) to find a set of strong reflections. These reflections were then used to determine an initial cell by least-squares methods. If this cell compared well with that estimated from precession photographs, then the derived orientation matrix was employed to locate strong reflections at θ angles between 4 and 4.5°. Least squares refinement of the positions of these new reflections gave an improved unit cell and higher angle reflections could be located. The procedure was repeated until an accurate unit cell and orientation matrix could be calculated using 25 reflections with θ > 8.5°.

The cell dimensions estimated from precession photographs and the average cell dimensions derived from diffractometer measurements on the crystals are given in Table 3.1.

Table 3.1: Estimated cell dimensions and average cell dimensions used in subsequent scaling and refinement procedures [a].

| complex | estimated cell dimensions (Å) [b] | average refined cell dimensions (Å) [c] |
|---|---|---|
| $Cu_2Lf$ | 156.03, 97.03, 55.91 | 155.90, 96.98, 56.04 |
| $Cu_2oxLf$ | 155.95, 97.07, 56.22 | 155.80, 96.85, 55.91 |

[a] Orthorhombic, $P2_12_12_1$, $\alpha = \beta = \gamma = 90°$ [b] From precession photographs. [c] From diffractometer orientation matrix.

3.1.3 *Data collection and processing:*

Medium resolution X-ray diffraction data were collected on each of the complexes with the CAD4 diffractometer using nickel-filtered $CuK_\alpha$ radiation. For $Cu_2Lf$, 5 crystals were required to collect data to 2.6 Å resolution, and for $Cu_2oxLf$, 4 crystals for data to 2.5 Å resolution. Intensities were measured using the limited-range step-scan method of Hanson *et al.* (1979), which minimises the X-ray exposure time and thus prevents unnecessary decay of the crystals. In this method, a symmetrical form of the peak is assumed and a limited number of steps across the top of the peak can be used to give a good estimate of the profile and the intensity (Hanson *et al.*, 1979). The scan does not need to be wide enough to cover background measurements as in the standard diffractometer protocol. Either (i) no

background measurements are made or (ii) one or two background measurements are made at positions between reciprocal lattice points. In the latter case, the background level for each individual reflection can be taken as either the measured background (in the case of a single measurement) or as the average of two measurements.

For the three data sets collected as a part of this study, the former method was employed, where no backgrounds were measured. During data processing, the backgrounds were estimated by taking the intensity measurements at the limits of the scan range for weak reflections and averaging them over different regions of reciprocal space (Baker & Dodson, 1980) (see the description of the program PREPRO below).

Data were collected in shells, generally starting at high values of the Bragg angle, $\theta$ (ie. at high resolution) and working down to low $\theta$ (low resolution). The high resolution data was collected first as it is generally of much weaker intensity and collection of this data from a fresh crystal ensures that it will be of the highest quality possible. Although there was virtually no overlap between shells for a given crystal, to assist with scaling the data from different crystals together, a small amount of overlap was allowed between crystals. Five standard reflections, chosen to cover the octant of reciprocal space from which data was being collected, were measured every 120 minutes of X-ray exposure time and these were used to provide an estimate of the decay of the crystals in the X-ray beam. Crystals were discarded after the initial intensities had fallen by between 30 - 40%.

Typically between 5000 and 10000 reflections could be measured from each crystal before it was discarded. At low $\theta$, where most of the reflections were of high intensity, moderately fast scan speeds could be employed and a large number of reflections measured before the intensity fell below 35%. However, at higher resolution ($\theta > 15°$) where the reflections are much weaker, slower scan speeds were used in order to give peak profiles from which a reasonable estimate of the integrated intensity could be derived. Consequently, fewer reflections could be measured from these crystals. Summaries of the collection of the medium resolution diffractometer data for the two complexes are given in Tables 3.2 and 3.3.

Table 3.2 : Diffractometer data collection summary for $Cu_2Lf$

| crystal | shell | theta range (deg) | scan speed [a] | collection scheme [b] | total reflections | cumulative decay (%) |
|---|---|---|---|---|---|---|
| 1 | 1 | 14.0-12.0 | 0.824 | 8 x 0.05° | 5297 | 16 |
| | 2 | 12.0-10.5 | 0.969 | 8 x 0.05° | 3029 | 26 |
| | 3 | 1.0-8.9 | 1.268 | 8 x 0.05° | 2578 | 36 |
| 2 | 4 | 8.9-10.5 | 0.969 | 8 x 0.04° | 2450 | 3 |
| | 5 | 1.0-8.9 | 1.268 | 8 x 0.04° | 3995 | 13 |
| | 6 | 14.0-17.2 | 0.250 | 8 x 0.04° | 657 | 19 |
| | 7 | 14.0-17.2 [c] | 0.250 | 8 x 0.04° | 1117 | 32 |
| 3 | 8 | 14.0-17.2 | 0.445 | 8 x 0.05° | 4763 | 31 |
| 4 | 9 | 14.0-17.2 | 0.445 | 8 x 0.04° | 2606 | 15 |
| | 10 | 14.0-17.2 [c] | 0.445 | 8 x 0.04° | 2472 | 36 |
| 5 | 11 | 14.0-17.2 | 0.499 | 8 x 0.05° | 3788 | 31 |

[a] Scan speed in degrees per minute. [b] Number of steps × step increment in degrees. The data were collected with h, k and l all positive. [c] Occasional faults with the diffractometer (such as water filter blockages or jammed shutters) meant that data collection of some shells had to be restarted and even though the theta range was unchanged, the shells were given different numbers.

Table 3.3 : Diffractometer data collection summary for $Cu_2OxLf$

| crystal | shell | theta range (deg) | scan speed [a] | collection scheme [b] | total reflections | cumulative decay (%) |
|---|---|---|---|---|---|---|
| 1 | 1 | 2.0-8.9 | 1.831 | 8 x 0.04° | 3968 | 10 |
| | 2 | 8.9-11.0 | 1.448 | 8 x 0.04° | 3363 | 17 |
| | 3 | 11.0-12.3 | 1.177 | 15 x 0.04° | 2808 | 20 |
| | 4 | 12.0-13.5 | 0.916 | 10 x 0.04° | 3110 | 33 |
| 2 | 5 | 15.0-16.2 | 0.634 | 8 x 0.04° | 4303 | 13 |
| | 6 | 13.0-15.0 | 0.634 | 8 x 0.04° | 6052 | 41 |
| 3 | 7 | 13.0-15.0 | 0.634 | 10 x 0.05° | 2524 | 10 |
| | 8 | 13.0-15.0 [c] | 0.634 | 10 x 0.04° | 1545 | 23 |
| | 9 | 16.2-17.0 | 0.445 | 10 x 0.04° | 1216 | 35 |
| 4 | 10 | 16.2-17.0 | 0.549 | 10 x 0.04° | 1157 | 12 |
| | 11 | 16.2-17.0 [c] | 0.549 | 10 x 0.04° | 2186 | 13 |
| | 12 | 17.0-18.0 | 0.499 | 10 x 0.05° | 1646 | 24 |

[a] Scan speed in degrees per minute. [b] Number of steps × step increment in degrees. The data were collected with h, k and l all positive except for crystal 4 which was collected as -h, k, -l.
[c] Occasional faults with the diffractometer (such as water filter blockages or jammed shutters) meant that data collection of some shells had to be restarted and even though the theta range was unchanged, the shells were given different numbers.

3.1.4 *Data reduction:*

The data were reduced in a three-step procedure (programs PREPRO, FITC3 and REDUCER), based on the method of Hanson *et al.* (1979). The programs used were based on originals from the University of Washington, Seattle and Washington University, St. Louis, which had been modified at Massey University (Baker & Dodson, 1980). The functions of each of the three steps are as follows:

(1) PREPRO

This program calculates a deterioration correction and an empirical absorption correction and prepares a background correction file. The five standard intensity control reflections are integrated by fitting the profile data of each to a Gaussian function and the values of the intensities obtained are then used to estimate the deterioration by fitting to a 3-term polynomial as a function of X-ray exposure time and intensity.

An estimate of the background for each reflection was made by fitting the observed background values of the very weak reflections to a 3-term polynomial in $\sin\theta$ (to allow for the change in background levels with resolution), and then averaging them as a function of the diffractometer angles phi and chi (Baker & Dodson, 1980). Reflections defined as weak were those with peak counts $\leq R \times$ background count (where R was typically between 1.1 and 2.0, depending upon the strength of the data). This procedure takes account of the variation of the backgrounds over the whole of reciprocal space. Among data collected at low theta (1.0 - 9.0°) there were a large number of intense reflections and an R value of 1.8 - 2.0 was required to yield enough weak reflections (at least 15 - 20 in each 10° phi-chi block). At higher values of theta (14.0 - 17.0°) the value of R was set as low as 1.2 as there were usually sufficient weak reflections to give a good estimate of the background. The averaged backgrounds were then ,moothed over adjacent 10° phi-chi blocks. As a final check, the smoothed, averaged backgrounds at a number of values of theta were compared to the peak profiles of weak reflections in the diffractometer output to see whether the estimated backgrounds corresponded with the observed backgrounds. If significant discrepancies were observed, then the value of R was altered to either include or exclude more of the stronger reflections in the averaging.

Finally an empirical absorption correction (method of North *et al.*, 1968) was calculated from an azimuthal scan. The X-ray transmission was plotted as a function of the diffractometer angle phi and smoothed by hand if necessary.

(2) FITC3

The steps making up the peak profiles of each reflection are fitted to a Gaussian function of the form

$$Y_i = C1 \exp - \frac{(x_i - C2)^2}{(C3)^3} \qquad [3.1]$$

where: $Y_i$ is the intensity at step i, C2 is the centroid of the peak (the value of ω, in degrees), $x_i$ is the ω position of step i, C1 is the peak height and C3 is the peak width.

The precision of the width parameter for weak reflections is quite low due to the poor statistical accuracy of counts (Hanson *et al.*, 1979) and also as a result of the crystal mosaicity. In the initial pass, only the strongest reflections (generally with a peak intensity/background ratio ≥ 2.0) are used to define the peak profiles in different regions of reciprocal space (the backgrounds as estimated by PREPRO are subtracted from these reflections before the fitting procedure). The profile parameters from the strong reflections are then fitted by least-squares refinement to a 7-term function for subsequent use in the data reduction program.

(3) REDUCER

This program uses the parameters and corrections determined by PREPRO and FITC3 and produces a reduced data file output as either intensities ($F^2$) or structure amplitudes (F). Firstly, the values for C3 are calculated from the relationship

$$C3 = \sum_i \sum_j a_i a_j A_{ij} - k\tan\theta \qquad [3.2]$$

where $a_i$ and $a_j$ are the direction cosines of the diffraction vector and $A_{ij}$ are the elements of a symmetric tensor, which describes the variation of the peak width throughout reciprocal

space, as calculated by FITC3. The ktanθ term accounts for the effect of spectral dispersion on the peak width. The fitting of the parameters obtained from strong reflections to the 7-term function allows an estimate of the C3 values for the weak reflections to be made. The C3 values are then fitted to equation [3.1] to determine the values for C1 and the relative integrated intensities which are proportional to C1 × C3.

The estimated background for each reflection is read from the phi-chi table generated by PREPRO, after correcting to the appropriate theta level, and is applied to the intensities, along with the deterioration and absorption corrections as well as a standard Lp correction. The final data sets were output as structure amplitudes (F).

### 3.1.5 *Data scaling:*

The structure amplitudes output by REDUCER were scaled in a two-step procedure with the programs INTAV3 and INTAV4 which calculate and apply scale factors based on common reflections. The data from $Cu_2$Lf and $Cu_2$oxLf were treated identically, as described below.

The shells collected for each crystal (see Tables 3.4 and 3.5) were scaled to the intensity level of the first shell using INTAV4, with scale factors based on the five standard intensity control reflections (Tables 1 - 7, Appendix II). This method assumes that the falloff in intensity of the five standards occurs at a roughly equal rate (which is not always the case) and that standard reflections chosen are representative of the data set as a whole (every effort is made to cover the widest possible area of the octant of reciprocal space from which data is being collected). These criteria cannot always be met. Specific radiation damage superimposed on a general decay can cause the intensities of different reflections to decline at different rates. Furthermore, for high resolution shells, no reflections may be strong enough to use as standards, and the standards chosen (from lower resolution shells) will not be truly representative.

Table 3.4 : Scale factors for $Cu_2Lf$

| crystal | shell | inter-shell [a] | inter-crystal [b] | common reflections [c] | $R_{in}$ [d] |
|---|---|---|---|---|---|
| 1 | 1 | 1.000 | 2.539 | 2840 | 0.024 |
| | 2 | 1.183 | | | |
| | 3 | 1.449 | | | |
| 2 | 4 | 1.000 | 4.891 | 3124 | 0.031 |
| | 5 | 1.162 | | | |
| | 6 | 0.539 | | | |
| | 7 | 0.575 | | | |
| 3 | 8 | — | 1.000 | 1315 | 0.040 |
| 4 | 9 | 1.000 | 1.439 | 658 | 0.101 |
| | 10 | 1.084 | | | |
| 5 | 11 | — | 1.399 | 940 | 0.064 |
| | | | | | $R_{merge}$ [e] = 0.032 |

[a] Calculated from the decrease in intensity of five reflections which were measured every 120 minutes of X-ray exposure time. These numbers are the square root of the values given in the tables in Appendix II. [b] Calculated from the reflections which were found to be common between the crystals. [c] Number of reflections which are common between a given crystal and all the others. [d] $R_{in}$ is the inter-crystal R-factor showing the agreement between a particular crystal and the others. [e] $R_{merge}$, the overall agreement factor between the crystals, is defined as $\Sigma|F - \overline{F}|/\Sigma\overline{F}$.

Table 3.5 : Scale factors for $Cu_2oxLf$

| crystal | shell | inter-shell | inter-crystal | common reflections | $R_{in}$ |
|---|---|---|---|---|---|
| 1 | 1 | 1.000 | 1.869 | 1161 | 0.055 |
| | 2 | 0.859 | | | |
| | 3 | 0.775 | | | |
| | 4 | 0.736 | | | |
| 2 | 5 | 1.000 | 1.000 | 3503 | 0.057 |
| | 6 | 1.094 | | | |
| 3 | 7 | 1.000 | 1.015 | 3115 | 0.069 |
| | 8 | 1.128 | | | |
| | 9 | 1.031 | | | |
| 4 | 10 | 1.000 | 0.806 | 212 | 0.066 |
| | 11 | 1.029 | | | |
| | 12 | 0.946 | | | |
| | | | | | $R_{merge}$ = 0.062 |

A more accurate method would be to allow a certain amount of overlap between the individual shells for a crystal, and then calculate scale factors based on the intensities of common reflections only. However, this method has a disadvantage in the case of proteins, where the data must be collected as fast as possible due to the decay of the crystals in the X-ray beam. At high resolution where the data is much weaker, a larger number of common reflections would be needed to give an accurate scale factor and a larger overlap of shells would be required. This would result in the need for more crystals for the data collection, leading to greater error in the final inter-crystal scaling and possibly a poorer quality final data set. Clearly a balance between the number of crystals required and the accuracy of the scale factors must be sought, and the use of standard reflections provides a quick and relatively straightforward method of estimating these scale factors to a reasonable degree of accuracy.

Once the shells in each crystal had been scaled together, inter-crystal scale factors were calculated based on common reflections using INTAV3 (column 4, Tables 3.4 & 3.5). This procedure resulted in final $R_{merge}$ values on $|F|$ ($R_{merge} = \Sigma |F - \overline{F}| / \Sigma \overline{F}$) of 0.032 and 0.062 for $Cu_2$Lf and $Cu_2$oxLf respectively (from 4844 and 3756 common reflections).

It should be noted that the value of $R_{merge}$ for $Cu_2$oxLf is significantly greater than that obtained for $Cu_2$Lf, as are the individual R-factors. Initially, a sigma cutoff of 1.0 on the $Cu_2$oxLf data was employed in INTAV3, meaning that reflections which had $I < 1\sigma_I$ were rejected. However, this resulted in the program being unable to find any common reflections between crystals 3 and 4. Subsequently, the sigma cutoff was lowered to zero (all reflections were treated as observed) and 300 common reflections were found, although the agreement was very poor. This data set, including crystal 4, was used in the initial stages of the structure refinement, although when it was merged with the high resolution synchrotron data (see Section 3.1.6), crystal 4 was omitted.

Finally the x-ray data files were converted into the LCF format (3 reflection records per line) for use in the PROLSQ refinement package. For $Cu_2$Lf there were 24387 unique reflections (21318 having $I > 2\sigma_I$) while for $Cu_2$oxLf there were 25037 reflections (21097 with $I > 2\sigma_I$). The details of the diffractometer data sets for the two complexes are given in Table 3.6. Figures 3.2a & b (page 53) show the total number of unique reflections measured

by the diffractometer from $Cu_2Lf$ and $Cu_2oxLf$ as a function of resolution ($4\sin^2\theta/\lambda^2$). From the two plots, it is clear that there is a marked falloff in the number of observed (with $I > 2\sigma_I$) reflections at about 3.1 and 2.7 Å for $Cu_2Lf$ and $Cu_2oxLf$ respectively.

Table 3.6: Data collection (diffractometer)

| | $Cu_2Lf$ | $Cu_2oxLf$ |
|---|---|---|
| number of crystals used | 5 | 4 |
| radiation, wavelength (Å) | $CuK_\alpha$, 1.5418 | $CuK_\alpha$, 1.5418 |
| scan type | step scan in ω | step scan in ω |
| absorption correction | azimuthal scan | azimuthal scan |
| total reflections measured | 29730 | 30338 |
| independent reflections | 24387 | 25037 |
| observed reflections [a] | 21318 (87%) | 21097 (84%) |
| $R_{merge}$ [b] | 0.032 | 0.062 |
| maximum resolution, $d_{min}$ | 2.60 Å | 2.56 Å |
| - theoretical number of reflections to $d_{min}$ | 25241 | 26604 |
| - completeness of data to $d_{min}$ | 97% | 94% |

[a] With $I > 2\sigma_I$. [b] $R_{merge} = \Sigma | F - \bar{F} | / \Sigma \bar{F}$

### 3.1.6 *High resolution oscillation photographic data:*

*$Cu_2Lf$:* Data to 2.1 Å resolution were collected at the SRS facility at Daresbury, UK, by the oscillation method (Arndt and Wonacott, 1977) using radiation of wavelength 0.88 Å produced by a 5 Tesla Wiggler magnet. The circulating current in the synchrotron was 130 mA. Two crystals, covering a rotation range of 90° about the a* axis were required to collect the full data set on 57 film packs of 3 films each. The oscillation angle was 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 digitised with an Optronics P1000 rotating drum microdensitometer and processed by a package of programs developed by Wonacott (1980). This involved the indexing of the films (IDXREF), the generation of a calculated diffraction pattern based on the cell parameters, the crystal missetting and the oscillation range (GENVEE) and the

alignment of this calculated pattern onto the observed (MOSBA1). Integrated intensities were obtained from each of the three films in the filmpacks (MOSBA2) and a standard Lorentz-polarisation correction made.

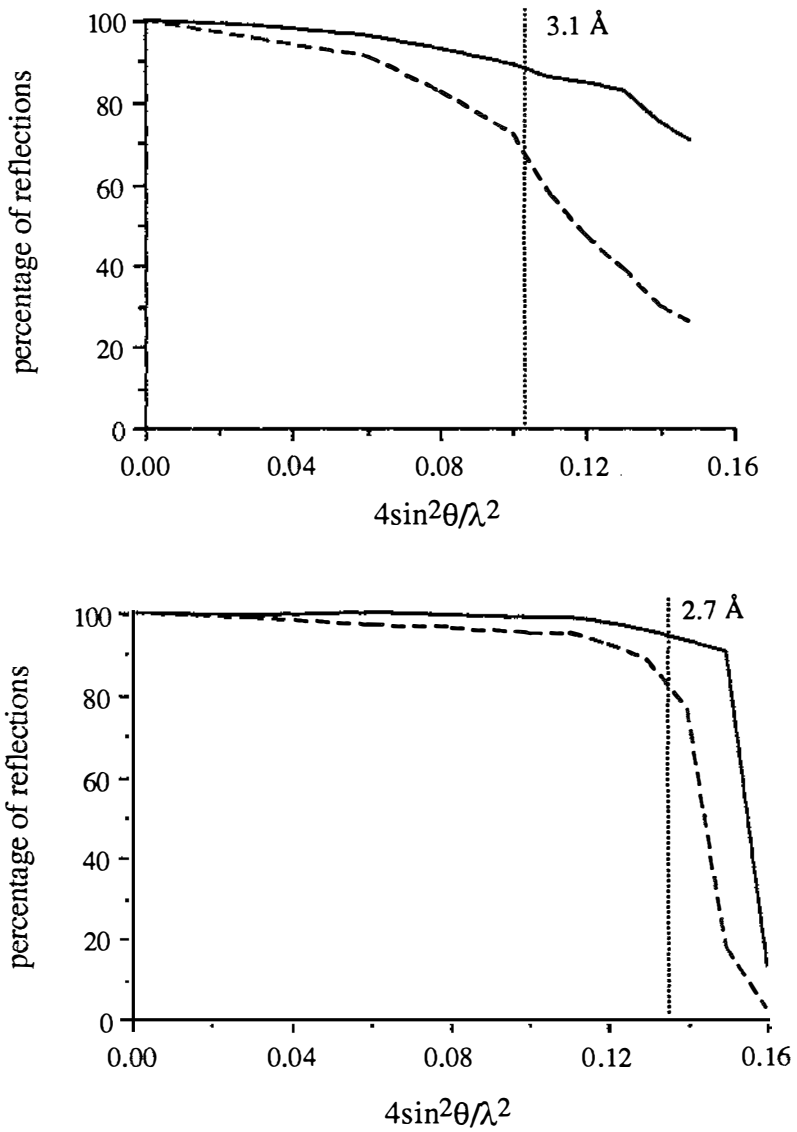



Figure 3.2: Plot of the percentage of unique reflections with $I > 1\sigma_I$ (——) and $I > 2\sigma_I$ (- - - -) as a function of resolution in units of $4\sin^2\theta / \lambda^2$ for the diffractometer data sets of (a) $Cu_2Lf$ and (b) $Cu_2oxLf$. The data were separated into thin shells and the percentages calculated from the total numbers of reflections in each shell.

An absorption correction was not applied to the data because as the reflections recorded on each filmpack come from only a very small region of reciprocal space, the variation in absorption between the two limits of the oscillation range will be negligible. Any differences

in absorption between adjacent filmpacks will be corrected for in the scaling process (Blundell & Johnson, 1976). A correction for the oblique incidence of the X-ray beam and the resulting increase in pathlength through the pack of three films was also made at this stage. Scale factors between the three films of each filmpack were calculated using common reflections (see Tables 8 and 9, Appendix II) and the data were scaled and merged to give 57 independent data files.

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, as a significant amount of the measured data can be lost at the scaling and merging stage if the partial reflections are not treated correctly. The postrefinement strategy consisted of initial refinement of the missetting angles, followed by the cell parameters and missetting angles together and finally the beam characteristics, cell dimensions and missetting. The 228,893 reflections, including 117,529 fully recorded and 64,087 partially recorded (the remainder rejected as being partially recorded at both ends of the range, too near the rotation axis or having calculated partialities less than zero), were scaled and merged using the CCP4 programs ROTAVATA and AGROVATA (using the algorithms of Fox & Holmes, 1966) into one data set of 44766 unique, non-zero reflections with an overall $R_{merge}$ of 0.096 on intensities ($R_{merge} = \Sigma |I - \bar{I}| / \Sigma \bar{I}$). Relevant statistics relating to the data collection can be found in Table 3.7 (page 57), with a more detailed summary in Tables 8, 9 and 10, Appendix II.

This data set was then scaled and merged ($R_{merge} = 0.089$ on intensities) with the diffractometer data, resulting in a final data set containing 46683 independent reflections with 32091 having $I > 2\sigma_I$. Figure 3.3 shows the percentage of the total number of unique reflections measured with $I > 1\sigma_I$ and with $I > 2\sigma_I$ as a function of $4\sin^2\theta/\lambda^2$. There is a sharp decrease in the percentage of observed reflections (with $I > 2\sigma_I$) at about 2.7 Å and the maximum effective resolution of the data (that used for least squares refinement) can be taken as 2.1 Å, where about 25% of the reflections can be deemed as observed (with $I > 2\sigma_I$).

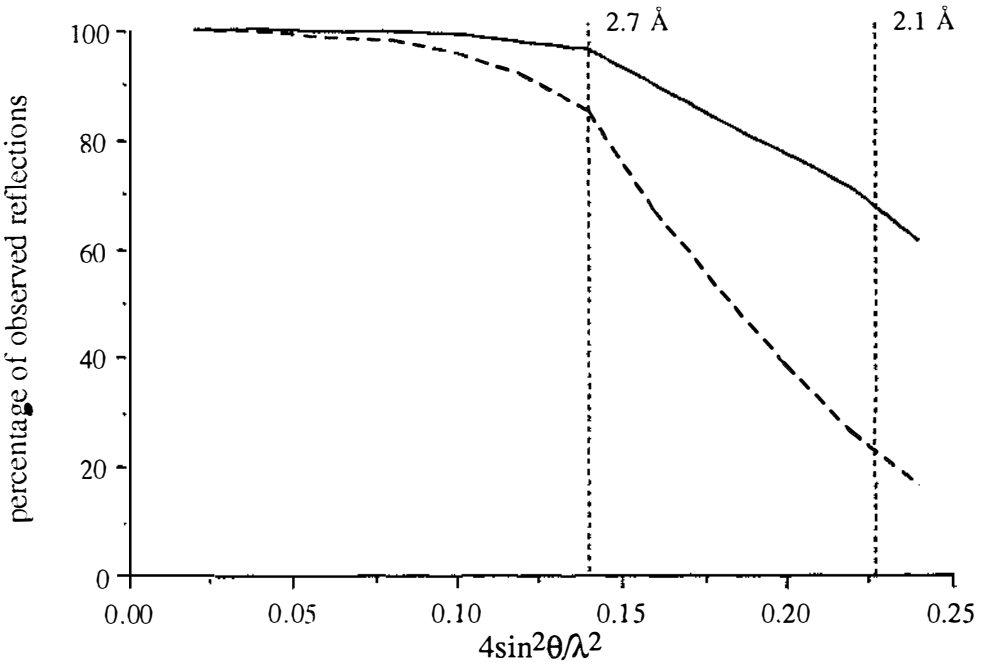



Figure 3.3: The percentage of the total number of unique reflections in each resolution shell (see Table 15 in Appendix II) in the final data set for $Cu_2Lf$, with $I > 1\sigma_I$ (——) and $I > 2\sigma_I$ (- - - -), as a function of $4\sin^2\theta / \lambda^2$.

*$Cu_2oxLf$:* A data set to 2.0 Å resolution was collected from two $Cu_2$oxLf crystals at Daresbury, again using radiation of 0.88 Å produced by the Wiggler. Fifty filmpacks of three films each were measured to cover 90° of rotation about the a* axis. A total of 240281 reflections were extracted from the films using the programs as described above. Of these, 176091 reflections were used in the scaling procedure, giving a final data set comprising 49994 unique, non-zero reflections with an $R_{merge}$ on intensities of 0.078.

Postrefinement of the cell parameters, crystal missetting and beam characteristics was attempted in order to obtain better estimates for the partially-recorded data. It was noted that several batches of data appeared to get much worse upon postrefinement, in that a large number of previously fully-recorded reflections became partials. The best possible data set from postrefinement was scaled using ROTAVATA and AGROVATA (CCP4 suite, Daresbury, UK), resulting in a an $R_{merge}$ of 0.086. This data was discarded in favour of the original data set obtained earlier, as the postrefinement clearly had not improved it. Statistics relating to the collection of the photographic data are summarized in Table 3.7 and given in more detail in Tables 11 - 14, Appendix II.

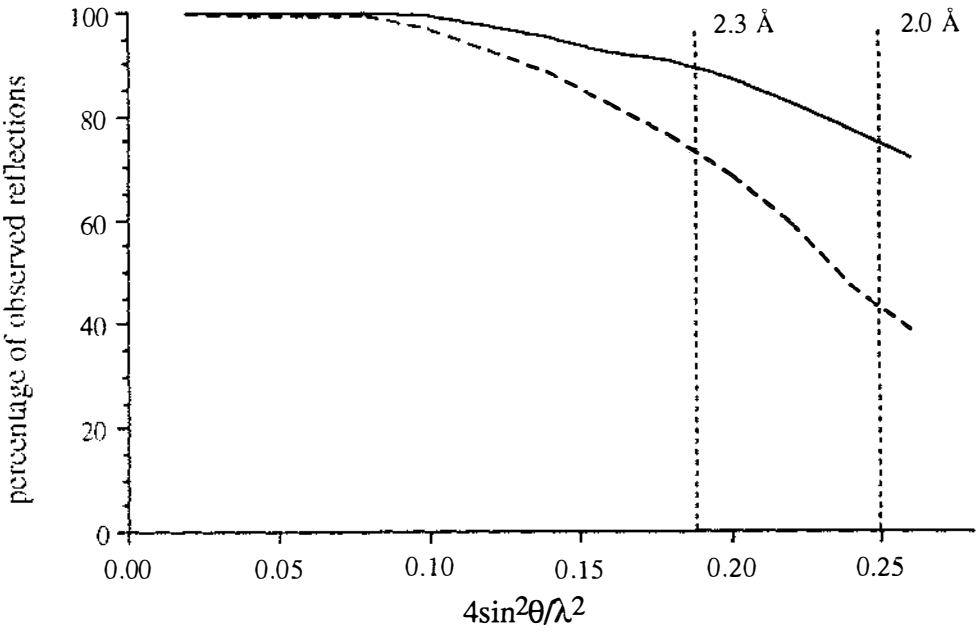



Figure 3.4: The percentage of the total number of unique reflections in each resolution shell (see Table 16 in Appendix II) in the final data set for $Cu_2$oxLf, with $I > 1\sigma_I$ (——) and $I > 2\sigma_I$ (- - - -) as a function of $4\sin^2\theta / \lambda^2$.

This data set was finally merged with the diffractometer data, although as previously mentioned, the reflections from crystal 4 were omitted on the premise that the photographic data covering crystal 4 in medium resolution range (between 2.5 and 2.75 Å) would be of much better quality than that derived from the diffractometer data, which was very weak and scaled to the remainder of the diffractometer data poorly. This resulted in a data set containing 51024 unique reflections (44000 with $I > 2\sigma_I$) with an $R_{merge} = 0.077$ on intensities.

Figure 3.4 shows the percentage of the unique data with $I > 1\sigma_I$ and $I > 2\sigma_I$ as a function of $4\sin^2\theta/\lambda^2$. There is a gradual fall in the numbers of observed reflection (with $I > 2\sigma_I$) at about 3.2 Å, although there are still at least 75% observed at 2.3 Å. The maximum effective resolution of the data (used for subsequent refinement) was taken at 2.1 Å (≈45% observed). Statistics on the final data sets for both $Cu_2$Lf and $Cu_2$oxLf are given in Table 3.8.

Table 3.7: Data collection (oscillation photography)

| | $Cu_2Lf$ | $Cu_2oxLf$ |
|---|---|---|
| wavelength of radiation | 0.88 Å | 0.88 Å |
| number of film packs | 57 | 50 |
| number of films per pack | 3 | 3 |
| estimated mosaic spread | 0.30° | 0.25° |
| average $R_{sca}$ [a] | 0.073 | 0.033 |
| average $R_{sym}$ [a] | 0.068 | 0.072 |
| $R_{merge}$ [a] | 0.096 | 0.078 |
| total reflections measured | 228893 | 240281 |
| independent reflections (non-zero) | 44766 | 49994 |

[a] $R = \Sigma |I - \bar{I}| / \Sigma \bar{I}$; $R_{sca}$ measures the agreement between reflections recorded on successive films of a given filmpack, $R_{sym}$ measures the agreement between symmetry-related reflections on the same film and $R_{merge}$ gives the overall agreement between intensities measured on different films.

Table 3.8: Statistics on final data sets

| | $Cu_2Lf$ | $Cu_2oxLf$ |
|---|---|---|
| Independent reflections | 46683 | 51024 |
| - total observed [a] | 32091 (69%) | 40144 (79%) |
| Maximum measured resolution, $d_{min}$ (Å) [b] | 2.04 | 1.98 |
| - theoretical number of reflections to $d_{min}$ | 52256 | 56830 |
| - overall completeness of data to $d_{min}$ | 89% | 90% |
| Maximum resolution used in refinement, $d_{ref}$ | 2.1 | 2.1 |
| - theoretical number of reflections to $d_{ref}$ | 47903 | 47634 |
| - total reflections to $d_{ref}$ | 44715 | 45394 |
| - completeness of data to $d_{ref}$ | 93% | 95% |
| - total observed to $d_{ref}$ [a] | 31818 (71%) | 37788 (83%) |
| Highest resolution shell (Å) | 2.40 to 2.10 | 2.30 to 2.00 |
| - total data in this shell | 12468 | 14223 |
| - observed data [a] | 4801 (33%) | 7598 (53%) |

[a] With $I > 2\sigma_I$. [b] The resolution of the reflection(s) measured at the highest value of $\theta$.

### 3.1.7 *$Cu_2Lf$ structure refinement:*

The crystals of $Cu_2Lf$ and $Cu_2oxLf$ were isomorphous with $Fe_2Lf$ with respect to their unit cell dimensions (see Tables 2.5 and 3.1 for a comparison), diffraction patterns (Plates 3.2a - d) and space group, $P2_12_12_1$. This being the case, it was probable that the tertiary structures of $Cu_2Lf$ and $Cu_2oxLf$ would be similar to that of $Fe_2Lf$ in terms of the overall folding of the polypeptide chain, the spatial orientation of the majority of the amino acid side-chains and the crystal packing. The major differences between the four structures were likely to be in the vicinity of the two metal binding sites and/or in the relative orientations of the four domains. In order to eliminate any bias in the electron density, therefore, all metal, anion, ligand and solvent atoms were removed from the starting model for both $Cu_2Lf$ and $Cu_2oxLf$.

*Refinement strategy:* Refinement of $Cu_2Lf$ was performed using the FFT version of the restrained least squares refinement program PROLSQ (Hendrickson & Konnert, 1980). The starting model was derived from the atomic coordinates of $Fe_2Lf$, refined at 2.2 Å resolution (M. Haridas, B.F. Anderson & E.N. Baker, manuscript in preparation). The two $Fe^{3+}$ ions were removed, as were the two carbonate anions and the solvent molecules. The side chains of the eight residues which served as the protein ligands were truncated to $C_\beta$. Thermal parameters for all atoms were set to a uniform value of 25 $Å^2$.

The initial R factor was 0.360 for data in the resolution range 5.0 and 2.1 Å (using 40785 reflections with $I > 0.5\sigma_I$). Restraints were placed on bond lengths, angle (1-3) distances, planar groups, chiral volumes, temperature factors, torsion angles and van der Waals contacts. The values of the constraints used are given in Table 3.9. Refinement proceeded in a series of "phases", each involving a number of cycles of least squares (typically 20 - 30) of atomic positions and individual isotropic temperature factors, followed by manual rebuilding of the model on an Evans & Sutherland PS 300 interactive graphics system using the program FRODO (Jones, 1978). The refinement strategy consisted of loosening the weighting on the geometrical constraints (see Table 3.9) and then slowly tightening them back to their initial values over 12-15 cycles. The crystallographic 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.

*The course of the refinement:* After 26 cycles (R = 0.301) the metal ions and protein ligands were added into the appropriate $2F_o$-$F_c$ density, with the anions built in to the corresponding $2F_o$-$F_c$ and $F_o$-$F_c$ density after a further 27 cycles (R = 0.276). 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 neighbouring groups and were not close to any parts of the structure whose conformation was in doubt. At regular intervals, solvent molecules with high B values or too-close contacts, as well as a number of poorly defined loop regions 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.

Table 3.9: Geometrical constraints applied in PROLSQ refinement of $Cu_2$Lf and $Cu_2$oxLf

| Geometrical variable | Initial constraint | Final constraint |
|---|---|---|
| Bonding distances | 0.05 Å | — [a] |
| Angle (1 - 3) distances | 0.04 Å | 0.045 Å |
| Planes | 0.02 Å | 0.02 Å |
| Torsion angles: | | |
| - planar (peptide omega) | 3° | 3° |
| - sidechain Chi1 | 15° | 15° |
| Chiral volumes | 0.15 $Å^3$ | 0.13 $Å^3$ [b] |
| van der Waals interactions: | | |
| - single torsion contacts | 0.5 Å | 0.5 Å |
| - multiple torsion contacts | 0.5 Å | 0.5 Å |
| Temperature factors: | | |
| - main chain bond | 4 $Å^2$ | 3 $Å^2$ |
| - main chain (1-3) angle | 6 $Å^2$ | 6 $Å^2$ |
| - side chain bond | 4 $Å^2$ | 4 $Å^2$ |
| Shift restraints: | | |
| - positional | 0.3 Å | 0.2 Å |
| - thermal | 3 $Å^2$ | 3 $Å^2$ |

[a] No constraints were applied to the metal and ligand bond distances in the final stages of the refinement.
[b] Chiral volume constraint altered in $Cu_2$Lf refinement only.

The carbonate anions were also treated in this way several times during the course of the refinement, by the calculation of omit maps with either one or both anions removed prior to the final 5 - 8 cycles of least squares refinement in a particular phase. It was noted that one result of this refinement was that the copper atoms tended to move away from the protein ligands and into the density previously occupied by the anion. In order to prevent the occurence of this type of movement, much tighter constraints were placed on the metal - ligand distances ($\sigma = 0.02$) during such least squares refinement cycles compared to phases where the anion was present. In addition, the target distance for the copper - histidine bond (which is directly opposite the carbonate position) was shortened by about 0.1 Å.

Constraints ($\sigma = 0.05$ Å) were placed on the Cu - ligand bond distances when they were first added to the structure and shifts in the bond lengths were monitored and the target values altered to take into account the direction of movement of these shifts. The constraints were loosened as refinement progressed and in the final phase, none were applied. Isotropic atomic temperature factors were refined for 9 - 12 cycles in each phase. Restraints were applied to the B values in order to prevent excessive shifts during refinement and to maintain the B values of adjacent atoms to within 5 $Å^2$ of each other. Temperature factors for the solvent molecules were treated in a similar way as the protein atoms. New water molecules were given B values of 25 $Å^2$ and monitored after refinement. Those which showed drastic increases in B values were checked on the graphics system and removed if in doubtful positions.

Some sugar residues of the two glycan chains could be located in later $F_O$-$F_C$ maps and five such residues were added to the structure. One N-acetylglucosamine (NAG) moiety was found attached to each of the side chains of Asn 137 and Asn 478. A fucose (FUC) molecule was identified attached covalently to each NAG via an $\alpha$-type linkage, $O_4$(NAG) - $C_1$(FUC), with a second NAG (located in the N-lobe), bonded to the asparagine-linked NAG with a $\beta$-lionkage ($O_6$ - $C_1$). The Asn - NAG, NAG - FUC and NAG - NAG linkage distances were constrained at 1.4 Å with the inter-sugar C1-On-Cn (in a 1 - n linkage) angle restrained at 114° (120° for the NAG - Asn linkages).

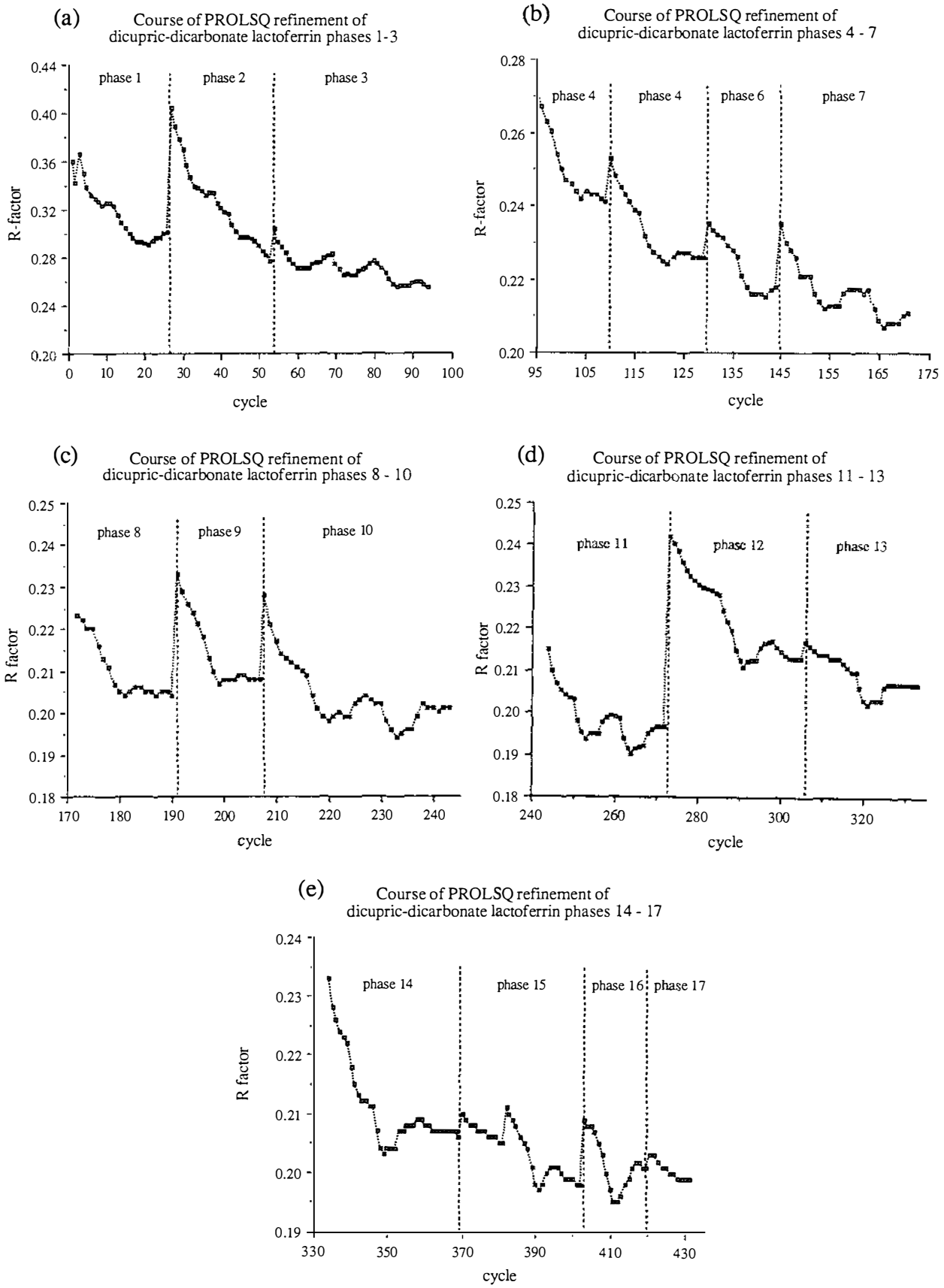



Figure 3.5: Plots of R-factor against least squares refinement cycle for phases (a) 1 -3, (b) 4 - 7, (c) 8 - 10, (d) 11 - 13 and (e) 14 - 17.

After about 430 cycles of PROLSQ refinement (17 phases of refinement and model building on a Evans and Sutherland PS300 graphics system) the R factor is 0.199 for data with $I > 0.5\sigma_I$ (43494 reflections) between 7.5 and 2.1 Å resolution. The course of the refinement is represented by Figures 3.5a - e and Table 3.10 summarises the refinement statistics. Tables 1 - 17 in Appendix III give more detailed statistics on the individual phases of the refinement. Some representative sections of the electron density for the final model are shown in Plates 3.5, 3.6b, 3.7, 3.8, 3.9 and 3.10.

### 3.1.8 *$Cu_2$oxLf structure refinement:*

*The initial model:* The refinement of $Cu_2$oxLf was also carried out using PROLSQ. The starting model for $Cu_2$oxLf was derived from the partially-refined atomic coordinates of $Cu_2$Lf with the two $Cu^{2+}$, two $CO_3^{2-}$, solvent molecules and sugar residues removed. The eight ligand sidechains, the anion-binding arginine (residues 121 and 465) and threonine (residues 117 and 461) sidechains were changed to alanine. The individual temperature factors were all given a uniform value of 25Å$^2$. The initial R factor was 0.293 for data between 5.0 and 2.8 Å (15607 reflections with $I > 0.5\sigma_I$), using only the medium resolution diffractometer data.

*The course of the medium resolution refinement:* After the first phase of refinement (22 cycles), the R factor dropped to 0.171 and at this stage the N-lobe copper atom and the four ligand sidechains were added into $2F_o$-$F_c$ density. The metal binding site in the C-lobe was left vacant, with the four ligand sidechains remaining truncated at $C_\beta$. After a further 25 cycles of least squares refinement (R = 0.177) the C-lobe copper and the four ligands were added into the appropriate $2F_o$-$F_c$ density.

At this stage, the electron density associated with the anions could be seen clearly and a (bi)carbonate and an oxalate were added into the corresponding density in the N- and C-terminal sites respectively, along with the anion-binding arginine and threonine residues.

The resolution was gradually increased to 2.5 Å (20064 reflections with $I > 0.5\sigma_I$) over 95 cycles of refinement (4 phases in total) and resulted in a model with an R factor of 0.170 with reasonable geometry (rms deviation from standard bond distances of 0.022 Å).

*The course of the high resolution refinement:* At this stage, refinement was switched to the oscillation film data set, the starting model being the partially-refined $Cu_2$oxLf model with deletions and alterations as described above for the first starting model derived from $Cu_2$Lf. The thermal parameters were once again assigned a uniform value of 25 $Å^2$. The initial value of the R factor was 0.243. The increase in R by about 7% cannot be attributed solely to the changes made in the positional and thermal parameters of the model, but is also associated with the biasing of the model towards the weaker diffractometer data. At 2.5 Å, only a small proportion (≈30% - see Figure 3.2b) of the reflections in the diffractometer data set were observed (with $I > 2\sigma_I$) and the positional and thermal parameters of the model refined to this data would reflect this inaccuracy. Upon changing to the photographic data, where the data at 2.5 Å resolution is much stronger (>80% observed - see Figure 3.4), the starting model derived from the medium resolution refinement will not adequately reflect the increased accuracy of the observed data and hence the $F_o$ and $F_c$ values deviate to a much greater extent.

The metal ions, metal ligand sidechains and the anions were added after 28 cycles of least squares refinement (Phase 5; R = 0.250; data from 5.0 to 2.5 Å) using both $2F_o$-$F_c$ and $F_o$-$F_c$ electron density maps. In addition, $F_o$-$F_c$ electron density corresponding to the positions of the asparagine-linked NAG moiety at residue 137 and the fucose residue linked to the NAG was observed and two such molecules were added to the model. No density for the second NAG in this site or the sugar residues in the C-lobe could be seen.

In order to hasten the process of water molecule location, the refined diferric and dicupric structures were superimposed upon each other and the solvent molecules which appeared in conserved positions (within ≈ 1.0 Å) were presumed to be common in the $Cu_2$oxLf structure also (see Section 3.2.5). The positions of 123 solvent molecules were checked in the $Cu_2$oxLf structure to make sure that they corresponded to positive $F_o$-$F_c$ electron density

peaks and made reasonable hydrogen bonds with either protein atoms or other water molecules. Sites for which there was no electron density or which made too-close contacts with regions of the protein structure were removed. Additional solvent molecules were added as previously described, bringing the total to 237.

Table 3.10 : Refinement statistics for $Cu_2Lf$ and $Cu_2oxLf$ [a]

| | $Cu_2Lf$ | $Cu_2oxLf$ |
|---|---|---|
| resolution limits (Å) | 7.5 - 2.1 | 7.5 - 2.1 |
| initial R factor (%) | 0.360 | 0.243 [a] |
| final R factor (%) | 0.199 | 0.212 |
| number of reflections used [b] | 43494 | 48224 |
| number of protein atoms | 5350 | 5323 |
| number of solvent molecules | 301 | 237 |
| other ions | 2 $Cu^{2+}$, 2 $CO_3^{2-}$ | 2 $Cu^{2+}$, 1 $CO_3^{2-}$, 1 $C_2O_4^{2-}$ |
| number of carbohydrate molecules | 3 NAG, 2 FUC | 1 NAG, 1 FUC |
| number of refinement cycles | 430 | 208 |
| rms deviations from ideality | | |
| bond length (Å) | 0.014 | 0.018 |
| angle distance (Å) | 0.058 | 0.064 |
| planarity (Å) | 0.013 | 0.013 |
| chiral volume ($Å^3$) | 0.190 | 0.212 |
| van der Waals contacts (Å) | 0.229 | 0.249 |
| average B value (protein atoms) | 44.6 | 35.6 |
| average B value (all atoms) | 45.7 | 36.1 |
| overall shift in position (final) (Å) | 0.003 | 0.003 |
| overall shift in B values (final) ($Å^2$) | 0.34 | 0.55 |

[a] For high resolution refinement (using the photographic data).

Prior to the 6th phase of refinement, the resolution range of the data was extended to 5.0 - 2.2 Å and the R factor increased to 0.282. After a further 32 cycles of least squares refinement (Phase 6; R = 0.225) and manual model rebuilding, the data set derived from the oscillation photographs was replaced by a data set obtained by the scaling and merging of the diffractometer and the photographic data sets. The R factor at the beginning of phase 7 was 0.235. Compared with the value at the end of the previous phase (and in light of the fact that

a further 25 solvent molecules had been added during the rebuilding stage), this indicated that the combined data set in the 5.0 - 2.2Å resolution range was essentially similar to the photographic data set. The resolution range was extended to 7.5 - 2.2 Å after 6 cycles (R = 0.223) and the model was refined for a further 25 cycles (Phase 7; R= 0.221). An eight phase (initial R = 0.223) of least squares refinement, in which the resolution was extended to 2.1 Å (48224 reflections), resulted the present model of $Cu_2$oxLf with an R factor of 0.212.

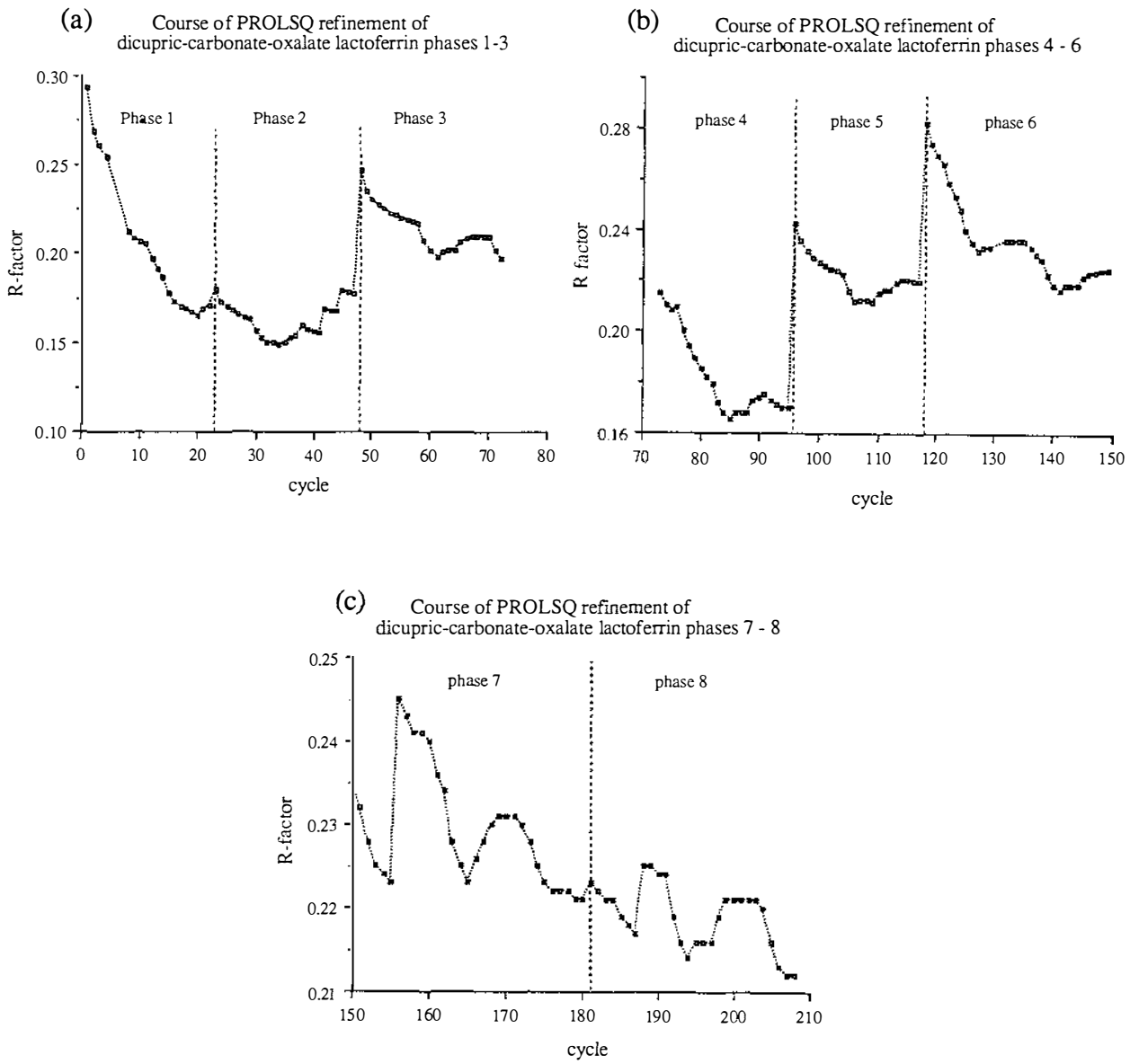



Figure 3.6: Plots of R-factor against least squares refinement cycle for $Cu_2$oxLf for phases (a) 1 -3, (b) 4 - 6 and (c) 7 - 8.

Constraints on the copper - ligand distances were applied in much the same way as for $Cu_2Lf$. Initial target values were chosen based on the metal - ligand distances observed when the ligands were first included in the model. The constraint on these bond distances was set at 0.05 Å. During the course of the refinement, the target values were altered in accordance with the observed shifts in the parameters, while the constraint was loosened to 0.10 Å. The refinement strategy as a whole was identical to that employed with $Cu_2Lf$, consisting essentially of a loosening of the geometry allowing a better fit to the observed data, followed by a slow tightening, interspersed with cycles of individual temperature factor refinement.

After 208 cycles of least squares refinement with PROLSQ, the R factor of the model was 0.212 for data with $I > 1.0\sigma_I$ (48224 reflections) between 7.5 and 2.1 Å resolution. Refinement statistics can be found in Table 3.10 and the course of the refinement is summarised in Tables 18 - 25 in Appendix III and Figures 3.6a - c.

## 3.2 Results; A: Dicupric-dicarbonato lactoferrin

### 3.2.1 *The model:*

The final model for $Cu_2Lf$ consists of 5350 protein atoms (691 amino acid residues), 301 solvent molecules, 62 carbohydrate atoms (three N-acetylglucosamine and two fucose residues), two $Cu^{2+}$ ions and two $CO_3^{2-}$ ions. Figure 3.7 shows a schematic representation of the polypeptide chain conformation ($C_\alpha$ plot) of the $Cu_2Lf$ molecule. It can clearly be seen that the molecule is divided into two halves, the N-terminal and C-terminal lobes. Each lobe can be further divided into two domains, as described for the $Fe_2Lf$ structure (section 1.2.5), the N1 domain consisting of residues 1 - 90 & 252 - 320; N2 residues 91 - 251; C1 residues 245 - 433 & 596 - 663 and C2 residues 434 - 595.

The copper binding sites are at the interface between these domains (one in each lobe) at the end of a deep solvent-filled cleft. Each $Cu^{2+}$ ion is approximately 12 - 15 Å from the protein surface. The overall folding of the polypeptide chain is virtually identical to that observed in the diferric molecule, the details of which have been previously described (Anderson *et al.*, 1989).

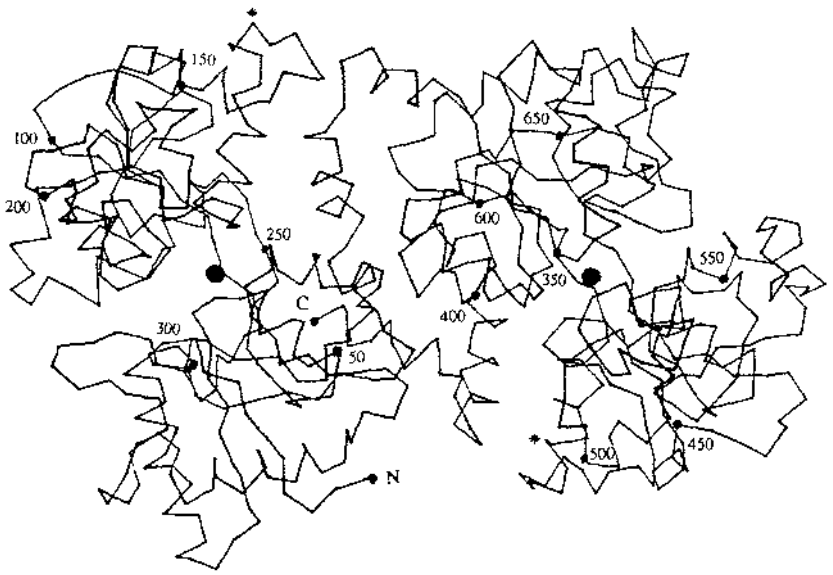



Figure 3.7: $C_\alpha$ plot of dicupric lactoferrin (N-lobe on the left and C-lobe on the right). The positions of the two $Cu^{2+}$ ions are given by ●. The N- and C-termini are indicated by N and C respectively, while the positions at which the glycan chains are attached are indicated by ∗. Every 50th residue is numbered.

The stereochemistry of the final model is close to ideal, with rms deviations of 0.014 Å and 0.058 Å from standard values of bond lengths and angle (1-3) distances respectively.

The maximum average error in the atomic positions is estimated, from a Luzzati plot (Figure 3.8) at about 0.30 Å, excluding the points at the lowest resolution, where the disordered solvent has not been adequately modelled, and at the highest resolution where a large proportion of the data (see Figure 3.3) is unobserved. The theoretical variation of R factor with resolution was calculated for non-centric data.

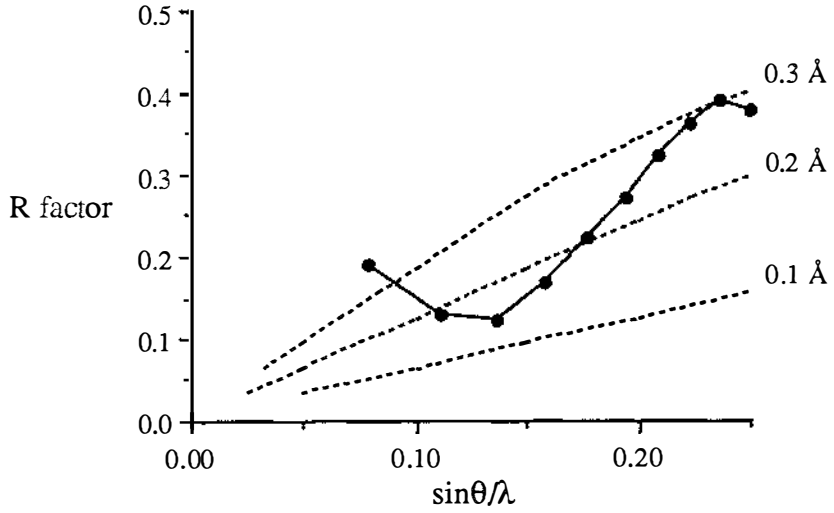



Figure 3.8: A Luzzati-type plot of resolution ($\sin\theta/\lambda$) against the crystallographic R factor (Luzzati, 1952) for $Cu_2Lf$. The theoretical variation of R with resolution for rms errors in atomic positions of 0.1, 0.2 and 0.3 Å, are shown for non-centric data.

The actual error is probably somewhat less than this in well-defined regions (the area around the binding sites and the secondary structure elements for example) but substantially more in the poorly-defined parts of the molecule, identified as the areas with high individual atomic temperature factors for mainchain atoms (see Table 3.14). Although every effort was made to determine their correct conformation with the use of omit maps, they remain ill-defined in relation to the rest of the structure.

Real space correlation coefficients for the individual residues were calculated using the real space fit routine implimented in the model building and refinement program O (Jones *et al.*, 1991). Figures 3.9a & b show the real space correlation coefficients for all atoms in the 691 residues of $Cu_2Lf$. The residues which show poorest correlation (< 0.60) between observed and calculated electron density include the N-terminus (residues 1-4) and external loops at 84 - 88, 220 - 222, 240 - 243, 280 - 285 and 291 - 294 in the N-lobe, and loops at 417 - 424,

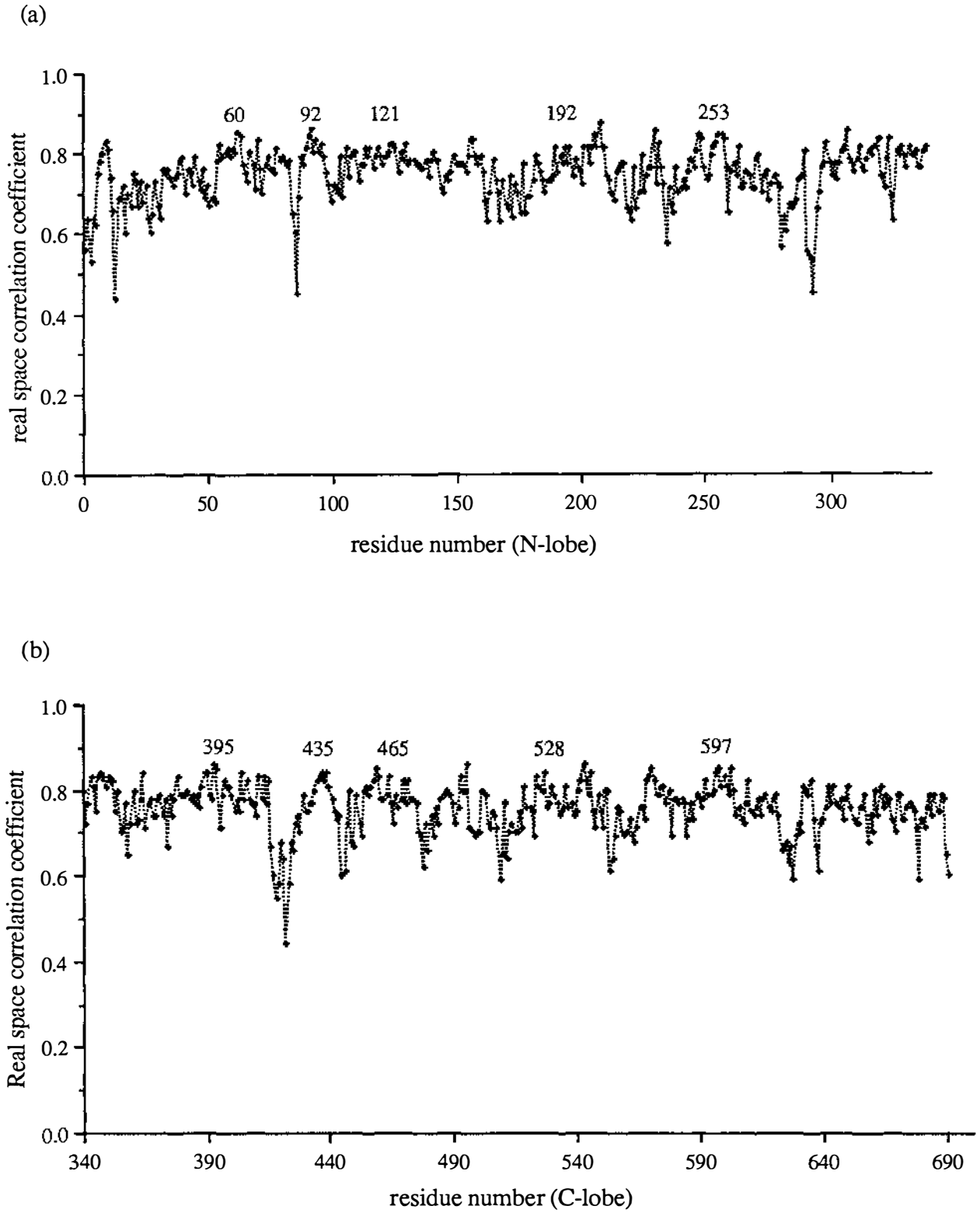



Figure 3.9: Plot of the real-space correlation coefficient for the fit of all atoms of each residue in the (a) N-lobe and (b) C-lobe of the final model of $Cu_2Lf$ to the final $2F_O$-$F_C$ electron density. The calculations were performed using the interactive graphics program O (Jones *et al.*, 1991).

445 - 446, 509 - 512, 623 - 628 and the C-terminus (residues 687 - 691). Analysis of the parts of the model which have high temperature factors (see Table 3.14B) and the regions which have the largest deviations in $C_\alpha$ and mainchain atom positions between $Cu_2Lf$ and $Fe_2Lf$ (see Table 3.20A) are consistent with the results from real space fitting of the model to the electron density. In addition, the well-defined regions of the model, including the metal and anion binding ligands, and those residues near the metal binding sites all show good correlation between observed and calculated density (correlation coefficients between 0.70 and 0.85). Some of these residues are indicated in Figures 3.9a & b.

Another way of estimating the error in the $Cu_2Lf$ model comes from a comparison with $Fe_2Lf$, the structure of which is essentially isomorphous with $Cu_2Lf$. The rms deviation in mainchain atom positions between $Cu_2Lf$ and $Fe_2Lf$, with the N-terminus (1 - 5) and the last 2 residues in the C-terminal helix removed is 0.34 Å. This figure includes the contributions from several poorly-ordered external loops (see Table 3.20 in Section 3.2.8) and from any changes in the positions of the mainchains of the metal and anion binding ligands and hence does not accurately reflect the overall error in the $Cu_2Lf$ model. Removal of these less-ordered regions results in an rms displacement of 0.28 Å. The additional removal of the residues directly involved in metal and anion binding has little effect on the overall rms displacement. This value, therefore, may correspond to the mean error in atomic positions for the moderate to well-defined areas of the model (≈ 90% of the polypeptide chain). Correlation of the rms deviations in mainchain atom positions with the average mainchain B values (see Section 3.2.4) gives an estimate of the mean error in atomic positions in the well-ordered regions of the structure (B between 20 and 40 $Å^2$) at ≈ 0.20 Å. The error in the ill-defined regions (B > 70 $Å^2$) is substantially greater at ≈ 0.60 - 0.70 Å.

Wang *et al.* (1985) and Baker (1988) have used a similar method to look at the structures of the two crystallographically independent molecules of chymotrypsinogen and azurin respectively. However, in both of these cases, the two molecules were in the same asymmetric unit and were not two isomorphous structures which had been crystallised and analysed independently, as is the case with $Cu_2Lf$ and $Fe_2Lf$. It should be noted that the value of 0.28 Å obtained as the rms deviation in mainchain atom positions between $Cu_2Lf$

and $Fe_2$Lf (excluding the N- and C-termini), although slightly higher than the estimated maximum error derived from the Luzzati plot, is not a significant increase. This is in contrast to the case with azurin (Baker, 1988) where the mean error from the refinement (from a Luzzati plot) was 0.15 Å, while the rms deviation in mainchain positions (excluding chain termini) was almost double at 0.27 Å. This difference is due primarily to the fact that the $Cu_2$Lf structure is not entirely independent of the $Fe_2$Lf structure, as the refined $Fe_2$Lf coordinates were used as the starting model for the $Cu_2$Lf structure analysis. In the azurin case, however, the two molecules were determined independently and the larger deviation in mainchain atom positions most probably reflects real differences between the two molecules.

### 3.2.2 *Polypeptide chain folding:*

A Ramachandran plot of the main chain conformational torsion angles ($\phi$ and $\psi$) (Figure 3.10) shows that virtually all lie within the allowed and partially-allowed regions of conformational space, typical of globular proteins. Two of the exceptions (residues 3 & 4 and 87) are in poorly-defined surface regions of the structure, where the poor quality of the electron density maps in these areas makes interpretation of the correct folding of the polypeptide chain very difficult.

Serine 191, although in a well-ordered part of the structure near the N-terminal metal binding site, has, nonetheless, unusual mainchain torsion angles ($\phi$ and $\psi$) of 63° and -174°, designated as an $\varepsilon$ conformation (Sibanda *et al.*, 1989). Although this appears to be an unfavourable conformation (for residues other than glycine), in that it does not lie within the allowed or partially-allowed regions, such a conformation for non-glycinyl residues in the $\varepsilon$ zone have been noted, including an arginine in the loop 130-137 in alcohol dehydrogenase (Englund *et al.*, 1976) and the active serine (Ser144) in the fungal triglyceride lipase from *Rhizomucor miehei* (Derewenda & Derewenda, 1991), where the four residue turn containing this serine has been classified as a $\beta$-hairpin type II' (Sibanda *et al.*, 1989). It is interesting to note that the five residues in $Cu_2$Lf of which Ser191 is the third residue (Tyr189-Phe190-Ser191-Tyr192-Ser193) have $\phi/\psi$ angles very similar to those in the

pentapeptide (of which Ser144 is the third residue) from *R. miehei* and although these residues in $Cu_2Lf$ are not involved in a hairpin turn, they join a section of polypeptide chain consisting of a number of loops (180 - 189) with the N-terminus of an α-helix (190 - 201), with a sharp reversal in the direction of the polypeptide chain.

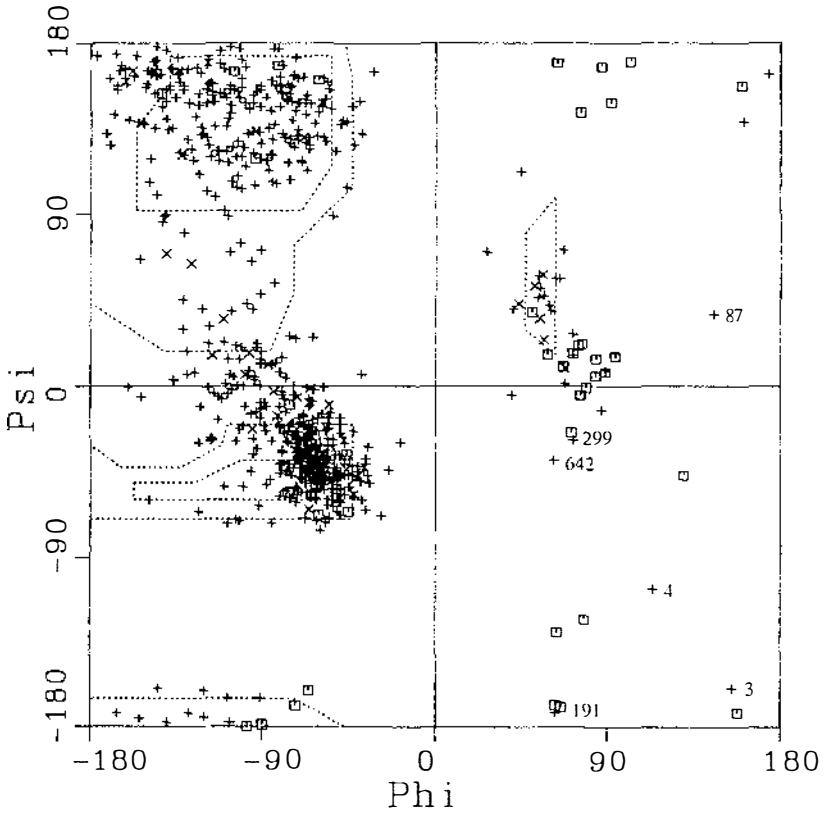



Figure 3.10: A Ramachandran plot of the mainchain torsion angles (ϕ and ψ) for $Cu_2Lf$. The allowed regions are enclosed by solid lines and the partially-allowed regions by dotted lines (Ramakrishnan & Ramachandran, 1965). Residues with ϕ/ψ angles in disallowed regions are indicated.

Morris *et al.* (1992), in their analysis of 462 protein structures from the Brookhaven Protein Data Bank also showed that the ε zone of the Ramachandran plot, although clearly not one of the highly favoured conformations, is an allowed zone, in that a significant number of non-glycine and non-proline residues adopt this conformation.

In $Cu_2Lf$, the residue in the C-lobe which corresponds in position to Ser191 is a glycine (Gly527), which has ϕ/ψ angles (66° and -170°) very similar to those for the serine. In all the other transferrins whose sequences are known, there are glycine residues at this position in both lobes. The Ser191 mainchain appears to have adopted a conformation similar to

Gly527 which is possibly a consequence of the folding in this region, as the two residues on both sides of Gly527 have $\phi/\psi$ angles virtually identical to those residues which surround Ser191. Although the next residue in this sequence is one of the metal-binding ligands, Tyr192 (528), and it is conceivable that the relative immobility of this residue when coordinated to a metal ion may help maintain the less-favourable conformation of this loop, analysis of the apoLf structure shows that both Ser191 and Gly527 have $\phi/\psi$ angles in the $\varepsilon$ zone (56 & -169°, and 67 & -160° respectively) and that the conformation of this turn is independent of the mobility of the tyrosine.

Residues 299 and 642 are leucines involved in $\gamma$-turns (see Figures 3.11a & b). These types of turns are generally characterised by mainchain conformational angles $\phi = 70°$, $\psi = -60°$ (Baker & Hubbard, 1984). In the case of Leu 299, these torsion angles are 68° and -29°, while the corresponding angles for Leu 642 are 60° and -36°. The residues to either side (298 & 300, 641 & 643) interact via 1-3 hydrogen bonds (Baker & Hubbard, 1984), the details of which are described later in Table 3.13A.

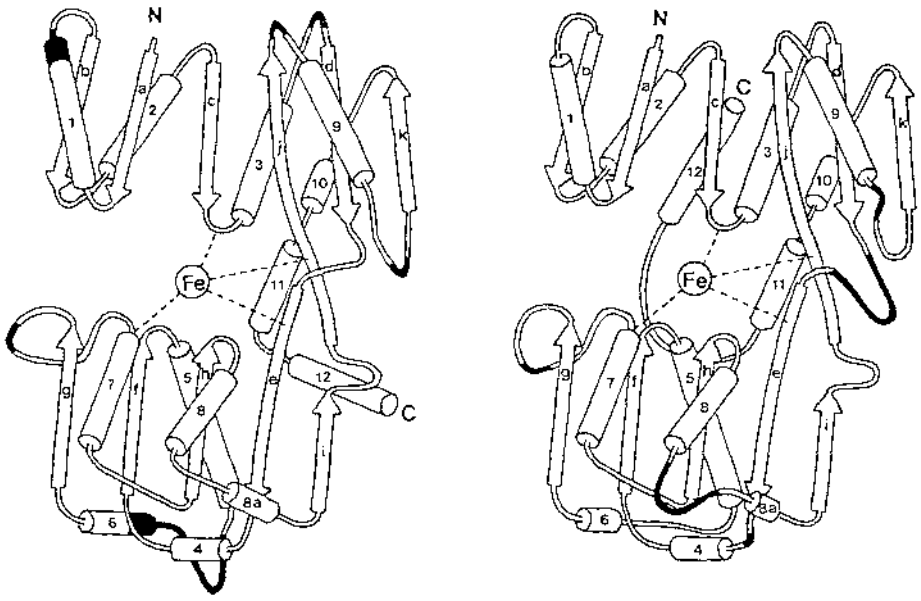
### 3.2.3 *Secondary structure and hydrogen bonding:*

*General description:* The secondary structure observed in $Cu_2Lf$ is the same as that described previously for $Fe_2Lf$ (Anderson *et al.*, 1989), in that about 24% is $\beta$-sheet and 41% helical, with the remainder as turns between $\beta$ strands or helices, regions of extended chain or loops. Table 3.11 lists the residues which make up the $\beta$-sheet and helical structure, while Figures 3.11a and Plates 3.3a & b give schematic representations of the individual lobes of the lactoferrin molecule with the helices shown as cylinders (coils in the photographs) and the $\beta$ strands as arrows.

The $\beta$-sheet in domains N1 and C1 consist of six strands in the order b-a-c-j-d-k (Figure 3.11b). There are four parallel strands (a,b,c,and d) pointed towards the interdomain cleft, all of which occur in the first part of the domain (from the N-terminus to residue 91) before the polypeptide chain crosses into domain 2. The other two domain 1 strands (j and k) are parallel to each other but antiparallel to strands a, b, c & d and point towards the exterior of the molecule. In the N2 and C2 domains there are four parallel strands pointed towards the

cleft (f,g,h and i/j) but only one antiparallel (strand e). This strand, beginning at residue 91 runs away from domain 1 towards the surface of the molecule.

(a)




(b)

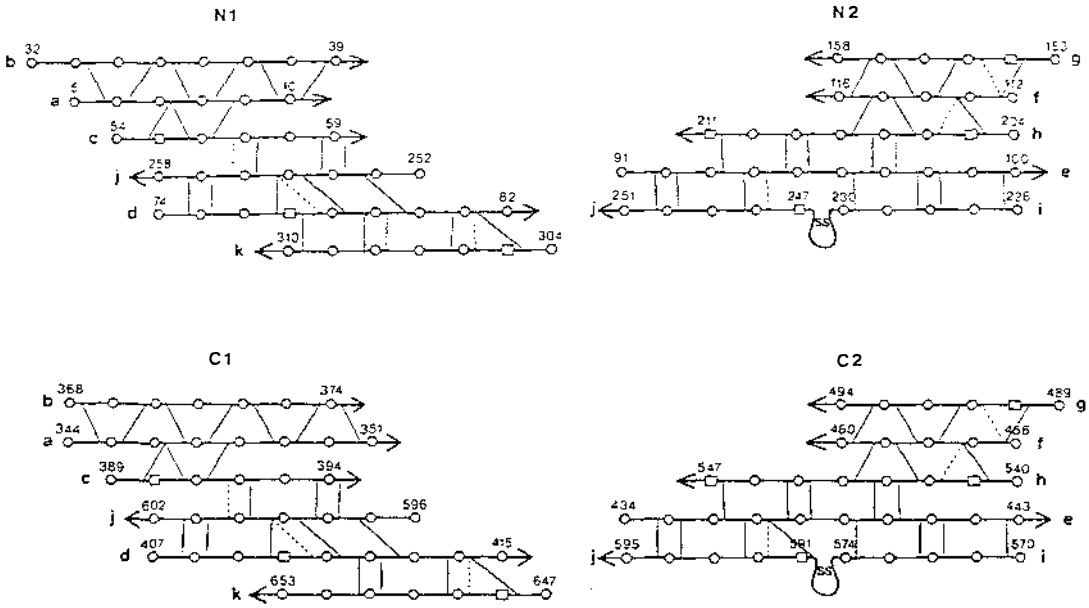



Figure 3.11: (a) Schematic representation of the secondary structure elements in human lactoferrin. Helices are shown as cylinders and β-strands as arrows. Insertions in one lobe or the other are indicated in black. (b) Hydrogen bonding pattern for the b-sheets of the four domains, N1, N2, C1 and C2 of lactoferrin. Hydrogen bonds are shown as thin lines, with positions where the distance is greater than 3.5 Å shown as broken lines. (Both figures adapted from Anderson *et al.* (1989).

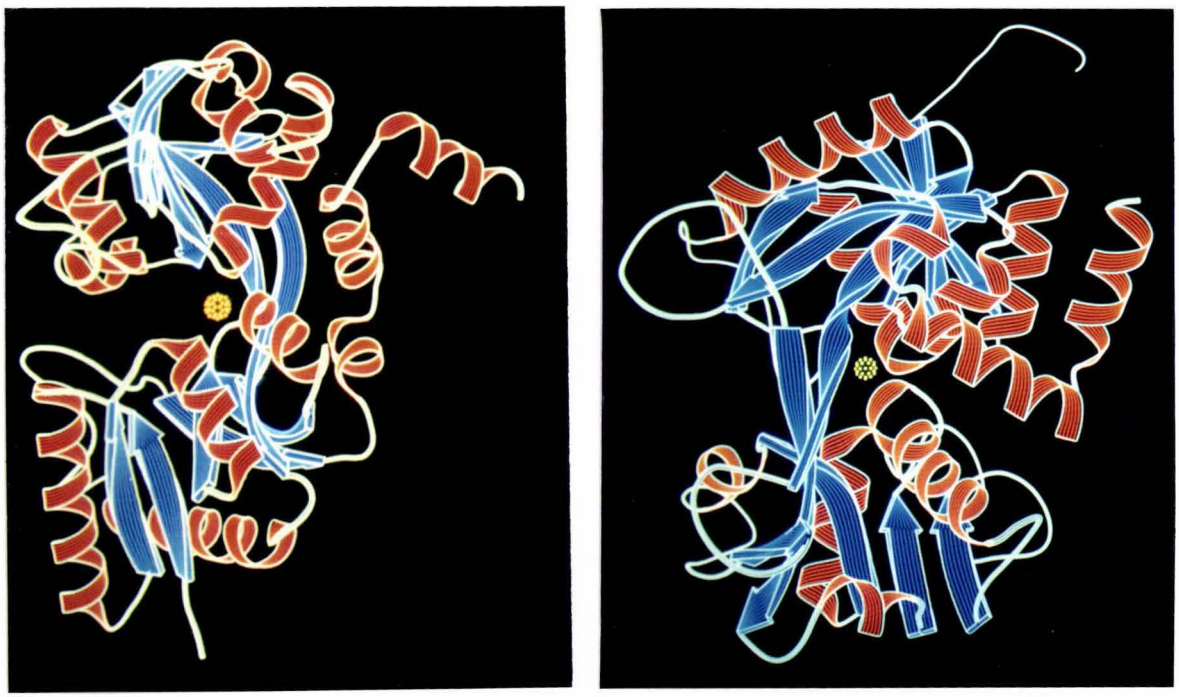


Plate 3.3: Richardson-type schematic representations of the polypeptide chains of the (a) N-lobe and (b) C-lobe of $Cu_2Lf$. The helices are shown in red and the β-strands in blue, with connecting pieces of random coil in white. The yellow spheres represent the positions of the copper atoms. These diagrams were generated with the RIBBON program of Priestle (1988).

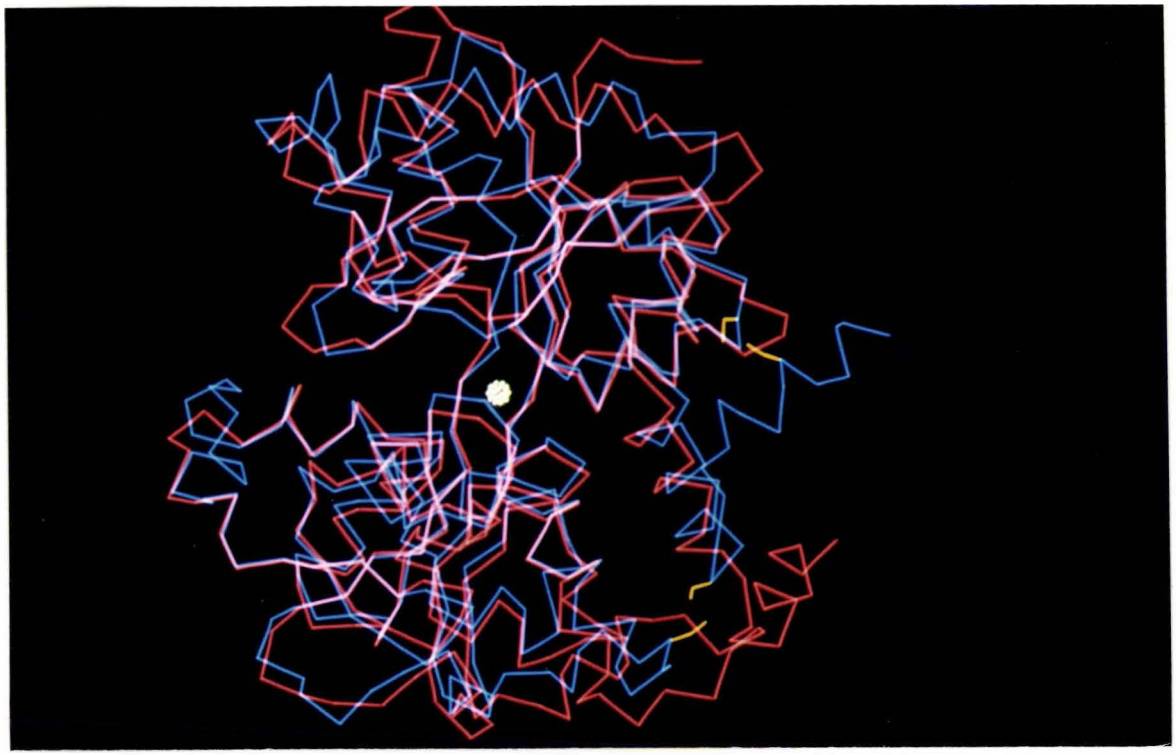


Plate 3.4: The N lobe of $Cu_2Lf$ (red) superimposed on the C-lobe (blue) on the basis of common secondary structure (helices and β-strands). The copper position in the C-lobe is shown as a white circle. Two disulphide bridges present only in the C-lobe are shown in yellow.

Table 3.11: Residue assignments for the secondary structure in $Cu_2Lf$

| β-strands | | | | helices | | | |
|---|---|---|---|---|---|---|---|
| N-lobe residues | C-lobe residues | domain | label [b] | N-lobe residues | C-lobe residues | domain | label [b] |
| 5 - 10 | 344 - 351 | N1, C1 | a | 13 - 31 | 351 - 365 | N1, C1 | 1 |
| 32 - 39 | 368 - 374 | N1, C1 | b | 41 - 53 | 376 - 388 | N1, C1 | 2 |
| 54 - 59 | 389 - 394 | N1, C1 | c | 60 - 69 | 395 - 403 | N1, C1 | 3 |
| 74 - 82 | 407 - 415 | N1, C1 | d | 105 - 110 | 449 - 454 | N2, C2 | 4 |
| 91 - 100 | 434 - 443 | N2, C2 | e | 121 - 136 | 462 - 481 | N2, C2 | 5 |
| 112 - 116 | 456 - 460 | N2, C2 | f | 144 - 152 | 484 - 488 | N2, C2 | 6 |
| 153 - 158 | 489 - 494 | N2, C2 | g | 191 - 203 | 527 - 539 | N2, C2 | 7 |
| 204 - 211 | 540 - 547 | N2, C2 | h | 212 - 218 | 548 - 554 | N2, C2 | 8 |
| 226 - 231 | 570 - 547 | [c] | i | 220 - 226 | 567 - 570 | N2, C2 | 8a |
| 246 - 258 | 590 - 602 | [c] | j | 263 - 279 | 605 - 621 | N1, C1 | 9 |
| 304 - 310 | 647 - 653 | N1, C1 | k | 315 - 321 | 658 - 664 | N1, C1 | 10 |
| | | | | 321 - 333 | 664 - 678 | N1, C1 | 11 |
| | | | | 334 - 344 | 680 - 691 | [d] | 12 |

[a] The β-strands and helices were defined by the observed hydrogen bonding patterns and subsequently checked using the program DSSP (Kabsch & Sander, 1983). [b] The β-strands and helices are labelled as for rabbit serum transferin (Bailey *et al.*, 1988) and diferric human lactoferrin (Anderson *et al.*, 1989). [c] Strand i and the first 5 residues of strand j are in the N2 & C2 domains while the latter 7 residues of strand j are in the N1 & C1 domains. [d] The N-lobe helix 12 joins the two lobes.

In domain 1 the first four parallel β-strands and their 3 connecting helices form an α/β type structure (Chothia & Lesk, 1976) with the helices antiparallel to the strands. This causes the N-termini of the helices to point in towards the interdomain cleft. The polypeptide chain then crosses behind the metal binding site (strand e) and a similar α/β structure makes up the bulk of domain 2. This domain, however, has three extra helices inserted relative to domain 1: helix 4 between strands e and f; helix 6 between helix 5 and strand g and helix 8a between helix 8 and strand i. Notwithstanding these insertions, the α/β structure in domain 2 once again has the effect of orienting helices 5, 7 and 8 in such a way that their N-termini all point towards the interdomain cleft. Finally, the chain crosses back into domain 1 to complete the folding.

The structure of each domain consists, therefore, of: (i) the two α/β sequences which serve to form the major part of each domain and also to direct the positively-charged N-termini of the helices in towards the interdomain cleft (it has been suggested (Anderson *et al.*,

1989) that this arrangement could have a role in attracting the anion into the binding site); (ii) the two strands (e and i/j) which run behind the metal binding site, each having one of the iron binding ligands (Tyr 92 (435) on strand e and His 253 (597) on strand i/j), each contributing to the β-sheets in the two domains and each containing inserted loops which may give some flexibility in the opening and closing of the lobe (Anderson *et al.*, 1989) and (iii) the final 80 or so residues which complete the folding of domain 1 and lead to the final helix (in the N-lobe, this is the connecting helix, while in the C-lobe it is the C-terminal helix).

*Differences in secondary structure between the two lobes:* The only major folding differences between the two lobes are in the orientation of the last helix (12) and a small number of insertions and deletions in the polypeptide chain. These are generally in the loop or turn regions between elements of secondary structure, although helices 1 and 6 in the N-lobe both have 3 or 4 extra residues than their counterparts in the C-lobe (helix 1 has a longer C-terminus and helix 6 has the insertion at the N-terminus in addition to an extra 5 residues in the loop connecting with helix 5). Two other significant insertions occur in the connecting peptide between strands d and e in the C-lobe, where there are an extra 10 residues, and an extra 8 residues between helices 8 and 8a, also in the C-lobe. Both of these extended loops (416 - 424 and 555 - 564) are very flexible (see Tables 3.14 and 3.20).

The other difference between the two halves of the molecule is in the orientation of the C-terminus of each lobe. In the N-lobe, helix 12 points away from the metal site, while in the C-lobe the loop between helix 11 and helix 12 has been extended by about 3 residues and helix 12 folds back behind the metal site to lie across the hydrophobic region between the two lobes, in the vicinity of helices 2 and 3. This difference can be clearly seen when the C-lobe is superimposed on to the N-lobe, as shown in Plate 3.4 (page 75). It is interesting to note that helix 12 in the C-lobe is stabilised by two disulphide bridges which are not present in the N-lobe (Cys 405 - Cys 686 and Cys 483 - Cys 677). The locations of these two bridges are indicated in Plate 3.4. Disulphide 483 - 677 is at the C-terminal end of helix 11 and helps to

hold this portion of the polypeptide chain in roughly the same position as in the N-lobe. The second disulphide bridges the end of helix 12 with a loop between helix 3 and strand d.

The rms deviation in the positions of the $C_\alpha$ atoms between the two lobes is 1.37 Å, substantially greater than that between the same lobes of different structures (see later in this section for a more detailed analysis). This reflects the sequence changes between the two lobes (as noted above), although there is a very good alignment of the secondary structure elements, supporting the view that the ancestor of lactoferrin (and the transferrins in general) was a single-sited protein, equivalent in size and with a folding pattern similar to one of the lobes of lactoferrin.

*Hydrogen bonds:* Hydrogen atoms were added to the all the mainchain and sidechain nitrogen atoms in the model in their calculated positions (with the exception of the $N\zeta$ atoms of lysine residues for which the hydrogen atom positions cannot be defined), using a standard N-H bond distance of 1.0 Å, with the groups assumed to be planar (Baker & Hubbard, 1984). Hydrogen atoms were not added to oxygen atoms as their positions cannot be unambiguously defined due to the free rotation about the C-O bonds in amino acids such as serine, threonine and tyrosine. Hydrogen bonds were tagged as being unfavourable if the O···H distances were less than 1.5 Å, greater than 2.5 Å (although it has been suggested that there could still be a significant level of interaction at distances beyond this (Baker & Hubbard, 1984)), or the N - H ··· O (N) angle was less than 90°. Potential hydrogen bonds were then tabulated into various types, including mainchain-mainchain (helix and sheet), mainchain-sidechain, sidechain-sidechain, mainchain-solvent and sidechain-solvent. Table 3.12 gives a summary of the occurence of hydrogen bonding in $Cu_2Lf$.

*Helices:* Analysis of the hydrogen bonding in the helical regions of $Cu_2Lf$ shows that while the majority contain essentially all α-type hydrogen bonds, there are some short pieces of $3_{10}$ helix (helices 4 and 8a in both lobes and helix 6 in the C-lobe), along with somewhat distorted structures such as helix 5, which has a wide loop at the N-terminus. This may have some function in anion binding as pointed out by Anderson *et al* (1989).

Table 3.12: The hydrogen bonding interactions in $Cu_2Lf$

| | | | distances (Å) [a] | | angles (deg) [a] | | |
|---|---|---|---|---|---|---|---|
| type | total | marginal [b] | O···H | O···X | C-O···H | C-O···X | N-H···O |
| m/c O···m/c HN | 375 | 34 | 2.04 | 2.87 | 137 | — | 150 |
| m/c O···s/c X | 109 | 3 | 1.97 | 2.85 [c] | 128 | 118 [c] | 145 |
| m/c NH···s/c O | 77 | 5 | 1.96 | 2.88 | 125 | — | 143 |
| m/c O···$O_w$ | 163 | 18 | — | 2.92 | — | 130 | — |
| m/c NH···$O_w$ | 34 | 9 | 2.11 | 3.00 | — | — | 147 |
| s/c O···$O_w$ | 108 | 10 | — | 2.78 | — | 121 | — |
| s/c NH···$O_w$ | 55 | 16 | 2.06 | 2.77 | — | — | 137 |
| s/c N···$O_w$ [d] | 18 | 1 | — | 2.85 | — | — | 127 [e] |
| s/c O···s/c NH | 30 | 1 | 2.00 | 2.89 | 123 | — | 147 |
| s/c O···s/c O | 21 | 4 | — | 2.73 | — | 122 | — |
| hydrogen bonds in secondary structure elements [f] | | | | | | | |
| helices | 236 | 20 | 2.09 | 2.99 | 147 | — | 148 |
| - α-type (1-5) | 152 | 9 | 2.04 | 2.96 | 145 | — | 152 |
| - $3_{10}$ (1-4) | 70 | 11 | 2.21 | 3.03 | 109 | — | 138 |
| β-sheet | 94 | 7 | 1.99 | 2.92 | 150 | — | 157 |

[a] X represents either nitrogen or oxygen hydrogen bond donors, depending on the bond in question. In the cases where a hydrogen atom has been added in its calculated position, no C-O···X angle is given. [b] These are interactions where O···H distance is between 2.4 and 2.6 Å. In the case of O - O or s/c N - O hydrogen bonds, these marginal bonds have O···O or N···O between 3.4 and 3.6 Å. [c] The O···X distance and the C-O···X angle correspond to m/c O···s/c O hydrogen bonds, while the other values relate to m/c O···s/c HN bonds. [d] Lysine residues only. [e] $C_\varepsilon$-N$_\zeta$···$O_w$ angle. [f] See also Tables 1 - 4 in Appendix IV.

The three main types of helix end distortions as described by Baker & Hubbard (1984) are observed in the helices of $Cu_2Lf$. These are: (1) $\alpha_N$ distortions at the N-termini involving a bifurcated hydrogen bond (a $3_{10}$ bond and an α-type bond), occurring in helices 2, 3, 6, 8 and 9 in the N-lobe and 3, 5, 8 and 12 in the C-lobe; (2) $\alpha_{C1}$ distortions where there is a change from α-type hydrogen bonding to $3_{10}$ bonds towards the C-terminus. This type of distortion is observed in helices 5, 6 and 10 in the N-lobe and 1, 8 and 11 in the C-lobe. In the case of helices 1 and 5, the hydrogen bond at the α-$3_{10}$ transition is not bifurcated at the nitrogen and the available carbonyl oxygen atoms are not satisfied by either sidechain or solvent hydrogen bonding; (3) $\alpha_{C2}$ distortions with a π (1-6) hydrogen bond followed by a

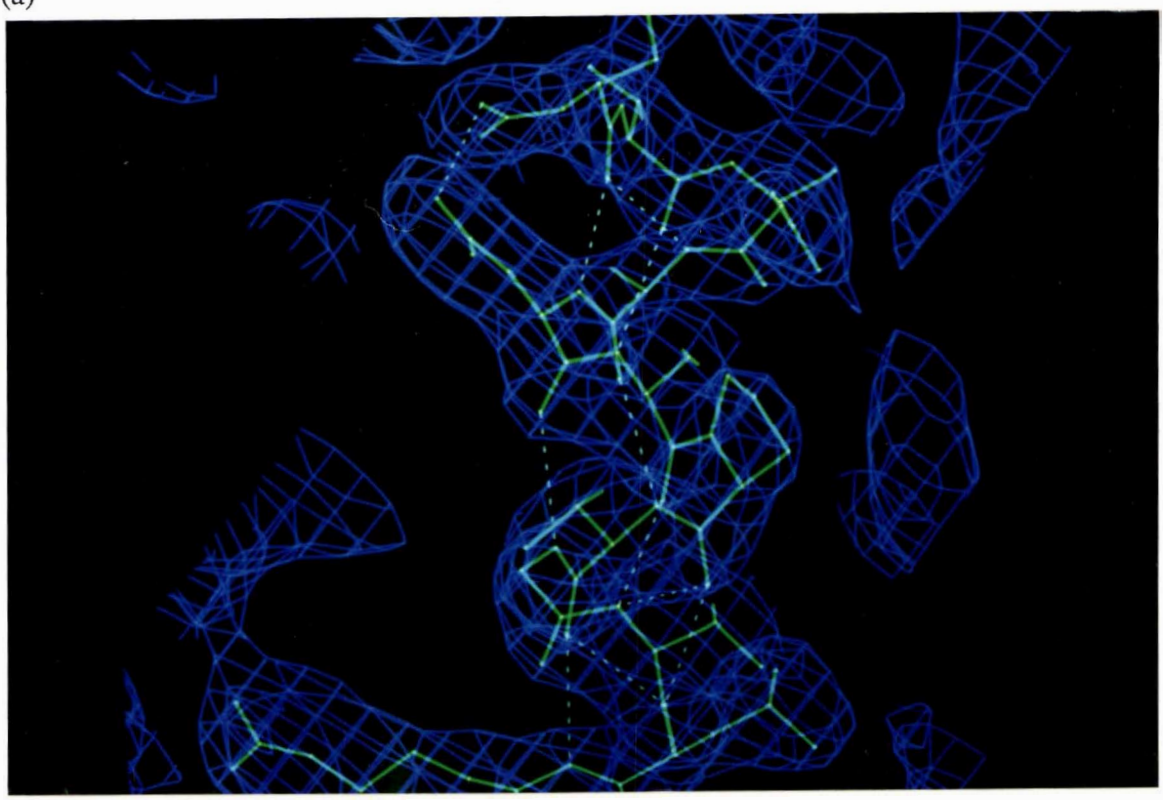
$3_{10}$ (2-5) bond at the C-terminus, as in helices 1, 2 and 7 in the N-lobe and 2, 3 and 7 in the C-lobe.

A full description of the hydrogen bond geometries in these helices can be found in Table 3, Appendix IV and Plate 3.5a shows a representative piece of helix in $2F_o$-$F_c$ electron density. The average O···H distance for all the helical mainchain-mainchain hydrogen bonds (O···HN distance) of 2.09 Å compares favourable with the result of 2.06 Å obtained by Baker & Hubbard (1984) in their analysis of 15 proteins obtained from the Protein Data Bank. The average C-O···H angle of 145°, virtually identical to that reported (143°), again indicates the consistently similar nature of the helical structures observed in globular proteins.

*β-strands:* The characteristics of the hydrogen bonds between the β-strands are given in Table 4, Appendix IV and a representative portion of a β-sheet is shown in Plate 3.5b. These were analysed using the same criteria as for the helices. A pattern similar to that previously described for $Fe_2$Lf was observed (Anderson *et al.*, 1989), with an average O···H distance of ≈ 1.97 Å. Again this is consistent with the value of 1.96 Å found by Baker & Hubbard (1984) for antiparallel β-strands. The hydrogen bonds, analysed as described above, are almost linear, with average N-H···O and C=O···H angles of 157° and 150° respectively.

*Turns:* The turns in $Cu_2$Lf have been analysed in terms of the observed φ/ψ angles and hydrogen bonding interactions, and of the 55 turns which can be readily identified and classified (see Table 5, Appendix IV), 52 can be designated as β-turns (1-4 turns) with the majority falling into categories typical of globular proteins, namely β-type I (29), II (9) and III (12) turns. In addition there are two type VI turns associated with two of the cis proline residues (71 and 141). The three remaining turns are γ-type turns, two of which (residues 298 - 300 and 461 - 463) can be described as having classical γ conformations, similar to those observed in thermolysin (Matthews, 1972) and the periplasmic ribose binding protein (RBP) from *Escherichia coli* (Mowbray & Cole, 1992). The third (residues 444 - 446) has φ/ψ angles of -96° and 52° for the second residue (Asp445), designated as an inverse γ-turn by Baker & Hubbard (1984), has a conformation similar to that observed in inverse γ-turns

(a)



(b)

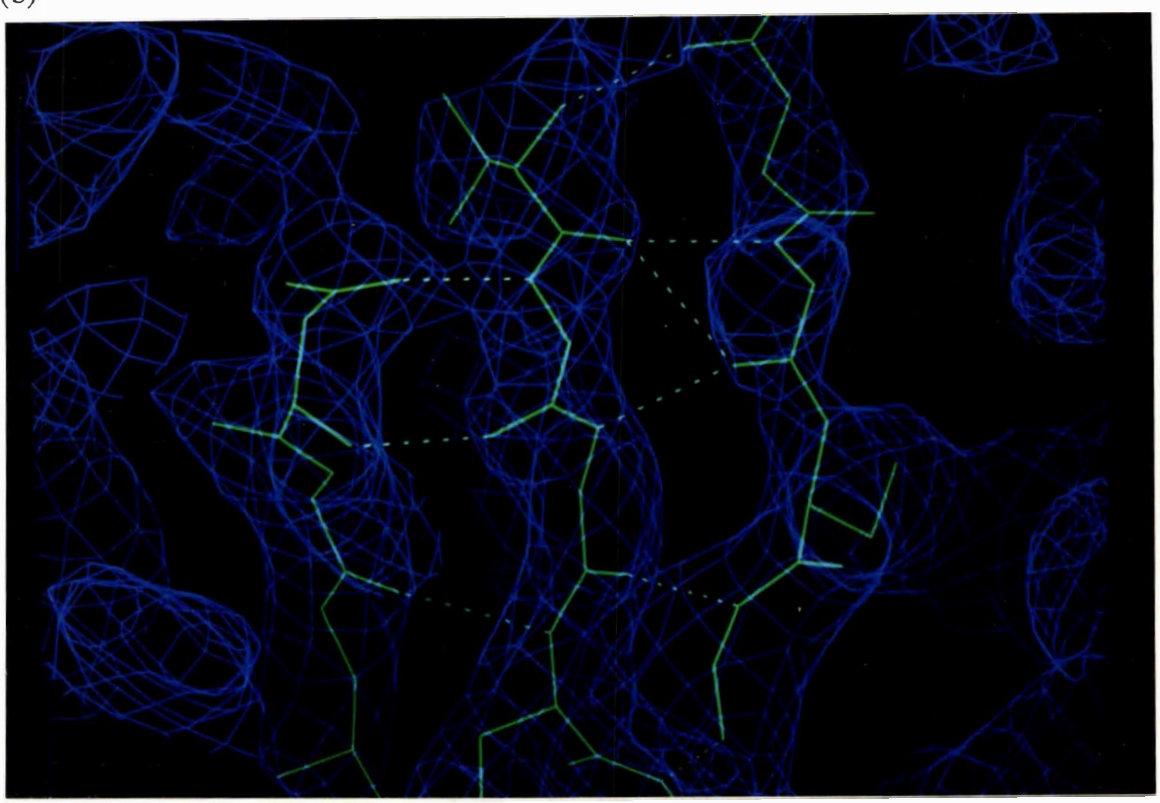


Plate 3.5: Representative sections of the $2F_o$-$F_c$ electron density associated with the final model of $Cu_2Lf$. Photograph (a) shows the structure of helix 5 in the N-lobe and photograph (b) shows three b-strands (from left to right, strands c, a and b) from the N-lobe.

in RBP (Mowbray & Cole, 1992). In this case of $Cu_2Lf$, however, although the O ··· H distance is favourable (2.36 Å), the angle at the carbonyl oxygen is ≈83°.

The former two γ-turns (at 299 and 462) are in equivalent positions in the N- and C-lobes of the protein, occurring as one of a series of loops in a portion of the polypeptide chain between an α-helix (helix 9 in Table 3.11) and a β-strand (strand k). Analysis of the atomic coordinates from the preliminary structures of $Fe_2rTf$ and $Fe_2OTf$, along with the N-terminal lobe of rTf, indicate similar turns at identical positions in the polypeptide chain, with φ/ψ angles near 70° and -50° respectively. It should be noted that these rare turns occur in a region which is highly conserved between most members of the transferrin family.

Table 3.13A: Characteristics of the two γ-turns in human lactoferrin

| | distance (Å) [a] | dihedral angles (deg) [a] | | mainchain torsion angles (deg) [b] | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1-3 hydrogen bond | O···HN (Å) | C–O···H (deg) | N–H···O (deg) | $\phi_1$ | $\psi_1$ | $\phi_2$ | $\psi_2$ | $\phi_3$ | $\psi_3$ |
| 298-300 | 1.67 | 90 | 165 | -85 | 122 | 72° | -28° | -157 | 162 |
| 641-643 | 1.51 | 96 | 162 | -91 | 127 | 62° | -39° | -138 | 157 |

[a] Hydrogen atoms were added to the coordinate file in calculated positions, using a standard N-H bond length of 1.0 Å. [b] The torsion angles with the subscripts 1 and 3 relate to the residues involved in the 1-3 hydrogen bonds, while the angles relating to the crucial second residue in the γ-turn are labelled with a subscript 2.

Table 3.13B: Additional stabilising hydrogen bonds in the vicinity of γ-turns at 299 and 642

| γ–turn | hydrogen bond | distance (Å)[a] | dihedral angles (deg) | | | |
|---|---|---|---|---|---|---|
| 298-300 | NH(298)...$O_\gamma$(291) | 1.95 | $C_\beta$-$O_\gamma$···H | 154 | N-H···$O_\gamma$ | 143 |
| | NH(299)···$O_{\varepsilon 1}$(15) | 2.32 | $C_\delta$–$O_{\varepsilon 1}$···H | 106 | N–H···$O_{\varepsilon 1}$ | 158 |
| | O(299)···$O_{\gamma 1}$(58) | 2.96 | $C_\beta$-$O_{\gamma 1}$···O | 94 | C-O···$O_{\gamma 1}$ | 166 |
| | O(299)···$O_w$(704) | 2.48 | C-O···$O_w$ | 125 | – | – |
| 641-643 | NH(641)···$O_\gamma$(636) | 2.17 | $C_\beta$-$O_\gamma$···H | 113 | N-H···$O_\gamma$ | 173 |
| | NH(642)···$O_{\varepsilon 2}$(354) | 1.96 | $C_\delta$–$O_{\varepsilon 2}$···H | 95 | N–H···$O_{\varepsilon 2}$ | 160 |
| | O(461)···$O_w$(713) | 3.24 | C-O···$O_w$ | 145 | – | – |
| | O(462)···$O_w$(706) | 2.92 | C-O···$O_w$ | 119 | – | – |
| | O(463)···$O_w$(713) | 2.92 | C-O···$O_w$ | 118 | – | – |

[a] Where a N-H bond exists, the hydrogen - oxygen distance is given is, while in the case of two oxygen atoms being hydrogen bonded, the oxygen - oxygen distance is given.

