A mutational analysis of the hinge region of the N-lobe of lactoferrin.

A thesis presented in partial fulfilment of the requirements for the degree of Master in Science in Biochemistry at Massey University.

Steven Christopher Shewry

1996

ABSTRACT

Lactoferrin is an 80kDa iron binding glycoprotein that is found as a major component of human milk, as well as in many other exocrine solutions. Lactoferrin binds reversibly, and with high affinity, 2 Fe^{3+} ions with 2 synergistic CO_3^{2-} ions. The crystal structure shows that the polypeptide chain is folded into two similar lobes, each binding one Fe^{3+} and CO_3^{2-} ion. In the metal free state, the N-terminal lobe has been found to adopt an open structure, with rotation occurring around two residues found in separate beta strands located at the back of the binding site in a region referred to as the hinge. A sequence alignment of these two strands over the greater transferrin family shows a very high level of conservation particularly of the two residues at the centre of this rotation (Pro 251 and Thr 90).

The N-terminal half of human lactoferrin (Lf_N) has been constructed, expressed and the crystal structure determined.

In an attempt to understand the importance of the conservation of these two residues, and their effect on binding, a mutational analysis was initiated. Oligonucleotide site-directed mutagenesis has been used to construct mutants in the cDNA encoding for human lactoferrin using the M13 bacteriophage. The mutant cDNA was transformed into a mammalian expression vector (pNUT). After transfection of the pNUT vector into baby hamster kidney cells (BHK), the mutant proteins were expressed and purified from the culture medium using a CM-sephadex ion-exchange column.

Absorption maxima and pH-dependent iron-release experiments were carried out on the mutants. The data shows that the mutants behave essentially the same as Lf_N, the exception being P251G which appears to release the iron over a shorter pH range. The reason for this is not yet fully understood.

The crystal structure of P251A in the iron-bound form was solved by molecular replacement using the structure of LfN as the starting model. The structure of P251A was refined using data between 20.0 and 2.0 Å. The current model has good geometry and has an R-factor of 18.6 %. Analysis of the structure shows that it is essentially identical to that of the LfN structure.

Although the structure of the iron-free form has not been determined, it appears that changes to the hinge region of the N-lobe of lactoferrin do not affect the iron-binding or structural characteristics of the protein.

Acknowledgements.

This thesis was made possible by the contributions of a number of people, to whom I wish to offer my appreciation.

Firstly to my supervisors, Professor Ted Baker and Associate professor John Tweedie for their advice and guidance throughout the course of this work. I am particularly grateful for Ted who despite a busy schedule still found time to read this thesis in fast time.

To all members of the protein crystallography and Twilight zone groups (past and present) for helping me come to terms with the hazy aspects of this work. Special thanks go to Heather Baker, Drs Gillian Norris, Clyde Smith, Jakki Cooney, Stanley Moore, Rick Faber, Maria Bewley, and to Neil Peterson for their kind ear during the difficult times.

Special mention to Dr Hale Nicholson for his help in the initial stages of the DNA work and for helping me to enter this mysterious world of nucleotides, and to Dr Bryan Anderson for his help with the confusing (I mean computing) part of the structure refinement.

Special thanks also to Dr Simon Brown for his help in the mathematical part of the iron release curves. Coming to terms with the equations is not an activity I would wish on anyone.

I am especially grateful to my family (my wife Philippa and children Jackson, Nathan, and Esther) for standing by me during what must have been a black hole experience for them. They still don't know what I was doing all that time. In particular to the youngest two who know no other way of life. May it get more interesting for them from now on.

Table of Contents

List of Figures.

List of Tables

Abbreviations

A280	Absorbance at 280 nm
A454	Absorbance at 454 nm
ApoLf	Iron free native lactoferrin
ApoLf$_N$	Iron free Lf$_N$
ATP	adenosine triphosphate
BHK	baby hamster kidney
bp	base pair
BRL	Bethesda Research Laboratories
BSA	bovine serum albumin
cDNA	complementary DNA
CM-sephadex	carboxymethyl sephadex
C-terminal	carboxy terminal
Cu$_2$Lf	copper saturated native lactoferrin
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
DG	deglycosylated
DGLf$_N$	deglycosylated Lf$_N$
dGTP	deoxyguanosine triphosphate
DHFR	dihydrofolate reductase
DMEM	Dulbeco's modification of Eagle's medium
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	mixture of dATP, dTTP, dCTP and dGTP
dsDNA	double stranded DNA
DTT	dithiothreitol
dUTP	deoxyuridine triphosphate
EDTA	ethylene diamine tetra-acetate
ELISA	enzyme linked immunoabsorbant assay
Endo F	endo β-N-acetylglucosamidase F1
F's	structure factors
Fcalc	calculated structure factors
Fobs	observed structure factors
F12	Hams-F12-medium
FCS	foetal calf serum
FeLf$_N$	iron saturated Lf$_N$
Fe$_2$Lf	iron saturated native lactoferrin

HEPES	N-2-hydroxethyl piperazine-N'-2-ethane sulfonic acid
hLf	human lactoferrin
hTr	human serum transferrin
I	intensities
IPA	isopropanol
IPTG	isopropyl β-D-thiogalactopyranoside
kb	kilobase
kDa	kilodalton
LB	Luria-Bertani
Lf	lactoferrin
Lf_N	amino terminal half of human lactoferrin
MIR	multiple isomorphis replacement
MTX	methotrexate
NAG	N-acetyl glucosamine
NMR	nuclear magnetic resanance
NTA	nitrilotriacetic acid
N-terminal	amino terminal
PAGE	polyacrylamide gel electrophoresis
PBP	periplasmic-binding protein
PBS	phosphate buffered saline
PEG	polyethylene glycol
pfu	plaque forming units
PMN	polymorphonuclear
PNGase F	peptide-N4-(N-acetyl-β-D-glucosaminyl) asparagine amidase F
RNA	ribonucleic acid
RNAse	ribonuclease
SDS	sodium dodecylsulphate
SDS-PAGE	sodium dodecylsulphate polyacrylamide gel electrophoresis
ssDNA	single stranded DNA
TAE	Tris acetate EDTA
TE	10 mM Tris, 1 mM EDTA
TEMED	N,N,N',- Tetramethylethylenediamine
Tris	Tris-(hydroxymethyl) aminomethane
UV	ultraviolet
X-gal	5-bromo-4-chlor-3-indoyl β-D-galactopyranoside

A. INTRODUCTION

A.1. Introduction.

For all plants and animals, and for virtually all microbes, life without iron is impossible. In the cell, where essential iron-containing enzymes and proteins are found in species ranging from bacteria to man, the thermodynamically stable form of iron is Fe(III). Free Fe^{3+} is, however, rapidly hydrolysed in solution to form the insoluble $Fe(OH)_3$ complex. This insolubility problem is overcome by proteins involved in storage and transport of iron.

Transferrin is the name given to the family of proteins involved in the solubilisation, sequestration, and transport of ferric iron. The roles of transferrins in general is coupled to their tight binding of Fe^{3+}. These roles involve controlling the levels of free iron in body fluids, preventing precipitation of $Fe(OH)_3$ by keeping the concentration of free Fe^{3+} at less than 10^{-18} M, and preventing free radical damage (free radical production is catalysed by free iron). The name "Transferrin" is taken literally to mean "transport of iron" (Baker, 1994) and the transferrin family includes at least four iron-binding proteins, serum transferrin, ovotransferrin, lactoferrin, and melanotransferrin. Serum transferrin is the iron transport protein found in the blood of vertebrates and some invertebrates. Ovotransferrin, found in the white of eggs, is identical to the serum transferrin of birds except in the carbohydrate moiety attached to the protein (Thibodeau *et al,* 1978). Melanotransferrin is a membrane-associated protein found on all cells, but expressed at high levels on melanoma cells (Rose *et al,* 1986). Lactoferrin, the subject of this study, is the iron-binding protein most commonly found in milk, but also in other exocrine secretions including tears, saliva, seminal fluid, cervical mucous, gastric fluid, nasal exudate, bronchial mucous, and hepatic bile (Masson *et al,* 1969; Weinburg, 1984). Lactoferrin is also found in the specific granules of mature neutrophils (Baggolini *et al,* 1970).

The transferrins are glycoproteins with a molecular weight of about 78,000 Da. The number and positioning of the carbohydrate chains differs from one transferrin to another. For example, human serum transferrin has two chains attached to the C-lobe of the molecule, while human lactoferrin also has two chains, but one attached to each of the lobes.

Melanotransferrin has additional amino acids on the C terminal end of the molecule thought to be involved in anchoring the protein to the cell membrane (Rose *et al*, 1986), and responsible in part for its higher molecular weight of 97,000 Da. Other transferrins of differing size are the crab serum transferrin (Cancer magister) with a molecular weight of about 150,000 Da, and the transferrin of Pyura stolonifera with a molecular weight of about 40,000 Da (Baker, 1994).

The transferrins are generally folded into two lobes each of which reversibly binds a ferric ion concomitantly with a bicarbonate anion. Two members of the family, however, melanotransferrin and the transferrin of the hornworm (Manduca sexta), bind only one ferric ion (Baker, 1994) as a result of amino acid sequence changes in one of the two binding sites. Apart from these, the metal and anion-binding amino acid residues are identical in both lobes of the transferrins, and are conserved in all members of the transferrin family.

Lactoferrin is found in the milk of some, but not all, species, being completely absent in the milk of the rat, rabbit and dog. The milk of the rat and rabbit contains significant levels of transferrin, while the milk of the dog contains neither protein (Masson and Heremans, 1971). Lactoferrin levels are highest in slow-growing species such as the human and guinea pig, and are low or absent in fast-growing species such as rabbits, rats and cattle (Weinburg, 1984). Lactoferrin expression also varies during lactation and this suggests that the gene is hormonally regulated. This is supported by the work of Pentecost and Teng (1982), who found that the major oestrogen-stimulated protein synthesised in the uterine tissue of the mouse is lactoferrin.

A.2. Biological roles of lactoferrin

For many years, a considerable amount of effort has been directed at determining the biological role(s) of lactoferrin. Many different roles have been postulated, but as yet, no single, primary one has been identified. Among those that have been suggested are (i) iron-withholding in order to starve potential pathogens of iron, (ii) a bactericidal function, (iii) binding to receptors in order to perform various functions, (iv) modulation of the inflammatory response, (v) acting as a growth promoter, and (vi) as a source of iron for nutritional reasons.

Many other functions have been associated with lactoferrin, but the evidence is somewhat dubious. The following are the most widely accepted potential functions.

A.2.1. Iron withholding.

The egg is a good example of iron-withholding defence. The developing embryo is provided with a large amount of iron (1 μg in the egg yolk) for use in DNA synthesis, electron transport, and other important functions, while a porous shell is needed in order for air to be available. This porous shell, however, allows microbial invaders to enter the egg, and this problem of microbial invasion is overcome by placing no iron (an essential element for bacteria) in the egg white, while also including a powerful iron-binding protein. This iron-binding protein was discovered by Osborne and Campbell (1900), and is called ovotransferrin.

