Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Lactoferrin–metal interactions: first crystal structure of a complex of lactoferrin with a lanthanide ion (Sm^{3+}) at 3.4 Å resolution

Ashwani K. Sharma and Tej P. Singh*

Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110 029, India

Correspondence e-mail: tps@medinst.ernet.in

Received 13 March 1999
Accepted 19 July 1999

PDB Reference: lactoferrin–Sm^{3+} complex, 1qjm.

Lactoferrin is an important member of the transferrin family. A characteristic property of transferrins is their ability to bind very tightly ($K_{app} \simeq 10^{20}$) but reversibly two Fe^{3+} ions. The structural consequences of binding a metal other than Fe^{3+} have been examined by crystallographic analysis at 3.4 Å resolution of mare samarium–lactoferrin (Sm_2Lf). The structure was refined to an R factor of 0.219 for 8776 reflections in the resolution range 17.0–3.4 Å. The samarium geometry (distorted octahedral coordination) is similar in both lobes. However, the anion interactions are quite different in the two lobes. In the N lobe, the anion is able to form only two hydrogen bonds instead of the four observed in the C lobe of Sm_2Lf and the six observed in Fe_2Lf. This is because Arg121, Thr117 and Gly124 have moved away from the anion as a consequence of the binding of the Sm^{3+} ion. The protein ligands in the binding cleft of Sm_2Lf show large displacements, but the overall protein structure remains the same. The binding of Sm^{3+} by lactoferrin shows that the protein is capable of sequestering ions of different sizes and charges, though with reduced affinity. This conclusion should be true of other transferrins also.

1. Introduction

Proteins of the transferrin family, which include serum transferrin and lactoferrin, control the levels of free iron in animals through their very tight ($K \simeq 10^{20}$) but reversible binding of iron. These proteins fold into two lobes which are connected by a 10–12 residue long peptide. Each lobe is further divided into two domains, forming a deep binding cleft between them. Each binding cleft can house one Fe^{3+} and the synergistic CO_3^{2-} ion. In addition to this very high affinity for Fe^{3+} and CO_3^{2-}, these proteins can also bind a wide variety of other cations and anions at the same specific sites. Metal ions of varying sizes and oxidation states ranging from +2 to +5 can be accommodated, including Cu^{2+}, Co^{3+}, Mn^{3+} (Ainscough *et al.*, 1979), VO^{2+} (Chasteen *et al.*, 1977), VO_2^{+} (Harris & Carrano, 1984), lanthanide cations (Luk, 1971; Zak & Aisen, 1988; O'Hara & Bersohn, 1982) and Th^{4+} (Harris *et al.*, 1981)

Crystallographic analysis of rabbit serum transferrin (Bailey *et al.*, 1988), diferric human lactoferrin (Anderson *et al.*, 1989; Haridas *et al.*, 1995), hen ovotransferrin (Kurokawa *et al.*, 1995), duck ovotransferrin (Rawas *et al.*, 1996), bovine lactoferrin (Moore *et al.*, 1997), mare lactoferrin (Sharma, Paramasivam *et al.*, 1999) and buffalo lactoferrin (Karthikeyan *et al.*, 1999*a*,*b*) have defined the location and nature of the iron sites. All these proteins have essentially the same bilobal structure, with one iron site in each lobe, deep in a cleft between two domains. The iron ligands (two Tyr, one His, one

Table 1
Crystallographic data statistics.

Space group	$P2_12_12_1$
a (Å)	80.5
b (Å)	103.2
c (Å)	113.2
V_m (Å^3 Da^{-1})	2.88
Total number of reflections collected	75612
Total number of unique reflections	8782
Solvent content (%)	57
Z	4
Resolution limits (Å)	17.0–3.4
Completeness (%)	85
Completeness in the highest resolution shell (3.8—3.4 Å) (%)	55
R_{sym} for all data	0.126
R_{sym} in the highest resolution shell	0.263
Overall $I/\sigma(I)$	11.1
$I/\sigma(I)$ for the last resolution shell	3.1

Table 2
Final statistics at the end of the refinement.