The residues involved in these two γ-turns are further stabilised by hydrogen bonds with the sidechains of adjacent residues and solvent molecules (Figures 3.12 a & b and Table 3.13A & B). Consequently, Leu 299 and Leu 642 are able to adopt this unusual conformation as a result of the sharp polypeptide chain reversal.

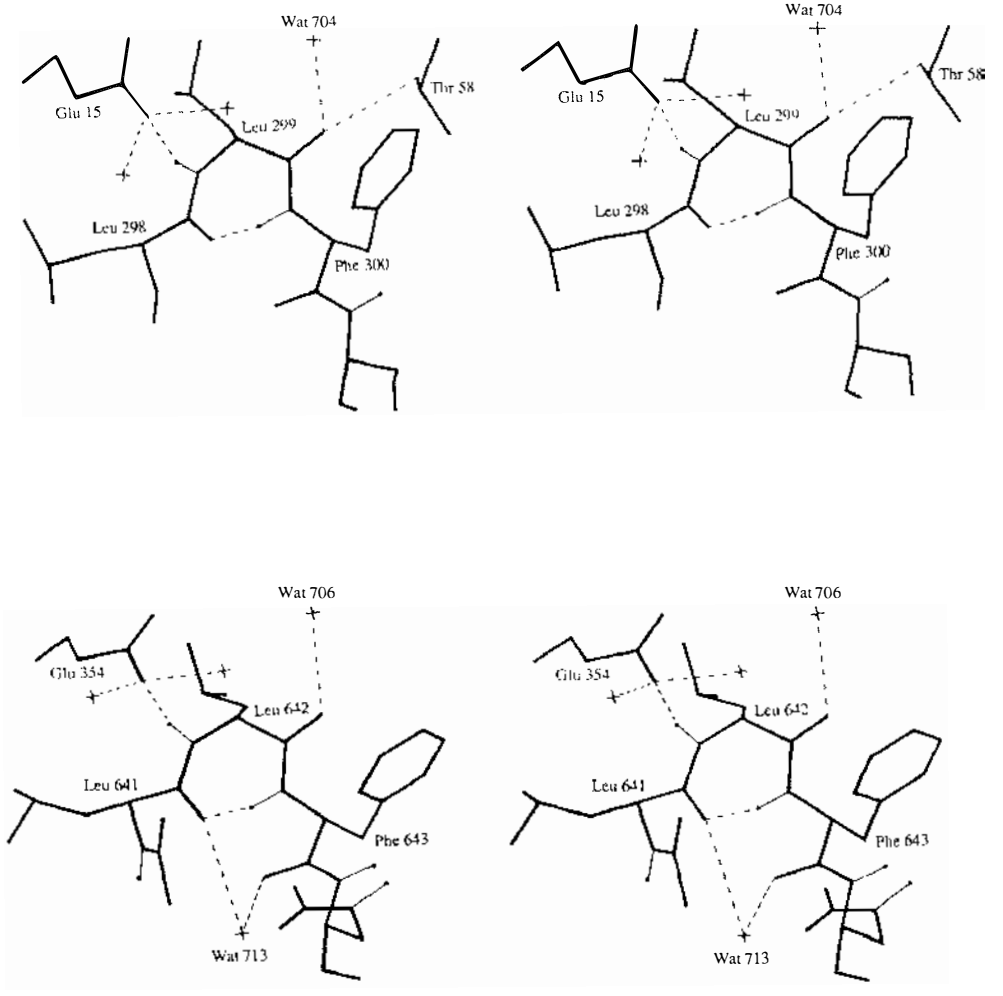



Figure 3.12: The two γ-turns at residues 299 and 642 and associated hydrogen bonds.

### 3.2.4 *Temperature factors:*

Figures 3.13a & b show the average mainchain temperature factors for $Cu_2Lf$ as a function of residue number for the N- and C-lobes respectively. The average mainchain B value of 44 $Å^2$ (see Table 3.14) reflects the weakness of the high resolution data and the fact

that all data were used in the least squares refinement. Analysis of the data indicates a sharp falloff in observed reflections at about 2.7 Å resolution (see Figure 3.3), with only about 25% of the reflections having $I > 2\sigma_I$ at the maximum resolution used for refinement (2.1 Å). This leads to much higher individual B values as the calculated structure factors are fitted to the weak observed data.

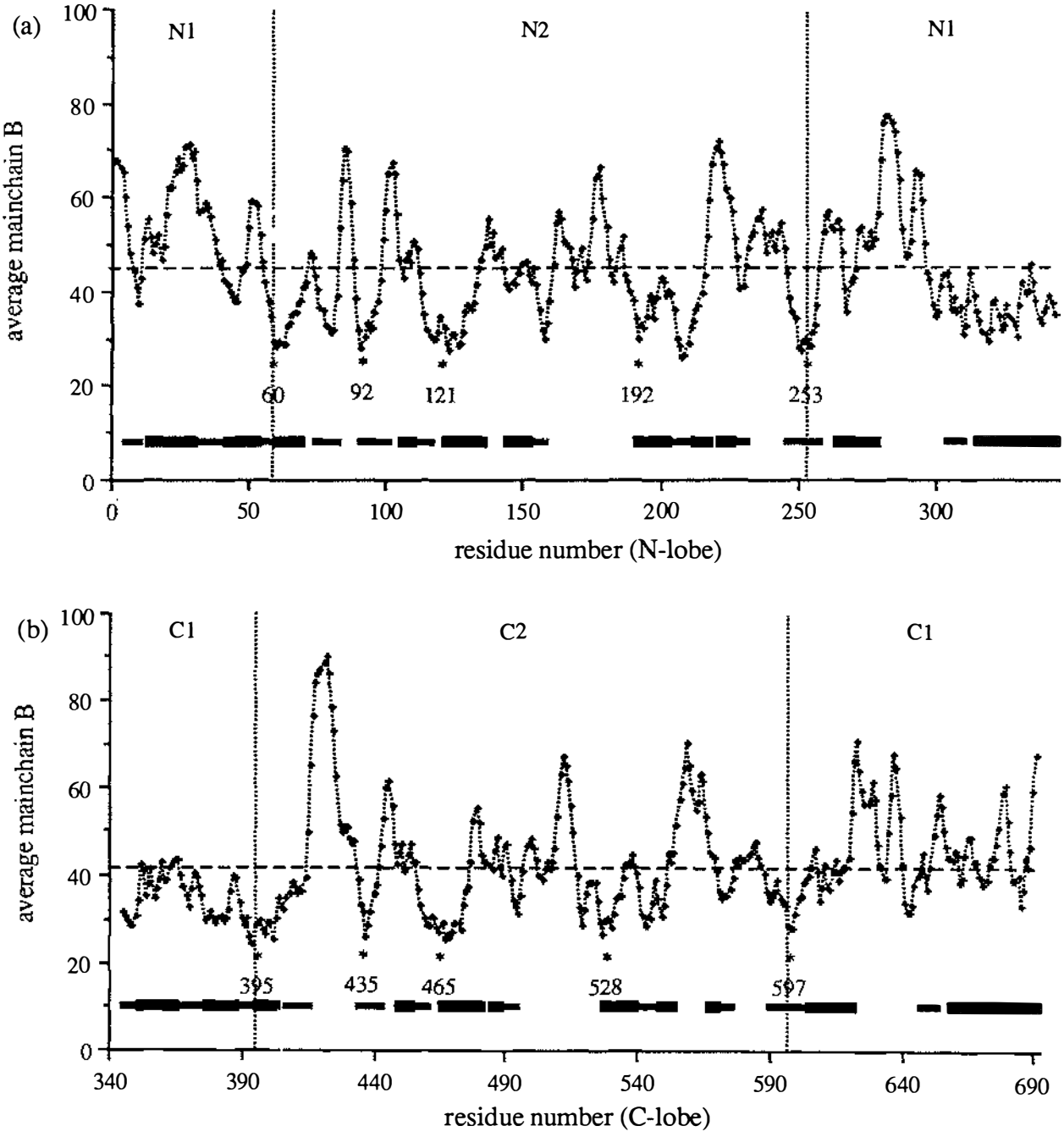



Figure 3.13: Plot of average mainchain B values against residue number for (a) the N-lobe and (b) the C-lobe of $Cu_2Lf$. The positions of the helices and β-strands are indicated by ▬ and — respectively. The *'s show the location of the four metal binding ligands in both lobes, along with the anion-binding Arg121 (N-lobe) and Arg465 (C-lobe).

The high B values are also consistent with the overall B estimated from a Wilson plot of $\ln(I_o / \Sigma f^2)$ as a function of $\sin^2\theta/\lambda^2$ (Wilson, 1942), where $I_o$ is the average intensity and $\Sigma f^2$ is the sum of the square of the scattering factors in small concentric shells of reciprocal space ($\Delta(\sin^2\theta/\lambda^2) \approx 0.0045$ Å$^{-2}$). This plot is shown in Figure 3.14. The gradient, -2B, gives an average B of about 43 Å$^2$, similar to the value obtained from least squares refinement.

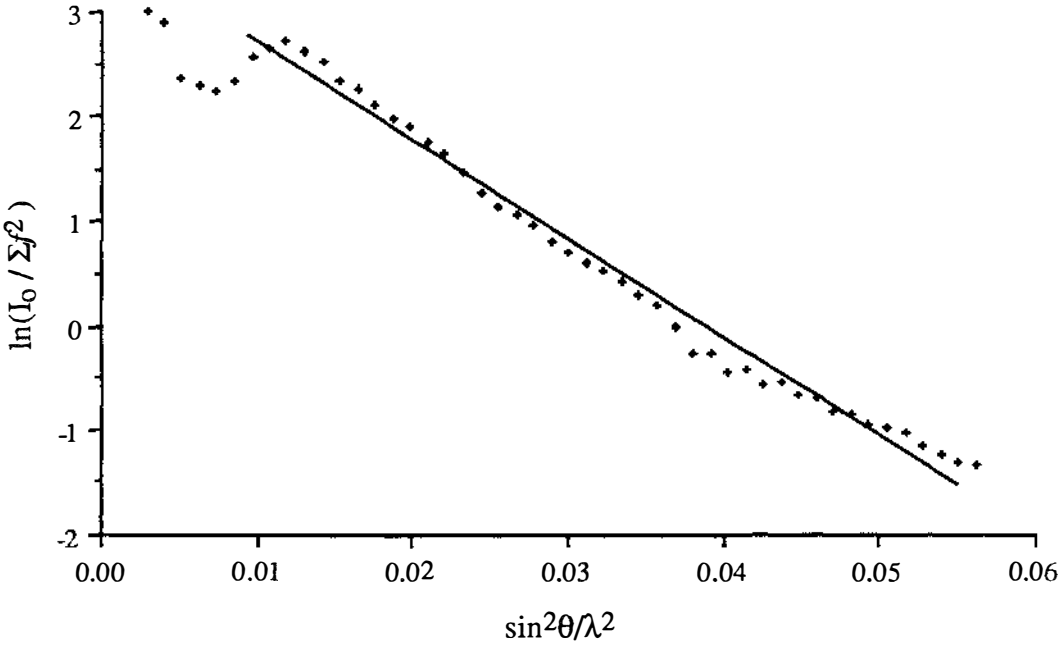



Figure 3.14: A Wilson plot for $Cu_2Lf$. The plot is based on 44715 reflections to 2.1Å resolution. The best fit line for points between 4.5 and 2.6Å is shown, giving an average B value of ≈ 43 Å$^2$.

In both lobes, the fluctuation of B values along the polypeptide chain appears to be quite regular and can be correlated to the observed variation in the secondary structure and the distribution of hydrogen bonding (discussed later). The residues with low B values correspond to the β-strands, the helices and the areas associated with the metal and anion binding ligands (as indicated in Figure 3.13). Most of the residues in these parts of the chain are internal and are stabilised by a large number of hydrogen bonding and other interactions with other protein atoms and ordered water molecules. The exceptions in the N-lobe are the initial 3 or 4 residues of β-strand a, where it connects to the poorly-ordered N-terminal

residues, the majority of helix 1 which is exposed to the external solvent, and helix 8a, a short piece of $3_{10}$ helix which also lies on the molecular surface. In the C-lobe, the secondary structure is quite well defined except for the last few residues of C-terminal helix 12.

Table 3.14A gives the average mainchain and sidechain atomic temperature factors for the whole molecule (excluding solvent and the glycan chains), individual lobes and the four domains. These results suggest that the N1 domain contains regions which are somewhat less ordered than the corresponding areas in the C1 domain, and, although it is not correct to generalise that the whole of the N1 domain is more mobile, a significant portion of it appears to be more flexible than the remainder. Figure 3.13 would indicate that it is the initial 50 or so residues which contribute most to the high B values and that another region in the second half of this domain (residues 280 - 300) also has B values significantly greater than the average. However, in the C1 domain, although the region corresponding to the latter (residues 620 - 640) also has high B values, the initial 50 residues have values less than or equal to the average. The reason for these differences is clear from analysis of the model: while the first 50 residues of the N-lobe are located on the exterior of the molecule, the same area in the C-lobe lies close to the N-lobe, in the inter-lobe contact area, and exposure to the external solvent is greatly reduced.

The regions with high B values (> 60 $Å^2$) are mostly external loops connecting the elements of secondary structure. In these parts, the majority of the hydrogen bonding interactions are with poorly-ordered solvent molecules and consequently the protein structure is flexible and not well defined. These ill-defined areas of the $Cu_2Lf$ model (listed in Table 3.14B), can be correlated with the rms deviation in mainchain atomic positions between $Cu_2Lf$ and $Fe_2Lf$, by analysing the average rms displacement as a function of temperature factor. The average mainchain temperature factors (ranging from 20 to 100 $Å^2$) were divided into 15 blocks each of 5 $Å^2$ (except the last which was a 10 $Å^2$ block between 90 and 100 $Å^2$), and the average mainchain rms deviation was calculated for the residues in each block (the N-terminus and a very poorly-ordered loop between residues 281 and 284 were omitted).

These results are given in Table 3.14C and shown diagrammatically in Figure 3.15a, and they show an expected increase in the rms difference in atomic positions with increasing B. Figure 3.14b is a plot of the average rms deviation in mainchain atom positions between $Fe_2Lf$ and $Cu_2Lf$ as a function of average mainchain temperature factor. From both of these graphs it can be clearly seen that as the atomic temperature factors increase, the positional parameters describing these atoms become less well-defined. This in turn can be linked with errors in the positional parameters of the atoms (as discussed in Section 3.2.1).

Table 3.14A: Average mainchain and sidechain temperature factors for $Cu_2Lf$ [a]

| | overall | N-lobe | C-lobe | N1 | N2 | C1 | C2 |
|---|---|---|---|---|---|---|---|
| average mainchain B | 43.7 | 45.2 | 42.2 | 46.6 | 43.7 | 43.0 | 41.3 |
| average sidechain B | 46.5 | 47.2 | 44.0 | 49.0 | 45.3 | 44.7 | 43.1 |

[a] Averages are for protein atoms only. Residues 1 - 240 make up the N-lobe and residues 341 - 691 the C-lobe. The residues making up the four domains are given in Section 3.2.1 on page 67.

Table 3.14B: Average temperature factors.for the flexible regions of $Cu_2Lf$ [a]

| N-lobe | average B value [b] | | C-lobe | average B value | |
|---|---|---|---|---|---|
| | main (Å2) | side (Å2) | | main (Å2) | side (Å2) |
| 1 - 6 | 89 | 94 | – | – | – |
| 20 - 36 | 64 | 70 | – | – | – |
| 83 - 87 | 66 | 85 | 416 - 425 | 84 | 82 |
| 100 - 104 | 68 | 62 | 445 - 447 | 64 | 74 |
| 137 - 139 | 58 | 66 | 479 - 480 | 56 | 66 |
| 162 - 164 | 60 | 68 | – | – | – |
| 176 - 180 | 63 | 73 | 510 - 516 | 58 | 81 |
| 218 - 226 | 63 | 75 | 555 - 567 | 61 | 69 |
| 280 - 287 | 75 | 79 | 622 - 631 | 62 | 74 |
| 293 - 295 | 71 | 76 | 635 - 639 | 62 | 77 |
| – | – | – | 678 - 680 | 64 | 71 |
| – | – | – | 690 - 691 | 69 | 78 |

[a] The residues have been aligned so that the residue ranges appearing on the same line are in corresponding positions in the two lobes. [b] The averages given are taken over all the residues in the specified ranges.

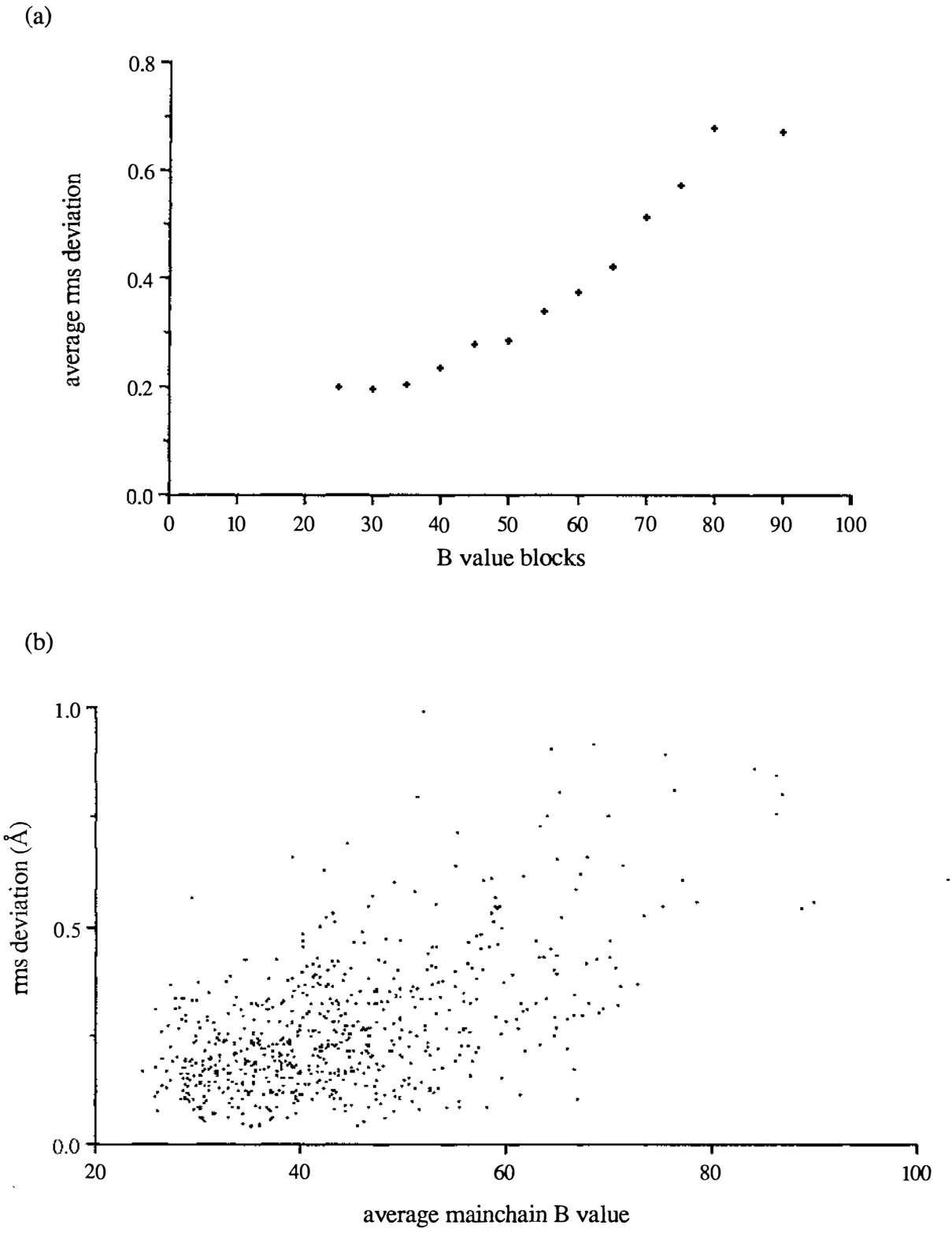



Figure 3.15: (a) The rms deviation between mainchain atoms (N, C, O and $C_\alpha$) in $Cu_2Lf$ and $Fe_2Lf$ averaged over 15 blocks of the average mainchain temperature factor, and plotted as a function of these B value blocks. The average rms deviation between mainchain atoms in the two structures for each residue is shown in (b) as a function of average mainchain B (only residues with rms deviations < 1.0 are shown).

Table 3.14C: Correlation of the average mainchain B values with the rms deviation in mainchain atomic positions between $Cu_2Lf$ and $Fe_2Lf$ [a]

| B value block (Å$^2$) | numbers in block | average B value in each block (Å$^2$) | average rms deviation (Å) |
|---|---|---|---|
| 20 - 25 | 1 | 24.7 | 0.20 |
| 25 - 30 | 62 | 28.3 | 0.20 |
| 30 - 35 | 106 | 32.4 | 0.20 |
| 35 - 40 | 134 | 37.5 | 0.23 |
| 40 - 45 | 135 | 42.6 | 0.28 |
| 45 - 50 | 85 | 47.5 | 0.28 |
| 50 - 55 | 50 | 52.2 | 0.34 |
| 55 - 60 | 47 | 57.2 | 0.37 |
| 60 - 65 | 24 | 63.1 | 0.42 |
| 65 - 70 | 21 | 67.2 | 0.51 |
| 70 - 75 | 9 | 71.2 | 0.57 |
| 75 - 80 | 4 | 76.5 | 0.68 |
| 80 - 85 | 0 | 0 | — |
| 85 - 90 | 6 | 87.3 | 0.67 |
| 90 - 100 | 0 | 0 | — |

[a] Excluding the first five residues at the N-terminus and residues 282 and 283, which had large rms deviations in mainchain atom positions.

### 3.2.5 *Solvent structure:*

The 301 solvent molecules were located independently of the $Fe_2Lf$ structure (i.e. no prior knowledge of the solvent positions in $Fe_2Lf$ were assumed). The water molecules were monitored during least squares refinement after they were added, and those which showed large increases in B value were either removed or their positions checked by calculating omit maps ($F_o$-$F_c$) with the water molecules in question being left out for 4-6 cycles of refinement. It was concluded that the majority of the 301 water molecules were in well-defined positions, in that they made good hydrogen bonded contacts with mainchain or sidechain atoms which would otherwise be uns .tisfied. As a test, several apparently well-ordered water molecules were removed and omit maps calculated (after removal of bias by 5 cycles of least squares refinement). Subsequent $F_o$-$F_c$ maps showed positive peaks at the positions of the omitted water molecules and they were added back into the structure in essentially the same positions as they occupied previously.

Some of the solvent in the interdomain clefts and a number of the external water molecules were involved in extended solvent networks, where they were only interacting with other water molecules. The positions of these solvent molecules are somewhat less well-defined, as, in general, they were only located with $F_o$-$F_c$ maps and had little or no $2F_o$-$F_c$ density associated with them. After 2 or 3 phases of refinement, a number of these water molecules were removed, omit maps calculated and they were relocated in the $F_o$-$F_c$ density which reappeared.

A summary of the solvent molecules in $Cu_2Lf$, in terms of their location in the model, how many hydrogen bonds they are involved in and their temperature factors, is given in Tables 3.15A & B. Of the 301 water molecules in $Cu_2Lf$, 234 can be classified as belonging to the first hydration sphere, in that they make at least one hydrogen bond to a protein atom. The remaining 67 water molecules (22%) are classified as second sphere solvent and make an average of 2.4 hydrogen bonds to other water molecules (both 1st and 2nd sphere). The majority of these second sphere solvent (≈60%) are involved in the solvent network in the two interdomain clefts.

Table 3.15A: Summary of the 301 solvent molecules in $Cu_2Lf$

| | |
|---|---|
| Number of solvent molecules | 301 |
| - hydrogen bonded to at least one protein atom (1st sphere solvent) | 234 |
| - 2nd sphere solvent | 67 |
| Total number of hydrogen bonds to solvent | 848 |
| - total 1st sphere hydrogen bonds | 687 |
| - average number / 1st sphere water | 2.9 |
| - total 2nd sphere bonds (to other solvent) | 161 |
| - average number / 2nd sphere water | 2.4 |

The results in Table 3.15B indicate that there is a small but consistent increase in the average B value for the solvent increases as the number of hydrogen bonds to the water molecules decreases. This has been observed previously in well-refined protein structures

(Baker, 1988 and references therein). It is clear that the more constrained the water molecules are (i.e. the more hydrogen-bonding partners they have), the lower is their temperature factor. The overall average temperature factor for the 301 water molecules in $Cu_2Lf$ is 57.3 $Å^2$.

Table 3.15B: Analysis of the solvent hydrogen bonds with respect to the individual temperature factors of the solvent molecules in $Cu_2Lf$

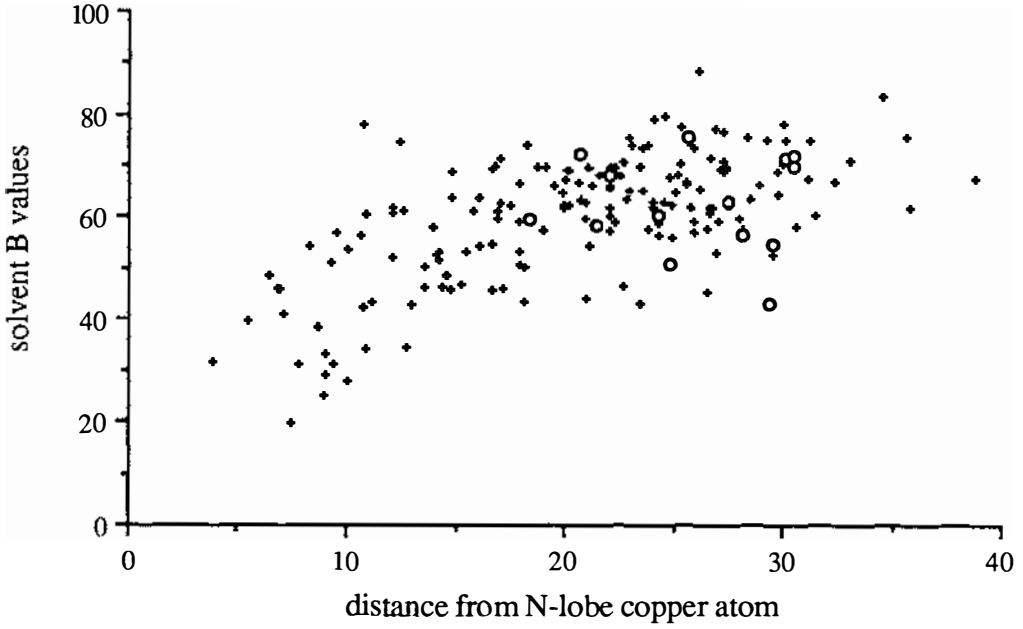
| number of hydrogen bonds | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| number of water molecules | 13 [a] | 59 | 61 | 65 | 43 | 60 |
| average B value ($Å^2$) | 62 | 64 | 59 | 57 | 55 | 49 |

[a] An $O_w \cdots X$ distance cutoff of 3.35 Å was used to classify interactions as hydrogen bonds, in conjunction with the angle at the donor atom (X). The 13 solvent molecules labelled as having 0 interactions made contacts with oxygen or nitrogen atoms at distances between 3.6 and 3.35 Å.

The water molecules were also classified according to which of the two lobes they were associated with, primarily in terms of the distance of each solvent molecule from the two copper atoms. Plots of their individual temperature factors as a function of the copper - water distance are shown in Figures 3.16a & b. In each lobe it was found that the copper - water distances ranged from ≈ 4 Å for those hydrogen bonded directly to the metal binding ligands (see later), to ≈ 25 - 30 Å for those water molecules on the molecular surface.

Solvent molecules closer than about 10 Å have temperature factors averaging 35 - 40 $Å^2$ and, with the exception of a number of second sphere solvent, are generally hydrogen bonded to 3 or 4 protein atoms. Further out from the metal atoms, between 10 and 25 Å, the variation in B values becomes wider, more so in the C- than the N-lobe. This can be linked to the increasing number of second sphere solvent in the interdomain clefts, in that water molecules hydrogen bonded to residues lining the cleft exhibit relatively low B values, while the second sphere solvent in the cleft have B values significantly higher. Those solvent molecules at distances greater than 25 Å generally have higher B values (averaging 60 - 70 $Å^2$) and the majority of these are located on the surface of the molecule.

(a)




(b)

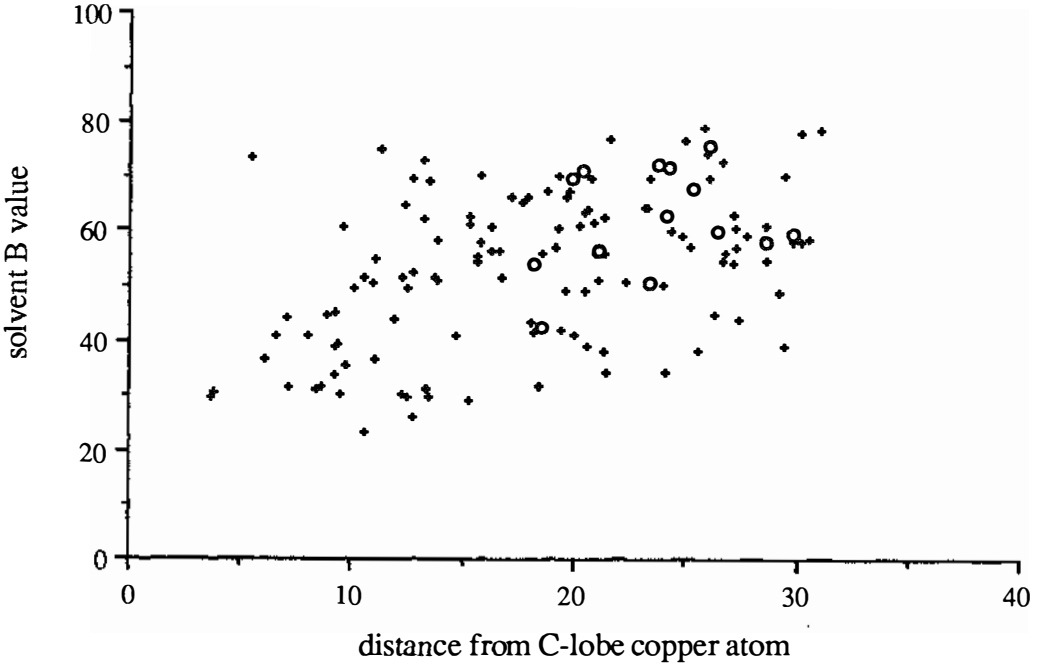


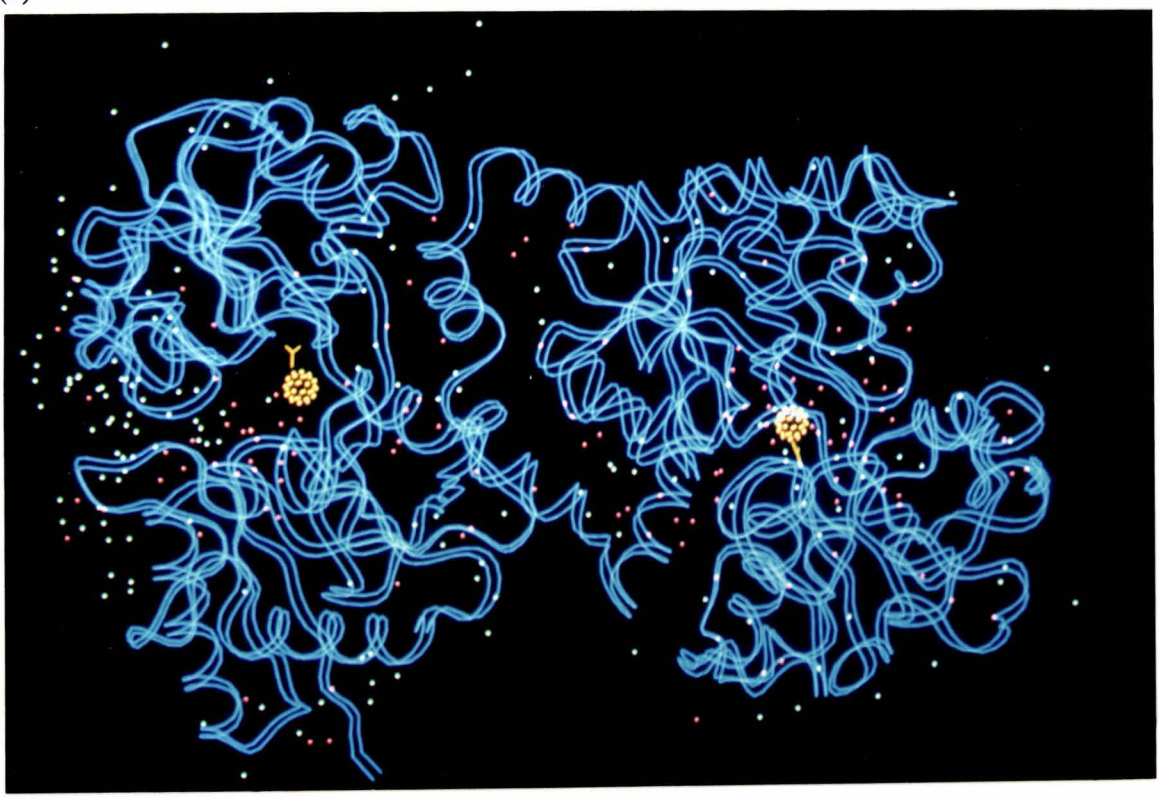

Figure 3.16: Plots of the B value for individual solvent molecules in the (a) N- and (b) C-lobes of $Cu_2Lf$ as a function of the distance of the solvent molecules from the closest copper atom. Solvent molecules within 30 Å of both copper atoms are indicated by o and all other solvent molecules are indicated by +.

In addition, there are 14 water molecules which are less than 30 Å from both copper atoms (indicated in Figures 3.16a & b by o ) and, upon inspection of a model of diferric lactoferrin, are found to be located near the interface between the two lobes. The B values for these 14 solvent molecules range from 43 to 75 $Å^2$ (average = 62 $Å^2$) and the majority of them are on the surface of the molecule. Analysis of the hydrogen bonding associated with these water molecules indicates an average of only 1.3 partners each. However, it should be noted that of these 14 water molecules, 11 are common to both $Cu_2$Lf and $Fe_2$Lf, in that their positions in the two molecules differ by less than 2 Å.

*Common solvent:* Least squares superposition of the $Cu_2$Lf and $Fe_2$Lf structures followed by analysis of the superimposed solvent structure showed that 123 out of 301 (Plate 3.6a) $Cu_2$Lf water molecules were in conserved positions (within 1.0 Å in the two structures) and that at least 40 of these solvent molecules could be designated as internal, in that they were either buried within the protein structure or located in the two interdomain clefts. These water molecules were characterised by relatively low B values (≈ 30 - 40 $Å^2$) and were hydrogen bonded to mainchain or sidechain atoms with low B values in well-defined electron density. Plate 3.6b shows some representative "common" solvent molecules. Another 58 "common" solvent molecules were about 3 Å below the approximate molecular surface, separated from external solvent molecules by protein atoms. Although they could not be rigorously defined as internal solvent, they did not interact with the external water molecules and had B values substantially lower than true external solvent molecules (an average B of ≈ 53 $Å^2$ compared with an average of ≈ 64 $Å^2$ for external solvent). A further 39 water molecules were between 1.0 and 2.0 Å from the positions of solvent molecules in $Fe_2$Lf. The average B value for the 162 common solvent is ≈50 $Å^2$.

Analysis of the individual temperature factors of the 162 common solvent molecules shows a good correlation between the B values and the number of hydrogen bonds associated with the water molecule (Table 3.16B), and also between the B values and the difference in position of the solvent molecules in the two structures (Table 3.16A and Figure

(a)



(b)

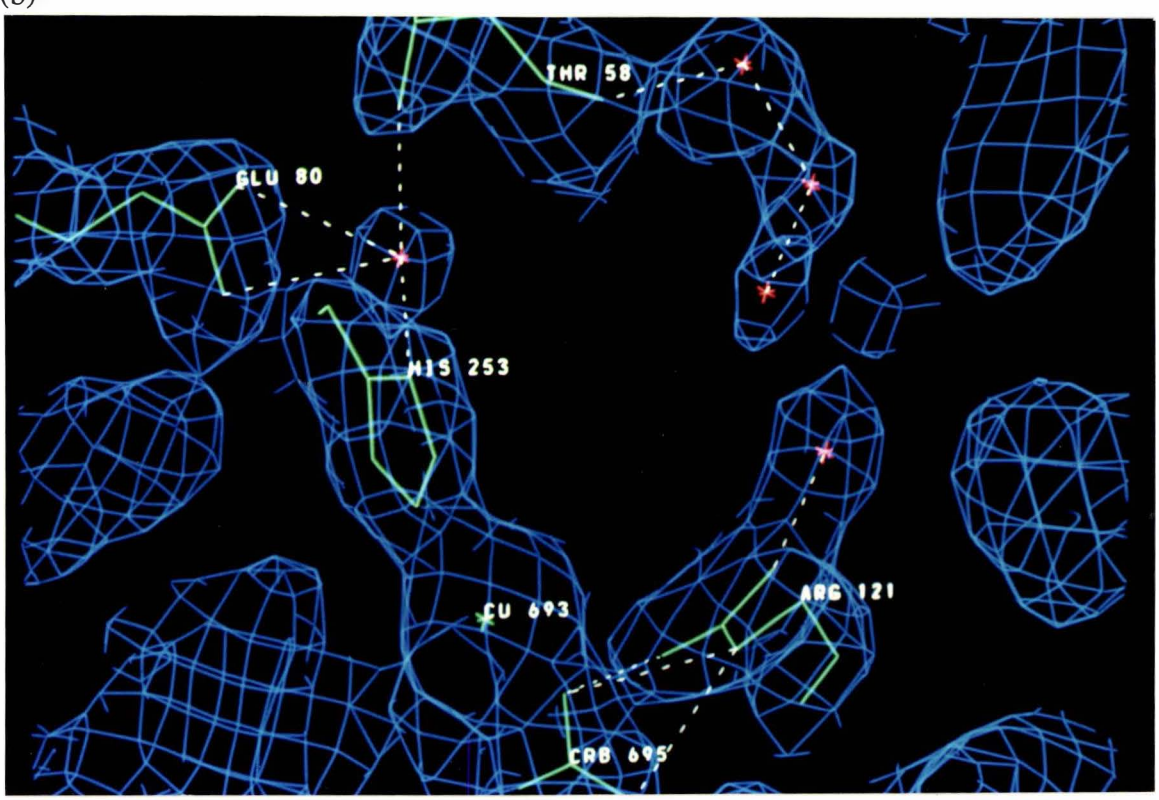



Plate 3.6: (a) The location of the solvent molecules in $Cu_2Lf$. The positions of solvent molecules common to $Cu_2Lf$ and $Fe_2Lf$ are shown in red, with the remaining unique water molecules in white; (b) Some representative "common" or "core" water molecules (red crosses) in their $2F_o$-$F_c$ electron density. These solvent molecules are in the vicinity of the N-lobe metal binding site of $Cu_2Lf$.

3.17). It is interesting to note that there is an increase in the average B value as the variation in the positions of the equivalent water molecules increases, and that the average number of hydrogen bonding partners (protein + solvent) decreases. This would indicate that the common water molecules with large differences in position between the two structures are generally those molecules which interact less with other atoms and are likely to be either at the molecular surface or 2nd sphere solvent (the number of interactions with other solvent molecules also increases).

Table 3.16A: Analysis of common solvent molecules in $Fe_2Lf$ and $Cu_2Lf$ with respect to the difference in position in the two structures

| | Distance between equivalent water molecules (Å) | | | |
|---|---|---|---|---|
| | 0.0 - 0.5 | 0.5 - 1.0 | 1.0 - 1.5 | 1.5 - 2.0 |
| internal solvent | 21 | 19 | 5 | 6 |
| - cleft | 18 | 16 | 4 | 6 |
| surface [a] | 41 | 17 | 6 | 3 |
| external solvent | 1 | 24 | 6 | 9 |
| Number of common solvent | 63 | 60 | 21 | 18 |
| Average B ($Å^2$) | 44 | 54 | 57 | 61 |
| Total hydrogen bonds | 191 | 153 | 47 | 30 |
| - average bonds/water molecule[b] | 3.0 | 2.6 | 2.2 | 1.7 |
| Total bonds to protein atoms [c] | 127 (66%) | 86 (56%) | 20 (43%) | 15 (50%) |
| Total bonds to solvent molecules | 64 | 67 | 27 | 15 |
| Number of 1st sphere water | 61 | 49 | 17 | 13 |
| Number of 2nd sphere water | 2 (3%) | 11 (18%) | 4 (19%) | 5 (28%) |

[a] Molecules within ≈ 3 Å of the molecular surface but not interacting with external solvent. [b] Based on the total number of water molecules in each range, although 8 water molecules out of 162 do not appear to interact with any protein or solvent atoms. [c] The numbers in parentheses are percentages of the total number in each range.

The large rms deviation in B values is primarily due to the difference between the two structures in the average B values for the common water molecules, the average value for $Fe_2Lf$ being ≈ 62 $Å^2$. This difference can be attributed primarily to the amount of least

squares refinement performed on the two structures (17 phases for $Cu_2Lf$ compared to 27 for $Fe_2Lf$).

Table 3.16B: Analysis of the common solvent in $Cu_2Lf$ and $Fe_2Lf$ with respect to the numbers of hydrogen bonds

| Number of hydrogen bonds | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| - number of water molecules | 13 | 31 | 38 | 42 | 27 | 16 |
| - average B value (Å$^2$) | 61 | 59 | 55 | 49 | 45 | 47 |

| | |
|---|---|
| Average B value for the 162 common water molecules (Å$^2$) | 50 |
| rms deviation in B values between $Cu_2Lf$ and $Fe_2Lf$ (Å$^2$) | 16 |

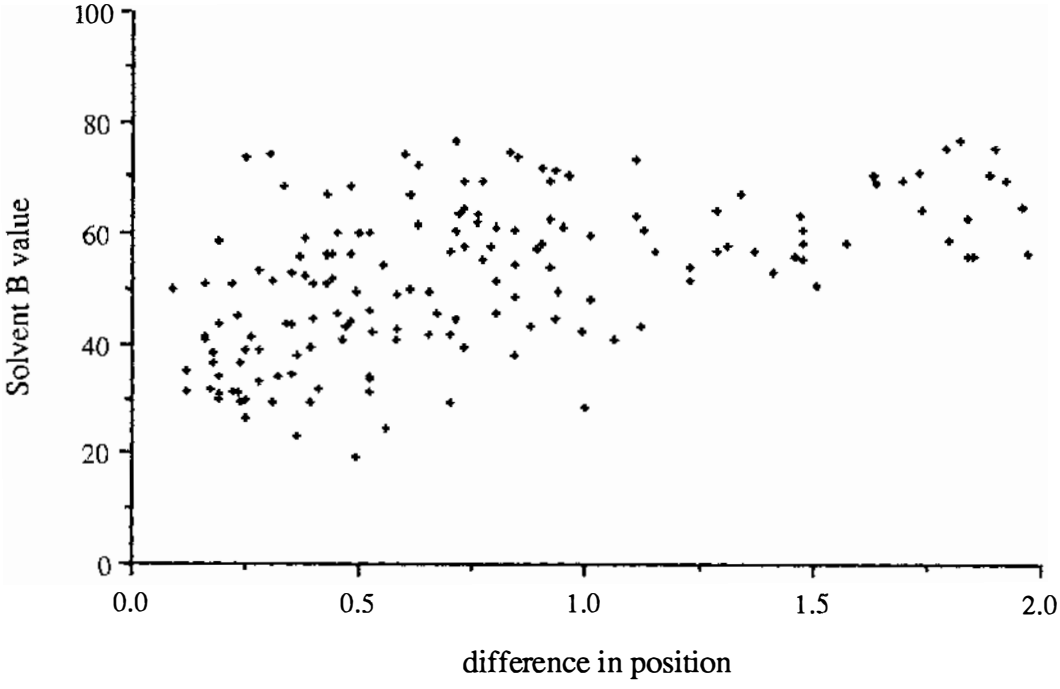
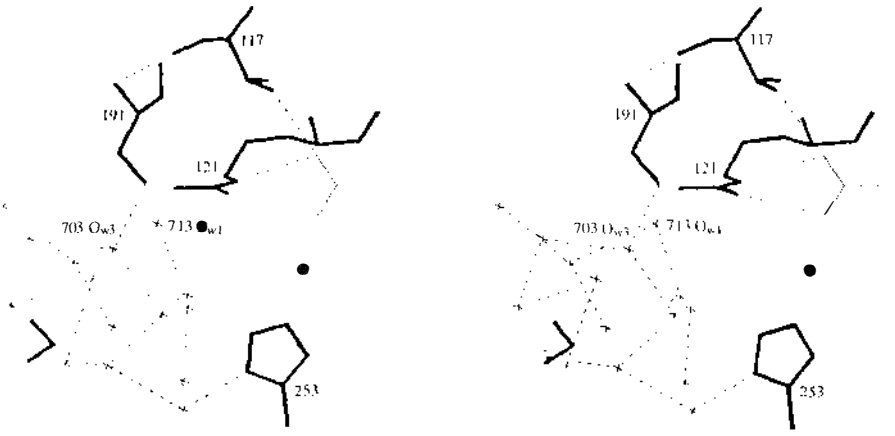



Figure 3.17: Plot of the individual temperature factor as a function of the difference in position between two equivalent solvent molecules in $Cu_2Lf$ and $Fe_2Lf$.

The solvent structure and hydrogen bonding patterns in the vicinity of the metal and anion binding sites is very highly conserved in both $Fe_2Lf$ and $Cu_2Lf$ (see Plate 3.11). Figures 3.18a & b show the solvent structure and hydrogen bonding network near the anion-binding arginine residues for both lobes of $Cu_2Lf$. In the N-lobe, Arg 121 is hydrogen bonded to Ser191 (2.80 Å) and a 1st sphere water molecule (715 $O_{w1}$) from the inter-domain cleft

solvent network, in addition to the (bi)carbonate anion (Figure 3.18a). The Ser191 sidechain also interacts with this extended network via 703 $O_{W3}$ which interacts with two 2nd sphere solvent molecules. In addition, there is a mainchain-mainchain interaction from this residue to Thr117.

(a)




(b)

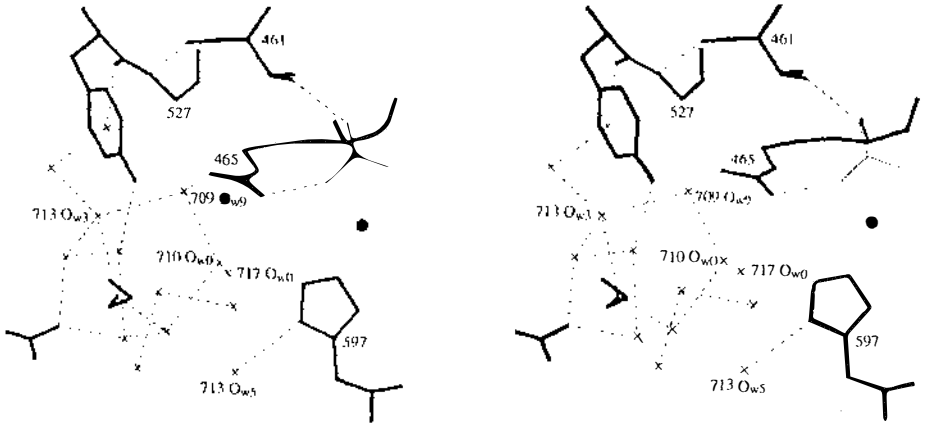



Figure 3.18: The solvent structure in the (a) N-lobe and (b) C-lobe sites of $Cu_2Lf$.

In the C-lobe, Arg465 is not hydrogen bonded to any other part of the protein and makes one hydrogen bond to each of two solvent molecules designated 709 $O_{W9}$ and 717 $O_{W0}$ (Figure 3.18b). The counterpart of Ser191 is a glycine (527) and hence is unable to interact with the sidechain of Arg465. However, the mainchain-mainchain hydrogen bond to the counterpart of Thr117 (461) is present.

Several water molecules appear to play a significant part in the hydrogen bonding interactions in the two binding sites. The active site histidines (His253 and His597) are bonded to the copper through $N_{\varepsilon 2}$ and hydrogen bonded to a water molecule (700 $O_{w2}$ and 713 $O_{w5}$ respectively) through $N_{\delta 1}$. Both of these water molecules interact with the solvent network via two or three other solvent molecules and are also hydrogen bonded to Glu 80 ($O_{e1}$) and Glu413 ($O_{e1}$) respectively. Comparison of the $Cu_2$Lf model with $Fe_2$Lf indicates that these water molecules are in conserved positions (within 0.49 and 0.48 Å respectively).

Another solvent molecule common to both sites and to both structures is hydrogen bonded to the phenolic oxygen of Tyr92 (Tyr453 in the C-lobe). These two waters are designated 700 $O_{w1}$ in the N-lobe and 712 $O_{w5}$ in the C-lobe. They make identical contacts in both lobes, namely to $O_{\delta 1}$ of Asp60 (395), $O_{\gamma 1}$ of Thr122 (466), the carbonyl oxygen of Pro251 (595), in addition to the $O_{\eta}$ atom of Tyr92 (435), and their positions are closely conserved between the $Cu_2$Lf and $Fe_2$Lf structures (0.41 and 0.19 Å respectively). Neither solvent molecule interacts with the interdomain solvent network.

The water molecule designated 700 $O_{w0}$ is present only in the C-lobe of $Cu_2$Lf and is in an equivalent position (within 0.24 Å) in $Fe_2$Lf where it bridges between Lys546 (discussed in Section 3.2.7) and Tyr528. It also interacts with the $O_{\eta}$ of Tyr415 and $O_{\delta 1}$ of Asn466. The equivalent position in the N-lobe is occupied by the guanidium group of Arg210 (discussed in Section 3.2.7).

Of the 37 water molecules (18 in the N-lobe, 19 in the C-lobe) within 10 Å of the corresponding copper atom, only two in each lobe are not in equivalent positions in $Fe_2$Lf. The average difference in the positions of the 33 common solvent molecules is ≈ 0.46 Å. Clearly, the solvent structure in the vicinity of the two metal binding sites in $Cu_2$Lf and $Fe_2$Lf is essentially similar, even though a some of the water molecules are involved in the solvent network and interact to only a very small extent with the protein molecule. The water molecules that do interact with protein atoms are constrained quite tightly by their hydrogen bonds, and the second sphere solvent molecules thus adopt positions dictated, at least in part, by the geometry of their hydrogen bonded contacts with the first hydration sphere. Obviously differences in the closure of the two domains would greatly affect the structure of

the water molecules in the interdomain cleft and in the area around the metal and anion binding sites and it is expected that metal ions or anions which influence the closure of the lobes would result in an interdomain solvent structure significantly different from that observed in $Fe_2$Lf and $Cu_2$Lf.

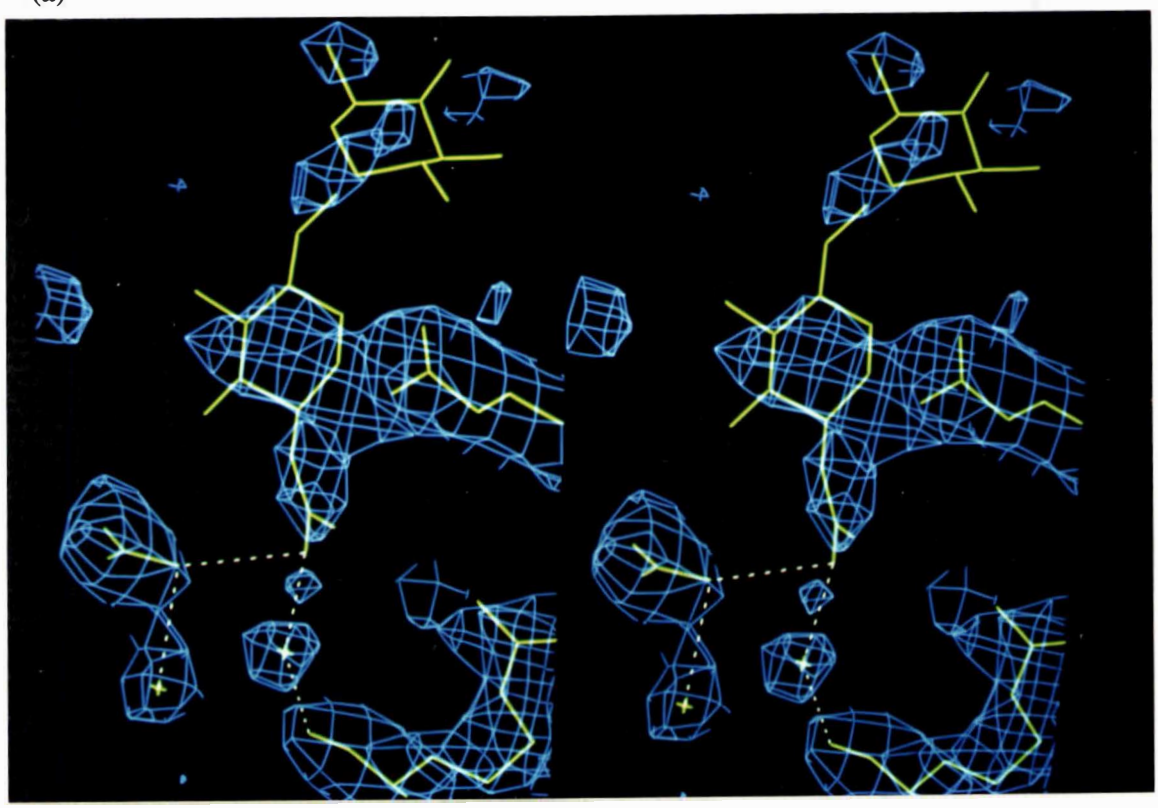
3.2.6 *The glycan chains:*

The first 3 sugar residues (two N-acetylglucosamine and one fucose) in the N-lobe and the first two (one NAG and one FUC) in the C-lobe are included in the model (based on the published carbohydrate sequence by Spik *et al.*, 1982) although the electron density for the second and third residue of each chain is very weak (Plates 3.7a & b). No interpretable density was visible beyond them, and the carbohydrate structure is clearly flexible and poorly ordered. This is reflected in the individual temperature factors of the atoms in the glycan residues, which range from 75 $Å^2$ to 90 $Å^2$ (the average for all five residues is ≈ 85 $Å^2$). The B values for the two asparagine-linked NAG residues are somewhat lower (75 - 80 $Å^2$) than the two FUC molecules and the second N-lobe NAG (85 - 90 $Å^2$). Although it is difficult to estimate the error in the positional parameters for the glycan residues, interpolation from the average mainchain rms deviation versus average mainchain B graph (see Figure 3.15a) would put the error in the range 0.6 - 0.7 Å.

The first NAG moiety (990 in the present numbering) in the N-lobe is covalently bound to $N_{\delta 2}$ of Asn137, with the second NAG (991) and the fucose (FUC992) residues bound to it by β- and α-linkages respectively (joined via the $O_4$ and $O_6$ positions). NAG990 is involved in two favourable hydrogen bonds (both via $O_7$) to $N_{\varepsilon 1}$ of Gln110 and the the mainchain carbonyl oxygen of Phe135, while FUC991 and NAG992 do not interact with any protein atoms. Two solvent molecules are hydrogen bonded to NAG990 (Plate 3.7a).

The C-lobe carbohydrate residues, NAG995 which is covalently bound to Asn478, and the α-linked FUC996, do not come within hydrogen bonding distance of any part of the

(a)



(b)

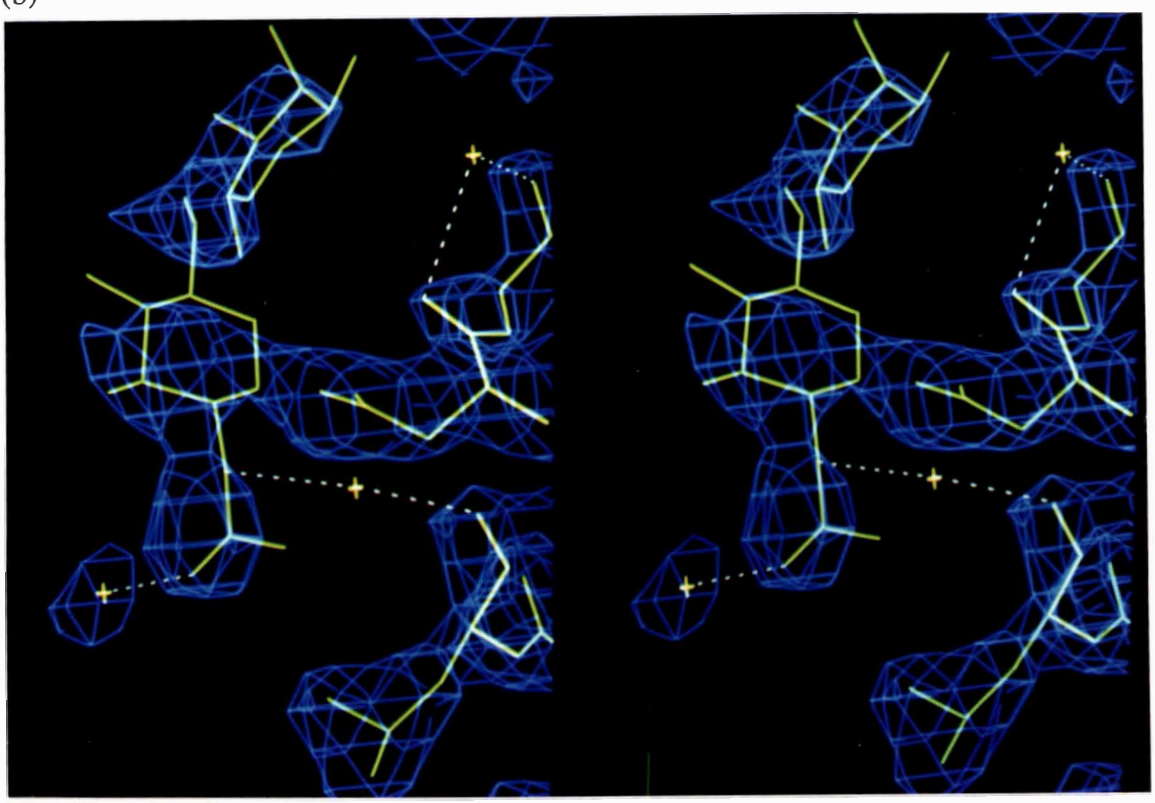


Plate 3.7: Stereo views of the 2Fo-Fc electron density associated with the asparagine-linked N-acetylglucosamine residues (centre) and the fucose residues (top) in the (a) N-lobe and (b) C-lobe of $Cu_2Lf$. Note that the electron density for the fucose residues is very weak, whereas that for the NAG is much better defined. The sidechain of Gln110 is at the bottom left in photograph (a).

protein molecule, although two solvent molecules are within hydrogen bonding distance of the NAG995 (Plate 3.7b), one via N2 (designated 726 $O_{w3}$) and bridging to the carbonyl oxygen of Leu475 and the other 726 $O_{w4}$ (via $O_7$) having no other hydrogen bonding interactions.

### 3.2.7 *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 $Cu_2$Lf electron density. The first $2F_o$-$F_c$ electron density map, calculated after the first phase of refinement, clearly showed the location of the ligands and 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 into the appropriate positions. Further refinement 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), Thr 117 (461) and the N-terminus of helix 5. As in the $Fe_2$Lf structure (Anderson *et al.*, 1989) this was interpreted as the site of the bound $CO_3^{2-}$ (or $HCO_3^-$) ion. In contrast to $Fe_2$Lf, 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 towards Thr 461 and helix 5, consistent with an anion bound in a bidentate fashion to the metal (Figure 3.19b). In the N-terminal site, however, the density was bifurcated in a manner suggesting monodentate coordination to the copper (Figure 3.19a). In addition, 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 (Tyr 92) bond (Figure 3.19a).

In the C-terminal site, the carbonate was initially fitted in a symmetrical bidentate fashion (Cu-O distances of 2.0 Å) that gave favourable hydrogen bonded contacts with four residues, Thr461, Arg465, and the mainchain of residues 467 and 468. The copper was placed such that distances to the side chains of Asp395, Tyr435, Tyr528 and His597 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 at least four hydrogen bonds with Thr117, Arg121, and the mainchain of residues 123 and 124. A bidentate carbonate could not be adequately fitted to the available density and would not interact favourably 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.

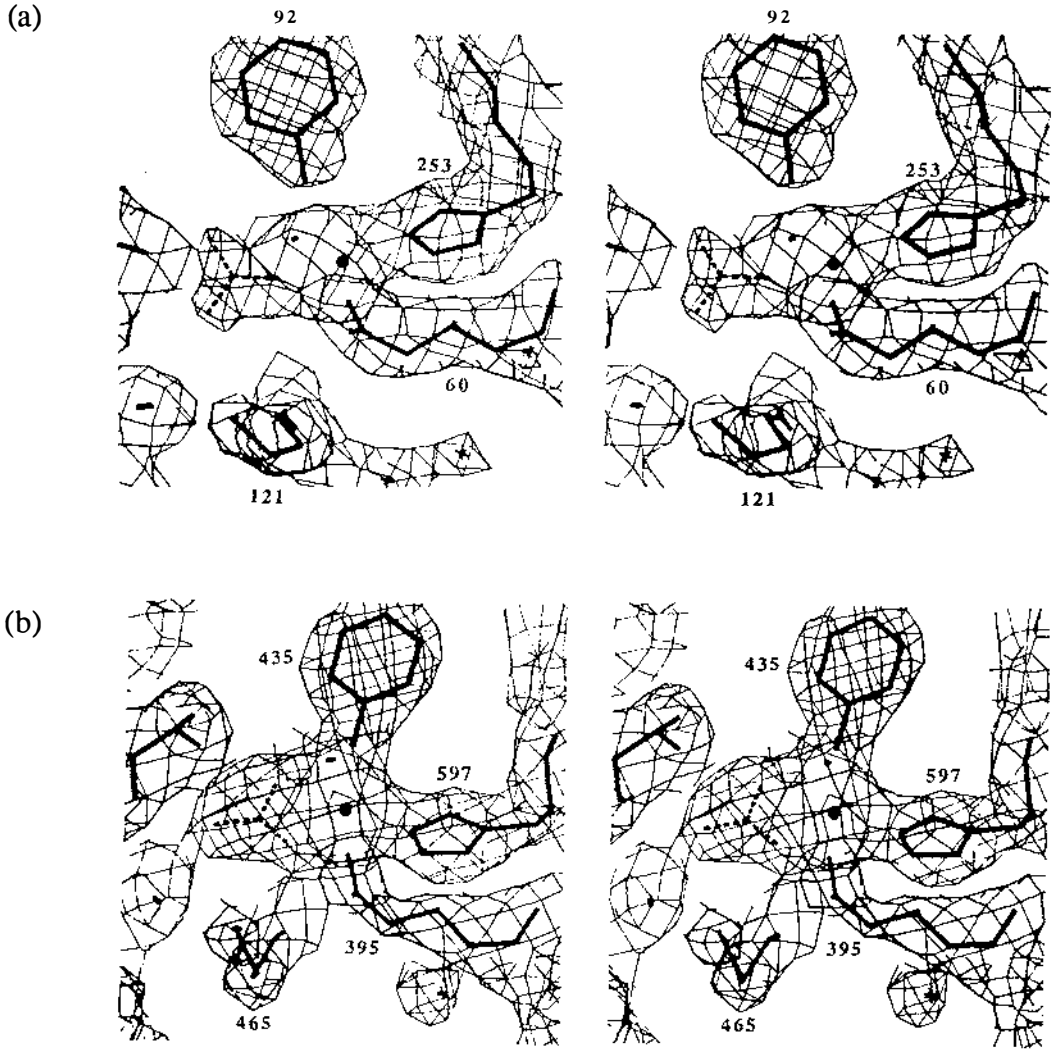



Figure 3.19: Stereo views of $2F_o$-$F_c$ electron density in (a) the N-lobe binding site and (b) the C-lobe binding site of $Cu_2Lf$, showing the density associated with three of the protein ligands (Asp60 (395), Tyr 92 (435) and His 253 (597), the copper atom and the carbonate anion. The location of the anion-binding arginine residues (Arg121 (465)) are also indicated. This map was calculated at the time the anions (dashed lines) were first built into the model (R = 0.276).

In the final model for $Cu_2Lf$, 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. The error in the Cu-ligand and Cu-anion bond distances is estimated at ≈ 0.2 Å. A list of relevant bond lengths and angles in the two copper sites are given in Tables 3.16 and 3.17 respectively, with atom labelling as in Figures 3.20a & b. Stereo views of the two sites are shown in Plates 3.8 and 3.9.

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 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 was consistently shifted towards to a longer value. The restraints on this distance were removed to allow the side chain to move to its most favoured position. The constraints on the other bond distances were maintained, however. Finally the restraints were removed from all the bond distances at the end of the refinement and very little change was observed.

*Hydrogen bonding:* The hydrogen bonding pattern around each anion is similar to those observed in the diferric structure (Anderson *et al.*, 1989). The C-terminal carbonate makes five favourable hydrogen bonded contacts (Plate 3.10b and Table 3.18), two with Arg465 ($N_{\eta 2}$ and $N_{\varepsilon}$), one with Thr461 ($O_{\gamma 1}$) and two with mainchain nitrogens from Ala467 and Gly468 (N-terminus of helix 5). In the N-terminal site, the anion is bound in a monodentate fashion, forming five hydrogen bonds (Plate 3.10a and Table 3.18), two to Arg121 (both to

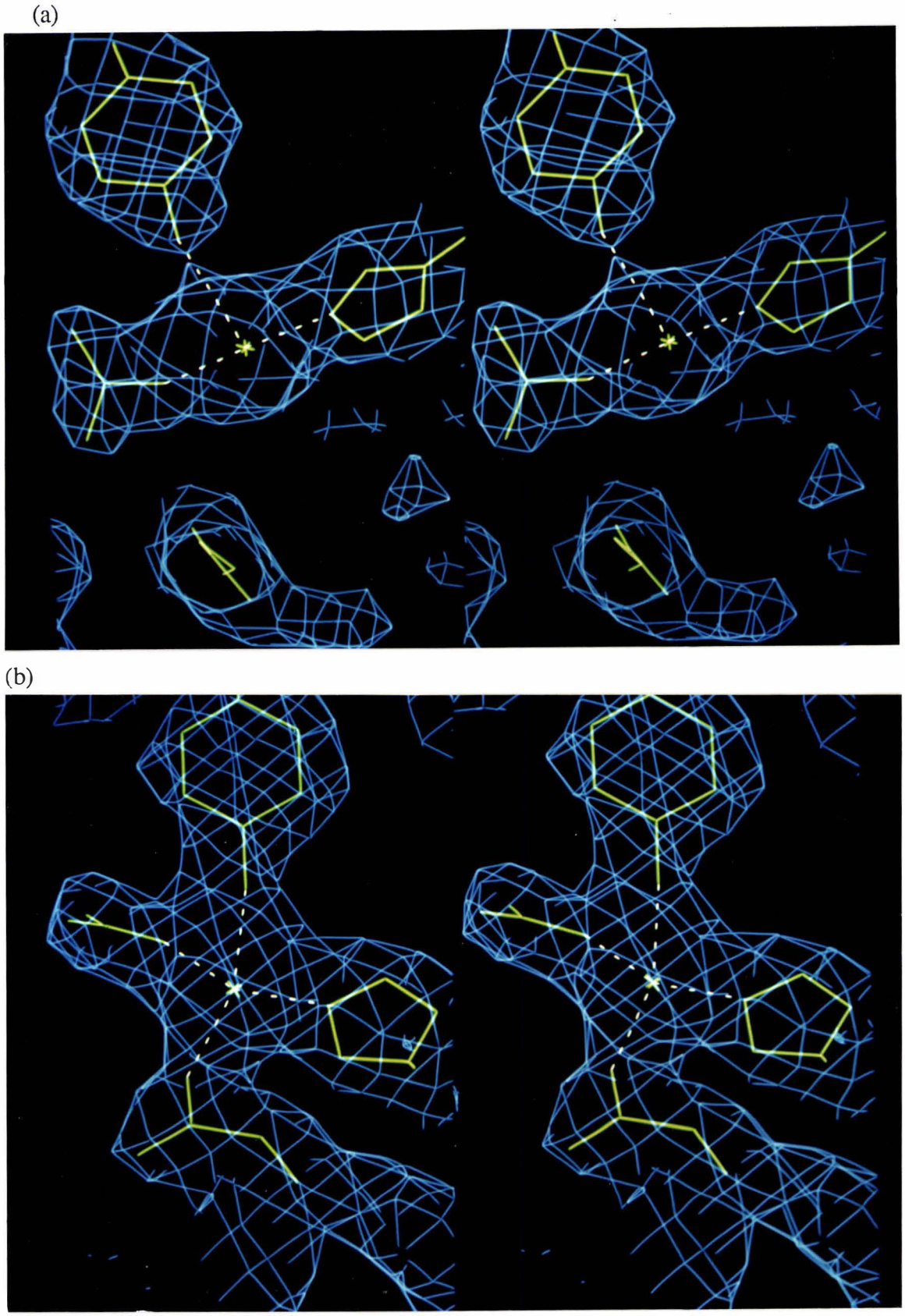
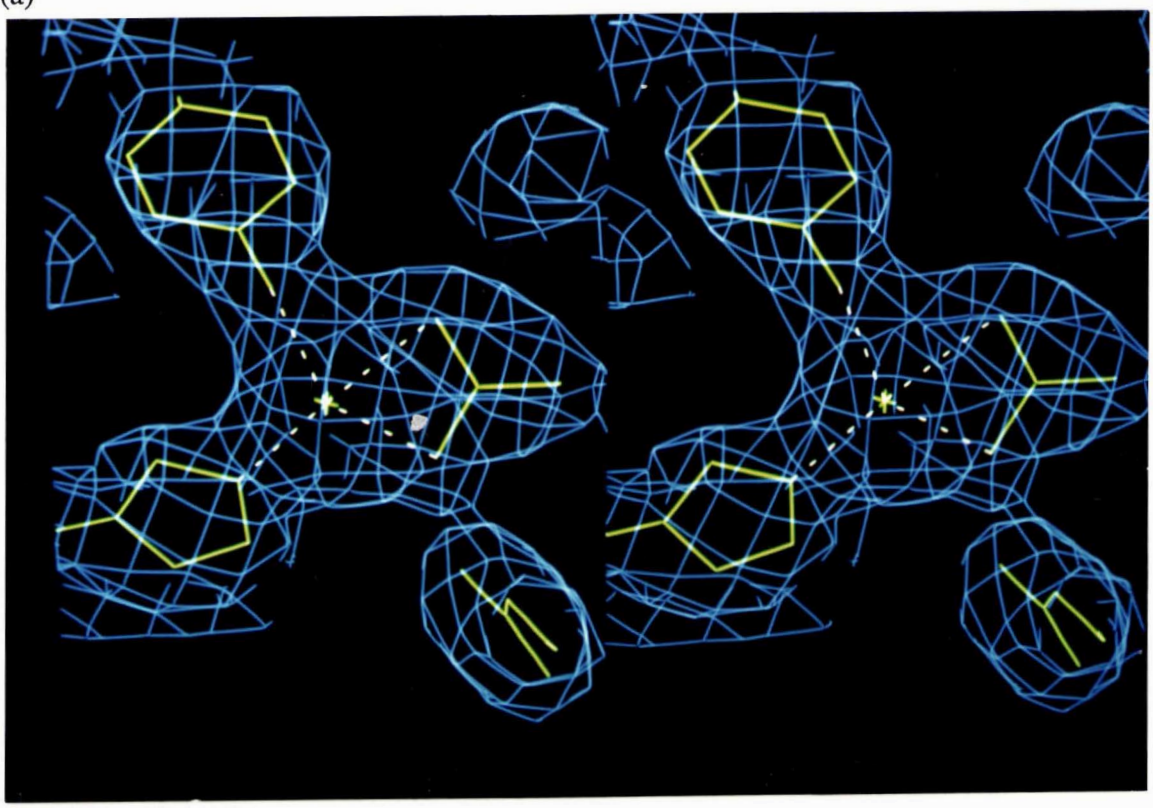



Plate 3.8: Stereo views of the $2F_O$-$F_C$ electron density in the N-lobe binding sites of $Cu_2$Lf. The final model of the $Cu_2$Lf structure is superimposed in green. Photograph (a) shows the copper atom, the monodentate carbonate anion, His253 and Tyr92. A rotation of 90° (b) allows the other two ligands, Tyr192 and Asp60, to be seen.

(a)



(b)

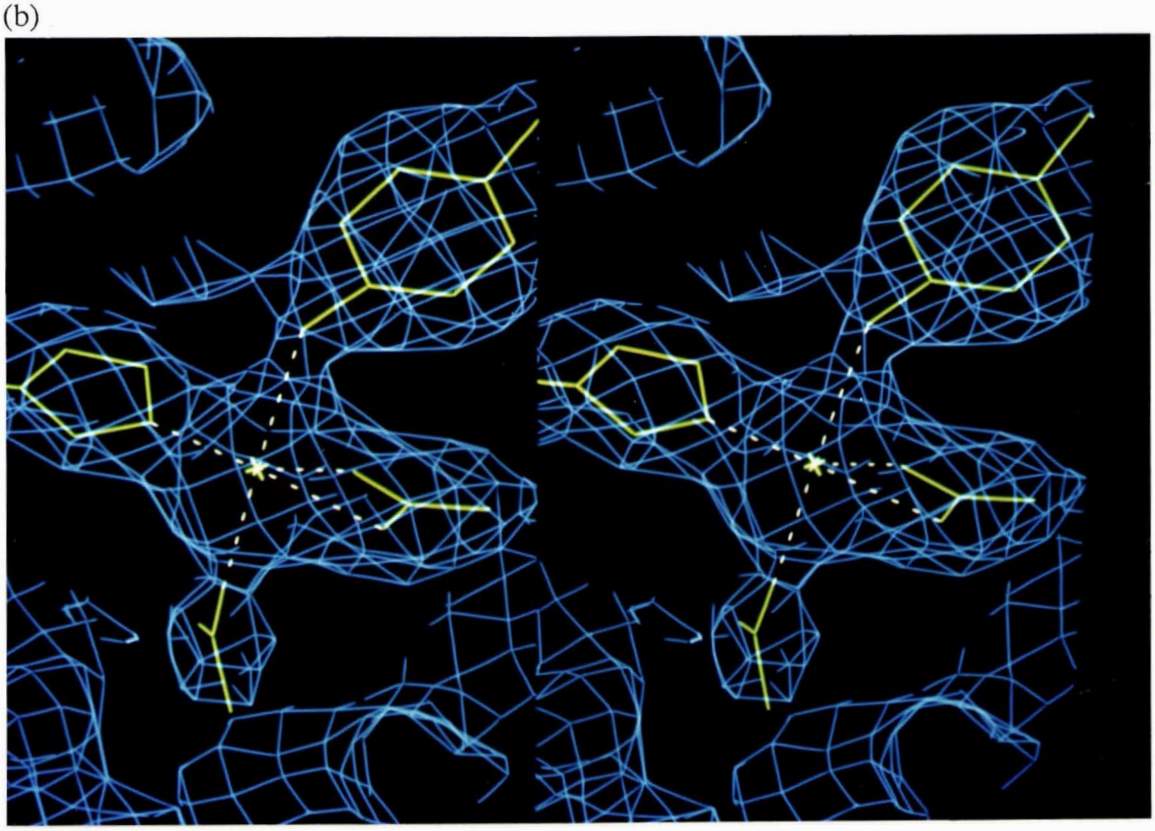


Plate 3.9: Stereo views of the $2F_o$-$F_c$ electron density in the C-lobe binding sites of $Cu_2$Lf. Photograph (a) shows the copper atom, the bidentate carbonate anion, His597 and Tyr435. A rotation of 90° (b) allows the other two ligands, Tyr528 and Asp395, to be seen.

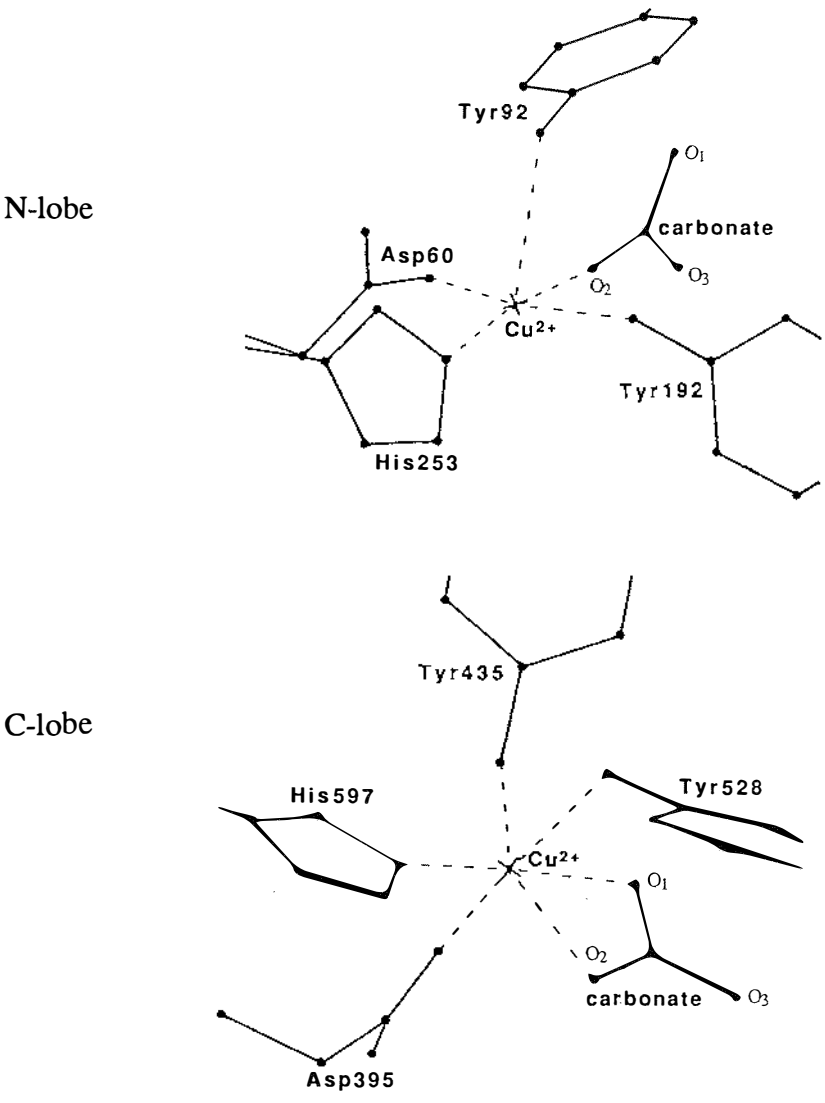



Figure 3.20: The coordination geometry and atom labelling scheme in the N- and C-lobe binding sites of $Cu_2Lf$.

the $N_\varepsilon$ atom of the guanidinium group), two to mainchain nitrogen atoms on helix 5 (residues 123 and 124) and one to the sidechain of Thr117.

The aspartic acid residues involved in metal binding (Asp60 and Asp395) are involved in three hydrogen bonded interactions (Tables 3 and 4, Appendix V)), one via the mainchain nitrogen to the solvent network, one from $O_{\delta 2}$ to the mainchain nitrogen of Gly62, and a bifurcated hydrogen bond at the $O_{\delta 2}$ atom to the mainchain and sidechain of Thr 122 (466 in the C-lobe). This latter interaction is one of the few inter-domain hydrogen bonds in the vicinity of the metal binding site (Asp60 and Asp395 are in the N1 and C1 domains, while

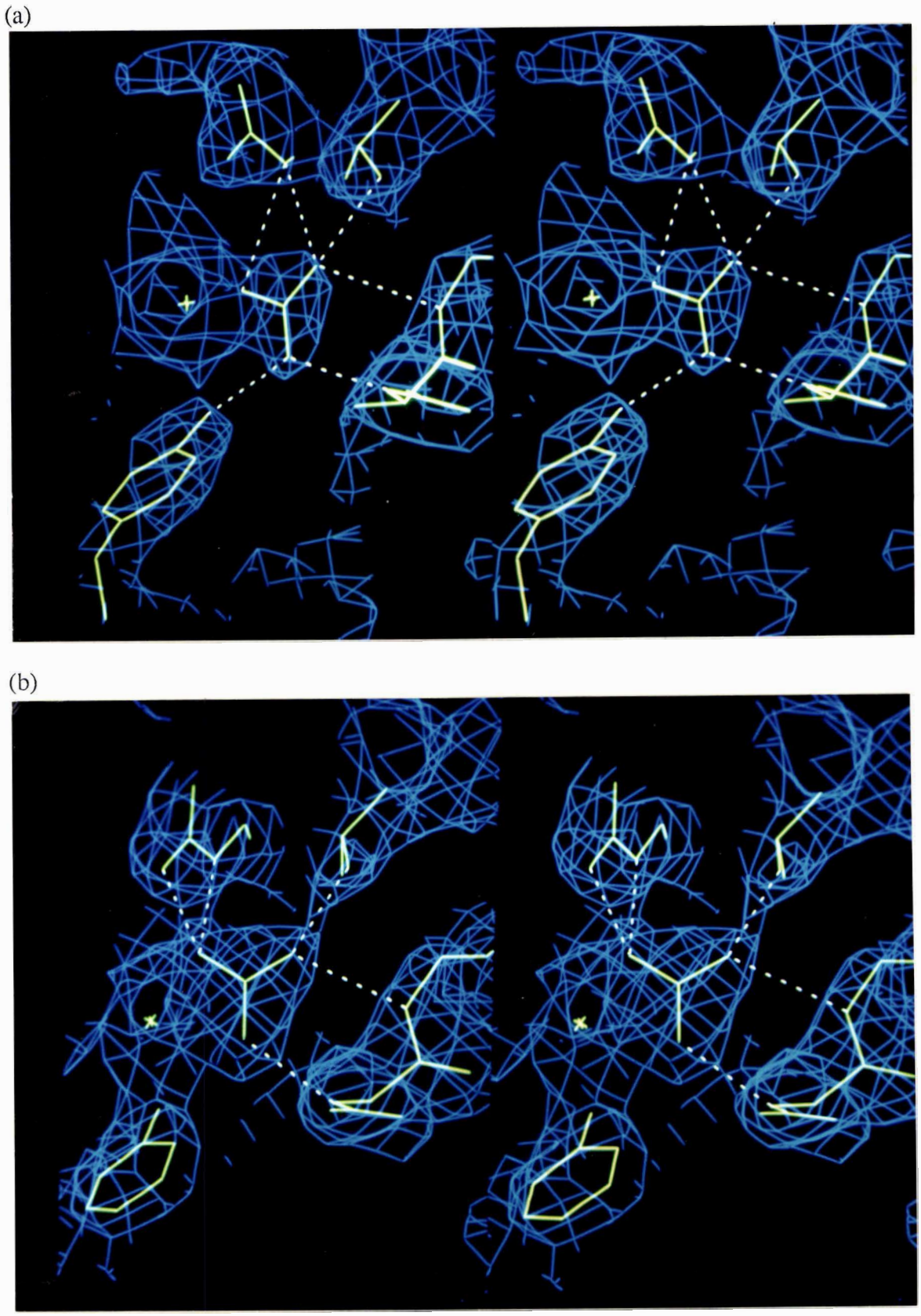



Plate 3.10: Hydrogen bonding interactions with the (a) N-lobe and (b) C-lobe carbonate ions. Note that in addition to the five hydrogen bonds to the anion binding residues (117, 121, 123 and 124, as listed in Table 3.18) the N-lobe anion also interacts with the $O_\eta$ of Tyr92, implying that either the tyrosine or the anion is protonated.

Table 3.17A: Copper coordination sphere ($Cu_2Lf$) - bond lengths

| Bond [a] | N-lobe site | C-lobe site |
|---|---|---|
| Cu - O60(395) | 2.0 Å | 2.0 Å |
| Cu - O92(435) | 2.8 Å | 2.1 Å |
| Cu - O192(528) | 2.0 Å | 2.4 Å |
| Cu - N253(597) | 2.0 Å | 2.1 Å |
| Cu - $O_1$(anion) | — | 2.2 Å |
| Cu - $O_2$(anion) | 1.9 Å | 2.0 Å |

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

Table 3.17B: Copper coordination sphere ($Cu_2Lf$) - bond angles

| Angle [a] | N-lobe site | C-lobe site |
|---|---|---|
| O60(395)-Cu-O92(435) | 84° | 82° |
| O60(395)-Cu-O192(528) | 158° | 170° |
| O60(395)-Cu-N253(597) | 112° | 82° |
| O60(395)-Cu-$O_1$(anion) | — | 84° |
| O60(395)-Cu-$O_2$(anion) | 85° | 81° |
| O92(435)-Cu-O192(528) | 83° | 106° |
| O92(435)-Cu-N253(597) | 102° | 106° |
| O92(435)-Cu-$O_1$(anion) | — | 72° |
| O92(435)-Cu-$O_2$(anion) | 76° | 134° |
| O192(528)-Cu-N253(597) | 84° | 102° |
| O192(528)-Cu-$O_1$(anion) | — | 92° |
| O192(528)-Cu-$O_2$(anion) | 75° | 89° |
| N253(597)-Cu-$O_1$(anion) | — | 166° |
| N253(597)-Cu-$O_2$(anion) | 159° | 113° |
| $O_1$(anion)-Cu-$O_2$(anion) | — | 63° |

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

the Thr122 and Thr466 are in the N2 and C2 domains), as most of the other hydrogen bonds in this region occur between groups in the same domain. It is likely that these hydrogen bonding interactions are important to the stability of the closed structure of the lobes, in that the aspartic acid carboxylate group intervenes between two helix N-termini (helices 3 and 5). Without this type of interaction, the positively-charged helix N-termini (Hol *et al.*, 1982) would tend to repel each other, thus decreasing the stability of the closed unliganded structure. These interactions do not exist in the open state (see Table 3, Appendix V), as the Asp60 and Thr122 residues in the N-lobe of the apoLf structure are about 10 Å apart.

Table 3.18: Hydrogen bonding interactions between the carbonate anions and protein residues in $Cu_2Lf$.

| | Hydrogen bonds | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) | Average O···D(Å) [a] |
|---|---|---|---|---|---|---|---|
| | $O_1$ - N (123) | 1.67 | 105 | 151 | 2.59 | 18 | |
| | $O_2$ - $N_\eta$ (121) | 2.49 | 95 | 138 | 3.30 | 30 | |
| N | $O_3$ - $N_\varepsilon$ (121) | 2.13 | 109 | 169 | 3.12 | 8 | 3.03 |
| | $O_3$ - N (124) | 2.65 [b] | 109 | 133 | 3.41 | 35 | |
| | $O_3$ - $O_{\gamma 1}$ (117) | — | 169 [b] | 115 [b] | 2.71 [b] | — | |
| | $O_1$ - N (467) | 1.98 | 137 | 137 | 2.79 | 29 | |
| C | $O_2$ - $N_\varepsilon$ (465) | 1.79 | 109 | 130 | 2.55 | 33 | 2.69 |
| | $O_2$ - $N_\eta$ (465) | 1.69 | 142 | 132 | 2.47 | 30 | |
| | $O_3$ - N (468) | 2.20 | 116 | 149 | 3.10 | 21 | |
| | $O_3$ - $O_{\gamma 1}$ (461) | — | 141 [b] | 99 [b] | 2.56 [b] | — | |

[a] The average value as calculated from the individual hydrogen bond distances. It gives a rough measure of the size of the anion binding pocket. [b] The O...H distance of 2.65 Å is too long to be a true hydrogen bond although the geometry at the hydrogen atom is favourable and some interaction between the two atoms is likely. [c] These values relate to the $C_1$-$O_3$...$O_{\gamma 1}$ and $C_\beta$-$O_{\gamma 1}$...$O_3$ angles and the $O_3$...$O_{\gamma 1}$ distance.

There are a number of interactions between groups in the two domains and residues belonging to the cross-over strands behind the metal binding sites. These are listed in Tables 3 and 4 in Appendix V, along with the interdomain contacts (hydrogen bonds and salt bridges). These interactions could also be important to the relative stabilities of the open and

closed forms, as several of these hydrogen bonds do not exist when the lobe is in an open conformation, as indicated in Appendix V.

*Differences in amino acid residues between the two binding sites:* The only changes in the sequence between the N-lobe and C-lobe in the vicinity of the metal and anion binding sites are at (i) the N-terminus of helix 7, where Ser191 in the N-lobe is a glycine (527) in the C-lobe and (ii) residue 210 (546), where an arginine in the N-lobe is replaced by a lysine. The sequence change at 191 (527) has been discussed previously in this chapter. In the latter case (Arg → Lys) although both residues are basic, they differ in the size of their sidechains and in the hydrogen bonds they make with the surrounding protein structure.

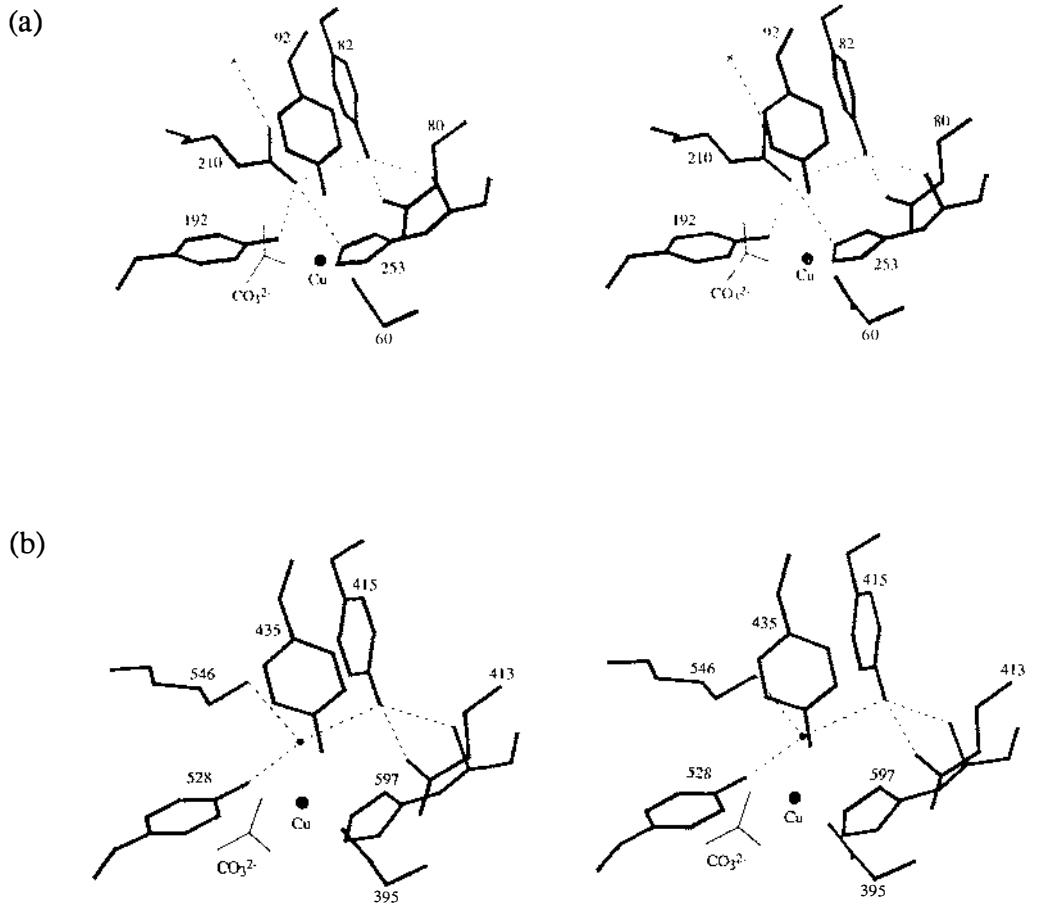



Figure 3.21: The hydrogen bonding interactions involving (a) Arg210 and (b) Lys546.

Arg210 is hydrogen bonded via $N_{\eta 2}$ to both $O_{\eta}$ of Tyr192 (N ··· O = 2.88 Å) and $N_{\varepsilon 2}$ of His253 (N ··· N = 3.20 Å), and in addition, the guanidinium group lies within ≈4 Å of the $O_{\eta}$ atom of Tyr92. In addition, there is a hydrogen bond between $N_{\eta 2}$ and $O_{\eta}$ of Tyr82 (N ··· O = 3.25 Å), which is another inter-domain interaction. In the C-lobe, Lys546 does not interact directly with any of the metal-binding ligands but is bridged to Tyr528 via a solvent molecule. Tyr415 (which corresponds to Tyr82) does not come within hydrogen bonding distance of Lys546. Figures 3.21a & b (previous page) show the hydrogen bonding interactions of Arg210 and Lys546.

It should be noted that even though there are distinct differences in the interactions involving Arg210 and Lys546, both residues are adjacent to the solvent-filled interdomain cleft. Although movement of the Tyr192 sidechain appears to be restricted by the interaction with Arg210, the arginine sidechain should be readily able to move away from the metal binding site to accommodate any alteration in the position of the tyrosine.

### 3.2.8 *Comparison 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 lack of structural changes beyond the metal 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 (Plates 3.11a & b). 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 initially estimated to be in the order of 0.3 - 0.8Å.

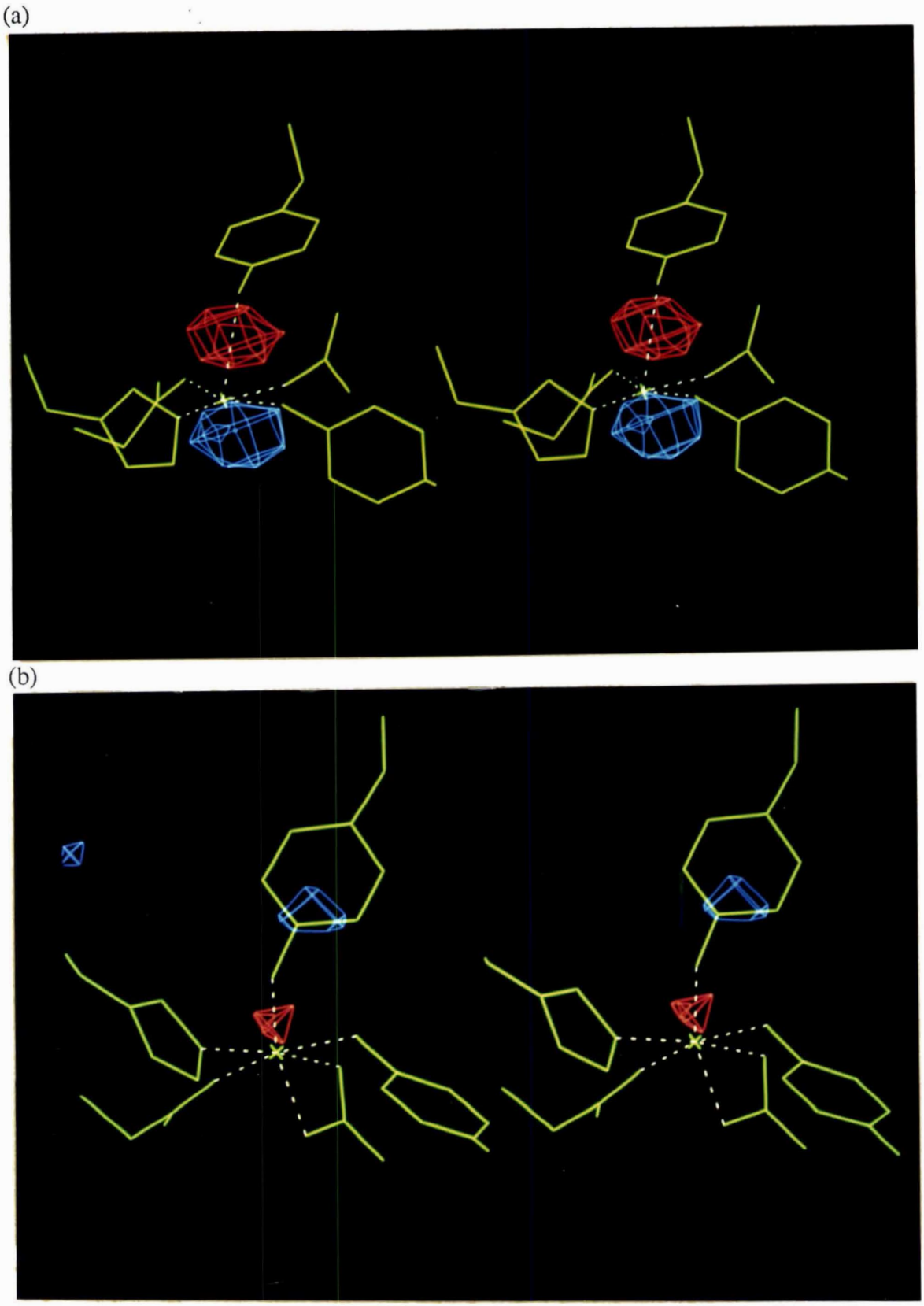



Plate 3.11 $Cu_2Lf$ - $Fe_2Lf$ difference electron density map (calculated with coefficient $\Delta F = F_{(Cu_2Lf)} - F_{(Fe_2Lf)}$ ) in the vicinity of (a) the N-lobe and (b) the C-lobe metal binding sites. Regions of negative density are shown in red and positive density in blue. The coordinates of $Cu_2Lf$ are superimposed in green.

*The polypeptide chain:* A series of superpositions between $Cu_2Lf$ and $Fe_2Lf$ (summarised in Table 3.19) also show the close similarity in the polypeptide fold between these two structures. The overall rms deviation in $C_\alpha$ positions for all 691 residues is 0.71 Å. The figure is exaggerated, however, by the large contributions from the first five residues at the N-terminus which have an average rms deviation in $C_\alpha$ positions of almost 5 Å. The overall rms value drops to 0.25 Å (for 590 $C_\alpha$ atoms) when the flexible regions of the structure (as defined in Table 3.20A) are excluded. The rms deviation in the mainchain atom positions is 0.28 Å (for 2360 atoms, again with the flexible regions removed). When the two lobes are superimposed individually in terms of the secondary structure elements ($C_\alpha$ atoms in the helices and sheets), the resulting rms deviations are 0.27 Å for the N-lobe (191 atoms) and 0.23 Å for the C-lobe (192 atoms). Clearly, the C-lobe of $Cu_2Lf$ seems to fit better to its counterpart in $Fe_2Lf$ than does the N-lobe, either because the substitution of $Cu^{2+}$ for $Fe^{3+}$ has induced changes in the extent closure of the N-lobe, or because of the greater flexibility of portions of the N-lobe compared with corresponding regions in the C-lobe, as noted earlier (Section 3.2.4) and as indicated by the results of superpositions of the individual domains of $Cu_2Lf$ on the corresponding domains of $Fe_2Lf$ (see Table 3.21). The latter possibility is in accord with the observation that the average mainchain temperature factors in the N-lobe are somewhat larger than those in the C-lobe (see Section 3.2.4), although differences in domain closure could still be important.

Table 3.19: Superposition of $Cu_2Lf$ on $Fe_2Lf$

| Atoms used in superposition | rms deviation (Å) | |
|---|---|---|
| | $C_\alpha$ (atoms) | main [a] (atoms) |
| all 691 residues | 0.71 (691) | 0.69 (2764) |
| - with N- and C-termini removed | 0.31 (682) | 0.34 (2728) |
| - with flexible regions removed [b] | 0.25 (590) | 0.28 (2360) |
| secondary structure elements only | 0.26 (383) | 0.28 (1532) |
| N-lobe onto N-lobe - secondary structure [c] | 0.27 (191) | 0.31 (764) |
| C-lobe onto C-lobe - secondary structure [c] | 0.23 (192) | 0.24 (768) |

[a] Mainchain: amide nitrogen, carbonyl oxygen, carbonyl carbon and alpha carbon. [b] As given in Table 3.20A. [c] The flexible regions in each lobe were excluded from the calculations.

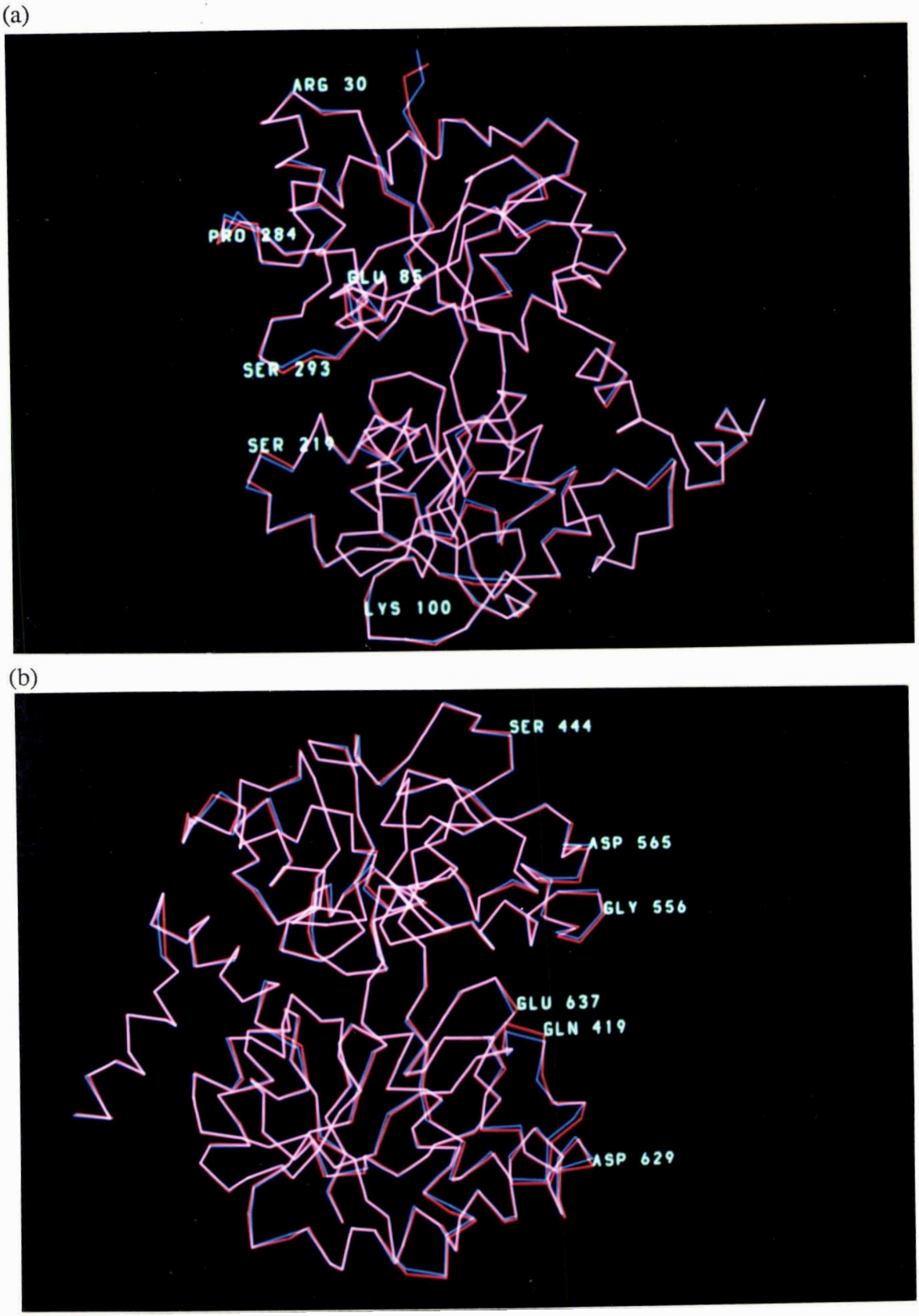



Plate 3.12: Superposition of (a) the N-lobe of $Cu_2Lf$ (blue) on to the N-lobe of $Fe_2Lf$ (red) and (b) the C-lobe of $Cu_2Lf$ (blue) on to the C-lobe of $Fe_2Lf$ (red). The areas of the two structures which appear to have the worst agreement are labelled in white (see also Table 3.20A).

Table 3.20A: Regions of the model with high rms deviations in $C_\alpha$ and mainchain atom positions between $Cu_2Lf$ and $Fe_2Lf$ [a]

| N-lobe | rms displacement | | C-lobe | rms displacement | |
|---|---|---|---|---|---|
| | $C_\alpha$ (Å) | main (Å) | | $C_\alpha$ (Å) | main (Å) |
| 1 - 5 | 2.50 | 2.29 | — | — | — |
| 20 - 30 | 0.48 | 0.54 | — | — | — |
| 83 - 87 | 0.67 | 0.71 | 416 - 424 | 0.68 | 0.67 |
| — | — | — | 480 - 482 | 0.73 | 0.65 |
| 176 - 177 | 0.61 | 0.55 | 511 - 513 | 0.55 | 0.58 |
| 218 - 225 | 0.58 | 0.60 | 554 - 560 | 0.71 | 0.66 |
| 280 - 287 | 0.64 | 0.76 | 623 - 631 | 0.51 | 0.51 |
| 291 - 296 | 0.55 | 0.51 | 636 - 637 | 0.56 | 0.51 |
| — | — | — | 654 - 655 | 0.47 | 0.52 |

a The residues in the N-lobe have been aligned with corresponding residues in the C-lobe (based on the sequence alignment in Anderson *et al.* (1989)). Where there is a gap, then the correponding residues do not have unduly high rms deviations and where there is an asterisk, there is no corresponding piece of polypeptide chain.

Table 3.20B: Numbers of residues with high rms deviations [a]

| | residues with rms deviations greater than: | | |
|---|---|---|---|
| | 0.35 Å | 0.50 Å | 1.0 Å |
| N-lobe | 77 | 31 | 7 |
| C-lobe | 64 | 25 | 0 |
| whole molecule | 141 | 56 | 7 |

a Between mainchain atom positions only.

Plates 3.12a & b show the results of the superposition of the refined $Cu_2Lf$ model onto the $Fe_2Lf$ model in terms of the $C_\alpha$ atoms. The majority of the $C_\alpha$ atoms in the $Cu_2Lf$ model occupy essentially the same relative positions as the corresponding atoms in $Fe_2Lf$. In the vicinity of the metal and anion binding sites, for example, the rms differences in $C_\alpha$ and mainchain atom positions between the two structures are quite low (see later).