Schade and Caroline (1944) demonstrated that of the 10 vitamins and 31 elements tested, only iron overcame the bacteriostatic activity of egg white. The bacteriostatic factor (ovotransferrin) suppressed activity of gram negative bacteria, gram positive bacteria, and also fungi. Alexander in 1948 discovered that babies that were breast fed suffered less severe cases of gastric enteritis than did bottle fed babies. A factor in human milk was therefore responsible for this reduced level of infection. It was later discovered (Bullen *et al*, 1972) that samples of milk with unsaturated iron-binding capacities had a powerful bacteriostatic effect on *E.coli* 0111/B4. This effect was abolished if the proteins involved were saturated with iron.

In vivo work (Murray *et al*, 1978) shows that iron indeed enhances the incidence of infection. Somali nomads have a high incidence of iron deficiency, due to their all-milk diet, and also show a low level of intestinal parasites. Murray *et al* fed the nomads daily for 30 days, either with placebo (aluminium hydroxide) as a control, or with iron (900 mg $FeSO_4$). The subjects were observed for fever above 38°C and symptoms of infection. 7/67 of the placebo group and 36/71 of the iron group suffered infections, suggesting that the host defence against these infections was better with an iron deficiency than with iron present.

The factor accounting for these observations (lactoferrin), was discovered, purified and studied around 1960 (Johansson, 1960) An antibacterial role for lactoferrin involving iron sequestering is also supported indirectly by the observation that both human and bovine lactoferrin are secreted in the apo form (Lonnerdal, 1985).

It seems that vertebrate hosts inhibit microbial growth by withholding iron, and do so by the use of proteins of the transferrin family, serum transferrin, ovotransferrin, and lactoferrin.

A.2.2. Bactericidal activity.

Although the bacteriostatic activity of lactoferrin, based on iron sequestering, appeared well established, Arnold *et al* (1977) demonstrated that in some cases this activity was not reversed by the addition of iron. This pointed to an alternative mechanism that could be involved in some cases. In this work, immunofluorescence studies showed apolactoferrin binding to the surface of susceptible bacteria. A possible explanation for this was that inhibition was due to lactoferrin blocking sites that were required for the transport of an essential nutrient. If this was so, then inhibition should be reversed by the removal of any surface lactoferrin, and any lactoferrin-inhibited cells should then be able to continue to grow. Arnold *et al* (1982) removed the lactoferrin by techniques that retained the viability of control bacteria. The experimental bacteria (*Streptococcus mutans*) however failed to regain viability upon treatment to remove the lactoferrin. In effect, it showed the irreversible bactericidal effect of lactoferrin upon *S.mutans*. Other supporting evidence comes from Valenti *et al* (1985 and 1987) who showed, by use of a dialysis membrane, that the lactoferrin needed to be in direct contact with the bacterial surface to exert its bactericidal effect.

In a recent report (Tomita *et al,* 1994), lactoferrin was found to agglutinate the protoplasts of *Micrococcus lutens*. This agglutination was lost by chemical modification of the basic residues of lactoferrin, indicating that electrostatic action is involved. Using phase contrast microscopy, and spectroscopy, Tomita *et al*, (1994) found that both the apo and iron loaded forms of lactoferrin inhibited the growth of the bacteria. This points to an antibacterial mechanism in addition to the iron witholding bacteriostatic mechanism, as the latter should be effective only with the apo form of lactoferrin. Ovotransferrin and human transferrin had no effect and it was concluded that the cationic charge from lysine or arginine residues in lactoferrin are needed to agglutinate the bacterial cells.

Bactericidal activity is also found in proteolytically-derived and synthetic cationic peptides corresponding to the N terminal regions of lactoferrin (Bellamy *et al,* 1992). The precise bactericidal mechanism remains unclear. It may involve direct interruption of the membrane by this region, or the blocking of essential receptors on the bacterial outer membrane surface.

To understand the mechanism of action of one of these peptides, Bellamy *et al,* (1993) investigated the cell binding properties of lactoferricin B (a peptide corresponding to the N terminal 25 amino acids of bovine lactoferrin) with both gram positive (*Bacillus subtilis*) and gram negative (*E.coli*) bacteria, and compared these properties with the rate of irreversible death of these cells. Lactoferricin B is highly cationic with 8/25 basic residues. Other antimicrobial peptides with similar cationic

features include magainins from frog skin; cecropins from the haemolymph of insects; and defensins from mammalian neutrophils (Westerhoff *et al,* 1989), (Hill *et al,* 1991).

^{14}C-labelled lactoferricin B bound to the cell surface of both gram negative and gram positive bacteria and the rate of binding was consistent with the rapid rate of killing observed. Cell-binding activity was pH dependent implying that the cationic property of the peptide was important. The optimum pH however was dependent on the particular bacterial strain (pH 6.0 for *E.coli*, pH 7.5 for *B.subtilis*). With each strain, the killing effect was maximum near the optimum pH for cell binding implying that the bactericidal effect of lactoferricin B is dependent upon cell binding. This reflects the earlier work by Valenti *et al* (1987) with lactoferrin.

Naidu *et al* (1993) generated a peptide from lactoferrin by pepsin hydrolysis. This also showed antimicrobial activity, but failed to inhibit ^{125}I-labelled lactoferrin from binding to *E.Coli* and 10 other species of the *Enterobacteriaceae*. This indicated that lactoferrin and this peptide may bind by different mechanisms, and cast some doubt on the relevance of peptides in elucidating the function of lactoferrin.

An interesting twist to this bacteriostatic/bactericidal role is the work by Yu and Schryvers (1993) showing that some pathogenic bacteria (*Neisseria meningitidis*, *N.gonorrhoeae*, and *Moraxella catarrhalis*) may work by inhibiting lactoferrin and transferrin. They found by SDS PAGE that these bacteria bound lactoferrin to their membrane, and that the receptors that bound human lactoferrin were different from the ones that bound human transferrin. They propose that these bacteria bind lactoferrin in order to acquire iron from it, therefore overcoming its bacteriostatic effect. It does not rule out, however, a defensive role against the bactericidal effect of lactoferrin.

A.2.3. Receptor binding.

One of the intriguing characteristics of lactoferrin is its ability to bind to a wide variety of cell types. Several researchers have investigated this binding, and have related it to possible functions for lactoferrin.

A putative intestinal receptor for lactoferrin in humans was initially proposed by Cox *et al* (1979) and was isolated from rabbit brush border cells in 1989 (Mazurier *et al,* 1989). This receptor was found to be a 100 kDa protein.

Lactoferrin has also been found to bind to the brush border cells of rhesus monkeys (Davidsson and Lönnerdal, 1988) under conditions in which these cells can accumulate iron from lactoferrin. Only human lactoferrin released its iron to these cells, bovine lactoferrin, human transferrin, and ovotransferrin having no such ability

(Davidsson and Lönnerdal, 1989). Lactoferrin receptors have also been reported in mouse small intestinal brush border cells (Hu *et al,* 1988).

The binding constant and number of receptors per cell have been calculated for caco-2-cells, a human colon carcinoma cell line with characteristics of the brush border cells (Iyer *et al,* 1993). Human lactoferrin was labelled either with ^{59}Fe, or with ^{125}I. The results showed that human lactoferrin in both forms bound to the receptor in a saturable manner. Scatchard analysis indicated that it was a single binding site with a Kd of 1.7×10^{-6}, and about 6×10^{5} binding sites per cell. At 37°C, lactoferrin was taken up by the cell. ^{59}Fe taken up as ^{59}Fe-Lf was not transported across the monolayer, but when ^{125}I-Lf was used, there was a continuous transport of ^{125}I associated with a major protein later confirmed to be lactoferrin using gel electrophoresis and ELISA. In conclusion, Fe_2Lf is taken up by the cell, iron is then released and used by the cell, and Lf in part is transported across the cell.

This receptor has been cloned, sequenced and expressed in a baculovirus system, and confirmed to bind hLf by ligand blotting. The protein consists of 351 amino acids with four potential glycosylation sites. There is a strong hydrophobic region of 29 amino acids possibly acting to anchor the receptor to the membrane (Iyer *et al,* 1994).

Lactoferrin receptors have also been described in other cell types including lymphocytes, parenchyma liver cells, macrophages, and monocytes. With human peripheral lymphocytes, Mazurier *et al* (1989) found that in these cells during the resting state there were no lactoferrin receptors present. However upon stimulation by phytohemagglutinin, lactoferrin receptors were visualised using ^{125}I-labelled lactoferrin as a probe. The receptor gave rise to 2 bands of 100 and 110 kDa.

A high affinity receptor has also been discovered in monocytes with a Kd of 1 nM (Yuen *et al,* 1993). This work found that both monocytes and lymphocytes showed a dose dependent response in interleukin(IL)-1B and IL-6 to increasing levels of lactoferrin. It appears that lactoferrin binds to these cells and this results in increased IL-1B and IL-6 levels.

McAbee and Esbensen (1991) showed that hepatocytes also bind and internalise lactoferrin. 80% of the lactoferrin was internalised by these cells after 60 min at 37°C. The functional consequences of lactoferrin binding to these cells remains unclear, although lactoferrin has been found, in vitro and in vivo, to inhibit the binding and uptake of apoE-bearing lipoproteins by parenchyma liver cells (Huettinger *et al,* 1988; Van Dijk *et al,* 1991). Arginine residues in lactoferrin were shown to be crucial for recognition of lactoferrin by the liver cells (Ziere *et al,* 1992). More recent work by the same group (Ziere *et al,* 1993), showed, by removal of the first 14 amino acids, that it is the four arginine residues at the N terminus that are important.

As previously outlined, the pathogenic bacteria *Neisseria meningitidis*, *N.gonorrhoeae*, and *Moraxella catarrhalis*, overcome the iron sequestering effect of lactoferrin by using receptors in their membranes to bind the individual lobes of lactoferrin (Yu and Schryvers, 1993). In the same way, lactoferrin can also bind to certain bacteria from the Enterobacteriaceae family. Gado *et al* (1991) identified porins as the lactoferrin binding site on these bacteria. This binding was reversible and had a low affinity. Yu and Schryvers (1993) found that it was both the C- and N-lobes of lactoferrin that were responsible for binding to bacterial receptors. This contrasts with transferrin, in which the bacterial transferrin receptor only binds to the C-lobe of human transferrin (Alcantara *et al*, 1993). The region of human lactoferrin involved in binding to human lymphocytes however, has been localised to domain 1 of the N-lobe and there is no observable binding of the C-lobe (Rochard *et al*, 1989), (Legrand *et al*, 1992). This shows that the interaction with lactoferrin is different for mammalian receptors than for bacterial receptors.

A.2.4. Inflammation.