Resolution limits (Å)	17.0–3.4
R factor for all data	0.226
R factor	0.219
R_{free}	0.316
Number of reflections used	8782
Number of protein atoms	5281
Ions	2 Sm^{3+}, 2 CO_3^{2-}
R.m.s. deviations from ideal values	
Bond lengths (Å)	0.018
Bond angles (°)	2.3
Dihedral angles (°)	25.5
Improper angles (°)	1.70

Table 3
Comparison of metal–ligand coordination distances between disamarium and diferric lactoferrins.

N lobe			C lobe		
	Bond length (Å)			Bond length (Å)	
Atoms	Sm_2Lf	Fe_2Lf	Atoms	Sm_2Lf	Fe_2Lf
Metal—O60	2.6	2.0	Metal—O395	2.9	2.2
Metal—O92	2.6	1.9	Metal—O433	2.8	1.9
Metal—O192	2.5	1.9	Metal—O526	2.8	1.9
Metal—N253	2.8	2.3	Metal—N595	2.9	2.5
Metal—O1	2.6	2.0	metal—O1	2.6	1.9
Metal—O2	2.6	1.9	Metal—O2	2.5	2.5

Asp and a bidentate carbonate anion) are the same in every case and binding geometries are also grossly similar.

A crystallographic analysis of human copper–lactoferrin (involving divalent Cu^{2+}; Smith *et al.*, 1992) has revealed that the copper coordination is different in the two binding sites. In the N lobe, the geometry is square pyramidal with a monodentate anion, while in the C lobe the geometry is a distorted octahedron with an asymmetrically bidentate anion. In spite of the above differences in the metal coordinations, the overall protein structure does not change significantly. Nevertheless, the observed differences between copper (a divalent cation) and iron (a trivalent cation) coordinations illustrate the potential danger in extrapolating the structural conclusions from one metal to another. Similarly, larger cations such as the lanthanide ions may not be accommodated in the binding clefts without significant perturbation in the protein structure, at least in and around the binding sites. In order to generalize the mechanism of metal ion–lactoferrin interactions and to correlate the structural changes in the lactoferrin molecule with the size and charge density of the metal ion, we have analysed the crystal structure of a disamarium lactoferrin (Sm_2Lf) at 3.4 Å resolution. Admittedly, it would have been desirable to work at a higher resolution. However, as repeated attempts to grow better diffracting crystals did not succeed, the effort here has been to extract maximum possible information from the available data through careful analysis.