The largest differences in $C_\alpha$ and mainchain positions between the two structures are at the N-terminus (rms deviations between 3 and 7 Å), which has been rebuilt substantially in both $Cu_2Lf$ and $Fe_2Lf$ (B.F. Anderson, personal communication) and which is very poorly defined by the electron density. In addition, external loop regions involving residues 83 -

87, 176 - 177, 218 - 222, 280 - 295, 416 - 424, 480 - 482, 511 - 513, 555 - 556 and 624 - 630 also show rms differences in $C_\alpha$ position well above the maximum average coordinate error in the $Cu_2Lf$ model (0.25 - 0.30 Å, see Figure 3.8). Figures 3.22a & b show the

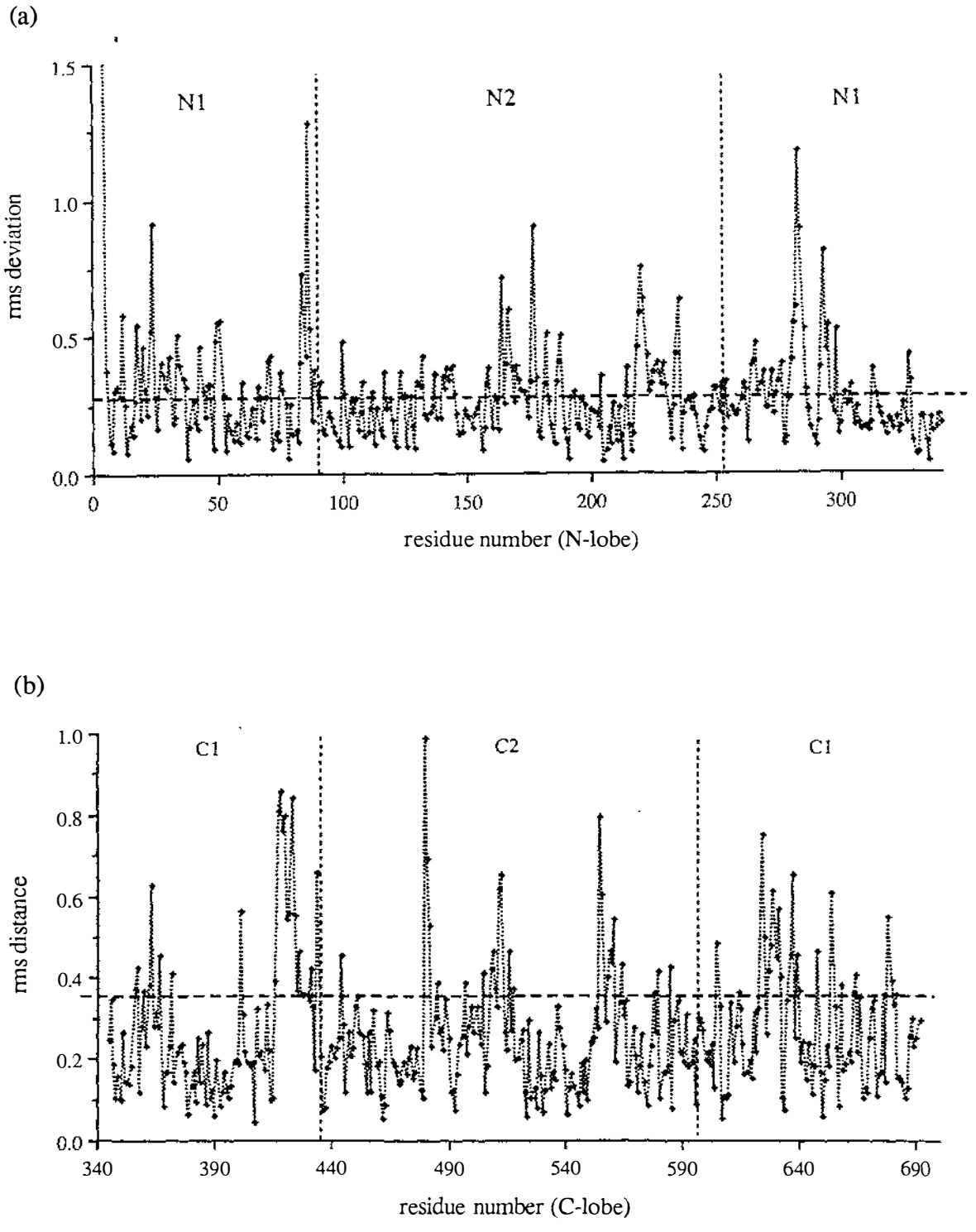



Figure 3.22: A plot of the rms deviation in $C_\alpha$ position between $Cu_2Lf$ and $Fe_2Lf$ for (a) the N-lobe and (b) the C-lobe.

variation of the rms deviation between the $C_\alpha$ positions in $Cu_2Lf$ and $Fe_2Lf$ as a function of residue number. It should be noted that the almost all the parts of the model which show the largest deviations in atomic positions between the two structures also have the lowest real space correlation coefficients as calculated with O (see Section 3.2.1).

These regions correlate well with the areas of the model which have previously been defined as being flexible or disordered based on the analysis of their temperature factors (see Figure 3.13a & b and Table 3.14B). The majority of these "flexible" regions occur on the surface of the molecule in areas exposed to the solvent. Superposition of the N- and C-lobes of $Cu_2Lf$ show that these poorly-defined regions are in similar positions in both halves of the molecule (see Plate 3.4) and that, apart from the N-terminus and the initial part of domain N1, there does not appear to be any significant difference in the flexibility of corresponding regions of the two lobes.

*The lobe and domain structure:* To analyse the relative domain movements between $Cu_2Lf$ and $Fe_2Lf$, a series of superpositions were carried out as follows: (1) the whole $Cu_2Lf$ molecule was superimposed on to $Fe_2Lf$, matching the $C_\alpha$ atoms in the N1 domain only (5 to 92 and 254 to 335); (2) the set of shifted $Cu_2Lf$ coordinates obtained from this superposition was then rotated to match the Cα atoms of the N2 domains (93 to 253), thus giving the relative rotation between domains N1 and N2. This step was repeated, rotating on to the C1 and C2 domains in turn, to give relative rotations between the N1 and the C1 and C2 domains. By repeating this process based on the N2 and C1 domains, a series of relative rotations between the four domains could be obtained. The results of these calculations are summarised in Table 3.21. It is clear that there is very little relative domain movement resulting from the substitution of copper(II) for iron(III).

The largest change appears to be in the N-lobe, where the difference in the relative orientations of the N1 and N2 domains between $Cu_2Lf$ and $Fe_2Lf$ is $\approx 0.6°$. This is larger than the corresponding C2-C1 rotation of 0.2° and could reflect a slight difference in the closure of the domain. However, this observed difference in the closure could be the result of the higher rms deviation in $C_\alpha$ positions between the N1 domains of the two structures.

The N1 domain itself does not superimpose onto the N1 domain in $Fe_2Lf$ as closely as the other three domains, possibly because of its somewhat greater mobility (discussed in Section 3.2.4).

Table 3.21: Domain relationships between $Cu_2Lf$ and $Fe_2Lf$ [a]

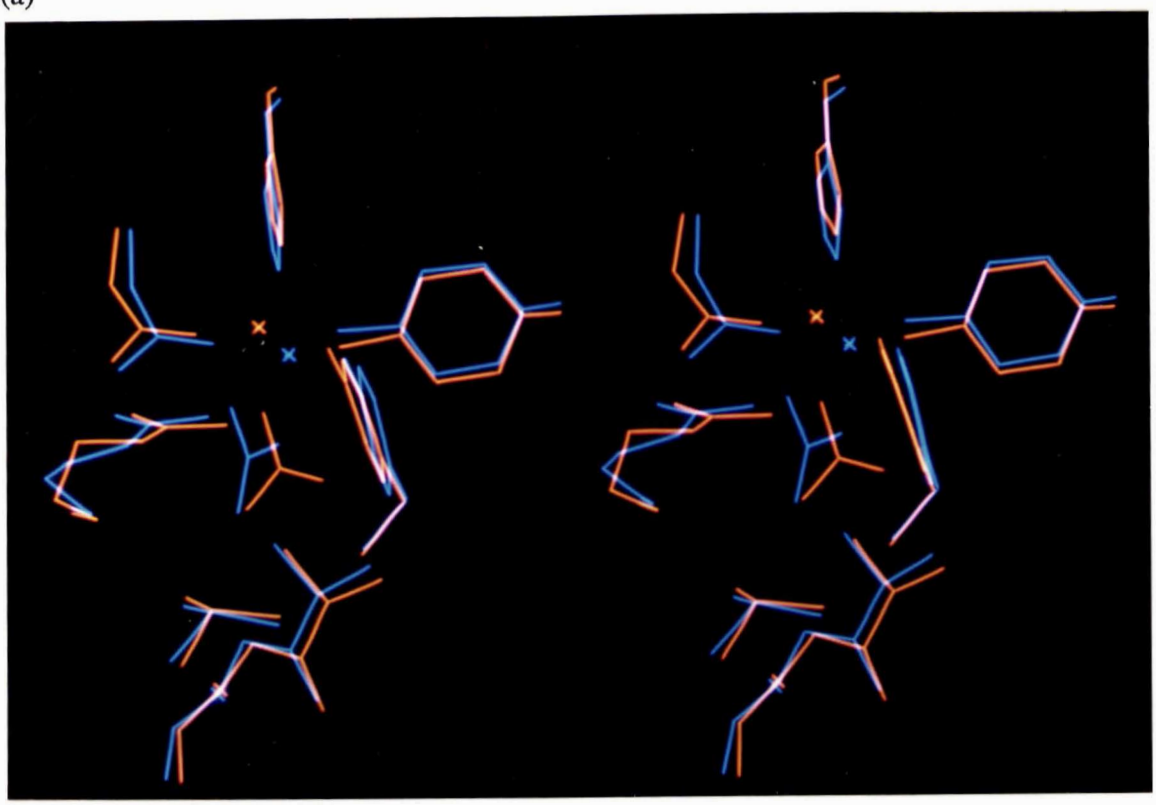
| | | $Fe_2Lf$ | | | |
|---|---|---|---|---|---|
| | | N1 | N2 | C1 | C2 |
| $Cu_2Lf$ | N1 | 0.29 Å | 0.6° | 0.3° | 0.1° |
| | N2 | — | 0.23 Å | 0.2° | 0.5° |
| | C1 | — | — | 0.22 Å | 0.2° |
| | C2 | — | — | — | 0.23 Å |

[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.

The N2, C1 and C2 domains have rms deviations of about 0.23 Å, which is essentially equal to the maximum estimated error in atomic positions in $Cu_2Lf$ (Section 3.2.1) and indicates that the 3-dimensional structures of these three domains are virtually identical. There is no reason to suppose that the structures of the N1 domains in the two structures are not identical to each other, although, as noted earlier, the first part of this domain does appear to be more flexible than the rest.

*The metal and anion binding sites:* In order to quantify the movements in the metal binding site upon substitution of $Fe^{3+}$ with $Cu^{2+}$, the separate N- and C-terminal lobes of $Cu_2Lf$ were superimposed on to the corresponding lobes of $Fe_2Lf$, using only secondary structure elements in the N2 and C2 domains (helices and β-strands, with the worst agreements removed and excluding helix 6 in the C-lobe and helices 8a in both lobes) in the least squares calculation. Where residues associated with metal or anion binding occurred in a portion of secondary structure, these too were omitted from the calculation. The results of these superpositions are shown in Plates 3.13a & b and later in Plates 3.19a & b.

(a)



(b)

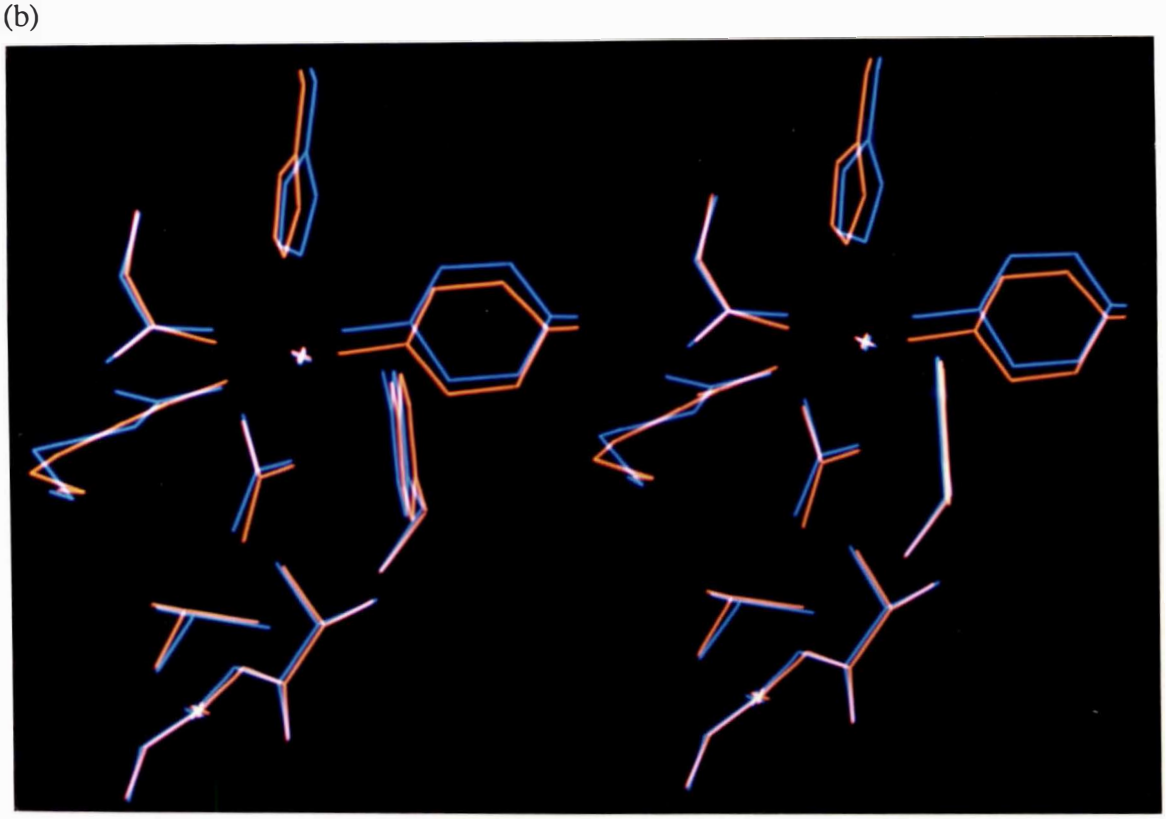


Plate 3.13: Stereoviews of the superposition of (a) the N-lobe and (b) the C-lobe metal and anion binding sites of $Cu_2Lf$ (red) on to the corresponding sites in $Fe_2Lf$ (blue). The metal ions and carbonate ions are in the centre of the photographs, with (clockwise from the top) His253 (597), Tyr92 (435), Tyr192 (528) (edge on), the N-terminus of helix 5, Thr117 (461), Arg121 (465) and Asp60 (395)..

Clearly, the copper atom in the N-lobe site has moved considerably more relative to the iron atom in $Fe_2Lf$ than has its counterpart in the C-lobe. Table 3.22A gives the rms displacements in the copper and anion positions between $Cu_2Lf$ and $Fe_2Lf$, based upon the superposition described above, while the rms differences in the ligand donor atoms and average sidechain positions (the x, y and z coordinates of all the sidechain atoms, excluding $C_\beta$, are averaged) are given in Table 3.22B. In the N-lobe site, the copper atom is about 1.1 Å from the equivalent iron position, having moved in a direction away from Tyr92 towards the interdomain cleft and His253. Movements of the sidechains of Asp60, and Tyr192 (the average displacement for the sidechain atoms ranges from 0.2 - 0.7 Å) allow the donor atoms to remain within coordinating distance of the copper. In general, the sidechains of these residues appear to move in accordance with the copper atom, i.e. in a direction away from Tyr92 and towards His253. This movement does not affect their hydrogen bonding interactions.

The sidechain of Tyr92 appears to be in an almost identical position to its counterpart in $Fe_2Lf$, the difference in the position of the $O_\eta$ atoms being only about 0.2 Å compared with values of 0.4 - 0.8 Å for the other ligands (Table 3.22B). The relative immobility of the sidechain of Tyr92 and the movement of the copper atom leads to the observed elongation of the Cu - O bond (2.0 Å in $Fe_2Lf$ compared to 2.8 Å in $Cu_2Lf$).

Table 3.22A: The differences in the positions of the metal ions and the anion atoms between $Cu_2Lf$ and $Fe_2Lf$, and between the two lobes of $Cu_2Lf$.

| | Between lobes of $Cu_2Lf$ [a] | N-lobe [b] | C-lobe [b] |
|---|---|---|---|
| metal ion | 1.2 Å | 1.1 Å | 0.2 Å |
| C | 0.63 Å | 0.8 Å | 0.1 Å |
| O1 | 0.90 Å | 1.2 Å | 0.1 Å |
| O2 | 0.98 Å | 1.1 Å | 0.2 Å |
| O3 | 0.49 Å | 0.2 Å | 0.3 Å |

[a] Superimposed using secondary structure elements only. [b] The N- and C-lobes of $Cu_2Lf$ were superimposed onto the corresponding lobes of $Fe_2Lf$ using the secondary structure.

The movement of the carbonate anion in the N-lobe site can best be described as a rotation about one of the oxygen atoms ($O_3$). Whereas in the N-lobe of $Fe_2Lf$ with its bidentate carbonate it is the $O_1$ atom of the anion which is in a trans position relative to His253, the movement of the copper atom and the detachment of the anion $O_1$ atom in the N-lobe of $Cu_2Lf$ puts the $O_2$ atom into the *trans* position with respect to the histidine (see Table 3.17). It is a combination of the movement of the anion coupled with that of the copper that allows the anion to become monodentate in $Cu_2Lf$ compared with bidentate in $Fe_2Lf$.

Table 3.22B: The rms displacement in donor atom and sidechain atom positions between $Cu_2Lf$ and $Fe_2Lf$ and between the two lobes of $Cu_2Lf$.

| | | N-lobe | | C-lobe | |
|---|---|---|---|---|---|
| | Between lobes of $Cu_2Lf$ | donor atoms [b] | sidechain atoms [c] | donor atoms | sidechain atoms |
| Asp60 ($O_{\delta 1}$) | 0.6 | 0.5 Å | 0.4 Å | 0.3 Å | 0.3 Å |
| Tyr92 ($O_\eta$) | 0.4 | 0.2 Å | 0.3 Å | 0.5 Å | 0.4 Å |
| Tyr192 ($O_\eta$) | 1.0 | 0.5 Å | 0.5 Å | 0.5 Å | 0.5 Å |
| His253($N_{\varepsilon 1}$) | 1.1 | 0.8 Å | 0.7 Å | 0.5 Å | 0.6 Å |
| Thr117($O_{\gamma 1}$) | 0.5 | 0.3 Å | 0.4 Å | 0.2 Å | 0.3 Å |
| Arg121 ($N_\varepsilon$) | 0.5 | 0.5 Å | 0.5 Å | 0.3 Å | 0.3 Å |
| ($C_\zeta$) | 0.3 | 0.5 Å | | 0.1 Å | |
| ($N_{\eta 1}$) | 0.4 | 0.5 Å | | 0.9 Å | |
| ($N_{\eta 2}$) | 0.8 | 0.4 Å | | 0.5 Å | |
| Ala123 (N) | 0.2 | 0.2 Å | — | 0.3 Å | — |
| Gly124 (N) | 0.2 | 0.2 Å | — | 0.2 Å | — |

[a] Equivalent residues in the C-lobe are Asp395, Tyr435, Tyr528, His597, Thr461, Arg465, Ala467 and Gly468. [b] The donor atoms are as listed in the first column. [c] All atoms excluding $C_\beta$ were used to calculate the average sidechain value.

In the C-terminal site, the copper atom has shifted to a much lesser extent ($\approx$ 0.3 Å) and the four ligand sidechains have moved slightly to keep in contact with the metal ion (between 0.3 and 0.5 Å). None of the four metal binding ligands in this site appear to be as fixed in position as is the case for Tyr92 in the N-lobe, although the sidechain of Asp395 is much less able to move due to the hydrogen bonding interactions with the C2 domain. The carbonate anion in this site occupies essentially the same position in both $Cu_2Lf$ and $Fe_2Lf$.

When the N-lobe of $Cu_2Lf$ is superimposed on the C-lobe, in terms of the helices and β-sheets in the N2 and C2 domains, similar positional differences are observed (this superposition is shown in Plate 3.22 in Section 3.4). Because of the way in which the superposition was done, the N-terminus of helix 5 and the sidechains of Arg121 and Arg465 are in almost identical positions (see Table 3.22B). The rotation of the carbonate ion is evident, as is the movement of the copper atom. Of the four metal binding ligands, the ones which move the most are Tyr192 and His253, where the $O_\eta$ and $N_{\varepsilon 2}$ positions differ by $\approx$1.0 Å respectively. Once again, the direction of movement of the copper is towards His253 which in turn makes the histidine sidechain move away.

## 3.3 Results; B: Dicupric-carbonate-oxalate lactoferrin

### 3.3.1 *The model:*

The final model for $Cu_2$oxLf contains 5323 protein atoms, 237 solvent molecules, 49 carbohydrate atoms (2 residues, an N-acetylglucosamine and a fucose, both in the N-lobe), two $Cu^{2+}$ ions and two anions, one each of carbonate and oxalate. The stereochemistry of the model is summarised in Table 3.10 and is very close to ideal, the rms deviations in bond distances being 0.018 Å, with an angle (1-3) distance of 0.064 Å. The present R-factor of 0.212 is for data with $I \geq 1.0\sigma_I$ between 7.5 and 2.1 Å resolution (48224 in total).

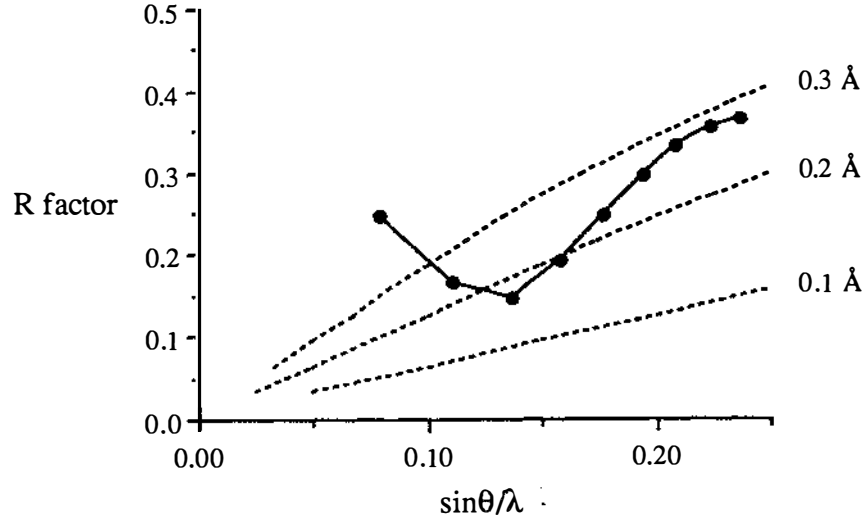



Figure 3.23: A Luzzati-type plot of resolution ($\sin\theta/\lambda$) against the crystallographic R factor (Luzzati, 1952) for $Cu_2$oxLf. The theoretical variation of R with resolution for rms errors in atomic positions of 0.1, 0.2 and 0.3 Å, are shown for non-centric data.

A Luzzati plot (Figure 3.23) gives an estimated average maximum coordinate error in the $Cu_2$oxLf model between 0.20 and 0.25 Å. As with $Cu_2$Lf, the error in the well-ordered regions would be somewhat less than this, while the external loops which connect the β-strands and helices will clearly have a much higher error associated with the positions of their atoms. Although it is difficult to judge the actual error in the atomic positions in the model (as noted in Section 3.1 2), the areas of the structure which are deeply buried (the metal and anion binding sites for example) would be expected to have lower errors than regions exposed to the solvent.

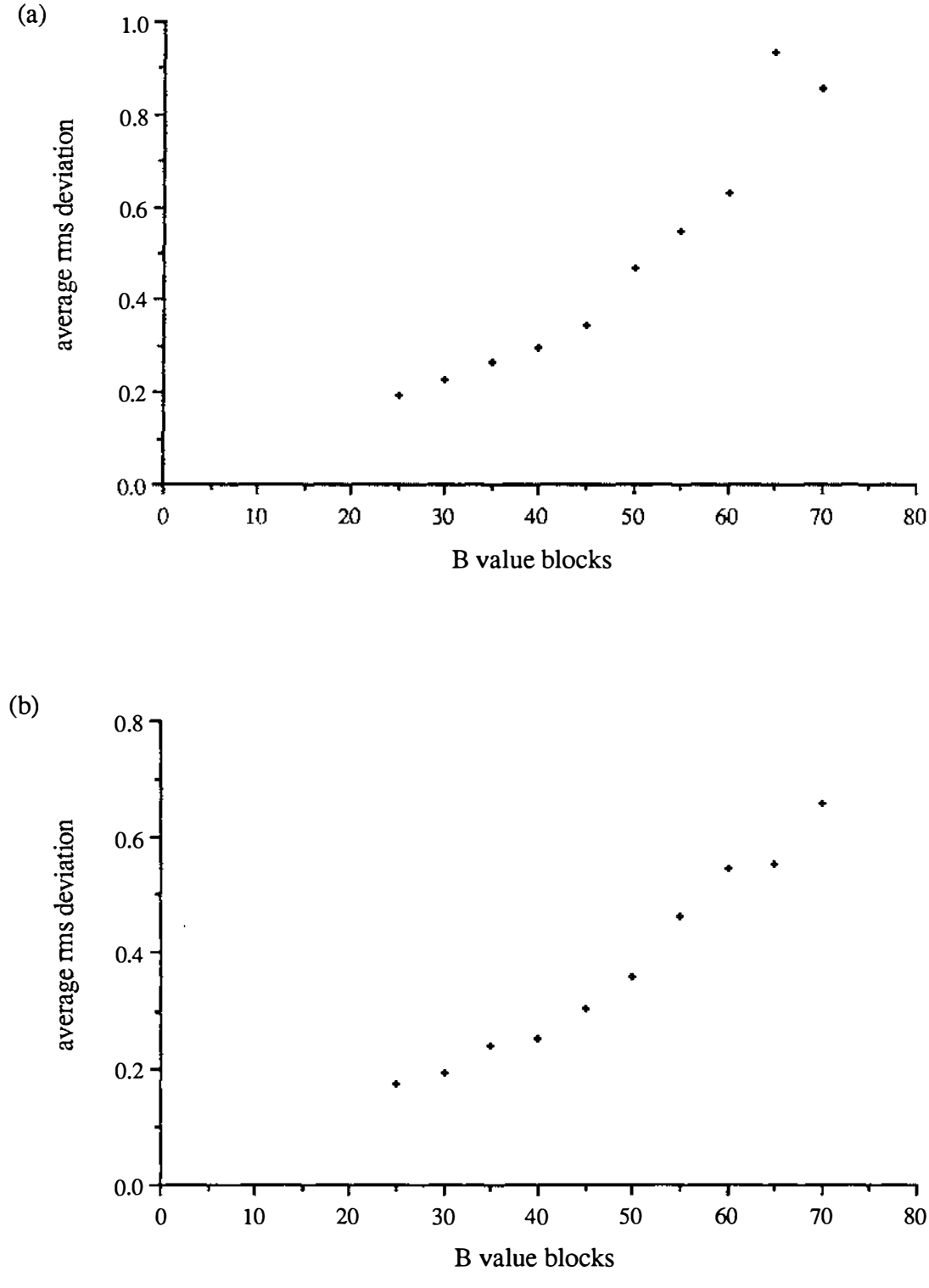



Figure 3.24: The rms deviation between mainchain atoms averaged over 15 blocks of the average mainchain temperature factor for (a) $Cu_2$oxLf and $Fe_2$Lf and (b) $Cu_2$oxLf and $Cu_2$Lf.

As discussed in Section 3.2.4, correlation of the average mainchain B values with the rms deviation in mainchain atom positions between two isomorphous structures gives an indication of the error in atomic positions in different regions of the molecule. $Cu_2$oxLf can be compared with two isomorphous structures, $Fe_2$Lf and $Cu_2$Lf. The results of superpositions of $Cu_2$oxLf on to each of these structures are summarised in Figures 3.24a &b and, as expected, they show that the more thermally mobile regions (with high B values) also have the highest rms deviations in mainchain atom positions. The average error in the well-defined regions of the structure (B values 20 - 40 $Å^2$) can be estimated at ≈ 0.15 - 0.20 Å, while in the ill-defined areas (B values > 60 $Å^2$) it may be as high as 0.5 - 0.6 Å.

3.3.2 *Polypeptide chain:*

The polypeptide chain folding and the secondary structure of $Cu_2$oxLf is essentially identical to that discussed previously for $Cu_2$Lf (section 3.2.2) and $Fe_2$Lf (Anderson *et al.*, 1987; 1989). A Ramachandran plot of the main chain conformational torsion angles for $Cu_2$oxLf is shown in Figure 3.25 and the observed distribution of $\phi/\varphi$ angles is essentially similar to $Cu_2$Lf (Figure 3.10). Once again , the N-terminus appears to be ill-defined, as do some of the external loops near residues 87, 283 and 637.

It is interesting to note that Ser191 again has unusual $\phi/\varphi$ angles as observed in $Cu_2$Lf (61° and -176° for $Cu_2$oxLf compared to 65° and -175° for $Cu_2$Lf). Additional comparisons with $Fe_2$Lf and apoLf show that these two structures also have $\phi/\varphi$ angles for Ser191 in this region of conformational space (this is discussed in Section 3.2.2).

Residues 299 and 642 once again have conformations typical of γ-turns involving 1-3 hydrogen bonds. The $\phi/\varphi$ angles (76 & -47° for Leu299 and 69 & -37° for Leu642) are similar to those observed in the two γ-turns in $Cu_2$Lf. As discussed in Section 3.2.3, these γ-turns, in identical positions in the two lobes, appear to be a characteristic of virtually all the transferrin structures determined to date. The inverse γ-turn at 445 is also present in $Cu_2$oxLf.

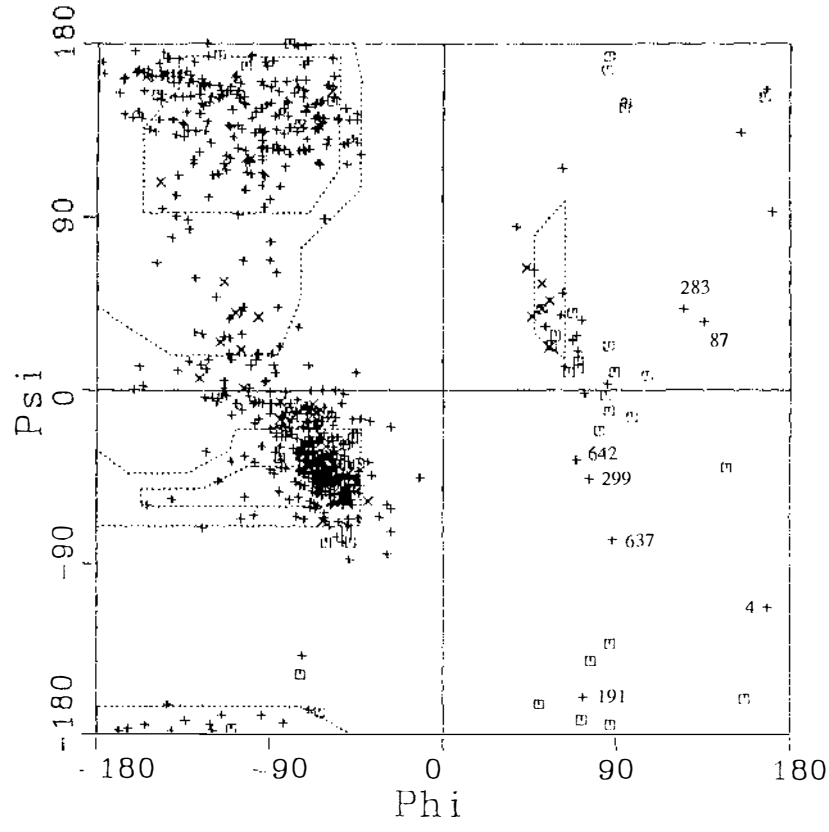



Figure 3.25: Ramachandran plot of the mainchain torsion angles for $Cu_2$oxLf.

### 3.3.3 *Temperature factors:*

The variation of mainchain temperature factors for $Cu_2$oxLf (Figures 3.26a & b) is similar to that seen for $Cu_2$Lf (see Figures 3.13a & b), in that the low values (< 40 Å$^2$) correspond in position to the secondary structure elements, while the areas with high average B values (> 60 Å$^2$) relate to the more flexible loops on the surface of the molecule which serve to join the β-strands and helices. Table 3.23 gives the average values for both mainchain and sidechain temperature factors for the whole $Cu_2$oxLf molecule and for the individual domains.

Table 3.23: Temperature factors for $Cu_2$oxLf

| | overall [a] | N-lobe | C-lobe | N1 | N2 | C1 | C2 |
|---|---|---|---|---|---|---|---|
| average mainchain B | 34.9 | 36.0 | 33.4 | 37.5 | 34.3 | 34.5 | 33.0 |
| average sidechain B | 36.7 | 37.7 | 34.9 | 39.5 | 35.8 | 35.7 | 34.1 |

[a] Averages are for protein atoms only. Residues 1 - 240 make up the N-lobe and residues 341 - 691 the C-lobe. The residues making up the four domains are the same as those given for $Cu_2$Lf in Section 3.2.1.

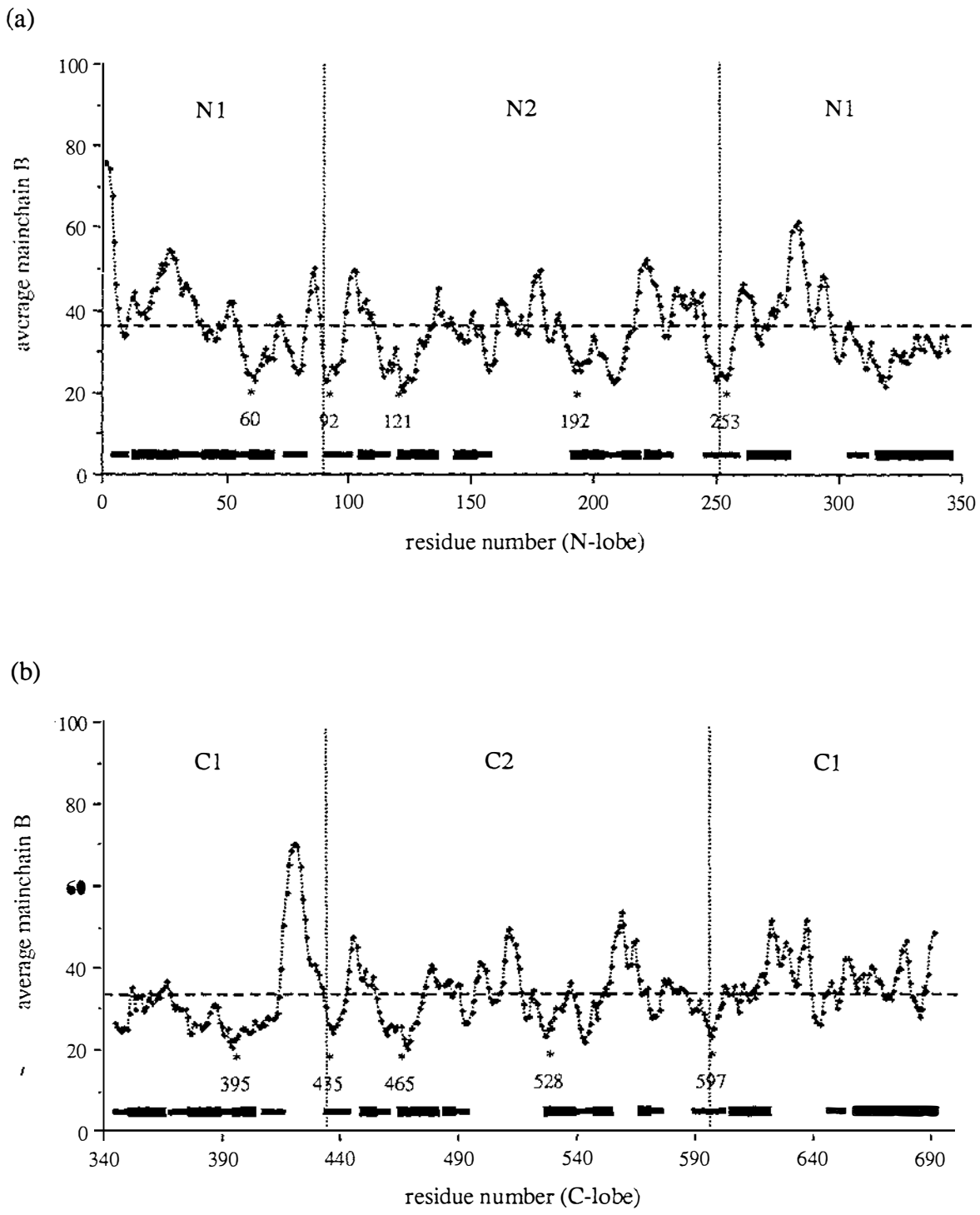



Figure 3.26: Plot of average mainchain B values against residue number for (a) the N-lobe and (b) the C-lobe of $Cu_2$oxLf. The positions of the helices and β-strands are indicated by ▬ and ▬ respectively. The *'s show the location of the four metal binding ligands in both lobes, along with the anion-binding Arg121 (N-lobe) and Arg465 (C-lobe).

3.3.4 *Solvent structure:*

As described earlier (Section 3.2.5), when the structure of $Cu_2Lf$ was superimposed on $Fe_2Lf$ by least squares fitting of the positions of the $C_\alpha$ atoms, 123 out of 301 $Cu_2Lf$ water molecules were found to be in conserved positions (within 1.0 Å) and to interact with the same protein atoms. It was assumed that this "core" set of solvent molecules (see Plate 3.6b) should also be common to $Cu_2oxLf$ and they were thus added to the list of atomic coordinates. The location of each of the 123 "core" water molecules was subsequently checked on the graphics system against $2F_o$-$F_c$ and $F_o$-$F_c$ electron density maps for which the solvent molecules had been omitted from the structure factor calculation, and molecules which did not appear to be in stereochemically sound positions or in good electron density were removed. A total of 101 of the "core" solvent molecules were retained in the structure. Plate 3.6b shows a selection of some of these "core" water molecules in their $2F_o$-$F_c$ density.

The remainder of the water molecules were located as described for $Cu_2Lf$, from the analysis of $2F_o$-$F_c$ and $F_o$-$F_c$ density maps over several phases of refinement and manual model rebuilding. This involved removing water molecules which had large increases in their B values upon refinement, which appeared in negative $F_o$-$F_c$ electron density or which were involved in unfavourable contacts with parts of the protein structure, and calculating omit maps ($F_o$-$F_c$) after further least squares refinement. Positions in which positive $F_o$-$F_c$ density reappeared were refitted with solvent molecules. In addition, several well-defined solvent molecules were deliberately omitted and their positions were re-evaluated in a similar way.

Of the water molecules present in the two inter-domain clefts, about 75% were common to both "core" solvent models. The pattern of hydrogen bonding in the N-lobe is virtually identical to that observed in $Cu_2Lf$, while the solvent structure in the vicinity of the C-lobe site has a number of changes with respect to $Cu_2Lf$ (see later in this chapter for a complete description).

3.3.5 *Glycan structure:*

The first 2 residues in the N-lobe glycan chain, the asparagine-linked N-acetylglucosamine and the α-linked fucose residue, are included in the $Cu_2$oxLf model and are shown in their $2F_o$-$F_c$ electron density in Plate 3.14. The $2F_o$-$F_c$ density is rather poorly defined for the FUC residue, although it is somewhat better for the NAG. Consequently, the individual B values for the atoms in the glycan chain are high (the average for the two residues is ≈78 Å$^2$).

These sugar molecules were located independently of the glycan chain in either $Cu_2$Lf (Section 3.2.6) or $Fe_2$Lf, in $F_o$-$F_c$ electron density maps. The NAG moiety was initially built after phase 5 of the refinement and the FUC residue was added in the next phase. No interpretable $F_o$-$F_c$ density was observed for the second NAG moiety. The carbohydrate residues occupy similar positions to those in the N-lobe of $Cu_2$Lf and $Fe_2$Lf, and interact with the protein molecule in an almost identical way. No electron density has been observed to date for the carbohydrate residues in the C-lobe site, although it is envisaged that the use of the high resolution photographic data in the structure refinement will allow these residues to be located.

3.3.6 *The metal and anion sites:*

A difference electron density map was calculated with coefficients $\Delta F = |F_{Cu_2oxLf}| - |F_{Cu_2Lf}|$, using phases from the refined coordinates of $Cu_2$Lf. This map indicated that the major changes upon substitution of carbonate by oxalate occurred in the C-terminal metal binding site (Plate 3.15). The map was essentially flat with the exception of a number of positive and negative peaks near the C-lobe copper at least 5 - 7 times greater than the rms deviation in the density. An area of positive density was found between the positions of the carbonate and Arg465, and two sets of negative and positive peaks were also associated with Arg465, and with Tyr398, which is adjacent to Arg465. Assuming the extra density near the anion corresponded to the two extra atoms of the oxalate ion, it was then possible to correlate the two sets of positive and negative peaks with movements of the Arg465 and Tyr398 side chains in a direction away from the metal binding site.

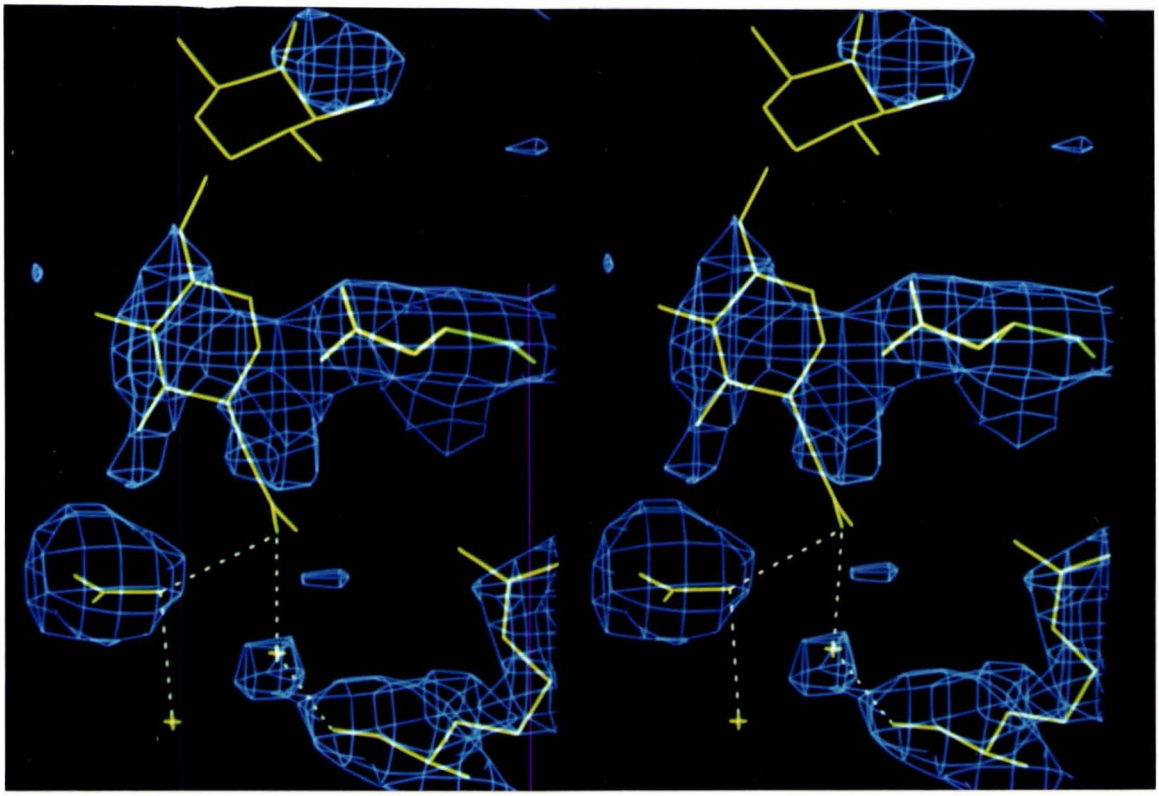


Plate 3.14: Stereo view of the $2F_O$-$F_C$ electron density associated with the carbohydrate residues in the N-lobe of $Cu_2$oxLf (Nag at the centre and FUC at the top). The interactions with Glu110 and a solvent molecule are almost identical to those found in $Cu_2$Lf (see Plate 3.7a)

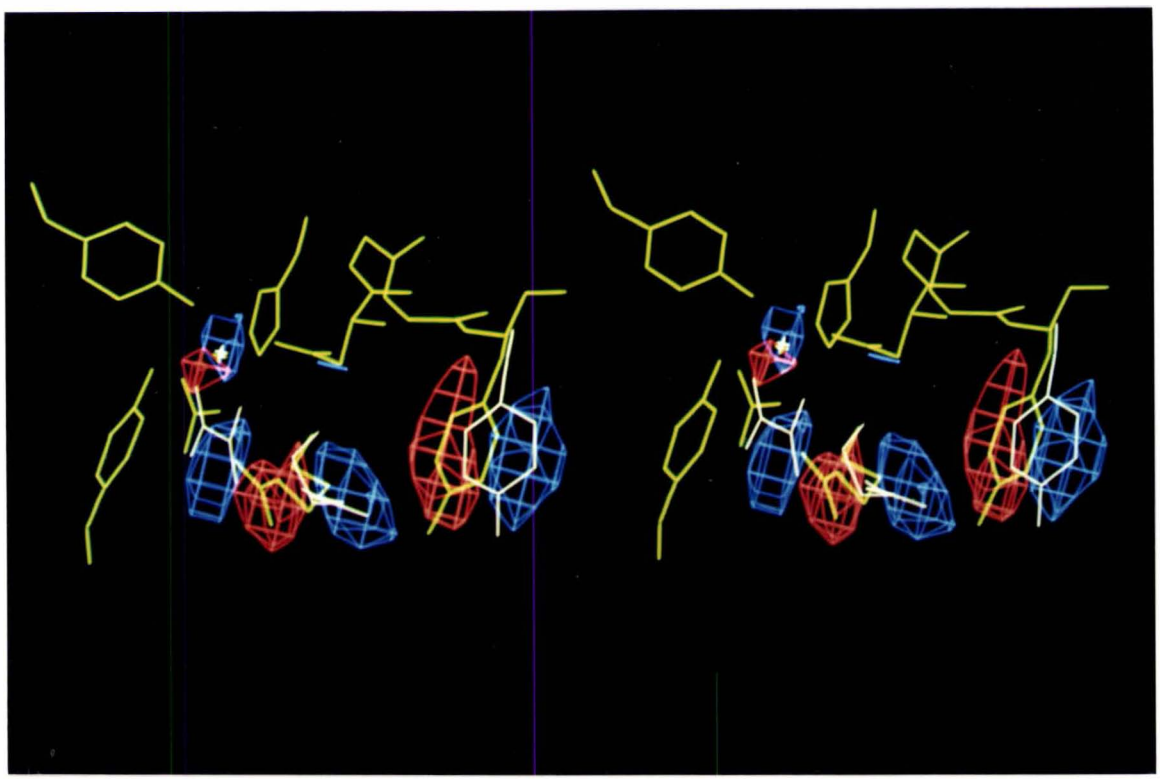


Plate 3.15: $Cu_2$oxLf - $Cu_2$Lf difference electron density map ($\Delta F = F_{(Cu_2oxLf)} - F_{(Cu_2Lf)}$) in the vicinity of the C-lobe metal binding sites. Regions of negative density are shown in red and positive density in blue. Atoms from the $Cu_2$Lf model are shown in green, with the copper atom, oxalate ion, Arg465 and Tyr398 of $Cu_2$oxLf superimposed in white. The extent of the movement of the arginine can be clearly seen.

At 5 times the rms deviation of the map, no features were discernible in the N-lobe binding site. Decreasing the contour level to 2 times the rms deviation resulted in a large number of positive and negative peaks which could not be clearly attributed to any changes in the positions of the metal ion, the anion or the metal binding ligands.

$2F_o$-$F_c$ and $F_o$-$F_c$ electron density maps calculated after the first phase of refinement showed the location of the ligands and the copper atoms and these were built into the appropriate positions (Figures 3.27a & b show the $2F_o$-$F_c$ map in the vicinity of the N- and C-lobe metal binding sites). Further refinement in the absence of any anions left a large piece of residual density at each site. A carbonate ion was added in the N-lobe, and it was reassuring to note that the shape and orientation of the anion density again suggested monodentate coordination of the anion (Figure 3.27a), as in $Cu_2$Lf. Once again it was involved in at least four hydrogen bonds with nearby parts of the protein molecule (described later).

In the C-terminal site, the anion density was bilobal in nature (Figure 3.27b) and too large to be accounted for by a bidentate carbonate ion. However, an oxalate ion, bound in a 1,2-bidentate fashion was found to fit almost perfectly into the available density, making at least five favourable hydrogen bonds. The oxalate anion was initially fitted with symmetrical coordination to the copper atom and although loose restraints were placed on all the copper - ligand bond distances, the copper-oxygen bonds to the oxalate showed no tendency to become unequal during the course of the refinement. The Cu - O bond distances are both ≈2.0 Å. It is this symmetry, along with the significantly larger chelate angle (≈85° for oxalate compared with ≈64° for carbonate), which accounts for the more regular octahedral geometry observed in the C-lobe site of this model compared with that observed in the C-lobe of $Cu_2$Lf (see Plate 3.9 and Tables 3.16 & 3.17).

The positions of the other ligands in the two binding sites appeared to be similar to those in $Cu_2$Lf; the majority of the copper - ligand distances were about 2.0 Å, although in the N-lobe site the Cu- $O_\eta$ (Tyr92) distance was once again lengthened to about 2.5 Å.

Restraints on the Cu - ligand and Cu-anion distances were treated in a similar way as for $Cu_2$Lf. All bond distances were initially restrained to target values near 2.0 Å, with the

exception of the Cu - Oη (92) bond which was not restrained at all. Small consistent shifts in the bond distances were accounted for with small alterations in the target values. The restraints were removed from all the bond distances near the end of the refinement and very little change was observed.

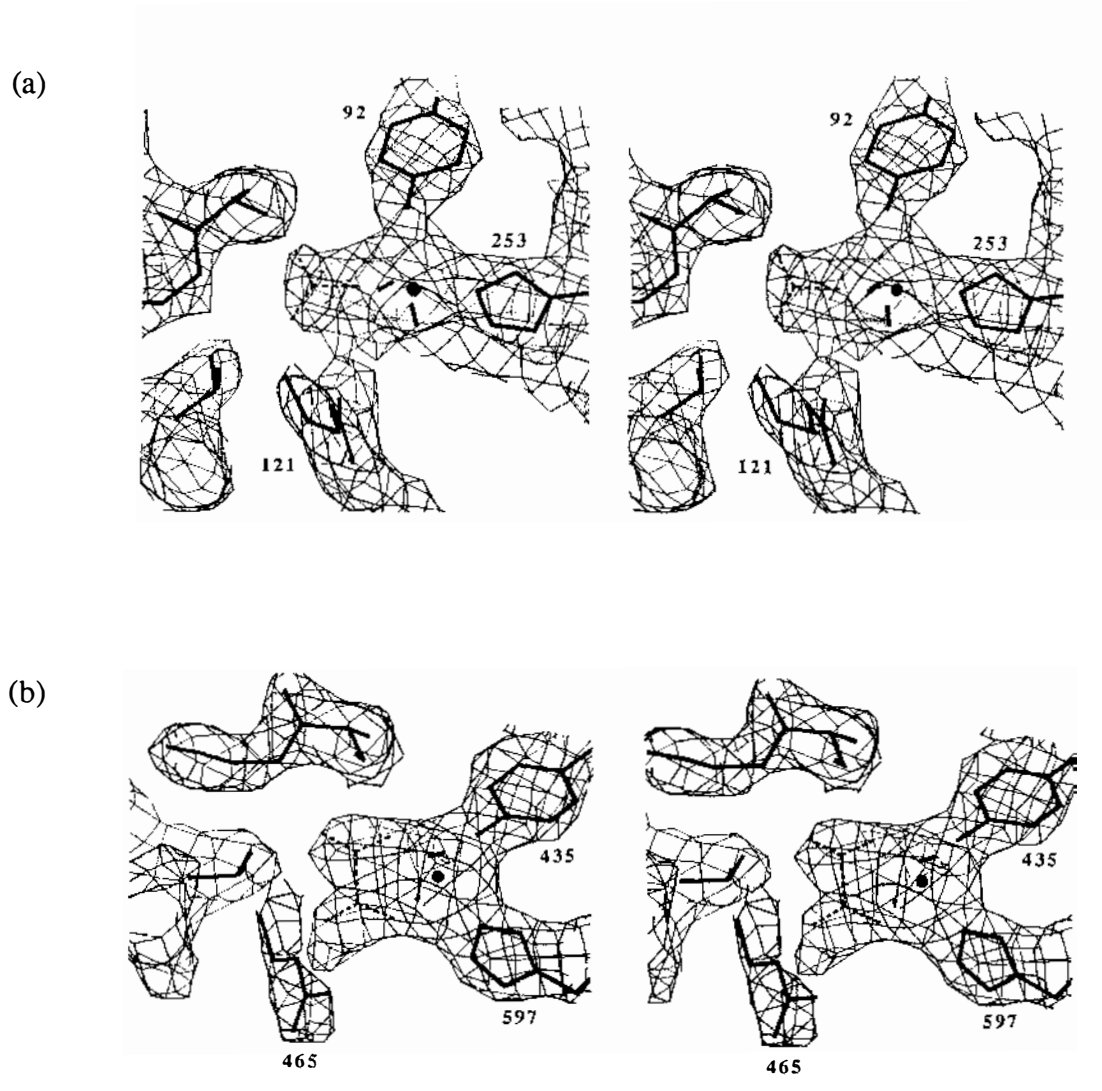



Figure 3.27: Stereo views of $2F_o$-$F_c$ electron density in (a) the N-lobe and (b) the C-lobe binding sites of $Cu_2$oxLf, showing the density associated with two of the protein ligands (Tyr 92 (435) and His 253 (597), the copper atom and the carbonate anion. This map was calculated at the time the anions (dashed lines) were first built into the model (R = 0.177 for data between 5.0 and 2.6 Å).

In the final model for $Cu_2$oxLf, the copper geometries in the two sites are again different, the N-site being approximately square pyramidal with a long (2.6 Å) apical bond (Tyr 92), as

for the N-site in dicupric lactoferrin described earlier, while the C-site is an almost regular octahedron with all six bonds of 2.0 - 2.2 Å. The error in the Cu-ligand and Cu-anion bond distances is estimated at ≈ 0.2 Å, based on the estimated error in the atomic coordinates in this area of the model (see Section 3.3.1).. A list of relevant bond lengths and angles in the two copper sites are given in Tables 3.24 and 3.25 respectively, with atom labelling as in Figures 3.28a & b. Stereo views of the two sites are shown in Plates 3.16a & b and 3.17a & b.

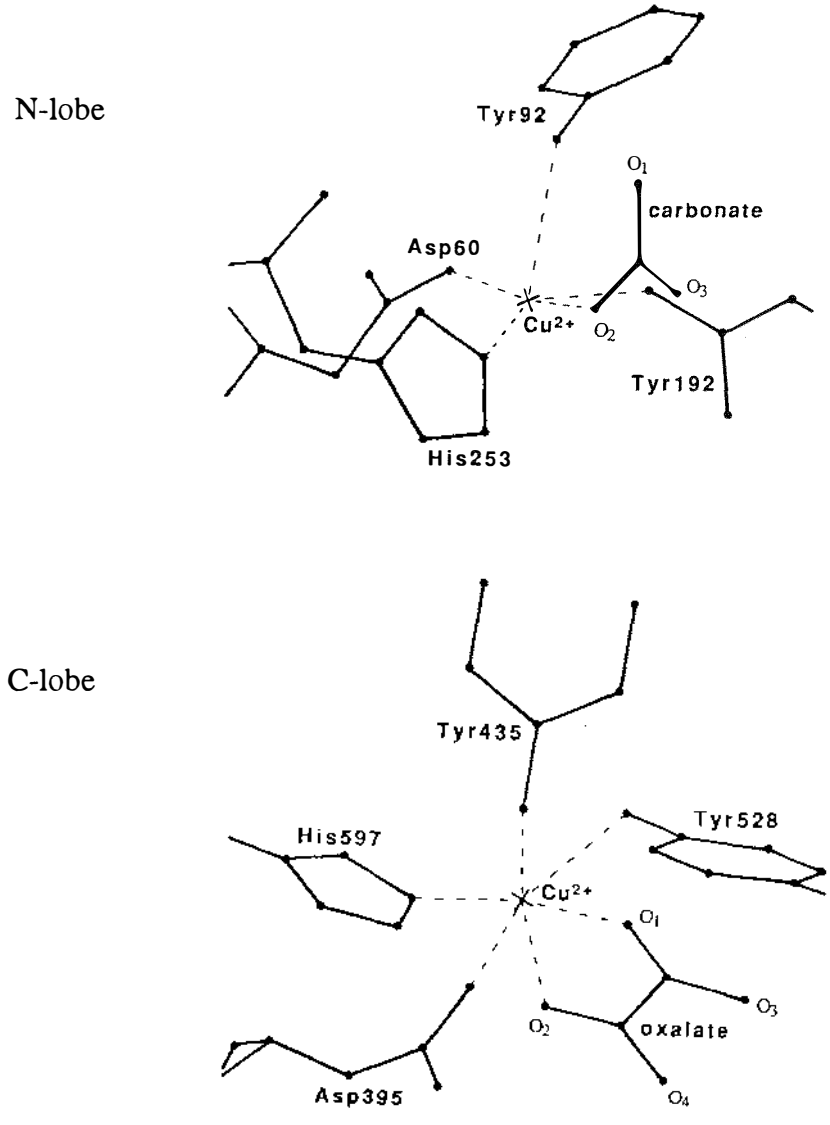



Figure 3.28: The coordination geometry and atom labelling scheme in the N- and C-lobe binding sites of $Cu_2$oxLf.

Table 3.24: Copper coordination sphere for $Cu_2$oxLf - bond lengths

| Bond [a] | N-lobe site | C-lobe site |
|---|---|---|
| Cu - O60(395) | 2.0 Å | 2.0 Å |
| Cu - O92(435) | 2.6 Å | 2.0 Å |
| Cu - O192(528) | 2.0 Å | 2.3 Å |
| Cu - N253(597) | 2.0 Å | 2.0 Å |
| Cu - O1(anion) | — | 2.0 Å |
| Cu - O2(anion) | 1.8 Å | 2.0 Å |

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

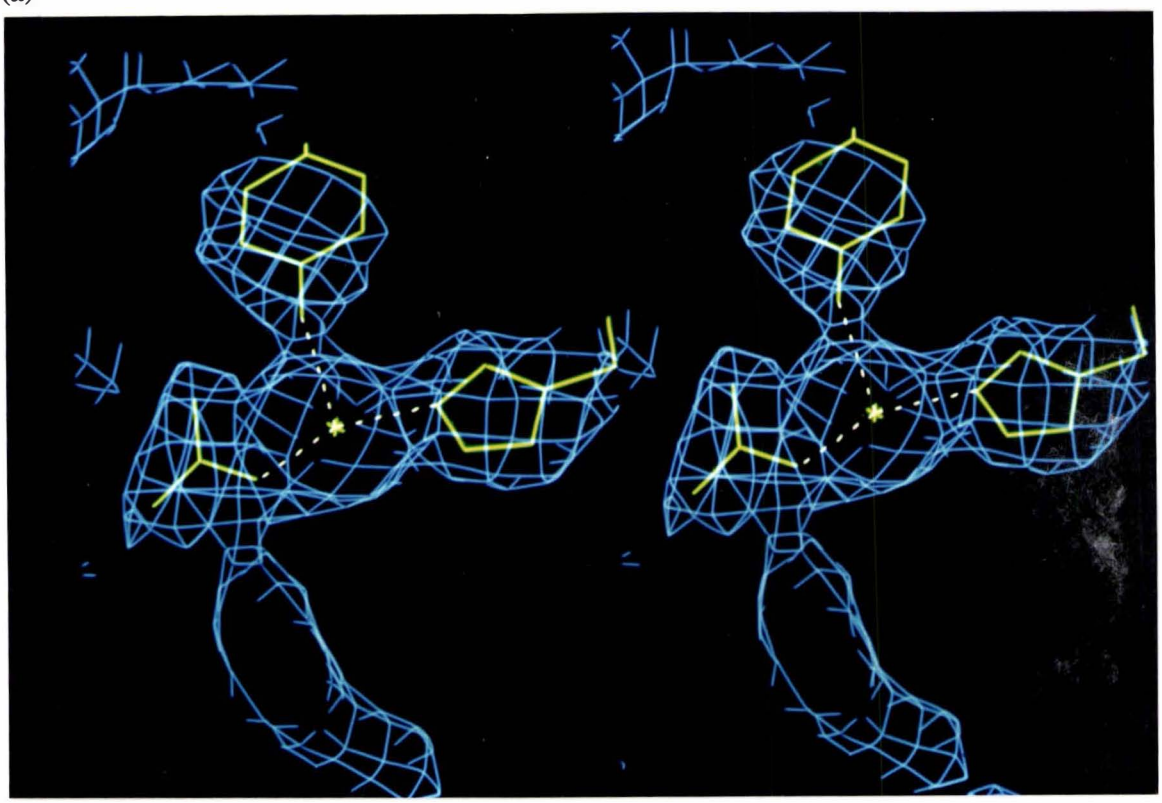
Table 3.25: Copper coordination sphere for $Cu_2$oxLf - bond angles

| Angle [a] | N-lobe site | C-lobe site |
|---|---|---|
| O60(395)-Cu-092(435) | 88° | 88° |
| O60(395)-Cu-O192(528) | 158° | 167° |
| O60(395)-Cu-N253(597) | 101° | 90° |
| O60(395)-Cu-O1(anion) | — | 87° |
| O60(395)-Cu-O2(anion) | 77° | 84° |
| O92(435)-Cu-O192(528) | 97° | 100° |
| O92(435)-Cu-N253(597) | 99° | 96° |
| O92(435)-Cu-O1(anion) | — | 92° |
| O92(435)-Cu-O2(anion) | 80° | 171° |
| O192(528)-Cu-N253(597) | 100° | 99° |
| O192(528)-Cu-O1(anion) | — | 83° |
| O192(528)-Cu-O2(anion) | 82° | 87° |
| N253(597)-Cu-O1(anion) | — | 171° |
| N253(597)-Cu-O2(anion) | 179° | 89° |
| O1(anion)-Cu-O2(anion) | — | 83° |

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

The hydrogen bonding patterns around the N-terminal (bi)carbonate (Table 3.26) are similar to those described for $Cu_2$Lf. Six hydrogen bonds are formed (Plate 3.18a), three with Arg121 (two with $N_\varepsilon$ and one with $N_{\eta 2}$), two with mainchain nitrogen atoms on the

(a)



(b)

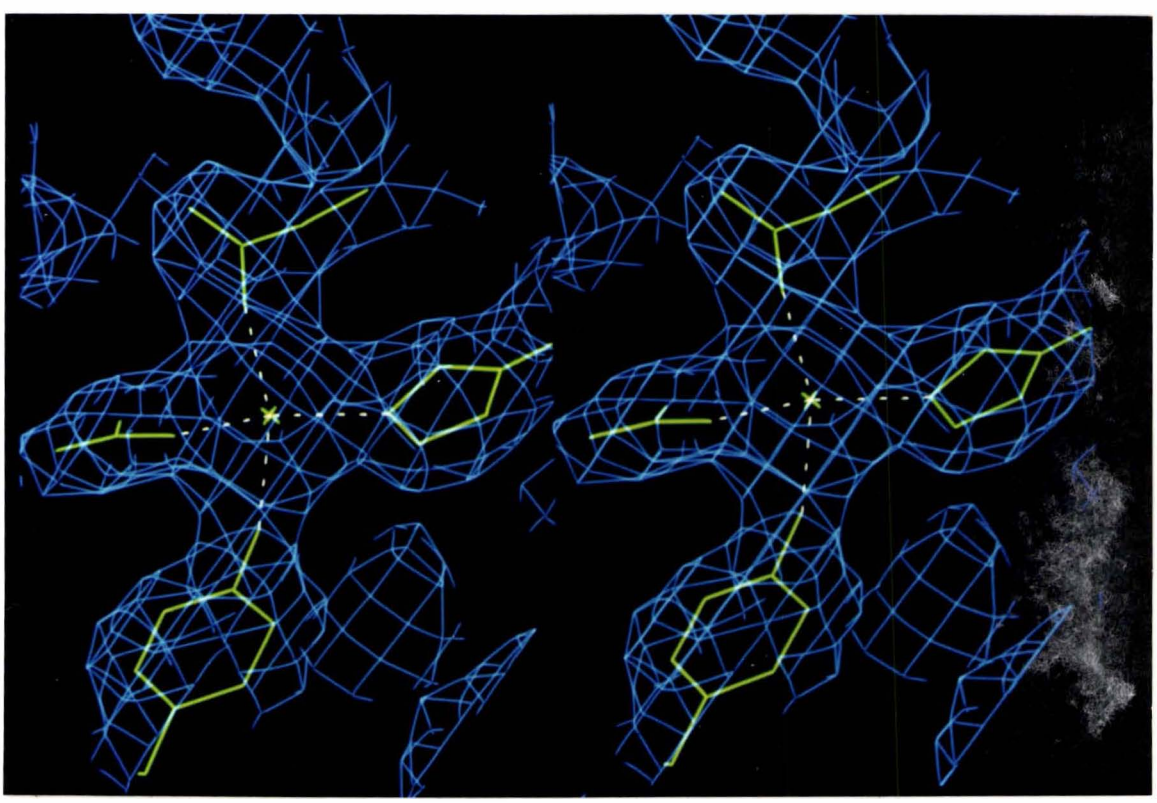
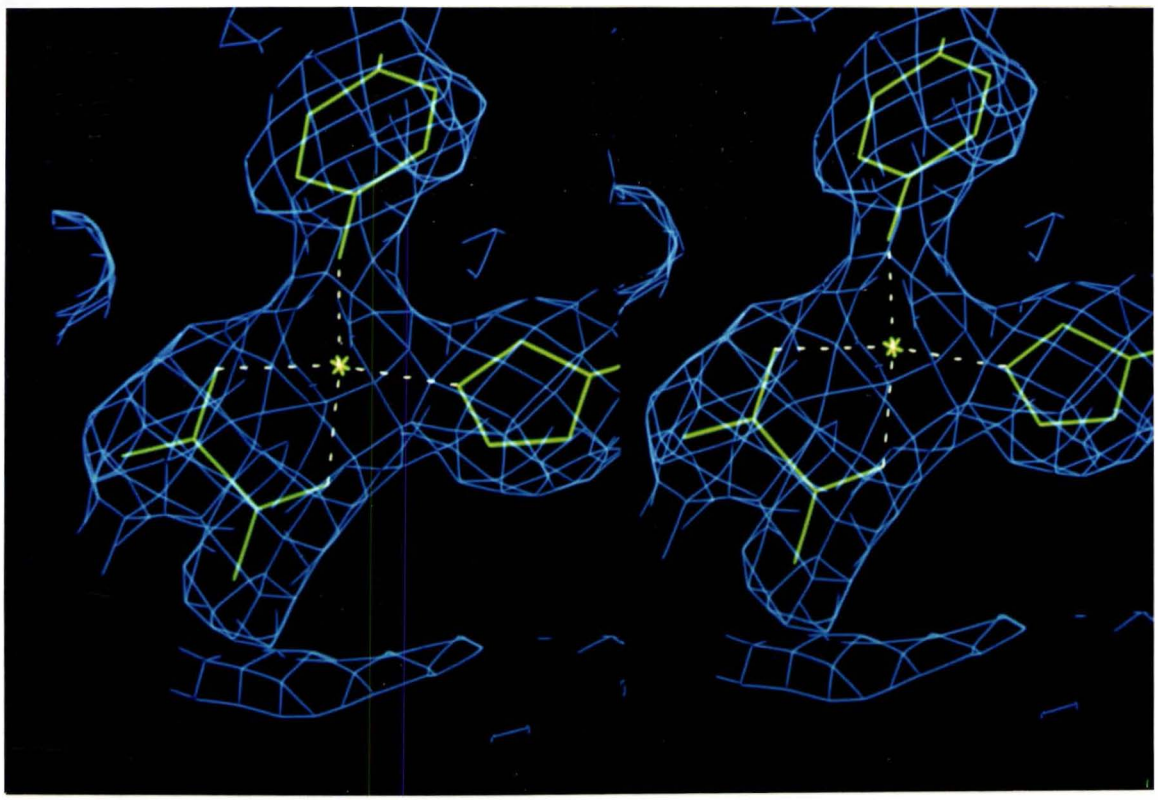


Plate 3.16: Stereo views of the $2F_O$-$F_C$ electron density in the N-lobe binding site of $Cu_2$oxLf, with the present model superimposed in green. Photograph (a) shows the copper atom, the monodentate carbonate ion, His253 and Tyr92. A rotation of 90° (b) allows the other two ligands, Tyr192 and Asp60, to be seen.

(a)



(b)

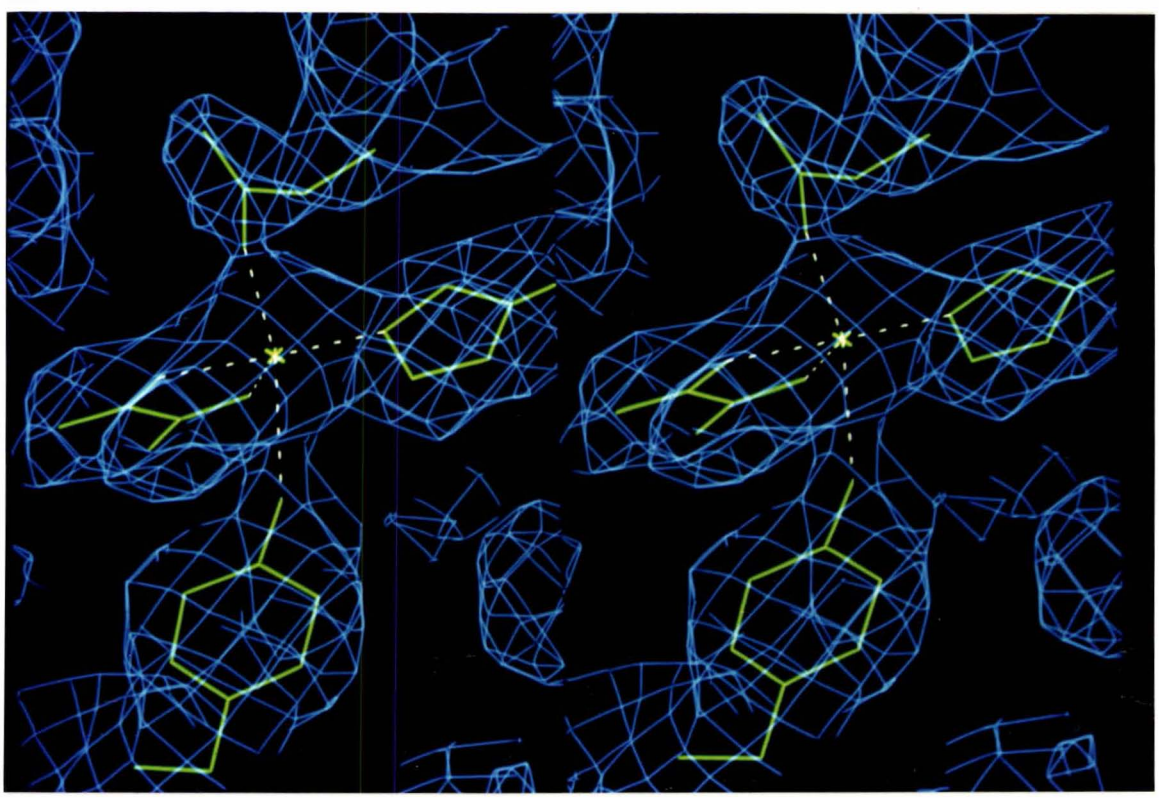


Plate 3.17: Stereo views of the $2F_O$-$F_C$ electron density in the C-lobe binding site of $Cu_2$oxLf. Photograph (a) shows the copper atom, the bidentate oxalate ion, His597 and Tyr435. A rotation of 90° (b) allows the other two ligands, Tyr528 and Asp395, to be seen.

N-terminus of helix 5 (123 and 124) and one with Thr117 ($O_{\gamma 1}$). The C-terminal oxalate anion is involved in six favourable hydrogen bonds (Plate 3.18b), two with the guanidinium group of Arg465 (one each with $N_{\varepsilon}$ and $N_{\eta 2}$), two with the mainchain peptide nitrogens of residues 467 and 468 and two with $O_{\gamma 1}$ of Thr461.

Table 3.26 Hydrogen bonding interactions between the carbonate and oxalate ions and the surrounding protein in $Cu_2$oxLf.

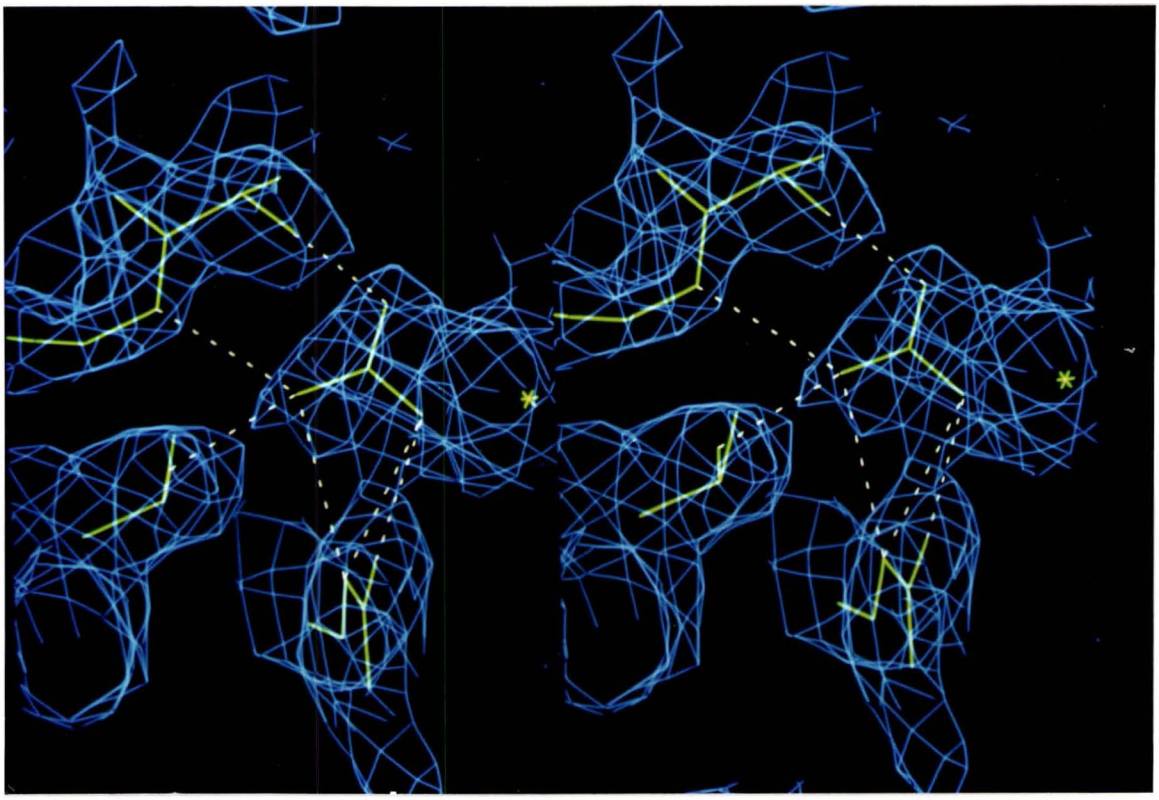
| | Hydrogen bonds | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) |
|---|---|---|---|---|---|---|
| | $O_1$ - N (123) | 1.67 | 105 | 151 | 2.59 | 18 |
| | $O_2$ - $N_{\varepsilon}$ (121) | 2.24 | 91 | 119 | 2.86 | 43 |
| | $O_2$ - $N_{\eta 2}$ (121) | 2.08 | 106 | 129 | 2.81 | 35 |
| N | $O_3$ - $O_{\gamma 1}$ (117) | — | 162 [a] | 111 [a] | 2.66 [a] | — |
| | $O_3$ - $N_{\varepsilon}$ (121) | 2.33 | 90 | 159 | 3.28 | 15 |
| | $O_3$ - N (124) | 2.31 | 122 | 127 | 3.03 | 37 |
| | $O_1$ - N (467) | 2.00 | 114 | 154 | 2.93 | 18 |
| C | $O_3$ - $O_{\gamma 1}$ (461) | — | 128 [a] | 111 [a] | 2.65 [a] | — |
| | $O_3$ - N (468) | 1.91 | 135 | 150 | 2.82 | 20 |
| | $O_4$ - $O_{\gamma 1}$ (461) | — | 113 | 92 | 3.02 | — |
| | $O_4$ - $N_{\varepsilon}$ (465) | 1.71 | 93 | 131 | 2.48 | 32 |
| | $O_4$ - $N_{\eta 2}$(465) | 1.44 | 99 | 131 | 2.30 | 24 |
| | $O_4$ - $O_{w4}$ (721) | — | 112 [b] | — | 2.70 [b] | — |

[a] These values relate to the $C_1$-$O_3$ ··· $O_{\gamma 1}$ and $C_{\beta}$-$O_{\gamma 1}$ ··· $O_3$ angles and the $O_3$ ··· $O_{\gamma 1}$ distance.
[b] These values relate to the $C_2$-$O_4$ ··· $O_w$ angle and the $O_4$ ··· $O_w$ distance

### 3.3.7 *Comparison with $Cu_2$Lf and $Fe_2$Lf:*

Because of the isomorphism of the $Cu_2$oxLf crystals with both $Fe_2$Lf and $Cu_2$Lf, the overall folding of the polypeptide chain was expected to be closely similar to both the diferric and dicupric structures. Difference maps calculated between $Cu_2$oxLf and $Cu_2$Lf, and $Cu_2$oxLf and $Fe_2$Lf suggested that the only significant differences between the three structures was in the vicinity of the metal and anion binding sites. These regions are shown in Plate 3.15 and Plates 3.19a & b respectively. The differences in the C-lobe site relate primarily to the movements of copper atom, Arg465 and Tyr398, and the incorporation of an

(a)



(b)

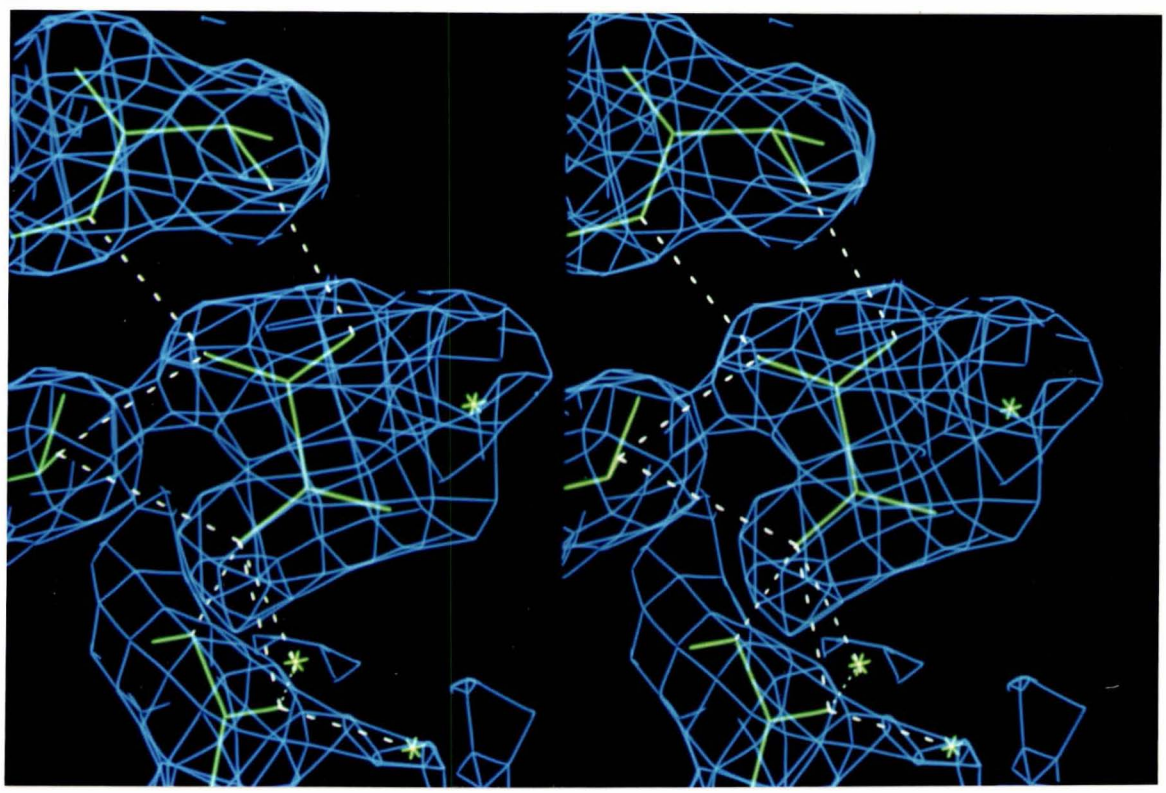
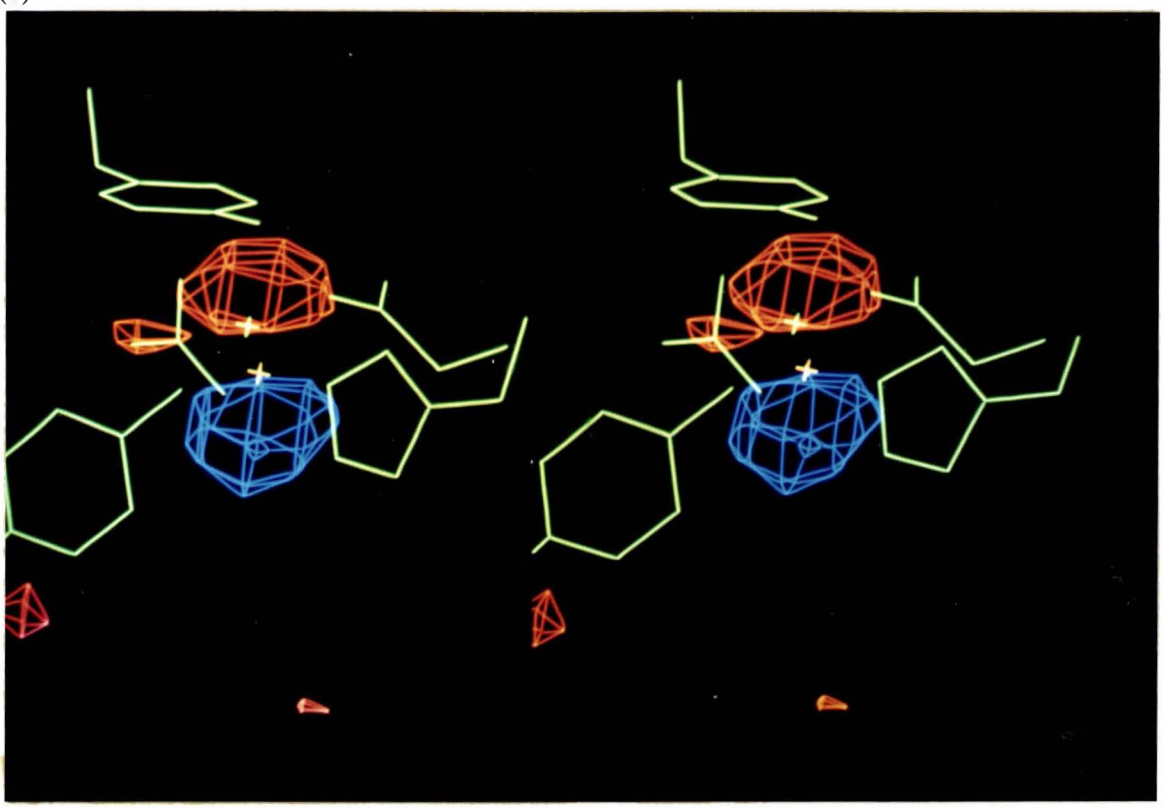


Plate 3.18: Hydrogen bonding interactions with (a) the N-lobe carbonate and (b) the C-lobe oxalate ions. In addition to 6 hydrogen bonds to protein atoms, the oxalate anion also interacts with a solvent molecule at bottom centre, through the $O_4$ atom.

(a)



(b)

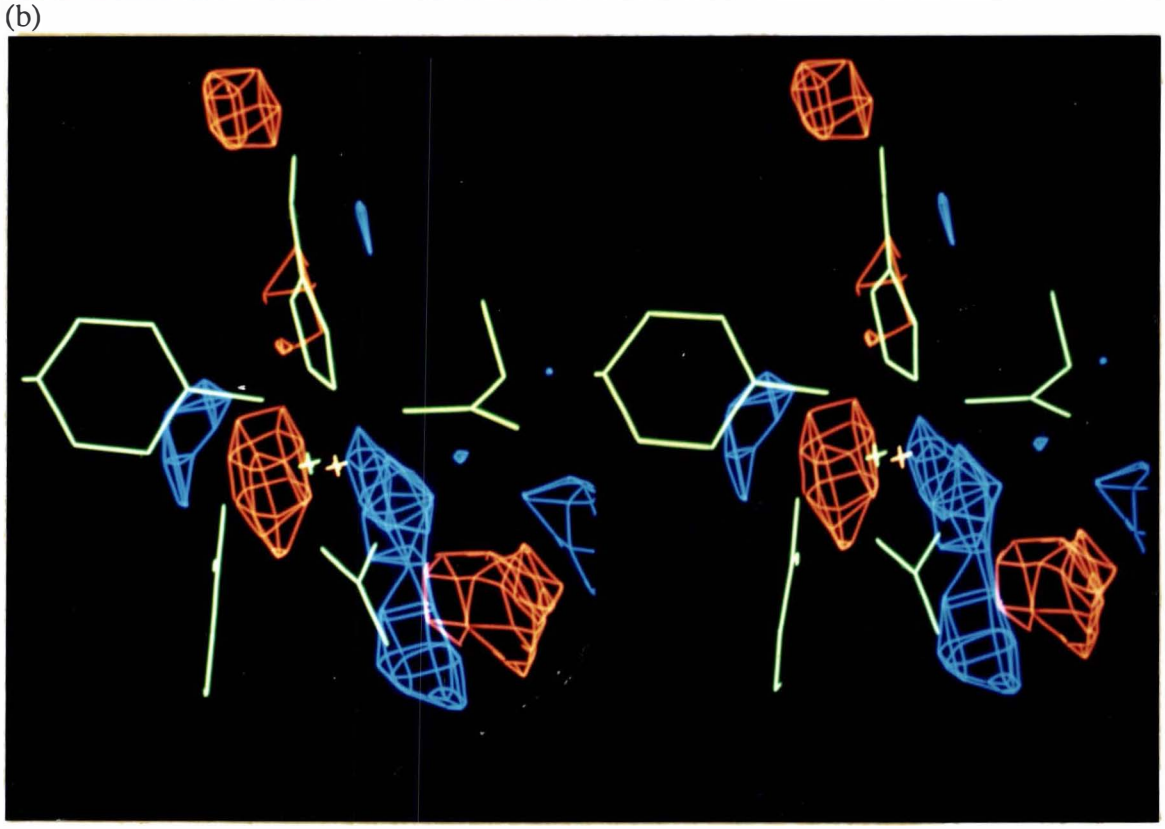


Plate 3.19: $Cu_2$oxLf - $Fe_2$Lf difference electron density map (calculated with coefficient $\Delta F = F_{(Cu_2oxLf)} - F_{(Fe_2Lf)}$ ) in (a) the N-lobe and (b) the C-lobe sites. Negative density are shown in red and positive density in blue. The $Fe_2$Lf model is superimposed in green, with the positions of the copper atoms in $Cu_2$oxLf given as red crosses (lower cross in photograph (a) and right-hand cross in photograph (b)).

oxalate into the anion binding site, as described above, while the changes observed in the N-lobe site relate only to the position of the metal ion. Relative to $Fe_2Lf$, the N-lobe metal ion appears to have moved about 1 Å from the iron position, while compared to $Cu_2Lf$, the movement of the metal ion is almost negligible. In the C-lobe, the metal ion is in roughly the same position as in the two other structures.

In general, the N-lobe metal and anion binding model for $Cu_2oxLf$ appear to be closely similar to that for the N-lobe site of $Cu_2Lf$, at least with respect to the hydrogen bonding interactions between the anion and the protein (Table 3.26). Hydrogen bonding interactions in other parts of the molecule, for example the interdomain contacts, are also very similar to those observed in $Cu_2Lf$ and $Fe_2Lf$.

*Polypeptide chain and domain structure:* Superposition of $Cu_2oxLf$ on to $Fe_2Lf$ and $Cu_2Lf$ results in rms deviations in $C_\alpha$ positions of 0.32 and 0.31 Å respectively. The largest rms deviations in $C_\alpha$ positions between $Cu_2oxLf$, $Fe_2Lf$ and $Cu_2Lf$ are in the external loop regions at residues 83 - 87, 218 - 222, 280 - 295, 416 - 424, 480 - 482 and 555-556, as described previously for $Cu_2Lf$. The deeply buried parts of the structure near the metal binding sites have rms deviations in the order of 0.1 - 0.3 Å.

Table 3.27: Domain relationships between $Cu_2oxLf$ and $Fe_2Lf$ [a]

| | | $Fe_2Lf$ | | | |
|---|---|---|---|---|---|
| | | N1 | N2 | C1 | C2 |
| | N1 | 0.27 Å | 0.9° | 0.5° | 0.2° |
| $Cu_2oxLf$ | N2 | — | 0.20 Å | 0.3° | 1.1° |
| | C1 | — | — | 0.22 Å | 0.8° |
| | C2 | — | — | — | 0.25 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, calculated as described in section 7.1.

Superpositions of the $Cu_2oxLf$ model on to both $Fe_2Lf$ and $Cu_2Lf$ (following the method described in detail in section 3.2.8), show that there is very little relative domain movement

resulting either from the substitution of copper(II) for iron(III) or for the substitution of oxalate for (bi)carbonate (see Tables 3.27 and 3.28). There is, however, a slightly larger movement of the C1 domain relative to the C2 domain (0.8° relative to $Fe_2Lf$ and 0.5 relative to $Cu_2Lf$, compared to only 0.2° between $Cu_2Lf$ and $Fe_2Lf$, Table 3.20), suggesting that the C-lobe, which binds the oxalate in place of the carbonate, may be very slightly more open than it is in either $Fe_2Lf$ or $Cu_2Lf$. The somewhat larger rotation (0.9°) between domains N1 and N2 for $Cu_2oxLf$ relative to $Fe_2Lf$, is also suggestive of a slightly more open N-lobe, in line with the results obtained from the superposition of $Cu_2Lf$ and $Fe_2Lf$ (where the rotation was about 0.6°). As the carbonate coordination is essentially the same in the N-lobes of both $Cu_2oxLf$ and $Cu_2Lf$, then this may reflect a slight opening of the lobe, as suggested previously (section 3.2.8), although it must be kept in mind that the atomic coordinates of $Cu_2Lf$ were used as the starting model in the refinement of $Cu_2oxLf$.

Table 3.28: Domain relationships between $Cu_2oxLf$ and $Cu_2Lf$

| | | $Cu_2Lf$ | | | |
|---|---|---|---|---|---|
| | | N1 | N2 | C1 | C2 |
| | N1 | 0.22 Å | 0.2° | 0.2° | 0.3° |
| $Cu_2oxLf$ | N2 | — | 0.18 Å | 0.0° | 0.6° |
| | C1 | — | — | 0.20 Å | 0.5° |
| | C2 | — | — | — | 0.19 Å |

*Metal and anion sites:* Superpositions of the separate N- and C-terminal lobes of $Cu_2oxLf$ onto the corresponding lobes of $Cu_2Lf$ and $Fe_2Lf$, using only the $C_\alpha$ atoms of the secondary structure elements in the least squares calculation (Plates 3.20a & b), show that the N-lobe metal and anion sites in the two structures are almost identical (Plate 3.20a), with the metal ion, the anion and the ligands occupying roughly the same positions (see Table 3.29B). The carbonate ion has rotated slightly about the $O_3$ atom but the size of this rotation is only about 5°. The combined effect of this rotation and the movement of the copper (relative to the iron in $Fe_2Lf$) still allows the copper to adopt a 5-coordinate geometry.

In the C-lobe of $Cu_2$oxLf, the copper atom has shifted ≈0.2 Å relative to $Cu_2$Lf and ≈0.3 Å relative to $Fe_2$Lf, while the four protein ligands occupy the same positions to within about 0.1 - 0.2 Å (Plate 3.20b). Tables 3.29A & B give the differences between $Cu_2$oxLf and $Cu_2$Lf in the positions of the metal atoms, carbonate anions and the ligand atoms directly bound to the copper atom.

Table 3.29A: The differences in the positions of the metal ions and the anion atoms between $Cu_2$oxLf and $Cu_2$Lf, and between the two lobes of $Cu_2$oxLf.

| | | Between $Cu_2$oxLf and $Cu_2$Lf | |
|---|---|---|---|
| | Between lobes of $Cu_2$oxLf | N-lobe [a] | C-lobe [a] |
| metal ion | 0.8 Å | 0.5 Å | 0.4 Å |
| C | — | 0.5 Å | — |
| O1 | — | 0.9 Å | — |
| O2 | — | 0.8 Å | — |
| O3 | — | 0.2 Å | — |

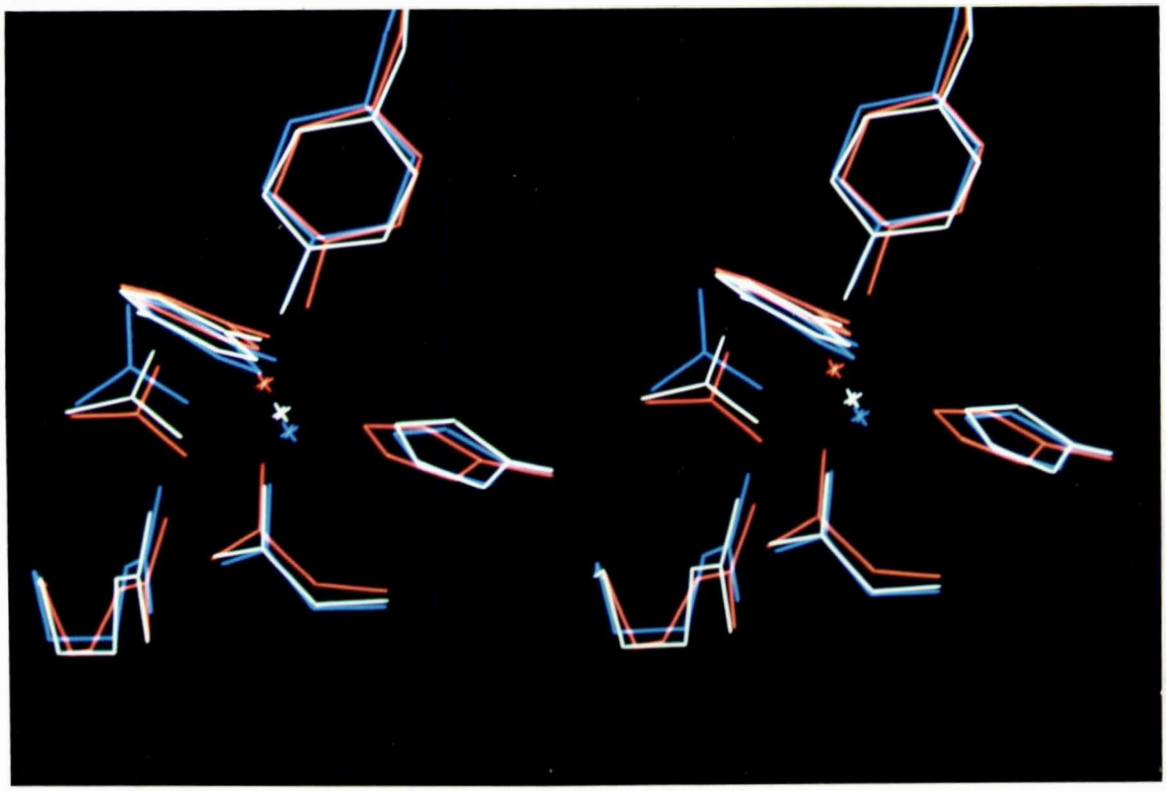
[a] Superimposed using secondary structure elements only. [b] The N- and C-lobes of $Cu_2$oxLf were superimposed onto the corresponding lobes of $Cu_2$Lf using the secondary structure.

Table 3.29B: The rms displacement in donor atom and sidechain atom positions between $Cu_2$oxLf and $Cu_2$Lf and between the two lobes of $Cu_2$oxLf.

| | | N-lobe | | C-lobe | |
|---|---|---|---|---|---|
| | Between lobes of $Cu_2$oxLf | donor atoms | sidechain atoms [c] | donor atoms | sidechain atoms |
| Asp60 (395) - $O_{\delta 1}$ | 0.7 Å | 0.2 Å | 0.2 Å | 0.4 Å | 0.4 Å |
| Tyr92 (435) - $O_\eta$ | 0.5 Å | 0.3 Å | 0.4 Å | 0.4 Å | 0.5 Å |
| Tyr192 (528) - $O_\eta$ | 0.8 Å | 0.5 Å | 0.5 Å | 0.5 Å | 0.6 Å |
| His253 (597) - $N_{\varepsilon 1}$ | 0.8 Å | 0.5 Å | 0.6 Å | 0.4 Å | 0.5 Å |
| Thr117 (461) - $O_{\gamma 1}$ | 0.5 Å | 0.3 Å | 0.3 Å | 0.2 Å | 0.2 Å |
| Arg121 (465) - $N_\varepsilon$ | 1.1 Å | 0.2 Å | 0.4 Å | 1.0 Å | 2.7 Å |
| $C_\zeta$ | 1.8 Å | 0.3 Å | | 1.9 Å | |
| $N_{\eta 1}$ | 2.1 Å | 0.6 Å | | 3.1 Å | |
| $N_{\eta 2}$ | 2.1 Å | 0.4 Å | | 1.9 Å | |
| Gly123 (467) - N | 0.1 Å | 0.1 Å | — | 0.2 Å | — |
| Ala124 (468) - N | 0.4 Å | 0.3 Å | — | 0.1 Å | — |

[a] The lobes were superimposed individually on the basis of the secondary structure elements, excluding a region of three residues centred on each of the 4 ligands in each lobe. The metal ions and the anions were not included in the calculation.

(a)



(b)

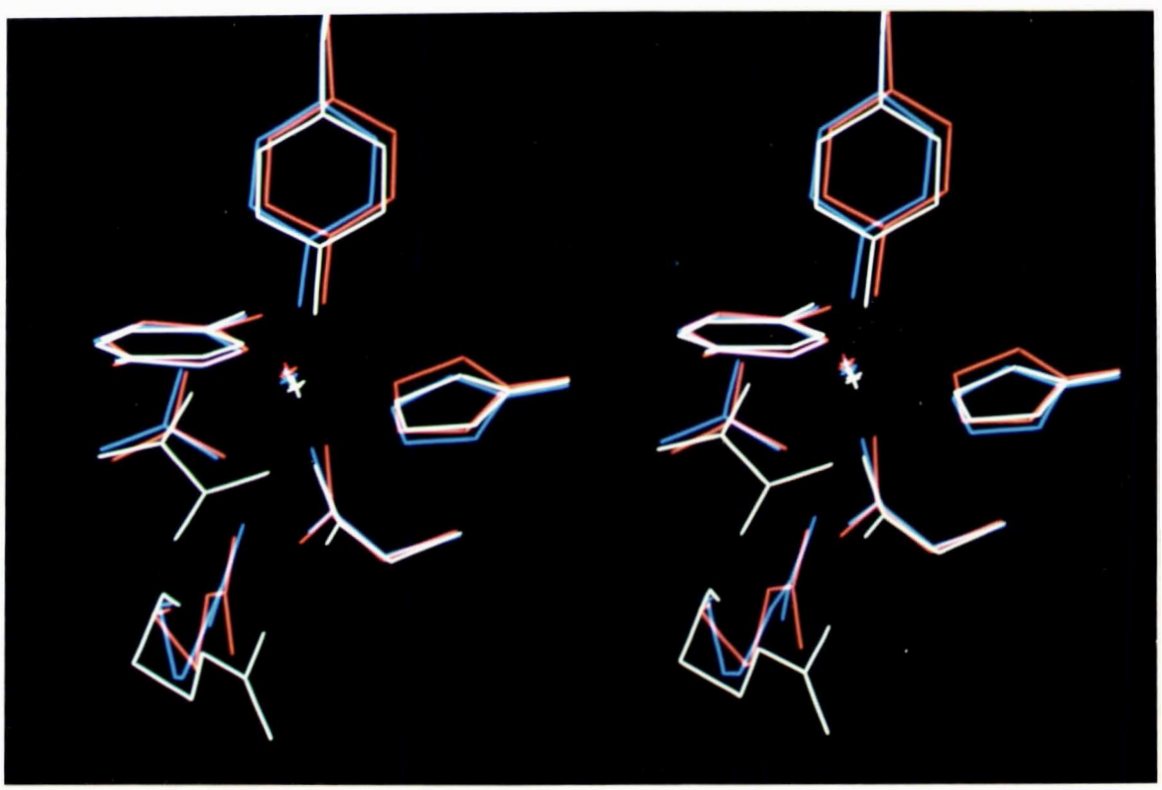


Plate 3.20: Stereo views of the superposition of the (a) N-lobe and (b) C-lobe metal binding sites of $Cu_2$oxLf (white) on to the corresponding sites in $Cu_2$Lf (blue) and $Fe_2$Lf (red). The photographs show, clockwise from the top, Tyr92 (435), His253 (597), Asp60 (395), Arg121 (465), the carbonate (in photograph (a)) and the oxalate (in photograph (b)) and Tyr192 (528).

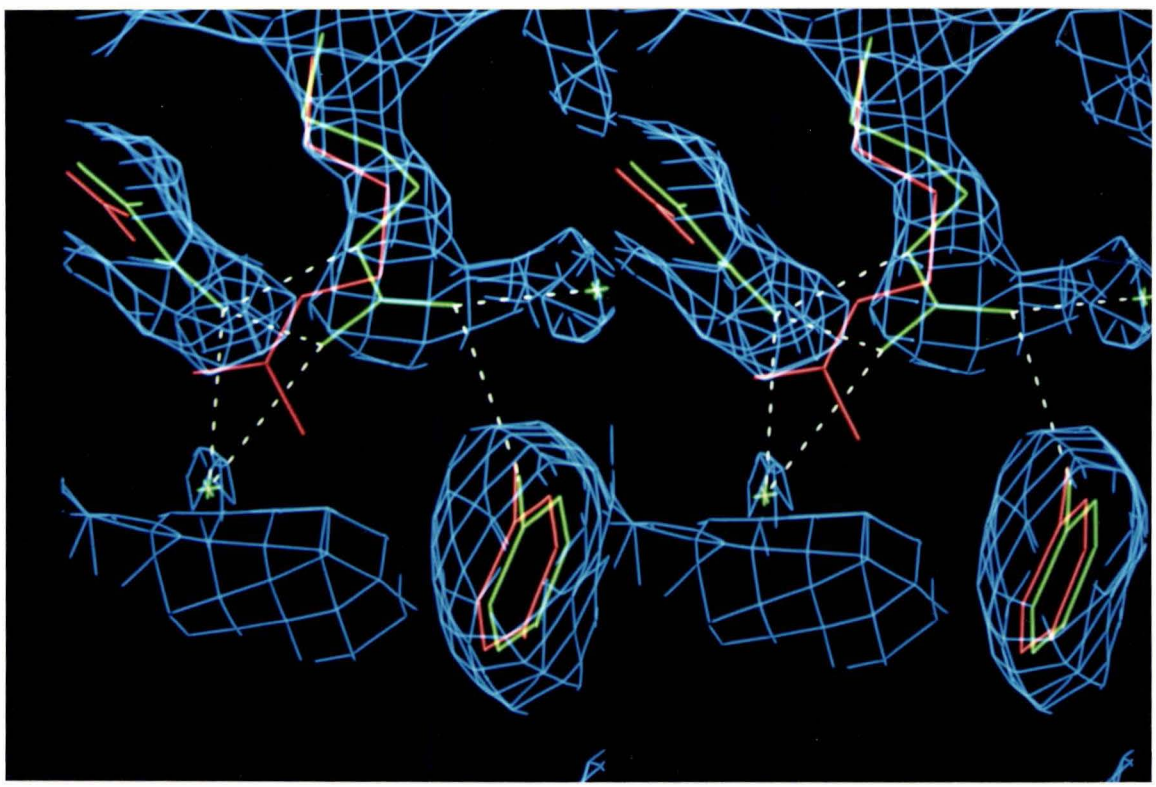


Plate 3.21: Stereo view of the superposition of Arg465 (centre) and Tyr526 (bottom right) in the C-lobe site of $Cu_2Lf$ (red) onto the corresponding site in $Cu_2oxLf$ (green). The carbonate and oxalate anions are shown edge on at the top left. The solvent molecule hydrogen bonded to the oxalate is at bottom centre.

The most evident difference between the C-lobe metal binding sites in the two structures is clearly related to the substitution of the carbonate anion by an oxalate. The incorporation of the larger oxalate anion results in the movement of the sidechain of Arg465 (the average movement in the four atoms of the guanidinium group, $N_\varepsilon$, $C_\zeta$, $N_{\eta 1}$ and $N_{\eta 2}$, is in the order of 2.7 Å) and brings the $N_{\eta 1}$ atom within hydrogen bonding distance of the $O_\eta$ atom of Tyr526. Plate 3.21 shows the relative movement of Arg465 and and the new hydrogen bonding interaction with Tyr526, resulting from the substitution of carbonate by oxalate.

The movement of the bulky arginine sidechain also causes the movement of the sidechain of Tyr398 by ≈1.5 Å. The resulting orientation of the $N_{\eta 1}$ of Arg465 and the $O_\eta$ of Tyr398 does not lead to the formation of a hydrogen bond, as the $O_\eta$ - $N_{\eta 1}$ distance is over 5.5 Å. Clearly the observed movement of the phenolate group is caused by van der Waals interactions with the approaching arginine.