Another function proposed for lactoferrin is modulation of the inflammatory response. Lactoferrin is found in the specific granules of neutrophils (Baggiolini *et al*, 1970), and during the inflammatory response, levels of lactoferrin in the granules decline, while plasma levels of lactoferrin increase (Malmquist *et al*, 1978). It seems that the neutrophils release the lactoferrin, although it is not known how this occurs. Boxer *et al* (1982), using immunohistological techniques, found that lactoferrin released from neutrophils bound to the polymorphonuclear (PMN) membrane and altered its surface properties by reducing the surface charge. Transferrin did not do this. This work correlates with the observation that during inflammation there is an increased adherence of the PMN cells to the endothelial cells which results in an amplification of the inflammatory response (Oseas *et al*, 1981). By binding to the PMN membrane, lactoferrin helps the PMN cells to adhere to the endothelial cells and therefore amplify the inflammatory response.

The release of lactoferrin from neutrophils has also been suggested to have an antibacterial role (van Snick *et al*, 1974). It was proposed that lactoferrin released from the neutrophils bound free iron, possibly derived from transferrin, and once iron-saturated, lactoferrin bound to macrophages which were then removed by the reticuloendothelial system. This bacteriostatic role of lactoferrin in neutrophils is supported by the finding that patients whose neutrophils lack the specific granules suffer from recurrent infections (Sanchez *et al*, 1992), but has been questioned by Baynes and Bezuida (1992) after they could not detect significant levels of lactoferrin

in the plasma.

A.2.5. Growth factor activity.

Transferrin is known to be involved in cell proliferation (Casey *et al*, 1989), and this raises the possibility that lactoferrin may also have a role in growth-promoting activity. This was proposed when human lactoferrin was found to stimulate the growth of all the human lymphoid cell lines tested, but not mouse lymphocyte cell lines, a human epithelial cell line, or a human fibroblast cell line (Hashizume *et al*, 1983). Human lactoferrin in the iron loaded form has also been found to stimulate the growth of phytohaemagglutinin-stimulated human peripheral blood lymphocytes. Moreover, receptors for lactoferrin appeared on the surface of these cells upon stimulation by phytohaemagglutinin (Mazurier *et al*, 1989).

In addition to its possible growth-promoting activity, lactoferrin has also been shown to indirectly inhibit the growth of certain cell lines. Lactoferrin is responsible for regulating the proliferation of the granulocyte-macrophage progenitor cells by decreasing the production and release of a colony-stimulating activity found in macrophages (Broxmeyer *et al* 1978) . Lactoferrin is known to bind monocytes and macrophages, and this binding activity correlates with the inhibitory effect on these cells.

A.2.6. Iron nutrition.

Another suggested role for lactoferrin is in the uptake of iron for nutritional value. To perform this function, lactoferrin must be able to withstand proteolysis in the gastrointestinal tract, and evidence for this resistance was provided in the work of Spik *et al* (1982). A nutritional role for lactoferrin is also supported by the finding that receptors for lactoferrin have been identified in the small intestines of a variety of species (Iyer and Lonnerdal 1993). Lactoferrin binding then apparently leads to the retention of the iron by these cells (Iyer *et al*, 1993).

This function is questioned by Davidsson *et al* (1994) who separated out lactoferrin from human milk by treatment with heparin-sepharose, and then measured the incorporation of ^{58}Fe into red blood cells 14 days after feeding infants aged 2-10 months old. Their results showed that less ^{58}Fe was incorporated into red blood cells with lactoferrin-containing milk, than with lactoferrin-deficient milk.

A.2.7. Other possible roles.

As mentioned earlier, lactoferrin has many other postulated functions, one of which is to protect monocytes from oxidation (Brighton *et al,* 1991). It is proposed that lactoferrin binds to a receptor on the monocyte preventing hydroxyl damage to these and neighbouring cells.

Whatever the function(s) of lactoferrin, it is generally accepted that most are linked in some way to the binding of iron.

A.3. Structure of lactoferrin.

A.3.1. Primary structure.

At the present time, the amino acid sequences of 17 proteins of the transferrin family have been determined either directly or from cDNA sequences. These are listed in Table A.1.

All the members of the transferrin family demonstrate a high degree of sequence similarity. This includes approximately 65% amino acid identity within the lactoferrins, and approximately 55% between the transferrins of higher animals and lactoferrin (Mead, 1992). The sequences of the insect transferrins and the membrane-associated melanotransferrin are less conserved. Melanotransferrin has a 40% amino acid sequence identity with the mammalian transferrins, while the insect transferrins have 20 - 30% amino acid sequence identity with the higher transferrins (Baker, 1994).

In addition to the amino acid sequences, the complete genomic DNA sequences for human transferrin (Schaeffer *et al,* 1987), chicken ovotransferrin (Jeltsch *et al,* 1987), and mouse lactoferrin (Shirsat *et al,* 1992), have been reported. Comparisons of these sequences reveal that although the sizes of the genes vary, this is due to variation in the sizes of the introns and not the exons (coding regions). The positions of the intron/exon boundaries are conserved among transferrins. The position of one of the conserved exon/intron boundaries is between the N and C-lobes of the protein, supporting the theory of gene duplication put forth by Bowman and colleagues (1988).

Table. A.1. Primary sequences of the transferrin family known to date.

Protein	Reference
Human Serotransferrin	Park *et al*, 1985, MacGillivray *et al*, 1983
Horse Serotransferrin	Carpenter *et al*, 1993;
Pig Serotransferrin	Baldwin and Weinstock, 1988
Rabbit Serotransferrin	Banfeld *et al*, 1991
Rat Serotransferrin	Schreiber *et al*, 1979
Xenopus Serotransferrin	Moskaitis *et al*, 1991
Human Lactoferrin	Rado *et al*, 1987, Metz-Boutigue *et al*, 1984
Bovine Lactoferrin	Mead and Tweedie, 1990
Goat Lactoferrin	Le Provost *et al*, 1994
Mouse Lactoferrin	Pentecost and Teng, 1987
Pig Lactoferrin	Alexander *et al*, 1992
Human Neutrophil Lactoferrin	Rado *et al*, 1987
Human Melanotransferrin	Rose *et al*, 1986
Chicken Ovotransferrin	Jeltsch and Chambon,1982; Williams *et al*, 1982
Atlantic Salmon Transferrin	Kvingedal *et al*, 1994
Hornworm Transferrin	Bartfeld and Law,1990
Cockroach Transferrin	Jamroz *et al*, 1993

A.3.2. Three-dimensional structure.

A.3.2.1. Iron-loaded lactoferrin.

The first crystallographic studies on transferrins date back more than 20 years (Magdoff-Fairchild and Low, 1970), and the first low resolution analysis of a transferrin was published in 1979 (Gorinsky *et al*,1979). This involved rabbit serum transferrin, and demonstrated the bilobal nature of the molecule.

In 1987, Anderson *et al* determined the structure of human lactoferrin at 3.2 Å resolution and this structure was subsequently refined at 2.8 Å resolution (Anderson *et al*, 1989). The polypeptide folding of this structure is shown in Fig A.1.

Fig. A.1. Drawing created by Ribbons (Richardson, 1985; Priestle, 1988) showing the structure of the full-length lactoferrin molecule with iron bound in both lobes.

The polypeptide chain is folded into two lobes, representing the N- and C-terminal halves of the molecule (residues 1-332 and 345-691 respectively). These are joined by a short, 3-turn, alpha helix, (residues 333-344) that consists largely of glutamyl, arginyl and alanyl residues. Both lobes have very similar polypeptide folding, consistent with their high level of amino acid sequence identity (about 40%), and each lobe contains a single iron-binding site.

Inspection of the folding (Fig A.1) shows that each lobe is further subdivided into two similar sized domains (of about 160 residues), with a deep cleft between them, which houses the iron-binding site (Fig A.2). The domains in the N-lobe of human lactoferrin have been labelled N1 (residues 1-90 and 252-332) and N2 (residues 91-251), with the equivalent C-lobe domains being C1 (residues 345-433 and 596-691) and C2 (residues 434-595).

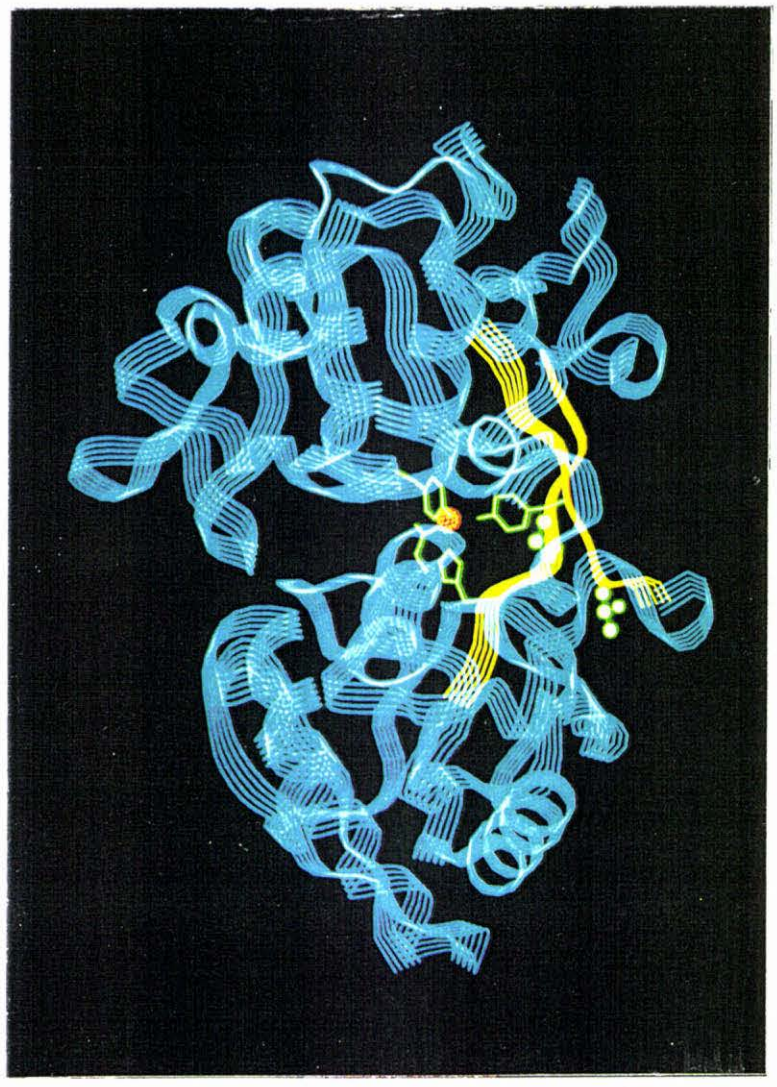

Fig. A.2. Drawing created by Ribbon (Richardson, 1985; Priestle, 1988) showing the structureof the half-length lactoferrin molecule (Lf_N) with iron bound. The side chains for the ligands are included, as are the side chain of two residues of this study, Thr 90 and Pro 251.