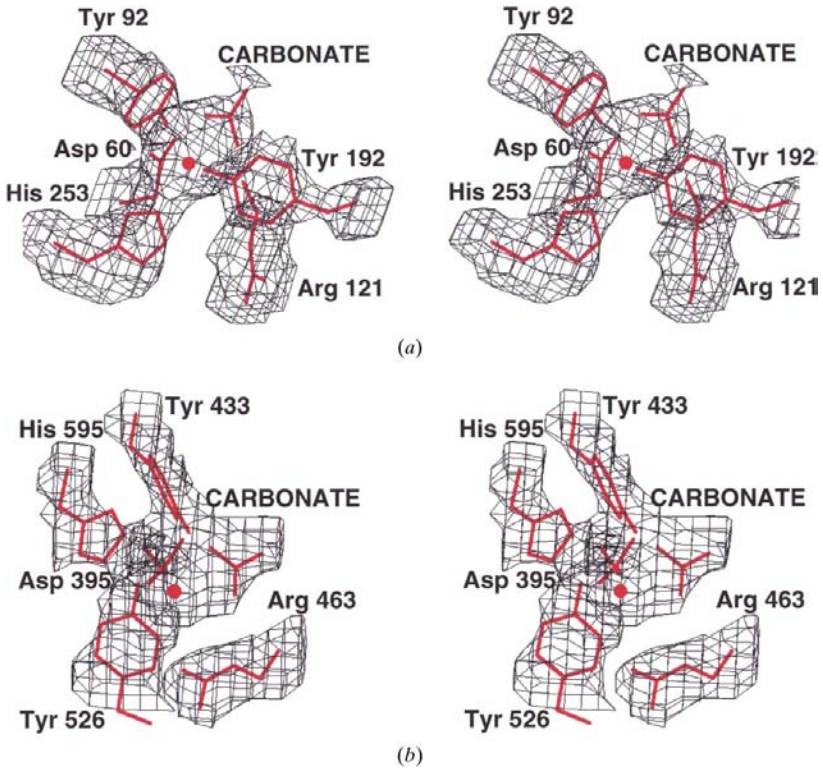

Figure 1
Stereoview of the $F_o - F_c$ electron-density map of (*a*) N-lobe metal and anion-binding site of disamarium mare lactoferrin, (*b*) C-lobe metal and anion-binding site of disamarium mare lactoferrin. The residues indicated in the figures were not included in phase calculations.

2. Materials and methods

2.1. Purification of lactoferrin

Mare milk was obtained from the National Research Centre on Equines, Hisar (India). The purification was carried out using a locally modified procedure (Sharma *et al.*, 1996). Diluted colostrum/milk was defatted by skimming. Skimmed milk was diluted twice with

0.05 *M* Tris–HCl pH 8.0. CM-Sephadex C-50 was added (7 g l^{-1}) and stirred slowly using a mechanical stirrer for 1 h. The gel was allowed to settle and the milk was decanted. The gel was washed with an excess of 0.05 *M* Tris–HCl pH 8.0, packed into a column (25 × 2.5 cm) and then washed with the same buffer containing 0.1 *M* NaCl which facilitated the removal of impurities. The lactoferrin was then eluted with the same buffer containing 0.25 *M* NaCl. The protein solution was dialysed against an excess of triple-distilled water. The protein was again passed through a CM-Sephadex C-50 column (10 × 2.5 cm) pre-equilibrated with 0.05 *M* Tris–HCl pH 8.0 and eluted with a linear gradient of 0.05–0.3 *M* NaCl in the same buffer. The protein was concentrated using an Amicon ultrafiltration cell. The concentrated protein was passed through a Sephadex G-100 column (100 × 2 cm) using 0.05 *M* Tris–HCl buffer pH 8.0. The purity of the lactoferrin was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE; Laemmli, 1970).

2.2. Preparation of Sm_2Lf

Iron was removed from the purified native lactoferrin using the procedure of Mason *et al.* (1968). The purified lactoferrin solution (1%) in 0.05 *M* Tris–HCl pH 8.0 buffer was dialyzed against an excess of 0.1 *M* citric acid with regular changes every 6 h. Finally, citric acid was removed by dialysis against an excess of distilled water with regular changes for 24 h at 277 K. The colourless apo-protein obtained was lyophilized. This mare apolactoferrin was used for the preparation of disamarium lactoferrin. Samarium oxide (Sm_2O_3) was added in excess to the solution of apolactoferrin in 0.1 *M* sodium bicarbonate–sodium citrate buffer pH 8.0 to saturate the protein with metal ions. The product formed was dialyzed against 0.025 *M* Tris–HCl pH 8.0 and lyophilized.