Although the solvent structure and the hydrogen bonding patterns in the N-lobe remains closely similar between $Cu_2Lf$ and $Cu_2oxLf$, the pattern observed in the C-lobe of $Cu_2oxLf$ has been altered relative to $Cu_2Lf$, particularly with respect to the solvent structure. The position vacated by the movement of $O_\eta$ (398) appears to be occupied by a solvent molecule (≈1.5 Å from the $O_\eta$ position in $Cu_2Lf$) unique to $Cu_2oxLf$ (designated $O_{w3}$ 713) which is hydrogen bonded to the $N_{\eta 1}$ atom of Arg465. In addition, there is second unique water molecule in $Cu_2oxLf$ (designated $O_{w4}$ 721) near the position occupied by the guanidinium group of Arg465 in $Cu_2Lf$. It is hydrogen bonded to the $N_{\eta 2}$ atom of Arg465 and is also directly hydrogen bonded to the oxalate anion via $O_4$. The majority of the other solvent molecules near residues 398, 465 and 526 are in conserved positions relative to $Cu_2Lf$, including one which is hydrogen bonded to $O_\eta$ (526) in both structures.

## 3.4 Discussion

### 3.4.1 *Effects of metal substitution*

*Overall structure:* The structural analysis of dicupric lactoferrin ($Cu_2Lf$) has demonstrated that the polypeptide chain conformation when lactoferrin binds $Cu^{2+}$ is the same as when $Fe^{3+}$ is bound, at least within experimental error. That is, when $Cu^{2+}$ binds to apolactoferrin in its "open" configuration (Anderson *et al.*, 1990), it is able to induce a similar domain closure as $Fe^{3+}$ in both halves of the molecule. If the substitution of $Cu^{2+}$ for $Fe^{3+}$ has induced any changes in the polypeptide conformation this should be reflected in the extent to which different parts of the structure superimpose. Thus a change in the relative orientation of the two lobes should cause the whole molecule to superimpose less well than do the individual lobes. Similarly a change in the closure of the two domains in a given lobe should cause the lobe to superimpose less well than its individual domains.

Superpositions reveal little difference between $Cu_2Lf$ and $Fe_2Lf$. The rms deviation when the whole molecules are superimposed is 0.26 Å compared with 0.27 Å and 0.23 Å for the N- and C-lobes (using only $C_\alpha$ atoms in the helices and β-strands in the least squares calculation), suggesting no significant differences in orientation of the two lobes. Likewise superpositions offer little suggestion of any differences in the closure of the domains in either lobe, when $Cu_2Lf$ is compared with $Fe_2Lf$. The N-lobes of the two structures agree to an rms deviation 0.27 Å ($C_\alpha$ atoms in the helices and β-sheets) compared with 0.29 Å and 0.23 Å for the N1 and N2 domains. The greatest differences in the polypeptide chain occur in the N1 domain where there are 18 $C_\alpha$ atoms differing in position by >0.5 Å (when $Cu_2Lf$ is superimposed on $Fe_2Lf$), compared with the N2 domain where there are only 7 $C_\alpha$ atoms with rms deviations greater than 0.5 Å. The C-lobe agrees even better, with an rms deviation of 0.23 Å, compared with 0.22 Å and 0.20 Å for the C1 and C2 domains.

Measurements of apparent differences in domain orientations have been presented in Section 3.2.8. These differences are extremely small and must be at the level of the experimental error. In the N-lobe of $Cu_2Lf$, the domains are calculated to be 0.6° more open than in $Fe_2Lf$, while in the C-lobe the difference is only ≈0.2°. Comparison of the N-lobe of $Cu_2$oxLf, however, which has a similar copper site, shows a 0.9° opening compared with

$Fe_2Lf$ (and a difference of only 0.3° with $Cu_2Lf$) suggesting that although the difference in domain opening on substitution of $Cu^{2+}$ for $Fe^{3+}$ is small, it may be real in the N-lobe at least.

The lesser change in the C-lobe compared with the N-lobe is consistent with observations from the apoLf structure where the C-lobe remains closed even though no metal is bound. Whatever the reason for this, the very close correspondence of the C-lobes of apoLf and $Fe_2Lf$ implies that this closed structure is particularly stable and well defined.

The N-lobe of $Cu_2Lf$ has few interdomain interactions. There are two salt bridges, Glu216 - Lys301 in the interdomain cleft and Asp217 - Lys296 near the mouth of the cleft. There are also 18 hydrogen bonds in the vicinity of the metal binding site (see Table 3, Appendix V), but 15 of these are at the inner end of the cleft (involving the two "cross-over" β-strands, e and j) and are not in a position to significantly stabilise the closed structure. The remaining three interactions include the two interdomain interactions involving Asp60 (see Section 3.2.7) and the bond between Tyr82 ($O_\eta$) and Arg210 ($N_{\eta 2}$). It should be noted that there is no significant difference in the nature of these interactions in $Fe_2Lf$ and $Cu_2Lf$. Excluding the interactions at the inner end of the cleft, there are thus only two salt bridges and three hydrogen bonds which can be regarded as true constraints on the two domains.

The C-lobe, on the other hand, has seven hydrogen bonding interactions between the two domains (see Table 4, Appendix V) which constrain the C1 and C2 domains. All of these interactions except one are conserved in $Fe_2Lf$ and apoLf, and the greater number of these interdomain interactions may explain why the "closed" structure (C-lobe) is more stable relative to the C-lobe appears to be more constrained than the N-lobe. There are, in addition, 11 hydrogen bonds involving the "cross-over" strands near the metal site (see Table 4, Appendix V), but as noted above, they would not help to stabilise the closed structure.

*Copper coordination:* The structural analysis of copper-lactoferrin shows that there is limited internal flexibility in each binding cleft, sufficient to allow some change in metal coordination. In the N-lobe, the metal ion is five coordinate, interacting with four protein ligands (Asp, His and 2 Tyr) and a monodentate (bi)carbonate ion, while in the C-lobe site

the coordination geometry is much more like that in $Fe_2Lf$, where the metal is bound to the four protein ligands and a bidentate anion. In both sites the principal change is a movement of the copper atom relative to the iron. This movement is greater in the N-lobe (1.1 Å) than in the C-lobe (0.3 Å), and it is the size of this movement in the N-lobe, when coupled with a small change in orientation of the anion, which results in the copper adopting a five coordinate geometry. In both cases the protein ligands move to remain in contact with the metal ion.

The copper(II) ion is generally considered to be a borderline "hard" acid, with O and N being the most dominant ligands (Hathaway, 1987). In small molecule complexes and in biological systems, 6-, 5- and 4-coordinate complexes are common, with a variety of different stereochemistries, ranging from octahedral ($K_2PbCu(NO_2)_6$), trigonal bipyramidal ($Cr(NH_3)_6CuCl_5$) and square pyramidal ($[Cu(NH_3)_4H_2O]SO_4$) to tetrahedral ($Cs_2CuCl_4$) (Hathaway & Billing, 1970). However, Cu(II) is unlike most other first row transition metals in that these geometries are seldom regular. In an octahedral field, the ninth electron in the $d^9$ outer electron configuration of the copper(II) ion has the option of entering either the $d_{z^2}$ or $d_{x^2-y^2}$ orbital. Hence the octahedral complex is degenerate and subject to Jahn-Teller distortion, which typically manifests itself as either elongation or compression of the bonds in an axial direction (tetragonal distortion). Energy calculations (Hathaway, 1987) suggest that elongated structures are more favourable than either regular or compressed geometries, which is consistent with their more frequent occurrence. Virtually every octahedral copper(II) structure known exhibits Jahn-Teller distortion to some degree, there being less than 20 complexes which can be deemed as having regular octahedral stereochemistry or compressed structures (Hathaway, 1987).

Examples of elongated structures include the $Cu(NH_3)_6^{2+}$ ion, which has 4 bonds at ≈2.1 Å and 2 at ≈2.6 Å (Huheey, 1978); the $[Cu(imidazole)_6]^{2+}$ ion with 4 bonds between 2.01 and 2.05 Å and 2 at 2.6 Å (McFadden *et al.*, 1975); the $[Cu(OH)_6]^{4-}$ ion (4 bonds at ≈1.96 Å and 2 at 2.81 Å; Dubler *et al.*, 1973) and $[Cu(imidazole)_4(OH_2)_2]F_2$ with 4 bonds between 1.98 Å and 2.03 Å and 2 at 2.64 Å (Vreugdenhil *et al.*, 1984). The existence long bonds to the axial ligands in complexes such as these has been termed "semi-coordination" (Hathaway

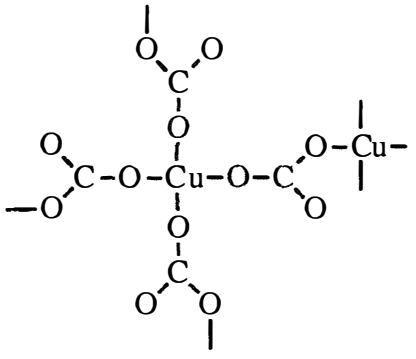
& Billing, 1970), implying weak bonding of these ligands to the copper(II) ion. Typically the elongated copper-ligand bonds are between 0.4 - 0.8 Å longer than the "normal" equatorial distances (≈2.0 Å for Cu-O and Cu-N).

Five-coordinate copper(II) can adopt either a trigonal bipyramidal or a square pyramidal geometry, both of which occur with about the same frequency (Hathaway, 1987). The two stereochemistries are generally subject to distortions both in the trigonal or square plane, and in the bond lengths to the axial ligands. The square pyramidal geometries are almost always characterised by an elongation of the 5th apical bond, typically by about 0.2 - 0.5 Å. Examples of this type of stereochemistry include the $[Cu(NH_3)_4H_2O]^{2+}$ ion, where the four nitrogen donors are arranged in a square plane with Cu-N distances of ≈2.03 Å and the water coordinated in an axial position at a distance of ≈2.5 Å (Hathaway, 1987); the $Cu(NH_3)_5^{2+}$ ion, with in-plane bonds between 2.01 - 2.05 Å and the apical bond ≈ 2.2 Å (Duggan *et al.*, 1980) and $[Cu(1,3\text{-pn})_2H_2O]SO_4$ (1,3-pn = 1,3-propylenediamine) where the four in-plane Cu-N bonds are all 2.0 Å and the Cu-O($H_2O$) bond is 2.54 Å (Hathaway & Billing, 1970).

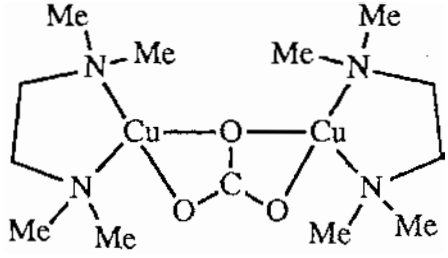
In this type of geometry (square pyramidal) the copper atom seldom lies in the plane of the four equatorial ligands, being usually displaced ≈ 0.2 Å out of the plane towards the more distant axial ligand. Largely as a consequence, the *trans*- equatorial ligands have bond angles between 150° and 170° (Hathaway, 1987).

Although 6-coordinate Cu(II) is as abundant as 5-coordinate (Hathaway, 1987), the copper geometry observed in the N-lobe of $Cu_2Lf$ (square pyramidal, with elongation of the apical ligand) is certainly more typical of Cu(II) than that seen in the C-lobe. Although the geometry in the C-site is clearly not regular octahedral (see Tables 3.16 & 3.17), the tetragonal distortion normally expected with 6-coordinate Cu(II) is not evident. Four of the bonds are 2.0 - 2.1 Å, while the Cu - $O_\eta$ (Tyr528) bond is ≈2.3 Å and the Cu - $O_1$ (anion) bond is ≈2.2 Å. These two bonds are *cis*- to each other however, whereas Jahn-Teller distortions typically involve the lengthening of bonds in a *trans*- arrangement. While it is possible that this could be a rare *cis*-distorted octahedral geometry (Hathaway, 1987), the estimated error in positional parameters in this part of the structure, although probably not greater than 0.2 Å, means that the differences observed in the copper bond lengths are not

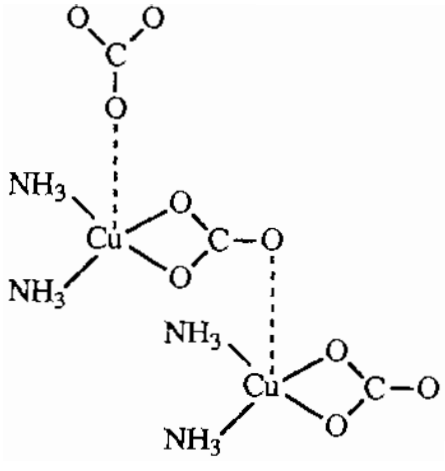
statistically significant. Either way, the N-site geometry would be expected to be more stable than the C-site geometry.



Monodentate $CO_3^{2-}$ coordination as in $K_2[Cu(CO_3)_2]$ and $Na_2[Cu(CO_3)_2]$.



Bridging bidentate coordination as in $[Cu_2(Me_4pn)_2Cl_2(CO_3)]$. The Cl- ions are bound to the Cu and are not shown.



Monodentate and bidentate coordination as in $[Cu(NH_3)_2CO_3]$.

Figure 3.29: Some typical modes of carbonate ion coordination in small molecule complexes.

This change from 6- to 5- coordination in the two lobes of $Cu_2Lf$ is facilitated by the change from bidentate to monodentate coordination of the synergistic carbonate ion. Small molecule complexes of Cu(II) with $CO_3^{2-}$, although not abundant, indicate that the carbonate can coordinate in a monodentate fashion, as in $K_2[Cu(CO_3)_2]$ (Farrand *et al.*, 1980) and $Na_2[Cu(CO_3)_2]$ (Healy & White, 1972), bidentate as observed in $[Cu_2(Me_4pn)_2Cl_2(CO_3)]$[§]

§ Where $Me_4pn$ = N,N,N',N'-tetramethylpropylenediamine

(Churchill *et al.*, 1979) or a combination of both as in $[Cu(NH_3)_2(CO_3)]$ (Meyer *et al.*, 1972). These stereochemistries are shown in Figure 3.29.

In addition , the carbonate ion can also be involved in a number of more complex bridging modes (Oldham, 1987). Although there is no evidence to suggest that monodentate coordination is favoured over bidentate, it has been observed with carboxylate ligands that the steric strain produced by bidentate chelation results in a preference for chelate rings with more than four members (Oldham, 1987). It is likely, therefore, that the monodentate binding of the anion observed in the N-lobe of $Cu_2Lf$ could be more energetically stable than bidentate coordination, and that the bidentate chelation in the C-lobe is influenced by steric restrictions imposed on the metal and anion by the surrounding protein (see later). In $Fe_2Lf$, however, the bidentate coordination observed in both sites is due to the preference of high spin Fe(III) for 6- rather than 5-coordinate geometries (Cotton & Wilkinson, 1988).

Protonation of metal-bound carbonate ions in small-molecule complexes generally leads to the loss of the carbonate as $CO_2$ (Oldham, 1987), so that very little is known about the coordination of bicarbonate to metal ions. Monodentate carbonate ions are particularly unstable when protonated. However, carbonate ions in a bridging mode can be converted to stable bridging bicarbonate ligands (Oldham, 1987).

*Origin of coordination differences between the two sites:* At first sight the most obvious difference in the protein structure of the two binding sites is the presence of Ser191 in the N-lobe, hydrogen bonded to the anion-binding Arg121. The equivalent residue in the C-lobe is Gly527, and Arg465 is not hydrogen bonded to any protein atom. It could be suggested that Ser191 might hold the arginine sidechain further from the helix 5 N-terminus, allowing the anion more freedom to move in the N-lobe, and thus adopt a monodentate coordination.

Superpositions show that this is not the case, however. When the C-lobe is superimposed on to the N-lobe, using just the N2 and C2 domains as the basis of the superposition (since it is these domains which define the anion binding site), it can be seen that the arginine sidechains and the N-terminus of helix 5 are in essentially identical positions (although the orientation of the guanidinium group varies slightly). This applies to both $Cu_2Lf$ and $Fe_2Lf$

(Plates 3.22a and b), ie. all four anion binding sites (the N- and C-sites for each of the two proteins) superimpose extremely closely. This is also shown by the distances between the arginine sidechain and the helix N-termini (the Arg $N_{\varepsilon}$ ··· Gly (N) (121 --- 123 and 465 --- 467) distances in $Fe_2$Lf are 6.4 Å and 6.3 Å for the N- and C-lobes respectively, while for $Cu_2$Lf, the equivalent distances are 6.3 Å and 6.5 Å) . In the $Fe_2$Lf structure the anion is also in the same place in both lobes (Plate 3.22b).

What, then, is the reason for the coordination differences? In both $Cu_2$Lf and $Fe_2$Lf the N-lobe cleft is more "open" than that of the C-lobe, the differences being 5.0° in the case of $Cu_2$Lf and 4.5° for $Fe_2$Lf. Although the effect must be small at the metal site, at the inner end of the cleft, it is sufficient to allow greater movement of the histidine ligand in the N-lobe than in the C-lobe (this movement can seen in Plate 3.22a). If the position adopted by His253 in the N-lobe is modelled on the C-lobe there is a clash with Asn644. Thus a possible scenario is as follows. The slightly greater space in the more open N-lobe cleft allows more movement in the metal ligands, and this in turn allows the copper atom to move further from the anion and Tyr92 (the His253 and Asp60 residues have to adjust to the new copper position). With the larger movement of the metal in the N-lobe, the anion is now able to adjust by a small rotation and take up a monodentate coordination mode, ie. it is suggested that the key difference is in the closure of the two domains over each site which allows the metal to move more, and the other ligands to adjust to it, but only if this is stereochemically desirable (it happens only with $Cu^{2+}$, not with $Fe^{3+}$).

*Functional implications:* 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 (5-coordinate, square pyramidal, compared with 6-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 (see Section 5.1).

(a)

(b)

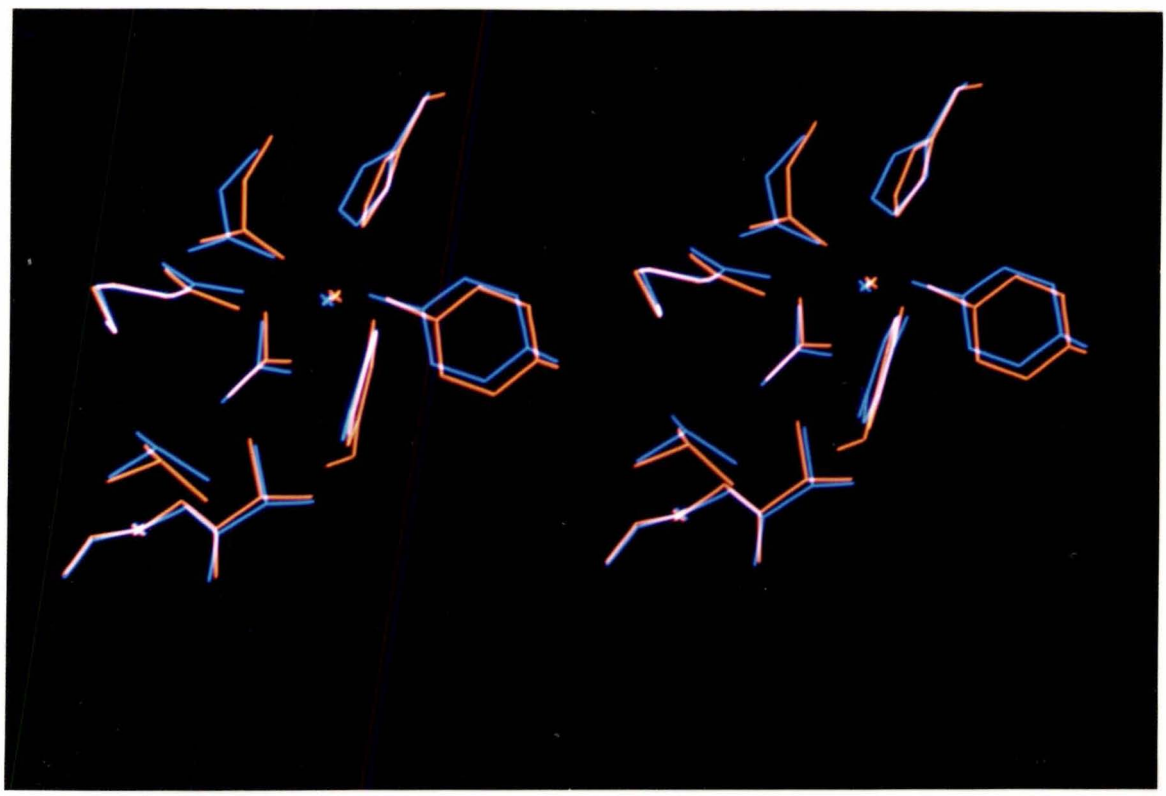


Plate 3.22: Stereoviews of the superposition of (a) the N-lobe (red) of $Cu_2Lf$ on to the C-lobe (blue) of $Cu_2Lf$ and (b) the N-lobe (red) of $Fe_2Lf$ on to the C-lobe (blue) of $Fe_2Lf$. The photographs show, clockwise from middle right, Tyr92 (435), Tyr192 (528), the helix 5 N-terminus, Thr117 (461), Arg121 (465), Asp60 (395) and His253 (597). The metal ions and the carbonate ions are in the centre.

A fundamental finding from this work, however, is that despite the differences in detail in metal coordination, when $Cu_2Lf$ is compared with $Fe_2Lf$, the protein structures are the same. Clearly, the metal ion does little to affect the overall protein structure. The only significant differences are in the positions and orientations of the metal binding ligands, and these appear to move in direct response to the observed movements of the metal ions. The differences in the binding of different metal ions (at least as far as iron and copper are concerned) are the result of a combination of the limited flexibility of the metal binding site (the ligands are capable of some movement) and the stereochemical restrictions of the metal ion.

It is anticipated that the same protein structure should be found when other metal ions of similar size to $Fe^{3+}$ are bound. This should apply, for example, to $Ga^{3+}$, $Cr^{3+}$, $Mn^{3+}$, $Co^{3+}$ and $Al^{3+}$, which all prefer octahedral geometries, are all classified as "hard" acids (all prefer O and N donors) and have ionic radii ranging from 0.535 Å to 0.645 Å. As noted previously, the relative strengths of binding of metal ions could be related to the geometrical preferences of the metal ion. Gallium(III) (ionic radius 0.62 Å) has a similar binding constant to $Fe^{3+}$ (radius 0.645 Å), with $Al^{3+}$ (radius 0.535 Å) somewhat lower. Zinc(II) has a similar radius (0.74 Å) to $Fe^{3+}$, but is more likely to adopt a coordination number less than six (Cotton & Wilkinson, 1988) and has a lower binding constant still. No binding constants are available for $Mn^{3+}$, $Co^{3+}$ or $Cu^{2+}$, although based on the stereochemical preferences of the former two (a preference for O and N donors, and, especially in the case of $Co^{3+}$, a strong preference for octahedral stereochemistry) and the small alterations in structure observed with the latter, it is reasonable to suggest that they would have association constants somewhere between $10^{15}$ and $10^{20}$.

It is likely that larger cations, such as the lanthanide and actinide ions, may not be accommodated in the same way as the smaller transition metals. That is, the larger size of the cation may require movements of the protein ligands which are not structurally possible within the confines of the closed structure or may require the inclusion of more ligands in the first coordination sphere. This can be shown, albeit rather crudely, by modelling a metal ion with ionic radius of $\approx$1 Å in the N-lobe site of $Cu_2Lf$, assuming that there is no increase in domain opening.

The larger metal ion can be accommodated by small movements of three of the four protein ligands (Asp60 and His253) away from the metal ion by ≈0.3 - 0.4 Å (Tyr92 is already at a distance of 2.8 Å and may be moved closer to the metal by ≈ 0.3 Å), but the required movement of the fourth (Tyr192) results in a very unfavourable contact between $O_\eta$ (Tyr192) and $N_{\varepsilon 2}$ of His253 (≈2.3 Å) (see Figure 3.30). These sidechain movements are accomplished by rotation about the $C_\alpha$ - $C_\beta$ bond and do not require the movement of the mainchain atoms. None of the other sidechains are in sterically unfavourable positions, although one of the two hydrogen bonds from the $O_{\delta 2}$ atom of Asp60 (to the $O_{\gamma 1}$ atom of Thr122) is lost, while that from $O_{\delta 2}$ to the peptide nitrogen of Thr122 is lengthened somewhat (2.91 Å to 3.25 Å).

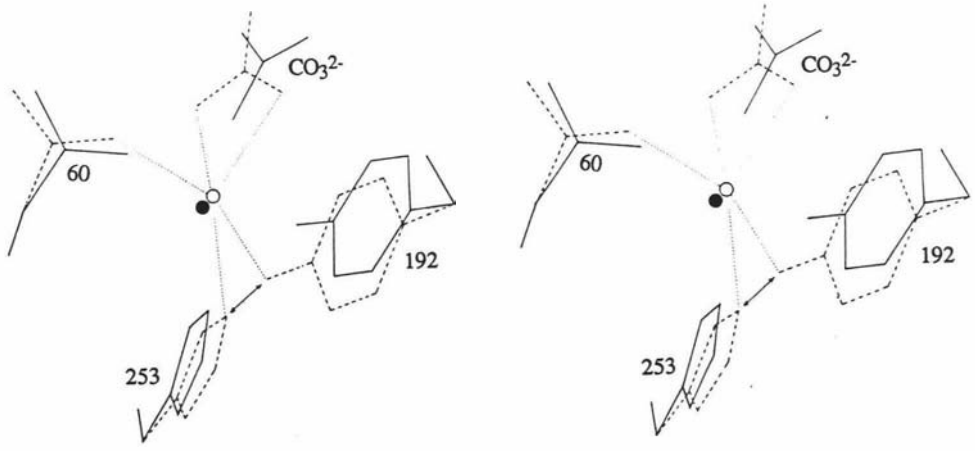



Figure 3.30: Substitution of a metal ion of ionic radius 1 Å into the N-site of $Cu_2Lf$. Only three of the metal binding ligands are shown (Asp60, Tyr192 and His253 - solid lines for copper and dashed lines for the larger cation). These sidechains have been moved away from the metal such that the metal-ligand bond lengths are about 2.4 Å. Tyr92 (not shown) has been moved closer to the metal ion (O - M distance ≈2.5 Å). In addition, the metal ion has been moved ≈0.2 Å and the carbonate ion rotated. The very short contact (≈2.3 Å) between $O_\eta$ (192) and $N_{\varepsilon 2}$ (253) is indicated.

Moreover, the increase in size of the metal ion means that the carbonate anion must also move. In order to overcome some unfavourable close contacts and maintain a hydrogen bonding pattern similar to that observed in $Fe_2Lf$ and $Cu_2Lf$, a reorientation of the anion so

that it coordinates in a bidentate fashion to the metal is required, and to accomplish this the metal ion must be moved closer to the carbonate by ≈0.2 Å.

Clearly, despite the limited conditions used in this modelling study, the binding of a lanthanide ion in the N-site of lactoferrin may be not be completely favourable without some degree of lobe opening.

In addition, lanthanide ions have a strong preference for coordination numbers greater than six (Hart, 1987). The effect of an increase in the coordination number of the metal ion is more difficult to judge, but due to the restricted space available in the metal binding sites in the closed lobe conformation, it is likely that increasing the coordination number would lead to a significant amount of steric crowding and the consequent opening of the lobe. It is conceivable that water molecules could coordinate to lanthanide ions with some rearrangement of the existing ligands. In the N-lobe metal binding site of $Cu_2Lf$ there are two solvent molecules (hydrogen bonded to His253 and Tyr92) which are within 4 -6 Å of the copper atom, while in the C-site there are three water molecules within 6 Å, two hydrogen bonded to His597 and Tyr435 and a third bridging between Tyr528 and Lys 546.

Spectroscopic studies (see Section 4.6.3) have implicated a third tyrosine residue in the binding of the lanthanide ions, based on the observed increase in the extinction coefficients of lanthanide-lactoferrin complexes relative to transition metal complexes. Transition metal complexes have $\Delta\varepsilon_1$ values of about 15000 - 20000 (where $\Delta\varepsilon_1$ is the extinction coefficient for the first equivalent of metal ion bound by the protein, see Chapter 4 for further discussion), while lanthanide complexes exhibit $\Delta\varepsilon_1$ values ranging from 20000 to 30000 (≈ 3/2 times greater than with transition metal ions). Analysis of the metal binding sites in $Fe_2Lf$ and $Cu_2Lf$ indicates that Tyr82 (415 in the C-lobe) is a possible candidate for lanthanide or actinide coordination. In both $Fe_2Lf$ and $Cu_2Lf$, the $O_\eta$ of these tyrosine residues are just under 6 Å from the metal ion, and although there do not appear to be any steric reasons why these residues could not move closer to the metal (ie. there are no protein atoms between the tyrosines and the metal), the movement required to bring the sidechain into coordinating position could not be achieved by simple rotations about bonds, but would involve the movement of the residue as a whole.

It should be noted that in bovine lactoferrin, the equivalent to Tyr415 is an arginine, which would not be involved in coordination to a metal ion. If Tyr415 in Lf is deprotonated by the lanthanide ions, then it is likely that the $\Delta\varepsilon_1$ values for the binding of the lanthanides to bLf will be lower than to Lf.

Although it is difficult to foresee what will happen with another metal based on the structural evidence from only iron and copper, the changes observed on going from Fe(III) to Cu(II) coordination in lactoferrin provide an answer to the question of which is more important in dictating how metal ions will bind, the protein or the metal ion itself. In the C-lobe, the metal binding ligands appear to be constrained by the more-closed nature of the domains, making this site consistent with the ideas of *rack-induced bonding* (Gray & Malmström, 1983). The copper atom in the C-lobe is not able to express its preferred geometry (tetragonally-distorted octahedral). Only the ligand positions are changed (although the extent of these changes is small) and the overall protein structure is not.

As discussed above, the more "open" nature of the domains in the N-lobe allows much more flexibility in the metal and anion binding sites of $Cu_2$Lf, such that the metal coordination is able to change to allow the copper to adopt one of its preferred geometries, square pyramidal. This is in marked contrast to proteins such as azurin, which provides a classic example of a metal binding site whose stereochemistry is determined by *rack-induced bonding*. In azurin, the copper is bound to two histidine residues (His46 and His117), Cys112, Met121 and the carbonyl oxygen of Gly45. The sidechain of Cys112 is highly constrained by two N-H ··· S hydrogen bonds, His46 interacts with the carbonyl oxygen of Asn10 via $N_{\varepsilon 2}$ and His117, while not involved in any hydrogen bonding, is sandwiched between the sidechains of Phe114 and Met13 (Baker, 1988).

In the N-lobe of $Cu_2$Lf, the majority of the metal binding ligands are not constrained by hydrogen bonds to protein atoms, with the exception of the Asp60, which is involved in interdomain interactions. Of the four ligands in the N-lobe binding site, Asp60 moves the least upon substitution of $Fe^{3+}$ by $Cu^{2+}$. The His253 sidechain is hydrogen bonded to a solvent molecule via $N_{\delta 1}$, while Tyr92 also interacts only with a water molecule. Tyr192, although hydrogen bonded to Arg210, is still capable of some movement, as the guanidinium

group of the arginine is not constrained. In this site, therefore, it appears that the observed coordination geometry of the copper is determined by both the stereochemical preferences of the metal ion and the flexibility of the protein structure. In the C-lobe site, although the ligands are no more restrained than those in the N-lobe (in terms of specific interactions) it appears to be the size of the site that limits the extent of adaptability, as noted earlier.

The conclusions made earlier concerning the limited flexibility of the metal binding ligands as a result of the binding of different metal ions and anions, and the observation that the change in the overall protein structure as a result is minimal, 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 internalisation should also occur, though whether the metal is then released will depend on the mechanism. Studies on the pH dependent release of Cu(II) from lactoferrin (Ainscough *et al.*, 1983), serum transferrin (Zweier & Aisen, 1977) and ovotransferrin (Zweier, 1980) shows that copper is lost at a pH higher than iron (for lactoferrin, copper is removed below pH 6 whereas iron does not begin to dissociate until below pH 4), although at physiological pH copper is 100% bound. If metal release is associated with a drop in pH inside a cell (Octave *et al.*, 1983), then copper will be released equally as well as iron. On the other hand, if low molecular weight chelators are involved (Pollack *et al.*, 1991), whether or not copper release takes place will also be dependent upon other factors such as the stability of $Cu^{2+}$-chelator complexes.

### 3.4.2 *Effects of anion substitution*

*The overall structure:* Superpositions of $Cu_2$oxLf on to both $Fe_2$Lf and $Cu_2$Lf, similar to those described for $Fe_2$Lf and $Cu_2$Lf earlier, show that once again there are only very small differences in the positions of the the mainchain atoms (rms deviation between mainchain

atoms = 0.30 Å, excluding the poorly ordered regions) and in the relative orientations of the four domains (between 0.3 and 1.0°). It is interesting to note that the largest relative domain movements in $Cu_2$oxLf compared to $Fe_2$Lf are 0.9° between the N1 and N2 domains, which opens the N-lobe slightly, and 1.5° between the N2 and C2 domains, while the C-lobe (domains C1 and C2) is more open by just over 0.6°. Compared to $Cu_2$Lf, the N2 - C2 and C1 - C2 domain movements of $Cu_2$oxLf are in the order of 0.6°, while the N1-N2 angular displacement is ≈ 0.3°.

These results suggest that the binding of copper(II) in the N-lobe again causes a slight opening of the N1 and N2 domains (as discussed earlier) and that replacement of carbonate by oxalate in the C-lobe site causes a slight opening of the C1 and C2 domains. The similar opening of the N-lobe domains in $Cu_2$oxLf compared to $Cu_2$Lf can be interpreted as an indication that the substitution of oxalate in the C-lobe of the molecule causes little disturbance to the N-lobe. Given these relative domain movements, and in the light of the structural studies on apoLf (where it has been shown that it is the N2 domain which moves upon lobe opening (Norris *et al.*, 1991)), it is probable that it is the C2 domain which moves upon oxalate binding, while the N1 and C1 domains represent an invariant "core" of the molecule.

*Metal and anion binding:* The interlocking sites model (Schlabach & Bates, 1975; see also Figure 2.1a) implies that dicarboxylate anions such as oxalate, malonate and maleate would coordinate to metal ions in a monodentate or perhaps a 1,1-bidentate fashion, ie. such that one end of the molecule interacts with the metal and the other end interacts with the protein. Dicarboxylic acids, however, are usually coordinated in a 1,2-bidentate fashion, as observed in small molecule studies on copper-oxalato species such as the $[Cu(C_2O_4)_2(H_2O)_2]^{2-}$ ion (Gleizes *et al.*, 1980), $[Cu(bipy)(C_2O_4)(OH_2)]\cdot 2H_2O$ and $[Cu(bipy)(C_2O_4)]\cdot 2H_2O$ (Fitzgerald *et al.*, 1982). In the first two complexes, the oxalate is symmetrically coordinated to the copper (Cu-O bond lengths between 1.93 and 1.97 Å), while the latter complex has the oxalate ion bridging between two copper atoms, with asymmetric bidentate 1,2-coordination (the Cu-O bond lengths are 1.99 and 2.32 Å).

Monodentate coordination is not common for dicarboxylic acids, and although the oxalate ion is capable of coordinating in a 1,1-bidentate mode, this type of chelation occurs more widely with the higher members of the dicarboxylate series (Oldham, 1987). Therefore, 1,1-bidentate and monodentate coordination in the transferrins, as in the interlocking sites model, seemed less likely than 1,2-bidentate coordination, given also the limited space available in the anion pocket.


1,1-bidentate chelation of an oxalate ion

1,2-bidentate chelation of an oxalate ion

Figure 3.31: Two modes of oxalate chelation to metal ions from small molecule structures.

Modelling of oxalate binding in lactoferrin suggested that the mode of oxalate coordination to the copper would indeed be 1,2-bidentate,. Not only did oxalate fit into the anion site more readily in this orientation, but its chelate angle of ≈ 80° appeared be more favourable than the 60° chelate angle of 1,1-bidentate coordination (Figure 3.31). These modelling experiments further suggested that oxalate should bind preferentially in the N-site anion pocket, as the space available to the anion appeared to be somewhat greater than for its counterpart in the C-lobe (Baker *et al.*, 1990).

This latter conclusion was consistent with the results of Zweier & Aisen (1977), who observed that oxalate bound only in the N-terminal site of hTf in the presence of Cu(II). The most surprising aspect of the structure of $Cu_2$oxLf, therefore, was the discovery that oxalate had bound preferentially in the apparently more structurally constrained C-terminal anion binding site. It should be noted, however, that this observation is in accord with results on cOTf (Zweier, 1981), for which oxalate could be bound to both sites with $Cu^{2+}$ as the associated metal ion, but with a five-fold greater affinity in the C-site than in the N-site.

The C-lobe anion pocket clearly has the ability to adapt to accept a larger molecule. The Arg465 sidechain, interacting only with the anion and a network of water molecules, is pushed aside to accommodate the oxalate. The solvent structure near the arginine is disrupted and the only other movement appears to be in the position of Tyr398, which is pushed away by the arginine sidechain. In fact, the movement of the guanidinium group brings the $N_{\eta 1}$ atom into hydrogen bonding distance of $O_{\eta}$ of Tyr526, thus stabilising the arginine in this new conformation.

The preferential binding of oxalate in the C-site raises an important question. If there is greater space in the N-lobe binding cleft, sufficient to allow the metal ion to move by ≈1 Å and the carbonate to reorient itself and still maintain a favourable hydrogen bonding network, why does an oxalate ion not bind preferentially to this site under the conditions of the substitution experiment?

One possible answer has been raised in section 3.4.1. With carbonate as the anion, the N-lobe copper ion is able to adopt a favourable square pyramidal geometry. The extra space in this site allows the copper to move more and the anion to adjust to remain coordinated to the metal while maintaining its hydrogen bonded interactions with the residues in the anion pocket. Oxalate binding in this site would have two consequences: (i) the oxalate would almost certainly bind in a 1,2-bidentate fashion, making the metal geometry octahedral rather than square pyramidal and this may be energetically less favourable, given the normal coordination geometry of copper(II) (see discussion earlier); (ii) oxalate binding would require the movement of the sidechain of Arg121, with possible adverse effects on the interactions made by this residue.

It appears, therefore, to be a question of stability; the Arg121 sidechain, as seen in $Cu_2Lf$ and $Fe_2Lf$, is in relatively stable position, in that the majority of its potential hydrogen bonding interactions are satisfied, by the anion ($N_{\varepsilon} \cdots O_3$, $N_{\eta 2} \cdots O_2$), the sidechain of Ser191 ($N_{\eta 1} \cdots O_{\gamma}$) and one solvent molecule ($N_{\eta 1} \cdots 715\ O_{w1}$). Movement into a new position some 2 Å away, caused by the substitution of carbonate by oxalate could result in a disruption of the hydrogen bonding and a lowering of the stability of the sidechain. A combination of the above factors may make carbonate substitution by oxalate less favourable

in the N-lobe of lactoferrin when copper is the bound metal (but it should be noted that this would not necessarily be the case for $Fe^{3+}$, because of its preference for octahedral coordination).

*Comparison with other transferrins:* The results of Zweier & Aisen (1977) and Zweier (1981), mentioned above, raise interesting questions related to the differences between the various members of the transferrin family.

With respect to oxalate binding at least, lactoferrin resembles ovotransferrin. However, as the C-lobe site of cOTf accepts oxalate in preference to the N-lobe (Zweier & Aisen, 1977), this appears to be inconsistent with the monocupric complexes of cOTf reported by Yamamura *et al.* (1984). The N-terminal site of cOTf has an estimated visible absorption maximum at 450 nm, while that of the C-terminal site is close to 430 nm. Dicupric ovotransferrin has a $\lambda_{max}$ near 440 nm, which could be viewed as an average of the values for the two sites. It should be noted, however, that this study did not involve the separation of the two lobes of cOTf but rather the $\lambda_{max}$ values were based on difference spectra between Cu-Fe-OTf complexes and $Fe_N$-OTf (N-site preferentially loaded).

Crystal field arguments (see Section 4.4.1) would predict that the site which has the lower $\lambda_{max}$ (the C-terminal site) contains the monodentate carbonate and, based on the arguments outlined for $Cu_2$Lf above, must represent a more favourable geometry for the copper. Oxalate, therefore, should not go into the C-lobe site of cOTf, but rather it should bind in the N-lobe. Clearly the arguments given earlier with respect to lactoferrin are an oversimplification; it is possible that subtle interactions in the vicinity of the binding sites, coupled with differences in the closure of the lobes could make one site more accommodating to a larger anion than the other, independent of the coordination preferences of the metal. Herein lies a problem, however, as a high resolution structure for cOTf is not as yet available.

The structure of rabbit serum transferrin (rSTf), on the other hand, is known (both the full molecule (Bailey *et al.*, 1988) and the N-lobe half (Sarra *et al.*, 1990)). Compared with $Fe_2$Lf, the two structures are very similar in terms of secondary structure, the overall

conformation of the lobes and the metal and anion binding sites. There are differences, however, in the orientation of the two lobes of the full molecule (Bailey *et al.*, 1988) and in the closure of the two domains (the N-lobe of rSTf is more closed than the C-lobe by about 2°, in contrast to the situation in hLf). As the hinge region of the N-lobe is in the general vicinity of the inter-lobe contact, it is also possible that differences in lobe orientations (and hence the interlobe contacts) could alter the stabilities of the "open" and "closed" forms of each lobe and thus affect the metal and anion binding properties of rSTf with respect to hLf.

One difference in the interdomain cleft of rSTf (and cOTf for that matter) is the absence of a serine at position 191 (lactoferrin numbering), which in lactoferrin hydrogen bonds to the sidechain of Arg121. In the structure of $FeSTf_N$, the sidechain of the arginine is hydrogen bonded to solvent molecules only (Sarra *et al.*, 1990), and may be more mobile than its counterpart in hLf. It is conceivable, therefore, that this arginine could move aside (as in the C-lobe of $Cu_2$oxLf) to accommodate an oxalate, and that this movement of the sidechain may be somewhat more favourable than in the case of hLf.

However, this does not explain why hTf does not bind oxalate at all in the C-terminal site when $Cu^{2+}$ is the associated metal ion. One interesting difference between the C-lobe sites of $Fe_2$Lf and $Fe_2$rSTf is that the arginine in rSTf (Arg455), equivalent to Arg465 in hLf is hydrogen bonded to Tyr512, which is the equivalent of Tyr526 in hLf. Modelling the movement of the arginine in the C-terminal site of rSTf causes this hydrogen bond to be broken. Whether or not this would be unfavourable is difficult to judge. The tyrosine sidechain makes no other contacts and it is readily able to move in accord with the arginine to maintain the hydrogen bonded contact. Similarly, the equivalent of Tyr398 (Tyr395 in rSTf) does not appear to be restricted in movement and could easily be pushed aside were the arginine to move.

Clearly, subtle differences in the hydrogen bonding and the solvent structure near the arginine could play an important role in determining the extent to which the sidechain can move, although the analysis of the $Fe_2$rSTf model does not reveal any structural reason why a movement similar to that seen in $Cu_2$oxLf is not possible.

# CHAPTER 4

## Spectroscopic studies

### 4.1 Introduction

4.1.1 *Spectroscopic studies on metal and anion binding:*

The substitution of metal ions as spectroscopic probes in the specific binding sites of the transferrins has produced some very important information relating to the nature of both the metal and anion binding sites. Among the most useful metal ions are $Cu^{2+}$, $Cr^{3+}$, $Mn^{3+}$, $Co^{3+}$ (Ainscough *et al.*, 1979) and $VO^{2+}$ (Campbell & Chasteen, 1977), notwithstanding $Fe^{3+}$ itself. The use of spectroscopic probes such as these, for example, have resulted in the identification of some of the ligands responsible for metal coordination (Section 1.2.3), estimation of the equilibrium binding constants for a series of metal ions (see Section 4.1.2) and identification of a number of possible synergistic anions other than carbonate (Section 2.2). A number of different methods have been employed such as electronic absorption (including fluorescence and circular dichroism) and Raman spectroscopy, resonance techniques (ESR and NMR), kinetic and calorimetric studies (described in Aisen & Harris, 1989) and more recently, X-ray absorption spectroscopy (Schneider *et al.*, 1984; Garratt *et al.*, 1986; 1992).

The techniques employed in this study are UV and visible absorption and difference spectroscopy, fluorescence and CD spectroscopy, and ESR.

*Electronic absorption:* The $Fe^{3+}$, $Cu^{2+}$, $Co^{3+}$, $Mn^{3+}$ and $Cr^{3+}$ complexes with cOTf, hTf and Lf have been extensively studied by electronic absorption spectroscopy (eg. Warner & Weber, 1953; Aisen *et al.*, 1969; Feeney & Komatsu, 1966; Tomimatsu *et al.*, 1976; Ainscough *et al.*, 1979) and their visible spectra are dominated by moderate to intense transitions in the 400 - 500 nm region (Table 4.1 and also Table 1 in Appendix I).

The bands given by $Fe^{3+}$ and $Cu^{2+}$ complexes can be attributed to ligand-metal charge transfer (LMCT) transitions which Gaber *et al.* (1974) have assigned as phenolate $\rightarrow$ metal ($p\pi \rightarrow d\pi^*$ for iron and $p\pi \rightarrow d\sigma^*$ for copper). Similar assignments have been made for the bands observed in the $Co^{3+}$ and $Mn^{3+}$ complexes (Ainscough *et al.*, 1979). These bands

result in intensely coloured solutions (Table 4.1) with extinction coefficients (ε) ranging from 4000 ($Fe^{3+}$) to 10000 ($Co^{3+}$ & $Mn^{3+}$) $M^{-1}$ $cm^{-1}$/metal ion. The chromium complexes however, exhibit two weak spin-allowed *d-d* transitions and the colour of the resulting solutions are somewhat less intense with ε values between 200 and 400 (Ainscough *et al.*, 1980). The spectra for these five transition metal complexes with human lactoferrin are shown in Figure 4.1.

Table 4.1: Visible absorption maxima for some first row transition metal complexes with human serum transferrin, ovotransferrin and human lactoferrin

| | position of visible $\lambda_{max}$ (nm) | | | |
|---|---|---|---|---|
| metal ion | hTf | cOTf | Lf | colour |
| $Fe^{3+}$ | 470 [a] | 462 [b] | 465 [c] | red |
| $Cu^{2+}$ | 440 [a] | 440 [b] | 438 [d] | yellow |
| $Cr^{3+}$ | 440 [e] | 435 [f] | 442 [d] | blue/grey |
| | 615 | 610 | 612 | |
| $Mn^{3+}$ | 430 [e] | 430 [g] | 435 [d] | brown |
| $Co^{3+}$ | 405 [e] | 415 [g] | 405 [d] | brown/yellow |

[a] Schade & Reinhart (1966). [b] Warner & Weber (1953). [c] Plowman (1979).
[d] Ainscough *et al.* (1979). [e] Aisen *et al.* (1969). [f] Tan & Woodwoth (1969).
[g] Feeney & Komatsu (1966)

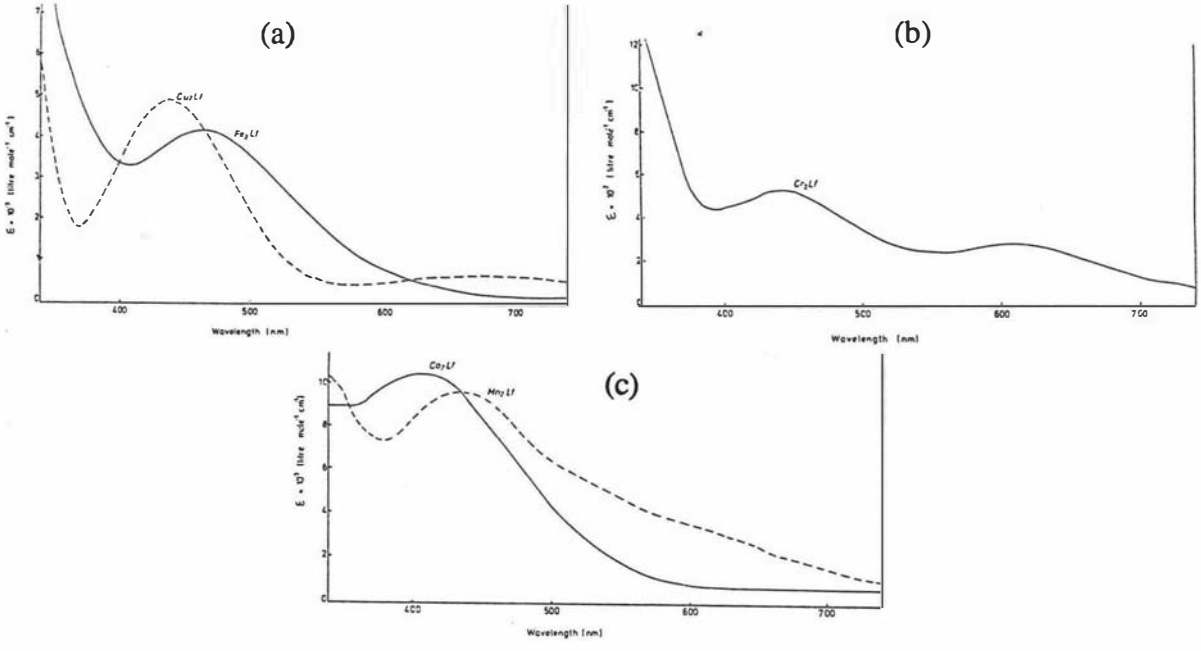



Figure 4.1: Visible absorption spectra of the a) $Fe^{3+}$ & $Cu^{2+}$; b) $Cr^{3+}$ and c) $Mn^{3+}$ & $Co^{3+}$ complexes of human lactoferrin (from Ainscough *et al.*, 1979).

The apo forms of the transferrins exhibit the very intense absorption band in the ultraviolet centred at about 280 nm, typical of all protein solutions, resulting from $\pi \rightarrow \pi^*$ transitions in the aromatic rings of tryptophan, tyrosine and phenylalanine residues. In addition, the binding of metal ions to the transferrins results in the appearance of two new bands in the UV region, typically near 240 nm and 295 nm, attributed to a deprotonation of the tyrosines and a resulting increase in the intensity of these $\pi \rightarrow \pi^*$ transitions (Wetlaufer, 1962). These bands are readily observed in a difference spectrum, obtained by using apo-protein as the reference and adding the metal ions to a second apo-protein solution, and have been extremely useful in determining the stoichiometry of metal binding (Luk, 1971; Teuwissen *et al.*, 1972; Pecoraro *et al.*, 1981; Harris *et al.*, 1981).

*Electron spin resonance:* This technique has proved invaluable in the characterisation of the metal-binding sites in the transferrins and the analysis of the effects of anion binding (for recent reviews see Chasteen, 1981, 1983 and Aisen & Harris, 1989).

ESR spectra are only observed when the system being studied contains atoms with unpaired electrons. Transition metals such as $Fe^{3+}$ ($d^5$), $Cu^{2+}$ ($d^9$), $Cr^{3+}$ ($d^3$) and $VO^{2+}$ ($d^1$) exhibit ESR spectra at liquid nitrogen temperatures when coordinated to the transferrins. In addition, resonance is also observed for the lanthanide ions $Eu^{2+}$ and $Gd^{3+}$ (both $f^7$).

The iron(III)-transferrin complexes have been extensively studied by ESR spectroscopy, although the spectra remain the most difficult to interpret due to the presence of five electrons ($d^5$, high spin). Figure 4.2 represents the ESR spectrum of a frozen solution of diferric lactoferrin and shows a strong line near 1500 gauss ($g' \approx 4.3$). A weaker resonance at 700 gauss ($g' \approx 9.7$) is also observed.

The ESR spectra of copper-substituted transferrins are typical of most $Cu^{2+}$ complexes, indicating that the $Cu^{2+}$ is in an axially symmetric field. As $Cu^{2+}$ is a $d^9$ species, there is only one free electron and the spectra observed are relatively simple to interpret with the signals near the free electron g-value of 2.0023. Two orientation-dependent resonances are observed, labelled $g_{\parallel}$ and $g_{\perp}$ in Figure 4.3.

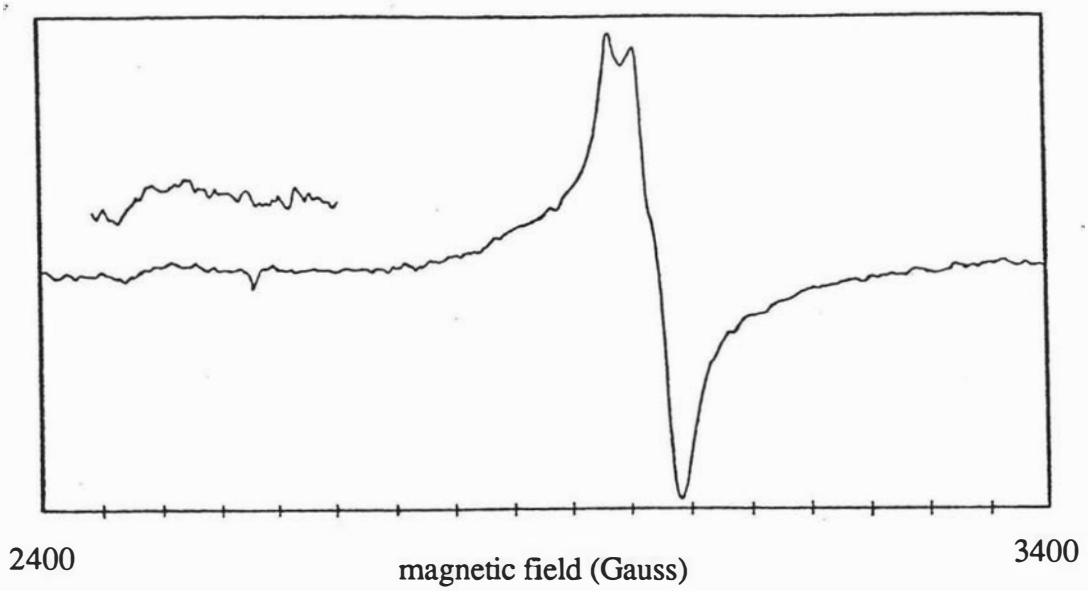



Figure 4.2: ESR spectrum of diferric human lactoferrin in 0.01M phosphate buffer, pH 7.6 at -160°C (Plowman, 1979).

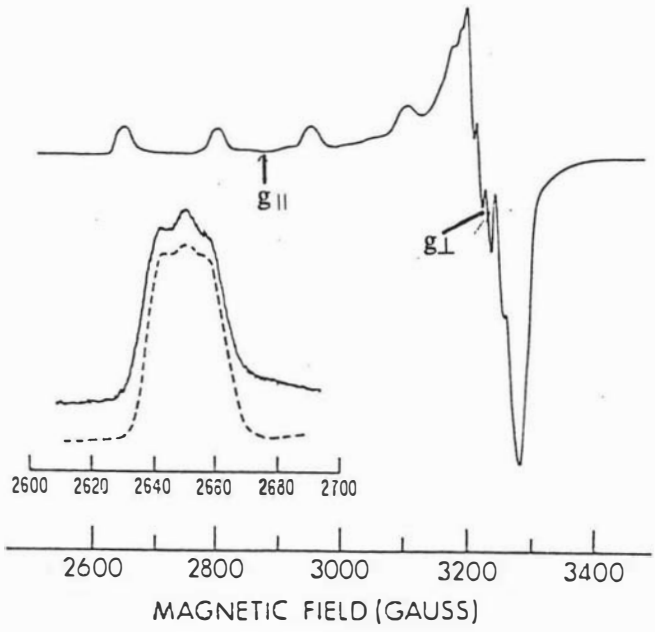



Figure 4.3: ESR spectrum of dicupric serum transferrin in 0.05 M Hepes/0.1 M KCl, pH 7.6. The positions of the $g_{||}$ and $g_{\perp}$ are shown, their values being 2.312 and 2.062 respectively. The hyperfine splitting constant A = 150 gauss and the $^{14}N$ superhyperfine splitting is ≈ 9 gauss (from Zweier & Aisen, 1977).

Isotopically-pure $^{63}Cu$ or $^{65}Cu$ are generally used to obtain sharp, intense copper-transferrin ESR and hyperfine splitting of the $g_{||}$ and $g_{\perp}$ resonances occurs due to coupling of

the spin of the unpaired electron with the nuclear moment of the copper nucleus. Both copper isotopes have nuclear moments (I) of 3/2, resulting in 4 hyperfine signals (Figure 4.3). In addition, $^{14}N$ superhyperfine splitting has been observed, superimposed on the lowest field hyperfine signal, and has been attributed to the presence of at least one nitrogen ligand in the copper coordination sphere (Zweier & Aisen, 1977).

Vanadium(IV), a $d^1$ metal ion, provides another useful probe into the metal binding sites of the transferrins. As for $Cu^{2+}$, the resonances of the unpaired electron are centred near the free-electron g value and the metal ion displays an almost axial symmetry. Naturally-occurring vanadium exists as the $^{51}V$ isotope (I = 7/2) in virtually 100% abundance. This leads to the splitting of both the $g_{||}$ and $g_{\perp}$ resonances into 8 hyperfine lines (Figure 4.4). Chasteen (1981) describes the vanadyl ESR spectra of a number of biological molecules in detail, including the members of the transferrin family. Vanadyl ESR has also been used to study the association of vanadium with haemoglobin (Cantley & Aisen, 1979) and ferritin (Chasteen & Theil, 1982).

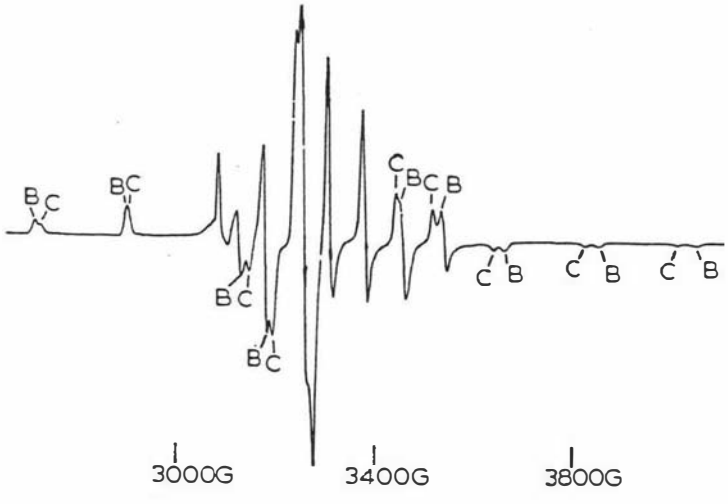



Figure 4.4: ESR spectrum of 0.3 mM vanadyl(IV)-transferrin in 0.1 M Hepes/0.025 M $NaHCO_3$, pH 7.5 at 77K. Microwave power and frequency, 5 mW and 9.112 GHz respectively (from Chasteen *et al.*, 1986a).

Gadolinium(III)-transferrin ESR spectra have been reported by two independent groups and although the spectra they observe are very similar, the results derived from them seem to

differ markedly (O'Hara & Koenig, 1986; Zak & Aisen, 1988). The spectra exhibit features ranging from g' = 13 to g' = 2, with the most prominent resonance at g' = 4.1 (Figure 4.5). This value for g' indicates deviation from the cubic symmetry usually associated with the $^8S_{7/2}$ ground state of $Gd^{3+}$ and is indicative of a rhombic distortion (O'Hara & Koenig, 1986), as in the case of $Fe^{3+}$ and $Cr^{3+}$ (Aisen *et al.*, 1969).

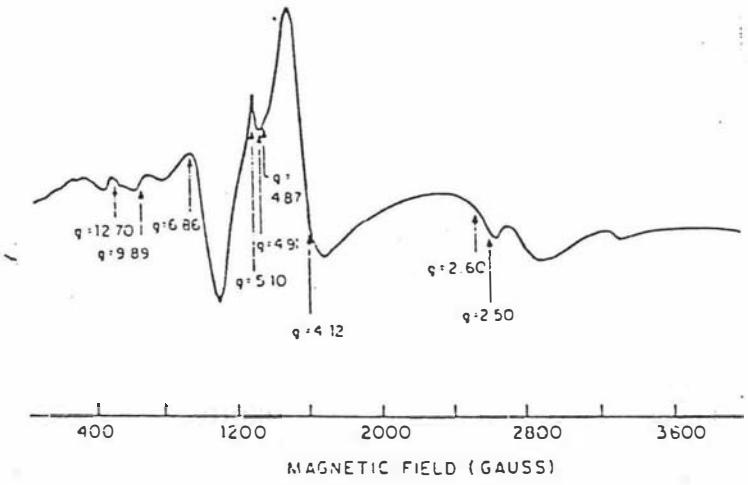



Figure 4.5 ESR spectrum of 5 mM gadolinium(III)-transferrin in 0.1 M Tris-HCL/0.02 M $NaHCO_3$, pH 8.5 at 77K. Microwave power and frequency, 10 mW and 9.5 GHz respectively. The major transition at g' = 4.12 is indicated (from O'Hara & Koenig, 1986).

*Circular dichroism:* The optical activity of the transferrins has been investigated using the technique of circular dichroism (CD) by a number of researchers (Nagy & Lehrer, 1972, Tomimatsu & Vickery, 1972, Brown & Parry, 1974, Prados *et al.*, 1975, Mazurier *et al.*, 1976). The CD spectra are measured in three distinct regions, due to the differences in intensity arising from different optically active chromophores. The 200 - 250 nm region (the far UV) generally shows one negative CD band arising from $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions of the polypeptide chain. Estimates of the amount of helical structure in a protein can be made from analysis of the shape of this band, with values of about 21% being obtained for hTf and 15% for cOTf (Tomimatsu & Vickery, 1972) and 16% for bLf (Brown & Parry, 1974). The value for hTf agrees well with the value of 24% obtained by the crystallographic

analysis of human lactoferrin (Anderson *et al.*, 1989), although the corresponding value for serum transferrin has not yet been reported.

In the region 250 to 350 nm, bands are observed due to the aromatic residues (tryptophan, tyrosine and phenylalanine) and tend to differ between various proteins. However, all members of the transferrin family display similar CD spectra in this region, although differences in intensity and band position have been observed upon the addition of metal ions to the proteins (Mazurier *et al.*, 1976).

The CD spectra of the metal-substituted transferrins in the visible region (350 - 800 nm) show features similar to their visible absorption spectra. Diferric lactoferrin for example, has one broad negative band centred at about 450 - 460 nm, which can be assigned as the phenolate → iron charge transfer band (see above).

*Fluorescence:* Emission from excited tryptophan residues dominates the fluorescence of the transferrins and energy which has been absorbed by other groups (tyrosine and phenylalanine) is usually transferred to the tryptophans before emission (Aisen & Harris, 1989). The presence of a transition metal or lanthanide ion causes some of this energy to be diverted into a metal-phenolate absorption band and lost as heat, resulting in a decrease in the amount of energy emitted (Lehrer, 1969). This phenomenon is known as quenching.

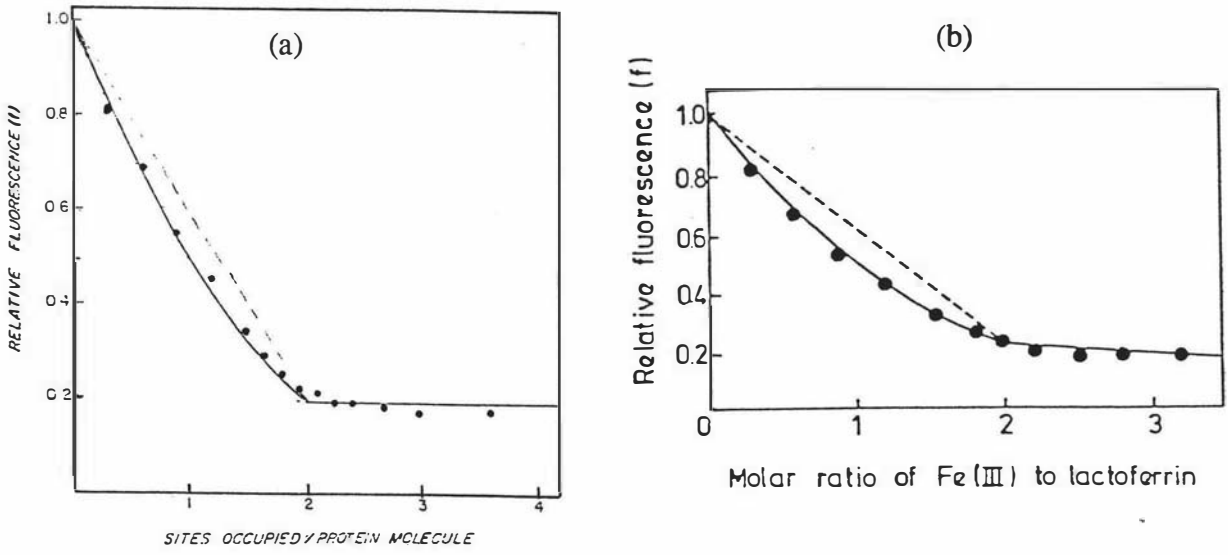



Figure 4.6: Fluoresence titrations of a) apotransferrin (Lehrer, 1969) and b) apolactoferrin (Ainscough *et al.*, 1980) with $Fe^{3+}$ (as FeNTA). Protein concentrations were $3{\cdot}57 \times 10^{-5}$ M and $1{\cdot}25 \times 10^{-5}$ M for hTf and Lf respectively.

The binding of metal ions can be investigated by monitoring the decrease in protein fluorescence as a function of the increasing metal/protein ratio (similar to that described using difference absorption spectroscopy). It has been observed that the fluorescence is quenched in a nonlinear fashion by the addition of $Fe^{3+}$ (Lehrer, 1969, Ainscough *et al.*, 1980) and that the slope of the titration curve is greater in the earlier half than in the later part (Figure 4.6a & b, previous page).

Fluorescence spectroscopy has also been employed to estimate the distance between the two metal binding sites in transferrin (Luk, 1971; O'Hara *et al.*, 1981) and to investigate the stability of the iron-ovotransferrin complex with respect to temperature and pH changes (Tan & Woodworth, 1970)

### 4.1.2 *Equilibrium binding constants:*

Warner and Weber (1953) were the first to address the question of how tightly the transferrins bound metal ions. They proposed a mechanism to account for observed dissociation curves for the $Fe^{3+}$ and $Cu^{2+}$ complexes of ovotransferrin in the presence of citrate and derived some equilibrium constants. Further studies by Aisen and colleagues and more recently by Harris have greatly extended this work to cover a wide range of metal ions including transition metals, group 13 metals and the lanthanides (see below).

*Equilibrium dialysis:* Aisen *et al.* (1978) have calculated conditional equilibrium binding constants for $Fe^{3+}$ using equilibrium dialysis methods similar to those used previously by Aasa *et al.* (1963). They obtained binding constants, $K^*_1 = 20.67$ and $K^*_2 = 19.38$ at pH 7.4 (ambient $CO_2$, $3.6 \times 10^{-4}$ atm, 0.14 mM $HCO_3^-$).

Because transferrin is a two-sited protein, it follows that there will be four specific site constants for the binding of metal ions to transferrin, depending on which site is occupied and which is vacant. Assuming the model depicted in Figure 4.7, it was found that at pH 6.7, the association constant for one site is much greater than that for the other and that binding would be virtually sequential, in that the first equivalent goes into one site only

(Aisen *et al.*, 1978). At physiological pH (7.4) however, this difference between sites is less marked, although the C-terminal site still has a higher affinity than the N-site. The authors maintain that the two sites are not equivalent or independent.

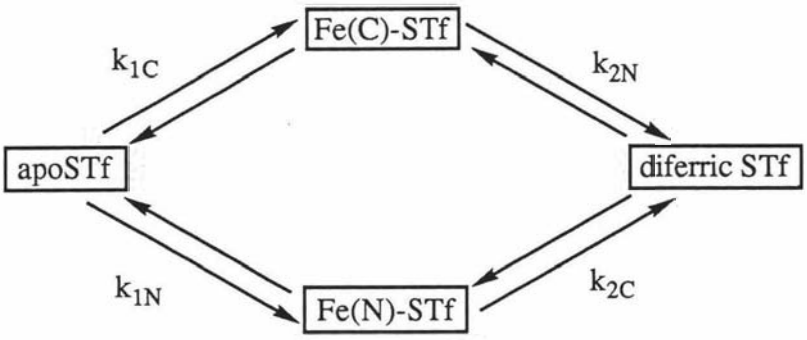



Figure 4.7: Binding scheme devised by Aisen *et al.* (1978). The subscripts C and N denote binding to the C- and N-terminal binding sites, so that $k_{1C}$ is the binding constant associated with the C site when the N site is empty and $k_{2C}$ is the C site constant when the N site is occupied. The values at pH 7.4 are: $k_{1C} = 4 \times 10^{20}$; $k_{1N} = 6.8 \times 10^{19}$; $k_{2C} = 1.6 \times 10^{20}$ and $k_{2N} = 2.8 \times 10^{19}$.

*UV difference spectroscopy:* Much of the recent development of methods for the calculation of transferrin binding constants for a variety of metal ions is due to Harris and co-workers. Harris & Pecoraro (1983) undertook a UV difference spectroscopic study of the complexation of gallium to transferrin, using both NTA and EDDA as competitive chelators to determine conditional binding constants for the two $Ga^{3+}$ ions ($K^*_1 = 19.53$ and $K^*_2 = 18.58$). By assuming that the binding constant would increase linearly with bicarbonate concentration, Harris and Pecoraro (1983) went on to calculate $\log K^*_1$ and $\log K^*_2$ values at $[HCO_3^-] = 0.14$mM and 27 mM. This assumption was later found to be valid only at low levels of bicarbonate (Harris, 1986a) so the $\log K^*$ values were corrected to "absolute" $\log K'$ values for gallium and iron binding to hTf. These values, independent of pH and $[HCO_3^-]$ are given at the bottom of Table 3, Appendix I. $\log K^*$ values for gallium binding to human lactoferrin were also estimated to compare with those for hTf (Harris, 1986a).

Harris (1983) also observed that the C-terminal site appeared to have a greater affinity for added metal ions in his study of the binding of $Zn^{2+}$ to monoferric transferrins. This has also been noted for $Nd^{3+}$ & $Sm^{3+}$ (Harris, 1986b), $Cd^{2+}$ (Harris & Madsen, 1988) and $Al^{3+}$

(Harris & Sheldon, 1990) binding to transferrin. It seems reasonable to assume, therefore, that the $K_1$ values are associated with the C-terminal site with the $K_2$ values corresponding to the N-terminal site.

A third method used to calculate binding constants was employed to estimate the value of $logK^*$ for the binding of the second equivalent of vanadate to hTf (Harris & Carrano, 1984) and rat transferrin (Chasteen *et al.*, 1986b). This involves titrating the protein in question with vanadate solution and plotting the average number of metal ions bound/protein molecule, $\bar{n}$, against the ratio of moles of metal added/moles of protein present, r. It is assumed that the first equivalent of metal binds quantitatively (the slope of the $\bar{n}$ versus r plot equals 1), then the curvature in the graph between r = 1 and 2 can be interpreted in terms of weaker binding to the second site (see Section 4.2.8 for a more detailed discussion of this method). The $logK^*$ values obtained with this technique compare favourably for hTf and rat Tf (6.5 and 6.36 respectively). They also agree remarkably well with a $logK^*_2$ value of 6.6 reported by Harris (1985) for the binding of vanadate to hTf, estimated from UV difference titrations in the presence of inorganic anions as perturbants. Whether or not this third technique results in values which are a true representation of the actual equilibria involved cannot be gauged from only one comparison. The assumption that has to be made concerning the quantitative binding of the first equivalent of metal may not be valid, as there is generally only a difference of about 1 to 1.5 log units between $K^*_1$ and $K^*_2$ for most of the metal-transferrin complexes which have been investigated, meaning that the binding of the first metal ion is only around 10 - 30 times stronger than the second. It remains however, a quick and relatively easy method for the estimation of binding constants and could be invaluable for the determination of $logK^*_2$ values for a series of metal ions where relative values can be just as important as absolute ones.

*Linear free energy relationships:* Conditional binding constants for the complexation of gallium can be directly compared to values for iron by the construction of a linear free energy relationships (LFER) for iron and gallium binding. LogK values for the binding of gallium to a series of small ligands are plotted against similar logK values for iron complexation. The

plots give a straight line from which an equation relating $\log K_{Fe}$ to $\log K_{Ga}$ can be derived. Hence if $\log K_{Ga}$ for a ligand is known, then $\log K_{Fe}$ for the same ligand can be calculated. This technique, only valid if the two metal ions in question have similar stereochemical preferences, has been employed in the evaluation of a number of logK's for metal complexation to the transferrins, such as $Fe^{3+}$ binding to hTf (Harris & Pecoraro, 1983) and Lf (Harris, 1986a), along with values for $Am^{3+}$ and $Cm^{3+}$ from LFER's with $Nd^{3+}$ and $Sm^{3+}$ (Harris, 1986b), and $Fe^{2+}$ from an $Fe^{2+}/Ni^{2+}$ LFER (Harris, 1986c). LogK* values for a number of other metal ions are given in Table 3, Appendix I.

## 4.2 Experimental Procedures

### 4.2.1 *Glassware and dialysis tubing:*

Glassware was washed with Pyroneg or DECON 50 and rinsed thoroughly with deionised water. It was subsequently treated with AR concentrated nitric acid to remove traces of metal ion contaminants, then rinsed several times with distilled deionised water. Cuvettes were washed with conc. $HNO_3$, rinsed with water and dried in an oven at about 100°C before use. ESR tubes were carefully washed with a mixture of concentrated $HNO_3$ and ethanol, then rinsed with water and oven-dried.

Dialysis tubing as obtained from the manufacturer was heated to 80°C in a solution containing 0.6M $NaHCO_3$ and 1mM EDTA (disodium salt) and held at this temperature for about 10 minutes. This process was repeated and the tubing was then rinsed in four changes of d.d. water. Finally the tubing was boiled in two changes of d.d. water and stored in a 50:50 ethanol:water solution at 4°C.

### 4.2.2 *Buffers:*

*Tris-HCl buffer:* This was prepared by dissolving the appropriate amount of tris(hydroxymethyl)aminoethane (SERVA Feinbiochemica) in d.d. water to give concentrations of 0.01 to 0.05M and adjusting to pH 7.7 - 8.2 with AR hydrochloric acid. $NaHCO_3$, and NaCl (Baker) were added to the Tris solution before titrating with the HCl.

*Phosphate buffer, 0.01M:* Prepared as two solutions - 0.01M $Na_2HPO_4 \cdot 2H_2O$ and 0.01M $NaH_2PO_4$ (Reidel-deHaen). The $HPO_4^{2-}$ solution was titrated with $H_2PO_4^-$ until the desired pH was reached.

*Acetate buffer, 0.2M, Ph 6.1:* This buffer was made by titrating a 0.2M solution of sodium acetate (BDH) with 0.2M AR acetic acid. The 0.2M acetic acid solution was prepared by dropwise addition of AR glacial acetic acid to d.d. water to a pH of 2.78 (based on $K_a$ acetic acid = $1.78 \times 10^{-5}$M).

*HEPES buffer, 0.01M:* Prepared by dissolving N-2-hydroxyethylpiperazine-N'-2-ethane sulphonic acid (U.S. Biochemical Corp.) in distilled water and adjusting to pH 7.4 - 7.8 with concentrated HCl.

4.2.3 *Preparation of metal ion solutions:*

*Transition metals:* Stock solutions of these ions were prepared from AR grade $NH_4VO_3$ (Reidel-deHaen), $VOSO_4$, $VCl_3$, $MnCl_2$ (BDH), $Fe(NO_3)_3 \cdot 9H_2O$ (Reidel-deHaen), $CoCl_2$ (Koch-Light), $CuCl_2 \cdot 2H_2O$ (BDH) and $Cu(NO_3)_2 \cdot 2½H_2O$ (Ajax, Australia) which were low in other metal ion impurities. Appropriate amounts of these compounds were dissolved in d.d. water to give solutions with concentrations around 0.01M. The concentrations of all the solutions, except vanadium, were determined by direct titration with EDTA using the indicators variamine blue B for Fe(III), catechol violet for Cu(II), xylenol orange for Co(II) and Eriochrome black T for Mn(II). (Vogel, 1961). The copper, cobalt and manganese solutions were added to the lactoferrin as prepared, but in the case of the ferric ion solution, one equivalent of NTA (disodium salt) was added to give a 1:1 Fe-NTA complex. The addition of $Fe^{3+}$ directly to transferrin has been shown to give low yields and complicated side reactions, probably due to the rapid hydrolysis of the ferric ion at physiological pH (Bates and Schlabach, 1973).

Standard EDTA solutions were made by dissolving an accurately weighed sample of the disodium salt (AR grade) in one litre of water to give a solution of concentration between 0.008 and 0.01M. The EDTA had previously been heated to around 80°C to quantitatively

remove any water. Once this precaution was taken, the EDTA could be used as a primary standard (Vogel, 1961).

The concentrations of vanadium (V) solutions were determined by reduction of vanadate to vanadyl(IV) by ferrous ammonium sulphate, followed by oxidation back to vanadate by $KMnO_4$, previously standardised with oxalate. The excess Fe(II) was destroyed by persulphate prior to the addition of $KMnO_4$. Persulphate ions react rapidly with $Fe^{2+}$ ions but only slowly with vanadyl, so that if the $KMnO_4$ titration is completed quickly, the loss of vanadyl is negligible. The vanadyl determinations were performed either by direct titration with $KMnO_4$ or by measuring the visible absorption at 750nm and using the extinction coefficient of 18.0 to calculate the concentration (Fitzgerald and Chasteen, 1974). The two methods agreed to within 0.5%. Vanadium(III) solutions could not be made up in air, as the $V^{3+}$ ion is very unstable with respect to oxidation. A sample of $VCl_3$ was transferred into a pre-weighed, nitrogen-flushed 20 ml Schlenk tube (Shriver, 1969) inside a nitrogen-filled glove bag. The sealed tube was weighed and the $VCl_3$ dissolved in a known volume of degassed d.d. water injected through a rubber septum. The final concentration was determined by titration of aliquots of V(III) with $KMnO_4$ under a steady stream of nitrogen gas to avoid air oxidation of the $V^{3+}$ ions.

*Lanthanide metals:* Standard stock solutions (0.01 - 0.02M) for the lanthanide ions, $La^{3+}$ through to $Lu^{3+}$ were prepared either by dissolving the appropriate metal nitrate in d.d. water or by dissolving the oxides in concentrated HCl. In the latter case, a cloudy suspension resulted to which about 2-3 ml of water was added. The suspension was then heated to boiling and titrated with HCl until a clear solution was obtained. Once the solution had cooled, it was made up to the appropriate volume with d.d. water. The final concentrations were determined by direct titration in acetate buffer (pH 6.1) with EDTA using xylenol orange indicator (Lyle and Rahman, 1963).

The pH of the solutions made up from the oxides were found to be between 1.5 and 3. The solutions made up by directly dissolving the metal nitrates in water were adjusted to pH 2 with concentrated $HNO_3$. Solutions of the lanthanides can be stored at this pH for a

considerable length of time as they are safe from hydrolysis and do not absorb $CO_2$ which could cause precipitation as the insoluble carbonates.

*Other metals:* Thorium(IV) solutions were prepared from spectroscopically pure $Th(NO_3)_4 \cdot xH_2O$ (Strem) by dissolving the solid in water. The concentrations were determined by direct titration at pH 3 with EDTA using pyrocatechol violet as the indicator (Vogel, 1961).

The aluminium solutions were made from either $AlCl_3 \cdot xH_2O$ or $(NH_4)Al(SO_4)_2 \cdot xH_2O$ and were standardised by back titration of excess EDTA with previously standardised $ZnCl_2$.

4.2.4 *Preparation and Purification of apolactoferrin:*

Lactoferrin was prepared from fresh, chilled colostrum, although frozen colostrum was used on occasion (the freeze/thaw treatment did not appear to have any detrimental effects on the metal binding properties of the protein).

Apolactoferrin was prepared following the method of Norris *et al.* (1989). One drop of β-mercaptoethanol was added to the colostrum (typically 100-500ml) and, in order to precipitate the immunoglobulins and caseins, AR $(NH_4)SO_4$ was slowly added to bring the solution to 2M. The resulting suspension was left to stir for one hour, then centrifuged at 12,000 r.p.m. for 20 minutes. Fat collected at the top of the tubes, with the precipitated caseins at the bottom. The supernatant between contained the crude lactoferrin and this was drawn off into more centrifuge tubes and spun for another 20 minutes at 12,000 r.p.m. The crude lactoferrin was pooled and dialysed for 3 hours against two changes of 5 litres of d.d. water containing 2 drops of β-mercaptoethanol. Finally, the crude protein solution was dialysed against two changes of 0.025M Tris-HCl/0.2M NaCl, pH 7.8. In some preparations, 5mM $Na_2EDTA$ was added to the second change of water and to the buffer to complex any metal ions which may have been present in the solutions (see Section 4.2.7).

A column was prepared by swelling about 8 gram of CM-50 Sephadex in d.d. water, degassing it under light vacuum and loading it into an acid-washed glass column (30 cm x 4 cm). The column was equilibrated with 0.025M Tris-HCl/0.2M NaCl, pH7.8 and the crude

protein solution loaded and washed thoroughly with several volumes of the buffer until the effluent showed no absorption at 280 nm. Lactoferrin bound tightly in a thin pale pink band at the top of the column (the colour was due to the residual iron content of the protein).

The lactoferrin was eluted with a 0.2-1.0M linear NaCl gradient, with the eluent being collected as sixty to seventy 10ml fractions with a Pharmacia fraction collector. Lactoferrin generally eluted as a single peak, although a smaller peak could sometimes be observed eluting just prior to the major band (Plowman, 1979). A typical elution curve is shown in Figure 4.8. The column was reswollen with 0.025M Tris-HCl/0.2M NaCl solution to be used again.

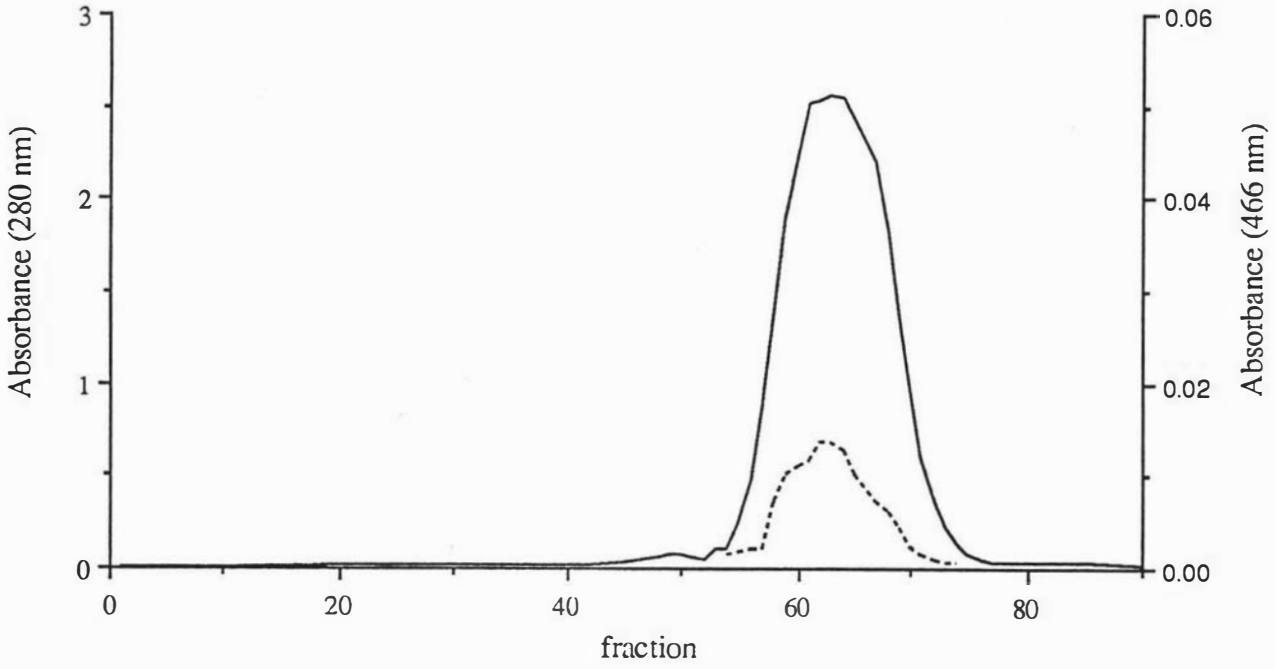



Figure 4.8: Apolactoferrin elution profile from a Sephadex CM-50 column using a 0.2 - 1.0 M NaCl gradient. The concentration of apolactoferrin (——) in the fractions was determined from the absorbance at 280 nm using a value of 10.9 for the absorbance of a 1% solution. The absorbance due to lactoferrin containing $Fe^{3+}$ (- - - -) was measured at 466 nm.

The concentration and iron saturation of the lactoferrin in each tube over the major peak were estimated from the absorbances at 280 and 466 nm using extinction coefficients for a 1% (10 mg/ml) solution of 10.9 and 0.51 respectively (Plowman, 1979). Iron saturation was typically below 10%. The purity of the isolated protein was checked by gel electrophoresis performed in denaturing conditions.

*Gel electrophoresis:* SDS/polyacrylamide gels were prepared according to Laemilli (1970), using a 4% stacking gel overlaid onto a 7% resolving gel (Norris *et al.*, 1989). The gel solutions, once mixed, were degassed under light vacuum and polymerised with ammonium persulphate.

Samples were taken from tubes covering the entire lactoferrin peak. From the tubes at the top of the peak, 10µl of the protein were mixed with 20µl d.d. water and 30µl of a solution containing bromothymol blue in plastic Eppendorf reaction vials. Thirty microlitres were taken from tubes on both sides of the peak and added to 30µl of indicator. The samples were denatured by heating to 100°C for 3 minutes, then spun at 6,000 r.p.m. for 1 minute.

Protein solution from both sides of the elution peak was loaded into the wells on the gel as 7 µl aliquots, while only 2 µl of solution from the central portion of the peak was injected. The reservoir on the electrophoresis apparatus was filled with a Tris/glycine buffer and a current of 10 mA applied for about 15 minutes (or until the dye-front had penetrated the resolving gel). This was increased to 20 mA and the gel run for 2 to 3 hours until the dye-front just disappeared off the bottom of the gel.

The gels were stained with Coomassie Brilliant Blue overnight and destained with a solution containing 7% acetic acid/7% methanol in water.

### 4.2.5 *Electronic spectra:*

*Electronic absorption:* Absorption spectra in the range 200 - 820 nm were recorded with a Hewlett-Packard 8452A diode array spectrophotometer interfaced with a IBM-compatible computer. The calibration of the spectrophotometer was checked with a holmium oxide filter at regular intervals and no fluctuation in wavelength was observed. The protein samples were contained in either 1.2 ml or 3.5 ml capped quartz cuvettes with a pathlength of 1 cm. The samples were maintained at a constant temperature with a circulating water bath connected to the cuvette holder.

*Circular dichroism:* Circular dichroism (CD) spectra were recorded on a Jobin/Yvon Mark V dichrograph linked to an Apple II computer. Light from a xenon arc source passes

through a double monochromator which produces a beam polarised in the horizontal plane. The beam then encounters a birefringence modulator which splits the light into left and right circularly polarised beams. These are focussed to a thin line in the sample cavity and a detector measures the absorption of the two beams ($A_L$ and $A_R$).

The sample cavity and xenon lamp housing were continuously flushed with dry oxygen-free nitrogen in order to prevent etching of the modulator and lamp window by ozone which would be generated if oxygen were present. Samples were contained in cylindrical quartz cuvettes with pathlengths of 1 cm for the visible region and 0.1 cm for the UV and far UV regions. The spectrum of the protein sample was recorded first, the zero being set at a wavelength at which there was no observed dichroic bands. The spectrum of the buffer was recorded next using the same settings as for the protein, and subtracted from the protein spectrum.

The wavelength range was split into three regions due to the different intensities of the dichroic bands associated with the peptide bonds and secondary structure elements (far UV), the aromatic residues and disulphide bridges (UV) and the charge transfer and *d-d* bands (visible). Ellipticities were calculated using the equation

$$[\Theta] = \frac{33.M.(A_L - A_R)}{100.l.c} \qquad [4.1]$$

where M is the molecular mass of the protein (taken as 80000 for lactoferrin), $A_L - A_R$ is the measured absortion difference, $l$ is the pathlength in dm (1 cm = 0.1 dm) and c is the concentration in g/ml. The numbers 33 and 100 are included in the equation for historical reasons.

*Fluorescence:* Fluorescence emission and excitation spectra were recorded with a Aminco SPF-500 Ratio spectrofluorometer in the ratio mode. Samples were contained in 3 ml capped quartz cuvettes. Light from a 250 Watt xenon arc lamp source is dispersed by a 600 line/mm grating to produce a monochromatic excitation beam. The emitted light, monitored at right

angles to the excitation beam, is dispersed by a similar grating in the emission monochromator and detected by a photomultiplier tube.

Apolactoferrin solutions of about $10^{-6}$M were used in all the fluorescence titrations. Apolactoferrin was found to have an absorption maximum in Tris-HCl buffer at $\lambda = 290$ nm, with an associated emission spectrum consisting primarily of a sharp peak near $\lambda = 340$ nm (Ainscough *et al.*, 1980). Titrations were performed by monitoring the fluorescence of the protein at this emission wavelength as incremental amounts of the appropriate metal ion solutions were added. Metal ions used in this study were Fe(III), Cu(II) and Ce(III).

A fluorescence titration with $Fe^{3+}$ was also performed on a $10^{-6}$M apomelanotransferrin (apoMTf) solution (kindly supplied by H.M. Baker) and a $10^{-6}$M apotransferrin (Sigma) solution as a comparison.

### 4.2.6 *Ultraviolet difference spectrophotometric titrations:*

Titrations of apolactoferrin with various metal ion solutions were performed using the HP 8452A spectrophotometer. Apolactoferrin solutions were loaded into either 1.2 or 2.3 ml quartz cuvettes to which the metal ion solutions were added directly. Typically the aliquots were in the order of 2 - 10 µl, with the total volume of metal ion added not usually greater than 50 µl. The mixtures were stirred thoroughly after each addition and sufficient time was allowed for complete mixing and reaction. An equivalent volume of d.d. water was added to a second apolactoferrin solution in a matching cuvette and the difference spectrum in the UV region (230 - 350 nm) was recorded using this second sample as the reference. The peak near 240 nm was monitored for at least 10 minutes until the absorbance values remained constant.

A second type of titration was performed where it was observed that the reaction between a particular metal ion and lactoferrin was slow. A number of samples of an apolactoferrin solution were loaded into plastic Eppendorf tubes and the metal ion aliquots were added directly to these. The solutions were then left for a period of 2 days to 1 week for the reaction to proceed to completion before the UV difference spectrum of each was recorded.

4.2.7 *Electron spin resonance spectroscopy:*

Electron spin resonance spectra were recorded at -160°C on a Varian E103A spectrometer fitted with a Varian E-257 variable temperature accessory, operating at about 9 GHz and 5 - 10 mW of microwave power. Typically 300 - 500 µl of the metalloprotein solution were loaded into glass tubes with a 2 mm internal diameter. The temperature of the sample was slowly lowered by passing a stream of oxygen-free nitrogen, cooled by liquid nitrogen, into the cavity and out over the sample tube. Spectral g values were calibrated with a 1,1-diphenyl-2-picrylhydrazyl (DPPH) internal standard which gives a sharp resonance at g = 2.0036. Spectral g values for the metalloprotein resonances were calculated using the equation:

$$g_{sample} = \frac{g_{std.} \times B_{std.}}{B_{sample}} \qquad [4.2]$$

where $g_{std.}$ = 2.0036 (for DPPH)

$B_{std.}$ = the observed magnetic field for the DPPH standard

and $B_{sample}$ = the observed magnetic field for the metal ion complex in question.

Protein concentrations were typically 10 - 15 mg/ml for copper, iron and gadolinium substituted lactoferrin, while vanadium(IV) lactoferrin spectra could be recorded with concentrations near 7 mg/ml due to the intensity of the $VO^{2+}$ signal.

4.2.8 *Preparation of metal and anion substituted lactoferrin complexes:*

*Dicupric dicarbonato lactoferrin ($Cu_2Lf$):* Apolactoferrin was prepared as described in section 3.1.4. Protein samples in which EDTA had been employed in the preparation ("EDTA treated") were dialysed against 5 - 6 changes of 0.025M Tris-HCl/0.01M $HCO_3^-$ /0.2M NaCl, pH 7.8 in an attempt to remove all of the EDTA before the addition of any metal. The "non-EDTA treated" protein was dialysed against 2 changes of 0.025M Tris-HCl/0.01M $HCO_3^-$/0.2M NaCl, pH 7.8 to decrease the ionic strength of the solution. Copper was added as either $CuCl_2$ or $Cu(NO_3)_2$ and the solution changed from colourless to yellow within seconds. Formation of the complex was monitored by UV/visible difference

spectrophotometric titration and it was observed in all cases that 2 $Cu^{2+}$ ions bound per protein molecule. The complex, $Cu_2Lf$, had a $\lambda_{max}$ at about 434 nm, attributed to a phenolate-copper charge transfer transition, with a weaker *d-d* transition near 670 nm.

*Dicupric-(carbonato-oxalato) lactoferrin ($Cu_2oxLf$):* This complex could be prepared in two ways, both involving a two-step procedure. The preparative routes are shown diagrammatically below (Figure 4.9). The starting material in both cases was "EDTA treated" apolactoferrin, prepared as described in section 3.1.4. The first route followed the protocol described by Ainscough *et al.* (1983), involving the addition of 2 $Cu^{2+}$ to the apoprotein in the presence of carbonate to form $Cu_2Lf$. To this solution was added a 50-fold molar excess of sodium oxalate (50 moles of oxalate/mole of protein). The resulting complex was monitored by electronic absorption spectroscopy and a shift to lower wavelength was noted. The new $\lambda_{max}$ was near 424 nm and conversion to this new complex was complete after 24 hours. The ESR spectrum of this complex was recorded and confirmed the identity of the complex as dicupric lactoferrin containing both carbonate and oxalate.

The second route involved the removal of the carbonate from "EDTA treated" apolactoferrin by a method similar to that described by Aisen *et al.* (1969). About 5 mls of a 10 mg/ml apolactoferrin solution was placed in a Schlenk tube and the pH was lowered to 3.5 by the addition of small aliquots of 1.0 M HCl. The solution was then subjected to a light vacuum (10 - 12 mm Hg) and vigorous bubbling was observed. After evacuation for 5 minutes, the system was flushed with dry, $CO_2$-free argon gas. This evacuation/flushing procedure was repeated for about 2 hours until it was deemed that all $CO_2$ and carbonate had been removed from the solution. The pH was then carefully raised to 7.8 by the addition of small amounts of $CO_2$-free $NH_3$, prepared by bubbling ammonia gas through distilled deionised water which had been evacuated and flushed in a similar way as described for the protein solution. The pH was monitored by an Orion pH meter fitted with a long, thin probe. A positive flow of argon gas was sustained through the side-arm of the Schlenk tube and out over the probe to exclude the diffusion of air into the solution. Sufficient sodium oxalate was added to the apoLf to give a solution 0.01M in oxalate.

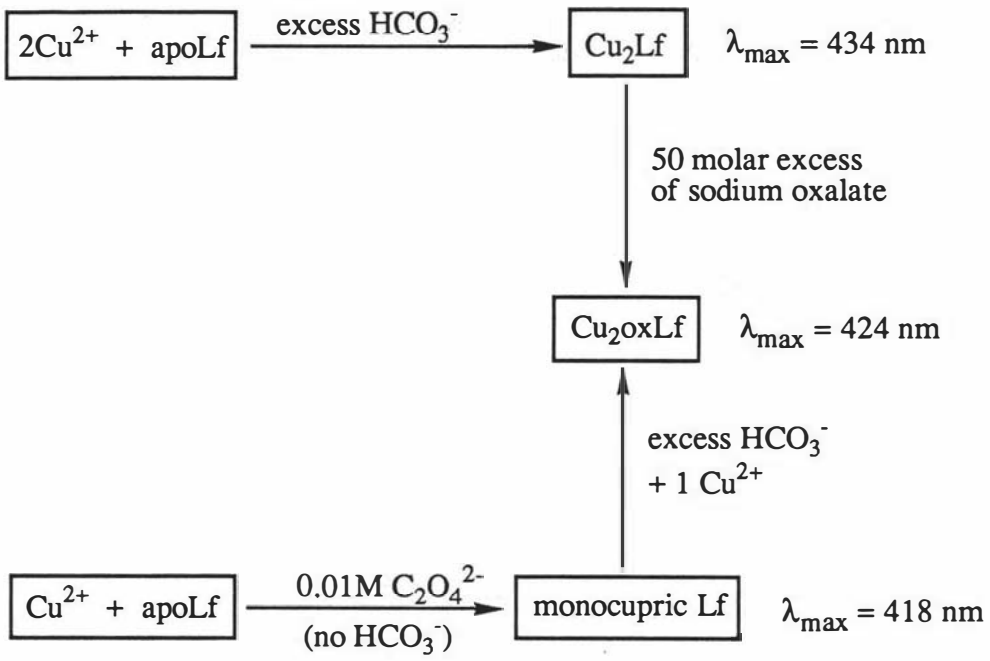



Figure 4.9: Preparation of dicupric-dicarbonato-lactoferrin ($Cu_2Lf$) and dicupric-oxalato-carbonato-lactoferrin ($Cu_2oxLf$).

A spectrophotometric titration was then performed by the addition of microlitre aliquots of $Cu^{2+}$ to 1 ml of the apoprotein in a sealed 1.2 ml quartz cuvette. It was observed that only one equivalent of copper appeared to bind under these conditions, even in the presence of excess $Cu^{2+}$. The $Cu_2oxLf$ complex was formed from the remainder of the carbonate free apoLf by adding one $Cu^{2+}$ under argon and then opening the solution to the air prior to the addition of the second equivalent of $Cu^{2+}$. The ESR spectrum of the complex formed by this second route was indistinguishable from that obtained by substitution of the carbonate by excess oxalate.

*Manganese(III) and cobalt(III) substituted lactoferrin:* These complexes were prepared as previously described by Ainscough *et al.* (1979), starting with $Mn^{2+}$ and $Co^{2+}$ solutions respectively. A sufficient amount of the appropriate metal ion solution was added to a 15mg/ml apoLf solution to fully saturate the two metal binding sites. For both $Co^{2+}$ and $Mn^{2+}$, the resulting solutions were colourless.

The manganese(II)-lactoferrin solution was allowed to stand at 4°C for a week, during which time a change from colourless to brown was observed. This has been attributed to a slow oxidation of the $Mn^{2+}$ to $Mn^{3+}$ by $O_2$ (Ainscough *et al.*, 1979).

The cobalt(II)-lactoferrin complex does not undergo spontaneous oxidation by atmospheric $O_2$, so a 5µl aliquot of a 5% $H_2O_2$ solution was added to the solution every two days over a period of about one week. A yellow/brown solution resulted.

*Aluminium-substituted lactoferrin:* Al(III) does not form a coloured complex with transferrin or ovotransferrin, although the difference absorption spectrum of the aluminium complexes reveal the usual two peaks in the ultraviolet region (Donovan & Ross, 1975, Cochran *et al.*, 1984) associated with deprotonation of the tyrosines involved in the binding sites.

Apolactoferrin was initially titrated with aluminium added as $Al^{3+}$ and the ultraviolet region monitored by difference absorption spectroscopy. The UV difference spectra that resulted proved to be different from difference spectra of transferrins substituted with metals such as gallium or holmium (Pecoraro *et al.*, 1981) in that the two peaks were very weak and an intense negative band at 260 nm was the dominant feature. No useful stoichiometric information could be gained from this difference titration and it appeared that the $Al^{3+}$ was interacting only weakly with the tyrosyl oxygen atoms. This could possibly be attributed to hydrolysis of $Al^{3+}$, as occurs with $Fe^{3+}$ when added as the free ion. It was noted that the peaks grew in intensity with increasing time, suggesting that the reaction of the $Al^{3+}$ with the protein was much slower than is the case with other metal ions.

In order to overcome this problem, the second method outlined in section 4.2.6 was employed, in which eight to 10 plastic Eppendorf tubes were each loaded with 1 ml of apolactoferrin solution and the aliquots of $Al^{3+}$ were added to these. The spectra obtained by this method closely resembled those reported for aluminium transferrin (Harris & Sheldon, 1990).

Further spectrophotometric titrations using this second method were performed adding aluminium as an Al-NTA complex. The NTA was used to maintain the aluminium in a

soluble form, as a precaution against the possible precipitation of aluminium hydroxide and aluminium carbonate. Comparative titrations with $Al^{3+}$ and Al-NTA were also performed on a metal-free human serum transferrin solution.

*Vanadium-substituted lactoferrin:* Vanadium(V)-lactoferrin solutions were prepared by adding $NH_4VO_3$ directly to an apolactoferrin solution. It has been proposed that the species which binds to the protein is the $VO_2^+$ cation (Harris & Carrano, 1984). The resulting solutions were colourless but the formation of a specific ternary complex could be monitored by UV difference spectroscopy, as previously described.

Vanadium(IV), although stable towards air oxidation in acidic solutions, is rapidly converted to the vanadate ion at physiological pH (Chasteen, 1981; Chasteen *et al.*, 1986a). In order to avoid this, $VOSO_4$ was added to apolactoferrin solution which had been thoroughly degassed by repeated evacuation and nitrogen flushing in a Schlenk tube. The protein solutions were transferred by syringe to a nitrogen-flushed 3.5 ml quartz cuvette for spectrophotometric titrations, or into an nitrogen-flushed system consisting of a 15 ml Thunberg tube joined to a quartz ESR tube for the recording of ESR spectra. In both cases, sufficient solid sodium bicarbonate was added to the apoLf solution before its removal from the Schlenk tube, as the evacuation/flushing steps remove virtually all of the carbonate which was initially present.

Vanadium(III)-lactoferrin complexes were prepared in a similar way to that described for the vanadium(IV) complex, although the metal ion solution had to be prepared under nitrogen as described in section 4.2.3.

*Lanthanide-substituted lactoferrins (and $Th^{4+}$):* The lanthanide ions were all added to apolactoferrin solutions as the aqueous ions, i.e. as, $Ce^{3+}$, $Gd^{3+}$ etc. In the first spectrophotometric titrations with the lanthanides, the apoprotein was in a Tris-HCl buffer at pH 7·8 containing 0.01M $HCO_3^-$. A white precipitate (possibly the insoluble lanthanide carbonate) was observed forming in the bottom of the cuvette. Nitrilotriacetate was employed in an attempt to maintain the metal ions in solution but the presence of this competitive chelator appeared to have a marked effect on the absorbance of the two ultraviolet

difference peaks. In order to overcome the problem of precipitation of the metal ions, the apolactoferrin solutions used for these particular titrations were dialysed in 0.025M Tris-HCl/0.05M NaCl, pH 7·4, with the only source of carbonate being atmospheric $CO_2$.

In order to quantify the effect that NTA has on the absorbance in the UV, neodymium and ytterbium solutions were prepared containing the following metal : NTA ratios:

1 : 0; 1 : 0.5; 1 : 1; 1 : 2; 1 : 3; 1 : 5

A sufficient amount of each metal ion solution was added to 1 ml samples of apolactoferrin to fully saturate one metal binding site. The solutions were left to stand at 4°C for two days to allow complete equilibration, then their UV difference spectra were recorded.

Thorium was added as the $Th^{4+}$ ion and, as noted above for the early lanthanide-Lf titrations, a white precipitate was observed in the cuvette. In subsequent titrations, bicarbonate was not added to the buffer, with atmospheric $CO_2$ the only source of the synergistic anion.

### 4.2.9 *Binding constants:*

Binding constants for the lanthanide ions were calculated using the method of Harris & Carrano (1984) which involves the spectrophotometric titration of the apoprotein with the metal ion in question, monitored at 244 nm (section 4.2.6). It is only possible to use this method when the binding constants are known to be small ( $<10^{10}$) and when the second binding constant is significantly less than the first.

The titrations were performed in tandem (section 4.2.) using 1 mg/ml apolactoferrin solutions in 0.025M Tris-HCl/0.05M NaCl, pH 7·4. A 3 ml sample of the apoprotein was loaded into a 3.5 ml quartz cuvette and 2 and 3 µl aliquots of the metal ion were added until a metal : protein ratio of 3 was reached (≈ 30µl). The UV difference spectrum was recorded after each addition. Metal ions used were $La^{3+}$, $Nd^{3+}$, $Sm^{3+}$, $Eu^{3+}$, $Tb^{3+}$, $Ho^{3+}$ and $Yb^{3+}$, with protein samples taken from the same preparation for each series of seven titrations. Results were plotted as $\Delta\varepsilon$ (244 nm) against the metal : protein ratio, r, where $\Delta\varepsilon$ is defined as the change in absorbance of the 244 nm peak divided by the total concentration of protein.

The binding constant can be calculated as follows:

1) the average number of metal ions bound per lactoferrin molecule, $\bar{n}$, is calculated for all points on the titration curve with the equation

$$\bar{n} = \frac{[\Delta\varepsilon_{obs} - (r \times \varepsilon_M)]}{(\Delta\varepsilon_1 - \varepsilon_M)} \qquad [4.3]$$

where $\Delta\varepsilon_{obs}$ = the observed absorptivity at 244 nm,
$\varepsilon_M$ = the molar absorptivity of the free metal in the buffer at 244 nm
$\Delta\varepsilon_1$ = the molar absorptivity of the 1 : 1 metal-lactoferrin complex as determined from the initial slope of the titration curve.

2) $K_{eq}$ is calculated from the equilibrium expression for the binding of the second equivalent of metal ion

$$K_{eq} = \frac{[M_2\text{-Lf}]}{[\text{M-Lf}][M^{3+}]} \qquad [4.4]$$

where $[M^{3+}]$ (free metal ion) = $(r - \bar{n})[Lf]_{tot}$ at r = 2
$[M_2\text{-Lf}] = (\bar{n} - 1)[Lf]_{tot}$ at r = 2 and
$[\text{M-Lf}] = [Lf]_{tot} - [M_2\text{-Lf}]$.

## 4.3 Results A: Transition metals and aluminium

### 4.3.1 *Electronic spectra:*

The electronic absorption spectra for the $Fe^{3+}$ complex of human lactoferrin (Figure 4.10), recorded primarily as a comparison with earlier work reported by Ainscough *et al.* (1979), exhibits a broad charge transfer band in the visible region (350 - 800 nm) due to the interaction of the $Fe^{3+}$ with tyrosine residues in the metal binding sites (Patch & Carrano, 1981). The extinction coefficient at 466 nm was 4160 (assuming a molecular weight of 80,000). This compares favourably with the results of Ainscough *et al.* (1979) (see Table 4.3). The position of the charge transfer band and the resultant colour of the complex is characteristic of the metal ion when bound specifically to lactoferrin.