The two iron-binding sites are extremely similar. In each case the iron atom is coordinated by four protein sidechains, 2 tyrosines, 1 histidine, and 1 aspartate, coming from separate parts of the polypeptide chain. In the N-lobe these ligands are Asp 60, Tyr 92, Tyr 192, and His 253, while in the C-lobe they are Asp 395, Tyr 435, Tyr 528, and His 597. A unique feature of transferrins is that a carbonate (or bicarbonate) anion is bound with each metal ion, and neither metal ion nor anion is bound strongly by the protein in the absence of the other. The relationship is thus said to be synergistic. In the binding site the carbonate anion is found to bind to the metal in bidentate fashion, thus completing the metal coordination, and to also form a bridge between the metal ion and a positively charged region of the protein. The latter comprises an arginine sidechain

(Arg 121 in the N-lobe, Arg 465 in the C-lobe) and the N terminus of an alpha helix, these being presented by one wall of the second domain (N2 or C2). The binding site is shown schematically in Fig A.3.

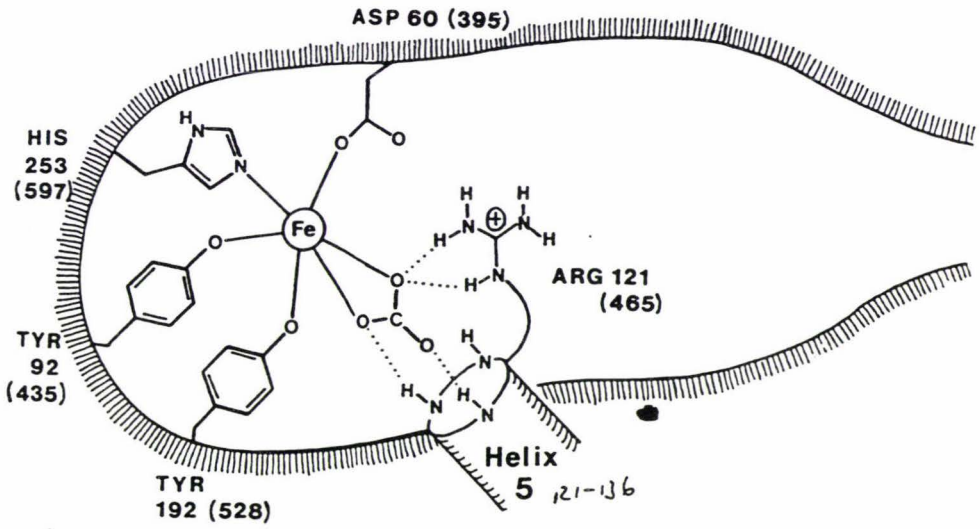

Fig A.3. Schematic diagram of the metal and anion binding site in lactoferrin. Residue numbers are for the N-lobe (with the C-lobe given in brackets). (From Baker, (1994) with permission).

With the knowledge of the lactoferrin structure, the structure of diferric rabbit serum transferrin has since been completed (Bailey *et al*, 1988). The overall folding pattern was similar to that of human lactoferrin, with similar domains, each based on a mixed beta sheet overlaid with alpha helices which pack against the two faces of the sheet. The main difference between the two structures was in the orientation of the two lobes. With the N-lobes superimposed, there is a 15 degree rotation of the C-lobe of transferrin compared to lactoferrin (Baker and Lindley ,1992).

Despite observed differences in function between rabbit serum transferrin and human lactoferrin, the iron-binding sites are similar. If the atoms of the iron-binding site are superimposed, the root mean square deviation between atomic positions for all the combinations of the N and C-lobes is less than 0.6 Å (Baker and Lindley, 1992). Therefore the functional differences between transferrin and lactoferrin are probably not concerned with the precise structure of the iron-binding sites. It has been suggested instead that the interactions between the two domains within a lobe may influence the functional differences between the transferrins and lactoferrins (Baker and Lindley, 1992).

The carbohydrate chains on transferrins are all N-linked through asparagine residues. Glycosylation sites vary in number from one in rabbit serum transferrin and chicken ovotransferrin, to four in bovine lactoferrin (human lactoferrin contains two sites). The sites are scattered over the surface of the molecule arguing against any direct functional role in iron-binding. The carbohydrate chains are heterogeneous with little defined structure observable by X-ray analysis. Except for certain species of fish (*Tinca tinca* and *Ctenopharybgodan idella*), all known vertebrate transferrins are glycoproteins (Stratil *et al,* 1983). The absence of carbohydrate chains in these species suggests that they are not important in the physiological functions of the proteins.

One intriguing question that has never been unequivocally answered is that of whether there is any biological advantage in having a bilobal molecule, with two iron sites. The bilobal structure has clearly arisen by gene duplication and fusion from a small ancestral protein of half the size (40kDa, a single lobe) with a single iron-binding site (Bowman *et al,* 1988). One possible explanation for the duplication has been advanced by Williams *et al* (1982) who found that isolated lobes of serum transferrin were rapidly lost from the bloodstream via the kidney. In this case only species that had evolved this bilobal form would have survived. Williams however proposes that a single lobed ancestor is unlikely, and instead, the ancestral protein may have been a bilobal form, or a membrane associated protein. An alternative possibility is that the two lobes have become differentiated in their properties (see later).

A.3.2.2. Apolactoferrin structure.

A proper understanding of iron binding and release requires that the nature of the associated conformational changes be defined. Although X-ray analysis gives only a static picture, comparisons of the structures of iron-loaded and iron-free strucutres are an essential element in reaching such an understanding.

The apolactoferrin structure was solved from protein in which the carbohydrate had been removed by enzymatic digestion using peptide N-glycosidase F (PGNaseF) and endoglycosidase F (Endo F), both isolated from *Flavobacterium meningosepticum* (Elder and Alexander, 1982). This deglycosylated protein had identical properties of iron-binding and release, and identical spectroscopic parameters to the native form. The crystals diffracted to 2.0 Å resolution (Norris *et al,* 1989). The structure was determined by molecular replacement using diffractometer data to 2.8 Å resolution, and refined to an R factor of 0.213 for data between 10 and 2.8 Å.

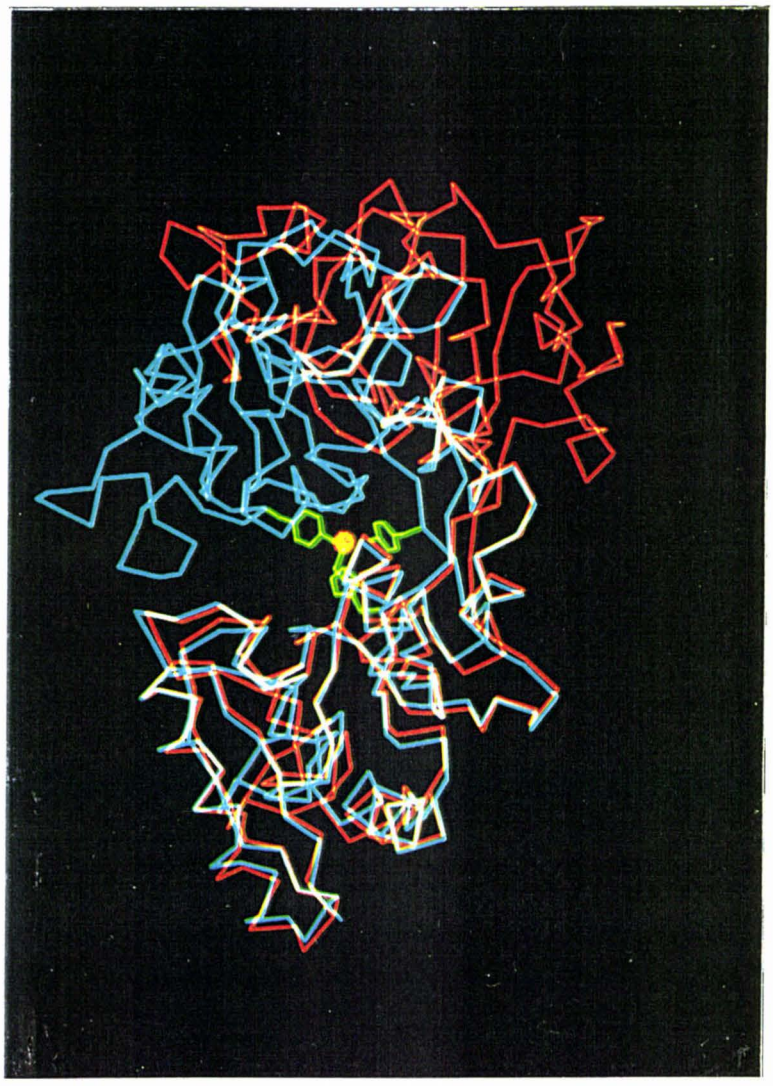

Fig. A.4. Superposition of the open and closed N-lobes in the full-length lactoferrin structure.

A striking feature of the X-ray analysis (Anderson *et al*, 1990), was the large conformational change seen for the N-lobe of the molecule (Fig A.4). One domain (N2) had rotated 54 degrees relative to the other domain (N1), resulting in a wide-open binding cleft. The equivalent movement was not seen in the C-lobe, however. Both sites (N and C-lobes) had lost their Fe^{3+} ion. This one-open, one-closed structure was unexpected and two theories to explain this were proposed, (i) that extra constraints in the C-lobe (specifically a disulphide bridge 483-677) inhibit opening of the C-lobe, and (ii) that an equilibrium exists between the open and closed forms in solution as seen for the periplasmic binding proteins (PBP's) (Oh *et al* 1993). Crystal packing could then have selected the closed form of the iron-free state. If this is so, then the energy difference between the open and closed forms of the iron free protein must be small. This second explanation was confirmed, firstly by solution X-ray scattering measurement (Grossmann *et al,* 1992), and secondly when Faber *et al* (1995) solved the

structure of a second crystal from the apoprotein in which both sites were open (Fig A.5). The C-lobe only opens by about 15 degrees, probably due to the extra constraints present in the C-lobe.