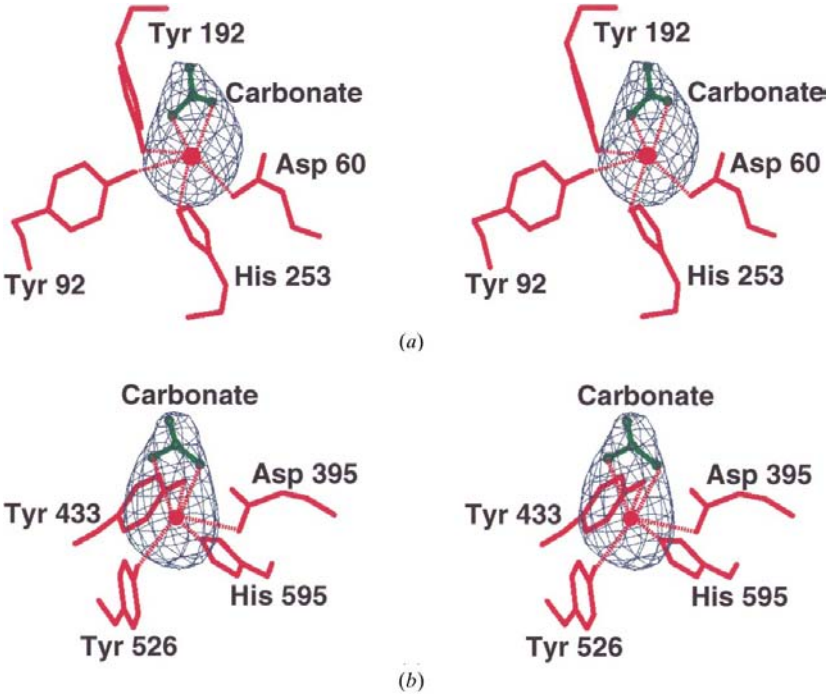

Figure 2
Stereoview of the $F_o - F_c$ electron-density map for metal and carbonate ions (*a*) in the N lobe, (*b*) in the C lobe. Contours are drawn at the 3.0σ level.

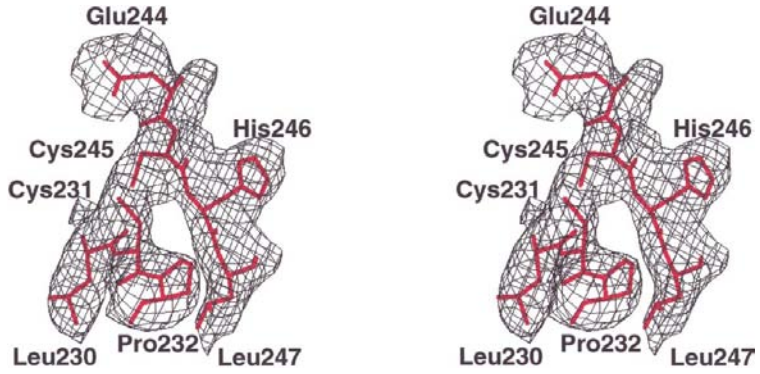

Figure 3
Stereoview of a section of the final $2F_o - F_c$ electron-density map contoured at 1.0σ.

2.3. Crystallization

The complex was immediately used for crystallization. Crystals suitable for X-ray diffraction were obtained by microdialysis. The protein solution, at a concentration of 40 mg ml^{-1} in 0.025 *M* Tris–HCl, was equilibrated against the same buffer containing 10%(*v*/*v*) ethanol pH 8.0. Crystallization experiments were carried out at 277 K. The crystals grew in two weeks to dimensions of 0.7 × 0.4 × 0.2 mm.

2.4. Data collection

Crystals were mounted in 0.7 mm diameter glass capillaries. X-ray intensities were measured at 288 K using a MAR Research imaging-plate scanner with a diameter of 300 mm. The crystal-to-detector distance was kept at 300 mm. Monochromatic Cu $K\alpha$ radiation was produced using a graphite monochromator mounted on a Rigaku RU-200 rotating-anode X-ray generator operating at 40 kV and 100 mA with a focal point of 0.3 × 3 mm. 110 1° rotation images were collected. The exposure time was fixed at 900 s. The data were collected using only one crystal, which diffracted to 3.4 Å resolution. The *HKL* package (Otwinowski, 1993; Minor, 1993), *MARXDS* and *MARSCALE* (Kabsch, 1988) were used for the determination of unit-cell parameters, data processing and scaling of the data. The details of crystallographic data and data collection are given in Table 1.

Table 4
The metal–ligand bond angles in disamarium lactoferrin.