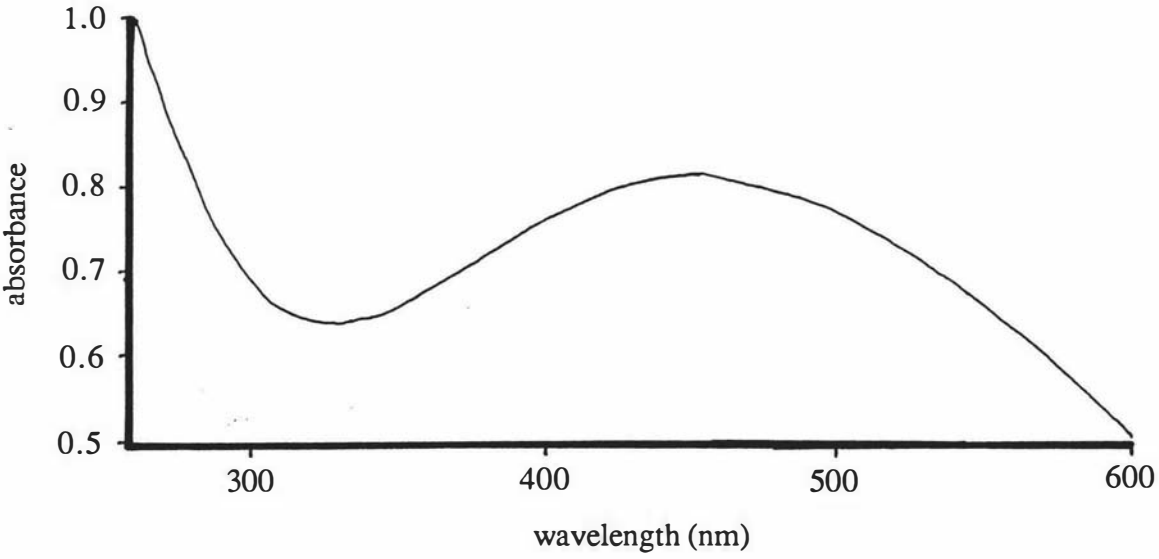



Figure 4.10: Electronic absorption spectra of $Fe_2Lf$ in the range 350 to 600 nm.

The visible absorption spectrum of dicupric lactoferrin ($Cu_2Lf$), is shown in Figure 4.11, along with the spectrum of the hybrid carbonate-oxalate complex, $Cu_2oxLf$ (Section 4.2.8). The position of the charge transfer transition for $Cu_2Lf$ is near 434 nm (the molar extinction coefficient, $\varepsilon = 4880$ L mol$^{-1}$ cm$^{-1}$), consistent with that found by Ainscough *et al.*, (1979), as is the presence of a weaker *d-d* band at about 660 nm. For the hybrid complex, $Cu_2oxLf$, the charge transfer band was shifted $\approx$ 10 nm, to 424 nm, as previously observed (Ainscough *et al.*, 1983), and there was a small shift in the position of the *d-d* transition to higher wavelength ($\approx$ 700 nm).

*Spectrophotometric titrations:* Ultraviolet difference spectrophotometric titrations were performed to estimate the stoichiometry of metal binding to Lf. It has already been established that metal ions such as $Fe^{3+}$, $Cu^{2+}$ and $Zn^{2+}$ bind to both sites of hTf and the study of Chung (1985) confirms that at least for these three metal ions the same is also true for Lf. Figures 4.12a & b show the observed difference spectra for $Fe^{3+}$ and $Cu^{2+}$ between 230 and 500 nm, while Figures 4.12c - e are the difference spectra recorded for V(III), V(IV) and V(V) respectively. It should be noted that the spectra obtained with $Fe^{3+}$ (added as Fe-NTA) and $Cu^{2+}$ are significantly different from those for the other three metal ions. This is primarily due to the high absorbance of NTA and free $Cu^{2+}$ in this region of the spectrum.

Figures 4.13a - d are representative titration curves for $Fe^{3+}$, $Cu^{2+}$, V(III) and V(IV) as plots of $\Delta\varepsilon$ (244 nm) against the metal ion:protein ratio r, where $\Delta\varepsilon$ is the absorbance divided by the analytical concentration of lactoferrin. In all cases two metals ions are bound per protein molecule, although in the case of $Fe^{3+}$ and $Cu^{2+}$, the continued increase in absorbance at 240/242 nm is the result of both the absorbance of NTA and $Cu^{2+}$ (as mentioned above) and also non-specific binding of $Fe^{3+}$ and $Cu^{2+}$ to tyrosine residues at places other than the two metal binding sites. By monitoring the titration at the visible $\lambda_{max}$, stoichiometries of 2:1 metal:protein are obtained for both of these metal ions.

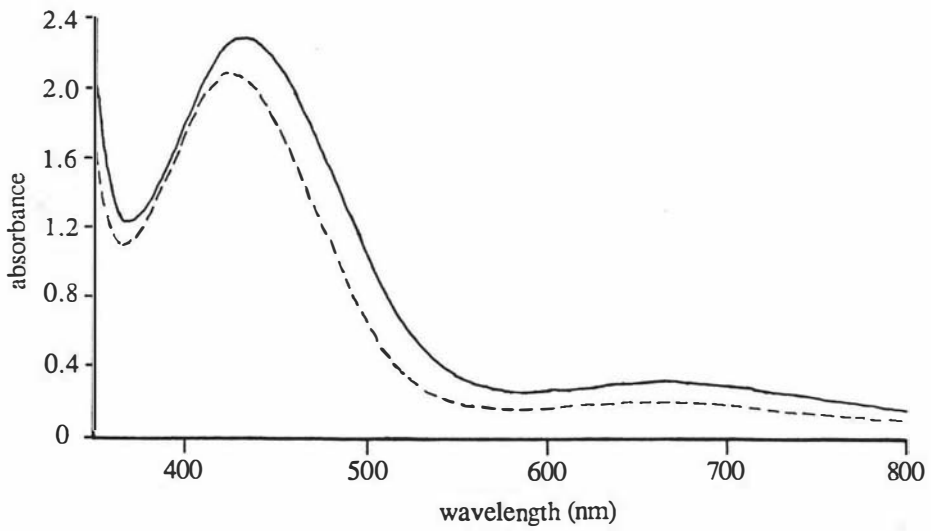



Figure 4.11: Electronic absorption spectra of $Cu_2Lf$ (———) superimposed upon the that of $Cu_2oxLf$ (- - - -) in the range 350 to 800 nm. The concentration of both complexes is 531 μM.

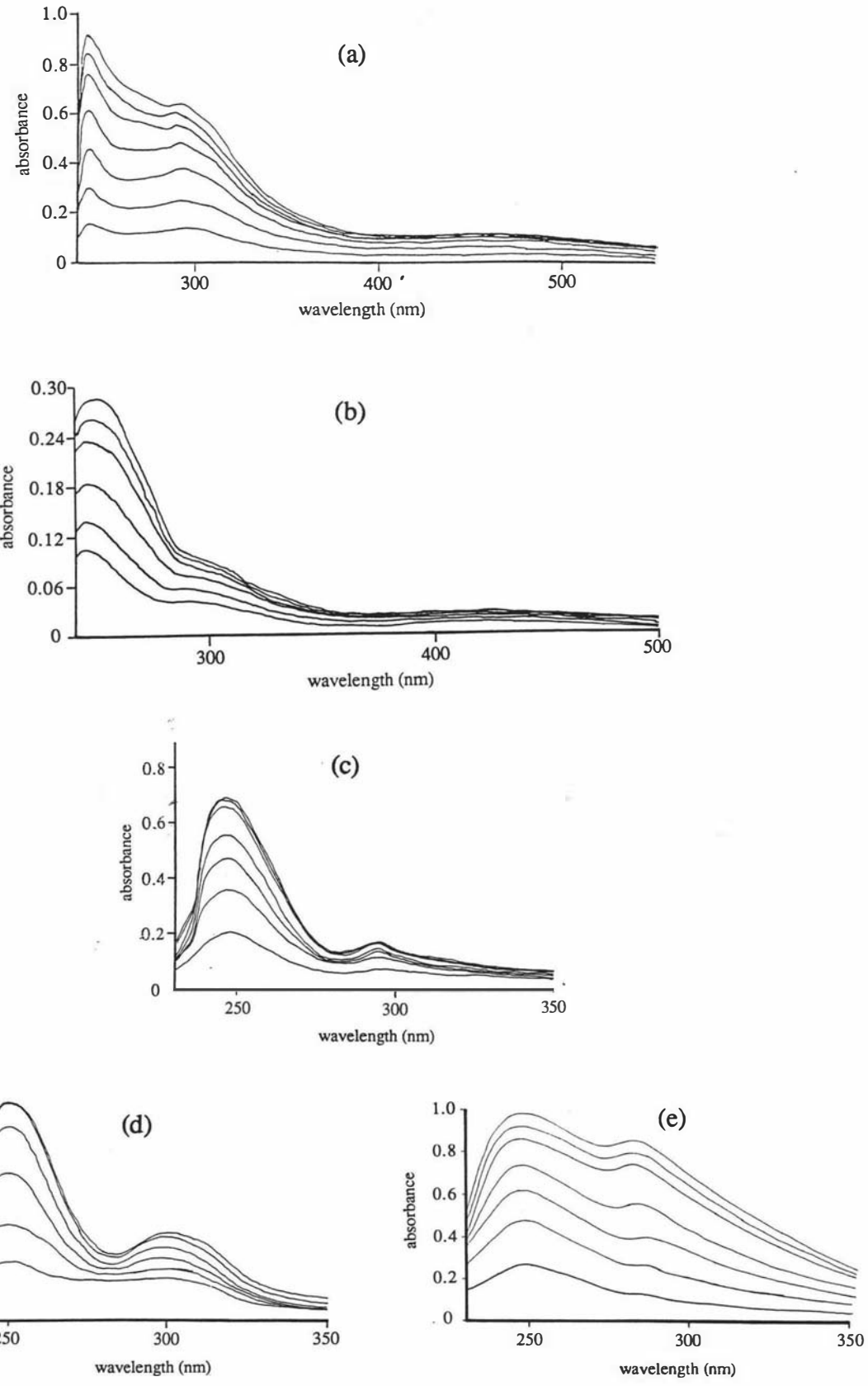



Figure 4.12: UV difference spectra of apoLf upon the addition of increments of a) $Fe^{3+}$, b) $Cu^{2+}$, c) $V^{3+}$, d) $VO^{2+}$ and e) $VO_2^{+}$. For clarity, only every second difference spectrum is shown.

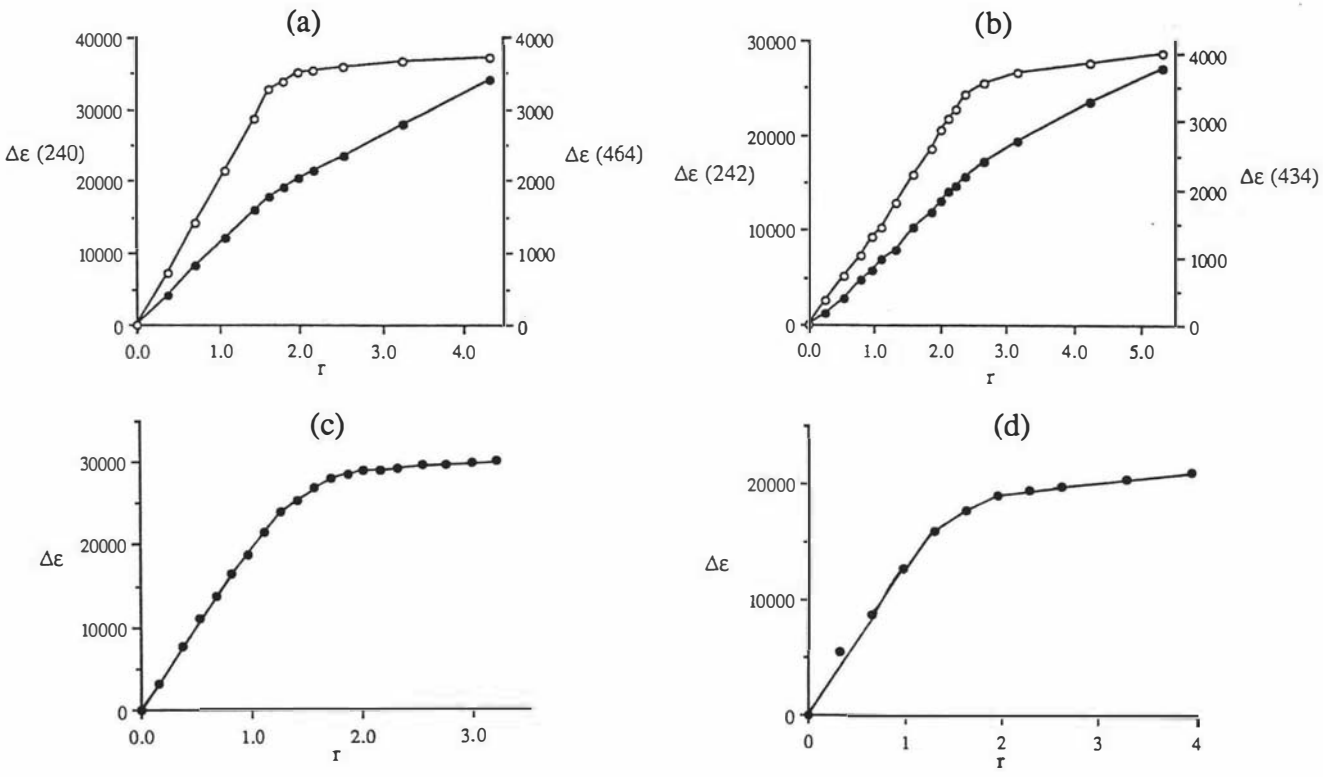



Figure 4.13: Titration curves from UV difference spectra of apoLf with a) $Fe^{3+}$, b) $Cu^{2+}$, c) $V^{3+}$ and d) $VO^{2+}$, plotted as a functionof Δε (245 nm) against r. For $Fe^{3+}$ and $Cu^{2+}$ where there are two plots shown, the filled circles (•—•—•) relate to the left hand scale (at 240/242 nm) and the open circles (o-o-o) relate to the right hand scale (464/434 nm).

Due to the high absorption of vanadate ($VO_2^+$) in the UV region (Harris & Carrano, 1984), the titration curve derived from the V(V) - Lf difference spectra also does not level off at a value of r near 2. As there is no visible band to monitor in this case, a correction for the absorbance of free vanadate is made by subtracting the absorbances for $VO_2^+$ in buffer from the value obtained in the spectrophotometric titration for r > 2. The results of the correction are also shown in Figure 4.14 (next page).

The UV difference spectra resulting from the addition of aluminium to Lf (Figure 4.15 - next page) proved to be different from difference spectra of transferrins substituted with metals such as iron or the lanthanides in that the peaks at 240 and 295 nm were very weak and an intense negative band at 260 nm was the dominant feature. This type of spectrum has also been observed in the difference spectra of hTf titrated with $Al^{3+}$ (Cochran *et al.*, 1984).

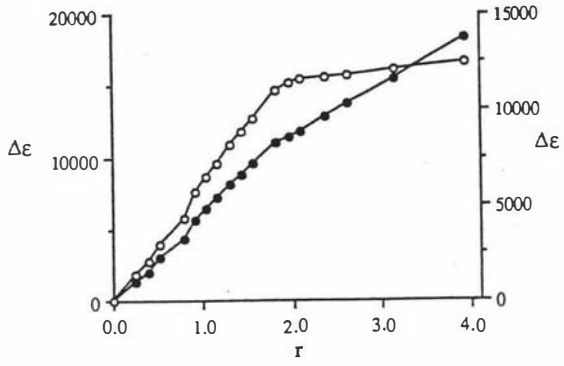



Figure 4.14: Titration curves from UV difference spectrophotometric titrations of apoLf with vanadate (•—•—• - left scale), plotted as a functionof Δε (250 nm) against r. The curve after correcting for free vanadate absorption as shown (o-o-o - right scale).

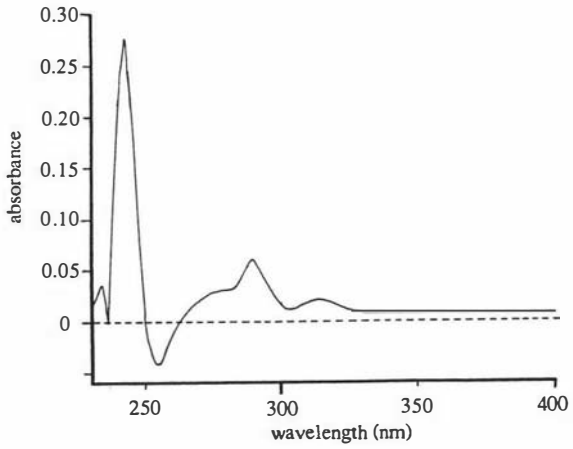



Figure 4.15: UV difference spectrum of the apoLf after the addition of $Al^{3+}$.

It was noted that the intensity of the absorption peak near 240 nm increased with time. In order to obtain a difference spectrum from which a reliable value for the stoichiometry of aluminium binding could be estimated, a titration involving several samples of an apoLf solution containing different amounts of either $Al^{3+}$ or a 1:2 Al:NTA complex was set up and left for 1 week to allow complete incorporation of the available metal ion. The results of this type of titration with both $Al^{3+}$ and $Al(NTA)_2$ are shown in Figure 4.16.

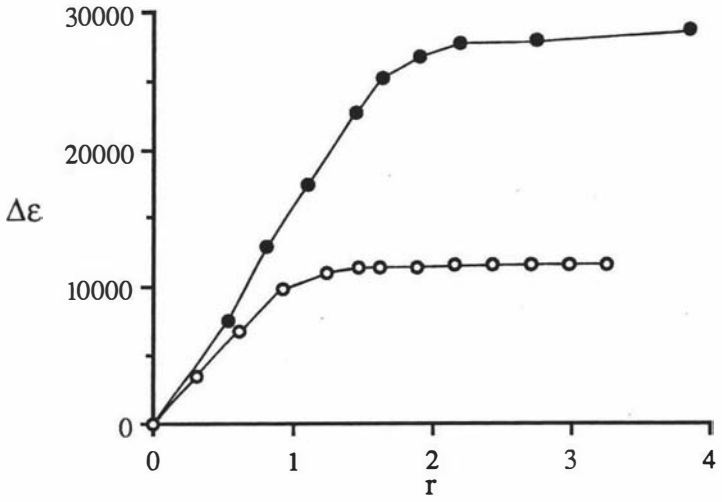



Figure 4.16: UV difference spectrophotometric titrations of apoLf with $Al^{3+}$ (• • •) and $Al(NTA)_2$ (o o o) plotted as Δε at 244 nm against the metal:protein ratio, r.

*Fluorescence titrations:* The results of titrations of apoLf with $Fe^{3+}$ and $Cu^{2+}$ are shown in Figures 4.17a & b. The apoLf solutions were excited at 288 nm and the fluorescence monitored at 340 nm. Titration curves are plotted as relative fluorescence (f) against the metal ion:protein ratio, r. The relative fluorescence values are calculated as the ratio of the observed absolute fluorescence at each addition of metal ion to the initial fluorescence of the apo protein, whose value was taken to be 1.0 (100%).

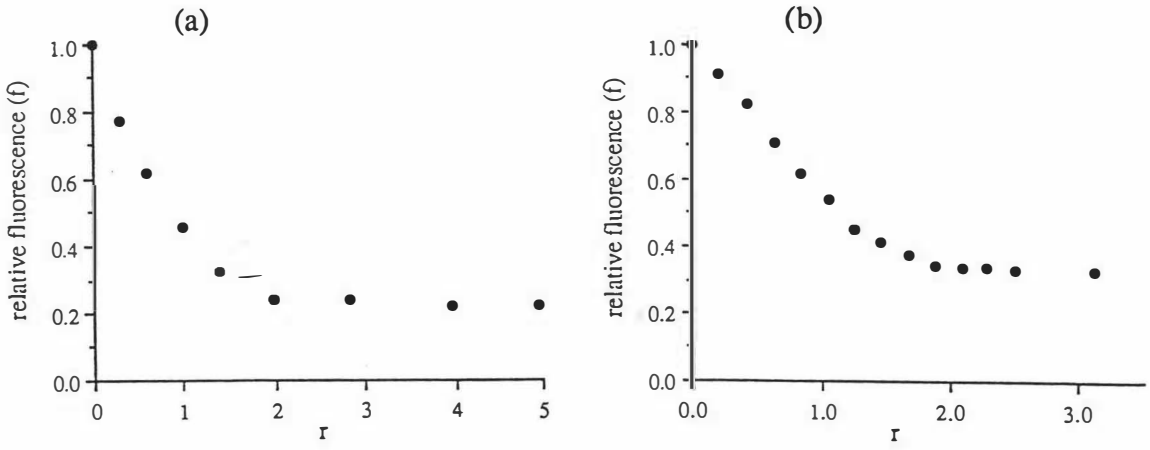



Figure 4.17: Fluorescence titration curves resulting from the addition of a) 10.15 mM $Fe^{3+}$, b) 9.93 mM $Cu^{2+}$ to apoLf. The protein solution concentrations were 11.6 μM for the iron and 9.52 μM for the copper titrations respectively. The samples were excited at 288 nm with a 10 nm bandpass and the emission was monitored at 340 nm, bandpass 4 nm.

Both the iron and copper titrations show that two metal ions bind per protein molecule, similar to results obtained with UV spectrophotometric difference titrations, with a total quenching of the intrinsic apoLf fluorescence of 79% and 66% for iron and copper respectively. As previously reported (Lehrer, 1969; Ainscough *et al.*, 1980), the graph between r=0 and r=2 is not a straight line but rather a curve with an initial gradient higher than the final gradient.

In addition to Lf, fluorescence titrations with iron(III) were also performed on apoTf and apoMTf as a direct comparison with lactoferrin and to establish the stoichiometry of iron binding to melanotransferrin. The results of these titrations are shown in figures 4.18a & b. They confirm that two $Fe^{3+}$ bind per molecule of hTf, with a quenching of over 75% (Lehrer, 1969) and that only one ferric ion is able to bind to MTf, the intrinsic fluorescence being quenched by only 50%.

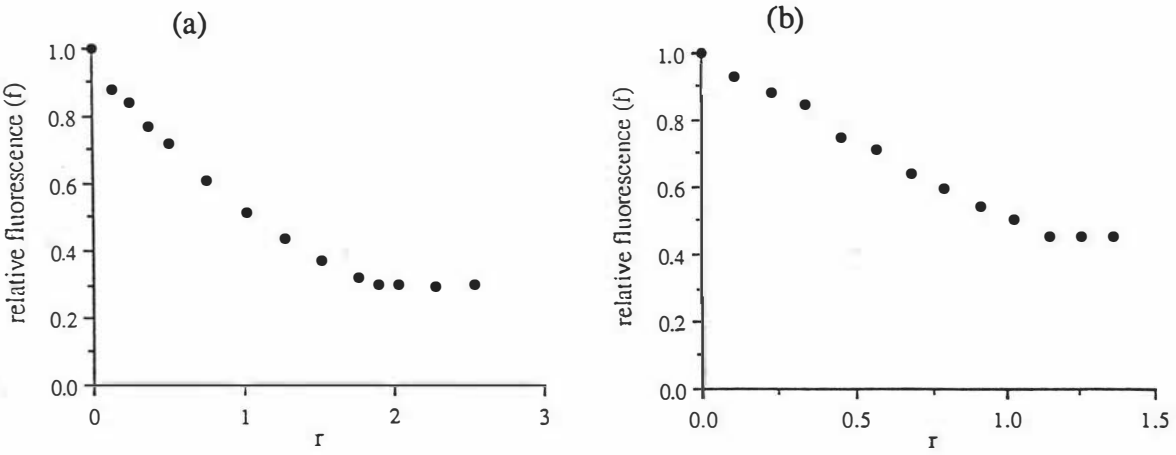



Figure 4.18: Titration curves for the addition of a) 0.946 mM $Fe^{3+}$ to 3.61 μM apoMTf and b) 10.15 mM $Fe^{3+}$ to 5.1 μM apoSTf. Both protein solutions were excited at 288 nm. The apoTf was monitored at 345 nm while apoMTf had a maximum emission at 334 nm.

*Circular Dichroism:* Circular Dichroism spectra of metal-free Lf and that of the iron(III) complex have been previously reported (Mazurier *et al.*, 1976). The $Fe^{3+}$ complex of bLf has also been investigated by CD spectroscopy (Brown & Parry, 1974), as have the $Cu^{2+}$ complexes with cOTf and hTf (Tomimatsu & Vickery, 1972; Prados *et al.*, 1975). The CD spectra for the $Fe^{3+}$, $Cu^{2+}$, $Mn^{3+}$ and $Co^{3+}$ complexes of Lf, in addition to that for apoLf,

are shown in Figures 4.19 - 4.23, with the positions and intensities of the major bands given in Table 4.2. The spectra were plotted as $[\Theta] \times 10^{-4}$ and $[\Theta] \times 10^{-3}$ deg.cm$^2$.dmole$^{-1}$ for the UV and visible regions respectively, although $[\Theta]_M$ values are tabulated as a comparison to the values reported by Tomimatsu & Vickery (1972).

Table 4.2: Peak positions and ellipticity values for apolactoferrin and four metal ion complexes of lactoferrin in the UV and visible absorption regions.

| | UV region (250 - 350 nm) | | | visible region (350 - 700) [a] | | | |
|---|---|---|---|---|---|---|---|
| metal | peak (nm) | $[\Theta]$ [b] | $[\Theta] \times 10^{-4}$ | peak (nm) | $[\Theta]$ | $[\Theta] \times 10^{-3}$ | $[\Theta]_M$ [c] |
| apo | 273 | -76770 | -7.68 | 466 | 536 | 0.54 | — |
| | 285 | -52220 | -5.22 | | | | |
| | 297 | -32140 | -3.21 | | | | |
| $Fe^{3+}$ | 275 | -71510 | -7.15 | 306 | 29450 | 29.5 | 14.7 |
| | 285 | -44360 | -4.44 | 324 | 27150 | 27.2 | 13.6 |
| | 297 | -9560 | -0.96 | 451 | -19890 | -19.9 | -9.95 |
| $Cu^{2+}$ | 272 | -87780 | -8.78 | 330 | 970 | 0.97 | 0.45 |
| | 287 | -65660 | -6.57 | 352 | -4140 | -4.14 | -2.1 |
| | 297 | -47100 | -4.71 | 410 | 8421 | 8.4 | 4.2 |
| | | | | 500 | -3280 | -3.3 | -1.6 |
| | | | | 628 | 11130 | 11.1 | 5.6 |
| $Mn^{3+}$ | 276 | -118710 | -11.87 | 405 | 18770 | 18.8 | 9.4 |
| | 285 | -99390 | -9.94 | 512 | 1170 | 1.17 | 0.59 |
| | 296 | -66260 | -6.63 | 620 | -3990 | -3.99 | -2.0 |
| $Co^{3+}$ | 271 | -74720 | -7.47 | 375 | 34350 | 34.4 | 17.2 |
| | 277 | -66130 | -6.61 | 455 | 7630 | 7.63 | 3.8 |
| | 285 | -52390 | -5.24 | 675 | -20820 | -20.8 | -10.4 |
| | 297 | -44660 | -4.47 | | | | |

[a] The visible CD for the copper complex was measured from 300 to 800 nm, while the spectrum for cobalt was measured from 350 to 750 nm. [b] The units of the ellipticity values are deg.cm$^2$.dmol$^{-1}$. [c] $[\Theta]_M = [\Theta]/2$ to give the ellipticity per dmole of metal ion.

All the spectra in the aromatic region (250 - 350 nm) are very similar, implying that the coordination environments of these four metal ions are similar. It is interesting to note that the spectrum obtained from apoLf is also similar, even though there are no metal ions bound. The spectra consist of a broad negative band at about 276 nm with two smaller bands (at

about 285 and 295 nm) superimposed. In fact, the 276 nm band is itself composed of at least two overlapping bands, the presence of which can be seen quite clearly in the copper and cobalt spectra. The iron, manganese and cobalt spectra (Figures 4.20, 4.22 and 4.23) have the 285 nm band as a shoulder on the side of the 276 nm band, while in the apo and copper spectra (Figures 4.19 and 4.21), this band is more clearly defined. As these bands are due primarily to the aromatic residues in the protein (Trp, Tyr and Phe in order of intensity), it is conceivable that differences in the interaction of the metal ions with the tyrosine ligands in the specific metal binding sites, or in the amount of tyrosine deprotonation, could result in either shifts in the position of the bands or in slight increases in intensity (Wetlaufer, 1962).

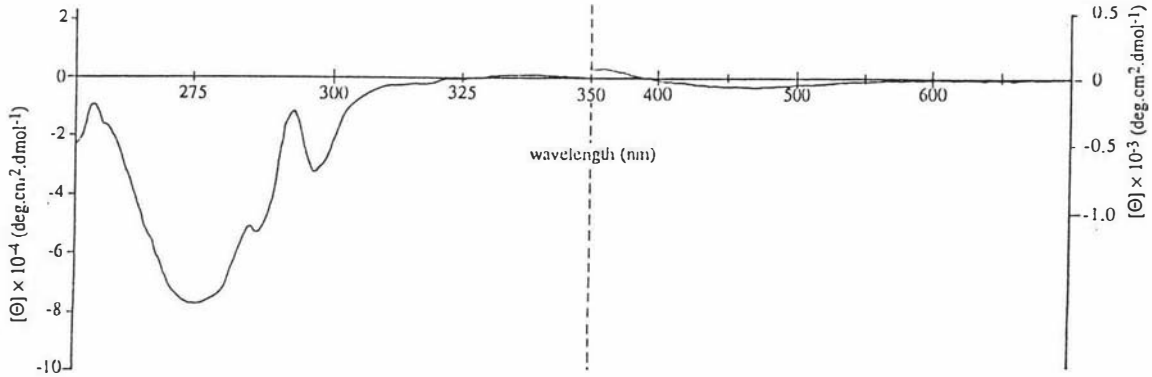



Figure 4.19: Circular dichroism spectra of apoLf in the aromatic (250 - 350 nm) and visible (350 - 700 nm) regions. Note the change in the wavelength scale at 350 nm. Molar ellipticities were calculated using equation 3.1. Protein concentration = 56.9 μM.

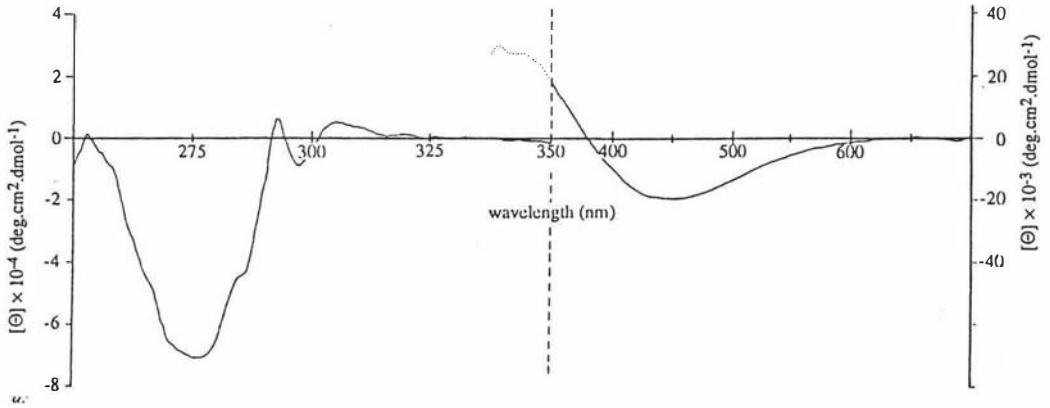



Figure 4.20: Circular dichroism spectra of $Fe_2Lf$ in the aromatic (250 - 350 nm) and visible (300 - 700 nm) regions. Note the change in scale at 350 nm. The part of the visible spectrum shown dotted is that portion of the spectrum measured between 300 and 350 nm. Protein concentration = 66.3 μM.

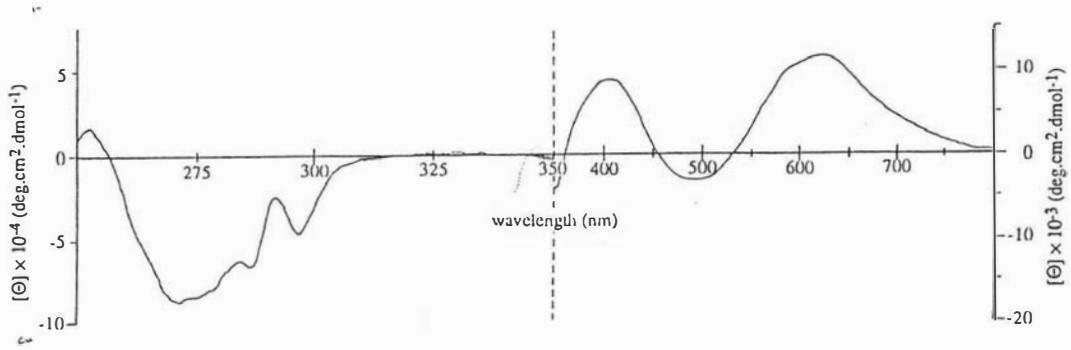



Figure 4.21: Circular dichroism spectra of $Cu_2Lf$ in the aromatic (250 - 350 nm) and visible (300 - 800 nm) regions. Note the change in scale at 350 nm. The part of the visible spectrum shown dotted is that portion of the spectrum measured between 300 and 350 nm. Protein concentration = 65.3 μM.

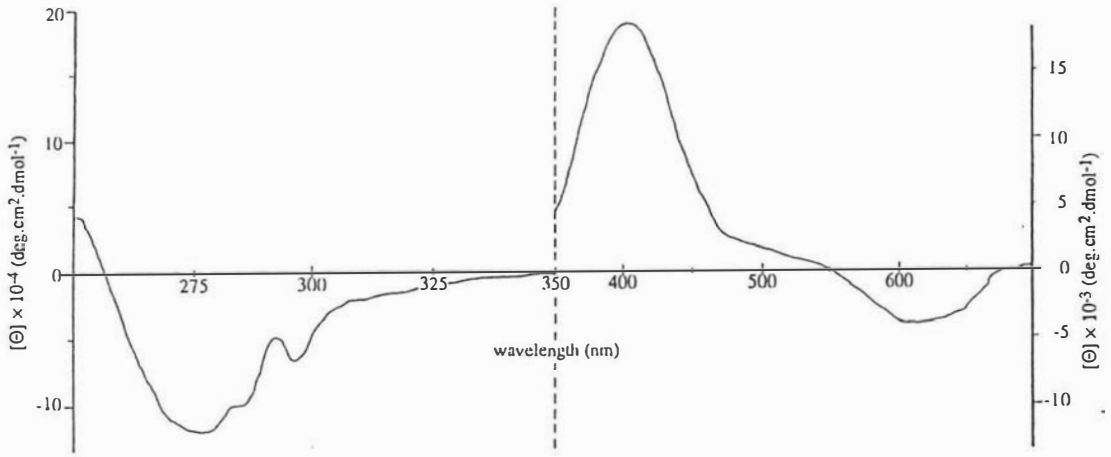



Figure 4.22: Circular dichroism spectra of $Mn_2Lf$ in the aromatic (250 - 350 nm) and visible (350 - 700 nm) regions. Note the change in scale at 350 nm. Protein concentration = 56.3 μM.

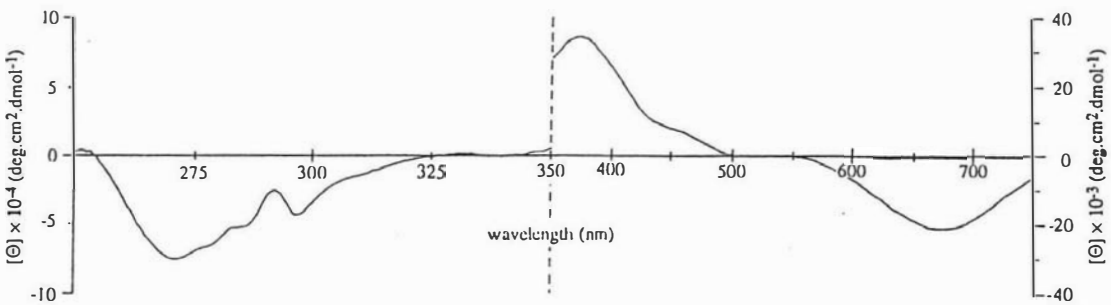



Figure 4.23: Circular dichroism spectra of $Co_2Lf$ in the aromatic (250 - 350 nm) and visible (350 - 750 nm) regions. Note the change in scale at 350 nm. Protein concentration = 147.5 μM.

The visible CD spectra for the four metal ion complexes all mimic their respective electronic absorption spectra, as the phenomenon responsible for both types of spectrum is the same. The spectra obtained for the manganese(III) and cobalt(III) complexes are similar, consisting of three bands with an intense peak centred at about 405 and 375 nm respectively. These correspond with the charge transfer bands at 435 and 405 nm reported for these complexes (Ainscough *et al.*, 1979). The extra bands observed in the Mn(III) spectrum at 512 and 620 nm are possibly due to *d-d* transitions resulting from Jahn-Teller distortion of the high-spin octahedral field into at least three excited terms. These have been observed in the $Mn_2Lf$ absorption spectrum as shoulders at 520 and 640 nm (Ainscough *et al.*, 1979).

The bands at 455 and 675 nm in the $Co_2Lf$ spectrum can be attributed to $^1A_{1g} \rightarrow {}^1T_{2g}$ and $^1A_{1g} \rightarrow {}^1T_{1g}$ transitions for low-spin Co(III) ions (assuming a regular octahedral geometry), which generally occur in this region and have energy separations of about 6000 - 8000 $cm^{-1}$ (Figgis, 1987). They are generally very weak ($\varepsilon \approx 100$) which is possibly why they were not observed in the absorption spectrum of $Co_2Lf$.

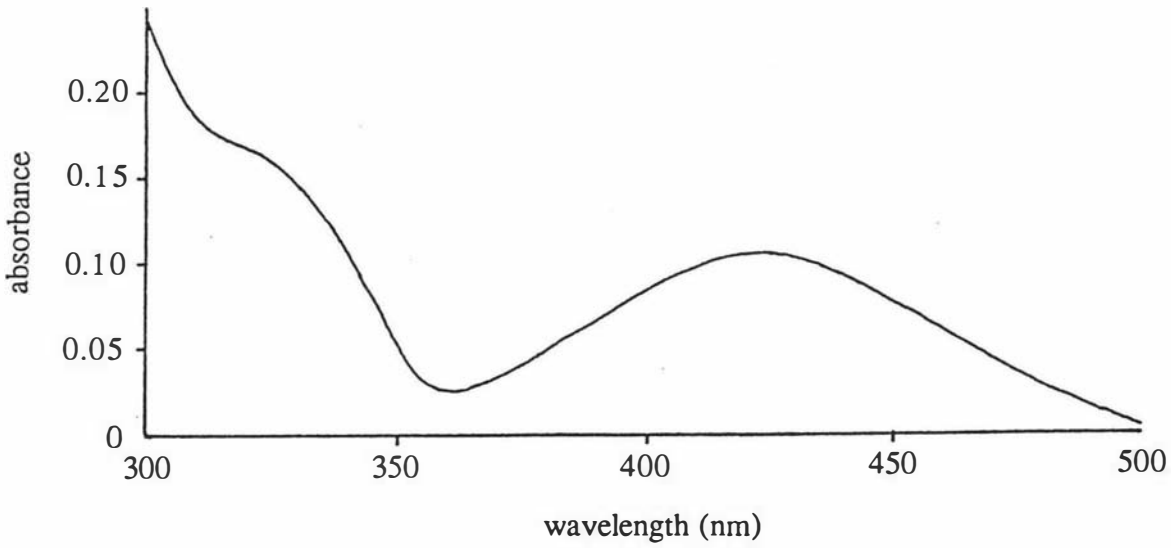



Figure 4.24: Difference spectrum of $Cu_2Lf$ between 300 and 500 nm, with apoLf as the reference. The protein concentration is 21.5 μM.

The spectrum for $Cu^{2+}$ shows the presence of at least three absorption bands at about 315, 430 and 680 nm. The band at 430 nm corresponds to the 434 nm ligand-metal charge transfer band observed in the electronic spectrum (see Figure 4.11), while the 680 nm band

can be assigned as a *d-d* transition arising from the coordination of at least one nitrogen donor atom (Ainscough *et al.*, 1979). The band at 315 nm is not observed in an absorption spectrum of $Cu_2Lf$ due to the presence of the intense 280 nm band attributed to the protein. However, a difference spectrum of $Cu_2Lf$ using apoLf as the reference indicates the presence of this absorption band at about 320 nm (Figure 4.24).

### 4.3.2 *Electron spin resonance spectroscopy:*

The relative stabilities of three oxidation states of lactoferrin-bound vanadium were investigated using ESR spectra. Vanadium(IV) is the only oxidation state of vanadium which is ESR active at liquid nitrogen temperature. Of the other easily accessible states, vanadium(V) ($d^0$) has no free electrons and vanadium(III) ($d^2$) is silent. When bound to the transferrins, V(IV) displays a characteristic ESR spectrum as shown in Figure 4.4. It should be noted that at physiological pH, any unbound vanadium(IV) is rapidly hydrolysed to species such as $VOOH^+$ and $(VOOH)_2^{2+}$ which are ESR silent (Chasteen, 1981), indicating that the observed signal is specific for vanadium bound to transferrin as $VO^{2+}$.

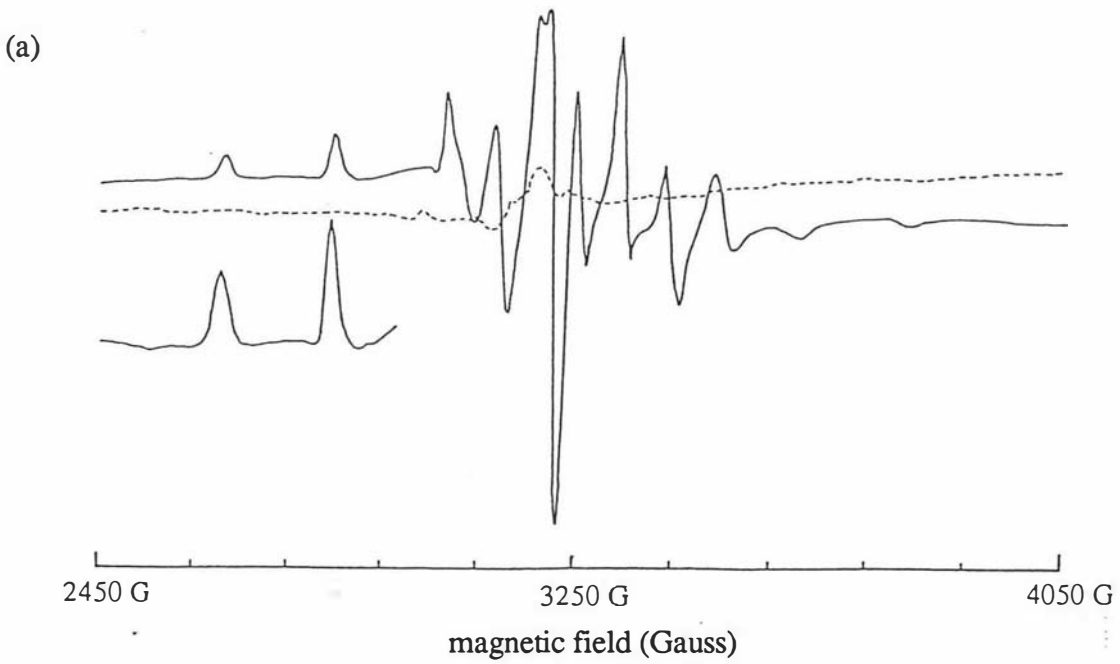



Figure 4.25: Electron spin resonance spectra of (a) $(VO^{2+})_2Lf$ under an argon atmosphere (——) and after exposure to air for 18 hours (- - - -). The receiver gain was $5 \times 10^3$ with a microwave frequency 8.985 GHz at 6 mW power. The sample temperature = -160°C (Figure continued overleaf).

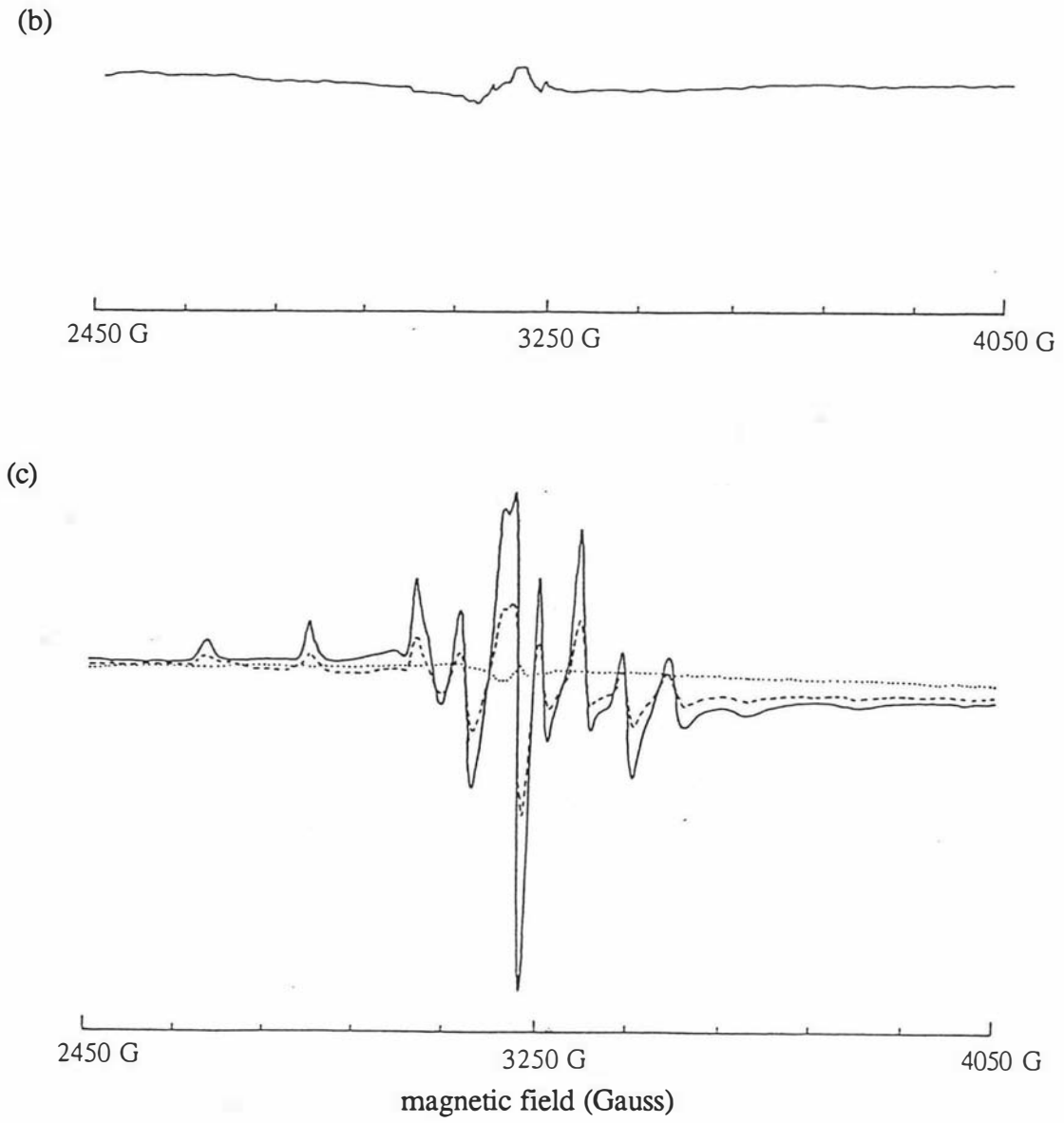



Figure 4.25 continued: Electron spin resonance spectra of $V^{3+}$-saturated Lf (b) with oxygen rigorously excluded and (c) upon opening the solution to the air. The formation of a characteristic $VO^{2+}$ ESR signal appeared and was recorded after five minutes (——), 30 minutes (- - - -) and 6 hours (·····). The microwave frequency is 8.985 GHz at 6 mW power for all spectra, with a gain of $5 \times 10^3$. Sample temperature = -160°C.

It has been reported previously that $VO^{2+}$ in the presence of human transferrin is unstable towards air oxidation (Chasteen *et al.*, 1986a) and a similar effect is observed with human lactoferrin (Figure 4.25a). The 2:1 $VO^{2+}$-Lf complex was prepared under argon and the ESR spectrum recorded. When samples of this complex were opened to the air (with $O_2$ bubbled through) or had a small volume of 5% $H_2O_2$ added, the results were similar: the

complete loss of ESR activity within 5 minutes for the $H_2O_2$-treated complex and one hour for the sample exposed to oxygen.

The ESR spectra of V(III)-lactoferrin in the absence and presence of oxygen are shown in Figures 4.25b & c (previous page). In the absence of oxygen, the spectrum is essentially flat, although the features discernible near 3000 G (Figure 4.25b) can be attributed to a small (2-3%) amount of V(IV) contamination in the vanadium trichloride used to prepare the standard $V^{3+}$ solution, as estimated by comparison with standard $VO^{2+}$ solutions. Once opened to the air (again with $O_2$ bubbled through) , the vanadium(III) is rapidly oxidised to vanadium(IV) and a strong $VO^{2+}$ signal is observed after 5 minutes incubation (Figure 4.25c). This signal decays over time and a flat line results once again as the vanadium(IV) is oxidised to vanadium(V). This latter result is also observed with hTf (Chasteen *et al.*, 1986a) where the half-life for the oxidation has been estimated to be between 5 and 13 minutes.

Electron spin resonance spectra of the copper-carbonate and copper carbonate-oxalate complexes, $Cu_2$Lf (Figure 4.26a) and $Cu_2$oxLf (Figure 4.26b), were used primarily as fingerprints to establish the nature of the complexes formed prior to crystallisation. The ESR spectrum of $Cu_2$Lf has been reported previously (Ainscough *et al.*, 1980) and is a typical axial copper(II) spectrum with $g_\perp = 2.060$, $A_\perp = 12$ Gauss and $g_{||} = 2.314$, $A_{||} = 155 \times 10^{-4}$ $cm^{-1}$, indicative of a $d_{x2-y2}$ ground state for the unpaired electron (Hathaway, 1987). The $g_\perp$ signal is overlaid with four hyperfine signals due to the coupling of the $^{65}Cu$ nucleus with the coordinated histidyl nitrogen (with nuclear spin, I = 1). These hyperfine signals are also observed in the $g_{||}$ region, where the splitting is much less resolved.

The ESR spectrum of $Cu_2$oxLf is similar to that previously described (Ainscough *et al.*, 1983) and shows the same features as that of $Cu_2$Lf, except that in the $Cu_2$Lf spectrum, the $A_{||}$ hyperfine are single bands (although, as observed in $Cu_2$Tf, Figure 4.3, there is the possibility of superhyperfine splitting of the lowest field band) while in $Cu_2$oxLf they are resolved into two signals, the set of four at higher field due to copper-carbonate site and those lower resulting from the copper site occupied by oxalate.($g_{||} = 2.345$, $A_{||} = 158 \times 10^{-4}$ $cm^{-1}$).

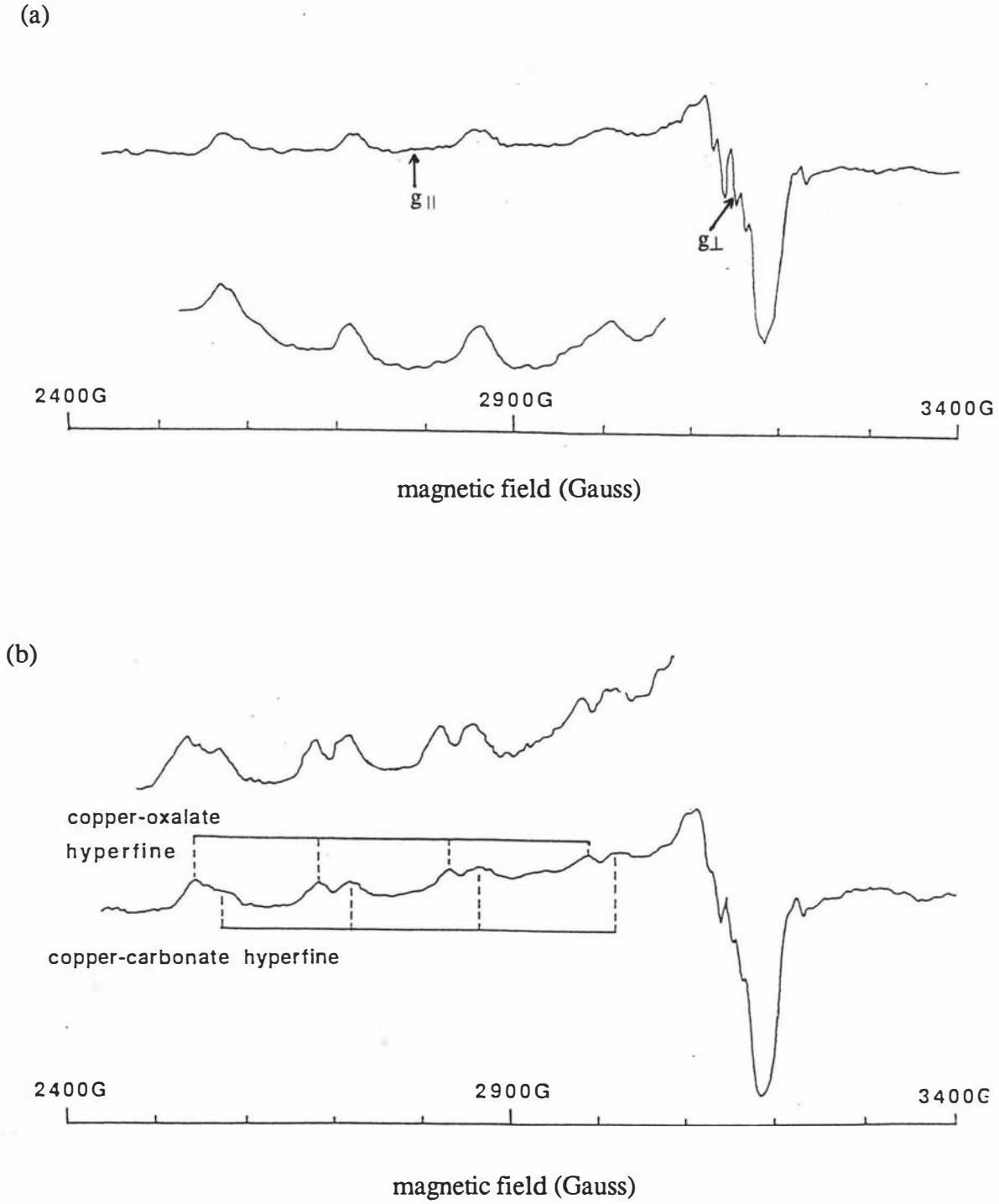



Figure 4.26: Electron spin resonance spectra of (a) $Cu_2Lf$ and (b) $Cu_2oxLf$ between 2400 and 3400 G, microwave frequency 8.999 GHz The sample temperature was -160°C. The nitrogen hyperfine signals are indicated on both spectra. The set at lower field in $Cu_2oxLf$ correspond to the oxalate-containing site, while the second set at higher field are due to the carbonate site.

## 4.4 Discussion A; transition metals and aluminium

### 4.4.1 *Iron and copper:*

*Electronic absorption:* The electronic absorption spectra recorded for the iron and copper complexes of human lactoferrin in the visible region are indistinquishable from those reported by Ainscough *et al.* (1979). In addition, a second charge transfer band at about 320 nm, not previously observed in the absorption spectrum due to its proximity to the intense 280 nm absorption band of the aromatic résidues in the protein, can be seen in both a UV difference spectrum and in the circular dichroism spectrum of $Cu_2$Lf in the visible region.

Electronic absorption and circular dichroism spectra of a copper(II)-D-*o*-tyrosine complex have shown the presence of at least two ligand-metal charge transfer bands at 330 and 400 nm respectively (Garnier-Suillerot *et al.*, 1981). These bands were assigned to tyrosine → copper charge transfer transitions between the $\pi_a$ and $\pi_b$ phenolate orbitals respectively, and the $d\sigma^*$ orbital on the Cu(II) ion. The bands correspond to the 320 nm band and the 430 - 440 nm band observed in the copper complex of lactoferrin (and all the transferrins in general). It should be noted that the higher energy phenolate - metal charge transfer band has also been observed in the circular dichroism spectrum of $Fe_2$Tf at about 360 nm (Nagy & Lehrer, 1972) and it can be assumed that it arises from a $\pi_b \rightarrow d\pi^*$ transition.

Using purely crystal-field arguments, it is possible to represent the copper(II) atomic orbitals in an energy diagram as shown in Figure 4.27. Assuming tetragonal distortion of an octahedral geometry such that the axial bonds are lengthened (the equatorial ligands are in the xy plane), then the energy of the $d_{x^2-y^2}$ orbital (relative to the same orbital in an octahedral field) will be greater while the $d_{z^2}$ will be lower. The "electron hole" will "float" to the highest energy orbital, the $d_{x^2-y^2}$. Although it is not possible to predict whether elongation or contraction of the axial bonds will occur, the situation shown in Figure 4.27 is consistent both with the assignment of the 430 nm band in the transferrins as a $\pi_b \rightarrow d_{x^2-y^2}$ transition (Garnier-Suillerot *et al.*, 1981) and with the observed positions of the $g_\perp$ and $g_{||}$ resonances in the $Cu_2$Lf ESR spectra (see above). This argument can be extended to cover the observed geometry in the N-lobe of $Cu_2$Lf, where the stereochemistry is square pyramidal with an

elongated axial ligand. The loss of one of the axial ligands (to go to 5-coordinate copper) will further split the $e_g$ orbitals, raising the $d_{x^2-y^2}$ still higher.

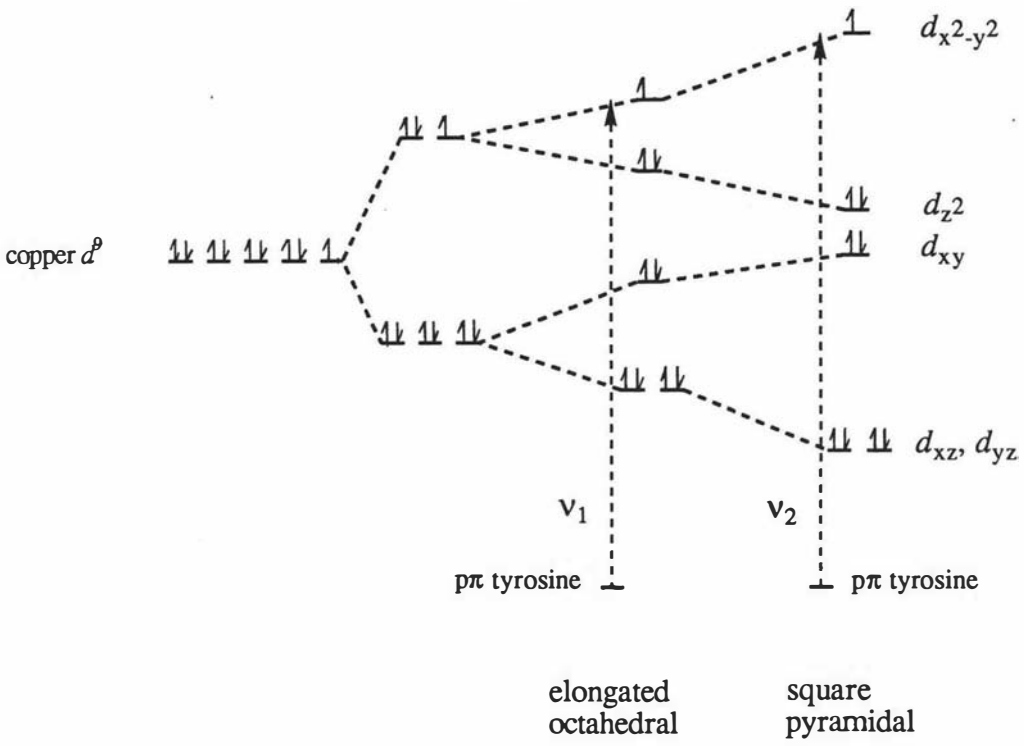



Figure 4.27: Crystal-field energy diagrams for $d^9$ Cu(II) in a regular octahedral field (left), upon tetragonal distortion (centre) and upon removal of one axial ligand to create a square pyramidal geometry (right). The energy of a charge transfer transition in the square pyramidal complex ($\nu_2$) will be greater than in the tetragonally distorted case ($\nu_1$), resulting in a $\lambda_{max}$ at a lower wavelength.

It is interesting to note that the extinction coefficients (designated $\Delta\varepsilon_1$ and $\Delta\varepsilon_{tot}$ - see Table 4.6) derived from the spectrophotmetric titration curves with iron and copper are significantly greater for the $Fe^{3+}$ than for $Cu^{2+}$. In light of the distorted 5-coordinate copper geometry seen in the N-lobe of $Cu_2Lf$, this is not surprising, in that the interaction of the copper with the apical Tyr92 ligand would be significantly weaker, possibly leading to lesser perturbation of the $\pi \rightarrow \pi^*$ transitions in the phenolate ring.

*Differences in copper coordination between the two sites:* Even though the monodentate carbonate observed in the N-lobe site of $Cu_2Lf$ was unexpected, the 5-coordinate geometry with the elongated apical bond which results is a favoured geometry for copper(II), more so than a regular octahedral geometry, due to the lack of cubic symmetry of the $3d^9$

configuration. The observed differences in copper coordination between the N- and C-lobe can be explained in terms of the differences in domain closure between the two binding sites (Section 3.5), in that the copper in the N-site is more able to adopt the square-pyramidal geometry than the metal in the C-site. However, the question remains as to why it has not been possible to clearly distinquish between the two different geometries by spectroscopic methods.

It seems reasonable to assume that the observed structural differences between the two copper sites should give rise to two different LMCT bands in the visible region. The effect of these differences is difficult to gauge, but it seems likely that the 434 nm band is composed of two slightly different transitions from the N- and C-sites. This is consistent with the results of studies on copper binding to cOTf, as noted in Section 3.4.1, and to the N-terminal half of Lf ($Lf_N$).

It has been shown that the Cu-$Lf_N$ complex has an absorption maximum at 422 nm (Day *et al.*, 1992) and this can be rationalised in terms of a simplified crystal-field model as given in Figure 4.27 (a regular octahedral crystal field is assumed, although as the geometries are rather distorted, this might not accurately reflect the true energies of the *d*-electrons in the valence shell). If the right-hand energy-level diagram is taken to represent the splitting in the N-lobe of $Cu_2Lf$ and in the Cu-$Lf_N$ complex, and the central diagram were to represent the situation in the C-lobe, then it can be seen that a decrease in coordination number from six to five (from a change from bidentate to monodentate coordination of the anion), would result in a decrease in energy of the $d_{z^2}$ orbital and an increase in the $d_{x^2-y^2}$. This would lead to a subsequent increase in energy of the tyrosinate ($\pi$) $\rightarrow$ copper ($d_{x^2-y^2}$) charge transfer transition, $\nu_1 < \nu_2$ (i.e. lower wavelength).

The difference between the sites cannot be readily detected in the electron spin resonance spectrum of $Cu_2Lf$, even though the ESR spectrum is usually very sensitive to the environment of the $Cu^{2+}$ ion. For example, in the spectrum of $Cu_2OTf$, differences in the $g_{\parallel}$ hyperfine signals are observed (Zweier, 1980), indicative once again of a difference in the coordination geometries in the two metal binding sites in this protein. These differences are enhanced as the pH increases, although at pH 7.4 they are still apparent.

When oxalate is introduced into one site of Lf, the differences between the two sites, at least in terms of the ESR experiment, are much more significant. A new set of hyperfine resonances in the $g_{||}$ region appear with parameters quite distinct from those of the carbonate-containing site ($g_{||} = 2.314$ and 2.345 for carbonate and oxalate respectively).

The N-lobe site of $Cu_2Lf$, being essentially square pyramidal, has a well defined equatorial set of ligands, Asp 60, Tyr 192, His 253 and the $O_2$ atom of the carbonate (Figure 4.28a) at distances of 2.0 - 2.1 Å. The copper atom sits about 0.1 Å above this plane in the direction of the apical ligand, Tyr92, which is at a somewhat greater distance of about 2.8 Å from the copper. This geometry is typical of the square pyramidal small-molecule copper(II) complexes which have been crystallised (Hathaway & Billing, 1970). Although at first glance the C-lobe site appears to be quite different, the equatorial donor set is still $NO_3$ and the bond angles are similar. In terms of the ESR experiment, where the spectrum is generally dependent only on the nature and geometry of the equatorial ligands, it would be predicted that these two sites might give rise to virtually identical ESR signals.

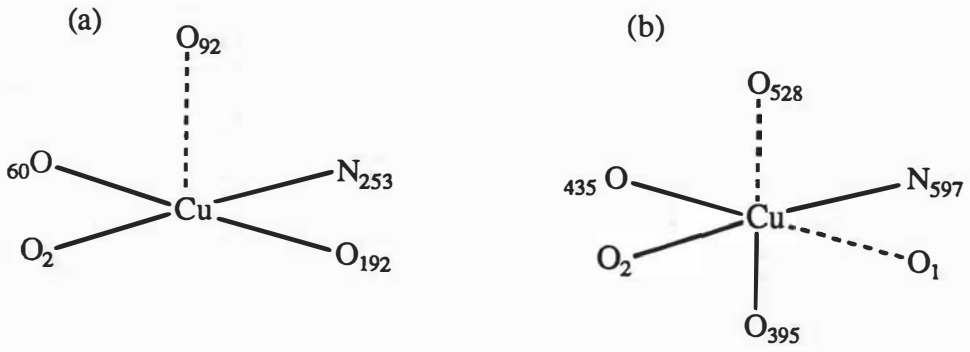



Figure 4.28: The coordination sphere of the (a) N-lobe and (b) C-lobe copper(II) ions in $Cu_2Lf$ arranged so that the longer Cu-$O_\eta$ (Tyr) bond is an axial position.

However, the ligand which is taken to be in the apical position in this case (Tyr435) does not exhibit the same elongation in the Cu - $O_\eta$ bond length as observed in the N-site. In fact, it is the other tyrosine ligand (Tyr528) which appears to have a longer Cu - $O_\eta$ bond of about 2.4 Å. By taking the equatorial plane through a different set of four ligands (Tyr 435, His 597 and the $O_1$ and $O_2$ atoms of the bidentate carbonate (with bond distances of ≈2.0 Å),

with the slightly longer Cu - Tyr 528 bond length (≈2.4 Å) in the apical position), the bond lengths in the equatorial plane are all similar at about 2.0 - 2.1 Å (Figure 4.28b). Once again, the donor set is $NO_3$ although the bond angles are now rather different, especially at the bidentate carbonate where the angle is only about 60°.

Clearly the two sites in $Cu_2$Lf are essentially equivalent, at least as far as the ESR experiment is concerned, although neither of the explanations given above in terms of the observed structure in the copper binding sites would indicate that this should be so. No account has been made of the fact that the ESR spectra are recorded at -160°C, where the water present in the interdomain cleft and in the vicinity of the anion binding arginines would have a much more ordered structure. It is conceivable that this could affect the conformation at the metal binding sites, in that the N-lobe copper may be forced to become 6-coordinate, or the C-lobe copper 5-coordinate. Such a solvent-induced conformation change is not unprecedented in light of the effect that freeze-drying has on the coordination number of copper in $Cu_2$OTf (Garratt *et al.*, 1992).

The substitution of carbonate by oxalate in the C-lobe also produces a shift in position of $\lambda_{max}$ in the absorption spectrum, from 434 to 424 nm. This results from an increase in charge density on the copper, attributed to a larger *d*-orbital splitting upon going from carbonate to oxalate, the latter being higher in the spectrochemical series (Moeller, 1982). This increased splitting raises the energy of $d\sigma^*$ orbitals and results in an increase in the energy of the $\pi_b \rightarrow d\sigma^*$ charge transfer transition (i.e. the resulting absorption band is at a shorter wavelength). Once again, the 424 nm band in $Cu_2$oxLf will result from the combination of two slightly different transitions in the N-and C-sites. It is interesting to note that a lactoferrin complex containing two $Cu^{2+}$ and two oxalate ions prepared in our laboratory by the complete removal of carbonate prior to the addition of a 50-fold molar excess of sodium oxalate and two $Cu^{2+}$ ions, exhibits a $\lambda_{max}$ at 420 nm (Shongwe *et al.*, 1992).

*Fluorescence:* Fluorescence titrations of apoLf with $Fe^{3+}$ (Ainscough *et al.*, 1980) and $Cu^{2+}$ (Ainscough *et al.*, 1983) have been reported previously and, together with those

determined in the course of this study, show a non-linearity of the titration curve between r=0 and r=2. It has been suggested (Lehrer, 1969; Ainscough *et al.*, 1980; Harris & Aisen, 1989) that if the metal ion in the binding site occupied first quenched all the intrinsic tryptophan fluorescence within a critical distance (estimated for $Fe^{3+}$ to be ≈21 Å by Lehrer, 1969, which is half the intersite distance in $Fe_2Lf$ given by Anderson *et al.*, 1989), the second equivalent of metal ion would quench somewhat less due to an overlap in the region between the two sites. This implies that at least one of the tryptophan residues may be within energy transfer distance of both iron atoms, while others are able to transfer their energy to only one of the metal centres.

Table 4.3: The distances from the iron and copper atoms to the ten tryptophan residues in $Fe_2Lf$ and $Cu_2Lf$ [a]

| N-lobe | $Fe_2Lf$ distances (Å) | | $Cu_2Lf$ distances (Å) | |
|---|---|---|---|---|
| | Fe 693 [b] | Fe 694 | Cu 693 [b] | Cu 694 |
| Trp8 | 16.7 | 46.6 | 15.7 | 46.2 |
| Trp22 | 20.1 | 50.4 | 19.4 | 50.5 |
| Trp125 | 13.1 | 47.3 | 13.6 | 47.2 |
| Trp138 | 22.7 | 45.8 | 23.1 | 45.5 |
| Trp268 | 20.9 | 38.1 | 20.4 | 37.9 |
| | 18.7 [c] | | 18.4 [c] | |
| C-lobe | | | | |
| Trp347 | 36.0 | 16.2 | 36.4 | 16.0 |
| Trp361 | 40.3 | 20.3 | 40.8 | 19.9 |
| Trp450 | 54.6 | 17.7 | 55.1 | 18.2 |
| Trp469 | 42.1 | 12.8 | 42.4 | 12.8 |
| Trp562 | 56.3 | 18.4 | 56.3 | 18.3 |
| | | 17.1 [c] | | 17.0 [c] |

[a] The average position of each Trp sidechain was used to calculate the distances given. [b] The residue numbers 693 and 694 refer to metal ions in the N- and C-terminal binding sites respectively. c The average distance between the tryptophan residues and the metals ions in the two lobes.

There are ten tryptophan residues in the sequence of lactoferrin (Anderson *et al.* (1989), five in the N-lobe (residues 8, 22, 125, 138 and 268) and a further five in the C-lobe (347,

361, 450, 469 and 562). Analysis of the structures of $Fe_2Lf$ and $Cu_2Lf$ (Table 4.3) show that in both cases, nine of these residues lie within this 21 Å limit (four in the N-lobe and five in the C-lobe) but that no tryptophan residue comes within 21 Å of both metal sites. The closest residues are Trp268 in the N-lobe (≈38 Å from the C-site metal) and Trp347 in the C-lobe (≈36 Å from the N-site metal), and these are well outside the critical limit.

The efficiency of long range Förster energy transfer can be given by the following equation (Lehrer, 1969)

$$E = \frac{(R_o/R)^6}{1 + (R_o/R)^6} \qquad [4.5]$$

were R is the distance from the donor (≈ 38 Å Trp268, for example) to the acceptor, $Fe^{3+}$ and $R_o$ is the critical distance for $Fe^{3+}$. In this case, E has a value of only 0.028, that is, only about 3% of the energy absorbed by the residue is transferred to the C-lobe iron atom. This is in contrast the energy transferred to the to N-lobe iron by this same residue, which is about 51% of that initially absorbed.

Obviously the partial quenching of tryptophan residues in one lobe by an added metal ion in the other lobe does not play an important role in determining the total amount of fluorescent energy quenched. The curvature in the fluorescence titration curves cannot be explained by this mechanism. However, what is clear from the analysis of the two structures is that the tryptophan residues in the C-lobe are on average closer to their corresponding metal centre than are those in the N-lobe (17 Å compared to about 18 Å), which would result in a more efficient transfer of energy from the excited tryptophans to the metal. The average E value for the five tryptophan residues in the N-lobe of $Fe_2Lf$ is 0.65, while the average for those in the C-lobe is 0.77. The corresponding values for $Cu_2Lf$ (where the critical distance, $R_o$, is 18 Å) are 0.58 and 0.49.

Hence, the observed non-linearity in the fluorescence titration curves results primarily from the differences in the proximity of the tryptophan residues to the two metal sites. This would suggest that it is the C-site which is occupied first upon the addition of both $Fe^{3+}$ and $Cu^{2+}$, in that the calculated quenching of the tryptophan fluorescence is greatest in the C-lobe for both of these metal ions. This seems a reasonable assumption, as a similar curvature of

the titration curve is also observed with hTf when titrated with FeNTA and it is the C-site of hTf which is preferentially loaded by FeNTA at physiological pH (Aisen & Harris, 1989).

Significant quenching of lactoferrin has been reported for a variety of M(II) and M(III) ions (Ainscough *et al.*, 1980) and in general, it appears that metal ions which give rise to intensely-coloured complexes quench the fluorescence to a greater extent than those metals which do not exhibit strong charge-transfer bands (Table 4.4). For example, $Mn^{3+}$ quenches about 65% of the fluorescence, while $Mn^{2+}$ quenches just over 20%.

Table 4.4: Relative fluorescence values of various metal-lactoferrin complexes (from Ainscough *et al.*, 1980)

| metal ion | f [a] | metal ion | f |
|---|---|---|---|
| Fe(III) | 0.22 | Co(II) | 0.77 |
| Cu(II) | 0.28 | Ni(II) | 0.81 |
| Mn(III) | 0.34 | Zn(II) | 0.82 |
| Co(III) | 0.47 | Cd(II) | 0.77 |
| Cr(III) | 0.63 | Hg(II) | 0.81 |
| Mn(II) | 0.78 | Pb(II) | 0.74 |

[a] f = fluorescence of the metal-protein complex relative to apolactoferrin.

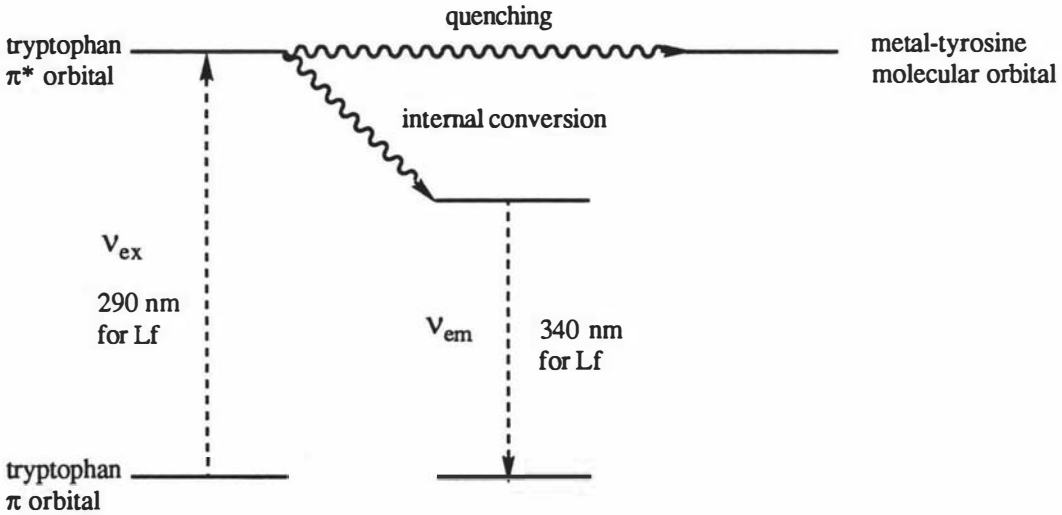



Figure 4.29: A simplified molecular energy diagram for transitions involving a tryptophan residue within energy transfer distance of a metal-tyrosine molecular orbital. Electrons are excited from the ground state of a tryptophan ($\pi$) into an excited state ($\pi^*$). In the absence of a metal ion in the specific binding site, internal conversion to a second excited state occurs followed by emission at lower energy (higher wavelength). If a metal ion is present, some of the energy of the excited electrons is transferred to the metal-tyrosinate molecular orbital (quenching), thus lowering the efficiency of the internal conversion process and decreasing the fluorescent energy output.

This suggests that the energy from the excited tryptophan residues is transferred into a molecular orbital resulting from the overlap of a tyrosine $\pi$ orbital with the $d\sigma^*$ orbital associated with these metal ions, when the energy of the LMCT transition is close to the energy of the emission. For the lactoferrin, the emission is at 334 nm, which corresponds almost exactly with the $\pi_b \rightarrow d\pi^*$ and the $\pi_b \rightarrow d\sigma^*$ transition observed in the iron and copper complexes of the transferrins respectively (dicussed earlier). The amount of energy transferred is dependent on the metal ion in question and the critical distance for long range energy transfer. Other factors, such as the energy of the LMCT transition and on the oxidation state of the metal may also affect the efficiency of the Förster transfer. Figure 4.29 represents the phenomenon of quenching with respect to the transferrins.

The fluoresence titration preformed on apoMTf proved to be very interesting, in that it showed that the protein was capable of binding only one $Fe^{3+}$ ion. Subsequent analysis of the structures of Lf and MTf suggest that in MTf a number of changes have occurred in the N-lobe site, and that this site is incapable of coordinating iron(III) (Baker *et al.*, 1992).

### 4.5.2 *Vanadium:*

Vanadium is able to bind specifically to lactoferrin in three oxidation states, +5 (as $VO_2^+$), +4 (as $VO^{2+}$) and +3. The stoichiometry of binding was observed to be 2:1 in all cases, as determined from UV difference spectrophotometric titrations. The difference spectra obtained from $V^{3+}$ and $VO^{2+}$ titrations resemble those for the lanthanides, in that the bands at 240 and 295 nm are sharp and clearly defined, although in intensity they are more similar to other transition metals ($\Delta\varepsilon_{tot}$ = 28,700, 17,300, 21,500 and 17,500 for $V^{3+}$, $VO^{2+}$, $Fe^{3+}$ and $Cu^{2+}$ respectively, while those for the $Ln^{3+}$ ions range between 30,000 to 60,000). Based on the intensities of these bands at full saturation and the associated $\Delta\varepsilon_1$ (19,580, 11,350, 12,600 and 9,850 - see Table 4.6), it seems reasonable to suggest that two $VO^{2+}$ and two $V^{3+}$ ions bind specifically in the two metal binding sites in a similar position to iron and that they each interact with at least two tyrosine residues. The stoichiometry of binding of vanadium(III) to hTf (Bertini *et al.*, 1982) has also been shown to be 2:1.

It is interesting to note that the extinction coefficients of the 245 nm bands decrease with increasing oxidation number of the vanadium species. This could be due to several factors including the association constants of the three cations, the decrease in the number of *d* electrons (from 2 for $V^{3+}$ to 0 for $VO_2^+$), the decrease in the overall charge on the cation (from +3 to +1) and the increase in the number of covalently-bound oxygen atoms (from 0 to 2).

*Stability of vanadium(III) bound to the transferrins:* Bertini *et al.* (1982) reported the preparation of an air-stable vanadium(III)-hTf complex with a 2:1 stoichiometry which they linked to the function of the protein as a stabiliser of metal ions in the +3 oxidation state. Although other workers in the area (Harris & Carrano, 1984) were not convinced by the arguments of Bertini and colleagues, the redox stability of transferrin-bound V(III) was reiterated by Bertini *et al.* (1985).

The 2:1 V(III)-Lf complex is not stable towards oxidation and is converted to V(IV) in a matter of minutes. The product, V(IV)-Lf is detectable by ESR only as an intermediate, as it too is oxidised to the stable V(V) complex. It is not clear why there should be this difference between transferrin and lactoferrin, as both have metal binding sites accessible to $O_2$, as seen by the oxidation of bound Mn(II) and Co(II) to Mn(III) and Co(III) respectively (Aisen *et al.*, 1969; Ainscough *et al.*, 1979). It seems very unlikely that V(III) ions would be stable towards oxidation unless the pH is very low (Vilas-Boas & Costa Pessoa, 1987).

Although it is difficult to estimate the pH experienced by the metal ion once it is bound in the specific binding site, it is assumed that it would not be very much different from the external solvent or the solvent in the interdomain cleft. The release of protons from the tyrosines (and possibly the bicarbonate) upon metal coordination could lower the local pH somewhat, although this would be offset by the presence of a basic sidechain such as the anion-binding arginines (Arg121 and 465 in the N- and C-terminal lobes) and a second basic sidechain (Arg210 and Lys 546) within 6.5 Å of the metal ion (Section 3.2.7).

*Vanadium(V) and vanadium(IV):* The spectra obtained from the $VO_2^+$ - Lf titrations differ from those obtained with iron and copper, in that the band at 240 nm has shifted to 252 nm and broadened and the 295 nm band is not well resolved. This is due to the significant absorbance of free vanadate in this region of the UV spectrum. When corrected for this, there is a distinct break at r=2, although, as noted by Chasteen *et al.* (1986a), the graph beyond this point is not completely flat. This can been attributed to additional V(V) binding at sites other than the specific metal binding sites.

The nature of the transferrin-bound V(V) could play an important role in the binding of the metal ion. Harris & Carrano (1984) concluded that at a pH greater than 9, V(V) probably existed as the dioxovanadium(V) ion ($VO_2^+$), while at a lower pH (between 6 and 7), one oxygen atom is protonated. In either case, there are two oxygen atoms covalently bound to the vanadium in a *cis* arrangement (Vilas-Boas & Costa Pessoa, 1987). This could account for the observed binding of V(V) to hTf in the absence of a suitable anion, where the *cis* oxygens fill the role of the anion (Harris & Carrano, 1984).

In the structure of $Fe_2$Lf, the carbonate anion binds in a bidentate fashion to the iron, occupying two octahedral positions separated by about 60° (i.e. a *cis* arrangement). It seems reasonable to assume, therefore, that the $VO_2^+$ ion will occupy essentially the same position as the $Fe^{3+}$, interact with the same four protein ligands but not require a carbonate anion as the two *cis* oxygen atoms attached to the vanadium would serve to bridge between the positively-charged anion pocket and the metal centre and hence stabilise the complex. Preliminary modelling studies suggest that it is possible a water molecule could fit between the $VO_2^+$ ion and the residues comprising the anion pocket, in roughly the same position as the third carbonate oxygen (Figure 4.30b, next page).

It is clear from ESR studies on the transferrins and other biological molecules (Chasteen, 1981), that the form of the bound vanadium(IV) ion is $VO^{2+}$. Of the vanadium(IV) species present in aqueous solution (see Equations 4.6, 4.7 and 4.8 - page 216), only the $VO^{2+}$ and the $VO(OH)_3^-$ ions are ESR active, the former stable at a pH below 3.5 in the absence of chelating agents, while the latter does not form until pH 11 (Chasteen, 1981). When bound

to the transferrins at physiological pH, the characteristic 7-line $VO^{2+}$ spectrum is observed, with any excess vanadyl being ESR silent.

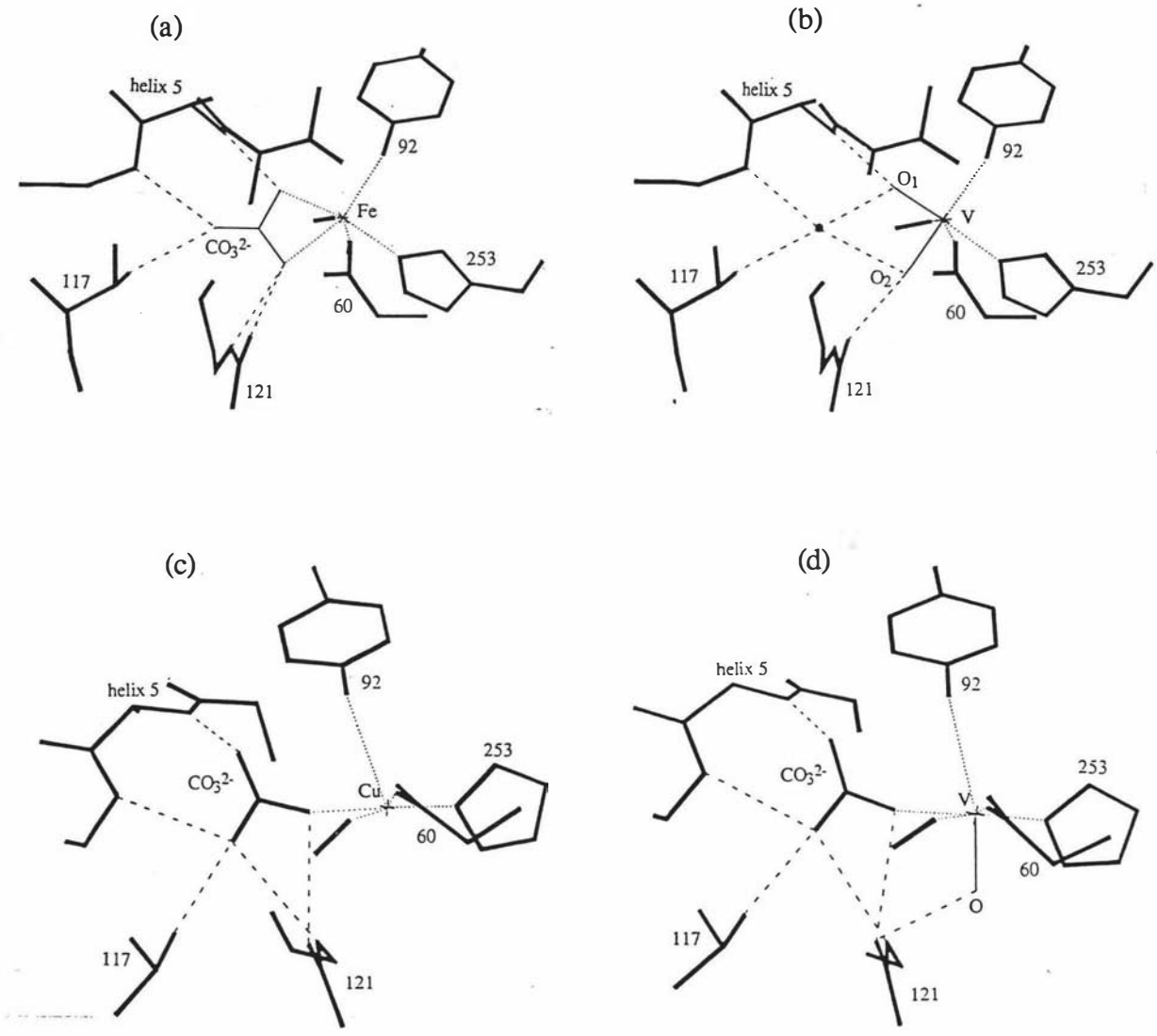



Figure 4.30: Schematic representations of a) the N-lobe metal binding site of $Fe_2Lf$ with an $Fe^{3+}$ and carbonate ion in their respective positions; b) $Fe_2Lf$ with a $VO_2^+$ moiety modelled into the position normally occupied by $Fe^{3+}$ and the carbonate anion. In addition, a water molecule is modelled in the position of the third (non-coordinated) carbonate oxygen to complete the hydrogen bonding pattern to the protein; (c) the N-lobe metal binding site of $Cu_2Lf$ with a $Cu^{2+}$ and monodentate carbonate ion in their respective positions; and (d) a $VO^{2+}$ ion modelled into the position normally occupied by $Cu^{2+}$, with the monodentate anion in virtually the same position. The vanadyl oxygen makes two hydrogen bonds, one with the Arg121 sidechain and one with the solvent network (not shown).

$$VO^{2+} + H_2O \overset{K}{\rightleftharpoons} VOOH^{+} + H^{+} \qquad K = 4.4 \times 10^{-6}\ M \qquad [4.6]$$

$$2VO^{2+} + 2H_2O \overset{K}{\rightleftharpoons} (VOOH)_2^{2+} + 2H^{+} \qquad K = 1.5 \times 10^{-7}\ M \qquad [4.7]$$

$$VO^{2+} + 2OH^{-} \overset{K_{sp}}{\rightleftharpoons} VO(OH)_{2\,(s)} \qquad K_{sp} = 1.1 \times 10^{-22}\ M^3 \qquad [4.8]$$

The binding of a $VO^{2+}$ ion to the transferrins (Figure 4.30d, previous page) can be modelled in the N-lobe site of $Cu_2Lf$ (Figure 4.30c) with a monodentate carbonate, the vanadium atom occupying the same position as the copper. This results in a 6-coordinate geometry with very little disruption to the surrounding structure. The hydrogen bonding interactions to the anion are almost identical to those seen with the 5-coordinate copper. In addition, the vanadyl oxygen, which occupies the 6th position (vacant in the case of copper), could possibly interact via hydrogen bonds to the Arg121 sidechain, as indicated in Figure 4.30d, or to solvent molecules.

It is interesting to note that vanadyl complexes typically have square pyramidal or bipyramidal structures, with the vanadium atom lying more or less in the equatorial plane and the vanadyl oxygen occupying an apical position (Vilas Boas & Costa Pessoa, 1987). In this respect, the $VO^{2+}$ model is consistent with the observed geometry in the N-lobe of $Cu_2Lf$, as the vanadyl oxygen is *trans* to the $O_\eta$ of Tyr92, which is in an apical position in $Cu_2Lf$. It is conceivable, therefore, that the replacement of either iron with a vanadyl ion would not significantly alter the interaction of the metal binding ligands or the amount of closure of the domains, as observed when copper is substituted.

## 4.5 Results;B; *f*-block elements

### 4.5.1 *Electronic spectra:*

All of the lanthanide-transferrin complexes previously prepared (Luk, 1971; Pecoraro *et al.*, 1981; Harris, 1986b) and the lactoferrin complexes prepared in this study, are colourless. The $Ln^{3+}$-Lf[§] complexes formed show strong bands in the UV near 245 and 290 nm upon the addition of the metal ions but no charge transfer bands are observed in the visible region.

However, an interesting result was obtained with cerium. Initially added as $Ce^{3+}$, this metal ion produced a colourless complex, similar to the other lanthanide ions. The visible absorption difference spectrum of this complex had a band at about 340 nm, tentatively assigned as a phenolate-cerium LMCT transition (Figure 4.31a), with a second band near 320 nm (appearing as a shoulder on the intense $\pi - \pi^*$ transition at ≈300 nm), also assigned as a phenolate-cerium(III) charge transfer band. Upon opening to the air, it was observed that over a period of several days the colour of the solution slowly changed to a deep red/brown with a $\lambda_{max}$ at about 442 nm (Figure 4.31b).

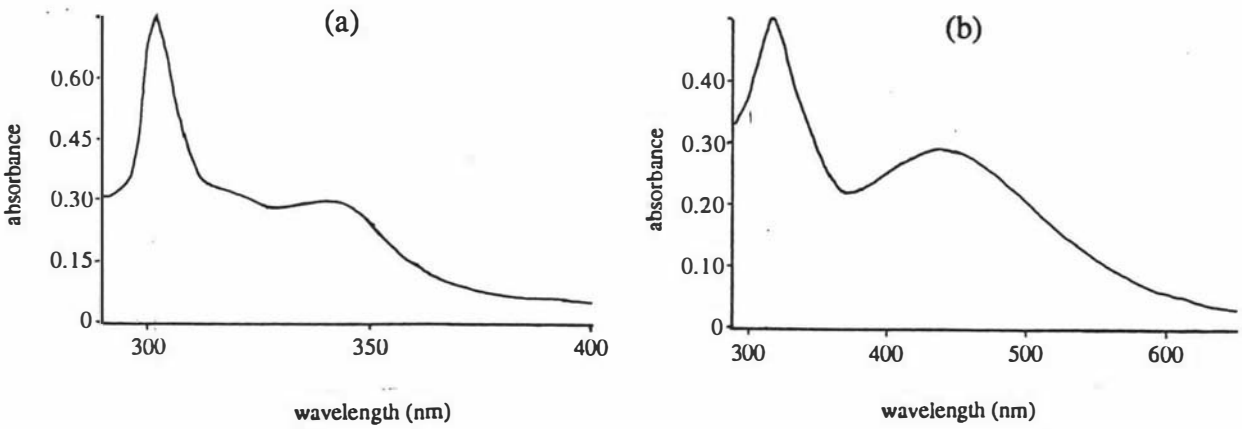



Figure 4.31: Electronic absorption spectrum of (a) $Ce(III)_2Lf$ in the range 290 to 400 nm and (b) $Ce(IV)_2Lf$ in the range 290 to 650 nm.

Absorbances for a 1% solution of this complex were calculated from known protein concentrations, assuming a 2:1 stoichiometry. The values obtained, 12.5 at 280 nm and 0.58

§ Ln or $Ln^{3+}$ are used when describing the interactions of the elements of the 4*f* block as a whole.

(4640 L $mol^{-1}$ $cm^{-1}$) at 442 nm, correspond closely to those reported by Ainscough *et al.* (1979) for four transition metal complexes of Lf. The results are summarised in Table 4.5 and corresponding values for the coloured transition metal complexes are given for comparison.

Table 4.5: Electronic absorption spectral data for complexes of human lactoferrin [a]

| complex | $A^{1\%}$ (280 nm) [b] | $\lambda_{max}$ (visible) (nm) | $A^{1\%}$ (visible) [c] | extinction coefficient, ε [d] |
|---|---|---|---|---|
| $Ce(IV)_2Lf$ | 12.5 | 442 | 0.58 | 4640 |
| $Fe_2Lf$ | 14.3 | 465 | 0.52 | 4160 |
| $Cu_2Lf$ | 12.5 | 434 | 0.61 | 4880 |
| $Co_2Lf$ | 13.6 | 405 | 1.28 | 10240 |
| $Mn_2Lf$ | 14.7 | 435 | 1.19 | 9520 |
| | | 520 | 0.60 | 4800 |
| | | 640 | 0.28 | 2240 |
| $Cr_2Lf$ | 12.2 | 442 | 0.064 | 510 |
| | | 612 | 0.035 | 280 |

[a] $A^{1\%}$ and $\lambda_{max}$ values for $Co_2Lf$, $Mn_2Lf$ and $Cr_2Lf$ were taken from Ainscough *et al.* (1979). [b] The corresponding value for apoLf is 10.9. The observed increase in the absorbance of this band is due to the increase in absorbance at 290 nm resulting from tyrosine deprotonation. [c] These values relate to the specific absorption band(s) in the visible region of the spectrum. [d] Assuming a molecular weight of 80,000 for Lf. Note that these values differ slightly from those given by Ainscough et al. (1979) as in this earlier work a molecular weight of 81,000 was assumed.

To test the possibility of an oxidation from $Ce^{3+}$ to $Ce^{4+}$, similar to that seen for $Mn^{2+}$ and $Co^{2+}$ bound to hTf (Feeney & Komatsu, 1966) and Lf (Ainscough *et al.*, 1979), a solution of apolactoferrin was degassed by alternately evacuating and flushing with $O_2$-free argon. The $HCO_3^-$ concentration was restored to 0.01M and $Ce^{3+}$ was added. After standing for 2 weeks at 4°C, the solution was still colourless. Upon exposure to the air, the complex turned red/brown within two days. A Ce(III)-lactoferrin complex prepared in a similar way was treated with a small aliquot of 5% $H_2O_2$ solution after standing sealed for two weeks and the formation of the red/brown colour was observed instantly. The visible absorption difference spectra of the coloured complexes formed in both of these experiments were indistinguishable from that shown in Figure 4.31b. The complex with $\lambda_{max}$ at 442 nm is, therefore, the result of the oxidation of Ce(III) to Ce(IV) in the presence of lactoferrin.

The stability of the coloured Ce(IV)-lactoferrin ($Ce_2Lf$) complex towards pH was investigated by carefully titrating a sample of $Ce_2Lf$ with aliquots of 0.1 M HCl, measuring the new pH after each addition. The results are given in Figure 4.32 as absorbance (442 nm) and the % saturation against pH. The absorbance of the charge transfer band decreases as the pH is lowered in such a way as to suggest that the cerium is being removed from the metal binding sites in what appears to be biphasic process, there being a clear break at about 50% saturation. It is known that the C-terminal site is more acid stable than the N-terminal site, with a biphasic metal release profile observed for iron (Mazurier & Spik, 1980) and it would appear that this is also the case for Ce(IV).

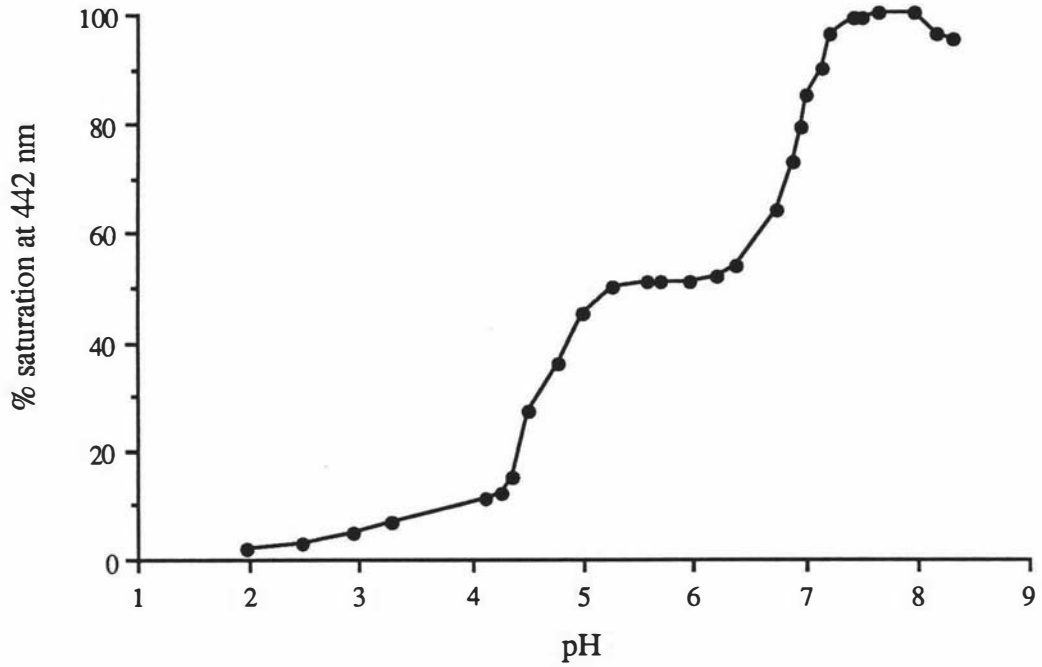



Figure 4.32: pH mediated cerium release for $Ce(IV)_2Lf$ in 0.03 M $HCO_3^-$. The initial pH of 8.32 was decreased by the addition of small increments (≈ 1 μL) of 0.1 M HCl to 3 mL of the metal-protein complex. The solution was stirred thoroughly and the pH measured with an Orion pH meter. The absorbance values were taken from the absorption spectrum recorded between 350 and 550 nm using Tris-HCl at pH 8.32 as the reference.

*Spectrophotometric titrations with the lanthanide ions and thorium(IV):* Difference spectra resulting from the addition of a number of the lanthanide ions ($La^{3+}$, $Ce^{3+}$, $Pr^{3+}$, $Nd^{3+}$, $Sm^{3+}$, $Eu^{3+}$, $Gd^{3+}$, $Tb^{3+}$, $Ho^{3+}$ and $Yb^{3+}$) to apoLf are represented by those of $Pr^{3+}$ and $Gd^{3+}$, as shown in Figures 4.33a & b (N.B. difference spectra for $Eu^{3+}$ and $Yb^{3+}$ addition

to apoLf are given in Figures 4.35a & b in section 4.5.2). Absorption bands in the UV at about 245 nm and 295 nm, characteristic of metal ion coordination to tyrosine residues (Pecoraro *et al.*, 1981; Harris, 1986b) were seen to increase upon the addition of each aliquot of metal ion solution. The spectra obtained for all the 4*f* metal ions are similar to one another, with strong, sharp peaks observed at 245 and 295 nm,.

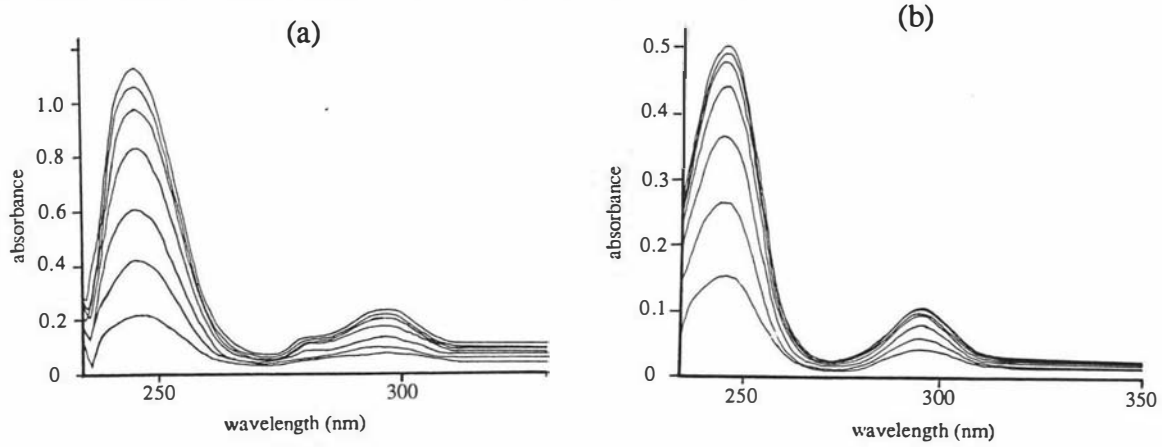



Figure 4.33: UV difference spectra of apoLf upon the addition of a) 2.32 mM $Pr^{3+}$ and b) 1.593 mM $Gd^{3+}$. Protein concentrations were 27.5 μM for Pr and 10.84 μM for Gd.

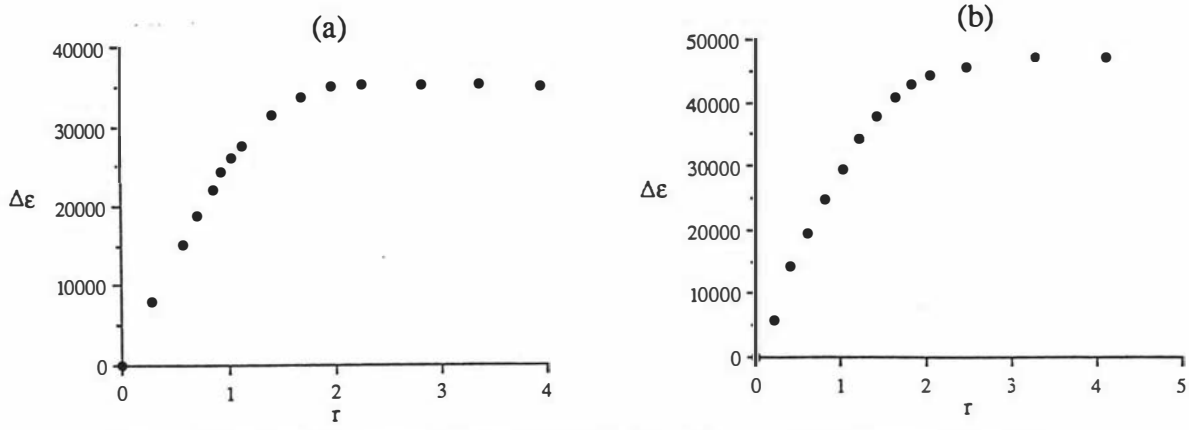



Figure 4.34: Titration curves for a) 1.07 mM $Pr^{3+}$ addition to 12.25 μM apoLf and b) 1.593 mM $Gd^{3+}$ addition to 10.84 μM apoLf.

Titration curves derived from these difference spectra are shown in Figures 4.34a & b for $Pr^{3+}$ and $Gd^{3+}$ as plots of Δε (244 nm) against the metal:protein ratio, r (those for the other lanthanide ions - $La^{3+}$, $Nd^{3+}$, $Sm^{3+}$, $Eu^{3+}$, $Tb^{3+}$, $Ho^{3+}$ and $Yb^{3+}$ - are given later in section 4.5.2, while $Ce^{3+}$ is discussed in Section 4.5.3). The use of Δε produces a normalised titration curve, enabling a direct comparison between results obtained from subsequent runs with either the same or different metal ion solutions. Both $Pr^{3+}$ and $Gd^{3+}$ bind to Lf in 2:1

ratios and the intensity of the difference peak at 245 nm indicates that these metal ions are interacting with at least two tyrosine residues in each of the two metal binding sites.

*Fluorescence spectroscopy:*

The titration curve obtained from cerium(III) quenching of the intrinsic apoLf fluorescence (Figure 4.35) indicates that between 1 and 2 metal ions are bound per lactoferrin molecule with a total quenching of about 37%. This indicates a poorer overlap of the excited tryptophan orbitals with the metal-tyrosinate molecular orbital and a smaller critical energy transfer distance. As in the case of $Fe^{3+}$ and $Cu^{2+}$, quenching occurs in a non-linear fashion consistent with the binding of the first equivalent of $Ce^{3+}$ preferentially into one of the metal binding sites (by analogy with $Fe^{3+}$ and $Cu^{2+}$, it is most propbably the C-site).

Fluoresence titrations were also attempted with $Tb^{3+}$ and $Th^{4+}$ and only very slight (≈5 - 10%) quenching was observed, consistent with the weak quenching reported for $Tb^{3+}$ with hTf (Luk, 1971), resulting from a lower value for the critical transfer distance for metal ions which do not exhibit LMCT transitions in the 300 - 400 nm region when bound to the transferrins.

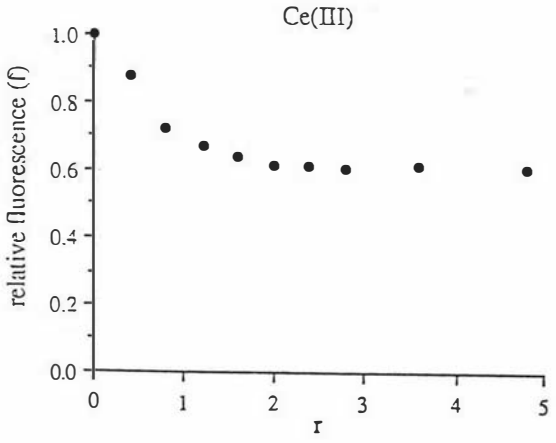



Figure 4.35: Fluorescence titration curve resulting from the addition of 9.52 mM $Ce^{3+}$ to 11.6 μM apoLf. The sample was excited at 288 nm with a 10 nm bandpass and the emission monitored at 340 nm, bandpass 4 nm.

The final solution following titration of apoLf with $Ce^{3+}$ (with a Ce:protein ratio of about 4:1) was left in the fluorescence cell exposed to air for 4 days until it was observed that a faint pink colour had developed. The emission spectrum was recorded after this time using

instrument setting identical to those employed during the titration. The amount of quenching of the original apoLf fluorescence had increased to about 50%, consistent with the increase in quenching on going from Mn(II) to Mn(III) (Ainscough *et al.*, 1980).