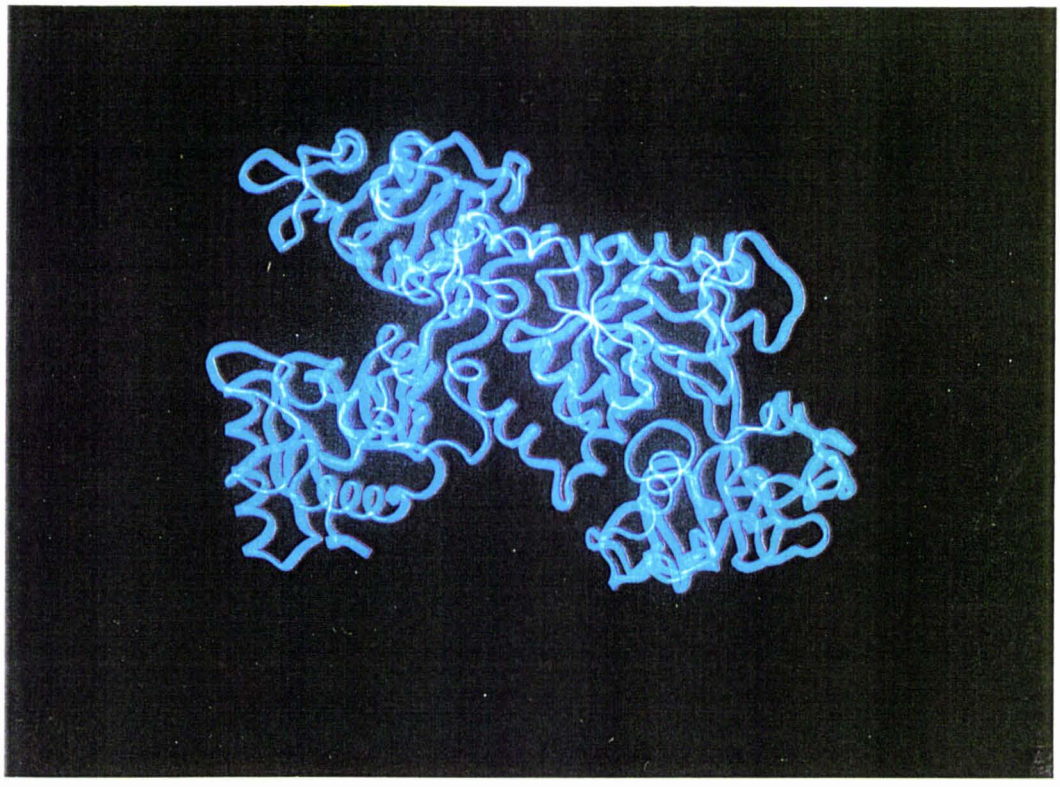

Fig. A.5. Structure of full-length lactoferrin with both the N- and C- lobes open.

In each lobe the conformational change that mediates the transition from closed to open form involves relative movement of the two domains. The axis of this movement is through a "hinge" in two backbone beta strands that connect the two domains and pass behind the iron-binding cleft (Fig A.6). The domains move as rigid bodies about this hinge, as is shown by superposition of the individual domains of Fe_2Lf on those of apoLf. When this is done, the rms deviation of the N2 domain (which rotates 54°) is only slightly greater that those for the other three domains (Anderson et al 1990); N1, 0.52 Å; N2, 0.65 Å; C1, 0.43Å; C2, 0.44Å. This means that the domain movement on iron release is a rigid body movement about a hinge. This rigid body movement for lactoferrin is the largest yet reported for any protein.

In serum transferrins, lactoferrins, and ovotransferrin, the C-lobe releases iron less readily, possibly due to its lesser flexibility, whereas in melanotransferrin and *M.sexta* transferrin, it does not bind iron at all. It may be that binding to the C-lobe has remained only in cases where a receptor mechanism exists to extract iron from the site and this idea again raises the question of the biological importance of the bilobal structure.

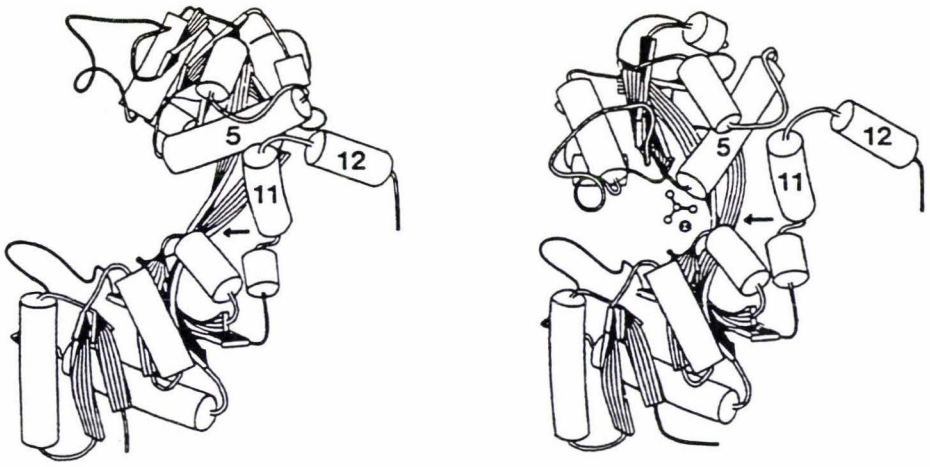

Fig A.6. Schematic diagram of the open (left) and closed (right) forms showing the action of the hinge in the N-lobe of lactoferrin. An arrow positions the part of the beta sheets involved in the "bending" of this region.

A.4. Iron-binding properties.

A.4.1. Introduction.

The orange-red color that develops when ferric iron is added to the transferrins is the main indicator of iron binding. The colour is due to an absorption band around 460nm, although the precise λmax varies among the transferrins. In addition to this band, there is another specific absorption near 295nm (Aisen *et al*, 1969).

One of the first studies to describe the iron-binding properties of lactoferrin was that of Aisen and Leibman (1972). These authors showed that the properties of transferrin and lactoferrin were very similar and suggested that the iron-binding ligands included 2 tyrosines and 1 nitrogen-containing amino acid, possibly histidine. These residues, and an additional aspartate residue, were confirmed when firstly the iron-loaded human lactoferrin structure (Anderson *et al*, 1987), and secondly the iron-loaded rabbit transferrin structure (Bailey *et al*, 1988) were determined.

A.4.2. Binding constant

Although the iron-binding ligands of the two proteins (transferrin and lactoferrin) are the same, there are differences of detail in their iron-binding properties. At pH 6.4, human lactoferrin (hLf) binds iron 300 times more strongly than transferrin (Tf) at pH 6.7 (Aisen and Leibman, 1972), and while hLf releases Fe^{3+} over pH range 4.0 to 2.5, transferrin releases it over the pH range 6.0 to 4.5 (Bullen *et al*, 1978).

The binding for both these proteins is represented by the equation

$$H_6\text{-Protein} + 2Fe^{3+} + 2HCO_3^- \text{ <---> } Fe_2\text{-Protein}(HCO_3^-)_2 + 6H^+$$

Because the binding of iron to any of the transferrins is fully reversible, it should be possible to determine an equilibrium binding constant. This is complicated, however, by the low solubility of Fe^{3+} in near-neutral solution (10^{-17}M at pH 7.0).

A method of equilibrium dialysis, using citrate as a competing iron complexing agent (Aasa *et al*, 1963), was used to overcome this low solubility problem, and therefore measure the equilibrium constant of the iron transferrin complex. This method analyses the amount of iron partitioned between two sides of a membrane with specific-sized pores. On one side only of the membrane is transferrin. Knowing the total iron present, and the amount of free iron (complexed to citrate) on each side, then the amount of iron bound to transferrin and hence the equilibrium constant can be calculated. Using this method, an equilibrium binding constant for lactoferrin at pH of 7.4, in air, was calculated to be 4.7×10^{22} M^{-1} (Harris and Aisen, 1989).

A.4.3. Anion binding.

A distinguishing feature of the transferrins is that an anion is bound with each metal ion. The anion found in vivo is CO_3^{2-} (or possibly HCO_3^-), although other anions than CO_3^{2-} are able to fulfil this role, a common feature of these synergistic anions being that they each contain a Lewis base close to a carboxylate group (Schlabach and Bates, 1975). Metal ion indicators, EPR experiments, NMR spectroscopy, and X-ray crystallography all indicate that the anion is directly attached to the metal ion in the transferrins (Harris and Aisen, 1989).

The ability of transferrin to bind two iron atoms per molecule, and the requirement for bicarbonate were first demonstrated by Schade *et al* (1949). The affinity for Fe^{3+} is weak in the absence of an anion (Bates and Schlabach, 1975), and that for the anion is weak in the absence of Fe^{3+}. Because of this feature, the relationship between the iron and anion is said to be synergistic.

A.4.4. Order of binding.

The next step in understanding the mechanism of iron-binding, was to determine the order of events when a transferrin binds iron. Whether metal or anion binding to transferrin occurs first is not completely clear, but using ^{13}C-NMR, Zweier et al (1981) observed that anions (oxalate and bicarbonate) bind weakly in the absence of metal ion, indicating that anion binding may precede metal binding.

Kinetic data also suggests that the anion binds first (Kojima and Bates, 1981). Using ferric chelate complexes, it has been proposed that there are five steps in the formation of the metal-anion-protein complex. These are (i) binding of the anion to the apoprotein, (ii) detachment of one or more ligands from the added ferric chelate, (iii) formation of a protein-anion-ferric-chelate complex, (iv) loss of the chelate ligand, and (v) conformational change involving the closing of the two domains over the ions. The closed conformation is locked together by the aspartate ligand, which plays a critical role in the metal-bound structure (Anderson *et al*, 1989). This sequence of events, which implies that the initial step involves binding of the anion to its site on the N2 (or C2) domain is shown schematically below (Fig A.7).

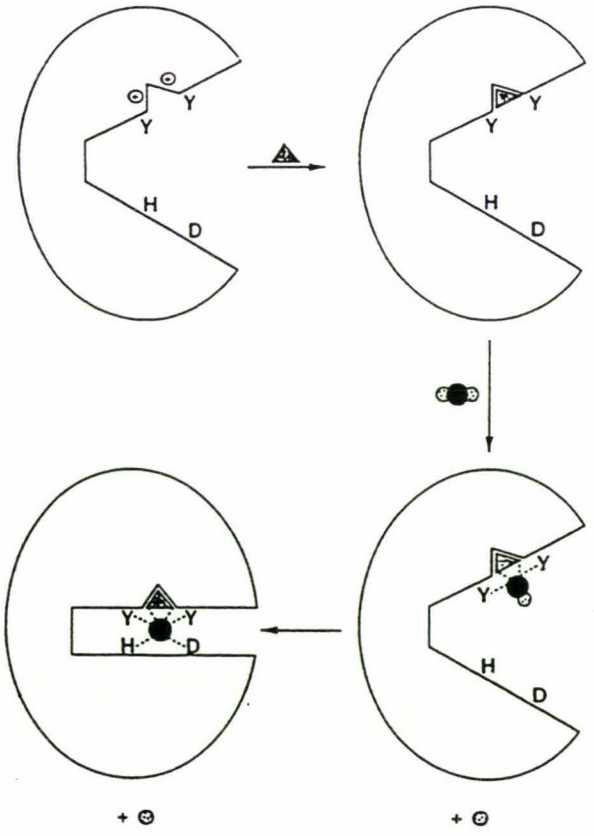

Fig. A.7. A schematic model of the steps involved in the uptake of iron by lactoferrin. (•) Iron; (▲) carbonate; Y, Tyr ligands; H, His ligand; D, Asp ligand; (O) chelate ligands. (From Baker (1994), with permission).