N lobe	Angle (°)	C lobe	Angle (°)
Asp60 OD1—Sm—Tyr92 OH	107	Asp395 OD1—Sm—Tyr433 OH	73
Asp60 OD1—Sm—Tyr192 OH	145	Asp395 OD1—Sm—Tyr526 OH	106
Asp60 OD1—Sm—His253 NE2	85	Asp395 OD1—Sm—His595 NE2	66
Asp60 OD1—Sm—O1	103	Asp395 OD1—Sm—O1	125
Asp60 OD1—Sm—O2	106	Asp395 OD1—Sm—O2	72
Tyr92 OH—Sm—Tyr192 OH	67	Tyr433 OH—Sm—Tyr526 OH	98
Tyr92 OH—Sm—His253 NE2	77	Tyr433 OH—Sm—His595 NE2	52
Tyr92 OH—Sm—O1	70	Tyr433 OH—Sm—O1	72
Tyr92 OH—Sm—O2	123	Tyr433 OH—Sm—O2	83
Tyr192 OH—Sm—His253 NE2	68	Tyr526 OH—Sm—His595 NE2	70
Tyr192 OH—Sm—O1	90	Tyr526O H—Sm—O1	120
Tyr192 OH—Sm—O2	108	Tyr526 OH—Sm—O2	141
His253 NE2— Sm—O1	147	His595 NE2—Sm—O1	124
His253 NE2—Sm—O2	151	His595 NE2—Sm—O2	83
O1—Sm—O2	57	O1—Sm—O2	53

2.5. Structure determination and refinement

The crystals of disamarium lactoferrin were isomorphous with native diferric lactoferrin crystals (Sharma, Paramasivam *et al.*, 1999). The refined atomic coordinates of the diferric lactoferrin structure were used as the starting model for refinement, after removing Fe^{3+} and CO_3^{2-} ions, the solvent molecules, protein side chains involved in metal binding (four in each binding site) and anion-binding protein side chains (Arg121 and Arg463, Thr117 and Thr459). All these side chains were mutated to Ala. The initial model was built using *O* (Jones *et al.*, 1991). Refinement was carried out using the rigid-body conjugate-gradient minimization, simulated-annealing and *B*-factor refinement protocols of *X-PLOR* (Brünger *et al.*, 1987) using the Engh and Huber parameter libraries (Engh & Huber, 1991). Initially, rigid-body minimization was performed by considering four rigid groups corresponding to four domains in the molecule. This reduced the *R* factor to 0.38. The sequence of the protein was compared with the map and the side chains were introduced into electron density without much ambiguity. The electron density improved after each cycle of refinement and the *R* factor dropped gradually. Repeated steps of model building using $2F_o - F_c$ and $F_o - F_c$ maps with *O* reduced the *R* factor to 0.28. In view of the limited resolution, a conservative *B*-factor refinement was performed. The global *B* factor was initially refined. This was followed by the refinement of one *B* factor per residue. So far, the model did not include the side chains of protein ligands and Sm^{3+}. The $F_o - F_c$ map calculated at this stage showed good electron density at the samarium-binding sites, to which Sm atoms and carbonate ions and protein side chains involved in metal binding were fitted in both lobes (Figs. 1*a* and 1*b*). The positions of Sm atoms and carbonate ions refined with distance constraints were checked by omit maps (Fig. 2*a* and 2*b*). Further model building and positional refinement was also carried out with distance constraints. This yielded an *R* factor of 21.9% (R_{free} = 32.6%). The entire data set, including the reflections which were set aside for calculating R_{free}, was then used in 40 cycles of positional refinement, which yielded a final *R* factor of 22.6%. A representative stereo plot of the final electron-density map ($2F_o - F_c$) is illustrated in Fig. 3. The quality of the protein model was assessed by the program *PROCHECK* (Laskowski *et al.*, 1993). A Ramachandran plot (Ramachandran & Sasisekaran, 1968) showed that 93% of the φ, ψ angles were in the most favoured and additionally allowed regions. A summary