### 4.5.2 *Binding Constants:*

The addition of $La^{3+}$, $Nd^{3+}$, $Sm^{3+}$, $Eu^{3+}$, $Tb^{3+}$, $Ho^{3+}$ and $Yb^{3+}$ to apolactoferrin solutions at pH 7.4 generated families of spectra represented by those obtained for $Eu^{3+}$ and $Yb^{3+}$ (Figures 4.36a & b). The results of these UV spectrophotometric titrations are shown in Figures 4.37a - g (filled circles) as plots of $\Delta\varepsilon$ (244 nm) against r. Table 4.6 (two pages over) lists the results of titrations of transition and lanthanide metals with Lf.

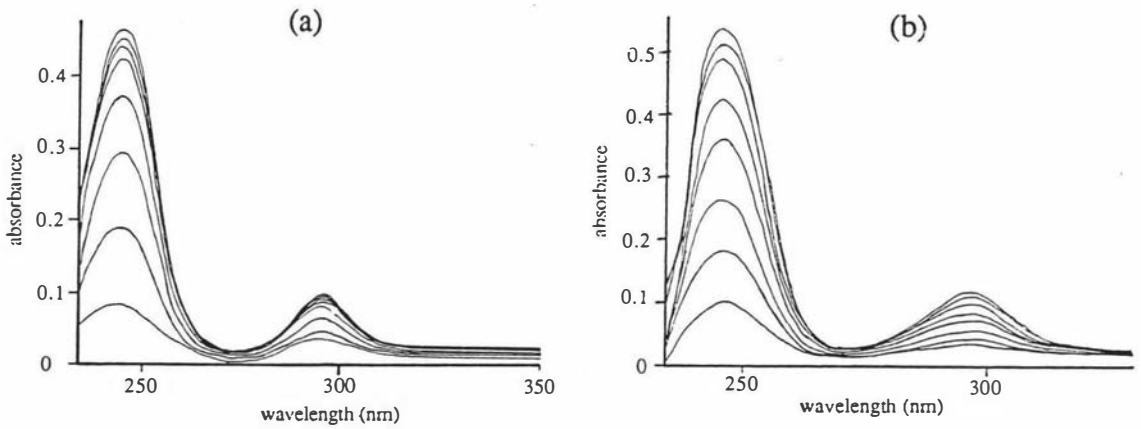



Figure 4.36: UV difference spectra of apoLf upon the addition of a) 0.810 mM $Eu^{3+}$ and b) 0.809 mM $Yb^{3+}$. The protein concentration for both titrations was 7.705 µM.

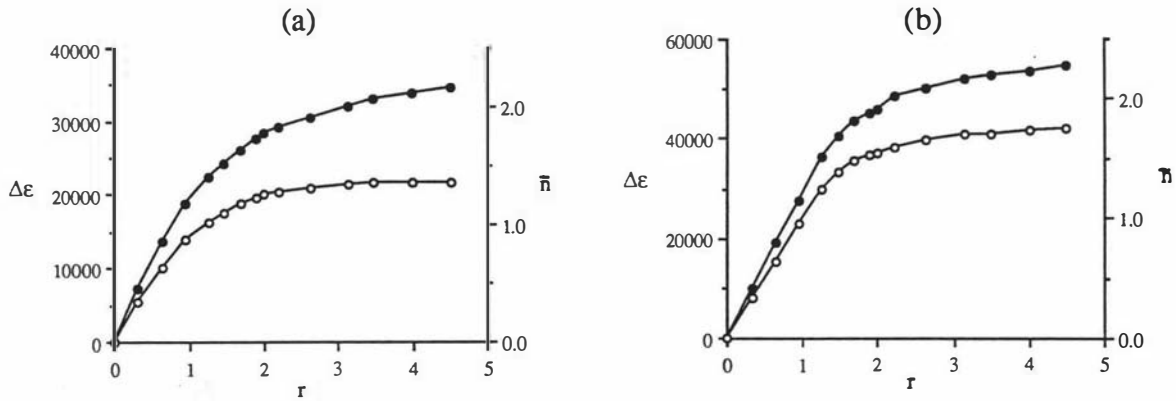



Figure 4.37: UV difference spectrophotometric titration curves for the addition of a) 0.810 mM $La^{3+}$ and b) 0.910 mM $Nd^{3+}$. The results are plotted as $\Delta\varepsilon$ at 244 nm versus the metal:protein ratio, r (•—•—•) and $\bar{n}$ vs r (o-o-o) where $\bar{n}$ = the average number of metal ions bound per protein molecule. Protein concentrations were 7.705 µM for both (continued overleaf).

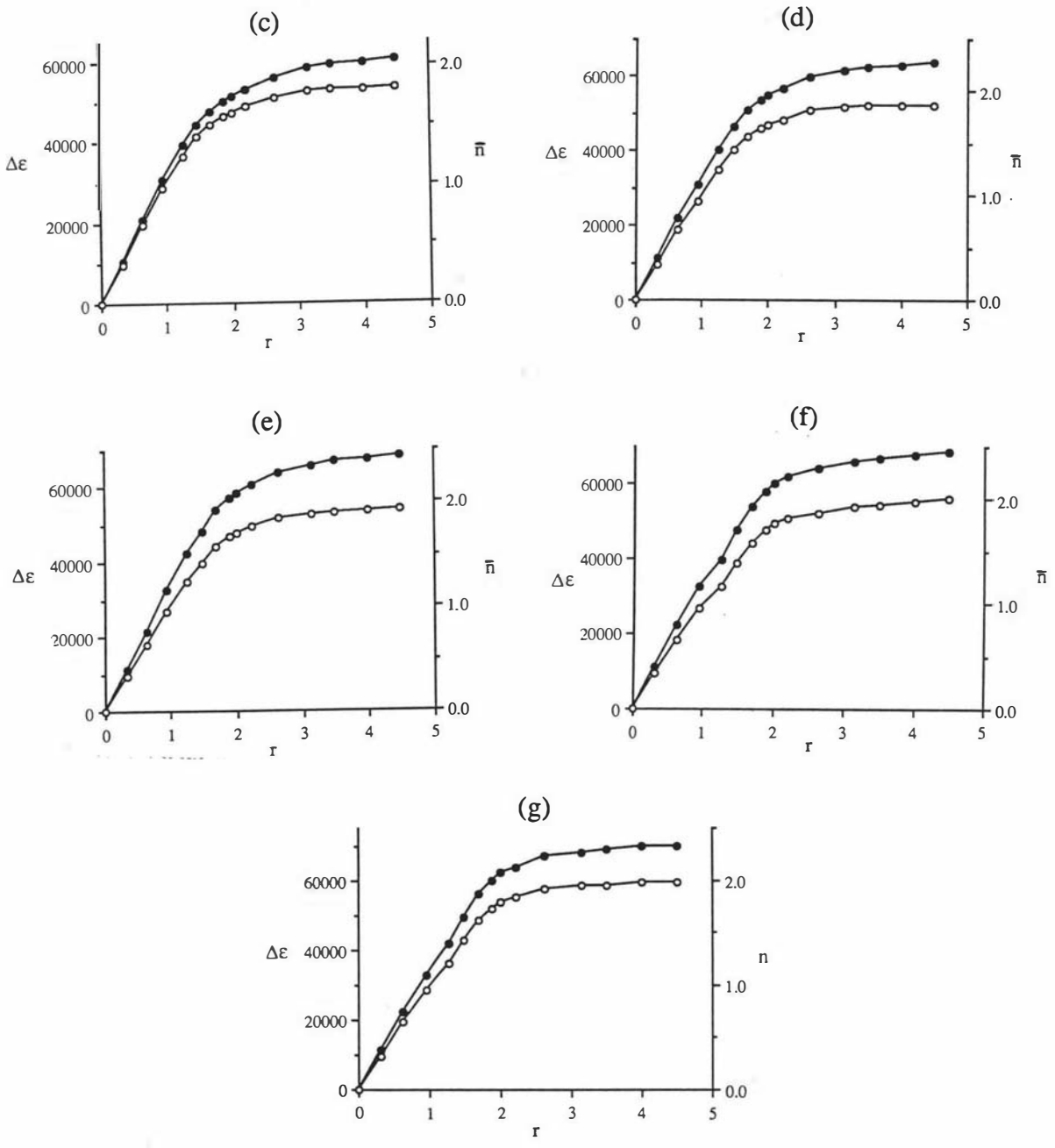



Figure 4.37 continued: UV difference spectrophotometric titration curves for the addition of c) 0.811 mM $Sm^{3+}$, d) 0.811 mM $Eu^{3+}$, e) 0.810 mM $Tb^{3+}$, f) 0.810 mM $Ho^{3+}$ and g) 0.809 mM $Yb^{3+}$. The results are plotted as Δε at 244 nm versus the metal:protein ratio, r (•—•—•) and n̄ vs r (o-o-o). Protein concentrations were all 7.705 μM.

All the lanthanide ions studied, with the exception of $La^{3+}$, gave Δε vs r plots which were linear between r=0 and r=1 and from about r=2 onwards. Between r=1 and r=2, the titration plots curve downwards to join the two linear portions, indicative of weak binding of the metal ions to the protein (Harris, 1986 ). From the first linear part (r=0 to r≈1) of the

titration curves, $\Delta\varepsilon_1$ can be estimated from the gradient and is defined as the molar absorptivity of a 1:1 metal:lactoferrin complex. Although the $\Delta\varepsilon_1$ values for some of the $Ln^{3+}$ ions are consistent with earlier values reported by Chung (1985) for Lf at ambient bicarbonate concentration, (21,560 and 29,120 $M^{-1}$ $cm^{-1}$ for $La^{3+}$ and $Nd^{3+}$ in the present study compared with 21,100 and 29,800 $M^{-1}$ $cm^{-1}$ as reported previously), the majority are somewhat different, being greater than those from the earlier study.

In addition, the $\Delta\varepsilon_1$ values for $Nd^{3+}$ and $Sm^{3+}$ binding to Lf are about 10,000 $M^{-1}$ $cm^{-1}$ greater than those obtained with human serum transferrin (Harris, 1986b)[§]. This may be due to differences between the two proteins, for example in the orientation of the metal-binding tyrosine sidechains which could result in differences in the overlap of the tyrosinate $\pi$ orbitals with the *d* orbitals of the metal.

Table 4.6 : Results of spectrophotometric titrations of Lf with various metal ions.

| Metal ion | Average $\Delta\varepsilon_1$ at 245 nm ($M^{-1}$ $cm^{-1}$) | Average $\Delta\varepsilon_{tot}$ at 245 nm ($M^{-1}$ $cm^{-1}$) | r [a] | number of titrations |
|---|---|---|---|---|
| $Fe^{3+}$ | 12 600 | 21 500 | 2 | 7 |
| $Cu^{2+}$ | 9 850 | 17 500 | 2 | 6 |
| $V^{3+}$ | 19 580 | 28 700 | 2 | 2 |
| $VO^{2+}$ | 11 350 | 17 300 | 2 | 4 |
| $VO_2^{+}$ | 6 700 | 11 300 | 2 | 5 |
| $Al^{3+}$ | 15 900 | 27 000 | 2 | 4 |
| $La^{3+}$ | 21 560 | 28 280 | 1 | 4 [b] |
| $Ce^{3+}$ | 25 750 | 35 900 | 2 | 3 |
| $Pr^{3+}$ | 24 900 | 34 900 | 2 | 1 |
| $Nd^{3+}$ | 29 120 | 45 930 | 2 | 4 [b] |
| $Sm^{3+}$ | 31 720 | 51 270 | 2 | 4 [b] |
| $Eu^{3+}$ | 32 470 | 54 620 | 2 | 4 [b] |
| $Gd^{3+}$ | 28 700 | 44 550 | 2 | 3 |
| $Tb^{3+}$ | 33 990 | 58 190 | 2 | 4 [b] |
| $Ho^{3+}$ | 28 600 | 60 250 | 2 | 4 [b] |
| $Yb^{3+}$ | 29 650 | 62 200 | 2 | 4 [b] |
| $Th^{4+}$ | 22 900 | 39 000 | 2 | 2 |

[a] Defined as the ratio of the moles of metal ion added to the moles of apoLf initially present. [b] Values are the result of 2 titrations performed on each of 2 different preparations of apoLf

§ It should be noted that $Nd^{3+}$ and $Sm^{3+}$ are the only two $Ln^{3+}$ ions in common between the work of Harris (1986b) and this present study.

Values of $\Delta\varepsilon_{tot}$ are estimated from the curve at r=2 and $\Delta\varepsilon_2$ is the difference between $\Delta\varepsilon_{tot}$ and $\Delta\varepsilon_1$. The $\Delta\varepsilon_2$ values obtained are all less than the corresponding $\Delta\varepsilon_1$. This has been attributed to competition for the second equivalent of metal ion by excess carbonate and it has been shown in the case of $Nd^{3+}$ that decreasing the bicarbonate concentration leads to an increase in the $\Delta\varepsilon_{tot}$ (Harris, 1986b). It was for this reason that the only source of bicarbonate anions ultimately used in the present metal-binding study was that arising from atmospheric $CO_2$. Carbonate cannot be eliminated completely of course, as it is required for metal binding.

The $\Delta\varepsilon$ vs r plots were converted to $\bar{n}$ vs r (Figures 4.37a - g; open circles) where $\bar{n}$ is defined as the average number of metal ions bound per lactoferrin molecule, as given by equation [4.3] and is the ratio of the observed extinction coefficient to the extinction coefficient for a 1:1 metal:Lf complex. By running duplicate titrations with the seven metal ions on protein solutions from the same preparation, the error in the $\bar{n}$ values from the titrations with different metals has been kept to a minimum. The only source of experimental error is in the measurement of the absorbances and in the determination of the concentrations of the metal ion solutions. Therefore, it can be assumed that the $\bar{n}$ values from different titrations with the same protein solution should be quite reliable with respect to one another. It is interesting to note that the values of $\bar{n}$ found for Lf are very similar to recent results obtained for lanthanide binding to hTf (Harris & Chen, 1991), where the $\bar{n}$ values ranged from 1.4 for the larger lanthanides to 1.9 for the smaller metal ions.

These curves also have two linear portions with a break near r=2. At r=1, all the lanthanide ions studied have $\bar{n}$ values near 1.0, indicating essentially quantitative binding of the first equivalent of metal ion. However, at r=2, the $\bar{n}$ values range from 1.25 ($La^{3+}$) to 1.8 ($Yb^{3+}$), reflecting competition by carbonato complexes and hydrolysis reactions, as well as the differences in the second association constant, $K_2$. A value for $K_2$ can be derived from the mass-balance relationship given by equation [4.3]. Table 4.7 lists the relevant $\Delta\varepsilon$ values and the derived values of $\bar{n}$ and $\log K_2$.

Table 4.7: Ionic radii and estimated binding constants for the second equivalent of a number of $Ln^{3+}$ ions to apoLf

| | $La^{3+}$ | $Nd^{3+}$ | $Sm^{3+}$ | $Eu^{3+}$ | $Tb^{3+}$ | $Ho^{3+}$ | $Yb^{3+}$ |
|---|---|---|---|---|---|---|---|
| Ionic radius (Å) [a] | 1.032 | 0.983 | 0.958 | 0.947 | 0.923 | 0.901 | 0.868 |
| $\Delta\varepsilon_1$ ($M^{-1}$ $cm^{-1}$) | 21560 | 29120 | 31720 | 32470 | 33990 | 34180 | 34680 |
| $\Delta\varepsilon_{tot}$ ($M^{-1}$ $cm^{-1}$) | 28280 | 45930 | 51270 | 54620 | 58190 | 60250 | 62200 |
| $\Delta\varepsilon_2$ ($M^{-1}$ $cm^{-1}$) | 6720 | 16810 | 19550 | 22150 | 24200 | 26070 | 27520 |
| $\varepsilon_M$ ($M^{-1}$ $cm^{-1}$) [b] | 1170 | 1770 | 1300 | 1200 | 1360 | 200 | 550 |
| r | 1 ? | 2 | 2 | 2 | 2 | 2 | 2 |
| $\bar{n}$ [c] | 1.25 | 1.55 | 1.6 | 1.67 | 1.70 | 1.76 | 1.79 |
| $\bar{n}_{max}$ [d] | 1.35 | 1.70 | 1.80 | 1.86 | 1.92 | 2.00 | 1.98 |
| $\log K_2$ | 4.76 | 5.55 | 5.69 | 5.90 | 6.00 | 6.24 | 6.38 |

[a] From Shannon, (1976). Values are for a coordination number of 6. [b] Calculated from the gradient of a $\Delta\varepsilon$ vs r plot of the metal ion solutions against buffer. In all cases the r values were calculated as if protein was present. [c] $\bar{n}$ = the average number of metal ions bound to the protein when 2 equivalents of metal ion have been added. [d] $\bar{n}_{max}$ = the final plateau value of $\bar{n}$ taken from Figures 4.37a - g.

There is a distinct correlation between the size of the lanthanide ions (represented by the ionic radius) and both n and $\log K_2$. As the ionic radius decreases ($La^{3+}$ - $Yb^{3+}$) the values for $\bar{n}$ and $\log K_2$ increase. In the case of the $\log K_2$ values, this increase is linear, as shown by Figure 4.38.

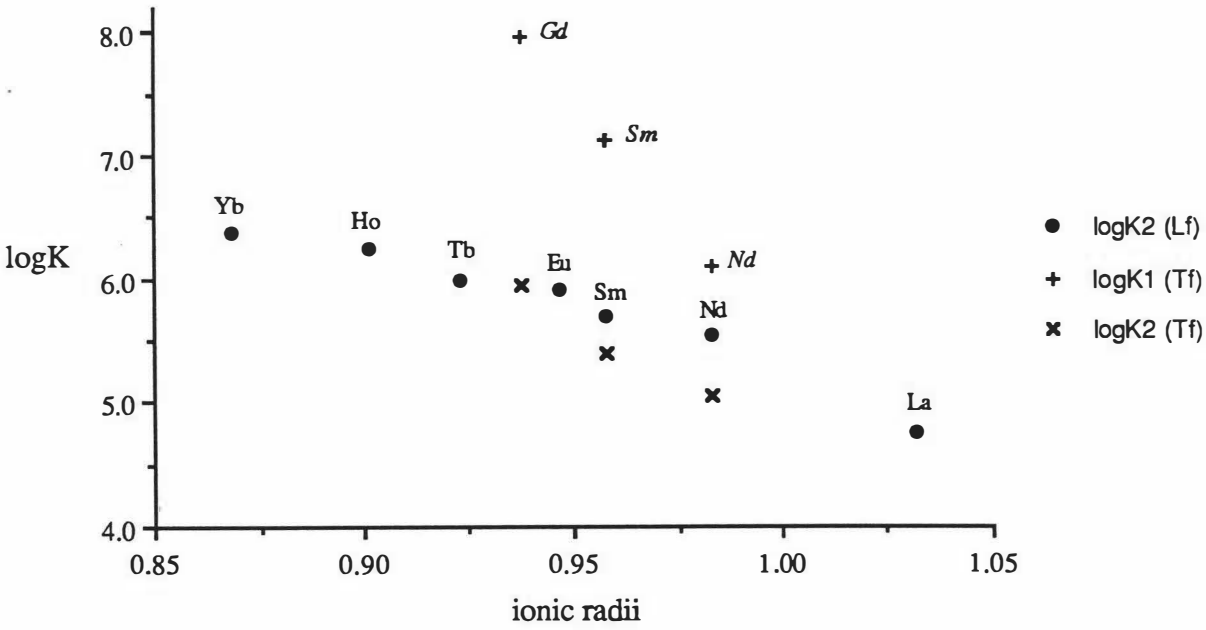



Figure 4.38: Plot of the estimated $\log K_2$ of the $Ln^{3+}$-Lf complexes against ionic radius. The $\log K_1$ and $\log K_2$ values for three $Ln^{3+}$-hTf complexes are shown as a comparison.

### 4.5.3 *The effects of anions on lanthanide binding:*

Spectrophotometric titrations of apoLf with cerium(III) appeared to be dependent on the method of protein purification used, at least in terms of the stoichiometry of metal binding. In some of the preparations, EDTA had been added in attempts to remove residual iron, while in others this step was omitted (see Section 4.2.7). Figure 4.39 shows the titration curves derived from the difference spectra generated by the addition of $Ce^{3+}$ to the "non-EDTA treated" apoLf (shown as • • •), to the "EDTA treated" protein after dialysis against two changes of Tris buffer (o o o) and after six changes of buffer (x x x). Clearly, the stoichiometry of binding is dependent upon the presence of EDTA, in that stoichiometries near 1:1 are obtained in the presence of trace quantities of EDTA. When the EDTA is completely removed from the protein, a stoichiometry of 2:1 is observed, similar to that seen with "non-EDTA treated" protein. It should be noted that there is essentially no alteration to the initial slope ($\Delta\varepsilon_1$) of the titration curve when EDTA is present or upon removal by exhaustive dialysis.

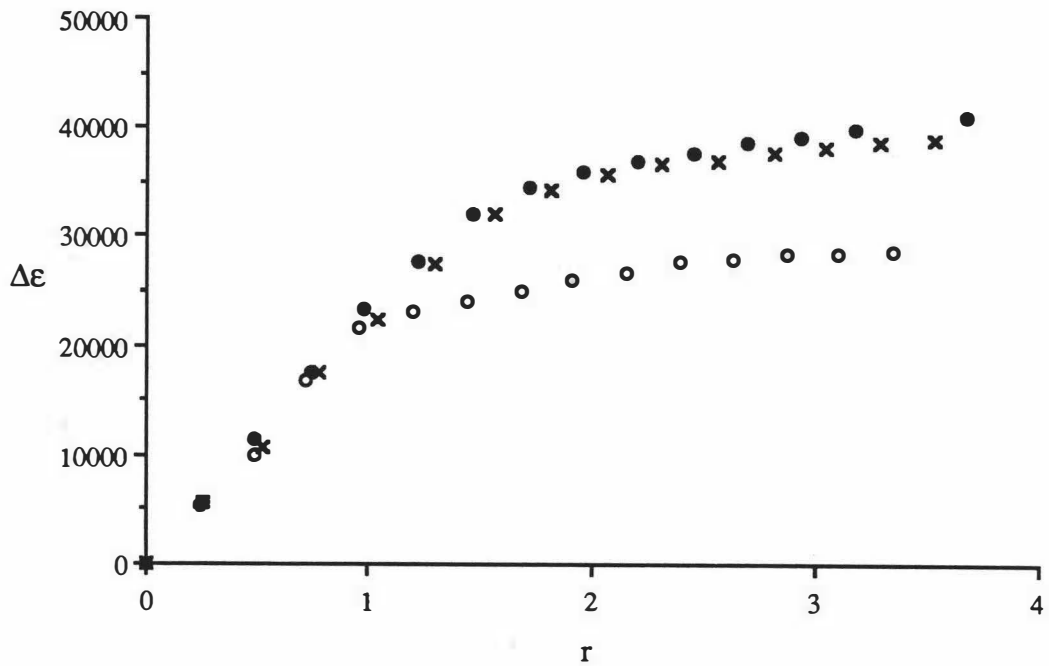



Figure 4.39: Titration curve resulting from the addition of $Ce^{3+}$ to "non-EDTA treated" apoLf (• • •), 'EDTA-treated' apoLf after two changes of Tris-HCl buffer (o o o) and after six changes of Tris-HCl buffer (x x x).

The effect of NTA as a competitive chelator can be clearly seen in the difference spectra for the addition of Th(IV) to apoLf. When the thorium was added as $Th^{4+}$ to apoLf in a Tris-HCl buffer containing 0.01 M in $HCO_3^-$, the graph shown in Figure 4.40 (filled circles) is obtained. The stoichiometry of binding appears to between 1 and 2 ($\Delta\varepsilon_1$ = 24000 and $\Delta\varepsilon_{tot}$ = 30000), resulting from competition by carbonate for the second equivalent of $Th^{4+}$. In fact, during the course of the experiment, it was noted that a precipitate was forming in the cuvette, which in hindsight can be attributed to insoluble $Th(CO_3)_2$.

In order to prevent this precipitation of thorium, NTA was added to the metal ion solution in a 2:1 ratio (using the protocol of Harris *et al.* (1981) who had reported that the presence of NTA did not compete with hTf for $Th^{4+}$) and the titration repeated. The resulting stoichiometry of the Th-Lf complex was 2:1 (Figure 4.40 - open circles) but with significantly lower values of $\Delta\varepsilon_1$ and $\Delta\varepsilon_{tot}$ (14000 and 26800 respectively). A third titration was attempted using $Th^{4+}$ and with ambient $CO_2$ as the only source of (bi)carbonate anions. Once again a 2:1 stoichiometry was achieved (Figure 4.40 - triangles) with extinction coefficients of 22900 and 39000 ($\Delta\varepsilon_1$ and $\Delta\varepsilon_{tot}$ respectively).

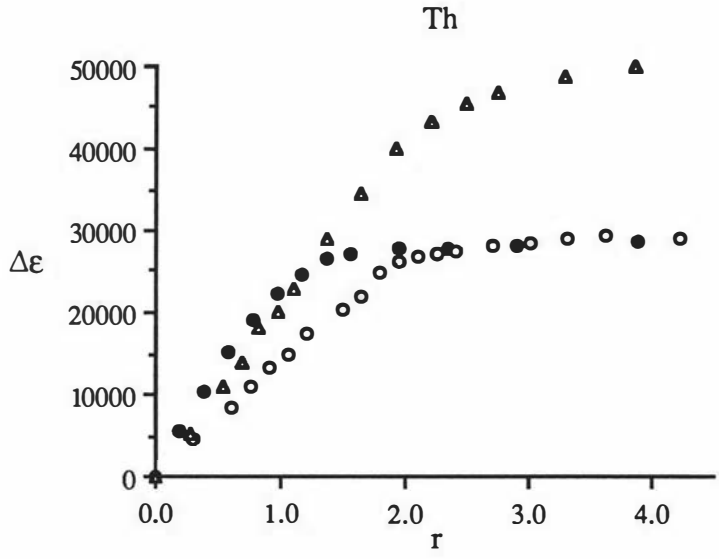



Figure 4.40: Titration curves resulting from the addition of $Th^{4+}$ to apoLf in a buffer containing 0.01 M $HCO_3^-$ (• • •), $Th(NTA)_2$ in the presence of ambient $CO_2$ (o o o) and $Th^{4+}$ at ambient $CO_2$ (Δ Δ Δ).

Clearly the presence of either carbonate or NTA has profound effects on thorium binding. To determine whether a similar effect occurred with the lanthanide ions, difference spectra of apoLf upon the addition of one equivalent of a series of $Nd^{3+}$:NTA and $Yb^{3+}$:NTA solutions were recorded (Figures 4.41a & b). When no NTA is present ($Ln^{3+}$:NTA ratio = 1:0; curve 1) the observed $\Delta\varepsilon$ (equivalent to $\Delta\varepsilon_1$ as only one metal ion is bound) is at a maximum. As the concentration of NTA in the metal ion solutions increases (1:0.5 - 1:5; curves 2 - 6), the value of $\Delta\varepsilon_1$ decreases. When this decrease in $\Delta\varepsilon_1$ is plotted as a function of the $Ln^{3+}$:NTA ratio (Figure 4.42), an approximately linear decrease results. A plateau is reached at a 1:3 metal:NTA ratio, possibly corresponding to the formation of stable tris-nitrilotriacetato neodymium(III) and ytterbium(III) complexes, where the addition of more chelator does not affect the absorption of the metal-protein complex.

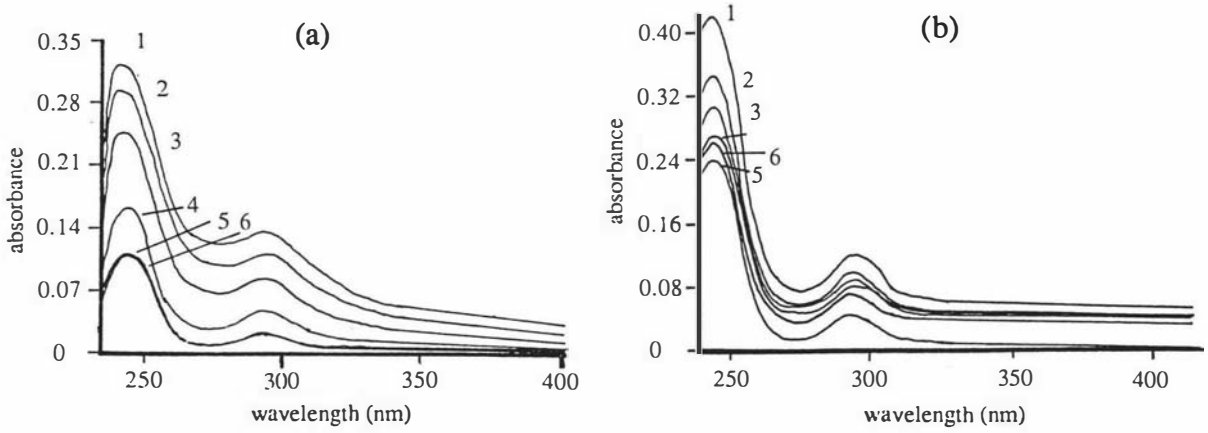



Figure 4.41: UV difference spectra resulting from the addition of one equivalent of a) $Nd(NTA)_x$ and b) $Yb(NTA)_x$ to apoLf, where x = 0, 0.5, 1, 2, 3 and 5. The curves are labelled 1 - 6. Note that in the case of Nd, curves 5 and 6 are coincident, while curve 4 is not shown for Yb.

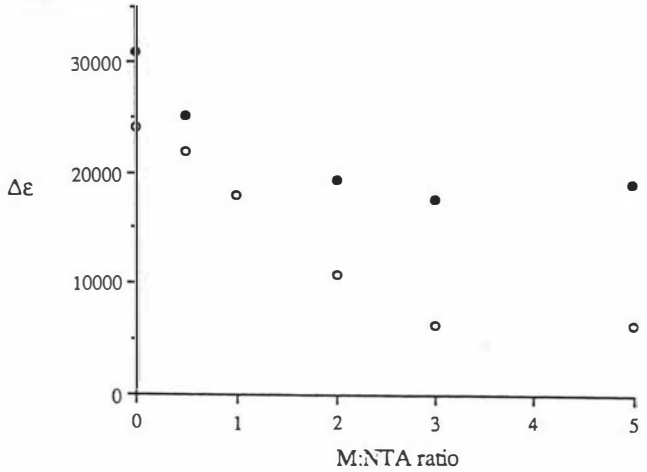



Figure 4.42: Plots of $\Delta\varepsilon$ as a function of the NTA concentration for $Nd(NTA)_x$ (o o o) and $Yb(NTA)_x$ (• • •).

4.5.4 *Electron spin resonance spectroscopy:*

The ESR spectra of $Gd^{3+}$-Tf (1:1 complexes) have been reported independently by two groups (O'Hara & Koenig, 1986; Zak & Aisen, 1988), the spectrum obtained by the latter group shown in Figure 2.13. The reported spectra are essentially identical and simulation calculations (O'Hara & Koenig, 1986) suggest that the $Gd^{3+}$ ions have a distorted, rhombic stereochemistry, similar to that proposed for $Fe^{3+}$ in $Fe_2$Tf (Aasa, 1970).

The spectrum for a 1:1 $Gd^{3+}$-lactoferrin complex is shown in Figure 4.43. The spectrum consists of a series of signals ranging from about g' =13 to g' =3 and appears to be similar to the ESR spectrum of $Gd^{3+}$-Tf complexes previously reported (O'Hara & Koenig, 1986; Zak & Aisen, 1988). It can best be described as being centred at g' = 4.94, with the major resonance extending about 380 - 400 Gauss either side. Superimposed on this are smaller resonances at g' = 4.62 and g' =5.02, while hyperfine structure, tentatively assigned as arising from the *x* component of the g-tensor (based on the calculations of O'Hara & Koenig, 1986), can be observed in the region from g' = 7 to g' = 13. Table 4.8 lists the major resonances, along with calculated g' values.

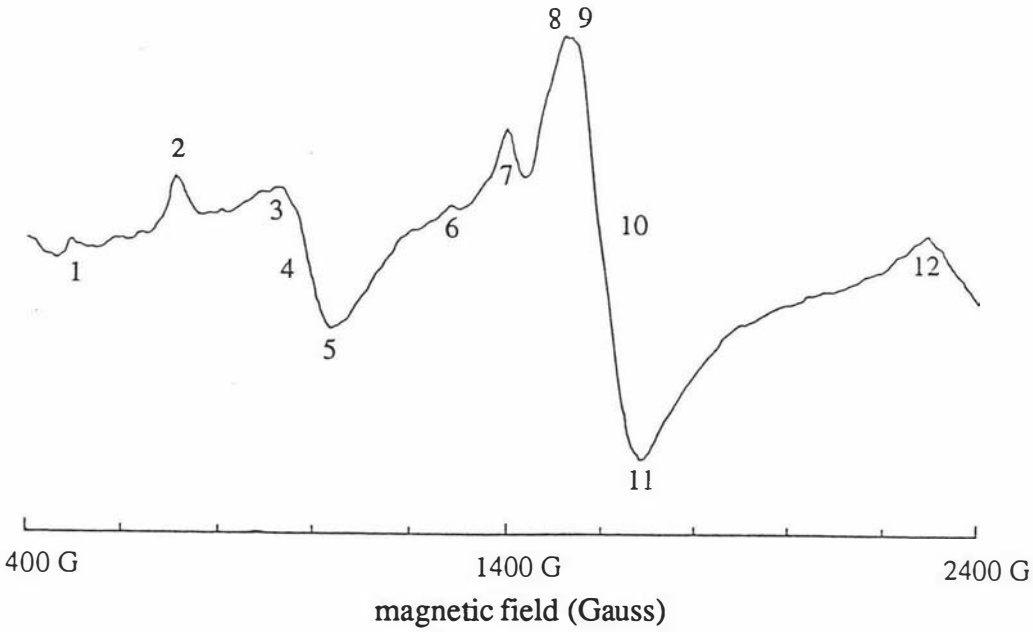



Figure 4.43: Electron spin resonance spectrum of a 1:1 Gd:Lf complex with the field set at 1400 G, scan range 2000 G. The sample temperature = -160°C. Microwave frequency is 8.995 GHz at 5 mW power and the receiver gain = $1.2 \times 10^4$. The values of g' indicated by numbers 1 - 12 are listed in Table 4.8 and were calculated using equation [4.1].

In addition to the 1:1 Gd-Lf complex, the ESR spectra of a series of gadolinium-lactoferrin complexes with varying metal:protein ratios were also recorded (Figure 4.44). The ratios used were 0.5:1, 1.5:1 and 2:1. It should be noted that at ratios greater than 0.5:1, a new band at g' = 4.55 appears. Clearly as the amount of metal ion present in the sample increases, the electronic environment of the two metal ions becomes slightly different.

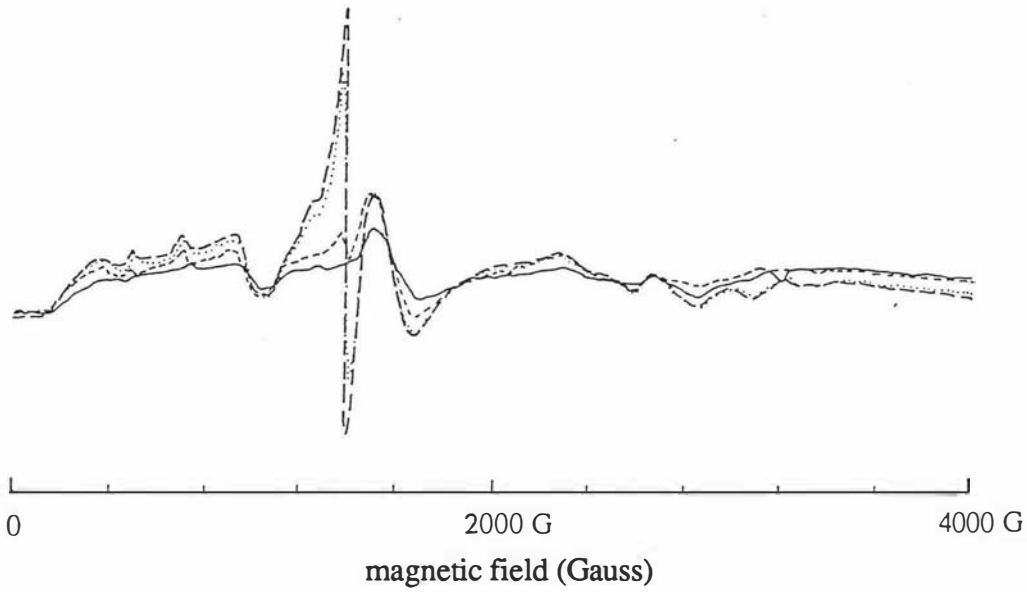



Figure 4.44: Electron spin resonance spectra of four Gd:Lf complexes with metal:protein ratios of 0.5:1 (——), 1:1 (- - -), 1.5:1 (·····) and 2:1 (— —). In all cases, the frequency was 8.990 GHz at 5 mW power with a receiver gain of $6.3 \times 10^3$. Sample temperature = -160°C.

Table 4.8: Major resonances and associated g' values for GdLf (1:1 metal:protein)

| position | resonant field B (Gauss) | observed g' |
|---|---|---|
| 1 | 495 | 12.97 |
| 2 | 715 | 9.04 |
| 3 | 915 | 7.02 |
| 4 | 990 | 6.48 |
| 5 | 1035 | 6.20 |
| 6 | 1280 | 5.02 |
| 7 | 1390 | 4.62 |
| 8 | 1515 | 4.24 |
| 9 | 1550 | 4.14 |
| 10 | 1595 | 4.02 |
| 11 | 1680 | 3.82 |
| 12 | 2290 | 2.80 |

## 4.6 Discussion; B: *f*-block elements

### 4.6.1 *Stoichiometry of binding:*

The stoichiometry of lanthanide binding to apoLf is, in all cases studied, 2:1. The titration curves, obtained from the sharp, intense UV difference spectra, all show a significant flattening beyond r=2, indicative of either negligible non-specific binding of the $Ln^{3+}$ ions or the absence of free $Ln^{3+}$ absorbance in the UV region.

Chung (1985) studied the binding of a variety of metal ions to both hTf and Lf, including a number of lanthanides ($La^{3+}$, $Pr^{3+}$, $Nd^{3+}$, $Sm^{3+}$, $Eu^{3+}$, $Gd^{3+}$, $Tb^{3+}$, $Ho^{3+}$ and $Er^{3+}$), and was able to estimate the stoichiometry of metal binding from the resultant difference spectra and titration curves. It was reported that in the case of Lf, that for the larger $Ln^{3+}$ ions ($La^{3+}$, $Pr^{3+}$ and $Nd^{3+}$), the stoichiometry was somewhat less than 2:1, although for those from $Sm^{3+}$ onwards, two metal ions bound per protein molecule. Re-examination of the results of Chung (1985) show that for $Pr^{3+}$ and $Nd^{3+}$, the interpretation of the titration curves was in error, in that to determine the stoichiometry of binding, best fit lines were drawn through the straight-line regions of the graph (between r=0 & r=1 and from r=2 to r=4) and the metal:protein ratio taken at the intercept between these lines. This method does not take into account the significant curvature of the titration curves between r=1 and r=2, attributed to weaker binding of the second equivalent of metal ion (Harris & Carrano, 1984). For $Pr^{3+}$ and $Nd^{3+}$, the curves of Chung (1985) do not completely flatten until r=2, indicating a stoichiometry near 2:1, as observed in the present study with these same metal ions.

### 4.6.2 *Electronic spectra of cerium-substituted lactoferrin:*

Perhaps the most interesting result comes from the studies on cerium(III) binding to Lf. Cerium(III) undergoes an oxidation to cerium(IV), facilitated either by atmospheric $O_2$ or by the addition of an oxidising agent such as $H_2O_2$. A shift in the position of an intense charge transfer band occurs, as observed in the oxidation of lactoferrin-bound Mn(II) ($\lambda_{max}$ = 310 nm) to Mn(III) ($\lambda_{max}$ = 435 nm), and an argument analogous to that for manganese can be used in an attempt to explain the results with cerium.

The manganese case, assuming regular octahedral symmetry, is shown diagrammatically in Figure 4.45. The left-hand diagram represents the *d*-orbital splitting of Mn(II), showing the ligand-metal charge transfer transition from tyrosinate pπ orbitals to a low-lying $t_{2g}$ Mn(II) orbital. Upon oxidation (right-hand diagram) the increase in charge on the metal ion draws the ligands closer, in turn perturbing the *d*-orbitals to a greater extent and causing an increase in the *d*-orbital splitting (Huheey, 1978). This results in a decrease in the energy of the charge transfer transition and a shift to higher wavelength of the charge transfer band.

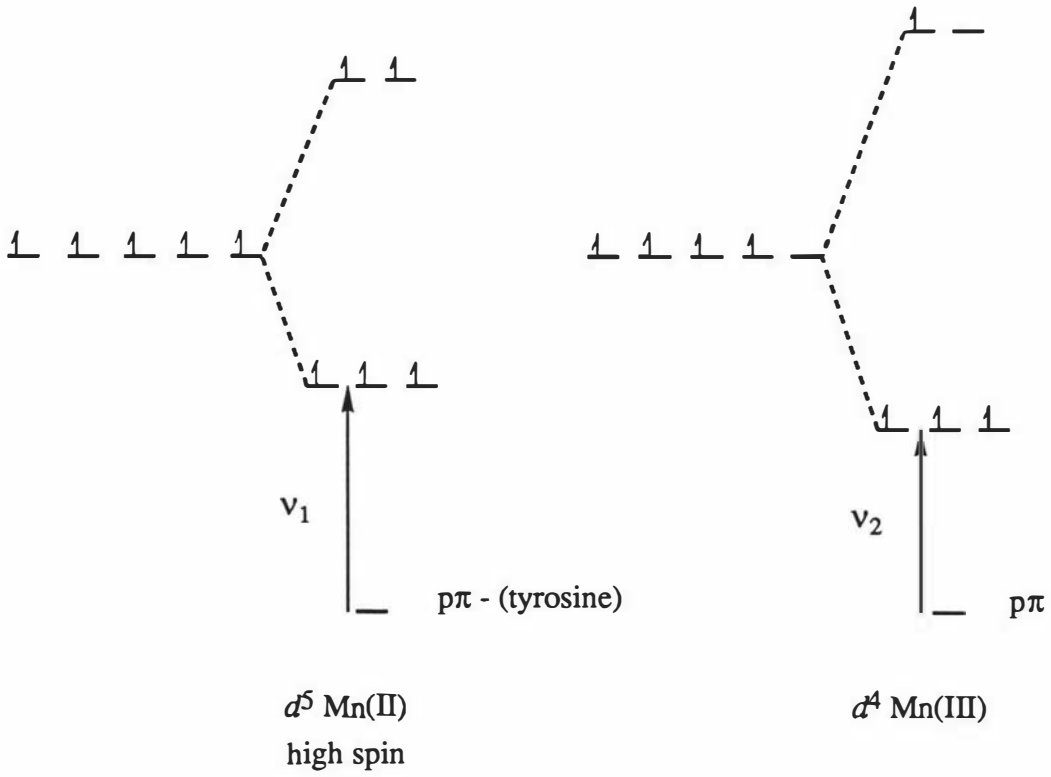



Figure 4.45: Crystal-field energy diagrams for high spin $d^5$ Mn(II) showing the LMCT transition from the pπ ligand orbital into a $t_{2g}$ orbital ($\nu_1$). Upon oxidation to $d^4$ Mn(III), the *d* orbital splitting is enhanced and the LMCT energy ($\nu_2$) is decreased such that $\nu_2 < \nu_1$ and the wavelength of the absorption band is increased.

In the case of cerium(III), with an electron configuration of $4f^1\ 5d^0\ 6s^0$, the 4*f* orbitals are the lowest-lying empty metal orbitals. However, the 5*d* orbitals are sufficiently close in energy to the 4*f* that three possibilities arise; (i) that the 340 nm transition results from a tyrosinate π molecular orbital into the lowest-lying 4*f* orbital; (ii) that the transition is from the ligand π to the $t_{2g}$ set of a purely *d*-character orbital or (iii) a certain amount of mixing of

the *d* and *f* orbitals could result, producing a metal-centred molecular orbital with partial *d* and *f* character.

In the purely *f*-orbital case, with the crystal-field split into three levels, $t_{1u}$, $t_{2u}$ and $a_{2u}$ (Figgis, 1987), the energy difference between the levels estimated is to be in the order of ≈ 100 $cm^{-1}$ (Moeller, 1973; Cotton & Wilkinson, 1988) due to the shielding of the 4*f* orbitals by the $5s^25p^6$ octet. Changes in the crystal field do not perturb the *f* orbitals to any great extent, once again as a consequence of the increased sheilding by the 5*s* and 5*p* orbitals, so that upon oxidation of Ce(III) to Ce(IV), the observed shift of 6800 $cm^{-1}$ cannot be accounted for if only the *f* orbitals are involved in the LMCT transitions.

It seems more reasonable to assume, therefore, that the band at 340 nm observed in Ce(III)-Lf is the result of a transition from the tyrosinate ligands into either a low-lying pure *d* orbital ($t_{2g}$) or a mixed *d/f* molecular orbital with mainly *d* character. Upon oxidation, the splitting of the metal-centred orbitals is increased (using the same argument as described above for the Mn(II)/Mn(III) case) leading to a lower energy transition (higher wavelength).

However, the presence of two bands in the difference spectrum of Ce(III)-Lf, at 320 nm and 340 nm, is difficult to explain. The difference in energy between the two bands is about 2,000 $cm^{-1}$. This is too small for the two bands to be the result of transitions from the ligand into the empty $t_{2g}$ and $^e_g$ sets, as in a purely *d*-electron metal complex where the observed splitting ranges from 12,000 to 45,000 (Huheey, 1978). However, for Ce(III) the effect of the lone 4*f*-electron is not known, and it is possible that this electron, in addition to those of the 5*s* and 5*p* orbitals (which are not normally occupied in transition metals), may be affording some shielding to the *d*-orbitals, thus lowering the splitting and resulting in two close transitions. On the other hand, the observed separation in the bands is too large to be the result of two ligand $\rightarrow f$ transitions. It is more likely that they are the result either from transitions into two *d/f* molecular orbitals close in energy, or perhaps two transitions of the type observed in $Cu_2Lf$, ie transitions from the $\pi_a$ and $\pi_b$ tyrosinate orbitals into a single *d* or *d/f* molecular orbital (Section 4.3.1). In the case of $Cu_2Lf$ the energy differences are in the order of 5000 $cm^{-1}$.

4.6.3 *Strength of $Ln^{3+}$ binding to lactoferrin:*

One result which is clearly evident in Table 4.6 is that the values of $\Delta\varepsilon_1$ for the lanthanide ions are significantly greater ($\approx$ 2 times) than the corresponding values for the transition metal ions. The intensity of the $\pi \rightarrow \pi^*$ transition responsible for the 245 nm band is dependent on two factors. Firstly, it is dependent on the average number of metal ions bound to the protein molecules as a whole. Clearly some molecules are likely to have one metal ion coordinated, some may have two while others may have none. The intensity of the 245 nm band is, therefore, a combination of the contributions from all the protein molecules in the sample. Secondly the number of deprotonated tyrosine residues will influence the intensity; the more tyrosine residues there are interacting, the greater the intensity. The first of these factors cannot account for such a large increase in $\Delta\varepsilon_1$, although the small increases observed across the *4f* group can be linked to differences in the association constants (see below). The differences in intensities observed between the transition metals and the lanthanides, therefore, could be interpreted in terms of a third tyrosine residue interacting with the $Ln^{3+}$ ions, either directly or indirectly.

As all the lanthanide ions used were in the +3 oxidation state, any variations in the $\Delta\varepsilon_1$ and $\Delta\varepsilon_{tot}$ values would be related the size of the metal ion and the coordination and stereochemical preferences. The ionic radius of the $Ln^{3+}$ ions decreases across the period (the well-known *lanthanide contraction*) due to the increasing nuclear charge being poorly shielded from one *4f* electron by the other *4f* electrons (Moeller, 1973). The stereochemistry observed for the members of the lanthanide series is a direct result of this decrease in size of the +3 ions. The larger ions, such as $La^{3+}$ and $Ce^{3+}$, commonly achieve coordination numbers of 9 or more (e.g. $La(NO_3)_2(DMSO)_4$), while the smaller ions at the other end of the period are generally octahedral (e.g. $[Er(NCS)_6]^{3-}$). Similarly, the anhydrous halides ($MCl_3$) of the larger ions are all 8- or 9-coordinate, while those from Tb to Lu are 7- or 6-coordinate (Cotton & Wilkinson, 1988).

Of those factors noted above, it is the ionic radius which should be the most important, as the stereochemistry is more or less restricted by the folding of the protein. Harris *et al.* (1981) reported a clear trend in the observed $\Delta\varepsilon$ values (at r = 2) of $M_2TF$ complexes (where

M = Pr, Nd, Eu, Tb, Ho, Er, Th and Pu) as a function of the ionic radius, in that $\Delta\varepsilon$ decreased as the size of the ion increased. A similar trend is observed in the binding of lanthanide ions to human lactoferrin, as shown in Figure 4.46. There are, however, some anomalous results which should be noted, most importantly that relating to Gd(III). The $\Delta\varepsilon_{tot}$ value associated with $Gd^{3+}$ is some 10 000 - 12 000 lower than would be expected if it were to fit into the same trend as the other lanthanides, while the value for $\Delta\varepsilon_1$ is about 3000 - 4000 lower. This could possibly be due to a change in the coordination number (Moeller, 1973) which is generally observed at this position in the lanthanides (the so-called *gadolinium break*, well documented relative to the stabilities of various Ln-ligand complexes where it generally manifests itself as a drop in the stability of the complex at or near $Gd^{3+}$). It seems likely that this change in the coordination number could either alter the interaction of the metal ion with the tyrosine ligands or decrease the actual number of tyrosine residues coordinating to the metal ion. This then would manifest itself as a decrease in the extinction coefficient as observed.

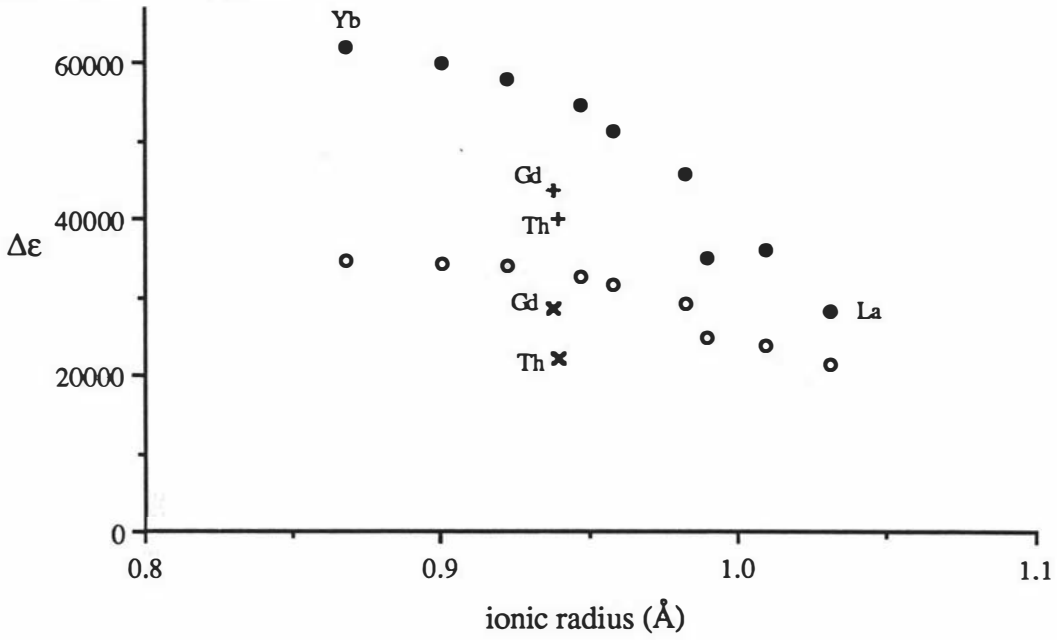
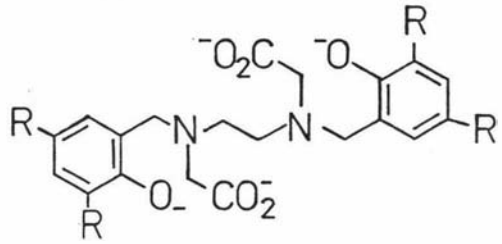



Figure 4.46: Plot of $\Delta\varepsilon_{tot}$ (•) and $\Delta\varepsilon_1$ (o) for the metal ions, from left to right, Yb, Ho, Tb, Eu, Sm, Nd, Pr, Ce and La, as a function of the ionic radius of the $Ln^{3+}$ ions. The values for Gd and $Th^{4+}$ are also shown as (+ for $\Delta\varepsilon_{tot}$ and x for $\Delta\varepsilon_1$).

It is conceivable therefore, that the larger lanthanide ions (Eu and above) could be coordinated to the four protein ligands established as being present in the metal binding sites

of lactoferrin (see section 1.2.5) and a third tyrosine which is within 6 Å of the metal binding site (Anderson *et al.*, 1989). Analysis of the tyrosine sidechains in the vicinity of the two metal binding sites in Lf, as noted in Section 3.4.1, indicate that Tyr82 in the N-terminal lobe and possibly Tyr415 in the C-terminal lobe (Tyr82 is equivalent to Tyr415) could be moved closer to the metal position and thus conceivably coordinate to a larger cation. As these residues are in a loop region between two pieces of secondary structure, it is possible that the required movement of the residues may not affect the stability of this piece of the structure to any large extent, although the significance of the steric effect that an increase from 6- to 7-coordination (as ligands moved closer together) would have on the stability of the metal binding site or on domain closure, is not known.

Difference spectrophotometric studies involving the use of the phenolate model compound, N,N'-di-(2-hydroxybenzyl)ethylenediamine-N,N'-diacetic acid (HBED) and its 3,5-dimethyl derivative (diMeHBED) (Figure 4.47) with the metal ions $Nd^{3+}$ and $Gd^{3+}$ also indicated that an increased number of tyrosine residues could be involved in the binding of lanthanide ions to Lf (Fitzpatrick, 1988).




(I) R = H

(II) R = Me

(I) : N,N'-di-(2-hydroxylbenzyl) ethylenediamine-diacetic acid (HBED)

(II): N,N'-di(3,5-dimethyl)-2-hydroxybenzyl) ethylenediamine-N,N'-diacetic acid (diMeHBED)

Figure 4.47: The structure of N,N'-di-(2-hydroxybenzyl)ethylenediamine-N,N'-diacetic acid (HBED; R = H) and its 3,5-dimethyl derivative (diMeHBED; R = methyl).

The ligand was titrated with $Nd^{3+}$ and $Gd^{3+}$ in 0.5M CAPS buffer (3-(cyclohexylamino)propane sulphonic acid) at pH 10.0 and difference spectra similar to those reported with EHPG (Pecoraro *et al.*, 1981; Chung, 1985) were obtained. From the titration

curves derived from the spectra, values for Δε/Tyr (the change in extinction coefficient associated with the deprotonation of one tyrosine residue) could be estimated.

Table 4.9: Some results from spectrophotometric titrations of two model compounds with copper, neodymium and gadolinium [a]

| | model compounds Δε / tyrosine | | human lactoferrin | Estimated total number of tyrosines (Tyr/metal) | |
|---|---|---|---|---|---|
| Metal ion | HBED | diMeHBED | Δε total (Lf) | HBED | diMeHBED |
| $Cu^{2+}$ | 5600 | 5825 | 17500 | 3.62 (1.6) | 3.0 (1.5) |
| $Nd^{3+}$ | 6450 | 7350 | 45930 | 7.1 (3.6) | 6.2 (3.1) |
| $Gd^{3+}$ | 6150 | 6900 | 44550 | 7.2 (3.6) | 6.5 (3.2) |

[a] Adapted from Fitzpatrick, 1988.

An estimate of the numbers of tyrosine residues interacting with each metal ion in lactoferrin can be made by dividing the Δε/phenolate values for the two ligands into the values for $\Delta\varepsilon_{tot}$ of $Nd_2Lf$ and $Gd_2Lf$, as estimated from the titration curves (Figures 4.37b and 4.33b). The results are given in Table 4.9. A value of around 3 Tyr/metal ion is estimated for both $Nd^{3+}$ and $Gd^{3+}$, consistent with earlier arguments related to the interaction of a third tyrosine residue.

A much clearer picture of the effect of ionic radius on the binding of the lanthanide ions to lactoferrin can be gained by comparing the average number of metal ions bound per protein molecule (n) after the addition of two equivalents of metal ion. It is apparent that the average number of metal ions bound increases with decreasing ionic radius (see Table 4.7), primarily as the result of a decrease in the steric crowding in the binding site and an increase in the lanthanide-carbonate solubility product ($K_{sp}$ ranges from $4 \times 10^{-34}$ for $La_2(CO_3)_3$ to $7.9 \times 10^{-32}$ for $Yb_2(CO_3)_3$, Martell & Smith, 1976), i.e. there is less competition for the free $Ln^{3+}$ ions from carbonate.

In calculating these n values, it was assumed that the binding of the first metal ion was stronger than the second and this is confirmed by analysis of the n vs r plots between r = 0 and 1, which all have gradients very close to 1.0 (i.e. the average number of metal ions

bound upon the addition of one equivalent of metal ion is 1.0). It is beyond r = 1, where the gradient of the titration curves begins to decrease, that weaker binding of the second equivalent is occurring. Harris & Carrano (1984) have used this part of the curve to estimate the second association constant ($\log K_2$) for the binding of vanadate to hTf, by using the calculated value of n to give the components of the mass-balance equation. More recently the same method has been used to compare the binding characteristics of a number of lanthanides to hTf (Harris & Chen, 1991). Although it is not possible to determine which of the metal binding sites is associated with the $K_1$ and $K_2$ constants for Lf, it has been shown that for hTf, the C-terminal site is linked to $K_1$, with $K_2$ corresponding to the N-site, at least for Nd(III), Sm(III) (Harris, 1986b) and Gd(III) (Harris & Chen, 1991) along with a variety of transition metal ions.

The $\log K_2$ values given in Table 4.7 should <u>not</u> be taken as absolute values but rather as relative values amongst the members of the lanthanide family. The values obtained do, however show a close similarity to the $\log K_2$ values reported by Harris (1986b) for $Nd^{3+}$ and $Sm^{3+}$ and more recently for $Gd^{3+}$ (Harris & Chen, 1991) binding to hTf, even though these values were obtained by a different method.

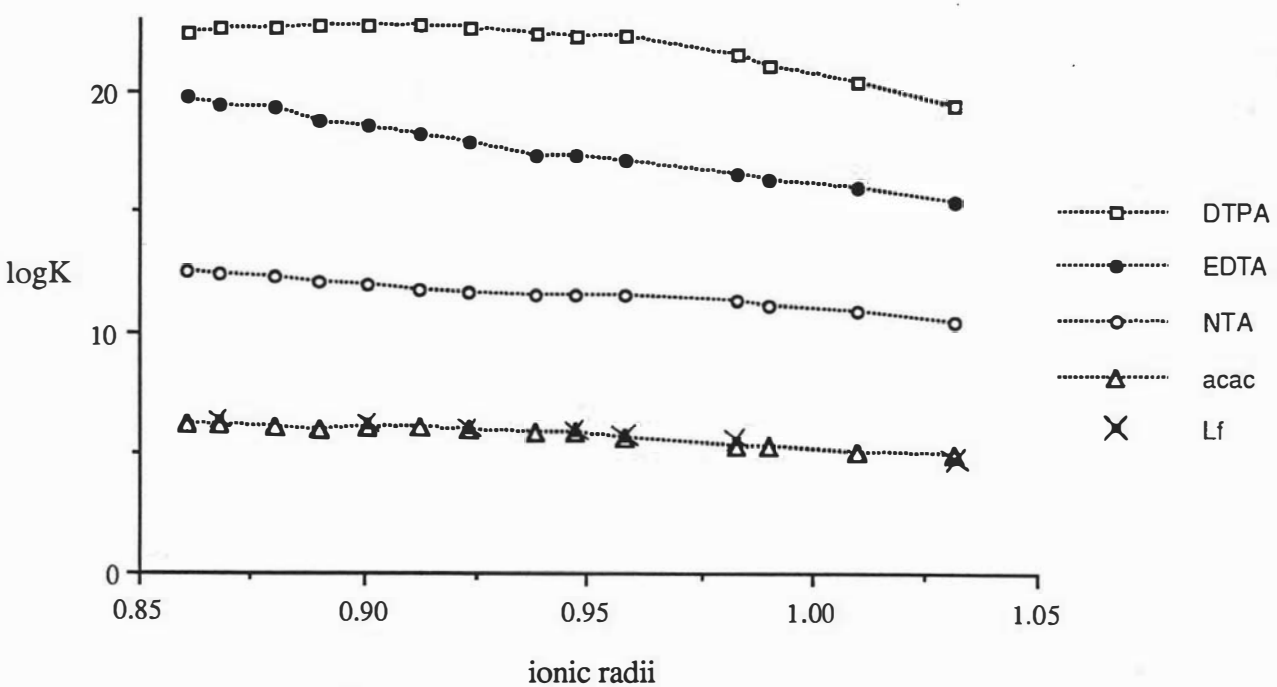



Figure 4.48: Plot of logK against atomic number for $Ln^{3+}$ complexes of EDTA, DTPA, NTA and acac, along with the $\log K_2$ values for Lf from this study.

Figure 4.48 shows a comparison of the logK values for lanthanide coordination to four organic chelating agents, EDTA, DTPA, NTA and acac and the seven $\log K_2$ values estimated here for Lf. Lactoferrin binds the lanthanide ions much more weakly than chelates such as EDTA, DTPA and NTA and appears to be much more like acetylacetone in both the magnitude of the association constants and the overall trend in the association constants.

*The effect of anionic chelators on f-block metal binding:* It is evident from Figure 4.48 that the presence of a competitive chelator such as EDTA or NTA in the spectrophotometric titrations of Lf with the lanthanide ions will have profound effects on metal uptake by the protein, causing changes either in the stoichiometry of binding or in the extent to which the metals ions can interact with the protein. EDTA, identified as a synergistic anion by Schlabach & Bates (1975), appears to have the greatest effect. As noted in section 4.2.4, the preparation of apolactoferrin from colostrum was carried out in essentially the same way throughout this study. A small amount of EDTA (5 mM) was added to the Tris-HCl buffer at the initial dialysis stage (after the ammonium sulphate precipitation step) to scavenge any adventitious metal ions and to produce essentially metal-free apoLf. Attempts were made to remove this later in the preparation prior to metal binding studies. In binding studies with iron and copper, the presence of trace amounts of EDTA had no effect, the binding proceeding at the same rate and to the same extent as with non-EDTA treated protein. However, cerium(III) binding to the "EDTA-treated" protein after dialysis against two changes of Tris-HCl gave a stoichiometry of binding near 1:1. When the apoLf was dialysed against a further 4-5 changes of buffer, normal binding of Ce(III) was observed with a stoichiometry of 2:1.

Analysis of the X-ray structure of apoLf has suggested that there may be an EDTA molecule in the open N-lobe binding cleft (G. E. Norris, personal communication). It was initially thought that this EDTA was bridging between the N1 and N2 domains but further analysis indicates that the EDTA interacts only with the N2 domain, by a series of hydrogen bonds.

That both $Fe^{3+}$ and $Cu^{2+}$ are able to bind efficiently to Lf in the presence of traces of EDTA shows that the inhibition of $Ce^{3+}$ binding is not a steric effect (with the EDTA blocking one site) but rather is related to the affinities of Lf and EDTA for the metal ions concerned. If the association constant for $Ce^{3+}$ binding to Lf is taken to be around $10^6$ (values between of $10^6$ and $10^7$ were obtained for $Nd^{3+}$ and $Sm^{3+}$ binding to hTf by Harris (1986b)) and the value for the Ce(III)-EDTA complex about $10^{16}$ (Moeller, 1973), then it is clear that a large proportion of one equivalent of the added $Ce^{3+}$ will exist as the EDTA complex (assuming one EDTA per lactoferrin molecule). In the case of iron, where the association constants for Lf and EDTA are much more similar ($\approx 10^{22}$ for Lf (Harris, 1986a) and $10^{25}$ for EDTA (Martell & Smith, 1976)), competition for $Fe^{3+}$ by EDTA will be significantly less than in the case of the lanthanides

In fact, iron is normally added as the Fe-NTA complex, in order to maintain the readily-hydrolysed $Fe^{3+}$ ion in a soluble form, and this leads to the formation of a short-lived quaternary intermediate (Bates, 1982), wherein the Fe-NTA complex is bound directly to the protein in the presence of (bi)carbonate, followed by the subsequent release of the NTA. Any EDTA present in the system would have little chance of coordinating the iron.

However, NTA cannot be used in the binding of the *f*-block metals to Lf to increase the solubility or prevent scavenging by EDTA, as its presence has a marked effect upon the binding of these weakly-coordinating metal ions. The results with thorium(IV) seem to justify this. In some of the earlier titrations of apoLf involving $Th^{4+}$, the metal was added as the $Th(NTA)_2$ complex following the proceedure of Harris *et al.* (1981). It had been reported that the NTA did not appear to have any effect on the binding of thorium to hTf, as judged by the $\Delta\varepsilon$ values and that even a 10-fold excess of the anion produced little perturbation in the difference spectra (Harris *et al.*, 1981). However, this was not the case with Lf.

A 1:2 Th:NTA complex gave a titration curve with an initial slope ($\Delta\varepsilon_1$) of about 14000 while the addition of $Th^{4+}$ (with no NTA present) to another sample of apoLf from the same preparation resulted in a $\Delta\varepsilon_1$ value near 23000. A similar result was observed with gadolinium, where the $\Delta\varepsilon_1$ values in the presence and absence of NTA were 17500 and

28000 respectively. This "NTA-effect" was quantified in the experiments where Nd and Yb with varying amounts of the anion were added to Lf to give 50% saturation of the protein. The presence of NTA has an effect on the observed $\Delta\varepsilon_1$ values of these metal ions and this effect is maximised at an NTA:$M^{3+}$ ratio of about 3:1. At this stoichiometry, a stable $Ln(NTA)_3$ complex is formed and additional NTA at this point would not coordinate the metal ions to the same extent.

Carbonate concentration also appears to be important in the binding of the lanthanides and actinides, as noted in the case of $Th^{4+}$ addition to apoLf in 0.01M $HCO_3^-$. The fact that the protein only appeared to bind 1·5 Th(IV) ions but with a $\Delta\varepsilon_1$ value closely similar to that observed with ambient $CO_2$, would indicate that the first equivalent of metal ion binds with little interference from the carbonate, and it is the second equivalent of $Th^{4+}$ which is is prevented from binding to the protein by carbonate.

It is interesting to note that when $Th(NTA)_2$ is used, although the observed $\Delta\varepsilon_1$ and $\Delta\varepsilon_{tot}$ values are lower, the stoichiometry is 2:1. Rather than inhibiting binding of the thorium, as does carbonate, the NTA still allows both equivalents of metal ion to bind, albeit in a much weaker fashion. It is possible that with NTA, a quaternary complex, as proposed by Bates (1982) is formed and the rearrangement in the metal binding site required to accommodate the NTA results in a weakening of the thorium - tyrosine interaction.

These results indicate quite clearly that NTA interferes in some way with the specific binding of weakly-coordinating metal ions such as the lanthanides and actinides and that it should be avoided when dealing with these metal ions. The draw-back to this, of course, is the increased risk of precipitation of the metal-carbonates or hydrolysis and precipitation of insoluble hydroxides (the lanthanide carbonates and hydroxides are quite insoluble, with $K_{sp}$ values between $10^{-32}$ and $10^{-34}$ for the carbonates and $\approx 10^{-20}$ for the hydroxides, Martell & Smith, 1976).

When taken together, these results indicate that the binding of lanthanide and actinide ions to lactoferrin are very sensitive to the conditions employed. It is clear that different anions influence the binding of these metal ions in different ways, whether by competing with the protein for the metal and lowering the apparent stoichiometry of binding, as in the case of

EDTA or carbonate, or by modulating the metal-tyrosine interaction while still allowing essentially complete binding to the two sites, as does NTA. In this respect, it must be pointed out once again that the $\log K_2$ values presented here may not accurately reflect the true nature of the association of the lanthanides with Lf. They do, however, give a relative measure of the strength of metal ion binding parallelling that observed in small molecule chelates and complexes (Moeller, 1973).

### 4.6.4 *Electron spin resonance of $Gd^{3+}$ binding:*

Although the majority of the data on lanthanide ion binding to the transferrins have come from the electronic absorption characteristics of both the protein and the metal ions themselves, the magnetic properties of the $f^7$ Gd(III) ion provide an ESR probe into the metal binding sites previously the domain of the transition metals. The ESR spectrum of the 1:1 Gd-Lf complex indicates that the metal ion is in an environment closely similar to that found in Tf, based on the similarity between this spectrum and those previously reported for a 1:1 Gd-Tf complex (Zak & Aisen, 1988). In hTf, this metal is most likely associated with the C-terminal binding site, as it has been determined that it is this site which corresponds with the $K_1$ association constant (Harris & Chen, 1991). Although there is no direct evidence linking $K_1$ with the C-site in Lf, the close similarity in metal binding and stability of the sites to pH (Brock, 1985) would suggest that this would most probably be the case.

Simulations of the transferrin spectrum suggested a rhombic environment for Gd(III), similar to that proposed for Fe(III). Bearing in mind the predominance of coordination numbers greater than six in the chemistry of the lanthanide ions (Moeller, 1973; Hart, 1987), it is quite possible that more than just the four iron ligands (in addition to the bidentate carbonate), may be involved when larger cations associate with the transferrins (as discussed earlier). Water molecules, for example, could also coordinate to the larger metal ions. The sites are readily accessible to solvent, as there are at least two water molecules within 3.5 Å of the metal atoms in both the $Fe_2$Lf and $Cu_2$Lf structures.

The ESR spectra of a series of Gd-Lf complexes with increasing amounts of metal ion show that the interaction of $Gd^{3+}$ with the protein is not as simple as with other metal ions.

The most striking feature of these spectra is the appearance of a sharp, intense signal at g' = 4.55, observed at a metal:protein ratio of 1:1 and above (position 7 in Figure 4.43) but absent at the lowest ratio (0.5:1 Gd:protein). There are two possible explanations for the emergence of this transition.

Firstly, the new, sharp peak could be due to non-coordinated $Gd^{3+}$ ions in solution. Free $Gd^{3+}$ ions in an aqueous environment (0.02M HEPES buffer) typically exhibit strong resonances centred at g = 2.0, resulting from the spherically symmetric electronic environment of the $^8S_{7/2}$ ground state. Even in other biological complexes of Gd(III), the major signals are generally in the g=2 region (O'Hara & Koenig, 1986), although a gadolinium(III)-bovine α-lactalbumin complex has two intense transitions at g = 3 and 5.8 (Musci *et al.*, 1986). As the Gd-Lf ESR spectra were recorded in non-coordinating HEPES buffer, it follows that any $Gd^{3+}$ ions not coordinated to the protein would exist as aquo ions and it therefore appears unlikely that the g = 4.55 resonance is due to free $Gd^{3+}$ ions.

The other possibility is that the $Gd^{3+}$ ions are in two slightly different environments. The resonance centred at g' = 4.02 (position 10 in Figure 4.43) appears to have reached a maximum at a ratio of 1:1 and undergoes no change as the ratio increases above this. This signal may be associated with the first equivalent of $Gd^{3+}$ bound to the protein. As the second equivalent binds, the g' = 4.55 resonance is observed to grow and it is conceivable that it is due to $Gd^{3+}$ in the second binding site. It should be noted that at the four different stoichiometries used, the hyperfine signals in the g' = 6.5 - 12 region continue to grow at a steady rate. These probably originate from spin-spin interactions which are essentially unaffected by the differences in environment responsible for the g' = 4.02 and 4.55 transitions.

## 4.6 Concluding remarks

The results of spectroscopic studies on the binding of vanadium and cerium to human lactoferrin suggest that rather than lactoferrin stabilising metal ions in the +3 state (Bertini *et al.*, 1982; 1985), the protein tends to convert metal ions into the highest stable oxidation state which can be maintained by the protein at physiological (or near-physiological) pH. In

aqueous solutions, it is the V(V) species which is the most stable, the other two oxidation states stabilised only by acidic pH or in the absence of oxidising agents.

With cerium, however, it is Ce(III) which is more stable in at pH 7. Cerium(III) bound to lactoferrin, therefore, behaves in much the same way as Mn(II) and Co(II), in that they all become much more unstable towards oxidation by $O_2$ than in aqueous solutions, changing readily to a higher oxidation state. This may be a consequence of the increase in negative charge in the binding site following deprotonation of the tyrosines, with the higher oxidation state becoming more favoured.

# CHAPTER 5

## Conclusions

5.1 *Metal binding:*

The presence of both 5- and 6-coordinate copper centres in $Cu_2Lf$ 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. Many other cations are also bound by the transferrins, often with differing sizes and coordination preferences. We may anticipate that some of these may also show significant stereochemical differences between the two sites even though the overall protein 3-dimensional structures may not change.

Indications of stereochemical differences between the two sites have previously been obtained from spectral studies. ESR studies on chromium-substituted hTf (Aisen *et al.*, 1969) and Lf (Ainscough *et al.*, 1980) indicate that there is a difference in the signals from the two sites, designated type 1 and type 2, and that, in the case of Lf, only one $Cr^{3+}$ ion can be replaced by $Fe^{3+}$. This apparent difference in the two chromium binding sites has been used by Harris (1977) to identify which sites showed the type 1 and type 2 spectra by selectively loading $Cr^{3+}$ into either metal binding site by alterations in the pH of the medium. Some additional spectroscopic studies which suggested differences between the two sites include: (i) the spectrophotometric titrations of apoTf with various lanthanide ions which established that the larger cations such as $Pr^{3+}$ and $Nd^{3+}$ bound to one metal binding site only, while other smaller cations (including $Eu^{3+}$ and $Tb^{3+}$) could bind to both (Luk, 1971). This was interpreted to mean that the space available in the binding sites was limited and that the sites were inequivalent for the binding of larger metal ions. As noted earlier, the later work of Harris (1986b), who determined both $\log K_1$ and $\log K_2$ values for $Nd^{3+}$ binding to hTf, shows (at least for the lanthanides with ionic radii less than or equal to that of $Nd^{3+}$) the conclusions of Luk were probably due to an error in the interpretation of the titration curves, as previously discussed (Section 4.6.1); (ii) aluminium binding studies on cOTf (Donovan & Ross, 1975), where it was observed that the addition of one $Al^{3+}$ ion increased the denaturation temperature of apoOTf by 5°, while a second equivalent of aluminium had no

effect; (iii) the observation of two different signals in the ESR spectra of vanadyl-transferrin complexes which were related to two magnetically inequivalent $VO^{2+}$ environments designated the A and B conformations (Chasteen *et al.*, 1977), and the pH-dependent nature of these conformations (this is discussed in more detail below); and (iv) the use of laser-induced excitation of transferrin-bound europium (O'Hara & Bersohn, 1982) which made it possible to distinguish the N- and C-terminal sites. It was suggested that the inequivalence of the two sites could be the result of differences in the charges of the liganding groups in the two sites, although the structural studies on both Lf (Anderson *et al.*, 1987; 1989) and rSTf (Bailey *et al.*, 1988) do not support this. The residues which coordinate to the metal are identical in both sites, as are the residues which hydrogen bond to the anion. It is conceivable (due to the larger size of the europium(III) ion) that it could be coordinated to other protein ligands in addition to those involved in iron binding, and that these ligands could be in different protonation states in the two sites.

The A and B conformations reported for the two vanadyl binding sites in hTf (Chasteen *et al.*, 1977) can be rationalised in terms of the observed coordination geometries in $Fe_2Lf$ and $Cu_2Lf$ (it should be noted that the term "conformation" was used by the authors to mean the geometrical arrangement of the ligands and any differences in their identity). Modelling of a $VO^{2+}$ ion in the N-lobe site of $Cu_2Lf$ (Section 4.5.2) indicates that there would probably be little rearrangement of the protein ligands or the carbonate ion upon vanadyl binding, resulting in a 6-coordinate geometry. In the more constrained C-lobe site, however, where the carbonate is not free to rotate, the incorporation of a $VO^{2+}$ ion would most likely result in a 7-coordinate environment and may lead to weaker interactions of some of the other ligands. This is a likely cause of the difference in the ESR signals from the two sites.

It was also reported that increasing the pH above 9 led to a change from the A conformation to the B, coupled to the ionisation of a non-coordinating protein group with a pK near 10, possibly arginine or lysine. The structures of lactoferrin and rabbit transferrin show that an arginine is involved in hydrogen bonding to the anion and it is possible that deprotonation of this group at high pH could alter the interactions with the anion which in turn could cause a movement of the anion and change the environment around the metal ion.

The different visible absorption maxima for the N- and C-substituted monocupric complexes of cOTf are also suggestive of a coordination difference (Yamamura *et al.*, 1984). Likewise, recent EXAFS studies on $Cu_2$OTf (Garratt *et al.*, 1992) have led to the conclusion that the average coordination number of the two $Cu^{2+}$ is 5.5, implying one 6-coordinate site and one 5-coordinate. Further, Garratt *et al.* (1992) suggest this could imply a monodentate anion in one site.

Most of the spectroscopic studies reported in this thesis show no indication of any non-equivalence in the two sites. The broad charge transfer band in the visible region does not allow resolution of this band into the separate contributions of the two sites. The ESR spectrum of $Cu_2$Lf shows that both $Cu^{2+}$ ions are in an axially symmetric field and that, once again, the contributions to the overall spectrum from the two sites cannot be distinguished. In fact the carbonate containing site in $Cu_2$oxLf has parameters ($g_{\|} = 2.310$, $A_{\|} = 155 \times 10^{-4}$ $cm^{-1}$) identical to those observed in $Cu_2$Lf ($g_{\|} = 2.314$, $A_{\|} = 155 \times 10^{-4}$ $cm^{-1}$). The studies with the larger lanthanide ions (where any restrictions due to the size of the cations could become evident) show that, in the case of the larger cations (from $La^{3+}$ to $Sm^{3+}$), the stoichiometry of binding to human lactoferrin is predominantly 2:1. This is apparently in contrast to serum transferrin (see above), where some of the larger metal ions bound only 1:1. In addition, the results with the lanthanide ions presented in this work conflict with the earlier work of Chung (1985). The limitations of the analysis of this earlier work have been discussed in Section 4.6.1.

One indication of inequivalence comes from the results obtained for the pH dependence of $Ce^{4+}$ release (from $Ce(IV)_2$Lf). This shows that there is a marked difference in the stability of the two sites towards acid, in that release of the metal ion is clearly a biphasic process in much the same way as is the release of iron (Mazurier & Spik, 1980). The ESR spectra of several gadolinium-lactoferrin complexes of different stoichiometries also suggest that there may be two different signals from the two sites, one which predominates at low Gd:Lf ratios, and the other which becomes more significant as this ratio increases. Similar differences have not been reported for serum transferrin.

Some indication of the spectroscopic properties of a single copper site in lactoferrin can be gained from the 1:1 copper complex of the recombinant half molecule of human Lf ($CuLf_N$) reported by Day *et al.* (1992). This complex has a visible absorption maximum at 424 nm, ≈10 nm lower than that for intact $Cu_2Lf$. Assuming that the copper site in the half molecule is not changed by separation of the two lobes, this suggests that the copper complex of the C-terminal half would have a $\lambda_{max}$ near 445 - 450 nm. These differences can be rationalised in terms of a crystal field argument for copper(II) (see Section 4.4.1). Following the same arguments, this would imply that it is the C-lobe of cOTf (with the lower $\lambda_{max}$) which has the 5-coordinate copper site and the N-lobe which has the 6-coordinate site.

5.2 *Anion binding:*

The monodentate (bi)carbonate in the N-terminal site of $Cu_2$oxLf serves as a confirmation of the results obtained with $Cu_2Lf$. Although the two structures were not determined completely independently (in that the starting model for the $Cu_2$oxLf refinement was the refined model of $Cu_2Lf$), care was taken to ensure that the electron density of the $Cu_2$oxLf in the vicinity of the two metal and anion binding sites was not biased towards the structure observed in $Cu_2Lf$. Therefore the two metal binding sites in $Cu_2$oxLf may be regarded as being independently determined, and there can be little doubt that the observed structure is the correct one.

The coordination observed for oxalate appears to contradict the classic "interlocking sites" model (Schlabach & Bates, 1975) in which the anion is bound to iron via the proximal electron donor group, while its carboxylate group interacts with the protein. In fact, in the case of oxalate, both the proximal ligand and the carboxylate bind to the metal ion. This can very likely be extended to other anions such as thioglycolate, for which EXAFS studies indicate that the sulphur ligand at least is directly bound to the iron (Schneider *et al.*, 1984). These arguments lead to a revised general model for anion binding in the transferrins, based on a synergistic anion with the structure described by Schlabach and Bates (1975). This is shown in Figure 5.1.

Another result emerging from the studies on $Cu_2Lf$ and $Cu_2oxLf$ is the importance of hydrogen bonding in helping to define the localised conformational changes which accompany metal or anion coordination. Clearly hydrogen bonds, and to some extent solvent molecules, do play an integral part in the stabilisation of various groups in the protein and can have quite marked effects on the overall function of the molecule. As discussed in Section 3.4.2, it is possible that the inability of oxalate in a 50 molar excess over the total protein to replace carbonate in the N-lobe site may be determined, in part, by the hydrogen bond between the sidechain of Arg121 and $O_\gamma$ of Ser191. The fewer interdomain interactions in the N-lobe (compared to the C-lobe) may also contribute, by allowing the domains to open somewhat more and the copper to adopt a more favourable geometry.

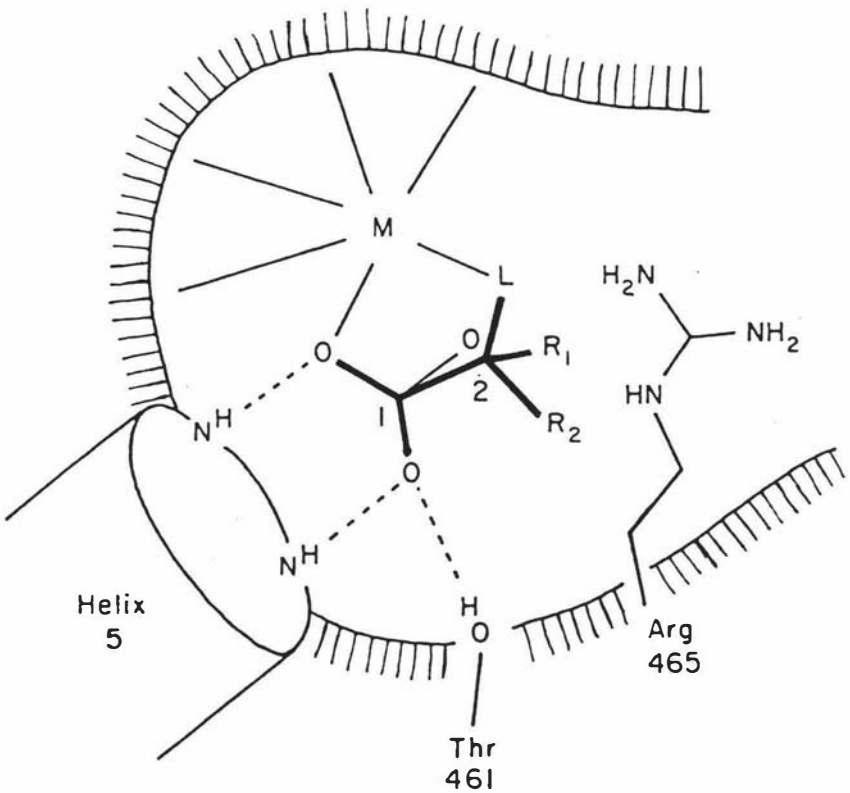



Figure 5.1: A revised model of the binding of a carboxylate synergistic anion to the transferrins (shown in the C-lobe binding site of human lactoferrin). The carboxylate group interacts with the metal, the N-terminus of helix 5 and Thr461, as for carbonate and oxalate. The rest of the anion (heavy lines) extends past the third carbonate position (the carbonate is shown with thin lines) and the electron donor group, L, can then bind to the metal. The sidechain of Arg465 is necessarily displaced away from the anion. Whether it can still hydrogen bond with the anion will depend on the nature of the substituents $R_1$ and $R_2$.

Analysis of the $Cu_2$oxLf structure beyond Tyr398 indicates that there would be little or no interference if the tyrosine sidechain were to be pushed further, perhaps by a larger movement of the arginine. It should then be possible to substitute even larger anions for carbonate so that they still provide a bridge between metal ion and the protein molecule. The binding of anions such as malonate, maleate, EDTA, NTA and glycine is well known (Schlabach & Bates, 1975) and it is quite conceivable that they can be readily accommodated in at least the C-terminal site of Lf by limited conformational changes involving only a small number of residues and a possible rearrangement of the solvent structure in the interdomain cleft. An anion the size of xylenol orange, a known synergist (Harris & Gelb, 1980), on the other hand could not be easily incorporated without some sort of large scale reorganisation of the structure, possibly a domain movement similar to that observed for apoLf (Anderson *et al.*, 1990; Norris *et al.*, 1991).

The affinity of anions other than carbonate may well depend on the extent to which their substituents (see Figure 5.1) alter the hydrogen bonding interactions in and around the binding site. Oxalate is the only anion with similar affinity to $CO_3^{2-}$, and the $Cu_2$oxLf structure shows that not only are the hydrogen bonds between the anion and arginine preserved, but that the displacement of residues beyond Arg465 also leads to new hydrogen bonding interactions. Anions substituted with non-hydrogen bonding groups on the proximal carbon would not be able to hydrogen bond to the arginine, and so may have lower affinity.

Campbell & Chasteen (1977), in their study of the ability of a variety of anions to facilitate $VO^{2+}$ binding to hTf found that anions could be classified in three ways: (i) Group I, non-synergistic, in that no metal binding was evident; (ii) Group II, those which produced B conformation spectra and (iii) Group III, those which produce both A and B spectra (the two conformations, A and B, as described by Chasteen *et al.*, 1977). Those in Group II include lactate, glycolate, thioglycolate, glyoxylate and salicylate, which have $R_1$ and $R_2$ substituents incapable of hydrogen bonding to the arginine. Group III anions on the other hand are dinegative anions containing two or more carboxylate groups (such as oxalate, malonate, nitrilotriacetate and maleate), having substituents which are capable of hydrogen bonding. In

the light of the structures of both $Fe_2Lf$ and $Cu_2oxLf$, it is possible, therefore, that the differences in the ESR spectra observed by Campbell & Chasteen (1977) between these two groups of anions, in addition to the arguments concerning changes in the coordination geometry of the metal (see earlier), could be also related to the ability of the anion to hydrogen bond to the arginine sidechain.

5.3 *$CO_3^{2-}$ or $HCO_3^-$; implications for the mechanism of metal binding and release:*

The origin of the protons released on metal complexation has long been a matter of debate (Gelb & Harris, 1980). It was reported that three protons are released when $Fe^{3+}$ binds to hTf. Two protons in each site must be released by deprotonation of the two tyrosine ligands. The third proton released from each site could come from deprotonation of $HCO_3^-$, implying that the form of the anion is $CO_3^{2-}$. However, the results with oxalate also reported by Gelb & Harris (1980) complicate the matter. It was found that in the presence of oxalate, similar results were observed, ie., three protons were released when $Fe^{3+}$ bound. These results could be interpreted in terms of a bicarbonate as the anion (irrespective of what metal is bound), as the third proton cannot come from the anion as oxalate is already deprotonated.

However, if the anion were bicarbonate, then it is difficult to identify where the third proton originates. Gelb & Harris (1980) suggest a water molecule as a possible source and although it is known from the structural analyses that there at least two solvent molecules within about 5 - 6 Å of the two metal ions, none are coordinated directly to the metal (as earlier thought). The bidentate chelation of the anion in $Fe_2Lf$ and the C-lobe of $Cu_2Lf$ are consistent with $CO_3^{2-}$, as conversion of a carbonate to bicarbonate generally results in either the breakup of the complex or monodentate coordination of $HCO_3^-$ (Oldham, 1987), ie., monodentate coordination of $HCO_3^-$ is favoured over bidentate.

The monodentate coordination of the anion in the N-lobe site of $Cu_2Lf$ is, therefore, consistent with $HCO_3^-$. Potential hydrogen bonding interactions in the binding site are also consistent with $HCO_3^-$; one non-coordinated oxygen, $O_1$, of the anion is 2.9 Å from $O_\eta$ of the apical ligand, Tyr 92, and the geometry of this interaction suggests a hydrogen bond. If

this is so it implies that either the anion or the Tyr sidechain must be protonated. This may also partially explain why the proton release titrations (Gelb & Harris, 1980) indicate that fewer protons are released when $Cu^{2+}$ is bound, compared with $Fe^{3+}$ (about 2 protons per $Cu^{2+}$ and 3 protons per $Fe^{3+}$).

A proposed mechanism for copper and iron binding to human lactoferrin, based on structural results and proton release studies, is given below. The process of iron binding was suggested by Anderson *et al.* (1990) as involving the initial binding of a bicarbonate anion in the anion pocket of apoLf, with one oxygen atom then essentially buried in the anion pocket while the other two project out into the interdomain cleft. This presents a negatively-charged surface to any incoming metal ion. The $Fe^{3+}$ ion binds to the anion and the two tyrosine residues, 92(435) and 192(528), which are in the vicinity, leading to the release of three protons per iron, two from the tyrosines and one from the bicarbonate. This interaction triggers the closure of the N2 domain over the N1 domain, leading to the coordination of His253(597) and Asp60(395), both of which are deprotonated at the pH used in binding experiments.

Iron binding:

$$2\ HCO_3^- + H_4Lf \rightleftharpoons (HCO_3^-)_2\text{-}H_4Lf$$

$$2\ Fe^{3+} + (HCO_3^-)_2\text{-}H_4Lf \rightleftharpoons (Fe^{3+})_2\text{-}(CO_3^{2-})_2\text{-}Lf + 6\ H^+$$

Copper binding:

$$2\ HCO_3^- + H_4Lf \rightleftharpoons (HCO_3^-)_2\text{-}H_4Lf$$

$$\text{N:}\quad Cu^{2+} + (HCO_3^-)_2\text{-}H_4Lf \rightleftharpoons Cu^{2+}\text{-}^*(HCO_3^-)_2\text{-}H_2Lf + 2\ H^+$$

$$\text{C:}\quad Cu^{2+} + Cu^{2+}\text{-}^*(HCO_3^-)_2\text{-}H_2Lf \rightleftharpoons (Cu^{2+})_2\text{-}^*(HCO_3^-)(CO_3^{2-})\text{-}Lf + 3\ H^+$$

* It is either the anion or Tyr92 which is protonated. As written, the anion is protonated.

In the case of copper, it is assumed that the initial stages of anion binding are identical, with the binding of carbonate in the two anion pockets. A $Cu^{2+}$ ion is attracted into the N-terminal site and binds to the anion and the tyrosines, releasing two protons (one from

tyrosine 192 and the other from either the anion or Tyr92). As the lobe closes, the copper adopts a square pyramidal geometry, resulting in the movements observed in the positions of the metal, anion and ligands.

The C-terminal copper atom should bind in much the same way, although the increased interdomain interactions do not allow the copper to take up the more favourable square pyramidal geometry. However, the asymmetry observed in the anion coordination during refinement, and the lengthening of the Cu - $O_\eta$ bond, could possibly reflect a movement towards a more distorted stereochemistry.

In summary, in $Fe_2$Lf the four tyrosines and both anions are deprotonated ($CO_3^{2-}$), while in the case of copper coordination (in the N-lobe) either one tyrosine or the anion is still protonated (the exact location of the hydrogen atom cannot be defined at the resolution of the structure determination). In the above mechanism, five protons are released upon the binding of two $Cu^{2+}$ ions in contrast to the four reported from proton release studies (Gelb & Harris, 1980). It is also possible, however, that the lengthening of the Cu - Tyr528 bond to 2.4 Å could imply that this tyrosine remains protonated, so accounting for the release of only one proton from this site. This is consistent with the lower $\Delta\varepsilon$ values at 245 nm for $Cu^{2+}$ binding to hLf compared with other transition metals. For example, both the $\Delta\varepsilon_1$ and $\Delta\varepsilon_2$ (see Section 4.5.2, page 224) for $Cu^{2+}$ are about half those for $Fe^{3+}$ and $V^{3+}$, suggesting only about half the number of tyrosine residues are deprotonated in the copper case.

It has been suggested that protonation of the carbonate and a change from bidentate to monodentate coordination could be an important step in metal release (Sarra *et al.*, 1990). It could be suggested, therefore, that the 5-coordinate structure observed in the N-lobe of $Cu_2$Lf and $Cu_2$oxLf may be an structure intermediate between full coordination (as in both lobes of $Fe_2$Lf and $Fe_2$rSTf, and the C-lobe of $Cu_2$Lf) and an empty binding site (as in apoLf). It is also feasible that protonation of either the two metal-binding tyrosine residues could initiate metal release, or it could even be a disruption of the inter-domain hydrogen bonds and salt bridges which leads to an opening of the two domains and a decrease in the coordination number of the metal.

5.4 *Additional metal-anion-lactoferrin complexes:*

The $Cu_2$oxLf complex was obtained by reacting the dicupric-dicarbonato complex with a 50 molar excess of oxalate, and showed that with this technique only one carbonate is able to be replaced. Two recent studies give results which further illuminate anion substitution: (i) it has been observed that the addition of oxalate to carbonate-free apoLf, followed by the addition of 2 equivalents of copper(II) produces a dicupric-dioxalato complex ($Cu_2ox_2Lf$) (Shongwe *et al.*, 1992), where clearly oxalate binding in the N-terminal anion site has taken place. It should be noted, however, that one of the oxalate anions (presumably that in the N-lobe site) is readily replaced by carbonate when the complex is exposed to air. The remaining oxalate, and for that matter the C-lobe oxalate in $Cu_2$oxLf, is not displaced under the same conditions; (ii) the structural analysis of a diferric-dioxalato lactoferrin complex ($Fe_2ox_2Lf$) indicates quite clearly that there is an oxalate anion bound in both anion sites in a 1,2-bidentate fashion (H. M. Baker, personal communication). This complex was prepared in essentially the same way as the afore-mentioned dicupric-dioxalate complex, in that both preparations were based on addition of oxalate to carbonate-free apoLf solutions.

These studies show that oxalate is unable to replace carbonate in the N-terminal site once the ternary copper-carbonate-lactoferrin complex ($Cu_2$oxLf) has formed but that the empty pocket is able to accommodate oxalate, albeit with somewhat reduced stability. This is also the case with cOTf when the associated metal is $Cu^{2+}$ (Zweier, 1980).

5.5 *Implications for future research:*

*Crystallographic studies:* The modelling studies with other transition metals such as vanadium (Section 4.4.2) and lanthanide ions (Section 3.4.1) are speculative at this stage. Although it has been shown that spectroscopic studies on the binding of larger cations to lactoferrin implicate a third tyrosine in metal binding (this thesis) and that at least in the N-lobe of hLf such a possibility seems to be structurally reasonable, it remains to be definitively proved with a crystal structure. Crystallisation trials using 0.01M phosphate buffer

(conditions identical to those which proved successful for $Fe_2Lf$, $Cu_2Lf$ and $Cu_2oxLf$) have been conducted with $Ce^{4+}$, $Gd^{3+}$ and $Th^{4+}$. However, the lower association constant for these metal ions with the protein, coupled with the high solubility products of the metal-phosphates, resulted in the loss of the metal from the protein and the appearance of apolactoferrin crystals. Trials using Tris-HCl and HEPES buffer have yet to produce crystals.

In order to confidently define the mode of binding of other synergistic anions (the model depicted in Figure 5.1 is based only on the results for oxalate, which could conceivably be a special case, as is carbonate) the structure of at least one more anion-substituted lactoferrin complex is required. The thioglycolate anion would be a good candidate, as the proximal ligand is very different from the carboxylate group. It is predicted that with either $Fe^{3+}$ or $Cu^{2+}$ as the associated metal ion, a metal-thioglycolate-lactoferrin complex would crystallise under the same conditions as $Fe_2Lf$, $Fe_2ox_2Lf$, $Cu_2Lf$, $Cu_2oxLf$ and $Cu_2ox_2Lf$.

*Spectroscopic studies:* As discussed previously, the association constants estimated for seven lanthanide ions should not be regarded as "absolute" constants, but rather as "relative" constants, in that they show the correlation between the size of the metal ions and the affinity of lactoferrin for the different metals. Research recently instigated in this laboratory uses the method of Harris & Carrano (1983) to estimate the equilibrium constants for metal ions with lactoferrin in the presence of competitive chelators in spectrophotometric titrations. Preliminary results on gadolinium(III) and gallium(III) binding to lactoferrin are very promising, and it is envisaged that the determination of the association constants for a range of metal ions will allow a direct comparison between serum transferrin and human lactoferrin. In addition, this work could be extended to include bovine lactoferrin, and from analysis of the structures of these three proteins, correlations between structural details and the affinity for different metal ions can perhaps be made.

In addition, the production of the half molecule fragments (both the N- and C-lobes) of lactoferrin in this laboratory will allow the differences between the two metal binding sites to be fully analysed by a variety of techniques. As noted previously, the electronic absorption

spectrum of the N-lobe fragment has shown one such difference. Structural studies at present under way on the iron-saturated N-lobe fragment may give further insight into the metal binding function of this protein, and the transferrin family of proteins as a whole.

*References*

Aasa, R., Malmstrom, B. G., Saltman, P. & Vanngard, T. (1963) *Biochim. Biophys. Acta 75*, 203-222. The specific binding of iron(III) and copper(II) to transferrin and conalbumin.

Aasa, R. (1970) *J. Chem. Phys. 52*, 3919-3930. Powder line shapes in the electron paramagnetic resonance spectra of high-spin ferric complexes.

Abola, J. E., Wood, M. K., Chweh, A., Abraham, D. & Pulsinelli, P. D. (1982) in *The biochemistry and Physiology of Iron* (P. Saltman & J. Hegenauer, Eds.) pp 27-34, Elsevier, New York. Structure of hen ovotransferrin at 5 Å resolution.

Ainscough, E. W., Brodie, A. M. & Plowman, J. E. (1979) *Inorg. Chim. Acta 33*, 149-153. The chromium, manganese, cobalt and copper complexes of human lactoferrin.

Ainscough, E.W., Brodie, A.M., Plowman, J.E., Bloor, S.J., Loehr, J.S. & Loehr, T.M. (1980) *Biochemistry 19*, 4072-4079. Studies on human lactoferrin by electron paramagnetic resonance, fluorescence and resonance Raman spectroscopy.

Ainscough, E.W., Brodie, A.M., McLachlan, S.J. & Ritchie, V.S. (1983) *J. Inorg. Biochem. 18*, 103-112. Spectroscopic studies on copper(II) complexes of human lactoferrin.

Aisen, P. Leibman, A. & Reich, H.A. (1966) *J. Biol. Chem. 241*, 1666-1671. Studies on the binding of iron to transferrin and conalbumin.

Aisen, P., Aasa, R. & Redfeild, A.C. (1969) *J. Biol. Chem. 244*, 4628-4633. The chromium, manganese and cobalt complexes of transferrin.

Aisen, P., Leibman, A. & Sia, C. (1972) *Biochemistry 11*, 3461-3464. Molecular weight and sub-unit structure of hagfish transferrin.

Aisen, P., Pinkowitz, R.A. & Leibman, A. (1974) *Ann. N.Y. Acad. Sci. 222*, 337-3 . EPR and other studies of the anion-binding sites of transferrin.

Aisen, P., Leibman, A. & Zweier, J.L. (1978) *J. Biol. Chem. 253*, 1930-1937. Stoichiometric and site characteristics of the binding of iron to human transferrin.

Aisen, P. (1980) in *Iron in Biochemistry and Medicine II* (A. Jacobs and M. Worwood, Eds.) pp 87-129, Academic Press, London. The transferrins.

Aisen, P. & Listowsky, I. (1980) *Annu. Rev. Biochem. 49*, 357-393. Iron transport and storage proteins.

Aisen, P. & Harris, D. C. (1989) in *Iron Carriers and Iron Proteins* (T. M. Loehr, Ed.) pp 241-351, VCH Publishers, New York. Physical biochemistry of the transferrins.

Aisen, P. (1989) in *Iron Carriers and Iron Proteins* (T. M. Loehr, Ed.) pp 355-371, VCH Publishers, New York. Physical biochemistry of the transferrins: Update, 1984-1989.

Al-Hilal, D., Baker, E., Carlisle, C.H., Gorinsky, B., Horsburgh, R.C., Lindley, P.F., Moss, D.S., Schneider, H. & Stimpson, R. (1976) *J. Mol. Biol. 108*, 255-257. Crystallization and preliminary X-ray investigation of rabbit plasma transferrin.

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. USA 84*, 1769-1773. Structure of human lactoferrin at 3.2Å resolution.

Anderson, B.F., Baker, H.M., Norris, G.E., Rice, D.W. & Baker, E.N. (1989) *J. Mol. Biol. 209*, 711-734. Structure of human lactoferrin: Crystallographic structure analysis and refinement at 2.8 Å resolution.

Anderson, B.F., Baker, H.M., Norris, G.E., Rumball, S.V., & Baker, E.N. (1990) *Nature (London) 344*, 784-787. Apolactoferrin structure demonstrates ligand-induced conformational change in transferrins.

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

Arnold, R.R, Cole, M.F. & McGhee, J.R. (1977) *Science 197*, 263-265. A bacterial effect for human transferrin.

Azari, P.R. & Feeney, R.E. (1958) *J. Biol. Chem. 232*, 293-302. Resistance of metal complexes of conalbumin and transferrin to proteolysis and chemical treatments.

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. Molecular structure of serum transferrin at 3.3 Å resolution.

Baker, E.N. & Rumball, S.V. (1977) *J. Mol. Biol. 111*, 207-210. Crystallographic data for human lactoferrin.

Baker, E. N. & Dodson, E. J. (1980) *Acta Cryst. A36*, 559-572. Crystallographic refinement of the structure of actinidin at 1.7 Å resolution by fast fourier least-squares methods.

Baker, E. N. & Hubbard, R. E. (1984) *Progr. Biophys. Mol. Biol. 44*, 97-179. Hydrogen bonding in globular proteins.

Baker, E. N., Anderson, B. F. & Rumball, S. V. (1987) *Trends Biochem. Sci. 12*, 350-353. Transferrins: Insights into structure and function from studies on lactoferrin.

Baker, E. N. (1988) *J. Mol. Biol. 203*, 1071-1095. Structure of azurin from *Alcaligenes denitrificans*. Refinement at 1.8 Å resolution and comparison of the two crystallographically independent molecules.

Baker, E. N., Anderson, B. F., Baker, H. M., Haridas, M., Norris, G. E., Rumball, S. V. and Smith, C. A. (1990) *Pure & Appl. Chem. 62*, 1067-1070. Metal and anion binding sites in lactoferrin and related proteins.

Baker, E. N., Anderson, B. F., Baker, H. M., Haridas, M., Jameson, G. B., Norris, G. E., Rumball, S. V. & Smith, C. A. (1991) *J. Biol. Macromolecules 13*, 122-129. Structure, function and flexibility in human lactoferrin.

Baker, E. N., Baker, H. M., Smith, C. A., Stebbins, M. R., Kahn, M., Hellström, K. E. & Hellström, I. (1992) *FEBS Lett. 298*, 215-219. Human melanotransferrin (p97) has only one functional iron-binding site.

Baldwin, D. A., Jenny, R. R. & Aisen, P. (1984) *J. Biol. Chem. 259*, 13391-13394. The effect of human serum transferrin and milk lactoferrin on hydroxyl radical formation from superoxide and hydrogen peroxide.

Baldwin, G. S. & Weinstock, J. (1988) *Nucleic Acids Res. 16*, 8720-8724. Nucleotide sequence of porcine liver transferrin.

Bartfeld, N. S. & Law, J. H. (1990) *J. Biol. Chem. 265*, 21684-21691. Isolation and molecular cloning of transferrin from the tobacco hornworm, *Manduca sexta.*

Bates, G. W. and Schlabach, M. R. (1973) *J. Biol. Chem. 248*, 3228-3232. The reaction of ferric salts with transferrin.

Bates, G. W. (1982) in *The biochemistry and Physiology of Iron* (P. Saltman & J. Hegenauer, Eds.) pp 3-18, Elsevier, New York. Metal ion and anion exchange reactions of serum transferrin: The role of quaternary complexes and conformational transitions.

Bertini, I., Canti, G. & Luchinat, C. (1982) *Inorg. Chim. Acta 67*, L21-L23. Preparation and characterization of the vanadium(III) derivative of transferrin.

Bertini, I., Luchinat, C. & Messori, L. (1983) *J. Am. Chem. Soc. 105*, 1347-1350. $^{205}$Tl as an NMR probe for the investigation of transferrin.

Bertini, I., Luchinat, C. & Messori, L. (1985) *J. Inorg. Biochem. 25*, 57-60. Spectral characterization of vanadium-transferrin systems.

Bertini, I., Luchinat, C., Messori, L., Scozzafava, A., Pellacani, G. C. & Sola, M. (1986a) *Inorg. Chem. 25*, 1782-1786. $^{13}$C NMR study of the synergistic anion in transferrins.

Bertini, I., Luchinat, C., Messori, L., Monnanni, R. & Scozzafava, A. (1986b) *J. Biol. Chem 261*, 1139-1146. The metal-binding properties of ovotransferrin. An investigation of cobalt(II) derivatives.

Bertini, I., Messori, L., Pellacani, G. C. & Sola, M. (1988) *Inorg. Chem. 27*, 761-762. Evidence of a metal-synergistic anion bond in thallium(III) transferrin.

Bezkorovainy, A. (1980) *Biochemistry of Non-heme Iron*, Chapter 4, pp 127-204, Plenum, New York. Chemistry and metabolism of the transferrins.

Blakeborough, P., Salter, D.N. and Gurr, M.I. (1983) *Biochem. J. 209*, 505-512. Zinc binding in cow's milk and human milk.

Bluard-Deconinck, J-M., Williams, J., Evans, R.W., Van Snick, J., Osinski, P.A. & Masson, P.L. (1978) *Biochem. J. 171*, 321-327. Iron-binding fragments from the N-terminal and C-terminal regions of human lactoferrin.

Blundell, T. L. & Johnson, L. N. (1976) *Protein Crystallography*, Academic Press, Oxford.

Brock, J.H. (1985) in *Metalloproteins Part 2: Metal proteins with non-redox roles* (P. M. Harrison, Ed) pp 183-262, Macmillan, London. Transferrins.