The role of the anion in binding is thought to be twofold. Firstly, it neutralises positive charges on the protein in the binding site which may otherwise repel the cation (Fe^{3+}), and secondly, by adding two more iron ligands (2 carbonate oxygens), it helps form the metal binding site.

A.4.5. Other metals.

While iron (as Fe^{3+}) is bound most strongly to the transferrins, other metal ions can also bind to lactoferrin (Ainscough *et al*, 1980). For one of these, Cu^{2+}, an X-ray crystallographic model of the complex with human lactoferrin has been determined (Smith *et al*, 1992). Although this shows a change in metal coordination from 6-coordinate to 5-coordinate in the N-lobe, the conformation of lactoferrin is still the same, suggesting that lactoferrin with metals other than Fe^{3+} bound, may still bind to receptors just as with Fe^{3+}.

A.4.6. Iron release.

In the mid 1970's, it was found that the pH dependence of iron dissociation from serum transferrin was biphasic, indicating that there was a difference between the two sites. $^{59}Fe^{3+}$ was found to be released completely from one site before the other (Princiotto and Zapolski, 1975), (Cannon and Chasteen, 1975). The N terminal site is acid labile, and the C terminal site is acid stable (Lestas ,1976), so that as the pH of human transferrin is lowered, iron is released from the N-lobe (pH 7 to 5.5) before it is released from the C-lobe (pH 5.5 to 4). The biphasic acid-induced iron release seen in serum transferrin is not seen in lactoferrin, however, in which both sites release iron together over the pH range 4.0-2.5.

The mechanism of iron release for the transferrins is not known, although it is thought to involve one or both of the two following proposals. In the first proposal, Kretchmar and Raymond (1986) showed that for serum transferrin iron was lost more readily from the N-terminal site and proposed that this may be due to the greater flexibility of this lobe. They suggest that conformational changes and electrostatic interactions may play a role in iron release. The interdomain interactions at the back of the binding pocket are known to differ between transferrin and lactoferrin (Baker and Lindley, 1992), and this may account for the difference in the pH at which dissociation occurs.

The second proposal for iron release in the transferrins suggests that a lowering of the pH would cause a protonation of the anion which would either disrupt the interaction between the anion and the anion-binding arginine residue (121 in the N-lobe of hLf, 465 in the C-lobe), or cause a change in anion coordination to the metal, from bidentate to the monodentate, as seen when copper binds to lactoferrin (Smith *et al*, 1992). Either mechanism could result in iron release. This second proposal involving the protonation of the anion, does not, however, explain why human serum transferrin and human lactoferrin release iron at different pH when they both contain the same ligands, Fe^{3+} and anion. Perhaps the environment around the binding site also plays an important role in this second scenario by altering the conditions under which protonation occurs.

Studies of half molecule fragments (one lobe only) suggest that the region at the back of the iron-binding site is important in iron release. The recombinant N-lobe of human lactoferrin differs in this region from the intact molecule, and it releases iron over the pH range 6 to 4.0 (Day *et al*, 1992). This is higher than that of the complete molecule which releases iron between 4.0 and 2.5 (Mazurier and Spik, 1980).

A.4.7. Differences between the lobes.

Given the existence of two similar binding sites on transferrins, four forms should exist, ie the apo form, two monoferric forms, and the diferric form. The existence of monoferric transferrin was first shown by electrophoresis (Aisen *et al*, 1966, Wenn and Williams, 1968) and routine separation is now made possible by electrophoresis of transferrin partially denatured by 6 M urea at pH 8.4. This method, due to Makey and Seal (1976), allows assesment of how much iron is bound to each lobe of the molecule under different conditions.

Using this same method, it was found that in fresh serum, the two sites of transferrin are not equally loaded with iron. Initially the N-lobe is preferentially loaded with iron, although continued incubation at 37°C distributes the iron evenly between the lobes. If the serum is stored at -15°C, then iron distribution is directed towards the C-lobe (Williams and Moreton, 1980).

In 1968, Fletcher and Huehns (1968) suggested that the two sites of iron-binding in transferrin may have different roles. They proposed that one site may be involved in the iron transport function often assigned to transferrin, while the other site may be involved in iron sequestering (antibacterial as for lactoferrin). This was followed by kinetic and thermodynamic work by Aisen *et al* (1978), in which they showed that the C terminal site of transferrin binds iron more strongly than the N terminal site. The

binding constant was 20 times greater. The release of iron also differed, being faster for the N terminal site (which is also more acid-labile, as noted earlier).

Studies of transferrin-receptor interactions (Bali and Aisen, 1991), show that the receptor specifically interacts with the C-lobe of transferrin to release the iron, whereas the N-lobe loses iron by a reduction in pH.

These differences in the two lobes of both human lactoferrin and human serum transferrin may be explained by the structure of lactoferrin (Anderson *et al,* 1990), where comparison of apolactoferrin with the iron-bound form shows that there is a difference in flexibility between the two lobes. It is thought that the greater flexibility of the N-lobe of lactoferrin may help in the binding and release of iron, and this would explain the greater thermodynamic stability and slower release of iron from the C-lobe compared with the N-lobe (Kretchmar and Raymond, 1986).

A.5. Conformational changes in lactoferrin.

A.5.1. Domain movements.

Lactoferrin shares a similar structure, topology, and binding site construction with the group II periplasmic binding proteins (Baker *et al,* 1987) (see Fig A.8). These proteins bind and transport small molecules (sugars, anions, amino acids) through the periplasmic space of gram negative bacteria, and interact with receptors in the bacterial cell wall. The structure of a number have been determined (Quiocho, 1990), and these studies show that they have about 300 residues in a two domain structure similar to a single lobe of the transferrins. In each domain, helices are packed on either face of a central sheet, and two similar extended polypeptides link the domains. Both domains provide ligands for the binding site, with one domain providing most of the groups thereby serving as the initial site of attachment (as seen in the transferrins). For the maltose binding protein (MBP), Spurlino *et al* (1991) have shown that maltose binds first to one domain in the open form.

The feature of lactoferrin structure and function that is addressed in this thesis involves the conformational change seen in the N-lobe of lactoferrin accompanying iron-binding and release. This movement is the largest yet seen in any such protein and involves a rigid body rotation of one domain relative to another, through an angle of 54° (Baker *et al,* 1991). It enables the protein to move between an open form (in which the binding cleft is wide open) and a closed form (in which the domains have closed over the bound metal ion).

Table A.2. lists all the proteins with domain motion for which both open and closed forms have been resolved by X-ray crystallography. Table A.3. lists proteins for which only one conformation is known, but in which a domain closure mechanism is thought to occur.

Table. A.2. Proteins with a hinge-type mechanism for domain closure for which open and closed structures have been determined.

Protein	Reference
Lactoferrin	Anderson *et al*, 1990
Tomato bushy stunt virus coat protein	Olson *et al*, 1983
Maltodextrin binding protein	Sharff *et al*, 1992
Lysine-arginine-ornithine binding protein	Oh *et al*, 1993
T4 lysozyme mutants	Faber and Mathews, 1991
Catabolite gene activator protein	Weber and Steitz, 1987
cAMP dependent protein kinase (catalytic domain)	Karlsson *et al*, 1993
Adenylate kinase	Schulz *et al*, 1990
Glutamate dehydrogenase	Stillman *et al*, 1993
Calmodulin	Meador *et al*, 1992, 1993

Table. A.3. Proteins with a hinge-type mechanism for which only one conformation has been determined.

Protein	Reference
Sulfate binding protein	Luecke and Quiocho, 1990
Phosphate binding protein	Pflugrath and Quiocho, 1988
Leucine binding protein	Gilliland and Quiocho, 1981
Galactose binding protein	Vyas *et al*, 1988, 1991
Arabinose binding protein	Sack *et al*, 1989
Transferrin	Sarra *et al*, 1990
Guanylate kinase	Stehle and Schulz, 1990
Porphobilinogen deaminase	Louie *et al*, 1992

An obvious question concerns why there is such a large domain movement, if the function of lactoferrin is only to sequester an atom of iron for bacteriostatic purposes. A possible answer for this could be the proposed receptor binding function of lactoferrin. Receptor binding is an important feature of the group II periplasmic binding proteins (Mowbray, 1992). For these proteins, the large conformational change is crucial for signal transduction in active transport (Jacobson *et al*, 1992), since the membrane bound receptors preferentially bind to the liganded, closed form of the maltodextrin binding protein (MBP). Furthermore, there are a number of mutations in the N and C domains of MBP (corresponding to the N1 and N2 domains of lactoferrin), at the opening to the cleft that affect the function of this protein.

Two of the periplasmic binding proteins have been analysed crystallographically in both the open and closed forms, ie the maltodextrin binding protein (Spurlino *et al*, 1991, Sharff *et al*, 1992), and the lysine-arginine-ornithine (LAO) binding protein (Oh *et al*, 1993). The nature of the domain movement is similar to that of lactoferrin. MBP has a 35° rotation of one domain relative to the other about an axis through the hinge region, while for LAOBP a movement of 52°, similar to that in the N-lobe of lactoferrin, is seen. This motion, as for lactoferrin (see below), involves only a few large torsion angle changes. In addition, two of the periplasmic binding proteins bind anions in a similar way to lactoferrin, ie the sulphate binding protein (Pflugrath and Quiocho, 1988), and the phosphate binding protein (Luecke and Quiocho, 1990). These two proteins have a similar fold to a single lobe of lactoferrin, and their anion binding sites coincide with the carbonate site in the transferrins. What is interesting is that there is a greater sequence similarity between the sulphate binding protein and lactoferrin

(15%), than between the sulphate and phosphate binding proteins (<10%) (Baker, 1994).