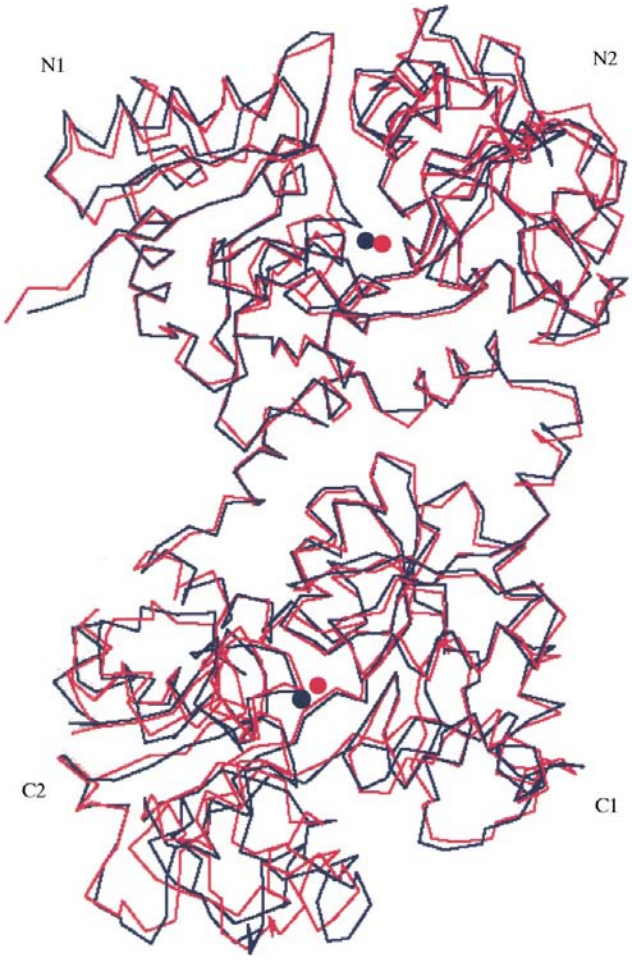

Figure 4
Superposition of C^{α} tracing of disamarium (dark blue) and diferric (red) mare lactoferrins. The r.m.s. deviation for the C^{α} atoms is 1.2 Å.

Table 5
Comparison of anion hydrogen-bonding distances between disamarium and diferric lactoferrins.

N lobe			C lobe		
	Bond length (Å)			Bond length (Å)	
Hydrogen bond	Sm_2Lf	Fe_2Lf	Hydrogen bond	Sm_2Lf	Fe_2Lf
O1···N (123)	3.1	3.0	O1···N (465)	3.1	2.7
O2···NE (121)	4.5	2.9	O2···NE (463)	3.3	2.9
O2···NH2 (121)	5.8	2.9	O2···NH2 (463)	4.3	2.8
O3···N (123)	2.9	3.4	O3···N (465)	2.8	3.2
O3···N (124)	3.4	3.1	O3···N (466)	2.9	3.0
O3···OG1 (117)	5.0	2.7	O3···N (459)	5.6	3.0

of the refined model and the relevant geometrical parameters is given in Table 2.

3. Results and discussion

3.1. General description

The final atomic coordinate set is comprised of 5281 protein atoms from 689 amino-acid residues, two Sm^{3+} ions and two CO_3^{2-} ions. The polypeptide fold is closely similar to diferric mare lactoferrin (Fe_2Lf; Sharma, Paramasivam *et al.*, 1999). The superposition of Sm_2Lf on Fe_2Lf gives a root-mean-square (r.m.s.) deviation of 1.2 Å for all 689 C^{α} atoms (Fig. 4). The average *B* factor for the Sm_2Lf structure is 39.6 $Å^2$, which is comparable with the values of 34.5 and 35.0 $Å^2$ observed in diferric mare lactoferrin (Sharma, Paramasivam *et al.*, 1999) and mare apolactoferrin (Sharma, Rajashankar *et al.*, 1999), respectively. The distribution of *B* values along the polypeptide chain follows the expected pattern, with regular secondary structures and internal regions having lower *B* values and external loops having higher *B* values. The average *B* values of the iron-binding ligands in the N and C lobes are 28.0 and 28.5 $Å^2$, respectively. There are only a few regions where the electron density is not well defined. The first two residues of the N-terminus are not clearly visible. The loops 83–88 in the N lobe and 418–424 in the C lobe are poorly defined.