Brock, J. H., Arzabe, F., Lampreave, F. & Piñeiro, A. (1976) *Biochim. Biophys. Acta 446*, 214-225. The effect of trypsin on bovine transferrin and lactoferrin.

Brown, E.M. & Parry, R.M. (1974) *Biochemistry 13*, 4560-4565. A spectroscopic study of bovine lactoferrin.

Bullen, J. J., Rogers, M. J. & Leigh, L. (1972) *Br. Med. J. 3*, 69-75. Iron proteins in milk and resistance to *Escherichia coli* infection in babies.

Bullen, J. J., Rogers, M. J. & Griffiths, E. (1974) in *Microbial Iron Metabolism. A Comprehensive Treatise* (J. B. Neilands, Ed.) pp 517-551, Academic Press, New York. Bacterial Iron metabolism in infection and immunity.

Butler, A., Danzitz, M.J. & Eckert, H. (1987) *J. Am. Chem. Soc. 109*, 1864-1865. $^{51}V$ NMR as a probe of metal ion binding in metalloproteins.

Campbell, R. F. & Chasteen, N. D. (1977) *J. Biol. Chem. 252*, 5996-6001. An anion binding study of vanadyl(IV) human serotransferrin. Evidence for direct linkage to the metal.

Cantley, L. C. & Aisen, P. (1979) *J. Biol. Chem. 254*, 1781-1784. The fate of cytoplasmic vanadium.

Carmichael, A. & Vincent, J. S. (1979) *FEBS letters 105*, 349-352. Low pH binding of vanadyl to bovine lactoferrin. Electron spin resonance studies.

Casey, J. D. & Chasteen, N. D. (1980) *J. Inorg. Biochem. 13*, 111-126. Vanadyl conalbumin. I. Metal binding site configurations.

Chasteen, N. D., White, L. K. & Campbell, R. F. (1977) *Biochemistry 16*, 363-368. Metal site comformational states of vanadyl(IV) human serotransferrin complexes.

Chasteen, N.D. (1981) in *Biological Magnetic Resonance*, Vol. 3 (L. J. Berliner and J. Reuben, Eds.) pp 53-119, Plenum Press, New York.

Chasteen, N.D. & Theil, E.C. (1982) *J. Biol. Chem. 257*, 7672-7677. Iron binding by horse spleen apoferritin. A vanadyl(IV) EPR spin probe study.

Chasteen, N.D. (1983) *Struct. Bonding 53*, 107-138. The biochemistry of vanadium.

Chasteen. N. D., Grady, J. K. & Holloway, C. E. (1986a) *Inorg. Chem. 25*, 2754-2760. Characterization of the binding, kinetics, and redox stability of vanadium(IV) and vanadium(V) protein complexes in serum.

Chasteen. N. D., Lord, E. M., Thompson, H. J. & Grady, J. K. (1986b) *Biochim. Biophys. Acta 884*, 84-92. Vanadium complexes of transferrin and ferritin in the rat.

Chung, T. D. Y. (1985) Ph.D. Thesis, University of California, Berkeley.

Churchill, M. R., Davies, G., El-Sayed, M. A., El-Shazly, M. F., Hutchinson, J. P., Rupich, M. W., Watkins, K. O. (1979) *Inorg. Chem., 18*, 2296-2300. Synthesis, physical properties and structure characterization of μ-carbonato-dichlorobis (N,N,N',N'-tetramethyl-1, 3 - propanediamine) dicopper (II), LCuCl($C0_3$)ClCuL, a diamagnetic initiator for the oxidative coupling of phenols by dioxygen.

Cochran, M., Coates, J. & Neoh, S. (1984) *FEBS Lett. 176*, 129-132. The competitive equilibrium between aluminium and ferric ions for the binding sites of transferrin.

Cotton, F. A. & Wilkinson, G. (1988) *Advanced Inorganic Chemistry*, 5th Edition, Wiley, New York.

Cox, T. M., Mazurier, J., Spik, G., Montreuil, J. & Peters, T. J. (1979) *Biochim. Biophys. Acta 588*, 120-128. Iron binding proteins and influx of iron across the duodenal brush border.

Davidsson, L., Lonnerdal, B., Sanstroem, B., Kunz, C. & Keen, C. L. (1989) *J. Nutri. 119*, 1461-1464. Identificatioin of transferrin as the major plasma carrier protein for manganese introduced orally or intravenously or aftert *in vitro* addition in the rat.

DeLucas, L. J., Suddath, F. L., Gams, R. A. & Bugg, C. E. (1978) *J. Mol. Biol. 123*, 285-286. Preliminary X-ray study of crystals of human transferrin.

Derewenda, Z. S. & Derewenda, U. (1991) *Biochem. Cell Biol. 69*, 842-851. Relationships among serine hydrolases: evidence for a common structural motif in triacylglyceride lipases and esterases.

Donovan, J. W. & Ross, K. D. (1975) *J. Biol. Chem. 250*, 6022-6025. Nonequivalence of the metal binding sites of conalbumin. Calorimetric and spectrophotometric studies of aluminium binding.

Dubler, E., Korber, P. & Oswald, H. R. (1973) *Acta Cryst. B29*, 1929-1933. The crystal structures of barium hexahydroxocuprate (II) and strontium hexahydroxocuprate, $Ba_2[Cu(OH)_6]$ and $Sr_2[Cu(OH)_6]$.

Duggan, M., Ray, N., Hathaway, B. J., Tomlinson, G., Brint, P. & Pelin, K. (1980) *J.C.S. Dalton Trans.*, 1342-1348. Crystal structures and electronic properties of ammine [tris(2-aminoethyl)amine]copper(II) diperchlorate and potassium penta-ammine-copper(II) tris(hexafluorophosphate).

Duffield, J. R. & Taylor, D. M. (1987) *Inorg. Chim. Acta 140*, 365-367. A spectroscopic study on the binding of plutonium(IV) and its chemical analogues to transferrin.

Eaton, S. S., Dubach, J., More, K. M., Eaton, G. R., Thurman, G. & Ambruso, D. R. (1989) *J. Biol. Chem. 264*, 4776-4781. Comparison of the electron spin echo envelope modulation (ESEEM) for human lactoferrin and transferrin complexes of copper(II) and vanadyl ions.

Eaton, S. S., Dubach, J., Eaton, G. R., Thurman, G. & Ambruso, D. R. (1990) *J. Biol. Chem. 265*, 7138-7141. Electron spin echo envelope modulation evidence for carbonate binding to iron(III) and copper(II) transferrin and lactoferrin.

Englund, H., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Soderberg, B.-O., Taipa, O., Branden, C.-I. & Akeson, A. (1976) *J. Mol. Biol. 102*, 27-59. Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution.

Eriksson, A. E., Kylsten, P. M., Jones, T. A. & Liljas, A. (1988) *Proteins 4*, 283-293. Crystallographic studies of inhibitor binding sites in human carbonic anhydrase II: A pentacoordinated binding of the $SCN^-$ ion to the zinc at high pH.

Evans, R. W. & Williams, J. (1978) *Biochem. J. 173*, 543-552. Studies on the binding of different iron donors to human serum transferrin and isolation of iron-binding fragments from the N- and C-terminal regions of the protein.

Evans, R. W. & Madden, A. D. (1984) *Biochem. Soc. Trans. 12*, 661-662. A low-molecular-weight iron-binding fragment from duck ovotransferrin.

Farrand, A., Gregson, A. K., Skelton, B. W. & White, A. H. (1980) *Aust J. Chem, 33*, 431-436. Crystal structure of the ferromagnetic polymer potassium bis(carbonato)cuprate(II).

Feeney, R. E. & Komatsu, S. K. (1966) *Struct. Bonding 1*, 149-206. The transferrins.

Figgis, B. N. (1987) in *Comprehensive Coordination Chemistry*, (G. Wilkinson, R. D. Gillard, J. A. McCleverty, Eds.) Vol 1, Theory and Background, pp 213-279, Pergamon Press, Oxford. Ligand field theory.

Fitzgerald, J. F. and Chasteen, N. D. (1974) *Anal. Biochem. 60*, 170-180. Determination of the vanadium content of protein solutions by electron paramagnetic resonance spectroscopy.

Fitzgerald, W., Foley, J., McSweeney, D., Ray, N., Sheahan, D., Tyagi, S., Hathaway, B. & O'Brien, P. (1982) *J. C. S. Dalton Trans.*, 1117-1121. Electronic properties and crystal structure of (2,2'-bipyridyl)-*catena*-μ-(oxalato-$O^1O^2$:$O^{1'}O^{2'}$)-copper(II) dihydrate and aqua(2,2'-bipyridyl)-(oxalato-$O^1O^2$)copper(II) dihydrate.

Fitzpatrick, M. (1988) BSc.(Hons) Thesis, Massey University.

Fox, G. C. & Holmes, K. C. (1966) *Acta Cryst. 20*, 886-891. An alternative method of solving the layer scaling equations of Hamilton, Rollett and Sparks.

Gaber, B. P., Schillinger, W. E., Koenig, S. H. & Aisen, P. (1970) *J. Biol. Chem. 245*, 4251-4255. Nuclear magnetic relaxation dispersion in protein solutions.

Gaber, B. P., Miskowski, V. & Spiro, T. G. (1974). *J. Am. Chem. Soc. 96*, 6868-6873. Resonance Rama scattering from iron(III)- and copper(II)-transferrin and an iron(III) model compound. A spectroscopic interpretation of the transferrin binding site.

Garnier-Suillerot, A., Albertini, J.-P., Collet, A., Faury, L., Pastor, J.-M. & Tosi, L. (1981) *J. C. S. Dalton Trans.*, 2544-2549. Spectrophotometric studies of the copper(II)-D-*o*-tyrosine complex. Assignment of the 330 nm dichroic band in copper(II) and iron(III) transferrins.

Garratt, R. C., Evans, R. W., Hasnain, S. S. & Lindley, P. F. (1986) *Biochem. J. 233*, 479-484. An extended X-ray absorption fine structure investigation of diferric transferrins and their iron-binding fragments.

Garratt, R. C., Evans, R. W., Hasnain, S. S., Lindley, P. F. & Sarra, R. (1992) *Biochem. J. 280*, 151-155. X.a.f.s. studies on chicken dicupric ovotransferrin.

Gelb, M. H. & Harris, D. C. (1980) *Arch. Biochem. Biophys. 200*, 93-98. Correlation of proton release and ultraviolet difference spectra associated with metal binding by transferrin.

Gilliland, G. L., & Quiocho, F. A. (1981) *J. Mol. Biol. 146*, 341-362. Structure of the L-arabinose-binding protein from *Escherichia coli* at 2.4 Å resolution.

Gleizes, A., Maury, F. & Galy, J. (1980) *Inorg. Chem. 19*, 2074-2078. Crystal structure of sodium bis(oxalato)cuprate dihydrate, $Na_2Cu(C_2O_4)_2{\cdot}2H_2O$. A deductive proposal for the structure of copper oxalate $CuC_2O_4{\cdot}xH_2O$ ($0 \leq x \leq 1$).

Glusker, J. P. (1991) *Advances in protein chemistry 42*, pp 1-76. Structural aspects of metal liganding to functional groups in proteins.

Graham, I. & Williams, J. (1975) *Biochem. J. 145*, 263-279. A comparison of glycopeptides from the transferrins of several species.

Gray, H. B. & Malmström, B. G. (1983) *Comments Inorg. Chem. 2*, 203-209. On the relationship between protein-forced ligand fields and the properties of the blue copper centers.

Greene F. C. & Feeney, R. E. (1968) *Biochemistry 7*, 1366-1371. Physical evidence for transferrins as single polypeptide chains.

Gorinsky, B., Horsburgh, C., Lindley, P. F., Moss, D. S., Parkar, M. & Watson, J. L. (1979) *Nature (London) 281*, 157-158. Evidence for the bilobal nature of diferric rabbit plasma transferrin.

Hanson, J. C., Watenpaugh, K. D., Seiker, L. & Jensen, L. H. (1979) *Acta Cryst. A35*, 616-621. A limited-range step-scan method for collecting X-ray diffraction data.

Harris, A. W. & Sephton, R. G. (1977) *Cancer Res. 37*, 3634-3638. Transferrin promotion of $^{67}Ga$ and $^{59}Fe$ uptake by cultured mouse myeloma cells.

Harris, D. C., Gray, G. A. & Aisen, P. (1974) *J. Biol. Chem. 249*, 5261-5264. $^{13}C$ nuclear magnetic resonance study of the spatial relation of the metal- and anion-binding sites of human transferrin.

Harris, D. C. (1977) *Biochemistry 16*, 560-564. Different metal-binding properties of the two sites of human transferrin.

Harris, D. C. and Gelb, M. H. (1980) *Biochim. Biophys. Acta 623*, 1-9. Binding of xylenol orange to transferrins. Demonstration of metal anion linkage.

Harris, W. R, Carrano, C. J., Pecoraro, V. L. & Raymond, K. N. (1981) *J. Am. Chem. Soc. 103*, 2231-2237. Siderophilin metal coordination. I. Complexation of thorium by transferrin: structure-function implications.

Harris, W. R. (1983) *Biochemistry 22*, 3920-3926. Thermodymanic binding constants of the zinc-human serum transferrin complex.

Harris, W. R. & Pecoraro, V. L. (1983) *Biochemistry 22*, 292-299. Thermodymanics of gallium binding by human transferrin.

Harris, W. R. & Carrano, C. J. (1984) *J. Inorg. Biochem. 22*, 201-218. Binding of vanadate to human serum transferrin.

Harris, W. R. (1985) *Biochemistry 24*, 7412-7418. Thermodynamics of anion binding to human serum transferrin.

Harris, W. R. (1986a) *Biochemistry 25*, 803-808. Thermodymanics of gallium binding by human lactoferrin.

Harris, W. R. (1986b) *Inorg. Chem. 25*, 2041-2045. Binding constants for neodymium(III) and samarium(III) with human serum transferrin.

Harris, W. R. (1986c) *J. Inorg. Biochem. 27*, 41-52. Estimation of the ferrous-transferrin binding constant based on studies of nickel(II)-transferrin.

Harris, W. R. & Madsen, L. J. (1988) *Biochemistry 27*, 284-288.

Harris, W. R. & Sheldon, J. (1990) *Inorg. Chem. 29*, 119-124. Equilibrium constants for the binding of aluminium to human serum transferrin.

Harris, W. R. & Chen, Y. (1991) *J. Inorg. Biochem. 43*, Abstract B016 from the 5th International Conference on Bioinorganic Chemistry, Oxford, UK. Equilibrium studies on the binding of gadolinium(III) to human serum transferrin.

Hart, F. A. (1987) in *Comprehensive Coordination Chemistry*, (G.Wilkinson, R. D. Gillard, J. A.McCleverty, Eds.) Vol 3, Main Group & Early Transition Elements, pp.1059-1127, Pergamon Press, Oxford. Scandium, yttrium and the lanthanides.

Hasnain, S. S., Evans, R. W., Garratt, R. C. & Lindley, P. F. (1987) *Biochem. J. 233*, 369-375. An extended-X-ray-absorption-fine-structure study of freeze-dried and solution ovotransferrin.

Hathaway, B. J. & Billing, E. E. (1970) *Coord. Chem. Rev. 5*, 143-207. The electronic properties and stereochemistry of mono-nuclear complexes of the copper(II) ion.

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

Healy, P. C. & White, A. H. (1972) *J.C.S. Dalton Trans.*, 1913-1917. Crystal structure and physical properties of anhydrous sodium copper carbonate.

Heaphy, S. & Williams, J. (1982) *Biochem. J. 205*, 611-617. The preparation and partial characterization of N-terminal and C-terminal iron-binding fragments from rabbit serum transferrin.

Hendrickson, W. A. & Konnert, J. H. (1980) in *Biomolecular structure, function conformation and evolution* (R. Srinivasan, Ed) Vol 1, pp. 43-57, Pergamon Press, Oxford.

Herzberg, O. & James, M. N. G. (1985) *Biochemistry 24*, 5298-5302. Common structural framework of the two $Ca^{2+}/Mg^{2+}$ binding loops of troponin C and other $Ca^{2+}$ binding proteins.

Hol, W. G. J., van Duijnen, P. T. & Berendson, H. J. C. (1978) *Nature (London) 273*, 443-446. The alpha-helix dipole and the properties of proteins.

Huebers, H. A., Huebers, E., Finch, C. A. & Martin, A. W. (1982) *J. Comp. Physiol. 148B*, 101-109. Characterization of an invertebrate transferrin from the crab *Cancer magister* (arthropoda)

Huebers, H. A. & Finch, C. A. (1987) *Physiol. Rev. 67*, 520-582. The physiology of transferrin and transferrin receptors.

Huheey, J. E. (1978) *Inorganic Chemistry. Principles of structure and reactivity*, 2nd Edition. Harper & Row, New York.

Jeltsch, J. M. & Chambon, P. (1982) *Eur. J. Biochem. 122*, 291-295. The complete nucleotide sequence of the chick ovotransferrin mRNA.

Jhoti, H., Gorinsky, B., Garratt, R. C., Lindley, P. F., Walton, A. R. & Evans, R. W. (1988) *J. Mol. Biol. 200*, 423-425. Crystallization and preliminary analysis of an 18,000 $M_r$ fragment of duck ovotransferrin.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268-272. A graphics model building and refinement system for macromolecules.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Cryst. A47*, 110-119. Improved methods for building protein models in electron density maps and the location of errors in these models.

Kabsch, W. & Sander, C. (1983) *Biopolymers 22*, 2577-2637. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.

Keung, W.-M., Azari, P. & Phillips, J. L. (1982) *J. Biol. Chem. 257*, 1177-1183. Structure and function of ovotransferrin. I. Production of iron-binding fragments from iron-ovotransferrin by the action of immobilised subtilisin. Purification and characterisation of the fragments.

Kojima, N. & Bates, G. W. (1981) *J. Biol. Chem. 256*, 12034-12039. The formation of $Fe^{3+}$-transferrin-$CO_3^{2-}$ via the binding and oxidation of $Fe^{2+}$.

Laemilli, U. K. (1970) *Nature (London) 227*, 680-685. Cleavage of structural proteins during the assembly of the head of bacteriophage T4.

Lee, M. Y., Huebers, H., Martin, A. W. & Finch, C. A. (1978) *J. Compt. Physiol. 127*, 349-354. Iron metabolism in the spider, *Dugesiella hentzi*.

Legrand, D., Mazurier, J., Metz-Boutigue, M-H., Jollés, J., Jollés, P., Montreuil, J. & Spik, G. (1984) *Biochim. Biophys. Acta 787*, 90-96. Characterization and localization of an iron-binding 18-kDa glycopeptide isolated from the N-terminal half of human lactotransferrin.

Lehrer, S. S. (1969) *J. Biol. Chem. 244*, 3613-3617. Fluorescence and absorption studies of the binding of copper and iron to transferrin.

Levitt, M. & Chothia, C. (1976) *Nature (London) 261*, 552-558. Structural patterns in globular proteins.

Lonnerdal, B. (1984) in *Iron Nutrition in Infancy and Childhood* (A. Stekel, Ed.) pp 95-117, Vevey/Raven Press, New York.

Luecke, H. & Quiocho, F. A. (1990) *Nature (London) 347*, 402-406. High specificity of a phosphate transport protein determined hy hydrogen bonds.

Luk, C. K. (1971) *Biochemistry 10*, 2838-2843. Study of the nature of the metal-binding sites and estimate of the distance between the metal-binding sites in transferrin using trivalent lanthanide ions as fluorescent probes.

Luzzati, V. (1952) *Acta Cryst. 5*, 802-810. Traitement statistiques des erreurs dans la determination des structures cristallines.

Lyle, S. J. and Rahman, M. M. (1963) *Talanta 10*, 1177-1182. Complexiometric titrations of yttrium and the lanthanons - 1.

McFadden, D. L., McPhail, A. T., Gross,P. M., Garner, C. D. & Maabs, F. E. (1975) *J. C. S. Dalton Trans.* 263-268. Crystal and molecular structure, electron spin resonance and electronic spectrum of hexakis(imidazole)copper(II) nitrate.

MacGillivray, R. T. A., Mendez, E., Shewale, J. G., Sinha, S. K., Lineback-Zins, J. & Brew, K. (1983) *J. Biol. Chem. 258*, 3543-3553. The primary structure of human serum transferrin.

Martell, A. E. & Smith, R. M. (1976) *Critical Stability Constants*, Plenum, New York.

Martin, A. W., Huebers, E., Huebers, H., Webb, J. & Finch, C. A. (1984) *Blood 64*, 1048-1052. A monosited transferrin from a representative deuterosome: The ascidian *Pyura stolonifera* (subphylum Urochordata).

Martin, R. B., Savory, J., Brown, S., Bertholf, R. L. & Wills, M. R. (1987) *Clin. Chem. 33*, 405-407. Transferrin binding of $Al^{3+}$ and $Fe^{3+}$.

Masson, P. L., Heremans, J. F. & Dive, C. (1966) *Clin. Chim. Acta 14*, 735-739. An iron binding protein common to many external secretions.

Matthews, B. W. (1972) Macromolecules 5, 818-819. The γ turn. Evidence for a new folded conformation in proteins.

Mazurier, J., Aubert, J.-P., Loucheux-Lefevre, M.-H. & Spik, G. (1976) *FEBS Lett. 66*, 238-242. Comparative circular dichroism studies of iron-free and iron-saturated forms of human serotransferrin and lactotransferrin.

Mazurier, J. & Spik, G. (1980) *Biochim. Biophys. Acta 629*, 399-408. Comparative study of the iron-binding properties of human transferrins. I. Complete and sequential iron saturation and desaturation of the lactotransferrin.

Mazurier, J., Metz-Boutigue, M.-H., Jollés, J., Spik, G., Montreuil, J. & Jollés, P. (1983) *Experientia 39*, 135-141. Human lactotransferrin: Molecular, functional and evolutionary comparisons with human serum transferrin and hen ovotransferrin.

Meares, C. F. & Ledbetter, J. E. (1977) *Biochemistry 16*, 5178-5180. Energy transfer between terbium and iron bound to transferrin: Reinvestigation of the distance between metal-binding sites.

Metz-Boutigue, M.-H., Jollés, J., Mazurier, J., Schoentgen, F., Legrand, D., Spik, G., Montreuil, J. & Jollés, P. (1984) *Eur. J. Biochem. 145*, 659-676. Human lactotransferrin: amino acid sequence and structural comparisons with other transferrins.

Meyer, M. H., Singh, P., Hatfield, W. E. & Hodgson, D. J. (1972) *Acta Cryst. B28*, 1607-1613. The crystal and molecualr structure of carbonatodiammine copper(II), $Cu(NH_3)_2CO_3$.

Mikami, B. & Hirose, M. (1990) *J. Biochem. 108*, 907-908. Crystallization of and preliminary crystallographic data for the N-terminal half-molecule of ovotransferrin.

Moeller, T. (1973) in *Comprehensive Inorganic Chemistry*, ( , Eds.), Vol. 4, pp1-101, Pergamon Press, Oxford. Lanthanides.

Moeller, T. (1983) *Inorganic Chemistry. A modern introduction.* Wiley, New York.

Montreuil, J., Mazurier, J., Legrand, D. & Spik, G. (1985) in *Proteins of Iron Storage and Transport*, (G. Spik, J. Montreuil, R. R. Crichton and J. Mazurier, Eds.) pp 25-38, Elsevier, New York. Human lactotransferrin: structure and function.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992) *Proteins 12*, 345-364. Stereochemical quality of protein structure coordinates.

Moskaitis, J. E., Pastori, R. L. & Schoenberg, D. R. (1991) *Nucleic Acids Res. 18*, 6135-6138. The nucleotide sequence of *Xenopus laevis* transferrin mRNA.

Mowbray, S. L. & Cole, L. B. (1992) *J. Mol. Biol. 225*, 155-175. 1.7 Å X-ray structure of the periplasmic ribose receptor from *Escherichia coli.*

Musci, G., Reed, G. H. & Berliner, L. J. (1986) *J. Inorg. Biochem. 26*, 229-236. An electron paramagnetic resonance study of bovine α-lactalbumin-metal ion complexes.

Nagy, B. & Lehrer, S. S. (1972) *Arch. Biochem. Biophys. 148*, 27-36. Circular dichroism of iron, copper and zinc complexes of transferrin.

Najarian, R. C., Harris, D. C. & Aisen, P. (1978) *J. Biol. Chem. 253*, 38-42. Oxalate and spin labelled oxalate as probes of the anion binding site of human transferrin.

Norris, G. E., Anderson. B. F., Baker, E. N., Baker, H. M., Gartner, A. L., Ward, J. & Rumball, S. V. (1986) *J. Mol. Biol. 191*, 143-145. Preliminary crystallographic studies on bovine lactoferrin.

Norris, G. E., Baker, H. M. & Baker, E. N. (1989) *J. Mol. Biol. 209*, 329-331. Preliminary crystallographic studies on human apolactoferrin in its native and deglycosylated forms.

North, A. C. T., Phillips, D. C. & Mathews, F. S. (1968) *Acta Cryst. A24*, 351-359. A semi-empirical method of absorption correction.

Octave, J.-N., Schneider, Y.-J., Trouet, A. & Crichton, R. R. (1983) *Trends Biochem. Sci. 8*, 217-220. Iron uptake and utilization by mammalian cells. I: Cellular uptake of transferrin and iron.

O'Hara, P. B., Yeh, S. M., Meares, C. F. & Bersohn, R. (1981) *Biochemistry 20*, 4704-4708. Distance between metal-binding sites in transferrin: Energy transfer from bound terbium(III) to iron(III) or manganese(III).

O'Hara, P. B. & Bersohn, R. (1982) *Biochemistry 21*, 5269-5272. Resolution of the two metal binding sites of human serum transferin by low-temperature excitation of bound europium(III).

O'Hara, P. B. & Koenig, S.H. (1986) *Biochemistry 25*, 1445-1450. Electron spin resonance and magnetic relaxation studies of gadolinium(III) complexes of human transferrin.

Oldham, C. (1987) in in *Comprehensive Coordination Chemistry*, (G.Wilkinson, R. D. Gillard, J. A.McCleverty, Eds.) Vol 2, Ligands, pp 435-459, Pergamon Press, Oxford. Carboxylates, squarates and related species.

Palmour, R. M. & Sutton, H. E. (1971) *Biochemistry 10*, 4026-4032. Vertebrate transferrins. Molecular weights, chemical composition and iron-binding studies.

Patch, M. G. & Carrano, C. J. (1981) *Inorg. Chim. Acta 56*, L71-L73. The origin of the visible absorption in metal transferrins.

Pecoraro, V. L., Harris, W. R., Carrano, C. J. & Raymond, K. N. (1981) *Biochemistry 20*, 7033-7039. Siderophilin metal coordination. Difference ultraviolet spectroscopy of di-, tri-, and tetravalent metal ions with ethylenebis[(o-hydroxyphenyl)glycine]

Pentecost, B. T. & Teng, C. T. (1987) *J. Biol. Chem. 262*, 10134-10139. Lactoferrin is the major estrogen inducible protein of mouse uterine secretions.

Pierce, A., Colavizza, D., Benaisser, M., Maes, P., Tartar, A., Montreuil, J. & Spik, G. (1991) *Eur. J. Biochem. 196*, 177-184. Molecular cloning and sequence analysis of bovine lactotransferrin.

Pflugrath, J. W. & Quiocho, F. A. (1985) *Nature (London) 314*, 257-260. Sulphate sequestered in the sulphate-binding protein of *Salmonella typhimurium* is bound solely by hydrogen bonds.

Pflugrath, J. W. & Quiocho, F. A. (1988) *J. Mol. Biol. 200*, 163-180. The 2 Å resolution structure of the sulfate-binding protein involved in active transport in *Salmonella typhimurium*.

Plowman, J. E. (1979) Ph.D. Thesis, Massey University, New Zealand.

Pollack, S., Weaver, J. & Zhan, H. (1991) Abstract from the *10th International Conference on Iron and Iron Proteins*, Oxford, U.K. Intracellular iron.

Prados, R., Boggess, R. K., Martin, R. B. & Woodworth, R. C. (1975) *Bioinorg. Chem. 4*, 135-142. Fe(III) and Cu(II) conalbumin visible circular dichroism spectra.

Priestle, J. P. (1988) *J. Appl. Crystallogr. 21*, 572-576. RIBBON: a stereocartoon drawing program for proteins.

Quarterman, J. (1983) *Proc. Nutr. Soc. 42*, 45A. Milk, lactoferrin and heavy metal absorption.

Ramakrishnan, C. & Ramachandran, G. N. (1965) *Biophys. J. 5*, 909-933. Stereochemical criteria for polypeptide and protein chain conformation.

Rawas, A., Moreton, K., Muirhead, H. & Williams, J. (1988) *J. Mol. Biol. 208*, 213-214. Preliminary crystallographic studies on duck ovotransferrin.

Rees, D. C., Lewis, M., Lipscomb, W. N. (1983) *J. Mol. Biol. 168*, 367-387. Refined crystal structure of carboxypeptidase A at 1.54 Å resolution.

Rogers, T. B., Børreson, T. & Feeney, R. E. (1978) *Biochemistry 17*, 1105-1109. Chemical modification of the arginines in transferrins.

Rose, T. M., Plowman, G. D., Teplow, D. B., Dreyer, W. J., Hellstrom, K. E. & Brown, J. P. (1986) *Proc. Natl. Acad. Sci. USA 83*, 1261-1265. Primary structure of the human melanoma-associated antigen p97 (melanotransferrin) deduced from the mRNA sequence.

Rypniewski, W. R., Breiter, D. R., Benning, M. M., Wesenberg, G., Oh, B.-H., Markley, J. L., Rayment, I. & Holden, H. M. (1991) *Biochemistry 30*, 4126-4131. Crystallization and structure determination to 2.5 Å resolution of the oxidised [2Fe-2S] ferredoxin isolated from *Anabaena* 7120.

Sack, J. S, Saper, M. A. & Quiocho, F. A. (1989) *J. Mol. Biol. 206*, 171-191. Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine.

Saper, M. A. & Quiocho, F. A. (1983) *J. Biol. Chem. 258*, 11057-11062. Leucine, isoleucine, valine-binding protein from *Escherichia coli*. Structure at 3.0 Å resolution and location of the binding site.

Sarra, R. & Lindley, P. F. (1986) *J. Mol. Biol. 188*, 727-728. Preliminary X-ray data for an N-terminal fragment of rabbit serum transferrin.

Sarra, R., Garratt, R., Gorinsky, B, Jhoti, H. & Lindley, P. (1990) *Acta Cryst. B46*, 763-771. High resolution X-ray studies on rabbit serum transferrin: Preliminary structure analysis of the N-terminal half-molecule at 2.3 Å resolution.

Schade, A. L. (1985) in *Proteins of Iron Storage and Transport* (G. Spik, J. Montreuil, R. R. Crichton and J. Mazurier, Eds.) pp 3-12, Elsevier, New York. Conalbumin and siderophilin as iron-binding proteins: A review of their discovery.

Schade, A. L. & Reinhart, R. W. (1966) *Prot. Biol. Fluids 14*, 75-81. Carbon dioxide in the iron and copper siderophilin complexes.

Schlabach, M. R. & Bates, G. W. (1975) *J. Biol. Chem. 250*, 2182-2188. The synergistic binding of anions and $Fe^{3+}$ by transferrin.

Schneider, D. J., Roe, A. L., Mayer, R. J. & Que, L., Jr. (1984) *J. Biol. Chem. 259*, 9699-9703. Evidence for synergistic anion binding to iron in ovotransferrin complexes from resonance Raman and extended X-ray absorption fine structure analysis.

Shannon, R. D. (1976) *Acta Cryst. A32*, 751-767. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides.

Shepard, W. E. B., Anderson, B. F., Lewandoski, D. A., Norris, G. E. & Baker, E. N. (1990) *J. Am. Chem. Soc. 112*, 7817-7819. Copper coordination geometry in azurin undergoes minimal change on reduction of copper(II) to copper(I).

Shepard, W. E. B. (1991) PhD Thesis, Massey University, New Zealand.

Shongwe, M. S., Smith, C. A., Ainscough, E. W., Baker, H. M., Brodie, A. M. & Baker, E. N. (1992) *Biochemistry 31*, 4451-4458. Anion binding by human lactoferrin: Results from crystallographic and physicochemical studies.

Shriver, D. F. (1969) *Manipulations of Air-sensitive Compounds*, McGraw-Hill, New York.

Sibanda, B. L., Blundell, T. L. & Thornton, J. M. (1989) *J. Mol. Biol. 206*, 759-777. Conformation of β-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering.

Sillen, L. G. & Martell, A. E. (1971) *Stability constants of metal-ion complexes*, Special Publication No. 25, The Chemical Society, London.

Spik, G., Bayard, B., Fournet, B., Strecker, G., Bouquelet, S. and Montreuil, J. (1975) *FEBS Lett. 50*, 296-299. Studies on glycoconjugates. LXIV. Complete structure of two carbohydrate units of human serum transferrin.

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. Primary structuer of the glycans from human lactotransferrin.

Stjernholm, R., Warner, F. W., Robinson, J. W., Ezekiel, E. & Katayama, N. (1978) *Bioinorg. Chem. 9*, 277-280. Binding of platinum to human transferrin.

Tan, A. T. & Woodworth, R. C. (1969) *Biochemistry 8*, 3711-3716. Ultraviolet difference specral studies of conalbumin complexes with transition metal ions.

Tan, A. T. & Woodworth, R. C. (1970) *J. Polymer Sci. (C) 30*, 599-606. Differences in absorption and emission properties of conalbumin and metal-substituted conalbumin.

Teuwissen, B., Masson, P. L., Osinski, P. & Heremans, J. F. (1972) *Eur. J. Biochem. 31*, 239-245. Metal-combing properties of human lactoferrin.

Tomimatsu, Y. & Vickery, L. E. (1972) *Biochim. Biophys. Acta 285*, 72-83. Circular dichroism studies of human serum transferrin and chicken ovotransferrin and their copper complexes.

Tomimatsu, Y. & Donovan, J. W. (1976) *FEBS Lett. 71*, 299-302. Spectroscopic evidence for perturbation of tryptophan in Al(III) and Ga(III) binding to ovotransferrin and human serum transferrin.

Tomimatsu, Y., Kint, S. & Scherer, J. R. (1976) *Biochemistry 15*, 4918-4924. Resonance Raman spectra of iron(III)-, copper(II)-, cobalt(III)-, and manganese(III)-transferrins and of bis(2,4,5-trichlorophenolato)diimidazolecopper(II) monohydrate, a possible model for copper(II) binding to transferrins.

Trapp, G. A. (1983) *Life Sciences 33*, 311-316. Plasma aluminium is bound to transferrin.

Vilas-Boas, & Costa Pessoa (1987) in *Comprehensive Coordination Chemistry*, (G.Wilkinson, R. D. Gillard, J. A.McCleverty, Eds.) Vol 3, Main Group & Early Transition Elements, pp. 453-583, Pergamon Press, Oxford. Vanadium.

Vogel, A. I. (1961) *Quantitative Inorganic Analysis including Instrumental Analysis*, 3rd Edition. Longmans, London.

Vreugdenhil, W. Birker, P. J. M. W. L., ten Hoeldt, R. W. M., Verschoor, G. C. & Reedijk, J. (1984) *J. C. S. Dalton Trans.*, 429-432. Tetrakis(imidazole)copper(II) fluorides. The crystal and molecular structure of diaquatetrakis(imidazole)copper(II) difluoride.

Vyas, N. K., Vyas, M. N. & Quiocho, F. A. (1983) *Proc. Natl. Acad. Sci. USA 80*, 1792-1796. The 3 Å resolution structure of a D-galactose-binding protein for transport and chemotaxis in *Escherichia coli*.

Wang, D. Bode, W. & Huber, R (1985) *J. Mol. Biol. 185*, 595-624. Bovine chymotrypsinogen A. X-ray crystal structure analysis and refinement of a new crystal form at 1.8 A resolution.

Warner, R. C. & Weber, I. (1953) *J. Am. Chem. Soc. 75*, 5094-5101. The metal combining properties of conalbumin.

Wetlaufer, D. B. (1962) *Advances in Protein Chemistry 17*, 304-390. Ultraviolet spectra of proteins and amino acids.

Williams, J. (1974) *Biochem. J. 141*, 745-752. The formation of iron-binding fragments of hen ovotransferrin by limited proteolysis.

Williams, J. (1975) *Biochem. J. 149*, 237-244. Iron-binding fragments from the carboxyl-terminal region of hen ovotransferrin.

Williams, J. (1982) *Trends Biochem. Sci. 7*, 394-397. The evolution of transferrin.

Williams, J., Grace, S. A. & Williams, J. M. (1982a) *Biochem. J. 201*, 417-419. Evolutionary significance of the renal excretion of transferrin half-molecule fragments.

Williams, J., Elleman, T. C., Kingston, I. B., Wilkins, A. G. & Kuhn, K. A. (1982b) *Eur. J. Biochem. 122*, 297-303. The primary structure of hen ovotransferrin.

Wilson, A. J. C. (1942) *Nature 150*, 152. Determination of absolute from relative X-ray intensity data.

Windle, J. J., Wiersma, A. K., Clark, J. R. & Feeney, R. E (1963) *Biochemistry* 2, 1341-1345. Investigation of the iron and copper complexes of avian conalbumins and human transferrins by electron paramagnetic resonance.

Winkler, F. K., Schutt, C. E. & Harrison, S. C. (1979) *Acta Cryst. A35*, 901-911. The oscillation method for crystals with very large unit cells.

Wonacott, A. J. (1980) *A suite of programs for on-line evaluation and analysis of integrated intensites on small angle rotation/oscillation photographs*, Cambridge, UK.

Yamamura, T., Hagiwara, S., Nakazato, K. and Satake, K. (1984) *Biochem. Biophys. Res. Commun. 119*, 298-304. Copper complexes at N- and C- site of ovotransferrin: quantitative determination and visible absorption spectrum of each complex.

Zak, O. & Aisen, P. (1988) *Biochemistry 27*, 1075-1080. Spectroscopic and thermodynamic studies on the binding of gadolinium(III) to human serum transferrin.

Zweier, J. L. & Aisen, P. (1977) *J. Biol. Chem. 252*, 6090-6096. Studies of transferrin with the use of $Cu^{2+}$ as an electron paramagnetic resonance spectroscopic probe.

Zweier, J. L. (1980) *J. Biol. Chem. 255*, 2782-2789. Electron paramagnetic resonance studies of the binding of copper to conalbumin. Probes of the structure and properties of the metal and anion binding sites.

Zweier, J. L., Wooton, J. B. & Cohen, J. S. (1981) *Biochemistry 20*, 3505-3510. Studies of anion binding by transferrin using carbon-13 nuclear magnetic resonance spectroscopy.

# Appendix I

## *Metal and anion binding*

Table 1: Metal ions which are known to bind to members of the transferrin family.

| | transferrin | | | ovotransferrin | | | lactoferrin | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| metal | $\lambda_{max}$ [a] | r [b] | ref | $\lambda_{max}$ | r | ref | $\lambda_{max}$ | r | ref | colour |
| $Al^{3+}$ | — | 2 | 1 | — | 2 | 2 | — | 2 | 3 | — |
| $V^{3+}$ | 370 | 2 | 4 | × | × | | × | 2 | 3 | light green |
| | $637^{d\text{-}d}$ | | | | | | | | | — |
| $VO^{2+}$ | $592^{d\text{-}d}$ | 2 | 5 | × | × | | 590 | 2 | 3 | — |
| $VO_2^{+}$ | — | 2 | 6 | × | × | | — | 2 | 3 | — |
| $Cr^{3+}$ | 440 | 2 | 7 | 435 | 2 | 8 | 442 | 2 | 9 | blue/green |
| | 615 | | | | | | | | | |
| $Mn^{2+}$ | × | × | | — | 2 | 8 | — | 2 | 9 | — |
| $Mn^{3+}$ | 430 | 2 | 7 | 430 | 2 | 10 | 435 | 2 | 9 | brown |
| $Fe^{3+}$ | 470 | 2 | 11 | 462 | 2 | 12 | 466 | 2 | 13 | pink/red |
| $Co^{2+}$ | × | × | | $550^{d\text{-}d}$ | 2 | 14 | × | 2 | 9 | — |
| $Co^{3+}$ | 405 | 2 | 7 | 415 | 2 | 10 | 405 | 2 | 9 | brown/yellow |
| $Ni^{2+}$ | — | 2 | 15 | — | 2 | 8 | — | 2 | 9 | — |
| $Cu^{2+}$ | 440 | 2 | 11 | 440 | 2 | 12 | 434 | 2 | 13 | yellow |
| | $685^{d\text{-}d}$ | | 16 | $660^{d\text{-}d}$ | | 16 | 677 | | 17 | |
| $Zn^{2+}$ | — | 2 | 18 | — | 2 | 12 | — | 2 | 9 | — |
| $Ga^{3+}$ | — | 2 | 19 | — | 2 | 20 | — | 2 | 21 | — |
| $Cd^{3+}$ | — | 2 | 15 | — | 2 | 8 | — | 2 | 9 | — |
| $Ce^{3+}$ | × | | | × | × | | 350 | 2 | 3 | — |
| $Ce^{4+}$ | ≈420 | | 22 | × | × | | 446 | 2 | 3 | brown/red |
| $Ln^{3+}$ [c] | — | 1 or 2 | 15 | × | × | | — | 2 | 3 | — |
| $Hf^{4+}$ | — | 2 | 23 | × | × | | × | × | | — |
| $Pt^{3+}$ | — | 2 | 24 | × | × | | × | × | | — |
| $Tl^{3+}$ | — | 2 | 25 | × | × | | × | × | | — |
| $Th^{4+}$ | — | 2 | 26 | × | × | | — | 2 | 3 | — |
| $Pu^{4+}$ | — | 2 [d] | 23 | × | × | | × | × | | — |

[a] Bands resulting from d-d transitions are indicated. All other bands are charge transfer. [b] r = the ratio of the maximum number of metal ions specifically bound per protein molecule. [c] $Ln^{3+}$ refers to the metals La - Lu and also $Y^{3+}$ and $Sc^{3+}$. [d] The titration curve is hard to interpret due to what the authors term "non-specific binding of nitrate".

× indicates no information is available. — indicates no visible absorption for the complex

References for Table 1:


1 Cochran, M., Coates, J. & Neoh, S. (1984) *FEBS Lett. 176*, 129-132.

2 Donovan, J. W. & Ross, K. D. (1975) *J. Biol. Chem. 250*, 6022-6025.

3 This thesis.

4 Bertini, I., Canti, G. & Luchinat, C. (1982) *Inorg. Chim. Acta 67*, L21-L23.

5 Chasteen, N. D., White, L. K. & Campbell, R. F. (1977) *Biochemistry 16*, 363-368.

6 Harris, W. R. & Carrano, C. J. (1984) *J. Inorg. Biochem. 22*, 201-218.

7 Aisen, P., Aasa, R. & Redfield, A. C. (1969) *J. Biol. Chem. 244*, 4628-4633.

8 Tan, A. T. & Woodworth, R. C. (1969) *Biochemistry 8*, 3711-3716.

9 Ainscough, E. W., Brodie, A. M., Plowman, J. E., Bloor, S. J., Loehr, J. S. & Loehr, T. M. (1980) *Biochemistry 19*, 4072-4079.

10 Tomimatsu, Y., Kint, S. & Scherer, J. R. (1976) *Biochemistry 15*, 4918-4924

11 Schade, A. L. & Reinhart, R. W. (1966) *Prot. Biol. Fluids 14*, 75- 81.

12 Warner, R. C. & Weber, I. (1953) *J. Am. Chem. Soc. 75*, 5094-5101.

13 Shongwe, M. S., Smith, C. A., Ainscough, E. W., Baker, H. M., Brodie, A. M. & Baker, E. N. (1992) *Biochemistry 31*,

14 Bertini, I., Luchinat, C., Messori, L., Monnanni, R. & Scozzafava, A. (1986b) *J. Biol. Chem. 261*, 1139-1146.

15 Chung, T. D. Y. (1985) Ph.D. Thesis, University of California, Berkeley.

16 Tomimatsu, Y. & Vickery, L. E. (1972) *Biochim. Biophys. Acta 285*, 72-83.

17 Ainscough, E. W., Brodie, A. M. & Plowman, J. E. (1979) *Inorg. Chim. Acta 33*, 149-153

18 Harris, W. R. & Stenback, J. Z. (1988) *J. Inorg. Biochem. 33*, 211-223.

19 Harris, W. R. & Pecoraro, V. L. (1983) *Biochemistry 22*, 292-299.

20 Aisen, P., Lang, G. & Woodworth, R. C. (1973) *J. Biol. Chem. 248*, 549-653.

21 Harris, W. R. (1986) *Biochemistry 25*, 803-808.

22 Yamamura, T., Ichimura, K., Tsuda, T., Hayashi, A., Taniguchi, T., Toi, R., Kihara, H. & Satake, K. (1988) *Nippon Kagaku Kaishi 4*, 452-458 (Japanese); cited in Chemical Abstracts 109: **34026x**.

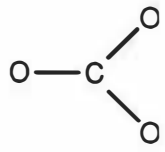
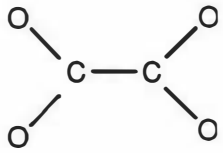
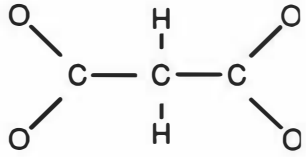
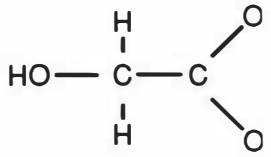
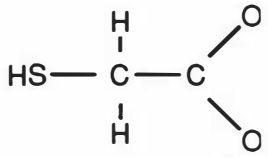
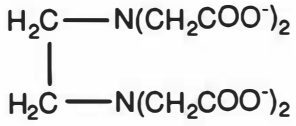
23 Duffield, J. R. & Taylor, D. M. (1987) *Inorg. Chim. Acta 140*, 365-367.

24 Stjernholm, R., Warner, F. W., Robinson, J. W., Ezekiel, E. & Katayama, N. (1978) *Bioinorg. Chem. 9*, 277-280.

25 Bertini, I., Luchinat, C. & Messori, L. (1983) *J. Am. Chem. Soc. 105*, 1347-1350.

26 Harris, W. R, Carrano, C. J., Pecoraro, V. L. & Raymond, K. N. (1981) *J. Am. Chem. Soc. 103*, 2231-2237.

Table 2: Some synergistic anions for the transferrins

| anion | $\lambda_{max}$ (nm) | Structure |
|---|---|---|
| carbonate | 465 [a] |  |
| oxalate | 465 [a]<br>476 [b]<br>482 [c] |  |
| malonate | 465 [a] |  |
| glycolate | 400-460 [a d] |  |
| thioglycolate | 505 [a] |  |
| nitrilotriacetate | 465 [a] | $N(CH_2COO^-)_3$ |
| EDTA | 515 [a]<br>490 [e] |  |

[a] Fe-anion-STf visible absorption maximum (Schlabach & Bates, 1975). [b] (Fe-$CO_3$)(Fe-$C_2O_4$)-Lf hybrid (Shongwe *et al.*, 1992). [c] Fe-$(C_2O_4)_2$-Lf (Shongwe *et al.*, 1992). [d] Broad shoulder. [e] These $\lambda_{max}$ values correspond to Fe-EDTA-STf and Fe-EDTA-OTf complexes respectively (Rogers, Feeney & Meares, 1977).

Table 3: "Conditional" equilibrium binding constants ($K^*_1$ and $K^*_2$) for the binding of various metal ions with transferrin and lactoferrin. The "absolute" constants for $Fe^{3+}$ and $Ga^{3+}$ are also given at the end of the table.

| Metal ion | protein | $[HCO_3^-]$ mM | pH | Effective $\log K^*_1$ | Effective $\log K^*_2$ | $K^*_1/K^*_2$ ($K^*_R$) | method [a] | ref [b] |
|---|---|---|---|---|---|---|---|---|
| $Fe^{3+}$ | Tf | ambient | 6.7 | 19.48 | 17.36 | 132 | E. D. | 1 |
| $Fe^{3+}$ | Tf | ambient | 7.4 | 20.67 | 19.38 | 19 | E. D. | 1 |
| $Fe^{3+}$ | Tf | 27 | 7.4 | 22.8 | 21.5 | 20 | calc. [c] | 2 |
| $Fe^{3+}$ | Lf | 5 | 7.4 | 22.36 | — | | LFER Ga/Fe | 3 |
| $Ga^{3+}$ | Tf | 5 | 7.4 | 19.53 | 18.58 | 9 | C. C. | 2 |
| $Ga^{3+}$ | Tf | ambient | 7.4 | 18.1 | 17.1 | 10 | calc. [c] | 2 |
| $Ga^{3+}$ | Tf | 27 | 7.4 | 20.3 | 19.3 | 10 | calc. [c] | 2 |
| $Ga^{3+}$ | Lf | 5 | 7.4 | 21.43 | 20.57 | 7 | C. C. | 3 |
| $Al^{3+}$ | Tf | 5 | 7.4 | 13.5 | 12.5 | 10 | C. C. | 4 |
| $Zn^{2+}$ | Tf | ambient | 7.4 | 5.7 | 4.3 | 25 | calc. [c] | 5 |
| $Zn^{2+}$ | Tf | 15 | 7.4 | 7.8 | 6.4 | 25 | C. C. | 5 |
| $Zn^{2+}$ | Tf | 27 | 7.4 | 8.0 | 6.6 | 25 | calc. [c] | 5 |
| $Ni^{2+}$ | Tf | 5 | 7.4 | 4.10 | 3.23 | 7 | C. C. | 6 |
| $Fe^{2+}$ | Tf | 5 | 7.4 | 3.2 | 2.5 | 5 | LFER Ni/Fe | 6 |
| $Cd^{2+}$ | Tf | 5 | 7.4 | 5.95 | 4.86 | 12 | C. C. | 7 |
| $VO_2^+$ | Tf | — | 7.4 | — | 6.5 | | UV | 8 |
| $VO_2^+$ | Tf | 20 | 7.4 | — | 6.36 | | UV | 9 |
| $VO_2^+$ | Tf | — | 7.4 | 7.45 | 6.6 | 7 | C. C. [d] | 10 |
| $Nd^{3+}$ | Tf | ambient | 7.4 | 6.09 | 5.04 | 11 | C. C. | 11 |
| $Sm^{3+}$ | Tf | ambient | 7.4 | 7.13 | 5.39 | 55 | C. C. | 11 |
| $Gd^{3+}$ | Tf | ambient | 7.4 | 6.83 | — | | ED | 12 |
| $Gd^{3+}$ | Tf | 27 | 7.4 | 8.87 | — | | ED | 12 |
| $Gd^{3+}$ | Tf | | | 7.96 | 5.94 | 105 | C. C. | 13 |
| $Am^{3+}$ | Tf | ambient | 7.4 | ≈6.5 | — | | LFER Nd/Am | 11 |
| $Cm^{3+}$ | Tf | ambient | 7.4 | ≈6.5 | — | | LFER Sm/Cm | 11 |
| $Fe^{3+}$ | Tf | — | 7.4 | 21.70 | 20.41 | 19 | calc. [e] | 3 |
| $Ga^{3+}$ | Tf | — | 7.4 | 19.71 | 18.76 | 9 | calc. [e] | 3 |

[a] The methods used to estimate $\log K^*$ values are as follows: E. D. - Equilibrium dialysis against citrate; C. C. - UV difference titrations incorporating various competitive chelators; LFER - linear free energy relationship (described in Chapter 2); UV - UV difference titration and estimation of $\log K^*_2$ from the curvature of the titration curve between r=1 and r=2. [b] References are listed over the page. [c] Calculated from the $\log K^*$ values at ambient $HCO_3^-$ by the method of Harris and Pecoraro (1983). [d] $\log K^*$ values were estimated from competition reactions with inorganic anions. [e] Absolute $\log K'$ values independent of pH and $[HCO_3^-]$ calculated from the conditional $\log K^*$.

1 Aisen, P., Leibman, A. & Zweier, J. (1978) *J. Biol. Chem. 253*, 1930

2 Harris, W. R. & Pecoraro, V. L. (1983) *Biochemistry 22*, 292

3 Harris, W. R. (1986) *Biochemistry 25*, 803

4 Harris, W. R. & Sheldon, J. (1990) *Inorg. Chem. 29*, 119

5 Harris, W. R. (1983) *Biochemistry 22*, 3920

6 Harris, W. R. (1986) *J. Inorg. Biochem. 27*, 41

7 Harris, W. R. & Madsen, L. J. (1988) *Biochemistry 27*, 284

8 Harris, W. R. & Carrano, C. J. (1984) *J. Inorg. Biochem. 22*, 201

9 Chasteen, N. D., Lord, E. M., Thompson, H. J. & Grady, J. K. (1986) *Biochim. Biophys. Acta. 884*, 84

10 Harris, W. R. (1985) *Biochemistry 24*, 7412

11 Harris, W. R. (1986) *Inorg. Chem. 25*, 2041

12 Zak, O. & Aisen, P. (1988) *Biochemistry 27*, 1075

13 Harris, W. R. & Chen, Y. (1991) *5th International Conference on Bioinorganic Chemistry*, Oxford, U.K. Abstract B016

## Appendix II

*X-ray data processing*

1) Calculation of scale factors:

Table 1: $Cu_2Lf$ - crystal 1

| | Standard reflections (Intensities) [a] | | | | | | |
|---|---|---|---|---|---|---|---|
| shell | 21 20 3 | 21 13 0 | 18 12 6 | 0 10 11 | 36 0 9 | sum [b] | S.F. [c] |
| 1 | 11105 | 13260 | 8850 | 6720 | 4810 | 44690 | 1.000 |
| 2 | 7600 | 9670 | 6500 | 4590 | 3640 | 32040 | 1.395 |
| 3 | 4880 | 6510 | 4470 | 2940 | 2430 | 21280 | 2.100 |

[a] Integrated intensities from the first block. [b] The first coefficient in the deterioration correction function. [c] These are scale factors on intensities as calculated from the ratio of sums in the previous column. As these were subsequently applied to F's, the square root of these values was used (Table 3.4).

Table 2: $Cu_2Lf$ - crystal 2

| | Standard reflections (Intensities) [a] | | | | | | |
|---|---|---|---|---|---|---|---|
| shell | 21 20 3 | 21 13 0 | 18 12 6 | 0 10 11 | 36 0 9 | sum [b] | S.F. [c] |
| 4 | 5400 | 6180 | 5110 | 4150 | 2950 | 23790 | 1.000 |
| 5 | 4020 | 4560 | 3810 | 3050 | 2140 | 17640 | 1.349 |
| 6 | 17850 | 21530 | 18100 | 14290 | 10020 | 81660 | 0.291 |
| 7 | 15140 | 19060 | 16090 | 12450 | 9260 | 72180 | 0.330 |

For footnotes, see Table 1.

Table 3: $Cu_2Lf$ - crystal 4

| | Standard reflections (Intensities) [a] | | | | | | |
|---|---|---|---|---|---|---|---|
| shell | 21 20 3 | 21 13 0 | 18 12 6 | 0 10 11 | 36 0 9 | sum [b] | S.F. [c] |
| 9 | 14450 | 15860 | 12230 | 9100 | 6750 | 58510 | 1.000 |
| 10 | 11490 | 14030 | 10710 | 7720 | 5820 | 49850 | 1.174 |

For footnotes, see Table 1.

Table 4: $Cu_2$oxLf - crystal 1

| | Standard reflections (Intensities) [a] | | | | | | |
|---|---|---|---|---|---|---|---|
| shell | 21 20 3 | 21 13 0 | 18 12 6 | 0 10 11 | 36 0 9 | sum [b] | S.F. [c] |
| 1 | 1420 | 1950 | 1420 | 770 | 720 | 6280 | 1.000 |
| 2 | 1890 | 2640 | 1980 | 1080 | 930 | 8520 | 0.737 |
| 3 | 2265 | 3270 | 2425 | 1300 | 1210 | 10470 | 0.600 |
| 4 | 2410 | 3670 | 2690 | 1510 | 1300 | 11580 | 0.541 |

[a] Integrated intensities from the first block. [b] The first coefficient in the deterioration correction function. [c] These are scale factors on intensities as calculated from the ratio of sums in the previous column. As these were subsequently applied to F's, the square root of these values was used (Table 3.5).

Table 5: $Cu_2$oxLf - crystal 2

| | Standard reflections (Intensities) [a] | | | | | | |
|---|---|---|---|---|---|---|---|
| shell | 21 20 3 | 21 13 0 | 18 12 6 | 0 10 11 | 36 0 9 | sum [b] | S.F. [c] |
| 5 | 4720 | 6460 | 4745 | 2430 | 2150 | 20505 | 1.000 |
| 6 | 3660 | 5500 | 4000 | 2115 | 1850 | 17125 | 1.197 |

For footnotes, see Table 4.

Table 6: $Cu_2$oxLf - crystal 3

| | Standard reflections (Intensities) [a] | | | | | | |
|---|---|---|---|---|---|---|---|
| shell | 21 20 3 | 21 13 0 | 18 12 6 | 0 10 11 | 36 0 9 | sum [b] | S.F. [c] |
| 7 | 3950 | 5090 | 4345 | 2790 | 1900 | 18075 | 1.000 |
| 8 | 2890 | 4170 | 3380 | 2250 | 1505 | 14195 | 1.273 |
| 9 | 3350 | 5180 | 4040 | 2670 | 1780 | 17020 | 1.062 |

For footnotes, see Table 4.

Table 7: $Cu_2$oxLf - crystal 4

| | Standard reflections (Intensities) [a] | | | | | | |
|---|---|---|---|---|---|---|---|
| shell | 21 20 3 | 21 13 0 | 18 12 6 | 0 10 11 | 36 0 9 | sum [b] | S.F. [c] |
| 10 | 5770 | 7450 | 7410 | 4640 | 3295 | 28565 | 1.000 |
| 11 | 5410 | 7175 | 7040 | 4320 | 3040 | 26985 | 1.059 |
| 12 | 6130 | 8590 | 8290 | 5240 | 3660 | 31910 | 0.895 |

For footnotes, see Table 4.

2) Oscillation photograph data processing:

Table 8: Oscillation data processing statistics for $Cu_2Lf$ crystal 1

| film | oscillation range (deg) | scale factors [a] B - A | C - A | number of reflections | number with $I > 3\sigma_I$ (%) | $R_{sym}$ [b] | $R_{sca}$ [b] |
|---|---|---|---|---|---|---|---|
| 2 | 0.04 - 2.04 | 2.3279 | 5.3362 | 5747 | 38 | 0.062 | 0.083 |
| 3 | 2.04 - 4.04 | 2.4402 | 5.5906 | 5894 | 37 | 0.060 | 0.093 |
| 4 | 4.04 - 6.04 | 2.3004 | 5.3204 | 5906 | 36 | 0.065 | 0.088 |
| 5 | 6.04 - 8.04 | 2.3059 | 5.0911 | 5899 | 33 | 0.073 | 0.089 |
| 6 | 8.04 - 10.04 | 2.3497 | 5.958 | 5880 | 31 | 0.069 | 0.088 |
| 7 | 10.04 - 12.04 | 2.2303 | 5.1177 | 5952 | 33 | 0.075 | 0.086 |
| 8 | 12.04 - 14.04 | 2.2285 | 5.1308 | 5808 | 25 | 0.058 | 0.092 |
| 21 | 48.00 - 49.50 | 2.3420 | 5.6497 | 4574 | 40 | 0.075 | 0.066 |
| 22 | 49.50 - 51.00 | 2.5588 | 6.4310 | 4321 | 39 | 0.066 | 0.070 |
| 23 | 51.00 - 52.50 | 2.4055 | 5.6899 | 4142 | 43 | 0.068 | 0.063 |
| 24 | 52.50 - 54.00 | 2.4062 | 5.7800 | 4174 | 39 | 0.071 | 0.066 |
| 25 | 54.00 - 55.50 | 2.3429 | 5.6411 | 4176 | 37 | 0.065 | 0.061 |
| 26 | 55.50 - 57.00 | 2.3593 | 5.8515 | 4269 | 38 | 0.057 | 0.063 |
| 27 | 57.00 - 58.50 | 2.4118 | 5.8210 | 4279 | 34 | 0.061 | 0.064 |
| 28 | 58.50 - 60.00 | 2.3671 | 5.7086 | 4283 | 30 | 0.063 | 0.068 |
| 32 | 82.00 - 83.00 | 2.3395 | 5.7548 | 3328 | 31 | 0.085 | 0.070 |
| 33 | 83.00 - 84.00 | 2.5433 | 5.9951 | 3388 | 28 | 0.062 | 0.071 |
| 34 | 84.00 - 85.00 | 2.4919 | 6.1941 | 3328 | 33 | 0.055 | 0.070 |
| 35 | 85.00 - 86.00 | 2.4126 | 5.7606 | 3431 | 33 | 0.045 | 0.071 |
| 36 | 86.00 - 87.00 | 2.5045 | 6.1633 | 3422 | 28 | 0.071 | 0.070 |
| 37 | 87.00 - 88.00 | 2.4319 | 6.1585 | 3403 | 28 | 0.098 | 0.077 |
| 38 | 88.00 - 89.00 | 2.5352 | 5.9009 | 3428 | 26 | 0.087 | 0.073 |
| 41 | 60.00 - 61.50 | 2.4068 | 6.3905 | 4311 | 38 | 0.059 | 0.067 |
| 42 | 61.50 - 63.00 | 2.3058 | 6.2863 | 4319 | 35 | 0.069 | 0.070 |
| 43 | 63.00 - 64.50 | 2.3158 | 6.4079 | 4339 | 38 | 0.065 | 0.067 |
| 44 | 64.50 - 66.00 | 2.8762 | 6.4497 | 4218 | 35 | 0.079 | 0.064 |
| 45 | 66.00 - 67.50 | 2.2820 | 5.3289 | 4360 | 32 | 0.065 | 0.071 |
| 46 | 67.50 - 69.00 | 2.3483 | 5.4805 | 4288 | 29 | 0.065 | 0.069 |
| 47 | 69.00 - 70.50 | 2.4331 | 5.7901 | 4248 | 33 | 0.067 | 0.071 |
| 48 | 70.50 - 72.00 | 2.4209 | 5.7727 | 4121 | 35 | 0.075 | 0.071 |
| 49 | 89.00 - 90.00 | 2.3196 | 5.4093 | 3388 | 39 | 0.058 | 0.089 |
| 51 | 28.00 - 30.00 | 2.3138 | 5.3246 | 5928 | 32 | 0.057 | 0.073 |
| 52 | 30.00 - 32.00 | 2.3037 | 5.5167 | 5937 | 32 | 0.061 | 0.072 |
| 53 | 32.00 - 34.00 | 2.4834 | 4.9671 | 5918 | 34 | 0.059 | 0.069 |
| 54 | 34.00 - 36.00 | 2.3263 | 4.8680 | 5922 | 31 | 0.079 | 0.064 |
| | | | | 160 329 | 34 | 0.067 | 0.073 |

[a] Calculated from reflections measured on all three films of a given filmpack. [b] $R = \Sigma |I - \bar{I}| / \Sigma I$; $R_{sca}$ measures the aggreement between reflections recorded on successive films of a given pack after the scale factors have been applied and $R_{sym}$ measures the agreement between symmetry-related reflections on the same film. Filmpacks marked with an asterisk (*) were omitted from subsequent scaling, merging and postrefinement proceedures due to the poor quality of the data and crystal deterioration.

Table 9: Oscillation data processing statistics for $Cu_2Lf$ crystal 2

| film | oscillation range (deg) | scale factors [a] B - A | C - A | number of reflections | number with $I > 3\sigma_I$ (%) | $R_{sym}$ [b] | $R_{sca}$ [b] |
|---|---|---|---|---|---|---|---|
| 61 | 14.00 - 16.00 | 2.4124 | 6.0555 | 3564 | 38 | 0.069 | 0.066 |
| 62 | 16.00 - 18.00 | 2.5307 | 6.2072 | 3541 | 39 | 0.075 | 0.071 |
| 63 | 18.00 - 20.00 | 2.5227 | 6.0624 | 3449 | 36 | 0.065 | 0.071 |
| 64 | 20.00 - 22.00 | 2.4916 | 6.1329 | 3417 | 36 | 0.076 | 0.075 |
| 65 | 22.00 - 24.00 | 2.4626 | 5.9916 | 3317 | 36 | 0.065 | 0.066 |
| 66 | 24.00 - 26.00 | 2.4622 | 5.9697 | 3355 | 33 | 0.078 | 0.071 |
| 67 | 26.00 - 28.00 | 2.4788 | 6.2263 | 3379 | 32 | 0.065 | 0.069 |
| 68 | 28.00 - 30.00 | 2.5359 | 6.2267 | 3241 | 31 | 0.077 | 0.073 |
| 71 | 72.00 - 73.00 | 2.4062 | 5.6081 | 5862 | 45 | 0.063 | 0.075 |
| 72 | 73.00 - 74.00 | 2.3883 | 5.7804 | 5868 | 38 | 0.082 | 0.083 |
| 74 | 75.00 - 76.00 | 2.3516 | 5.6388 | 5861 | 39 | 0.065 | 0.078 |
| 75 | 76.00 - 77.00 | 2.3973 | 5.8366 | 5904 | 34 | 0.078 | 0.083 |
| 76 | 77.00 - 78.00 | 2.1998 | 5.2174 | 5917 | 32 | 0.057 | 0.077 |
| 77 | 78.00 - 79.00 | 2.2956 | 5.4687 | 5947 | 31 | 0.060 | 0.079 |
| 78 | 79.00 - 80.00 | 2.3565 | 5.5278 | 5942 | 28 | 0.064 | 0.076 |
| | | | | 68 564 | 35 | 0.069 | 0.074 |

[a] Calculated from reflections measured on all three films of a given filmpack. [b] $R = \Sigma |I - \bar{I}| / \Sigma I$; $R_{sca}$ measures the aggreement between reflections recorded on successive films of a given pack after the scale factors have been applied and $R_{sym}$ measures the agreement between symmetry-related reflections on the same film. Filmpacks marked with an asterisk (*) were omitted from subsequent scaling, merging and postrefinement proceedures due to the poor quality of the data and crystal deterioration.

Table 10: Range of oscillation data collected from each $Cu_2Lf$ crystal

| | phi range | | phi range |
|---|---|---|---|
| Crystal 1 | 0.04 - 14.04 | Crystal 2 | 14.00 - 30.00 |
| | 28.00 - 36.00 | | 72.00 - 80.00 |
| | 48.00 - 72.00 | | |
| | 82.00 - 90.00 | | |
| Total number of reflections collected | | 228 893 | |
| Average percentage with $I > 3\sigma_I$ (%) | | 34 | |

Table 11: Oscillation data processing statistics for $Cu_2$oxLf crystal 1

| film | oscillation range (deg) | scale factors [a] B - A | C - A | number of reflections | with $I > 3\sigma_I$ (%) | $R_{sym}$ [b] | $R_{sca}$ [b] |
|---|---|---|---|---|---|---|---|
| 101 | 0.00 - 2.50 | 2.066 | 4.604 | 6217 | 48 | 0.040 | 0.059 |
| 102 | 2.50 - 5.00 | 2.025 | 4.412 | 6714 | 44 | 0.042 | 0.064 |
| 103 | 5.00 - 7.50 | 2.031 | 4.692 | 6785 | 46 | 0.039 | 0.062 |
| 104* | 7.50 - 10.00 | 2.118 | 4.326 | 6517 | 42 | 0.039 | 0.084 |
| 105* | 10.00 - 12.50 | 2.169 | 4.540 | 6867 | 45 | 0.039 | 0.057 |
| 106* | 12.50 - 15.00 | 2.071 | 4.554 | 6671 | 43 | 0.040 | 0.079 |
| 107* | 15.00 - 17.50 | 2.071 | 4.378 | 6814 | 42 | 0.041 | 0.086 |
| 108* | 17.50 - 20.00 | 2.031 | 4.297 | 6797 | 41 | 0.040 | 0.091 |
| 109* | 20.00 - 22.50 | 2.061 | 4.551 | 6761 | 40 | 0.039 | 0.102 |
| 110* | 22.50 - 25.00 | 1.961 | 4.251 | 6788 | 36 | 0.041 | 0.089 |
| 111* | 25.00 - 27.50 | 2.102 | 4.483 | 6745 | 36 | 0.044 | 0.098 |
| 112* | 27.50 - 30.00 | 2.090 | 4.866 | 6963 | 38 | 0.042 | 0.116 |
| 113* | 30.00 - 32.50 | 2.097 | 4.864 | 6881 | 35 | 0.039 | 0.097 |
| 114* | 32.50 - 35.00 | 2.171 | 4.421 | 6949 | 31 | 0.048 | 0.105 |
| 115* | 35.00 - 37.50 | 2.172 | 4.522 | 6862 | 36 | 0.038 | 0.117 |
| 116* | 37.50 - 40.00 | 2.190 | 4.513 | 6884 | 35 | 0.040 | 0.104 |
| 134 | 7.50 - 10.00 | 2.122 | 4.337 | 6538 | 43 | 0.039 | 0.085 |
| 135 | 10.00 - 12.50 | 2.169 | 4.535 | 6901 | 45 | 0.039 | 0.056 |
| 136* | 12.50 - 15.00 | 2.072 | 4.579 | 6689 | 43 | 0.040 | 0.077 |
| 137* | 15.00 - 17.50 | 2.070 | 4.385 | 6831 | 42 | 0.041 | 0.086 |
| 138* | 17.50 - 20.00 | 2.034 | 4.310 | 6817 | 42 | 0.041 | 0.090 |
| total reflections and average R factors [c] | | | | 33,155 | 45 | 0.40 | 0.065 |

[a] Calculated from reflections measured on all three films of a given filmpack. [b] $R = \Sigma |I - \bar{I}| / \Sigma I$; $R_{sca}$ measures the agreement between reflections recorded on successive films of a given pack after the scale factors have been applied and $R_{sym}$ measures the agreement between symmetry-related reflections on the same film. Filmpacks marked with an asterisk (*) were omitted from subsequent scaling, merging and postrefinement proceedures due to the poor quality of the data and crystal deterioration. c Total reflections and average R factors are calculated only for the filmpacks used in subsequent refinement

Table 12: Oscillation data processing statistics for $Cu_2$oxLf crystal 2

| film | oscillation range (deg) | scale factors [a] B - A | C - A | number of reflections | number with $I > 3\sigma_I$ (%) | $R_{sym}$ [b] | $R_{sca}$ [b] |
|---|---|---|---|---|---|---|---|
| 201 | 62.00 - 63.25 | 2.271 | 5.240 | 4017 | 51 | 0.029 | 0.100 |
| 202 | 63.25 - 64.50 | 2.274 | 5.278 | 4023 | 49 | 0.031 | 0.070 |
| 203 | 64.50 - 65.75 | 2.213 | 5.126 | 3986 | 49 | 0.028 | 0.046 |
| 204 | 65.75 - 67.00 | 2.257 | 5.223 | 3919 | 49 | 0.033 | 0.040 |
| 205 | 67.00 - 68.25 | 2.318 | 5.389 | 3915 | 48 | 0.033 | 0.031 |
| 206 | 68.25 - 69.50 | 2.251 | 5.379 | 3893 | 46 | 0.032 | 0.050 |
| 207 | 69.50 - 70.75 | 2.261 | 5.284 | 3794 | 45 | 0.032 | 0.082 |
| 208 | 70.75 - 72.00 | 2.271 | 5.144 | 3824 | 46 | 0.032 | 0.055 |
| 209 | 70.00 - 73.25 | 2.092 | 4.620 | 3759 | 44 | 0.037 | 0.070 |
| 210 | 73.25 - 74.50 | 2.191 | 4.681 | 3587 | 45 | 0.038 | 0.061 |
| 211 | 74.50 - 75.75 | 2.136 | 4.734 | 3625 | 48 | 0.038 | 0.033 |
| 212 | 75.75 - 77.00 | 2.152 | 4.794 | 3691 | 46 | 0.034 | 0.106 |
| 213 | 77.00 - 78.25 | 2.140 | 4.684 | 3618 | 45 | 0.035 | 0.114 |
| 214 | 78.25 - 79.50 | 2.109 | 4.632 | 3688 | 45 | 0.037 | 0.082 |
| 215 | 79.50 - 80.75 | 2.112 | 4.633 | 3626 | 45 | 0.034 | 0.107 |
| 216 | 80.75 - 82.00 | 2.209 | 4.614 | 3691 | 42 | 0.032 | 0.167 |
| 217 | 82.00 - 83.25 | 2.339 | 5.473 | 3701 | 49 | 0.027 | 0.046 |
| 218 | 83.25 - 84.50 | 2.272 | 5.360 | 3575 | 49 | 0.030 | 0.038 |
| 219 | 84.50 - 85.75 | 2.285 | 5.352 | 3629 | 49 | 0.034 | 0.031 |
| 220 | 85.75 - 87.00 | 2.354 | 5.570 | 3687 | 49 | 0.035 | 0.039 |
| 221 | 87.00 - 88.25 | 2.216 | 5.043 | 3749 | 46 | 0.036 | 0.038 |
| 222 | 88.25 - 89.50 | 2.180 | 5.298 | 3777 | 44 | 0.040 | 0.051 |
| 223 | 89.50 - 90.75 | 2.243 | 5.214 | 3748 | 44 | 0.033 | 0.049 |
| 224 | 90.75 - 92.00 | 2.309 | 5.611 | 3703 | 45 | 0.035 | 0.072 |
| Total reflections and average R factors | | | | 90 225 | 47 | 0.034 | 0.066 |

Table 13: Oscillation data processing statistics for $Cu_2$oxLf crystal 3

| film | oscillation range (deg) | scale factors [a] B - A | C - A | number of reflections | number with $I > 3\sigma_I$ (%) | $R_{sym}$ [b] | $R_{sca}$ [b] |
|---|---|---|---|---|---|---|---|
| 301 | 24.00 - 26.50 | 2.209 | 5.009 | 6629 | 50 | 0.029 | 0.058 |
| 302 | 26.50 - 29.00 | 2.219 | 5.102 | 6623 | 50 | 0.029 | 0.051 |
| 303 | 29.00 - 31.50 | 2.157 | 4.658 | 6743 | 50 | 0.031 | 0.057 |
| 304 | 31.50 - 34.00 | 2.146 | 4.777 | 6823 | 46 | 0.030 | 0.061 |
| 305 | 34.00 - 36.50 | 2.123 | 4.785 | 6911 | 47 | 0.031 | 0.065 |
| 306 | 36.50 - 39.00 | 2.158 | 4.837 | 6964 | 44 | 0.030 | 0.066 |
| 307 | 39.00 - 41.50 | 2.114 | 4.854 | 6923 | 42 | 0.032 | 0.089 |
| 308 | 41.50 - 44.00 | 2.159 | 4.875 | 6846 | 43 | 0.031 | 0.078 |
| 309 | 44.00 - 45.50 | 2.245 | 4.866 | 4654 | 43 | 0.028 | 0.092 |
| 310 | 45.50 - 47.00 | 2.223 | 5.263 | 4607 | 44 | 0.029 | 0.057 |
| 311 | 47.00 - 48.50 | 2.205 | 5.169 | 4630 | 41 | 0.031 | 0.057 |
| 312 | 48.50 - 50.00 | 2.224 | 4.992 | 4524 | 44 | 0.031 | 0.081 |
| 313 | 50.00 - 51.50 | 2.175 | 4.958 | 4559 | 44 | 0.030 | 0.037 |
| 314 | 51.50 - 53.00 | 2.152 | 4.993 | 4543 | 43 | 0.031 | 0.069 |
| 315 | 53.00 - 54.50 | 2.277 | 5.010 | 4453 | 41 | 0.031 | 0.090 |
| 316 | 54.50 - 56.00 | 2.268 | 5.020 | 4334 | 42 | 0.029 | 0.048 |
| 317 | 56.00 - 57.50 | 2.245 | 5.059 | 4220 | 46 | 0.029 | 0.085 |
| 318 | 57.50 - 59.00 | 2.127 | 4.774 | 4226 | 46 | 0.029 | 0.052 |
| 319 | 59.00 - 60.50 | 2.112 | 4.702 | 4211 | 47 | 0.031 | 0.048 |
| 320 | 60.50 - 62.00 | 2.117 | 4.789 | 4257 | 45 | 0.030 | 0.037 |
| 321 | 62.00 - 63.50 | 2.151 | 4.764 | 4262 | 45 | 0.030 | 0.052 |
| 322 | 63.50 - 65.00 | 2.133 | 4.811 | 4271 | 44 | 0.030 | 0.084 |
| 323 | 65.00 - 66.50 | 2.202 | 5.088 | 4316 | 42 | 0.029 | 0.073 |
| 324 | 66.50 - 68.00 | 2.200 | 5.065 | 4317 | 40 | 0.032 | 0.062 |
| Total reflections and average R factors | | | | 124 846 | 45 | 0.030 | 0.065 |

Table 14: Range of data collected from each $Cu_2$oxLf crystal

| | phi range | | | | phi range |
|---|---|---|---|---|---|
| Crystal 1 | 0.00 - 12.50 | Crystal 2 | 62.00 - 92.00 | Crystal 3 | 24.00 - 68.00 |
| Total number of reflections collected | | | 248 226 | | |
| Average percentage with $I > 3\sigma_I$ (%) | | | 46 | | |

3) Statistical analysis of the final data sets:

Table 15: Statistics for the final $Cu_2Lf$ data set (from the merging of the diffractometer and photographic data sets) as a function of resolution

| resolution [a] ($4\sin^2\theta/\lambda^2$) | $N_{ref}$ (theoretical) [b] | $N_{ref}$ (measured) | $N_{ref}$ with $I > 1\sigma_I$ | $N_{ref}$ with $I > 2\sigma_I$ | % with $I > 2\sigma_I$ |
|---|---|---|---|---|---|
| 0.02 | 1255 | 1460 | 1458 | 1457 | 99.8 |
| 0.04 | 2294 | 2517 | 2511 | 2499 | 99.3 |
| 0.06 | 2971 | 3162 | 3147 | 3120 | 98.7 |
| 0.08 | 3518 | 3705 | 3681 | 3637 | 98.2 |
| 0.10 | 3837 | 4149 | 4107 | 3974 | 95.8 |
| 0.12 | 4413 | 4523 | 4405 | 4142 | 91.6 |
| 0.14 | 4797 | 4752 | 4566 | 4025 | 84.7 |
| 0.16 | 5154 | 4765 | 4271 | 3189 | 66.9 |
| 0.18 | 5487 | 4742 | 3931 | 2437 | 51.4 |
| 0.20 | 5801 | 4785 | 3680 | 1839 | 38.4 |
| 0.22 | 6098 | 4643 | 3273 | 1202 | 25.9 |
| 0.2403 [c] | 6480 | 3480 | 2117 | 570 | 16.4 |
| | 52259 | 46683 (89%) [d] | 41147 | 32091 | 68.7 [e] |

[a] The resolution is divided into shells in units of $4\sin^2\theta/\lambda^2$, the numbers of reflections given being the totals in each shell. The first shell is from $4\sin^2\theta/\lambda^2 = 0$ to 0.02. [b] The theoretical number of reflections in each shell were calculated using the equation in Blundell and Johnson (1976), pp 248 - 249. It should be noted that for space groups with high symmetry, this value is an underestimate of the true value. [c] The maximum recorded resolution is 2.04 Å. [d] The number of reflections measured to 2.1 Å (the highest resolution used in the structure refinement) was 44715 and the completeness of the data to this resolution is 93% (47903 theoretical). [e] There are 31818 observed data with $I > 2\sigma_I$ (71.1%) to a resolution of 2.1 Å.

Table 16: Statistics for the final $Cu_2$oxLf data set (from the merging of the diffractometer and photographic data sets) as a function of resolution

| resolution [a] ($4\sin^2\theta/\lambda^2$) | $N_{ref}$ (theoretical) | $N_{ref}$ (measured) | $N_{ref}$ with $I > 1\sigma_I$ | $N_{ref}$ with $I > 2\sigma_I$ | % with $I > 2\sigma_I$ |
|---|---|---|---|---|---|
| 0.02 | 1248 | 1407 | 1403 | 1398 | 99.4 |
| 0.04 | 2281 | 2505 | 2498 | 2486 | 99.2 |
| 0.06 | 2954 | 3163 | 3146 | 3129 | 98.9 |
| 0.08 | 3498 | 3703 | 3687 | 3628 | 98.0 |
| 0.10 | 3968 | 4197 | 4151 | 4050 | 96.5 |
| 0.12 | 4388 | 4585 | 4493 | 4241 | 92.5 |
| 0.14 | 4770 | 4803 | 4557 | 4222 | 87.9 |
| 0.16 | 5125 | 4827 | 4440 | 3952 | 81.9 |
| 0.18 | 5456 | 4897 | 4414 | 3711 | 75.8 |
| 0.20 | 5768 | 4914 | 4262 | 3329 | 67.8 |
| 0.22 | 6064 | 4814 | 3952 | 2829 | 58.8 |
| 0.24 | 6346 | 4645 | 3559 | 2188 | 47.1 |
| 0.255 [a] | 4963 | 2564 | 1821 | 981 | 38.3 |
| | 56830 | 51024 (90%) [b] | 46383 | 40144 | 78.7 [c] |

[a] The maximum recorded resolution is 1.98 Å. [b] The number of reflections measured to 2.1 Å (the highest resolution used in the structure refinement) was 45394 and the completeness of the data to this resolution is 95% (47634 theoretical). [c] There are 37788 observed data with $I > 2\sigma_I$ (83.2%) to a resolution of 2.1 Å.

# Appendix III

*Structure refinement statistics*

## 1) $Cu_2Lf$

Table 1: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 1 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ [c] | $B_{tot}$ [c] |
|---|---|---|---|---|---|---|---|---|
| 1 [d] | 1 | 0.360 | 0.019 | 0.013 | 0.190 | 0.226 | 0.046 | — |
| 2 | 1 | 0.342 | 0.021 | 0.013 | 0.191 | 0.225 | 0.036 | — |
| 3 | 4 | 0.366 | 0.034 | 0.017 | 0.226 | 0.231 | 0.032 | — |
| 4 | 4 | 0.350 | 0.044 | 0.021 | 0.259 | 0.238 | 0.025 | — |
| 5 | 4 | 0.339 | 0.049 | 0.023 | 0.284 | 0.243 | 0.021 | — |
| 6 | 2 | 0.332 | 0.040 | 0.021 | 0.266 | 0.241 | 0.030 | — |
| 7 | 2 | 0.330 | 0.036 | 0.019 | 0.258 | 0.240 | 0.023 | — |
| 8 | 2 | 0.317 | 0.032 | 0.018 | 0.251 | 0.239 | 0.019 | — |
| 9 | 1 | 0.324 | 0.025 | 0.016 | 0.242 | 0.236 | 0.025 | — |
| 10 | 1 | 0.326 | 0.022 | 0.015 | 0.238 | 0.236 | 0.020 | — |
| 11 | 1 | 0.326 | 0.020 | 0.015 | 0.230 | 0.235 | 0.017 | — |
| 12 [e] | 4 | 0.324 | 0.032 | 0.017 | 0.240 | 0.247 | 0.021 | — |
| 13 | 4 | 0.315 | 0.040 | 0.019 | 0.264 | 0.251 | 0.015 | — |
| 14 | 4 | 0.309 | 0.044 | 0.021 | 0.280 | 0.255 | 0.013 | — |
| 15 | 4 | 0.305 | 0.052 | 0.024 | 0.309 | 0.258 | 0.014 | — |
| 16 | 4 | 0.300 | 0.057 | 0.026 | 0.330 | 0.262 | 0.012 | — |
| 17 | 4 | 0.296 | 0.060 | 0.027 | 0.345 | 0.265 | 0.010 | — |
| 18 | 2 | 0.293 | 0.056 | 0.027 | 0.335 | 0.266 | 0.009 | — |
| 19 | 2 | 0.293 | 0.053 | 0.027 | 0.331 | 0.268 | 0.008 | — |
| 20 | 2 | 0.292 | 0.052 | 0.027 | 0.329 | 0.269 | 0.008 | — |
| 21 | 1 | 0.291 | 0.041 | 0.024 | 0.296 | 0.267 | 0.014 | — |
| 22 | 1 | 0.294 | 0.036 | 0.022 | 0.281 | 0.266 | 0.009 | — |
| 23 | 1 | 0.296 | 0.033 | 0.021 | 0.272 | 0.265 | 0.007 | — |
| 24 | 1 | 0.296 | 0.024 | 0.017 | 0.246 | 0.260 | 0.015 | — |
| 25 | 1 | 0.300 | 0.021 | 0.015 | 0.232 | 0.257 | 0.010 | — |
| 26 | 1 | 0.301 | 0.020 | 0.015 | 0.223 | 0.254 | 0.008 | — |

[a] The coordinates of diferric lactoferrin with the iron atoms, ligands, carbonate anions and water molecules removed, was used as the initial model. [b] Unit weights applied to distance, planar groups, torsion angles, chiral groups, van der Waals contacts and temperature factors. [c] $D_{tot}$ and $B_{tot}$ are the rms shifts in the positional and thermal parameters respectively. [d] Photographic data used for cycles 1 - 11.
[e] Photographic and diffractometer data scaled together for all subsequent cycles.

Table 2: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 2 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.403 | 0.017 | 0.013 | 0.233 | 0.254 | 0.031 | — |
| 2 | 1 | 0.388 | 0.019 | 0.012 | 0.253 | 0.257 | 0.023 | — |
| 3 | 1 | 0.378 | 0.020 | 0.012 | 0.268 | 0.260 | 0.020 | — |
| 4 | 4 | 0.370 | 0.035 | 0.016 | 0.302 | 0.264 | 0.028 | — |
| 5 | 4 | 0.356 | 0.044 | 0.010 | 0.332 | 0.268 | 0.022 | — |
| 6 | 4 | 0.346 | 0.049 | 0.022 | 0.351 | 0.271 | 0.019 | — |
| 7 | 2 | 0.338 | 0.040 | 0.021 | 0.331 | 0.271 | 0.017 | — |
| 8 | 2 | 0.337 | 0.036 | 0.020 | 0.322 | 0.271 | 0.014 | — |
| 9 | 2 | 0.334 | 0.035 | 0.020 | 0.319 | 0.273 | 0.013 | — |
| 10 | 1 | 0.331 | 0.025 | 0.016 | 0.286 | 0.270 | 0.016 | — |
| 11 | 1 | 0.333 | 0.021 | 0.015 | 0.271 | 0.268 | 0.012 | — |
| 12 | 1 | 0.333 | 0.020 | 0.014 | 0.263 | 0.268 | 0.011 | — |
| 13 | 1 | 0.325 | 0.019 | 0.014 | 0.238 | 0.263 | 0.015 | — |
| 14 | 1 | 0.321 | 0.019 | 0.013 | 0.225 | 0.260 | 0.012 | — |
| 15 | 1 | 0.318 | 0.019 | 0.013 | 0.217 | 0.257 | 0.011 | — |
| 16 | 4 | 0.316 | 0.032 | 0.016 | 0.250 | 0.262 | 0.019 | — |
| 17 | 4 | 0.307 | 0.040 | 0.019 | 0.284 | 0.268 | 0.015 | — |
| 18 | 4 | 0.301 | 0.044 | 0.021 | 0.313 | 0.272 | 0.012 | — |
| 19 | 2 | 0.297 | 0.036 | 0.020 | 0.306 | 0.273 | 0.011 | — |
| 20 | 2 | 0.297 | 0.033 | 0.019 | 0.304 | 0.274 | 0.009 | — |
| 21 | 2 | 0.296 | 0.032 | 0.019 | 0.305 | 0.275 | 0.008 | — |
| 22 | 1 | 0.295 | 0.023 | 0.016 | 0.278 | 0.273 | 0.013 | — |
| 23 | 1 | 0.293 | 0.020 | 0.015 | 0.266 | 0.272 | 0.009 | 1.13 |
| 24 | 1 | 0.289 | 0.019 | 0.014 | 0.259 | 0.271 | 0.008 | 1.10 |
| 25 | 1 | 0.285 | 0.019 | 0.013 | 0.237 | 0.266 | 0.007 | 1.07 |
| 26 | 1 | 0.281 | 0.018 | 0.013 | 0.225 | 0.263 | 0.007 | 1.02 |
| 27 | 1 | 0.276 | 0.018 | 0.013 | 0.217 | 0.260 | 0.007 | 0.97 |

[a] Copper atoms and protein ligands included in refinement. Anions omitted. [b] Unit weights applied to all parameters.

Table 3: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 3 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 4 | 0.303 | 0.032 | 0.020 | 0.191 | 0.223 | 0.020 | — |
| 2 | 4 | 0.293 | 0.038 | 0.021 | 0.196 | 0.224 | 0.014 | — |
| 3 | 4 | 0.288 | 0.042 | 0.022 | 0.207 | 0.226 | 0.012 | — |
| 4 | 8 | 0.284 | 0.052 | 0.025 | 0.229 | 0.229 | 0.015 | — |
| 5 | 8 | 0.278 | 0.059 | 0.027 | 0.245 | 0.231 | 0.012 | — |
| 6 | 8 | 0.274 | 0.064 | 0.019 | 0.269 | 0.233 | 0.010 | — |
| 7 | 4 | 0.270 | 0.056 | 0.029 | 0.244 | 0.234 | 0.010 | — |
| 8 | 4 | 0.271 | 0.053 | 0.028 | 0.235 | 0.235 | 0.008 | — |
| 9 | 4 | 0.271 | 0.051 | 0.028 | 0.231 | 0.237 | 0.007 | — |
| 10 | 2 | 0.270 | 0.039 | 0.024 | 0.202 | 0.235 | 0.013 | — |
| 11 | 2 | 0.274 | 0.034 | 0.022 | 0.190 | 0.234 | 0.008 | — |
| 12 | 2 | 0.275 | 0.032 | 0.021 | 0.183 | 0.233 | 0.006 | — |
| 13 | 1 | 0.275 | 0.023 | 0.018 | 0.176 | 0.234 | 0.014 | — |
| 14 | 1 | 0.279 | 0.021 | 0.016 | 0.173 | 0.234 | 0.009 | — |
| 15 | 1 | 0.281 | 0.019 | 0.015 | 0.171 | 0.234 | 0.007 | — |
| 16 | 8 | 0.282 | 0.036 | 0.018 | 0.205 | 0.240 | 0.019 | — |
| 17 | 8 | 0.274 | 0.047 | 0.022 | 0.232 | 0.245 | 0.014 | — |
| 18 | 8 | 0.269 | 0.055 | 0.025 | 0.251 | 0.249 | 0.011 | — |
| 19 | 4 | 0.265 | 0.050 | 0.025 | 0.234 | 0.248 | 0.007 | — |
| 20 | 4 | 0.266 | 0.048 | 0.025 | 0.228 | 0.248 | 0.005 | — |
| 21 | 4 | 0.265 | 0.047 | 0.025 | 0.226 | 0.250 | 0.005 | — |
| 22 | 2 | 0.265 | 0.037 | 0.023 | 0.204 | 0.246 | 0.011 | — |
| 23 | 2 | 0.268 | 0.033 | 0.021 | 0.194 | 0.244 | 0.006 | — |
| 24 | 2 | 0.269 | 0.031 | 0.020 | 0.188 | 0.242 | 0.005 | — |
| 25 | 1 | 0.270 | 0.023 | 0.017 | 0.185 | 0.242 | 0.013 | — |
| 26 | 1 | 0.274 | 0.020 | 0.016 | 0.182 | 0.241 | 0.008 | — |
| 27 | 1 | 0.276 | 0.019 | 0.015 | 0.180 | 0.240 | 0.006 | — |
| 28 | 1 | 0.274 | 0.018 | 0.013 | 0.179 | 0.240 | 0.007 | 0.94 |
| 29 | 1 | 0.270 | 0.017 | 0.013 | 0.178 | 0.240 | 0.006 | 0.91 |
| 30 | 4 | 0.267 | 0.019 | 0.016 | 0.202 | 0.245 | 0.015 | — |
| 31 | 4 | 0.261 | 0.035 | 0.018 | 0.224 | 0.249 | 0.011 | — |
| 32 | 4 | 0.257 | 0.039 | 0.020 | 0.240 | 0.252 | 0.009 | — |
| 33 | 2 | 0.254 | 0.032 | 0.019 | 0.227 | 0.252 | 0.009 | — |
| 34 | 2 | 0.255 | 0.019 | 0.018 | 0.223 | 0.252 | 0.006 | — |
| 35 | 2 | 0.255 | 0.018 | 0.018 | 0.222 | 0.252 | 0.006 | — |
| 36 | 1 | 0.255 | 0.021 | 0.016 | 0.205 | 0.248 | 0.011 | — |
| 37 | 1 | 0.258 | 0.019 | 0.014 | 0.197 | 0.246 | 0.007 | — |
| 38 | 1 | 0.259 | 0.018 | 0.014 | 0.191 | 0.245 | 0.006 | — |
| 39 | 1 | 0.259 | 0.017 | 0.013 | 0.188 | 0.244 | 0.005 | 0.85 |
| 40 | 1 | 0.257 | 0.017 | 0.013 | 0.185 | 0.243 | 0.005 | 0.82 |
| 41 | 1 | 0.254 | 0.017 | 0.013 | 0.183 | 0.242 | 0.005 | 0.80 |

[a] Carbonate anions present in the refinement cycles. [b] Unit weights applied to all parameters.

Table 4: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 4 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.273 | 0.017 | 0.016 | 0.170 | 0.236 | 0.019 | 0.93 |
| 2 | 1 | 0.267 | 0.016 | 0.015 | 0.171 | 0.235 | 0.012 | 0.87 |
| 3 | 1 | 0.263 | 0.016 | 0.015 | 0.171 | 0.236 | 0.009 | 0.84 |
| 4 | 4 | 0.260 | 0.027 | 0.016 | 0.193 | 0.240 | 0.016 | — |
| 5 | 4 | 0.254 | 0.033 | 0.018 | 0.212 | 0.143 | 0.011 | — |
| 6 | 4 | 0.250 | 0.036 | 0.019 | 0.226 | 0.246 | 0.009 | — |
| 7 | 2 | 0.247 | 0.029 | 0.018 | 0.214 | 0.246 | 0.009 | 0.81 |
| 8 | 2 | 0.246 | 0.027 | 0.018 | 0.211 | 0.247 | 0.007 | 0.81 |
| 9 | 2 | 0.244 | 0.026 | 0.018 | 0.211 | 0.248 | 0.006 | 0.79 |
| 10 | 1 | 0.242 | 0.019 | 0.015 | 0.196 | 0.245 | 0.011 | 0.64 |
| 11 | 1 | 0.244 | 0.017 | 0.014 | 0.189 | 0.243 | 0.007 | 0.67 |
| 12 | 1 | 0.243 | 0.016 | 0.013 | 0.185 | 0.241 | 0.006 | 0.67 |
| 13 | 1 | 0.243 | 0.016 | 0.013 | 0.182 | 0.240 | 0.006 | 0.67 |
| 14 | 1 | 0.242 | 0.016 | 0.013 | 0.179 | 0.239 | 0.005 | 0.67 |
| 15 | 1 | 0.241 | 0.016 | 0.012 | 0.177 | 0.238 | 0.005 | 0.67 |

[a] Twenty-two water molecules added. [b] Unit weights applied to all parameters.

Table 5: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 5 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.253 | 0.016 | 0.013 | 0.168 | 0.232 | 0.014 | 0.83 |
| 2 | 1 | 0.248 | 0.016 | 0.012 | 0.169 | 0.232 | 0.009 | 0.77 |
| 3 | 1 | 0.245 | 0.015 | 0.012 | 0.169 | 0.232 | 0.008 | 0.73 |
| 4 | 1 | 0.243 | 0.015 | 0.012 | 0.169 | 0.232 | 0.007 | 0.71 |
| 5 | 1 | 0.241 | 0.015 | 0.012 | 0.168 | 0.233 | 0.007 | 0.68 |
| 6 | 1 | 0.239 | 0.015 | 0.012 | 0.168 | 0.233 | 0.006 | 0.67 |
| 7 | 4 | 0.238 | 0.025 | 0.015 | 0.190 | 0.238 | 0.015 | — |
| 8 | 4 | 0.232 | 0.031 | 0.017 | 0.207 | 0.242 | 0.011 | — |
| 9 | 4 | 0.229 | 0.033 | 0.019 | 0.220 | 0.245 | 0.009 | — |
| 10 | 2 | 0.227 | 0.027 | 0.018 | 0.211 | 0.245 | 0.008 | 0.68 |
| 11 | 2 | 0.226 | 0.025 | 0.017 | 0.208 | 0.245 | 0.006 | 0.68 |
| 12 | 2 | 0.225 | 0.024 | 0.017 | 0.208 | 0.246 | 0.005 | 0.66 |
| 13 | 1 | 0.224 | 0.018 | 0.015 | 0.195 | 0.242 | 0.010 | 0.52 |
| 14 | 1 | 0.226 | 0.017 | 0.014 | 0.188 | 0.241 | 0.007 | 0.54 |
| 15 | 1 | 0.227 | 0.016 | 0.013 | 0.184 | 0.239 | 0.006 | 0.55 |
| 16 | 1 | 0.227 | 0.016 | 0.013 | 0.181 | 0.238 | 0.005 | 0.54 |
| 17 | 1 | 0.227 | 0.016 | 0.012 | 0.179 | 0.237 | 0.005 | 0.54 |
| 18 | 1 | 0.226 | 0.015 | 0.012 | 0.177 | 0.237 | 0.005 | 0.54 |
| 19 | 1 | 0.226 | 0.015 | 0.012 | 0.176 | 0.237 | 0.004 | 0.53 |
| 20 | 1 | 0.226 | 0.015 | 0.012 | 0.175 | 0.237 | 0.004 | 0.53 |

[a] Thirty-six new water molecules; none deleted; 58 in total. [b] Unit weights applied to all parameters.

Table 6: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 6 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.235 | 0.015 | 0.012 | 0.170 | 0.233 | 0.012 | — |
| 2 | 1 | 0.233 | 0.015 | 0.012 | 0.169 | 0.233 | 0.009 | — |
| 3 | 1 | 0.232 | 0.015 | 0.012 | 0.169 | 0.233 | 0.007 | — |
| 4 | 1 | 0.231 | 0.015 | 0.012 | 0.168 | 0.233 | 0.007 | 0.65 |
| 5 | 1 | 0.229 | 0.015 | 0.012 | 0.168 | 0.233 | 0.006 | 0.62 |
| 6 | 1 | 0.228 | 0.015 | 0.012 | 0.168 | 0.233 | 0.006 | 0.59 |
| 7 | 4 | 0.226 | 0.025 | 0.014 | 0.190 | 0.238 | 0.015 | — |
| 8 | 4 | 0.221 | 0.030 | 0.017 | 0.208 | 0.241 | 0.010 | — |
| 9 | 4 | 0.218 | 0.033 | 0.018 | 0.222 | 0.244 | 0.008 | — |
| 10 | 2 | 0.216 | 0.027 | 0.017 | 0.212 | 0.244 | 0.008 | 0.61 |
| 11 | 2 | 0.216 | 0.024 | 0.017 | 0.210 | 0.244 | 0.005 | 0.60 |
| 12 | 2 | 0.216 | 0.024 | 0.017 | 0.209 | 0.244 | 0.005 | 0.57 |
| 13 | 1 | 0.215 | 0.018 | 0.015 | 0.195 | 0.241 | 0.010 | 0.45 |
| 14 | 1 | 0.217 | 0.017 | 0.014 | 0.189 | 0.240 | 0.006 | 0.46 |
| 15 | 1 | 0.218 | 0.016 | 0.013 | 0.185 | 0.238 | 0.005 | 0.47 |

[a] Twenty-three new water molecules; 2 deleted; 79 in total. [b] Unit weights applied to all parammeters.

Table 7: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 7 [a]

| Cycle | x-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.235 | 0.018 | 0.021 | 0.226 | 0.236 | 0.015 | — |
| 2 | 1 | 0.230 | 0.019 | 0.024 | 0.241 | 0.237 | 0.011 | — |
| 3 | 1 | 0.228 | 0.018 | 0.026 | 0.249 | 0.237 | 0.009 | — |
| 4 | 1 | 0.226 | 0.023 | 0.023 | 0.237 | 0.239 | 0.010 | 0.87 |
| 5 | 1 | 0.221 | 0.018 | 0.018 | 0.212 | 0.238 | 0.010 | 0.60 |
| 6 | 1 | 0.221 | 0.016 | 0.015 | 0.200 | 0.237 | 0.007 | 0.56 |
| 7 | 4 | 0.221 | 0.024 | 0.020 | 0.236 | 0.241 | 0.014 | — |
| 8 | 4 | 0.216 | 0.028 | 0.025 | 0.263 | 0.244 | 0.010 | — |
| 9 | 4 | 0.214 | 0.019 | 0.028 | 0.283 | 0.247 | 0.008 | — |
| 10 | 2 | 0.212 | 0.023 | 0.030 | 0.280 | 0.246 | 0.008 | — |
| 11 | 2 | 0.213 | 0.021 | 0.031 | 0.282 | 0.246 | 0.005 | — |
| 12 | 2 | 0.213 | 0.020 | 0.032 | 0.285 | 0.246 | 0.005 | — |
| 13 | 1 | 0.213 | 0.015 | 0.030 | 0.272 | 0.242 | 0.011 | — |
| 14 | 1 | 0.216 | 0.014 | 0.030 | 0.266 | 0.239 | 0.007 | — |
| 15 | 1 | 0.217 | 0.014 | 0.029 | 0.263 | 0.237 | 0.005 | — |
| 16 | 1 | 0.217 | 0.015 | 0.020 | 0.228 | 0.238 | 0.011 | 0.53 |
| 17 | 1 | 0.217 | 0.015 | 0.017 | 0.213 | 0.237 | 0.006 | 0.50 |
| 18 | 1 | 0.216 | 0.015 | 0.015 | 0.204 | 0.237 | 0.004 | 0.49 |
| 19 | 4 | 0.217 | 0.023 | 0.020 | 0.238 | 0.241 | 0.015 | — |
| 20 | 4 | 0.212 | 0.026 | 0.024 | 0.264 | 0.244 | 0.010 | — |
| 21 | 4 | 0.209 | 0.028 | 0.028 | 0.283 | 0.247 | 0.008 | — |
| 22 | 2 | 0.207 | 0.022 | 0.029 | 0.282 | 0.246 | 0.008 | — |
| 23 | 2 | 0.208 | 0.020 | 0.031 | 0.284 | 0.246 | 0.005 | — |
| 24 | 2 | 0.208 | 0.020 | 0.032 | 0.287 | 0.246 | 0.004 | — |
| 25 | 1 | 0.208 | 0.017 | 0.021 | 0.241 | 0.243 | 0.013 | 0.44 |
| 26 | 1 | 0.210 | 0.016 | 0.017 | 0.222 | 0.241 | 0.007 | 0.44 |
| 27 | 1 | 0.211 | 0.016 | 0.015 | 0.211 | 0.240 | 0.005 | 0.44 |

[a] Thirty-four new water molecules; 113 in total. [b] Unit weights applied to all parameters.

Table 8: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 8 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 [c] | 0.223 | 0.014 | 0.020 | 0.228 | 0.238 | 0.013 | — |
| 2 | 1 | 0.222 | 0.013 | 0.023 | 0.237 | 0.236 | 0.009 | — |
| 3 | 1 | 0.220 | 0.013 | 0.024 | 0.242 | 0.235 | 0.007 | — |
| 4 | 1 [d] | 0.220 | 0.019 | 0.011 | 0.248 | 0.238 | 0.016 | 0.71 |
| 5 | 1 [d] | 0.216 | 0.020 | 0.008 | 0.246 | 0.240 | 0.007 | 0.62 |
| 6 | 1 [d] | 0.213 | 0.021 | 0.007 | 0.243 | 0.241 | 0.005 | 0.57 |
| 7 | 4 | 0.211 | 0.027 | 0.012 | 0.247 | 0.243 | 0.012 | — |
| 8 | 4 | 0.207 | 0.031 | 0.015 | 0.253 | 0.246 | 0.008 | — |
| 9 | 4 | 0.205 | 0.032 | 0.017 | 0.258 | 0.248 | 0.007 | — |
| 10 | 2 [e] | 0.204 | 0.023 | 0.023 | 0.223 | 0.246 | 0.011 | — |
| 11 | 2 [e] | 0.205 | 0.021 | 0.026 | 0.208 | 0.245 | 0.006 | — |
| 12 | 2 [e] | 0.206 | 0.020 | 0.028 | 0.200 | 0.245 | 0.005 | — |
| 13 | 1 [d] | 0.206 | 0.021 | 0.012 | 0.236 | 0.246 | 0.016 | 0.49 |
| 14 | 1 [d] | 0.205 | 0.022 | 0.008 | 0.244 | 0.246 | 0.005 | 0.47 |
| 15 | 1 [d] | 0.205 | 0.021 | 0.007 | 0.244 | 0.246 | 0.003 | 0.45 |
| 16 | 1 [d] | 0.204 | 0.015 | 0.016 | 0.250 | 0.242 | 0.014 | 0.58 |
| 17 | 1 [e] | 0.205 | 0.016 | 0.017 | 0.260 | 0.240 | 0.006 | 0.61 |
| 18 | 1 [f] | 0.205 | 0.016 | 0.012 | 0.234 | 0.241 | 0.006 | 0.51 |
| 19 | 1 | 0.204 | 0.015 | 0.013 | 0.213 | 0.239 | 0.005 | 0.49 |

[a] Twenty-two new water molecules 135 in total. [b] Unit weights applied to all parameters unless otherwise stated. [c] A weight of 2.0 applied to distance. [d] A weight of 2.0 applied to planes. [e] Weights of 2.0 and 1.5 for distances and planes. [f] Weights of 2.0, 1.5 and 1.5 for planes, distance and chiral volumes.

Table 9: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 9 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.233 | 0.016 | 0.012 | 0.190 | 0.237 | 0.015 | — |
| 2 | 1 | 0.229 | 0.016 | 0.012 | 0.185 | 0.236 | 0.011 | — |
| 3 | 1 | 0.226 | 0.016 | 0.012 | 0.186 | 0.236 | 0.009 | — |
| 4 | 1 | 0.224 | 0.015 | 0.012 | 0.180 | 0.236 | 0.007 | 0.73 |
| 5 | 1 | 0.221 | 0.015 | 0.012 | 0.178 | 0.236 | 0.007 | 0.65 |
| 6 | 4 | 0.218 | 0.025 | 0.015 | 0.198 | 0.240 | 0.015 | — |
| 7 | 4 | 0.213 | 0.030 | 0.018 | 0.215 | 0.243 | 0.010 | — |
| 8 | 4 | 0.210 | 0.032 | 0.019 | 0.227 | 0.246 | 0.008 | — |
| 9 | 2 | 0.207 | 0.026 | 0.018 | 0.217 | 0.245 | 0.008 | — |
| 10 | 2 | 0.208 | 0.024 | 0.018 | 0.215 | 0.246 | 0.005 | — |
| 11 | 2 | 0.208 | 0.023 | 0.018 | 0.214 | 0.240 | 0.004 | — |
| 12 | 1 | 0.208 | 0.018 | 0.015 | 0.201 | 0.242 | 0.010 | 0.54 |
| 13 | 1 | 0.209 | 0.016 | 0.014 | 0.194 | 0.241 | 0.006 | 0.51 |
| 14 | 1 | 0.209 | 0.016 | 0.013 | 0.190 | 0.239 | 0.005 | 0.49 |
| 15 | 1 | 0.208 | 0.016 | 0.013 | 0.187 | 0.238 | 0.004 | — |
| 16 | 1 | 0.208 | 0.015 | 0.013 | 0.185 | 0.237 | 0.004 | — |
| 17 | 1 | 0.208 | 0.015 | 0.013 | 0.184 | 0.237 | 0.003 | — |

[a] Thirty-seven new water molecules; 4 deleted; 168 in total. [b] Unit weights applied to all parameters.