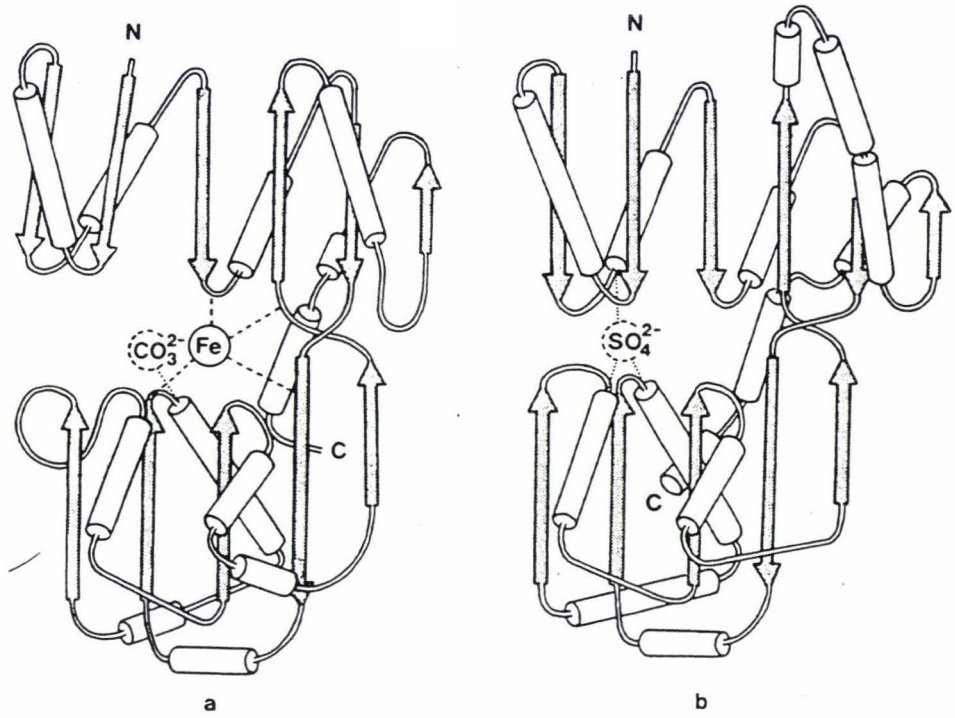

Fig. A.8. Comparison of the folding pattern for the N-lobe of lactoferrin (a) and for the sulphate binding protein (b). (From Baker (1994), with permission).

A.5.2. The hinge of lactoferrin.

The large rigid body motion in lactoferrin is made possible by two extended polypeptide strands that run behind the binding cleft connecting the two domains of each lobe.

In Fig A.9, the hinges in the N-lobe of human lactoferrin are indicated. Gerstein *et al* (1993) described the motion as a screw motion by fixing the origin so that translation was minimal upon domain movement. In this description, the N2 domain translates only 1.0Å with the rotation axis passing very close to Thr 90 and Pro 251, the two residues Anderson *et al* (1990) identified as the centre of the hinges; the axis is within 1.4Å of the Cα of the Thr 90, and 2.4Å of the Cα of Pro 251. Therefore the motion of N2 relative to N1 involves almost no translation, and is instead a pure rotation about the central residue in each hinge (Thr 90 and Pro 251). This is illustrated in Fig. A.10.

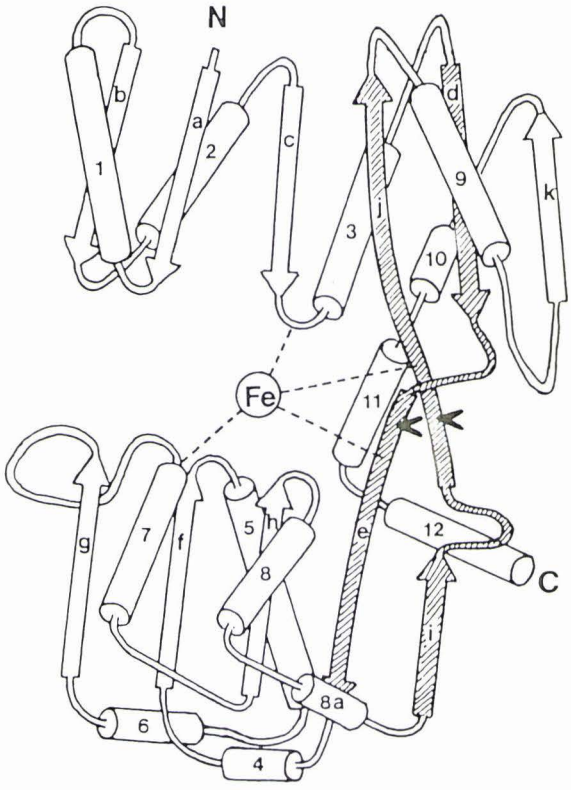

Fig. A.9. Polypeptide folding pattern for the N-lobe of lactoferrin indicating the two beta strands responsible for the position of the hinge. (From Baker (1994), with permission).

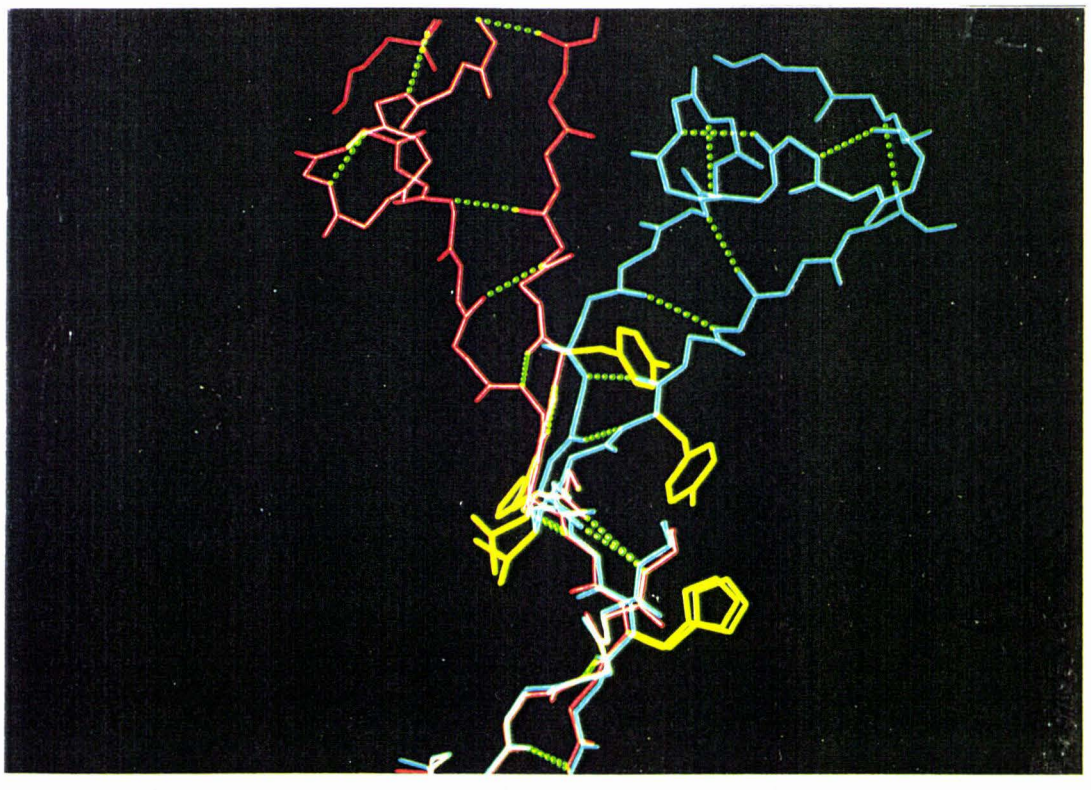

Fig. A.10. Main chain atoms only of the open (red) and closed (blue) structures of the two beta strands involved in the hinge of lactoferrin.

In the first hinge, the domain movement is associated with torsion angle changes in ψ (90) and ϕ (91), these changes being coupled across a peptide bond (Fig A.11). In total these produce a rotation of 81° , which, although greater than the overall 54° rotation seen in the N-lobe, is compensated by small opposing changes in the torsions of neighbouring residues (Gerstein *et al*, 1993).

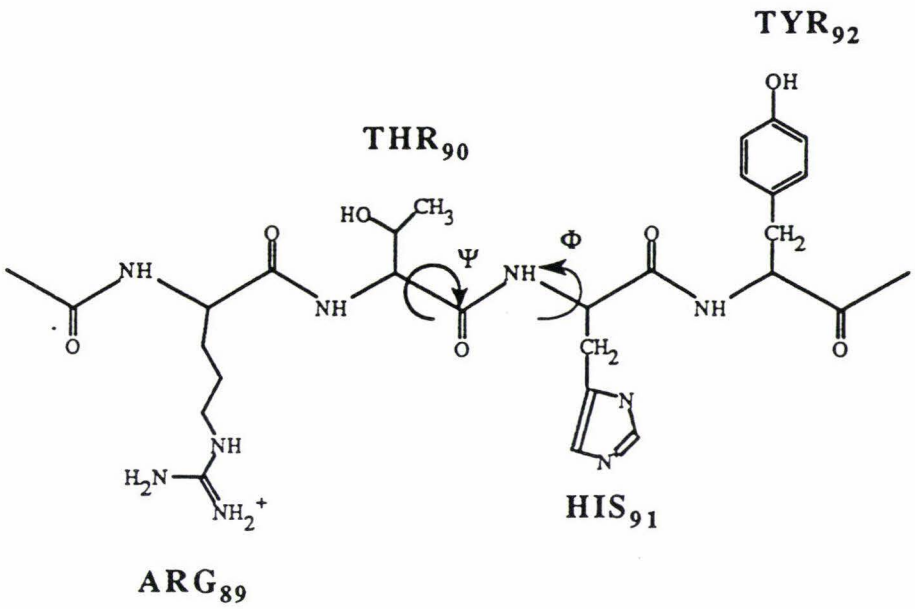

Fig. A.11. Residues 89 - 92 of the first beta strand involved in the hinge motion of the N-lobe of lactoferrin. The phi and psi angles responsible for the motion are indicated.

The second hinge is not so simple. Three torsion angles have changes greater that 20°; ie ϕ (250), ψ (250), and ϕ (251), with the central angle (ψ (250)) changing by 33 degrees. These angles are also coupled across a peptide bond to produce a rotation of 49° (Fig A.12).

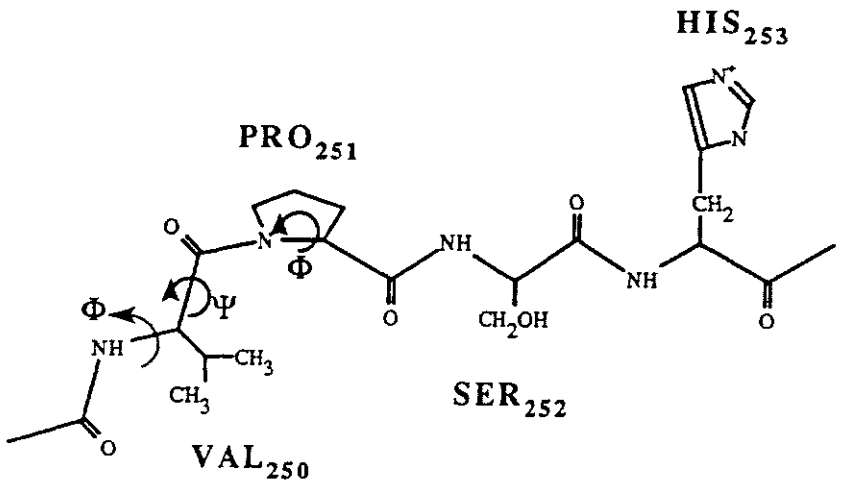

Fig. A.12. Residues 250 - 253 of the second beta strand involved in the hinge motion of the N-lobe of lactoferrin. The phi and psi angles involved in the motion are indicated.

The axis of the overall 54° rotation is closely aligned with the rotational axis of these principal torsion angle changes. This means that the local motion of the hinge reflects the overall domain movement of the molecule.