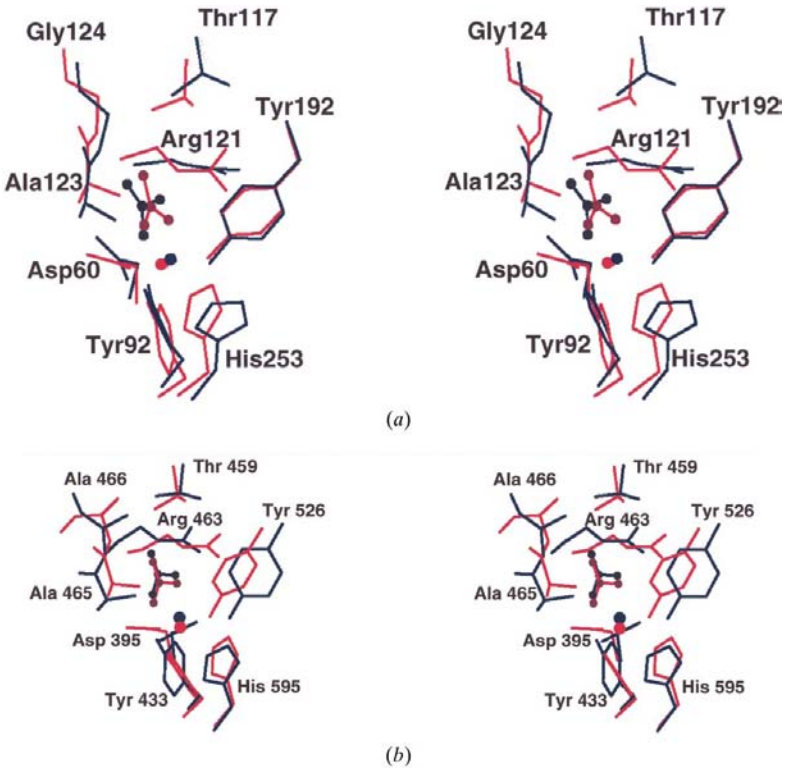

Figure 5
(*a*) Superposition of the metal- and anion-binding site in the N lobe of disamarium (dark blue) and diferric (red) mare lactoferrins. In disamarium lactoferrin, samarium is displaced by 0.7 Å from the position of the Fe atom in the diferric form. Arg121 and Thr117 have shifted away substantially from the carbonate anion and do not interact with the carbonate anion. (*b*) Superposition of the metal- and anion-binding site in the C lobe of disamarium (dark blue) and diferric (red) mare lactoferrins. In disamarium lactoferrin, samarium is displaced by 1.5 Å from the position of the Fe atom in the diferric form. Arg463 and Thr459 do not interact with the carbonate anion.

3.2. Metal and anion sites

The samarium geometries (distorted octahedral coordinations) are similar in the two sites. A list of coordination distances is given in Table 3 and stereoviews of the two sites superimposed on the corresponding portions of the iron-binding sites of Fe_2Lf are shown in Figs. 5(*a*) and 5(*b*). The Sm atom has been found to be shifted by 0.7 Å in the N lobe and more drastically by 1.5 Å in the C lobe from the corresponding positions of the Fe^{3+} ion in the diferric mare lactoferrin structure. In general, the metal–ligand coordination distances are larger in Sm_2Lf compared with the diferric structure (Table 3). The metal–ligand bond angles are given in Table 4. As seen from Fig. 5, the protein ligands have rearranged themselves with considerable shifts in both sites to accommodate the larger samarium cation.

The carbonate anion interacts with samarium symmetrically in both lobes. The anion hydrogen bonding in the N lobe is weaker than that observed in the C lobe (Table 5). It is observed that the Arg121 (463 in C lobe) has shifted away substantially from the carbonate ion. The distance between O2 and the arginine guanidinium group has increased to 4.5 Å. The corresponding distance in Fe_2Lf is 2.9 Å. The distance between O3 and Gly124 N is also significantly increased. Similarly, O3 and Thr117 OG1 are distant from each other in Sm_2Lf, whereas the corresponding distance in Fe_2Lf is a hydrogen-bonding distance. These values clearly show that the overall anion–protein interactions are considerably weaker in Sm_2Lf than in Fe_2Lf.