Table 10: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 10 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.228 | 0.016 | 0.013 | 0.178 | 0.232 | 0.016 | 0.98 |
| 2 | 1 | 0.221 | 0.016 | 0.013 | 0.178 | 0.233 | 0.011 | 0.73 |
| 3 | 1 | 0.217 | 0.016 | 0.013 | 0.178 | 0.233 | 0.009 | 0.64 |
| 4 | 1 | 0.214 | 0.016 | 0.013 | 0.178 | 0.234 | 0.007 | — |
| 5 | 1 | 0.213 | 0.016 | 0.013 | 0.178 | 0.234 | 0.006 | — |
| 6 | 1 | 0.212 | 0.016 | 0.013 | 0.179 | 0.234 | 0.006 | — |
| 7 | 1 | 0.211 | 0.018 | 0.014 | 0.189 | 0.232 | 0.007 | 0.58 |
| 8 | 1 | 0.210 | 0.018 | 0.015 | 0.193 | 0.231 | 0.005 | 0.55 |
| 9 | 4 | 0.209 | 0.028 | 0.018 | 0.214 | 0.235 | 0.015 | — |
| 10 | 4 | 0.204 | 0.033 | 0.020 | 0.232 | 0.239 | 0.010 | — |
| 11 | 4 | 0.201 | 0.036 | 0.022 | 0.246 | 0.241 | 0.007 | — |
| 12 | 4 | 0.199 | 0.037 | 0.023 | 0.250 | 0.245 | 0.008 | — |
| 13 | 2 | 0.198 | 0.029 | 0.021 | 0.233 | 0.245 | 0.008 | — |
| 14 | 2 | 0.199 | 0.025 | 0.023 | 0.244 | 0.245 | 0.006 | — |
| 15 | 2 | 0.200 | 0.025 | 0.021 | 0.134 | 0.246 | 0.004 | — |
| 16 | 2 | 0.199 | 0.025 | 0.020 | 0.230 | 0.247 | 0.004 | — |
| 17 | 1 | 0.199 | 0.019 | 0.017 | 0.213 | 0.244 | 0.010 | — |
| 18 | 1 | 0.202 | 0.017 | 0.015 | 0.205 | 0.242 | 0.005 | — |
| 19 | 1 | 0.203 | 0.017 | 0.014 | 0.200 | 0.241 | 0.004 | — |
| 20 | 1 | 0.204 | 0.017 | 0.014 | 0.196 | 0.240 | 0.003 | 0.48 |
| 21 | 1 | 0.203 | 0.016 | 0.014 | 0.194 | 0.240 | 0.003 | 0.46 |
| 22 | 1 | 0.202 | 0.016 | 0.013 | 0.192 | 0.239 | 0.003 | 0.44 |
| 23 | 4 | 0.202 | 0.027 | 0.017 | 0.211 | 0.244 | 0.014 | — |
| 24 | 4 | 0.198 | 0.032 | 0.019 | 0.228 | 0.247 | 0.009 | — |
| 25 | 4 | 0.196 | 0.034 | 0.021 | 0.240 | 0.250 | 0.007 | — |
| 26 | 2 | 0.194 | 0.028 | 0.020 | 0.231 | 0.249 | 0.007 | — |
| 27 | 2 | 0.195 | 0.026 | 0.019 | 0.228 | 0.249 | 0.003 | — |
| 28 | 2 | 0.196 | 0.025 | 0.019 | 0.228 | 0.250 | 0.003 | — |
| 29 | 1 | 0.196 | 0.017 | 0.020 | 0.235 | 0.245 | 0.012 | — |
| 30 | 1 | 0.199 | 0.016 | 0.021 | 0.237 | 0.243 | 0.006 | — |
| 31 | 1 | 0.201 | 0.016 | 0.016 | 0.215 | 0.242 | 0.008 | 0.47 |
| 32 | 1 | 0.201 | 0.016 | 0.015 | 0.206 | 0.241 | 0.005 | 0.42 |
| 33 | 1 | 0.201 | 0.016 | 0.014 | 0.200 | 0.241 | 0.004 | 0.40 |
| 34 | 1 | 0.200 | 0.016 | 0.014 | 0.197 | 0.240 | 0.008 | — |
| 35 | 1 | 0.200 | 0.016 | 0.013 | 0.194 | 0.239 | 0.003 | — |
| 36 | 1 | 0.201 | 0.016 | 0.013 | 0.193 | 0.239 | 0.003 | — |

[a] Ten new water molecules; 6 deleted; 172 in total. [b] Unit weights applied to all parameters.

Table 11: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 11 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.215 | 0.016 | 0.013 | 0.184 | 0.237 | 0.017 | 0.80 |
| 2 | 1 | 0.210 | 0.016 | 0.013 | 0.184 | 0.237 | 0.011 | 0.67 |
| 3 | 1 | 0.207 | 0.016 | 0.013 | 0.183 | 0.236 | 0.009 | 0.59 |
| 4 | 1 | 0.205 | 0.016 | 0.013 | 0.183 | 0.237 | 0.007 | — |
| 5 | 1 | 0.204 | 0.016 | 0.013 | 0.183 | 0.237 | 0.006 | — |
| 6 | 1 | 0.204 | 0.016 | 0.013 | 0.182 | 0.237 | 0.006 | — |
| 7 | 4 | 0.203 | 0.026 | 0.016 | 0.204 | 0.241 | 0.015 | — |
| 8 | 4 | 0.198 | 0.031 | 0.018 | 0.222 | 0.245 | 0.010 | — |
| 9 | 4 | 0.195 | 0.033 | 0.020 | 0.236 | 0.248 | 0.008 | — |
| 10 | 2 | 0.194 | 0.027 | 0.019 | 0.226 | 0.248 | 0.008 | — |
| 11 | 2 | 0.195 | 0.025 | 0.018 | 0.224 | 0.248 | 0.005 | — |
| 12 | 2 | 0.195 | 0.024 | 0.018 | 0.223 | 0.249 | 0.004 | — |
| 13 | 1 | 0.195 | 0.019 | 0.016 | 0.209 | 0.245 | 0.010 | — |
| 14 | 1 | 0.198 | 0.017 | 0.014 | 0.202 | 0.244 | 0.006 | — |
| 15 | 1 | 0.199 | 0.016 | 0.014 | 0.197 | 0.242 | 0.004 | — |
| 16 | 1 | 0.199 | 0.016 | 0.014 | 0.194 | 0.241 | 0.004 | 0.48 |
| 17 | 1 | 0.199 | 0.016 | 0.013 | 0.192 | 0.240 | 0.003 | 0.45 |
| 18 | 1 | 0.199 | 0.016 | 0.013 | 0.191 | 0.240 | 0.003 | 0.42 |
| 19 | 4 | 0.194 | 0.031 | 0.019 | 0.229 | 0.248 | 0.009 | — |
| 20 | 4 | 0.192 | 0.033 | 0.020 | 0.242 | 0.251 | 0.007 | — |
| 21 | 2 | 0.190 | 0.027 | 0.019 | 0.232 | 0.250 | 0.007 | — |
| 22 | 2 | 0.192 | 0.025 | 0.019 | 0.230 | 0.251 | 0.003 | — |
| 23 | 2 | 0.192 | 0.024 | 0.018 | 0.229 | 0.251 | 0.003 | — |
| 24 | 1 | 0.192 | 0.019 | 0.016 | 0.214 | 0.247 | 0.009 | — |
| 25 | 1 | 0.195 | 0.017 | 0.015 | 0.207 | 0.245 | 0.005 | — |
| 26 | 1 | 0.196 | 0.017 | 0.014 | 0.202 | 0.244 | 0.004 | — |
| 27 | 1 | 0.197 | 0.016 | 0.014 | 0.198 | 0.243 | 0.003 | 0.40 |
| 28 | 1 | 0.197 | 0.016 | 0.013 | 0.196 | 0.242 | 0.003 | 0.38 |
| 29 | 1 | 0.197 | 0.016 | 0.013 | 0.194 | 0.241 | 0.003 | 0.37 |

[a] Twelve water molecules added; 4 deleted; 180 in total. [b] Unit weights applied to all paramaters.

Table 12: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 12 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 [c] | 1 | 0.241 | 0.025 | 0.015 | 0.205 | 0.232 | 0.028 | — |
| 2 | 1 | 0.240 | 0.024 | 0.015 | 0.203 | 0.232 | 0.025 | — |
| 3 | 1 | 0.239 | 0.024 | 0.014 | 0.201 | 0.232 | 0.023 | — |
| 4 | 1 | 0.236 | 0.023 | 0.014 | 0.198 | 0.232 | 0.019 | — |
| 5 | 1 | 0.234 | 0.022 | 0.014 | 0.196 | 0.232 | 0.016 | — |
| 6 | 1 | 0.233 | 0.022 | 0.014 | 0.196 | 0.232 | 0.014 | — |
| 7 | 1 | 0.232 | 0.021 | 0.014 | 0.196 | 0.232 | 0.013 | — |
| 8 | 1 | 0.231 | 0.021 | 0.014 | 0.196 | 0.232 | 0.012 | — |
| 9 | 1 [d] | 0.230 | 0.018 | 0.020 | 0.234 | 0.231 | 0.014 | — |
| 10 | 1 [d] | 0.230 | 0.017 | 0.024 | 0.254 | 0.230 | 0.012 | — |
| 11 | 1 [d] | 0.229 | 0.017 | 0.026 | 0.265 | 0.230 | 0.011 | — |
| 12 | 1 [d] | 0.229 | 0.017 | 0.027 | 0.273 | 0.230 | 0.010 | — |
| 13 | 1 | 0.228 | 0.019 | 0.020 | 0.236 | 0.231 | 0.012 | 1.31 |
| 14 | 1 | 0.224 | 0.019 | 0.017 | 0.222 | 0.232 | 0.008 | 1.17 |
| 15 | 1 | 0.222 | 0.019 | 0.015 | 0.215 | 0.233 | 0.007 | 1.07 |
| 16 | 4 | 0.220 | 0.032 | 0.019 | 0.240 | 0.239 | 0.020 | — |
| 17 | 4 | 0.215 | 0.039 | 0.021 | 0.261 | 0.244 | 0.013 | — |
| 18 | 4 | 0.213 | 0.042 | 0.023 | 0.276 | 0.249 | 0.010 | — |
| 19 | 2 | 0.211 | 0.034 | 0.022 | 0.261 | 0.248 | 0.010 | — |
| 20 | 2 | 0.212 | 0.031 | 0.021 | 0.258 | 0.247 | 0.007 | — |
| 21 | 2 | 0.213 | 0.030 | 0.020 | 0.256 | 0.248 | 0.006 | — |
| 22 | 1 [d] | 0.213 | 0.019 | 0.025 | 0.283 | 0.242 | 0.018 | — |
| 23 | 1 [d] | 0.216 | 0.017 | 0.027 | 0.295 | 0.240 | 0.009 | — |
| 24 | 1 [d] | 0.217 | 0.016 | 0.029 | 0.300 | 0.238 | 0.007 | — |
| 25 | 1 | 0.217 | 0.016 | 0.029 | 0.303 | 0.237 | 0.006 | — |
| 26 | 1 | 0.217 | 0.018 | 0.021 | 0.256 | 0.237 | 0.011 | 1.01 |
| 27 | 1 | 0.216 | 0.018 | 0.018 | 0.238 | 0.236 | 0.006 | 0.92 |
| 28 | 1 | 0.215 | 0.018 | 0.016 | 0.228 | 0.237 | 0.005 | 0.86 |
| 29 | 1 | 0.214 | 0.018 | 0.015 | 0.222 | 0.237 | 0.005 | 0.81 |
| 30 | 1 | 0.213 | 0.018 | 0.015 | 0.217 | 0.237 | 0.005 | — |
| 31 | 1 | 0.213 | 0.018 | 0.015 | 0.214 | 0.237 | 0.004 | — |
| 32 | 1 | 0.213 | 0.018 | 0.015 | 0.212 | 0.237 | 0.004 | — |
| 33 | 1 | 0.213 | 0.018 | 0.014 | 0.210 | 0.237 | 0.004 | — |

[a] Sixty-five water molecules added; 2 deleted; 243 in total. Five carbohydrate residues were also added at this stage, three in the N-lobe and two in the C-lobe. [b] Unit weights applied to distance, planar groups, torsion angles, chiral groups and van der Waals contacts unless otherwise stated. [c] The resolution range was extended to 7.5 Å. [d] Weights of 2.0 applied to bond distances.

Table 13: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 13 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.217 | 0.019 | 0.014 | 0.211 | 0.237 | 0.014 | — |
| 2 | 1 | 0.216 | 0.019 | 0.014 | 0.206 | 0.237 | 0.009 | — |
| 3 | 1 | 0.215 | 0.018 | 0.014 | 0.204 | 0.236 | 0.007 | — |
| 4 | 1 | 0.214 | 0.018 | 0.014 | 0.203 | 0.236 | 0.006 | — |
| 5 | 1 | 0.214 | 0.018 | 0.014 | 0.203 | 0.237 | 0.005 | — |
| 6 | 1 | 0.214 | 0.018 | 0.014 | 0.202 | 0.236 | 0.005 | — |
| 7 | 1 | 0.213 | 0.018 | 0.014 | 0.202 | 0.237 | 0.005 | — |
| 8 | 1 | 0.213 | 0.018 | 0.014 | 0.201 | 0.237 | 0.004 | — |
| 9 | 1 | 0.213 | 0.018 | 0.014 | 0.201 | 0.237 | 0.004 | 0.85 |
| 10 | 1 | 0.212 | 0.018 | 0.014 | 0.201 | 0.237 | 0.004 | 0.75 |
| 11 | 1 | 0.211 | 0.018 | 0.014 | 0.201 | 0.237 | 0.004 | 0.74 |
| 12 | 1 | 0.210 | 0.018 | 0.014 | 0.200 | 0.237 | 0.004 | 0.69 |
| 13 | 4 | 0.210 | 0.029 | 0.017 | 0.221 | 0.243 | 0.017 | — |
| 14 | 4 | 0.206 | 0.035 | 0.020 | 0.239 | 0.248 | 0.011 | — |
| 15 | 4 | 0.203 | 0.038 | 0.021 | 0.254 | 0.251 | 0.008 | — |
| 16 | 2 | 0.202 | 0.031 | 0.020 | 0.247 | 0.251 | 0.009 | — |
| 17 | 2 | 0.203 | 0.029 | 0.020 | 0.245 | 0.251 | 0.005 | — |
| 18 | 2 | 0.203 | 0.028 | 0.020 | 0.245 | 0.251 | 0.004 | — |
| 19 | 1 | 0.203 | 0.021 | 0.017 | 0.269 | 0.248 | 0.012 | — |
| 20 | 1 | 0.206 | 0.019 | 0.016 | 0.278 | 0.246 | 0.006 | — |
| 21 | 1 | 0.207 | 0.018 | 0.015 | 0.281 | 0.245 | 0.005 | — |
| 22 | 1 | 0.207 | 0.018 | 0.015 | 0.242 | 0.244 | 0.005 | 0.65 |
| 23 | 1 | 0.207 | 0.018 | 0.015 | 0.227 | 0.243 | 0.004 | 0.62 |
| 24 | 1 | 0.207 | 0.018 | 0.014 | 0.219 | 0.242 | 0.003 | 0.58 |
| 25 | 1 | 0.207 | 0.018 | 0.014 | 0.214 | 0.242 | 0.003 | 0.57 |
| 26 | 1 | 0.207 | 0.018 | 0.014 | 0.211 | 0.242 | 0.003 | — |
| 27 | 1 | 0.207 | 0.018 | 0.014 | 0.209 | 0.242 | 0.003 | — |
| 28 | 1 | 0.207 | 0.018 | 0.014 | 0.207 | 0.241 | 0.003 | — |

[a] No water molecules added; 243 in total. [b] Unit weights applied to all parameters.

Table 14: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 14 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.233 | 0.021 | 0.014 | 0.206 | 0.240 | 0.024 | — |
| 2 | 1 | 0.228 | 0.020 | 0.014 | 0.204 | 0.239 | 0.016 | — |
| 3 | 1 | 0.226 | 0.020 | 0.014 | 0.202 | 0.240 | 0.014 | — |
| 4 | 1 | 0.224 | 0.020 | 0.014 | 0.202 | 0.240 | 0.011 | — |
| 5 | 1 | 0.223 | 0.020 | 0.014 | 0.202 | 0.240 | 0.009 | — |
| 6 | 1 | 0.222 | 0.020 | 0.014 | 0.202 | 0.240 | 0.008 | 1.26 |
| 7 | 1 | 0.218 | 0.019 | 0.014 | 0.203 | 0.241 | 0.007 | 1.14 |
| 8 | 1 | 0.215 | 0.019 | 0.014 | 0.203 | 0.241 | 0.007 | 1.04 |
| 9 | 1 | 0.213 | 0.019 | 0.013 | 0.204 | 0.241 | 0.006 | 0.95 |
| 10 | 1 [c] | 0.212 | 0.019 | 0.014 | 0.236 | 0.242 | 0.006 | — |
| 11 | 1 [c] | 0.212 | 0.018 | 0.014 | 0.252 | 0.243 | 0.006 | — |
| 12 | 1 [c] | 0.211 | 0.017 | 0.014 | 0.262 | 0.243 | 0.005 | — |
| 13 | 4 | 0.211 | 0.031 | 0.018 | 0.283 | 0.250 | 0.018 | — |
| 14 | 4 | 0.207 | 0.037 | 0.021 | 0.306 | 0.254 | 0.012 | — |
| 15 | 4 | 0.204 | 0.040 | 0.023 | 0.325 | 0.258 | 0.009 | — |
| 16 | 2 | 0.203 | 0.033 | 0.021 | 0.290 | 0.256 | 0.09 | — |
| 17 | 2 | 0.204 | 0.031 | 0.021 | 0.276 | 0.256 | 0.006 | — |
| 18 | 2 | 0.204 | 0.030 | 0.021 | 0.270 | 0.256 | 0.005 | — |
| 19 | 1 [d] | 0.204 | 0.020 | 0.023 | 0.238 | 0.253 | 0.013 | — |
| 20 | 1 [d] | 0.207 | 0.018 | 0.025 | 0.226 | 0.252 | 0.006 | — |
| 21 | 1 [d] | 0.207 | 0.017 | 0.025 | 0.217 | 0.250 | 0.006 | — |
| 22 | 1 [d] | 0.208 | 0.016 | 0.025 | 0.211 | 0.249 | 0.005 | — |
| 23 | 1 | 0.208 | 0.017 | 0.019 | 0.216 | 0.246 | 0.01 | — |
| 24 | 1 | 0.208 | 0.017 | 0.016 | 0.215 | 0.244 | 0.006 | — |
| 25 | 1 | 0.209 | 0.017 | 0.015 | 0.211 | 0.242 | 0.005 | — |
| 26 | 1 | 0.209 | 0.016 | 0.014 | 0.208 | 0.241 | 0.004 | 0.73 |
| 27 | 1 | 0.208 | 0.016 | 0.014 | 0.205 | 0.240 | 0.004 | 0.68 |
| 28 | 1 | 0.208 | 0.016 | 0.014 | 0.203 | 0.239 | 0.004 | 0.63 |
| 29 | 1 | 0.207 | 0.016 | 0.013 | 0.201 | 0.239 | 0.004 | 0.60 |
| 30 | 1 | 0.207 | 0.016 | 0..013 | 0.200 | 0.239 | 0.003 | — |
| 31 | 1 | 0.207 | 0.016 | 0.013 | 0.198 | 0.238 | 0.003 | — |
| 32 | 1 | 0.207 | 0.016 | 0.013 | 0.197 | 0.238 | 0.003 | — |
| 33 | 1 | 0.207 | 0.016 | 0.013 | 0.196 | 0.238 | 0.003 | — |
| 34 | 1 | 0.207 | 0.016 | 0.013 | 0.195 | 0.237 | 0.003 | 0.57 |
| 35 | 1 | 0.207 | 0.016 | 0.013 | 0.194 | 0.237 | 0.003 | 0.54 |
| 36 | 1 | 0.206 | 0.016 | 0.013 | 0.194 | 0.237 | 0.003 | 0.51 |

[a] Fourty-two water molecules added; none deleted; 285 in total. [b] Unit weights applied at all parameters unless otherwise stated. [c] A weight of 1.5 applied to bond distances. [d] Weights of 1.5 applied to bond distances and chiral volumes.

Table 15: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 15 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.210 | 0.016 | 0.013 | 0.178 | 0.237 | 0.012 | — |
| 2 | 1 | 0.209 | 0.016 | 0.013 | 0.173 | 0.236 | 0.008 | — |
| 3 | 1 | 0.208 | 0.016 | 0.013 | 0.170 | 0.236 | 0.006 | — |
| 4 | 1 | 0.208 | 0.016 | 0.013 | 0.169 | 0.236 | 0.005 | — |
| 5 | 1 | 0.207 | 0.016 | 0.013 | 0.168 | 0.237 | 0.005 | — |
| 6 | 1 | 0.207 | 0.016 | 0.013 | 0.167 | 0.237 | 0.004 | 0.61 |
| 7 | 1 | 0.207 | 0.016 | 0.013 | 0.167 | 0.237 | 0.004 | 0.56 |
| 8 | 1 | 0.206 | 0.016 | 0.013 | 0.166 | 0.237 | 0.004 | 0.52 |
| 9 | 1 | 0.206 | 0.016 | 0.013 | 0.166 | 0.237 | 0.004 | 0.49 |
| 10 | 1 | 0.206 | 0.016 | 0.013 | 0.165 | 0.237 | 0.004 | — |
| 11 | 1 | 0.205 | 0.016 | 0.013 | 0.165 | 0.237 | 0.003 | — |
| 12 | 1 | 0.205 | 0.016 | 0.013 | 0.165 | 0.237 | 0.003 | — |
| 13 c | 1 | 0.211 | 0.016 | 0.013 | 0.165 | 0.250 | 0.009 | — |
| 14 | 1 | 0.210 | 0.016 | 0.013 | 0.166 | 0.251 | 0.007 | — |
| 15 | 1 | 0.209 | 0.016 | 0.013 | 0.166 | 0.251 | 0.006 | — |
| 16 | 1 | 0.208 | 0.016 | 0.013 | 0.167 | 0.251 | 0.006 | 0.76 |
| 17 | 1 | 0.206 | 0.016 | 0.013 | 0.167 | 0.252 | 0.005 | 0.65 |
| 18 | 1 | 0.205 | 0.016 | 0.013 | 0.167 | 0.251 | 0.005 | 0.60 |
| 19 | 4 | 0.204 | 0.020 | 0.023 | 0.181 | 0.253 | 0.013 | — |
| 20 | 4 | 0.201 | 0.023 | 0.025 | 0.226 | 0.255 | 0.008 | — |
| 21 | 4 | 0.198 | 0.026 | 0.027 | 0.241 | 0.256 | 0.007 | — |
| 22 | 2 | 0.197 | 0.025 | 0.023 | 0.221 | 0.255 | 0.006 | — |
| 23 | 2 | 0.198 | 0.023 | 0.022 | 0.216 | 0.254 | 0.006 | — |
| 24 | 2 | 0.200 | 0.022 | 0.022 | 0.214 | 0.253 | 0.006 | — |
| 25 | 1 | 0.201 | 0.020 | 0.016 | 0.208 | 0.249 | 0.006 | — |
| 26 | 1 | 0.200 | 0.020 | 0.016 | 0.204 | 0.247 | 0.006 | — |
| 27 | 1 | 0.200 | 0.019 | 0.015 | 0.202 | 0.245 | 0.005 | — |
| 28 | 1 | 0.199 | 0.019 | 0.015 | 0.200 | 0.244 | 0.005 | 0.57 |
| 29 | 1 | 0.198 | 0.019 | 0.015 | 0.199 | 0.243 | 0.005 | 0.53 |
| 30 | 1 | 0.197 | 0.019 | 0.015 | 0.197 | 0.242 | 0.005 | 0.49 |
| 31 | 1 | 0.197 | 0.018 | 0.014 | 0.195 | 0.241 | 0.005 | — |
| 32 | 1 | 0.196 | 0.018 | 0.014 | 0.194 | 0.240 | 0.005 | — |
| 33 | 1 | 0.196 | 0.018 | 0.014 | 0.194 | 0.240 | 0.005 | — |

[a] Twenty-three water molecules added; none deleted; 308 in total. The first five residues were removed during manual rebuilding. [b] Unit weights applied at all parameters. [c] Residues 4 and 5 were added into the appropriate $2Fo\text{-}Fc$ and $Fo\text{-}Fc$ density at this point and refinement continued.

Table 16: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 16 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.209 | 0.018 | 0.015 | 0.197 | 0.239 | 0.018 | |
| 2 | 1 | 0.208 | 0.017 | 0.015 | 0.196 | 0.239 | 0.016 | |
| 3 | 1 | 0.208 | 0.017 | 0.015 | 0.195 | 0.239 | 0.015 | |
| 4 | 1 | 0.207 | 0.016 | 0.014 | 0.195 | 0.239 | 0.013 | 0.51 |
| 5 | 1 | 0.205 | 0.015 | 0.014 | 0.194 | 0.239 | 0.010 | 0.47 |
| 6 | 1 | 0.203 | 0.015 | 0.014 | 0.193 | 0.238 | 0.007 | 0.44 |
| 7 | 4 | 0.200 | 0.019 | 0.021 | 0.199 | 0.243 | 0.014 | |
| 8 | 4 | 0.197 | 0.022 | 0.023 | 0.216 | 0.245 | 0.009 | |
| 9 | 4 | 0.195 | 0.024 | 0.025 | 0.231 | 0.246 | 0.008 | |
| 10 | 2 | 0.195 | 0.023 | 0.022 | 0.221 | 0.245 | 0.007 | |
| 11 | 2 | 0.196 | 0.021 | 0.021 | 0.210 | 0.244 | 0.007 | |
| 12 | 2 | 0.198 | 0.020 | 0.020 | 0.217 | 0.233 | 0.006 | |
| 13 | 1 | 0.199 | 0.018 | 0.016 | 0.210 | 0.239 | 0.006 | |
| 14 | 1 | 0.201 | 0.018 | 0.016 | 0.206 | 0.237 | 0.006 | |
| 15 | 1 | 0.202 | 0.017 | 0.015 | 0.203 | 0.235 | 0.005 | |
| 16 | 1 | 0.202 | 0.017 | 0.015 | 0.200 | 0.234 | 0.005 | 0.43 |
| 17 | 1 | 0.201 | 0.017 | 0.014 | 0.198 | 0.234 | 0.005 | 0.40 |
| 18 | 1 | 0.201 | 0.016 | 0.014 | 0.196 | 0.234 | 0.004 | 0.38 |

[a] No water molecules added; seven deleted; 301 in total. Residues 1, 2 and 3 were added during phase 15 manual rebuilding. [b] Unit weights applied at all parameters.

Table 17: Hendrickson-Konnert refinement of dicupric lactoferrin (2.1Å) - phase 17 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.203 | 0.016 | 0.014 | 0.196 | 0.233 | 0.009 | |
| 2 | 1 | 0.203 | 0.016 | 0.014 | 0.196 | 0.233 | 0.007 | |
| 3 | 1 | 0.202 | 0.017 | 0.014 | 0.195 | 0.233 | 0.005 | |
| 4 | 1 | 0.201 | 0.016 | 0.013 | 0.194 | 0.232 | 0.005 | 0.39 |
| 5 | 1 | 0.201 | 0.015 | 0.013 | 0.194 | 0.232 | 0.005 | 0.38 |
| 6 | 1 | 0.200 | 0.015 | 0.013 | 0.193 | 0.231 | 0.004 | 0.36 |
| 7 | 1 | 0.200 | 0.015 | 0.013 | 0.193 | 0.231 | 0.004 | 0.35 |
| 8 | 1 | 0.199 | 0.015 | 0.013 | 0.192 | 0.230 | 0.004 | |
| 9 | 1 | 0.199 | 0.015 | 0.013 | 0.191 | 0.230 | 0.004 | |
| 10 | 1 | 0.198 | 0.014 | 0.013 | 0.191 | 0.229 | 0.003 | |
| 11 | 1 | 0.198 | 0.014 | 0.013 | 0.190 | 0.229 | 0.003 | 0.34 |

[a] No water molecules added; none deleted; 301 in total. [b] Unit weights applied at all parameters.

B) *$Cu_2$oxLf*

Table 18: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.5Å) - phase 1[a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 [c] | 1 | 0.293 | 0.029 | 0.022 | 0.251 | 0.243 | 0.033 | — |
| 2 | 1 | 0.268 | 0.024 | 0.027 | 0.263 | 0.245 | 0.027 | — |
| 3 | 1 | 0.261 | 0.023 | 0.030 | 0.273 | 0.248 | 0.022 | — |
| 4 | 1 | 0.254 | 0.022 | 0.032 | 0.281 | 0.251 | 0.019 | — |
| 5 | 1 | 0.249 | 0.022 | 0.030 | 0.275 | 0.251 | 0.025 | 2.07 |
| 6 | 1 | 0.234 | 0.022 | 0.029 | 0.268 | 0.253 | 0.023 | 2.00 |
| 7 | 1 | 0.222 | 0.022 | 0.027 | 0.261 | 0.254 | 0.020 | 1.91 |
| 8 | 1 | 0.212 | 0.022 | 0.025 | 0.258 | 0.255 | 0.019 | — |
| 9 | 1 | 0.208 | 0.023 | 0.020 | 0.246 | 0.256 | 0.014 | — |
| 10 | 1 | 0.206 | 0.023 | 0.019 | 0.239 | 0.257 | 0.011 | — |
| 11 | 1 [d] | 0.205 | 0.023 | 0.018 | 0.286 | 0.260 | 0.011 | 1.73 |
| 12 | 1 [d] | 0.197 | 0.022 | 0.017 | 0.309 | 0.262 | 0.010 | 1.51 |
| 13 | 1 [d] | 0.191 | 0.021 | 0.017 | 0.322 | 0.262 | 0.009 | 1.35 |
| 14 | 4 | 0.186 | 0.035 | 0.021 | 0.326 | 0.269 | 0.020 | — |
| 15 | 4 | 0.178 | 0.042 | 0.024 | 0.333 | 0.274 | 0.013 | — |
| 16 | 4 | 0.173 | 0.045 | 0.026 | 0.340 | 0.277 | 0.011 | — |
| 17 | 2 | 0.170 | 0.036 | 0.024 | 0.316 | 0.276 | 0.011 | 0.87 |
| 18 | 2 | 0.169 | 0.033 | 0.023 | 0.305 | 0.275 | 0.008 | 0.82 |
| 19 | 2 | 0.167 | 0.032 | 0.023 | 0.299 | 0.276 | 0.007 | 0.77 |
| 20 | 1 [e] | 0.165 | 0.022 | 0.023 | 0.321 | 0.272 | 0.015 | — |
| 21 | 1 [e] | 0.169 | 0.020 | 0.023 | 0.332 | 0.269 | 0.009 | — |
| 22 | 1 [e] | 0.171 | 0.019 | 0.023 | 0.337 | 0.267 | 0.007 | — |

[a] The coordinates of dicupric lactoferrin with the copper atoms, ligands, carbonate anions and water molecules removed, was used as the initial model. [b] Unit weights applied to distances, planar groups, torsion angles, chiral volumes and van der Waals contacts unless otherwise stated. [c] Resolution limits initially 5.0Å - 2.7Å. [d] Weights of 2.0 were applied to distances and planes respectively. [e] Weights of 2.0 and 1.5 were applied to distances and planes respectively.

Table 19: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.5Å) - phase 2 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 [c] | 1 | 0.179 | 0.020 | 0.018 | 0.274 | 0.263 | 0.016 | 0.99 |
| 2 | 1 | 0.173 | 0.020 | 0.017 | 0.254 | 0.263 | 0.011 | 0.81 |
| 3 | 1 | 0.170 | 0.020 | 0.016 | 0.244 | 0.263 | 0.009 | 0.72 |
| 4 | 1 | 0.168 | 0.020 | 0.016 | 0.237 | 0.263 | 0.008 | 0.65 |
| 5 | 1 | 0.166 | 0.020 | 0.016 | 0.233 | 0.263 | 0.007 | 0.60 |
| 6 | 1 | 0.164 | 0.020 | 0.015 | 0.230 | 0.262 | 0.006 | 0.56 |
| 7 | 4 | 0.163 | 0.032 | 0.019 | 0.254 | 0.269 | 0.017 | — |
| 8 | 4 | 0.156 | 0.038 | 0.022 | 0.273 | 0.273 | 0.011 | — |
| 9 | 4 | 0.152 | 0.041 | 0.023 | 0.287 | 0.276 | 0.009 | — |
| 10 | 2 | 0.150 | 0.033 | 0.022 | 0.278 | 0.275 | 0.009 | 0.61 |
| 11 | 2 | 0.150 | 0.030 | 0.021 | 0.275 | 0.275 | 0.006 | 0.58 |
| 12 | 2 [d] | 0.149 | 0.027 | 0.024 | 0.304 | 0.275 | 0.007 | — |
| 13 | 1 | 0.143 | 0.022 | 0.019 | 0.272 | 0.271 | 0.012 | 0.46 |
| 14 | 1 | 0.153 | 0.020 | 0.017 | 0.256 | 0.269 | 0.007 | 0.44 |
| 15 | 1 | 0.154 | 0.020 | 0.016 | 0.247 | 0.267 | 0.006 | 0.43 |
| 16 [e] | 1 | 0.160 | 0.020 | 0.016 | 0.241 | 0.267 | 0.010 | — |
| 17 | 1 | 0.158 | 0.020 | 0.016 | 0.238 | 0.268 | 0.007 | — |
| 18 | 1 | 0.156 | 0.020 | 0.016 | 0.235 | 0.268 | 0.006 | — |
| 19 | 1 | 0.155 | 0.020 | 0.016 | 0.233 | 0.267 | 0.006 | — |
| 20 [f] | 1 | 0.169 | 0.020 | 0.015 | 0.230 | 0.266 | 0.010 | — |
| 21 | 1 | 0.168 | 0.020 | 0.015 | 0.228 | 0.266 | 0.007 | — |
| 22 | 1 | 0.168 | 0.020 | 0.015 | 0.227 | 0.265 | 0.006 | — |
| 23 [g] | 1 | 0.178 | 0.020 | 0.015 | 0.225 | 0.265 | 0.008 | — |
| 24 | 1 | 0.177 | 0.020 | 0.015 | 0.223 | 0.264 | 0.007 | — |
| 25 | 1 | 0.177 | 0.019 | 0.015 | 0.222 | 0.264 | 0.006 | — |

[a] The four protein ligands and a copper atom were added to the N-lobe metal binding site prior to this phase of refinement. No ligands, metal ions or anions were included in the C-lobe site at this stage. [b] Unit weights applied to distance, planar groups, torsion angles, chiral volumes and van der Waals contacts unless otherwise stated. [c] Resolution limits initially 5.0Å - 2.8Å. [d] Weights of 2.0 and 1.5 applied to distances and planes respectively, for 1 cycle. [e] Resolution limits reset to 5.0 - 2.75 Å. [f] Resolution limits reset to 5.0 - 2.7 Å. [g] Resolution limits reset to 5.0 - 2.6 Å.

Table 20: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.5Å) - phase 3 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 [c] | 1 | 0.246 | 0.024 | 0.015 | 0.217 | 0.265 | 0.021 | — |
| 2 | 1 | 0.235 | 0.025 | 0.015 | 0.220 | 0.266 | 0.014 | — |
| 3 | 1 | 0.230 | 0.025 | 0.015 | 0.221 | 0.267 | 0.011 | — |
| 4 | 1 [d] | 0.227 | 0.022 | 0.022 | 0.258 | 0.269 | 0.014 | — |
| 5 | 1 | 0.225 | 0.024 | 0.018 | 0.239 | 0.268 | 0.010 | — |
| 6 | 1 | 0.223 | 0.025 | 0.017 | 0.232 | 0.271 | 0.008 | — |
| 7 | 1 | 0.221 | 0.025 | 0.016 | 0.230 | 0.271 | 0.007 | — |
| 8 | 1 | 0.220 | 0.025 | 0.016 | 0.228 | 0.271 | 0.007 | — |
| 9 | 1 | 0.219 | 0.025 | 0.016 | 0.228 | 0.271 | 0.006 | — |
| 10 | 1 | 0.218 | 0.025 | 0.016 | 0.227 | 0.272 | 0.006 | — |
| 11 | 4 | 0.217 | 0.042 | 0.020 | 0.267 | 0.279 | 0.021 | — |
| 12 | 4 | 0.207 | 0.051 | 0.023 | 0.296 | 0.285 | 0.013 | — |
| 13 | 4 | 0.201 | 0.055 | 0.025 | 0.317 | 0.289 | 0.010 | — |
| 14 | 2 | 0.198 | 0.045 | 0.024 | 0.297 | 0.288 | 0.011 | — |
| 15 | 2 | 0.201 | 0.041 | 0.023 | 0.291 | 0.288 | 0.006 | — |
| 16 | 2 | 0.201 | 0.040 | 0.023 | 0.289 | 0.289 | 0.005 | — |
| 17 | 1 | 0.201 | 0.029 | 0.019 | 0.264 | 0.284 | 0.014 | — |
| 18 | 1 | 0.207 | 0.026 | 0.018 | 0.253 | 0.282 | 0.007 | — |
| 19 | 1 | 0.208 | 0.025 | 0.017 | 0.247 | 0.281 | 0.005 | — |
| 20 | 1 | 0.209 | 0.025 | 0.016 | 0.243 | 0.280 | 0.005 | — |
| 21 | 1 | 0.209 | 0.025 | 0.016 | 0.240 | 0.279 | 0.004 | — |
| 22 | 1 | 0.210 | 0.024 | 0.016 | 0.238 | 0.279 | 0.004 | — |
| 23 | 1 | 0.210 | 0.024 | 0.016 | 0.236 | 0.279 | 0.004 | |
| 24 | 1 | 0.202 | 0.024 | 0.016 | 0.235 | 0.278 | 0.004 | |
| 25 | 1 | 0.196 | 0.023 | 0.015 | 0.233 | 0.277 | 0.004 | |

[a] The N-lobe anion was added, along with the sidechains of Arg 121 and Thr 117. The ligands, metal ion and an oxalate were added into the C-lobe site. The sidechains of Arg 465 and Thr 461 were also included. [b] Unit weights applied to all parameters. [c] Resolution limits initially 5.0Å - 2.6Å. [d] A weight of 2.0 applied to distances for 1 cycle.

Table 21: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.5Å) - phase 4

| Cycle | X-ray weight [a] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 [b] | 1 | 0.215 | 0.027 | 0.016 | 0.233 | 0.278 | 0.013 | — |
| 2 | 1 | 0.210 | 0.028 | 0.017 | 0.237 | 0.279 | 0.008 | — |
| 3 | 1 | 0.208 | 0.028 | 0.017 | 0.239 | 0.280 | 0.006 | — |
| 4 [c] | 1 | 0.210 | 0.027 | 0.017 | 0.239 | 0.281 | 0.006 | |
| 5 | 1 | 0.200 | 0.026 | 0.017 | 0.238 | 0.281 | 0.005 | 1.58 |
| 6 | 1 | 0.194 | 0.025 | 0.017 | 0.237 | 0.281 | 0.005 | 1.38 |
| 7 | 1 | 0.189 | 0.024 | 0.016 | 0.237 | 0.280 | 0.005 | 1.24 |
| 8 | 1 | 0.186 | 0.024 | 0.016 | 0.236 | 0.281 | 0.005 | 1.14 |
| 9 | 1 | 0.182 | 0.023 | 0.016 | 0.235 | 0.281 | 0.005 | 1.05 |
| 10 | 4 | 0.180 | 0.039 | 0.021 | 0.264 | 0.287 | 0.019 | — |
| 11 | 4 | 0.172 | 0.046 | 0.024 | 0.294 | 0.292 | 0.012 | — |
| 12 | 4 | 0.168 | 0.049 | 0.026 | 0.311 | 0.296 | 0.004 | — |
| 13 | 2 | 0.166 | 0.039 | 0.024 | 0.295 | 0.295 | 0.010 | — |
| 14 | 2 | 0.168 | 0.037 | 0.024 | 0.291 | 0.295 | 0.005 | — |
| 15 | 2 | 0.168 | 0.036 | 0.023 | 0.289 | 0.296 | 0.004 | — |
| 16 | 1 | 0.168 | 0.027 | 0.020 | 0.267 | 0.291 | 0.013 | — |
| 17 | 1 | 0.173 | 0.024 | 0.018 | 0.257 | 0.289 | 0.006 | — |
| 18 | 1 | 0.174 | 0.024 | 0.017 | 0.251 | 0.287 | 0.005 | — |
| 19 | 1 | 0.175 | 0.023 | 0.017 | 0.247 | 0.286 | 0.004 | 1.00 |
| 20 | 1 | 0.173 | 0.023 | 0.017 | 0.244 | 0.285 | 0.004 | 0.92 |
| 21 | 1 | 0.171 | 0.023 | 0.017 | 0.242 | 0.284 | 0.004 | 0.86 |
| 22 | 1 | 0.170 | 0.022 | 0.016 | 0.240 | 0.284 | 0.004 | — |
| 23 | 1 | 0.170 | 0.022 | 0.016 | 0.238 | 0.283 | 0.003 | — |

[a] Unit weights applied to distance, planar groups, torsion angles, chiral volumes and van der Waals contacts unless otherwise stated. [b] Resolution limits initially 5.0Å - 2.6Å. [c] Resolution limits reset to 5.0Å - 2.5Å.

Table 22: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.5Å) - phase 5 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.2426 | 0.021 | 0.015 | 0.235 | 0.277 | 0.025 | — |
| 2 | 1 | 0.2354 | 0.022 | 0.015 | 0.232 | 0.274 | 0.017 | — |
| 3 | 1 | 0.2315 | 0.022 | 0.015 | 0.231 | 0.272 | 0.014 | — |
| 4 | 1 | 0.2289 | 0.021 | 0.015 | 0.230 | 0.270 | 0.012 | — |
| 5 | 1 | 0.2270 | 0.021 | 0.015 | 0.229 | 0.269 | 0.010 | — |
| 6 | 1 | 0.2255 | 0.021 | 0.015 | 0.229 | 0.068 | 0.009 | — |
| 7 | 1 | 0.2244 | 0.021 | 0.015 | 0.228 | 0.268 | 0.008 | — |
| 8 | 1 | 0.2234 | 0.021 | 0.015 | 0.228 | 0.267 | 0.008 | — |
| 9 | 4 | 0.2226 | 0.036 | 0.019 | 0.264 | 0.273 | 0.019 | — |
| 10 | 4 | 0.2155 | 0.043 | 0.022 | 0.292 | 0.277 | 0.013 | — |
| 11 | 2 | 0.2116 | 0.036 | 0.022 | 0.283 | 0.276 | 0.009 | — |
| 12 | 2 | 0.2124 | 0.034 | 0.021 | 0.282 | 0.276 | 0.007 | — |
| 13 | 2 | 0.2121 | 0.033 | 0.021 | 0.283 | 0.277 | 0.006 | — |
| 14 | 1 | 0.2116 | 0.025 | 0.018 | 0.259 | 0.274 | 0.012 | — |
| 15 | 1 | 0.2150 | 0.022 | 0.017 | 0.249 | 0.271 | 0.007 | — |
| 16 | 1 | 0.2162 | 0.022 | 0.016 | 0.243 | 0.270 | 0.006 | — |
| 17 | 1 | 0.2166 | 0.021 | 0.016 | 0.239 | 0.269 | 0.005 | 2.52 |
| 18 | 1 | 0.2190 | 0.022 | 0.016 | 0.236 | 0.268 | 0.006 | 1.43 |
| 19 | 1 | 0.2201 | 0.022 | 0.016 | 0.233 | 0.268 | 0.006 | 0.92 |
| 20 | 1 | 0.2203 | 0.022 | 0.015 | 0.231 | 0.267 | 0.005 | — |
| 21 | 1 | 0.2199 | 0.022 | 0.015 | 0.230 | 0.266 | 0.005 | — |
| 22 | 1 | 0.2196 | 0.022 | 0.015 | 0.228 | 0.266 | 0.005 | — |

[a] Photographic data with resolution limits 5.0Å - 2.5Å was used in all subsequent refinement. The 8 ligands and Arg 121 and 465 were changed to Ala and the metal ions and anions were removed. [b] Unit weights applied to all parameters.

Table 23: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.1Å) - phase 6 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.282 | 0.019 | 0.014 | 0.201 | 0.258 | 0.023 | — |
| 2 | 1 | 0.274 | 0.019 | 0.014 | 0.198 | 0.256 | 0.017 | — |
| 3 | 1 | 0.269 | 0.019 | 0.014 | 0.196 | 0.254 | 0.014 | — |
| 4 | 1 | 0.266 | 0.019 | 0.013 | 0.195 | 0.253 | 0.012 | 1.12 |
| 5 | 1 | 0.259 | 0.019 | 0.013 | 0.194 | 0.253 | 0.011 | 0.94 |
| 6 | 1 | 0.253 | 0.019 | 0.013 | 0.194 | 0.252 | 0.010 | 0.85 |
| 7 | 4 | 0.248 | 0.034 | 0.017 | 0.225 | 0.258 | 0.020 | — |
| 8 | 4 | 0.240 | 0.040 | 0.020 | 0.250 | 0.062 | 0.014 | — |
| 9 | 4 | 0.235 | 0.044 | 0.022 | 0.268 | 0.266 | 0.011 | — |
| 10 | 2 | 0.232 | 0.035 | 0.021 | 0.252 | 0.264 | 0.011 | — |
| 11 | 2 | 0.233 | 0.032 | 0.020 | 0.247 | 0.264 | 0.008 | — |
| 12 | 2 | 0.233 | 0.031 | 0.020 | 0.246 | 0.265 | 0.007 | — |
| 13 | 1 | | | | | | | 0.84 |
| 14 | 1 | | | | | | | 0.81 |
| 15 | 1 | 0.236 | 0.020 | 0.015 | 0.209 | 0.255 | 0.007 | — |
| 16 | 1 | 0.236 | 0.019 | 0.014 | 0.206 | 0.254 | 0.006 | — |
| 17 | 1 | 0.236 | 0.019 | 0.014 | 0.203 | 0.251 | 0.005 | — |
| 18 | 1 | 0.236 | 0.019 | 0.014 | 0.201 | 0.252 | 0.005 | 0.80 |
| 19 | 1 | 0.233 | 0.019 | 0.014 | 0.200 | 0.252 | 0.005 | 0.76 |
| 20 | 1 | 0.230 | 0.019 | 0.014 | 0.199 | 0.251 | 0.005 | 0.74 |
| 21 | 4 | 0.228 | 0.033 | 0.017 | 0.229 | 0.257 | 0.018 | — |
| 22 | 4 | 0.222 | 0.039 | 0.020 | 0.253 | 0.262 | 0.011 | — |
| 23 | 4 | 0.219 | 0.043 | 0.022 | 0.269 | 0.266 | 0.009 | — |
| 24 | 2 | 0.216 | 0.034 | 0.020 | 0.254 | 0.264 | 0.009 | — |
| 25 | 2 | 0.218 | 0.031 | 0.020 | 0.250 | 0.264 | 0.005 | — |
| 26 | 2 | 0.219 | 0.030 | 0.019 | 0.248 | 0.265 | 0.004 | — |
| 27 | 1 | 0.219 | 0.020 | 0.017 | 0.228 | 0.260 | 0.012 | — |
| 28 | 1 | 0.222 | 0.020 | 0.015 | 0.219 | 0.257 | 0.007 | — |
| 29 | 1 | 0.223 | 0.020 | 0.015 | 0.213 | 0.256 | 0.005 | — |
| 30 | 1 [c] | 0.224 | 0.018 | 0.018 | 0.232 | 0.254 | 0.007 | — |
| 31 | 1 [c] | 0.224 | 0.017 | 0.020 | 0.240 | 0.253 | 0.005 | — |
| 32 | 1 [c] | 0.225 | 0.017 | 0.021 | 0.244 | 0.252 | 0.005 | — |

[a] One hundred and twenty-eight water molecules common to both $Fe_2Lf$ and $Cu_2Lf$ were added and checked; Twenty-three of these were deleted. Eighty-nine new water molecules added; 194 in total. [b] Unit weights applied to all parameters. [c] A weight of 1.5 applied to bond distances.

Table 24: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.1Å) - phase 7 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.235 | 0.018 | 0.020 | 0.243 | 0.250 | 0.015 | — |
| 2 | 1 | 0.232 | 0.018 | 0.020 | 0.241 | 0.250 | 0.012 | — |
| 3 | 1 | 0.228 | 0.018 | 0.020 | 0.239 | 0.250 | 0.010 | — |
| 4 | 1 [c] | 0.225 | 0.017 | 0.018 | 0.240 | 0.250 | 0.008 | — |
| 5 | 1 [c] | 0.224 | 0.016 | 0.017 | 0.241 | 0.251 | 0.007 | — |
| 6 | 1 [c] | 0.223 | 0.015 | 0.016 | 0.241 | 0.251 | 0.006 | — |
| 7 [d] | 1 | 0.245 | 0.015 | 0.016 | 0.238 | 0.249 | 0.017 | — |
| 8 | 1 | 0.243 | 0.016 | 0.016 | 0.237 | 0.248 | 0.014 | — |
| 9 | 1 | 0.241 | 0.016 | 0.016 | 0.236 | 0.248 | 0.011 | — |
| 10 | 1 | 0.241 | 0.016 | 0.016 | 0.235 | 0.247 | 0.009 | 0.79 |
| 11 | 1 | 0.240 | 0.016 | 0.015 | 0.235 | 0.247 | 0.007 | 0.76 |
| 12 | 1 | 0.236 | 0.016 | 0.015 | 0.234 | 0.246 | 0.006 | 0.72 |
| 13 | 4 | 0.234 | 0.021 | 0.017 | 0.244 | 0.250 | 0.016 | — |
| 14 | 4 | 0.228 | 0.028 | 0.020 | 0.249 | 0.251 | 0.013 | — |
| 15 | 4 | 0.225 | 0.032 | 0.025 | 0.254 | 0.253 | 0.010 | — |
| 16 | 2 | 0.223 | 0.029 | 0.023 | 0.254 | 0.253 | 0.009 | — |
| 17 | 2 | 0.226 | 0.027 | 0.021 | 0.254 | 0.253 | 0.008 | — |
| 18 | 2 | 0.228 | 0.024 | 0.020 | 0.253 | 0.253 | 0.007 | — |
| 19 | 1 | 0.230 | 0.020 | 0.019 | 0.248 | 0.252 | 0.007 | — |
| 20 | 1 | 0.231 | 0.020 | 0.019 | 0.245 | 0.252 | 0.006 | — |
| 21 | 1 | 0.231 | 0.020 | 0.019 | 0.243 | 0.252 | 0.005 | — |
| 22 | 1 | 0.231 | 0.020 | 0.018 | 0.241 | 0.252 | 0.005 | — |
| 23 | 1 | 0.230 | 0.020 | 0.018 | 0.239 | 0.252 | 0.005 | 0.73 |
| 24 | 1 | 0.228 | 0.019 | 0.017 | 0.236 | 0.252 | 0.005 | 0.68 |
| 25 | 1 | 0.225 | 0.019 | 0.017 | 0.235 | 0.251 | 0.005 | 0.64 |
| 26 | 1 | 0.223 | 0.018 | 0.017 | 0.235 | 0.249 | 0.004 | 0.61 |
| 27 | 1 [e] | 0.222 | 0.018 | 0.016 | 0.231 | 0.250 | 0.004 | — |
| 28 | 1 [e] | 0.222 | 0.019 | 0.015 | 0.225 | 0.250 | 0.004 | — |
| 29 | 1 [e] | 0.222 | 0.019 | 0.014 | 0.220 | 0.250 | 0.003 | — |
| 30 | 1 | 0.221 | 0.019 | 0.014 | 0.219 | 0.249 | 0.003 | — |
| 31 | 1 | 0.221 | 0.018 | 0.013 | 0.219 | 0.249 | 0.003 | — |

[a] Twenty-eight water molecules added; 3 deleted; 219 in total. [b] Unit weights applied to all parameters. [c] A weight of 2.0 was applied to both distances and planes. d The resolution range was extended to 7.5 - 2.1 Å

Table 25: Hendrickson-Konnert refinement of dicupric-oxalato lactoferrin (2.1Å) - phase 8 [a]

| Cycle | X-ray weight [b] | R | Distance rms | Planar rms | Chiral rms | van der Waals rms | $D_{tot}$ | $B_{tot}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.223 | 0.018 | 0.014 | 0.206 | 0.250 | 0.011 | — |
| 2 | 1 | 0.222 | 0.018 | 0.014 | 0.205 | 0.250 | 0.008 | — |
| 3 | 1 | 0.221 | 0.018 | 0.014 | 0.204 | 0.250 | 0.007 | — |
| 4 | 1 | 0.221 | 0.018 | 0.014 | 0.204 | 0.250 | 0.006 | 0.71 |
| 5 | 1 | 0.219 | 0.018 | 0.014 | 0.203 | 0.250 | 0.006 | 0.62 |
| 6 | 1 | 0.218 | 0.018 | 0.014 | 0.203 | 0.250 | 0.006 | 0.57 |
| 7 | 1 | 0.217 | 0.018 | 0.014 | 0.202 | 0.250 | 0.005 | 0.53 |
| 8 [c] | 1 | 0.225 | 0.018 | 0.014 | 0.201 | 0.250 | 0.005 | — |
| 9 | 1 | 0.225 | 0.017 | 0.014 | 0.201 | 0.250 | 0.005 | — |
| 10 | 1 | 0.224 | 0.017 | 0.014 | 0.201 | 0.250 | 0.005 | — |
| 11 | 4 | 0.224 | 0.037 | 0.017 | 0.225 | 0.256 | 0.019 | — |
| 12 | 4 | 0.219 | 0.041 | 0.020 | 0.248 | 0.260 | 0.012 | — |
| 13 | 4 | 0.216 | 0.044 | 0.021 | 0.264 | 0.264 | 0.009 | — |
| 14 | 2 | 0.214 | 0.035 | 0.020 | 0.253 | 0.263 | 0.010 | — |
| 15 | 2 | 0.216 | 0.032 | 0.020 | 0.251 | 0.258 | 0.005 | — |
| 16 | 2 | 0.216 | 0.032 | 0.019 | 0.250 | 0.255 | 0.004 | — |
| 17 | 1 | 0.216 | 0.023 | 0.017 | 0.271 | 0.253 | 0.013 | — |
| 18 | 1 | 0.219 | 0.021 | 0.015 | 0.278 | 0.252 | 0.007 | — |
| 19 | 1 | 0.221 | 0.020 | 0.015 | 0.281 | 0.251 | 0.005 | — |
| 20 | 1 | 0.221 | 0.020 | 0.015 | 0.246 | 0.251 | 0.005 | — |
| 21 | 1 | 0.221 | 0.020 | 0.015 | 0.233 | 0.251 | 0.004 | — |
| 22 | 1 | 0.221 | 0.020 | 0.015 | 0.227 | 0.250 | 0.004 | — |
| 23 | 1 | 0.221 | 0.020 | 0.015 | 0.223 | 0.250 | 0.004 | 0.59 |
| 24 | 1 | 0.220 | 0.019 | 0.015 | 0.220 | 0.250 | 0.004 | 0.57 |
| 25 | 1 | 0.216 | 0.019 | 0.015 | 0.217 | 0.250 | 0.004 | 0.55 |
| 26 | 1 | 0.213 | 0.019 | 0.014 | 0.215 | 0.250 | 0.003 | — |
| 27 | 1 | 0.212 | 0.018 | 0.014 | 0.213 | 0.250 | 0.003 | — |
| 28 | 1 | 0.212 | 0.018 | 0.013 | 0.212 | 0.249 | 0.003 | — |

[a] Eighteen new water molecules added; none deleted; 237 in total. [b] Unit weights applied to all parameters.
[c] Resolution range extended to 7.5 - 2.1 Å.

## Appendix IV

*Hydrogen bonding in the helices, β-sheets and turns in dicupric lactoferrin*

Hydrogen bonds were identified as described in Section 3.2.3. Helices and β-sheets were defined in terms of the observed hydrogen bonding patterns, while turns were identified by their mainchain torsion angles (ϕ and ψ). Tables 1 and 2 summarise the helix and β-sheet hydrogen bonds, while the complete lists of the hydrogen bonding observed in the helices and sheets are given in tables 3 and 4. Table 5 lists the β- and γ-turns in $Cu_2Lf$.

Table 1: Summary of the hydrogen bonding in the helices of $Cu_2Lf$

| | Average distances and angles in each helix * | | | |
|---|---|---|---|---|
| N-lobe | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
| 1 | 2.05 (2.06) | 143 (137) | 154 (154) | 18 (17) |
| 2 | 1.91 (2.01) | 148 (136) | 155 (149) | 17 (21) |
| 3 | 2.04 (2.13) | 145 (125) | 159 (142) | 14 (27) |
| 5 | 2.03 (2.05) | 137 (139) | 145 (146) | 24 (23) |
| 6 | 2.15 (2.13) | 154 (139) | 153 (142) | 19 (26) |
| 7 | 1.97 (2.04) | 146 (140) | 156 (149) | 18 (23) |
| 8 | 1.93 (1.99) | 144 (137) | 154 (147) | 17 (23) |
| 9 | 2.01 (1.95) | 149 (145) | 153 (151) | 18 (19) |
| 10 | 1.87 (2.01) | 152 (137) | 148 (144) | 21 (25) |
| 11 | 2.16 (2.18) | 143 (130) | 146 (145) | 24 (25) |
| 12 | 2.18 (2.23) | 151 (140) | 157 (147) | 16 (23) |
| average [a] | 2.05 (2.07) | 146 (136) | 152 (149) | 19 (22) |
| C-lobe | | | | |
| 1 | 2.03 (2.07) | 146 (136) | 153 (147) | 18 (23) |
| 2 | 2.13 (2.15) | 143 (136) | 152 (143) | 19 (26) |
| 3 | 2.16 (2.19) | 137 (124) | 142 (141) | 26 (27) |
| 5 | 2.19 (2.17) | 139 (139) | 144 (148) | 25 (22) |
| 7 | 2.06 (2.04) | 149 (144) | 156 (151) | 17 (20) |
| 8 | 2.07 (2.00) | 155 (135) | 157 (156) | 16 (16) |
| 9 | 2.04 (2.09) | 141 (135) | 154 (151) | 18 (20) |
| 10 | 1.79 (1.79) | 148 (148) | 151 (151) | 19 (19) |
| 11 | 2.02 (2.11) | 148 (136) | 155 (149) | 17 (21) |
| 12 | 2.10 (2.13) | 146 (133) | 153 (149) | 19 (22) |
| average [a] | 2.10 (2.10) | 143 (135) | 150 (148) | 20 (22) |
| overall averages | | | | |
| a-type | 2.07 | 145 | 151 | 20 |
| all bonds | 2.09 | 136 | 149 | 22 |

* Average values given for α-type hydrogen bonds, while the values for all the hydrogen bonds in a particular helix given in parentheses. The hydrogen bonded interactions used to determine the averages were those with NH···O distances less than or equal to 2.5 Å (Table 3). [a] For the N- and C-lobes as a whole.

Table 2: Summary of the hydrogen bonding in the β-sheets of $Cu_2Lf$

| | Average distances and angles in each strand | | | | |
|---|---|---|---|---|---|
| N-lobe | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) |
| b - a | 1.90 | 148 | 150 | 2.89 | 21 |
| a - c | 2.14 | 148 | 148 | 3.01 | 22 |
| c - j | 2.10 | 161 | 156 | 3.04 | 16 |
| d - j | 1.94 | 153 | 168 | 2.92 | 8 |
| d - k | 1.95 | 143 | 155 | 2.87 | 17 |
| e - j | 2.04 | 140 | 155 | 2.97 | 17 |
| e - i | 1.98 | 158 | 163 | 2.94 | 12 |
| e - h | 1.85 | 151 | 147 | 2.83 | 23 |
| f - g | 2.26 | 150 | 151 | 3.16 | 20 |
| f - h | 1.90 | 146 | 147 | 2.96 | 22 |
| average [a] | 2.00 | 150 | 157 | 2.93 | 16 |
| C-lobe | | | | | |
| b - a | 2.04 | 153 | 159 | 2.98 | 14 |
| a - c | 1.88 | 156 | 154 | 2.82 | 17 |
| c - j | 1.83 | 158 | 159 | 2.79 | 13 |
| d - j | 1.84 | 161 | 159 | 2.94 | 12 |
| d - k | 1.90 | 146 | 160 | 2.85 | 13 |
| e - j | 1.89 | 148 | 145 | 2.73 | 23 |
| e - i | 1.93 | 148 | 154 | 2.86 | 17 |
| e - h | 2.21 | 148 | 154 | 3.11 | 19 |
| f - g | 2.14 | 150 | 140 | 2.96 | 28 |
| f - h | 1.99 | 143 | 147 | 2.84 | 23 |
| average [a] | 1.97 | 151 | 155 | 2.88 | 17 |
| overall [b] | 1.98 | 151 | 156 | 2.91 | 16 |

[a] For all hydrogen bonds in each lobe, excluding those interaction with unfavourable distances and/or geometries (Table 4). [b] Excluding the unfavourable interactions.

Table 3: Hydrogen bonding in the helices of dicupric human lactoferrin

| N-terminal lobe | | | | | | C-terminal lobe | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| helix 1 | hydrogen bond [a] | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) | helix 1 | hydrogen bond [a] | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
| 12 - 16 | α | 2.28 | 102 | 152 | 20 | 351 - 355 | α | 2.51 | 139 | 138 | 31 |
| 13 - 17 | α | 1.63 | 155 | 149 | 19 | 352 - 356 | α | 2.18 | 139 | 154 | 18 |
| 14 - 18 | α | 2.38 | 124 | 160 | 14 | 353 - 357 | α | 2.30 | 147 | 155 | 17 |
| 15 - 19 | α | 1.73 | 149 | 169 | 7 | 354 - 358 | α | 2.04 | 128 | 141 | 26 |
| 16 - 20 | α | 1.87 | 131 | 137 | 29 | 355 - 359 | α | 1.63 | 169 | 175 | 3 |
| 17 - 21 | α | 2.04 | 153 | 171 | 6 | 356 - 360 | α | 2.18 | 139 | 155 | 17 |
| 18 - 22 | α | 1.73 | 169 | 165 | 9 | 357 - 361 | α | 2.06 | 133 | 142 | 26 |
| 19 - 23 | α | 1.98 | 133 | 158 | 15 | 358 - 361 | $3_{10}$ | 2.32 | 97 | 116 | 46 |
| 20 - 24 | α | 2.48 | 151 | 142 | 27 | 358 - 362 | α | 1.78 | 157 | 149 | 20 |
| * 21 - 24 | $3_{10}$ | 2.55 | 96 | 131 | 36 | 359 - 362 | $3_{10}$ | 2.40 | 96 | 110 | 51 |
| 21 - 25 | α | 1.96 | 162 | 157 | 15 | 359 - 363 | α | 1.58 | 165 | 166 | 8 |
| 22 - 26 | α | 2.37 | 158 | 152 | 20 | 362 - 365 | $3_{10}$ | 1.82 | 123 | 160 | 13 |
| * 23 - 27 | α | 2.67 | 123 | 172 | 6 | | | | | | |
| 24 - 28 | α | 2.39 | 130 | 129 | 36 | | | | | | |
| 26 - 29 | $3_{10}$ | 2.24 | 114 | 147 | 23 | | | | | | |
| 26 - 31 | π | 1.77 | 144 | 155 | 16 | | | | | | |
| 27 - 30 | $3_{10}$ | 2.20 | 91 | 127 | 37 | | | | | | |
| average values | | 2.05 | 143 | 154 | 18 | average values | | 2.03 | 146 | 153 | 18 |
| helix 2 | | | | | | helix 2 | | | | | |
| 41 - 44 | $3_{10}$ | 2.21 | 117 | 122 | 41 | 376 - 380 | α | 2.21 | 129 | 139 | 28 |
| 41 - 45 | α | 1.87 | 159 | 159 | 13 | 377 - 381 | α | 1.83 | 140 | 146 | 22 |
| 42 - 46 | α | 1.74 | 136 | 146 | 22 | * 378 - 381 | $3_{10}$ | 2.51 | 91 | 113 | 49 |
| 43 - 47 | α | 1.64 | 137 | 162 | 11 | 378 - 382 | α | 2.01 | 152 | 162 | 12 |
| 44 - 48 | α | 1.92 | 152 | 170 | 7 | 379 - 383 | α | 2.20 | 143 | 166 | 10 |
| 45 - 49 | α | 1.97 | 174 | 167 | 9 | * 380 - 348 | α | 2.53 | 133 | 139 | 30 |

Table 3 cont.: Hydrogen bonding in the helices of dicupric human lactoferrin

| helix 2 cont. | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|
| 46 - 50 | α | 2.17 | 140 | 155 | 17 |
| 47 - 51 | α | 2.08 | 138 | 124 | 39 |
| 48 - 51 | $3_{10}$ | 2.06 | 106 | 136 | 31 |
| 48 - 53 | π | 2.11 | 132 | 154 | 18 |
| 49 - 52 | $3_{10}$ | 2.36 | 107 | 146 | 24 |
| average values | | 1.91 | 148 | 155 | 17 |
| **helix 3** | | | | | |
| 60 - 63 | $3_{10}$ | 2.04 | 108 | 132 | 33 |
| 60 - 64 | α | 1.87 | 154 | 147 | 22 |
| 61 - 64 | $3_{10}$ | 2.41 | 91 | 117 | 46 |
| 61 - 65 | α | 2.16 | 149 | 159 | 15 |
| 62 - 66 | α | 2.35 | 136 | 172 | 6 |
| 63 - 67 | α | 1.71 | 140 | 154 | 17 |
| 64 - 68 | α | 2.10 | 137 | 130 | 34 |
| 65 - 68 | $3_{10}$ | 2.14 | 106 | 133 | 33 |
| * 65 - 69 | α | 2.53 | 152 | 143 | 27 |
| 66 - 69 | $3_{10}$ | 2.43 | 106 | 134 | 33 |
| average values | | 2.04 | 145 | 159 | 14 |
| **helix 4** | | | | | |
| 105 - 108 | $3_{10}$ | 2.00 | 116 | 158 | 14 |

| helix 2 cont. | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|
| 381 - 384 | $3_{10}$ | 2.41 | 98 | 128 | 37 |
| 381 - 385 | α | 2.44 | 135 | 134 | 33 |
| 382 - 385 | $3_{10}$ | 2.03 | 106 | 127 | 36 |
| 382 - 386 | α | 2.06 | 157 | 165 | 10 |
| 383 - 388 | π | 2.17 | 164 | 120 | 42 |
| average values | | 2.13 | 143 | 152 | 19 |
| **helix 3** | | | | | |
| 395 - 398 | $3_{10}$ | 2.49 | 117 | 127 | 39 |
| 395 - 399 | α | 2.39 | 143 | 141 | 27 |
| 396 - 399 | $3_{10}$ | 2.37 | 91 | 123 | 41 |
| 396 - 400 | α | 2.50 | 137 | 131 | 35 |
| 397 - 400 | $3_{10}$ | 2.50 | 94 | 136 | 32 |
| * 397 - 401 | α | 2.52 | 137 | 159 | 15 |
| 398 - 402 | α | 1.85 | 145 | 158 | 14 |
| 399 - 403 | α | 2.36 | 111 | 125 | 40 |
| 400 - 404 | α | 1.84 | 150 | 155 | 16 |
| 401 - 405 | α | 2.04 | 134 | 144 | 24 |
| 402 - 407 | π | 1.74 | 149 | 167 | 9 |
| 403 - 406 | $3_{10}$ | 1.97 | 98 | 149 | 21 |
| average values | | 2.16 | 137 | 142 | 26 |
| **helix 4** | | | | | |
| 449 - 452 | $3_{10}$ | 1.88 | 112 | 155 | 16 |
| 450 - 453 | $3_{10}$ | 2.45 | 117 | 167 | 9 |

Table 3 cont.: Hydrogen bonding in the helices of dicupric human lactoferrin

| helix 5 | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|
| 121 -125 | α | 1.96 | 122 | 142 | 25 |
| 121 - 126 | π | 1.82 | 157 | 158 | 14 |
| 122 - 127 | π | 2.11 | 175 | 160 | 14 |
| 125 - 129 | α | 1.96 | 162 | 160 | 13 |
| 126 - 130 | α | 2.28 | 125 | 143 | 26 |
| 127 - 131 | α | 1.71 | 134 | 139 | 26 |
| 128 - 132 | α | 2.24 | 140 | 141 | 28 |
| * 129 - 132 | $3_{10}$ | 2.52 | 97 | 125 | 49 |
| * 129 - 133 | α | 2.53 | 143 | 114 | 49 |
| 132 - 135 | $3_{10}$ | 2.10 | 131 | 131 | 33 |
| 133 - 136 | $3_{10}$ | 2.24 | 107 | 142 | 26 |
| average values | | 2.03 | 137 | 145 | 24 |
| **helix 6** | | | | | |
| 144 - 147 | $3_{10}$ | 1.99 | 121 | 138 | 29 |
| 145 - 148 | $3_{10}$ | 2.10 | 101 | 119 | 43 |
| 145 - 149 | α | 2.15 | 150 | 140 | 28 |
| 146 - 149 | $3_{10}$ | 2.33 | 95 | 126 | 38 |
| 146 - 150 | α | 1.85 | 160 | 169 | 7 |
| 147 - 151 | α | 2.51 | 133 | 135 | 32 |
| 149 - 152 | $3_{10}$ | 2.01 | 98 | 142 | 25 |
| 149 - 153 | α | 2.07 | 171 | 168 | 8 |
| average values | | 2.15 | 154 | 153 | 19 |

| helix 5 | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|
| 465 - 469 | α | 2.32 | 134 | 129 | 36 |
| 465 - 470 | π | 2.12 | 150 | 163 | 12 |
| * 466 - 469 | $3_{10}$ | 2.60 | 101 | 143 | 27 |
| 466 - 471 | π | 2.17 | 158 | 166 | 10 |
| 469 - 473 | α | 2.06 | 159 | 157 | 15 |
| 470 - 474 | α | 2.16 | 106 | 132 | 33 |
| 471 - 475 | α | 2.25 | 152 | 153 | 18 |
| 472 - 476 | α | 2.34 | 139 | 157 | 16 |
| 473 - 477 | α | 2.04 | 137 | 148 | 22 |
| 474 - 478 | α | 2.24 | 124 | 137 | 30 |
| 476 - 479 | $3_{10}$ | 2.05 | 106 | 148 | 22 |
| 476 - 480 | α | 2.27 | 174 | 159 | 15 |
| 477 - 481 | α | 2.05 | 125 | 127 | 36 |
| average values | | 2.19 | 139 | 143 | 26 |
| **helix 6** | | | | | |
| 483 - 485 | 1-3 | 2.08 | 103 | 140 | 28 |
| 484 - 487 | $3_{10}$ | 2.02 | 121 | 129 | 35 |
| * 485 - 488 . | $3_{10}$ | 2.66 | 96 | 164 | 11 |

Table 3 cont.: Hydrogen bonding in the helices of dicupric human lactoferrin

| helix 7 | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|
| 190 - 194 | α | 1.85 | 150 | 129 | 34 |
| 191 - 195 | α | 1.85 | 152 | 165 | 10 |
| 192 - 196 | α | 2.31 | 123 | 141 | 28 |
| 193 - 197 | α | 2.31 | 147 | 151 | 20 |
| 194 - 197 | $3_{10}$ | 2.47 | 98 | 117 | 46 |
| 194 - 198 | α | 1.74 | 160 | 175 | 3 |
| 195 - 199 | α | 2.32 | 133 | 153 | 19 |
| 196 - 200 | α | 1.62 | 146 | 171 | 20 |
| 197 - 201 | α | 1.76 | 158 | 166 | 9 |
| 198 - 203 | π | 2.27 | 165 | 122 | 41 |
| 199 - 202 | $3_{10}$ | 1.93 | 112 | 144 | 24 |
| average value | | 1.97 | 146 | 156 | 18 |
| **helix 8** | | | | | |
| 212 - 215 | $3_{10}$ | 2.16 | 114 | 125 | 39 |
| 212 - 216 | α | 1.71 | 157 | 163 | 11 |
| 213 - 217 | α | 1.81 | 147 | 151 | 18 |
| 214 - 218 | α | 2.27 | 129 | 148 | 22 |
| average values | | 1.93 | 144 | 154 | 17 |
| **helix 8a** | | | | | |
| * 223 - 226 | $3_{10}$ | 2.72 | 118 | 127 | 39 |
| 224 - 227 | $3_{10}$ | 2.45 | 112 | 160 | 14 |
| **helix 9** | | | | | |
| 263 - 267 | α | 1.82 | 160 | 168 | 8 |
| 264 - 268 | α | 2.06 | 130 | 145 | 24 |
| 265 - 269 | α | 1.90 | 154 | 168 | 8 |

| helix 7 | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|
| * 526 - 530 | α | 2.56 | 132 | 119 | 45 |
| 527 - 531 | α | 1.69 | 157 | 164 | 10 |
| 528 - 532 | α | 2.17 | 153 | 170 | 7 |
| 529 - 533 | α | 1.73 | 143 | 166 | 9 |
| 530 - 534 | α | 2.01 | 144 | 150 | 20 |
| 531 - 535 | α | 2.47 | 145 | 147 | 24 |
| 532 - 535 | $3_{10}$ | 2.27 | 112 | 123 | 40 |
| 532 - 536 | α | 2.09 | 169 | 167 | 9 |
| 533 - 537 | α | 2.27 | 132 | 125 | 39 |
| 534 - 539 | π | 1.91 | 174 | 142 | 25 |
| 535 - 538 | $3_{10}$ | 1.75 | 114 | 153 | 18 |
| average values | | 2.06 | 149 | 156 | 17 |
| **helix 8** | | | | | |
| 547 - 550 | $3_{10}$ | 1.88 | 131 | 147 | 21 |
| 548 - 552 | α | 2.01 | 166 | 153 | 18 |
| 549 - 553 | α | 2.12 | 143 | 161 | 13 |
| 550 - 553 | $3_{10}$ | 1.92 | 101 | 143 | 24 |
| 551 - 554 | $3_{10}$ | 2.06 | 133 | 178 | 2 |
| average values | | 2.07 | 155 | 157 | 16 |
| **helix 8a** | | | | | |
| 567 - 570 | $3_{10}$ | 1.96 | 138 | 162 | 12 |
| 568 - 571 | $3_{10}$ | 2.30 | 127 | 160 | 14 |
| **helix 9** | | | | | |
| 605 - 609 | α | 2.13 | 152 | 146 | 23 |
| 606 - 610 | α | 2.01 | 140 | 157 | 16 |
| 607 - 611 | α | 2.13 | 145 | 159 | 15 |

Table 3 cont.: Hydrogen bonding in the helices of dicupric human lactoferrin

| helix 9 cont. | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) | helix 9 cont. | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 266 - 270 | α | 1.99 | 170 | 163 | 11 | 608 - 612 | α | 2.10 | 146 | 168 | 8 |
| 267 - 271 | α | 2.03 | 137 | 159 | 14 | 609 - 613 | α | 1.96 | 157 | 165 | 10 |
| 268 - 272 | α | 1.66 | 161 | 165 | 9 | 610 - 614 | α | 1.98 | 147 | 151 | 19 |
| 269 - 273 | α | 1.84 | 167 | 160 | 13 | 611 - 615 | α | 1.91 | 141 | 158 | 14 |
| * 270 - 273 | $3_{10}$ | 2.56 | 102 | 123 | 42 | 612 - 616 | α | 1.92 | 134 | 156 | 16 |
| 270 - 274 | α | 2.23 | 170 | 172 | 6 | * 613 - 617 | α | 2.54 | 129 | 113 | 49 |
| 271 - 275 | α | 2.41 | 130 | 137 | 30 | 614 - 617 | $3_{10}$ | 2.42 | 98 | 152 | 20 |
| 273 - 276 | $3_{10}$ | 1.72 | 100 | 128 | 34 | 614 - 618 | α | 2.50 | 140 | 148 | 23 |
| 273 - 277 | α | 1.95 | 163 | 163 | 11 | 615 - 618 | $3_{10}$ | 2.35 | 111 | 125 | 40 |
| 274 - 278 | α | 1.89 | 135 | 125 | 37 | 616 - 620 | α | 2.19 | 116 | 154 | 18 |
| 275 - 279 | α | 2.34 | 109 | 113 | 48 | 617 - 621 | α | 1.61 | 131 | 128 | 32 |
| average values | | 2.01 | 149 | 153 | 18 | average values | | 2.04 | 141 | 154 | 18 |
| helix 10 | | | | | | helix 10 | | | | | |
| 315 - 319 | α | 1.65 | 157 | 153 | 17 | 658 - 662 | α | 1.73 | 167 | 159 | 14 |
| 316 - 320 | α | 1.61 | 159 | 159 | 13 | 659 - 663 | α | 1.68 | 158 | 167 | 8 |
| 317 - 321 | α | 2.36 | 140 | 133 | 33 | 660 - 664 | α | 1.97 | 120 | 127 | 36 |
| 318 - 321 | $3_{10}$ | 2.42 | 91 | 130 | 36 | | | | | | |
| average values | | 1.87 | 152 | 148 | 21 | average values | | 1.79 | 148 | 151 | 19 |
| helix 11 | | | | | | helix 11 | | | | | |
| 321 - 325 | α | 1.97 | 144 | 141 | 26 | 664 - 668 | α | 1.98 | 156 | 167 | 8 |
| 322 - 326 | α | 2.17 | 133 | 148 | 22 | 665 - 669 | α | 1.85 | 142 | 146 | 22 |
| 323 - 327 | α | 2.38 | 125 | 126 | 39 | 666 - 670 | α | 2.08 | 134 | 149 | 21 |
| 324 - 327 | $3_{10}$ | 2.02 | 96 | 114 | 45 | 667 - 671 | α | 2.04 | 153 | 147 | 22 |
| 324 - 328 | α | 2.11 | 159 | 135 | 31 | 668 - 671 | $3_{10}$ | 2.47 | 91 | 118 | 46 |
| 325 - 328 | $3_{10}$ | 2.49 | 91 | 135 | 33 | 668 - 672 | α | 2.08 | 155 | 172 | 5 |
| 325 - 329 | α | 2.19 | 148 | 167 | 9 | 669 - 673 | α | 2.06 | 152 | 165 | 10 |

Table 3 cont.: Hydrogen bonding in the helices of dicupric human lactoferrin

| helix 11 cont. | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) | helix 11 cont. | hydrogen bond | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | H-N···O (deg) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 326 - 330 | α | 1.84 | 149 | 172 | 5 | 670 - 674 | α | 2.09 | 147 | 172 | 6 |
| 328 - 331 | $3_{10}$ | 2.31 | 100 | 155 | 18 | 671 - 675 | α | 1.76 | 138 | 134 | 30 |
| 328 - 332 | α | 2.45 | 144 | 131 | 36 | 672 - 676 | α | 2.27 | 153 | 139 | 29 |
| 329 - 332 | $3_{10}$ | 2.03 | 138 | 167 | 9 | 674 - 677 | $3_{10}$ | 2.34 | 108 | 147 | 23 |
| | | | | | | 675 - 678 | $3_{10}$ | 2.35 | 104 | 137 | 31 |
| average values | | 2.16 | 143 | 146 | 24 | average values | | 2.02 | 148 | 155 | 17 |
| helix 12 | | | | | | helix 12 | | | | | |
| 334 - 338 | α | 1.95 | 149 | 166 | 9 | 680 - 683 | $3_{10}$ | 2.37 | 100 | 123 | 41 |
| 336 - 340 | α | 1.89 | 143 | 145 | 23 | 680 - 684 | α | 1.94 | 146 | 156 | 16 |
| 337 - 340 | $3_{10}$ | 2.34 | 104 | 119 | 43 | 681 - 685 | α | 1.63 | 146 | 174 | 4 |
| 337 - 341 | α | 2.13 | 176 | 175 | 4 | 682 - 686 | α | 1.91 | 131 | 143 | 25 |
| 338 - 342 | α | 2.51 | 137 | 145 | 25 | 683 - 687 | α | 2.41 | 130 | 125 | 40 |
| 339 - 343 | α | 2.22 | 142 | 157 | 16 | 684 - 687 | $3_{10}$ | 2.16 | 111 | 143 | 25 |
| 340 - 343 | $3_{10}$ | 2.43 | 108 | 118 | 45 | 684 - 688 | α | 2.41 | 161 | 171 | 6 |
| 340 - 344 | α | 2.35 | 160 | 151 | 21 | 685 - 689 | α | 2.06 | 152 | 152 | 19 |
| | | | | | | 687 - 690 | $3_{10}$ | 2.03 | 99 | 150 | 20 |
| | | | | | | 687 - 691 | α | 2.37 | 153 | 149 | 22 |
| average values | | 2.18 | 151 | 157 | 16 | average values | | 2.10 | 146 | 153 | 19 |

[a] Hydrogen bond types: π = 1-6, α = 1-5, $3_{10}$ = 1-4.

Table 4: Hydrogen bonding in the β-strands of dicupric human lactoferrin [a]

| N-terminal lobe | | | | | | C-terminal lobe | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| strand b - a | O···H N (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) | strand b - a | O···H N (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) |
| 6N - 33O | 1.84 | 155 | 165 | 2.82 | 10 | 345N - 368O | 1.83 | 153 | 166 | 2.81 | 9 |
| 6O - 35N | 1.86 | 130 | 147 | 2.75 | 22 | 345O - 370N | 2.08 | 161 | 173 | 3.08 | 5 |
| 8N - 35O | 2.07 | 168 | 137 | 2.88 | 30 | 347N - 370O | 2.10 | 151 | 175 | 3.10 | 3 |
| 8O - 37N | 1.98 | 161 | 161 | 2.95 | 12 | 347O - 372N | 1.78 | 159 | 134 | 2.58 | 30 |
| 10N - 37O | 1.75 | 155 | 175 | 2.75 | 3 | 349N - 372O | 2.03 | 124 | 156 | 2.97 | 16 |
| * 10O - 39N | 2.64 | 120 | 115 | 3.19 | 49 | 349O - 374N | 2.42 | 163 | 151 | 3.33 | 21 |
| | | | | | | 351N - 374O | 2.03 | 163 | 160 | 2.99 | 14 |
| strand a - c | | | | | | strand a - c | | | | | |
| 7O - 55N | 1.68 | 145 | 167 | 2.67 | 8 | 346O - 390N | 1.53 | 147 | 158 | 2.49 | 13 |
| 7O - 56N | 2.42 | 146 | 149 | 3.32 | 22 | 346O - 391N | 2.34 | 151 | 146 | 3.22 | 24 |
| 9N - 56O | 2.32 | 153 | 128 | 3.04 | 37 | 348N - 391O | 1.78 | 171 | 157 | 2.74 | 15 |
| strand c - j | | | | | | strand c - j | | | | | |
| 57O - 256N | 2.06 | 152 | 152 | 2.98 | 19 | 392O - 600N | 2.07 | 177 | 159 | 3.03 | 14 |
| 59O - 254N | 2.36 | 170 | 157 | 3.30 | 16 | 394O - 598N | 1.80 | 152 | 150 | 2.72 | 19 |
| 59N - 254O | 1.89 | 160 | 159 | 2.85 | 14 | 394N - 598O | 1.63 | 145 | 169 | 2.62 | 7 |
| strand d - j | | | | | | strand d - j | | | | | |
| 75O - 257N | 1.86 | 150 | 156 | 2.80 | 16 | 408O - 601N | 2.25 | 155 | 157 | 3.20 | 16 |
| 75N - 257O | 1.84 | 169 | 177 | 2.84 | 2 | 408N - 601O | 2.01 | 169 | 160 | 2.96 | 13 |
| 77N - 255O | 1.90 | 150 | 169 | 2.88 | 7 | 410N - 599O | 1.53 | 169 | 156 | 2.47 | 15 |
| 78O - 255N | 1.66 | 160 | 165 | 2.64 | 9 | 411O - 599N | 1.73 | 168 | 175 | 2.73 | 3 |
| 78N - 255O | 2.45 | 138 | 171 | 3.44 | 7 | | | | | | |
| 80N - 253O | 1.94 | 153 | 168 | 2.93 | 8 | 413N - 597O | 1.66 | 158 | 175 | 2.66 | 3 |

Table 4 continued: Hydrogen bonding in the β-strands of dicupric human lactoferrin

| strand d - k | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) | strand d - k | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 77O - 310N | 1.99 | 122 | 168 | 2.97 | 8 | 410O - 653N | 1.95 | 126 | 170 | 2.94 | 7 |
| 79O - 308N | 1.97 | 150 | 145 | 2.85 | 24 | 412O - 651N | 2.09 | 157 | 159 | 3.05 | 14 |
| 79N - 308O | 1.83 | 153 | 172 | 2.82 | 5 | 412N - 651O | 2.09 | 146 | 161 | 3.05 | 13 |
| 81O - 305N | 2.09 | 131 | 154 | 3.03 | 18 | 414O - 648N | 1.63 | 144 | 144 | 2.51 | 22 |
| | | | | | | * 414O - 649N | 2.58 | 111 | 151 | 3.49 | 21 |
| 81N - 306O | 1.87 | 160 | 138 | 2.69 | 28 | 414N - 649O | 1.74 | 155 | 167 | 2.72 | 8 |
| strand e - j | | | | | | strand e - j | | | | | |
| 90O - 252N | 2.15 | 123 | 158 | 3.10 | 15 | | | | | | |
| 92O - 250N | 1.92 | 142 | 152 | 2.84 | 18 | 435O - 594N | 1.70 | 141 | 154 | 2.63 | 16 |
| 92N - 250O | 1.88 | 157 | 163 | 2.85 | 11 | 435N - 594O | 1.96 | 134 | 136 | 2.76 | 29 |
| 94O - 247N | 1.83 | 153 | 149 | 2.74 | 20 | 437O - 591N | 2.26 | 140 | 122 | 2.92 | 41 |
| 94O - 248N | 2.41 | 139 | 162 | 3.38 | 13 | * 437O - 592N | 2.56 | 126 | 158 | 3.50 | 16 |
| 94N - 248O | 2.02 | 128 | 148 | 2.92 | 22 | 437N - 592O | 1.63 | 175 | 168 | 2.62 | 7 |
| strand e - i | | | | | | strand e - i | | | | | |
| 96O - 230N | 2.22 | 165 | 155 | 3.15 | 17 | 439O - 574N | 2.50 | 151 | 145 | 3.37 | 25 |
| 98O - 228N | 2.04 | 158 | 162 | 3.00 | 12 | 441O - 572N | 1.73 | 155 | 156 | 2.68 | 15 |
| 98N - 228O | 2.07 | 163 | 169 | 3.06 | 7 | 441N - 572O | 1.95 | 161 | 156 | 2.89 | 16 |
| 100N - 226O | 1.58 | 146 | 164 | 2.56 | 10 | 443N - 570O | 1.55 | 123 | 160 | 2.51 | 12 |
| strand e - h | | | | | | strand e - h | | | | | |
| * 93O - 211N | 2.54 | 142 | 115 | 3.09 | 48 | | | | | | |
| 95O - 209N | 2.40 | 144 | 130 | 3.13 | 36 | 438O - 545N | 2.48 | 151 | 151 | 3.39 | 21 |
| 95N - 209O | 1.72 | 162 | 159 | 2.68 | 13 | 438N - 545O | 2.09 | 163 | 170 | 3.08 | 7 |
| 97O - 207N | 1.53 | 145 | 166 | 2.51 | 9 | 440O - 543N | 2.35 | 133 | 134 | 3.14 | 32 |
| 97N - 207O | 1.76 | 161 | 165 | 2.74 | 10 | 440N - 543O | 1.91 | 145 | 159 | 2.86 | 14 |

Table 4 continued: Hydrogen bonding in the β-strands of dicupric human lactoferrin

| strand f - g | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) | strand f - g | O···HN (Å) | C-O···H (deg) | N-H···O (deg) | O···N (Å) | H-N···O (deg) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 112O - 154N | 2.35 | 133 | 137 | 3.15 | 31 | 456O - 490N | 2.20 | 130 | 128 | 2.93 | 36 |
| 114O - 157N | 2.02 | 156 | 156 | 2.96 | 16 | | | | | | |
| 114N - 155O | 2.50 | 142 | 154 | 3.43 | 19 | 458N - 491O | 2.43 | 140 | 134 | 3.21 | 33 |
| 116N - 157O | 2.16 | 168 | 158 | 3.11 | 15 | 460N - 493O | 1.78 | 179 | 158 | 2.73 | 14 |
| strand f - h | | | | | | strand f - h | | | | | |
| 113O - 205N | 1.75 | 123 | 150 | 2.66 | 19 | 457O - 541N | 1.50 | 141 | 161 | 2.47 | 11 |
| * 113O - 206N | 2.64 | 131 | 128 | 3.36 | 38 | 457O - 542N | 2.29 | 146 | 147 | 3.17 | 23 |
| 115O - 208N | 1.87 | 171 | 162 | 2.84 | 11 | 459O - 544N | 1.82 | 161 | 164 | 2.79 | 10 |
| 115N - 206O | 2.07 | 159 | 149 | 2.97 | 21 | 459N - 542O | 2.35 | 124 | 115 | 2.92 | 47 |

[a] The majority of the hydrogen bonds in the two lobes are in equivalent positions and are given on the same line. Where there is no equivalent hydrogen bond in one of the lobes, a gap has been left. The asterisks indicate interactions with unfavourable distances and/or geometries.

Table 5: Turns in Cu2Lf (β- and γ-type turns) [a]

| N-lobe turn | start | end | type | C-lobe turn | start | end | type |
|---|---|---|---|---|---|---|---|
| 1 | Leu69 | Tyr72 | VI | 24 | Ser365 | Ser368 | I' |
| 2 | Thr84 | Gln87 | I | 25 | Asp424 | Cys427 | I |
| 3 | Leu109 | Leu112 | II | 26 | Asn426 | Asp429 | III |
| 4 | Gly118 | Arg121 | II | 27 | Cys427 | Arg430 | I |
| 5 | Gly140 | Glu143 | VI | 28 | Arg442 | Asp445 | III |
| 6 | Val158 | Ala161 | II | 29 | Ser444 | Thr446 | γ-turn |
| 7 | Asp162 | Phe166 | α-turn | 30 | Asp445 | Leu448 | I |
| 8 | Phe166 | Leu169 | III | 31 | Val453 | Lys456 | II |
| 9 | Pro167 | Cys170 | III | 32 | Ala462 | Arg465 | II |
| 10 | Cys170 | Cys173 | I | 33 | Ala494 | Ser497 | II |
| 11 | Gly177 | Lys180 | I | 34 | Asp498 | Ser501 | I |
| 12 | Asn179 | Ala182 | II | 35 | Ser501 | Cys504 | III |
| 13 | Leu184 | Glu187 | III | 36 | Asn502 | Ala505 | I |
| 14 | Glu187 | Phe190 | I | 37 | Cys504 | Cys507 | I |
| 15 | Cys231 | Asn234 | I | 38 | Asp510 | Gly513 | I |
| 16 | Pro238 | Lys241 | I | 39 | Asn515 | Val518 | II |
| 17 | Lys241 | Asp244 | III | 40 | Asn520 | Glu523 | III |
| 18 | Phe242 | Cys245 | I | 41 | Glu523 | Tyr526 | I |
| 19 | Gln287 | Gly290 | I | 42 | Thr554 | Asn557 | I' |
| 20 | Pro292 | Gln295 | I | 43 | Glu560 | Ala563 | III |
| 21 | Leu298 | Phe300 | γ-turn | 44 | Ala563 | Leu566 | I |
| 22 | Lys301 | Ala304 | II | 45 | Cys575 | Gly578 | I |
| 23 | Pro311 | Ile314 | I | 46 | Pro582 | Glu585 | III |
| | | | | 47 | Glu585 | Ser588 | III |
| | | | | 48 | Ala586 | Cys589 | I |
| | | | | 49 | Arg602 | Lys605 | III |
| | | | | 50 | Gly621 | Gly624 | I |
| | | | | 51 | Gly624 | Cys627 | I |
| | | | | 52 | Asp626 | Lys630 | α-turn |
| | | | | 53 | Cys632 | Gln635 | I |
| | | | | 54 | Ser636 | Lys639 | III |
| | | | | 55 | Leu641 | Phe643 | γ-turn |
| | | | | 56 | Asn644 | Thr647 | I |
| | | | | 57 | Leu653 | Lys656 | I |

[a] β-turns have hydrogen bonds linking residue i to residue i+3 and γ-turns link residue i to residue i + 2. The presence of α-turns, linking i toi+4 are also noted.

## Appendix V

*Hydrogen bonding at the metal and anion sites in $Cu_2Lf$*

Table 1: Hydrogen bonds in the vicinity of the N-lobe metal and anion binding sites (not including hydrogen bonds which specifically involve the anion)

A: Hydrogen bonds involving metal and anion binding residues

| | O···HN (Å) | O···N ($O_1$···$O_2$) (Å) [a] | C-O···H ($C_1$-$O_1$···$O_2$) (deg) [a] | N-H···O ($C_2$-$O_2$···$O_1$) (deg) [a] | H-N···O (deg) | comments |
|---|---|---|---|---|---|---|
| 60 N - 722 $O_{w1}$ | 2.15 | 3.09 | — | 155 | 17 | |
| 60 $O_{\delta1}$ - 700 $O_{w1}$ | — | 3.54 | 117 | — | — | |
| 60 $O_{\delta2}$ - 62 N | 1.88 | 2.82 | 121 | 155 | 16 | |
| 60 $O_{\delta2}$ - 122 N | 2.21 | 3.05 | 108 | 141 | 27 | inter-domain |
| 60 $O_{\delta2}$ - 122 $O_{\gamma1}$ | — | 3.02 | 103 | 113 | — | inter-domain |
| 92 N - 250 O | 1.87 | 2.84 | 159 | 164 | 11 | β (strands e - j) |
| 92 O - 250 N | 1.87 | 2.79 | 145 | 152 | 18 | β (strands e - j) |
| 92 $O_\eta$ - 700 $O_{w1}$ | — | 2.67 | 107 | — | — | |
| 117 O - 191 N | 2.29 | 3.12 | 148 | 140 | 28 | |
| 121 O - 125 N | 1.98 | 2.88 | 123 | 148 | 22 | α (helix 5) |
| 121 O - 126 N | 1.89 | 2.81 | 160 | 152 | 18 | π (helix 5) |
| 121 $N_{\eta1}$ - 191 $O_\gamma$ | 2.06 | 2.92 | 111 | 143 | 25 | |
| 121 $N_{\eta1}$ - 715 $O_{w1}$ | 2.00 | 2.71 | — | 126 | 37 | |
| 121 $N_{\eta2}$ - 191 $O_\gamma$ | 2.39 | 3.16 | 145 | 133 | 33 | |
| 123 O - 700 Ow9 | — | 2.55 | 134 | — | — | |
| 124 O - 116 $N_{\delta1}$ | 2.54 | 3.15 | 143 | 120 | 45 | |
| 192 N - 729 $O_{w9}$ | 2.69 | 3.06 | — | 102 | 59 | poor geometry at N |
| 192 O - 196 N | 2.23 | 3.10 | 126 | 144 | 25 | α (helix 7) |
| 192 $O_\eta$ - 210 $N_{\eta2}$ | 1.95 | 2.88 | 102 | 154 | 17 | |
| 253 N - 82 O | 2.32 | 3.25 | 127 | 154 | 4 | β (strands d - j) |
| 253 O - 80 O | 1.89 | 2.87 | 154 | 166 | 9 | β (strands d - j) |
| 253 $N_{\delta1}$ - 700 $O_{w2}$ | 1.60 | 2.60 | — | 176 | 3 | |

B: Hydrogen bonds involving nearby residues

| | | | | | | |
|---|---|---|---|---|---|---|
| 82 $O_\eta$ - 80 $O_{\varepsilon2}$ | — | 2.69 | 117 | 142 | — | |
| 210 $N_{\eta1}$ - 702 $O_{w8}$ | 2.14 | 2.94 | — | 136 | 30 | |
| 210 $N_{\eta2}$ - 82 $O_\eta$ | 2.36 | 3.35 | 96 | 175 | 4 | inter-domain |

[a] For hydrogen bonds other than NH···O, $C_1$ and $O_1$ relate to the first-named residue and $C_2$ and $O_2$ are for the second-named residue (protein or solvent).

Table 2: Hydrogen bonds in the vicinity of the C-lobe metal and anion binding sites (not including hydrogen bonds which specifically involve the anion)

A: Hydrogen bonds involving metal and anion binding residues

| | O···HN (Å) | O···N ($O_1$...$O_2$) (Å) [a] | C-O···H ($C_1$-$O_1$...$O_2$) (deg) [a] | N-H···O ($C_2$-$O_2$...$O_1$) (deg) [a] | H-N···O (deg) | comments |
|---|---|---|---|---|---|---|
| 395 N - 702 $O_{w0}$ | 1.73 | 2.53 | — | 134 | 29 | |
| 395 $O_{\delta1}$ - 712 $O_{w5}$ | — | 3.15 | 100 | — | — | |
| 395 $O_{\delta2}$ - 397 N | 2.11 | 3.07 | 129 | 159 | 14 | |
| 395 $O_{\delta2}$ - 466 N | 1.83 | 2.75 | 114 | 151 | 19 | inter-domain |
| 395 $O_{\delta2}$ - 466 $O_{\gamma1}$ | — | 3.36 | 112 | 105 | — | inter-domain |
| 435 N - 594 O | 1.98 | 2.79 | 133 | 136 | 29 | β (strands e - j) |
| 435 O - 594 N | 1.75 | 2.67 | 140 | 151 | 19 | β (strands e - j) |
| 435 $O_{\eta}$ - 712 $O_{w5}$ | — | 2.66 | 108 | — | — | |
| 461 O - 527 N | 2.08 | 3.00 | 129 | 152 | 19 | |
| 461 $O_{\gamma1}$ - 462 N | 2.32 | 2.87 | 94 | 113 | 48 | poor geometry at N |
| 465 N - 462 O | 2.43 | 3.40 | 130 | 162 | 13 | |
| 465 O - 469 N | 2.08 | 2.85 | 141 | 132 | 33 | α (helix 5) |
| 465 O - 470 N | 2.02 | 2.96 | 144 | 156 | 16 | π (helix 5) |
| 465 $N_{\eta1}$ - 709 $O_{w9}$ | 2.04 | 3.04 | — | 175 | 3 | |
| 465 $N_{\eta1}$ - 710 $O_{w0}$ | 2.63 | 3.53 | — | 143 | 28 | long bond |
| 467 O - 710 $O_{w5}$ | — | 2.91 | 116 | — | — | |
| 468 O - 460 $N_{\delta1}$ | 1.72 | 2.64 | 168 | 152 | 18 | |
| 528 N - 709 $O_{w9}$ | 2.56 | 3.52 | — | 160 | 14 | long bond |
| 528 O - 532 N | 2.06 | 3.06 | 150 | 172 | 5 | π (helix 7) |
| 528 $O_{\eta}$ - 700 $O_{w0}$ | — | 2.74 | 116 | — | — | |
| 597 N - 415 $O_{\eta}$ | 2.47 | 3.42 | 136 | 159 | 15 | |
| 597 O - 413 N | 1.56 | 2.56 | 156 | 171 | 5 | β (strands d - j) |
| 597 $N_{\delta1}$ - 713 $O_{w5}$ | 2.01 | 2.99 | — | 166 | 9 | |

B: Hydrogen bonds involving nearby residues

| | | | | | | |
|---|---|---|---|---|---|---|
| 415 O - 430 $N_{\eta1}$ | 2.53 | 3.42 | 158 | 148 | 23 | bifurcated |
| * 415 O - 430 $N_{\eta2}$ | 2.51 | 3.41 | 104 | 150 | 22 | bifurcated |
| 415 $O_{\eta}$ - 413 $O_{\varepsilon2}$ | — | 2.70 | 108 | 159 | — | |
| 415 $O_{\eta}$ - 644 $O_{\delta1}$ | — | 3.15 | 109 | 144 | — | inter-domain |
| 415 $O_{\eta}$ - 700 $O_{w0}$ | — | 3.23 | 108 | — | — | |
| * 546 N - 549 $O_{\gamma1}$ | 2.51 | 3.42 | 124 | 151 | 21 | |
| 546 O - 549 N | 2.41 | 3.30 | 103 | 147 | 24 | α (helix 8) |
| 546 $N_{\zeta}$ - 700 $O_{w0}$ | — | 2.86 | 94 [b] | — | — | |

Hydrogen bonds labelled with an asterisk (*) indicate interactions which are not present in $Fe_2$Lf. [a] For hydrogen bonds other than NH···O, $C_1$ and $O_1$ relate to the first-named residue and $C_2$ and $O_2$ are for the second-named residue (protein or solvent). [b] The $C_{\varepsilon}$-$N_{\zeta}$...$O_w$ angle.

Table 3: Interdomain hydrogen bonds and interactions with the two cross-over β-strands (e and j) in the N-lobe of $Cu_2Lf$

| | O···HN (Å) | O···N ($O_1$···$O_2$) (Å) [a] | comments [b] | other structures |
|---|---|---|---|---|
| **A: Inter-domain contacts** | | | | |
| 60 $O_{\delta 2}$ - 122 N | 2.21 | 3.05 | cleft | Fe |
| 60 $O_{\delta 2}$ - 122 $O_{\gamma 1}$ | — | 3.02 | cleft | Fe |
| 82 $O_{\eta}$ - 210 $N_{\eta 2}$ | 2.36 | 3.35 | cleft | Fe |
| Glu216 - Lys301 | — | — | salt bridge - cleft | Fe |
| Asp217 - Lys296 | — | — | salt bridge - mouth of cleft | Fe |
| **B: Contacts involving strands e and j in the cross-over region** | | | | |
| 90 O - 252 N | 2.08 | 3.03 | strand e - strand j | Fe / apo |
| 93 N - 211 $O_{\varepsilon 1}$ | 2.30 | 3.27 | strand e - strand h (s/c - N2) | Fe / apo |
| 93 O - 211 $O_{\varepsilon 1}$ | — | 3.37 | strand e - strand h (s/c - N2) | Fe |
| 94 N - 248 O | 1.96 | 2.85 | strand e - domain N2 | Fe / apo |
| 94 O - 247 N | 1.86 | 2.80 | strand e - domain N2 | Fe / apo |
| 95 N - 209 O | 2.42 | 3.13 | strand e - strand h (N2) | Fe / apo |
| 95 O - 209 N | 1.64 | 2.59 | strand e - strand h (N2) | Fe / apo |
| 59 N - 254 O | 1.91 | 2.89 | strand c (N1) - strand j | Fe / apo |
| 59 O - 254 N | 2.30 | 3.24 | strand c (N1) - strand j | Fe / apo |
| 77 N - 255 O | 1.81 | 2.80 | strand d (N1) - strand j | Fe / apo |
| 78 N - 255 O | 2.43 | 3.41 | strand d (N1) - strand j | Fe / apo |
| 78 O - 255 N | 1.68 | 2.66 | strand d (N1) - strand j | Fe / apo |
| 80 O - 252 $O_{\gamma}$ | — | 2.84 | domain N1 - strand j (s/c) | Fe |
| 252 O - 319 $O_{\eta}$ | — | 2.58 | strand j - domain N1 | Fe |
| 252 $O_{\gamma}$ - 319 $O_{\eta}$ | — | 3.50 | strand j - domain N1 | Fe |

[a] For hydrogen bonds other than NH···O, $O_1$ and $O_2$ relate to the first- and second-named residues. [b] s/c denotes sidechain contacts.

Table 4: Interdomain hydrogen bonds and interactions with the two cross-over β-strands (e and j) in the C-lobe of $Cu_2Lf$

| | O···HN (Å) | O···N ($O_1$···$O_2$) (Å) [a] | comments [b] | other structures |
|---|---|---|---|---|
| **A: Inter-domain contacts** | | | | |
| 352 $O_{\varepsilon 2}$ - 522 $N_{\delta 2}$ | 1.68 | 2.64 | mouth of cleft | Fe |
| 353 $O_{\varepsilon 1}$ - 521 $O_{\eta}$ | — | 3.22 | mouth of cleft | Fe / apo |
| 375 O - 520 $N_{\delta 2}$ | 1.57 | 2.51 | mouth of cleft | Fe / apo |
| 395 $O_{\delta 2}$ - 466 N | 1.83 | 2.75 | cleft | Fe / apo |
| 395 $O_{\delta 2}$ - 466 $O_{\gamma 1}$ | — | 3.36 | cleft | Fe |
| 529 $O_{\gamma 1}$ - 638 O | — | 3.01 | mouth of cleft | Fe |
| 560 $O_{\varepsilon 1}$ - 638 $O_{\gamma 1}$ | — | 2.64 | mouth of cleft | — |
| **B: Contacts involving strands e and j in the cross-over region** | | | | |
| 418 N - 433 $O_{\varepsilon 2}$ | 2.36 | 2.94 | domain C1 - strand e (s/c) | Fe |
| 434 N - 596 $O_{\delta 1}$ | 2.16 | 2.94 | strand e - strand j (s/c) | Fe / apo |
| 438 N - 545 O | 2.09 | 3.07 | strand e - strand h (C2) | Fe / apo |
| 438 O - 545 N | 2.41 | 3.33 | strand e - strand h (C2) | Fe / apo |
| 392 O - 600 N | 2.11 | 3.07 | strand c (C1) - strand j | Fe / apo |
| 394 N - 598 O | 1.64 | 2.64 | strand c (C1) - strand j | Fe / apo |
| 394 O - 598 N | 1.86 | 2.75 | strand c (C1) - strand j | Fe / apo |
| 410 N - 599 O | 1.54 | 2.49 | strand d (C1) - strand j | Fe / apo |
| 411 O - 599 N | 1.66 | 2.65 | strand d (C1) - strand j | Fe / apo |
| 413 O - 596 $N_{\delta 2}$ | 2.28 | 2.98 | strand d (C1) - strand j (s/c) | Fe / apo |
| 596 O - 662 $O_{\eta}$ | — | 2.49 | strand j - domain C1 | Fe |

[a] For hydrogen bonds other than NH···O, $O_1$ and $O_2$ relate to the first- and second-named residues. [b] s/c denotes sidechain contacts.