The principal torsion angle changes in the lactoferrin hinge occur in normally allowed regions of the Ramachandran plot (Fig A.13) and therefore involve only low energy transitions, indicating that the open and closed states can occur in dynamic equilibrium as proposed for the periplasmic binding proteins (Oh *et al*, 1993).

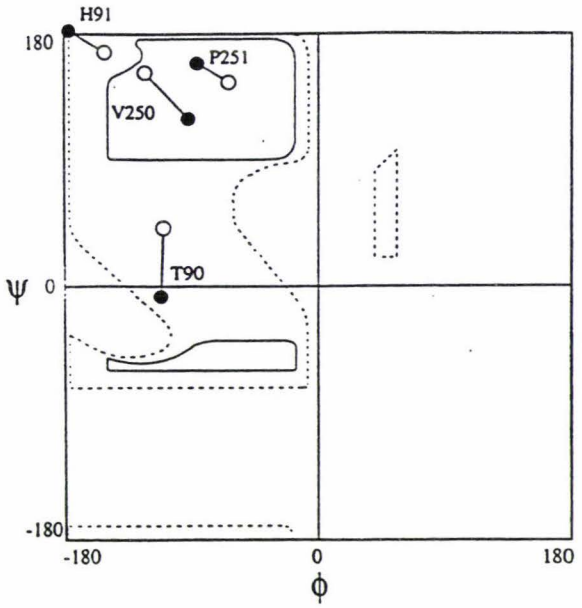

Fig. A.13. Ramachandran plot of the main chain torsion angle changes seen in the N-lobe of lactoferrin. Phi and psi angles are shown with open circles for the open conformation, and with filled circles for the closed conformation.

In the ribose and galactose binding proteins however, there is a residue (215 in RBP, 236 in GBP) that has an unfavourable conformation. This is one of only two Ramachandran violations in these proteins, and the conformation is necessary for the correct action of the hinge in these proteins (Mowbray, 1992). A buried aspartate helps maintain this conformation and therefore the correct action of the hinge. In human lactoferrin, the one Ramachandran violation is not associated with the hinges.

The location of the hinges with respect to the ligands seems likely to be important in lactoferrin. Tyr 92 (a ligand) is 2 residues away from Thr 90 (human lactoferrin numbering), and His 253 is two residues away from Pro 251. This allows the hinges to split these two ligands, with Tyr 92 moving with the N2 domain, and His 253 with the N1 domain. This provides the N2 domain with 4 of the 6 iron-binding groups (2 from the carbonate anion together with Tyr 92 and Tyr 192). The N2 domain is thought to serve as the initial site of attachment. This is supported by recent work by Lindley et al (1993) on the analysis of a proteolytic fragment of duck ovotransferrin corresponding to its N2 domain. In this one-domain molecule, the iron is bound to the carbonate ion and to two tyrosine residues just as in the intact molecule. Presumably

once the iron is bound to the N2 domain, domain closure occurs to allow the iron to complete its coordination to His 253 and Asp 60, and the aspartate helps lock the two domains together with an interdomain hydrogen bond (Anderson *et al*, 1989).

It has been proposed that the apo-protein is in a dynamic equilibrium, with the N1 and N2 domains "flexing" about the hinge, and that the iron then stabilises the closed form as the bound ligand does in the periplasmic binding proteins. The importance of the hinge in this mechanism would be in directing the two domains together into the correct position by restraining the number of conformations in the open state.

In most proteins, the main chain is buried beneath layers of side chains leaving little freedom for large torsion angle changes. In the hinge region of lactoferrin,however, the main chain is free from steric constraint, with the mainchain atoms of Thr 90, His 91, and Pro 251 making no contacts with the rest of the protein. The absence of main chain packing constraints with the rest of the protein, and the absence of internal hydrogen bonds is also found in hinged mechanisms in lactate dehydrogenase (Gerstein and Chothia, 1991) and in adenylate kinase (Gerstein *et al*, 1993) and appears to be a crucial structural requirement for hinged motion.

A.5.3. Amino acid sequences in the hinges.

The sequences of the hinge regions in transferrins are listed in Table A.4. and Table A.5. and display a remarkable degree of sequence conservation.

In the N-lobe, the hinges are located at Thr 90 in one "backbone" strand and Pro 251 in the other. Elsewhere in the two strands, a number of residues are conserved either because they are ligands (Tyr 92 and His 253) or because they contribute to the hydrophobic core of one domain or the other (the sequences AVAVV in the first strand, and AVV in the second strand). The most intriguing conservation, however, is that of Thr 90 and Pro 251.

Table A.4. Amino acid sequence in the two hinge strands of the N-terminal lobe, for all known transferrin sequence.

Protein	N-lobe	
	1st Hinge Strand	2nd Hinge Strand
	90	251
Human Serotransferrin	EDPQTFYYAVAVVKK	HLAQVPSHTVVAR
Horse Serotransferrin	TEPQTHYYAVAVVKK	YLASIPSHAVVAR
Pig Serotransferrin	DNPGTHYYAVAVVKK	YLAQVPSHAVVAR
Rabbit Serotransferrin	ENPKTFYYAVALVKK	HLARVPSHAVVAR
Rat Serotransferrin	EHRQTHYLAVAVVKK	YLARIPSHAVVAR
Xenopus Serotransferrin	TETDTCYYAVAVVKK	NLAKVPAHAVLTR
Human Lactoferrin	RQPRTHYYAVAVVKK	HLARVPSHAVVAR
Bovine Lactoferrin	ESPQTHYYAVAVVKK	HLAQVPSHAVVAR
Goat Lactoferrin	KSPQTHYYAVAVVKK	HLAQVPSHAVVAR
Mouse Lactoferrin	EQPRTHYYAVAVVKN	HLAQVPSHAVVSR
Pig Lactoferrin	ENPGTYYYAVAVVKK	HLARVPSHAVVAR
Human Melanotransferrin	QEVGTSYYAVAVVRR	HLARVPAHAVVVR
Chicken Ovotransferrin	EGSTTSYYAVAVVKK	NWARVAAHAVVAR
Atlantic Salmon Transferrin	EDSDTCYYAVAVAKK	HLARVPAHAVVSR
Hornworm Transferrin	PDAPFRYEAVIVVHK	SWAARPWQGLIGH
Cockroach Transferrin	PDEEFRYEAVCVIHK	IWAARPWQGYMAN

Thr 90 is totally conserved, except in the two insect transferrins. It is difficult to see why this should be so unless this residue is important for the function of transferrins in general. The obvious conclusion is that Thr 90 has some as yet unsuspected role in helping to define the conformational change, since it is at the exact location of one of the hinges. The only apparent interaction made by the sidechain of Thr 90 is a hydrogen bond with the C-terminal helix of the full-length molecule (90 OG1 - O=C 691), this being present in apolactoferrin but not in Fe_2Lf.

Pro 251 is also totally conserved, except in chicken ovotransferrin. Proline residues are known for their restraining role in the polypeptide conformation, and are present in the hinge regions of several bacterial binding proteins where their presence may modulate the motion of the hinge (Mowbray ,1992).

Table A.5. Amino acid sequence of the two hinge strands of the C-terminal lobe, for all known transferrin sequences.

Protein	C-lobe 1st Hinge Strand	2nd Hinge Strand 595
Human Serotransferrin	DTPEAGYFAVAVVKK	HLAMAPNHAVVSR
Horse Serotransferrin	DTPEEGYHAVAVVKS	YLARAPNHAVVSR
Pig Serotransferrin	NTPEKGYLAVAVVKK	HLARAPNHAVVAR
Rabbit Serotransferrin	KAPEEGYLSVAVVKK	HLAKAPNHAVVSR
Rat Serotransferrin	QSDVFPKGYYAVAVV	HLAQAPNHVVVSR
Xenopus Serotransferrin	SQAKGNYYAVAIVKK	NLAEVPAHAVVTL
Human Lactoferrin	DRPVEGYLAVAVVRR	HLAMAPNHAVVSR
Bovine Lactoferrin	LRPTEGYLAVAVVKK	HLAVAPNHAVVSR
Goat Lactoferrin	LRPTEGYLAVAVVKK	YLAVAPNHAVVSR
Mouse Lactoferrin	NRPVEGYLAVAAVRR	HLAIAPNHAVVSR
Pig Lactoferrin	HRPTGGYFAVAVVRK	HLAVAPSHAVVSR
Human Melanotransferrin	EDSSNSYYVVAVVRR	NLAQIPPHAVMVR
Chicken Ovotransferrin	ERPA SYFAVAVARK	NLAEVPTHAVVVR
Atlantic Salmon Transferrin	PGEASSYYAVAVAKK	HLAKVPAHAVITR
Hornworm Transferrin	GELKTPNYAVAVVKK	
Cockroach Transferrin	GEHGSLYYAVAVVRK	

In the C-lobe there is no conserved residue equivalent to Thr 90 in the first strand, although there is a totally conserved proline equivalent to Pro 251 in the second strand. In the C-lobe, however, the exact location of the hinge is not yet known.

Oh et al (1993) have suggested that the residues in the hinge do not play any part in the domain movement. The conservation of hinge residues seen in lactoferrin sequences argues against this view, however. In order to test this apparent contradiction, an obvious approach is to mutate the two conserved residues in the hinge (Thr 90 and Pro 251) and to characterise the resulting proteins by iron-binding studies and x-ray crystallography.

A.6. Aims of this project.

The aims of this project were to address the roles of the two "hinge" residues, Thr 90 and Pro 251, by site-directed mutagenesis and characterisation of the resultant mutant proteins.

The cDNA from human lactoferrin had previously been cloned into a suitable expression vector and expression had been shown to occur (Day, 1993). The complete structure for the half-length molecule had been solved by molecular replacement by Day (1993) using the Fe_2Lf structure as a starting model (Anderson *et al*, 1989). Bates (1994) had also cloned a segment of the human lactoferrin cDNA encoding amino acids 250 to 688 into the M13 bacteriophage.

In order to introduce amino acid changes into the cDNA for human lactoferrin, the method of Kunkel (1989) was chosen. Changes made in the M13 bacteriophage containing this segment were to be cloned into the pNUT vector containing the half-length lactoferrin cDNA (pNUT:Lf_N) constructed by Day (1993). This would then be used in tissue culture to produce the recombinant proteins. After purification of the recombinant protein, their iron-release properties would be examined, and the structural effects of mutation determined by x-ray crystallography.

The changes at Pro 251 were designed to allow varying degrees of flexibility into this region. For this reason, Gly, Ala and Val were chosen. Asp was also chosen as this is convenient using a degenerate oligonucleotide, and this change also introduces a negative charge into this region.

The changes at Thr 90 involved a change to alanine in order to determine the importance of the proposed hydrogen bond between the side chain hydroxyl group of Thr 90 and C-terminal end of the polypeptide chain.