The present work suggests that the combined binding strength of samarium and carbonate ions is weaker than that observed for the combination of iron and carbonate ions in Fe_2Lf. The pH-dependent metal binding and release experiments have consistently indicated the release of samarium ion

at a slightly higher pH value (pH 3–4) than ferric ion (pH 2–3). It is remarkable that the protein ligands are able to rearrange to accommodate a larger samarium cation in the binding cleft without perturbing the overall structure of the protein. Although the binding clefts in lactoferrin molecule are well designed for Fe^{3+}, the protein side chains involved in the metal and anion bindings are flexible enough to allow the entry of various cations and anions of different sizes and charges. Thus, these results indicate a wider sequestering capability of the lactoferrin molecule. This conclusion should also be true for other transferrins. One important implication is that transferrins carrying different metal ions could bind to receptors, as the protein structure is unchanged.

We thank the department of Biotechnology for financial assistance. AKS thanks CSIR for the award of a fellowship.

References

Ainscough, E. W., Brodie, A. M. & Plowman, J. E. (1979). *Inorg. Chim. Acta*, **33**, 145–153.

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

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

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Chasteen, N. D., White, L. K. & Campbell, R. F. (1977). *Biochemistry*, **16**, 363–368.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Haridas, M., Anderson, B. F. & Baker, E. N. (1995). *Acta Cryst.* D**51**, 629–646.

Harris, W. R. & Carrano, C. J. (1984). *J. Inorg. Biochem.* **22**, 201–218.

Harris, W. R., Carrano, C. J., Pecoraro, V. L. & Ramond, K. N. (1981). *J. Am. Chem. Soc.* **103**, 2231–2237.

Jones, T. A., Zou, J., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Karthikeyan, S., Paramasivam, M., Yadav, S., Srinivasan, A. & Singh, T. P. (1999*a*). *Acta Cryst.* D**55**, 1792–1798.

Karthikeyan, S., Paramasivam, M., Yadav, S., Srinivasan, A. & Singh, T. P. (1999*b*). *Acta Cryst.* D**55**. Submitted.

Kurokawa, H., Mikami, B. & Hirose, M. (1995). *J. Mol. Biol.* **254**, 196–207.

Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–682.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Luk, C. K. (1971). *Biochemistry*, **10**, 2838–2843.

Mason, P. L. & Heremans, J. F. (1968). *Eur. J. Biochem.* **6**, 579–584.

Minor, W. (1993). *XDISPLAYF Program*, Purdue University, West Lafayette, Indiana, USA.

Moore, S. A., Anderson, B. F., Groom, C. R., Haridas, M. & Baker, E. N. (1997). *J. Mol. Biol.* **274**, 222–236.

O'Hara, P. B. & Bersohn, R. (1982). *Biochemistry*, **21**, 5269–5272.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Ramachandran, G. N. & Sasisekaran, V. (1968). *Adv. Protein Chem.* **23**, 283–438.

Rawas, A., Muirhead, H. & Williams, J. (1996). *Acta Cryst.* D**52,** 631–640.

Sharma, A. K., Karthikeyan, S., Kaur, P., Yadav, M. P. & Singh, T. P. (1996). *Acta Cryst.* D**52**, 1196–1198.

Sharma, A. K., Paramasivam, M., Srinivasan, A., Yadav, M. P. & Singh, T. P. (1999). *J. Mol. Biol.* **289**, 303–317.

Sharma, A. K., Rajashankar, K. R., Yadav, M. P. & Singh, T. P. (1999). *Acta Cryst.* D**55**, 1152–1157.

Smith, C. A., Anderson, B. F., Baker, H. M. & Baker, E. N. (1992). *Biochemistry*, **31**, 4527–4533.

Zak, O. & Aisen, P. (1988). *Biochemistry*, **27**, 1075–1080.