Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of mare apolactoferrin: the N and C lobes are in the closed form

A. K. Sharma,[a] K. R. Rajashankar,[b] M. P. Yadav[c] and T. P. Singh[a]*

[a]Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110 029, India, [b]Institute of Physiological Chemistry c/o DESY Notkestrasse 85, 22603, Hamburg, Germany, and [c]National Research Centre on Equines, Hisar 125 001, India

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

Received 14 December 1998
Accepted 8 March 1999

PDB Reference: mare apolactoferrin, 1b7u.

The structure of mare apolactoferrin (MALT) has been determined at 3.8 Å resolution by the molecular-replacement method, using the structure of mare diferric lactoferrin (MDLT) as the search model. The MDLT structure contains two iron-binding sites: one in the N-terminal lobe, lying between domains N1 and N2, and one in the C-terminal lobe between domains C1 and C2. Both lobes have a closed structure. MALT was crystallized using the microdialysis method with 10%(*v*/*v*) ethanol in 0.01 *M* Tris–HCl. The structure has been refined to a final *R* factor of 0.20 for all data to 3.8 Å resolution. Comparison of the structure of MALT with that of MDLT showed that the domain arrangements in these structures are identical. However, the structure of MALT is very different to the structures of human apolactoferrin (HALT) and duck apo-ovotransferrin (DAOT), in which the domain associations differ greatly. In HALT, the N lobe adopts an open conformation while the C lobe is in the closed form. On the other hand, in DAOT both the N and the C lobes adopt the open form. These results indicate the domain arrangements in these proteins to be an important structural feature related to their specific biological functions. Based on the structures of MALT, HALT and DAOT, it can be stated that the native apoproteins of the transferrin family adopt three forms: (i) with both the N and the C lobes in closed forms, as observed in MALT, (ii) with the N lobe open and the C lobe closed, as observed in HALT, and (iii) with both the N and the C lobes open, as found in DAOT. All these proteins attain a convergent form when iron is bound to them, suggesting an efficient and unique form of iron binding. The interface between the N and C lobes, which is formed by N1–C1 contact in the core of the molecule, does not change significantly.

1. Introduction

Lactoferrin is a member of the transferrin family of iron-binding proteins. These glycoproteins of molecular weight ~80 kDa have specific binding sites for two Fe^{3+} ions together with two CO_3^{2-} ions per molecule. They consist of a single polypeptide chain folded into two homologous N and C lobes. The N and the C lobes each contain about 345 amino acids and are made up of two domains N1, N2 and C1, C2, respectively, with the iron-binding site situated in the interdomain cleft. The crystal structures of human lactoferrin (Haridas *et al.*, 1995), buffalo lactoferrin (Karthikeyan *et al.*, 1999), bovine lactoferrin (Stanley *et al.*, 1997), mare lactoferrin (Sharma *et al.*, 1999), rabbit serum transferrin (Bailey *et al.*, 1988), hen ovotransferrin (Kurokawa *et al.*, 1995) and duck ovotransferrin (Rawas *et al.*, 1996) have provided structural informa-

tion regarding the organization of these proteins in the presence of iron, the environment of the iron-binding sites and the interdomain interactions. In the presence of iron, interdomain interactions result in the closed forms of these proteins. However, the crystal structures of duck apoovotransferrin (DAOT; Rawas *et al.*, 1997) and human apolactoferrin (HALT; Norris *et al.*, 1991) showed different domain association behaviours. In the former, both lobes adopt open conformations, while in the latter the N lobe has an open structure whereas the C lobe prefers a closed conformation. Here, we report the structure of iron-free mare lactoferrin (apolactoferrin) at 3.8 Å resolution. In contrast to duck apoovotransferrin and human apolactoferrin, in mare apolactoferrin both lobes are found to be in the closed conformation, similar to its iron-saturated form (Sharma *et al.*, 1999).

2. Experimental

2.1. Purification of mare lactoferrin

Mare's milk was obtained from the National Research Centre on Equines, Hisar, India. Purification was carried out using a modified version of the procedure of Law & Reiter (1977). 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 for 1 h using a mechanical stirrer. 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 in a column (25 × 2.5 cm) and washed with same buffer containing 0.1 *M* NaCl, which facilitated the removal of impurities. The lactoferrin was then eluted with 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 purified protein was used for the preparation of apolactoferrin.

2.2. Preparation of mare apolactoferrin

The iron was removed from the purified lactoferrin using the procedure of Mason *et al.* (1968). The purified iron-saturated lactoferrin solution (1%) in 0.05 *M* Tris–HCl buffer at pH 8.0 was dialyzed against an excess of 0.1 *M* citric acid with regular changes after 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 apoprotein was obtained.

2.3. Protein crystallization

Crystals of mare apolactoferrin (MALT) were obtained using the microdialysis method with a protein concentration ranging from 40 to 50 mg ml^{-1} in 0.01 *M* Tris–HCl, dialysing against the same buffer containing 10%(*v*/*v*) ethanol at pH 8.5. All crystallization experiments were carried out at 277 K. The colourless crystals grew in two weeks to dimensions 0.3 × 0.25 × 0.2 mm.

Table 1
Summary of data-collection statistics.

Crystal size (mm)	0.3 × 0.25 × 0.2
System	Orthorhombic
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 100.4, b = 77.4, c = 102.6
Resolution range (Å)	32–3.8
V (Å^3)	7.9 × 10^5
V_m (Å^3 Da^{-1})	2.49
Total number of reflections collected	101096
Total number of unique observed reflections	7887
R_{sym}	0.12
Solvent content (%)	51
Z (number of molecules in the unit cell)	4
Resolution limit (Å)	3.8
Completeness (%)	99.1
Last resolution shell (Å)	3.8–4.09
Completeness in last resolution shell (%)	98.6

2.4. Data collection

The crystals diffracted to 3.8 Å resolution. The data were collected on the EMBL Hamburg outstation beamline BW-7B at DESY with λ = 0.8345 Å using a MAR Research MAR 345 image-plate detector with 1.5° rotation for each image. The crystals were frozen at 100 K in a nitrogen-gas stream with 30% MPD in the buffer serving as cryoprotectant. The data were processed using the *HKL* package (Otwinowski, 1993; Minor, 1993). The data were complete to 99.2% with an R_{sym} of 0.12. The V_m value was 2.49 Å^3 Da^{-1} for one monomer per asymmetric unit, which is within the normal range of values found in protein crystals (Matthews, 1968) and corresponds to a solvent content of 51%. Details of the crystallographic data collection and processing are given in Table 1.

2.5. Structure determination

The structure of MALT was determined using the molecular-replacement method by employing the *AMoRe* program (Navaza, 1994) incorporated in the *CCP*4 package (Collaborative Computational Project, Number 4, 1994) and using MDLT as a model. Initially, structure solution was attempted using HALT and DAOT as search models. These models did not produce the solution, thus indicating that MALT possesses a different structure to HALT and DAOT. The cross-rotation function was calculated using diffraction data in the resolution range 12–5 Å with a sphere radius of 30 Å. The first 50 rotation-solution peaks were used to compute the translation function. The first peak of the rotation solution appeared as a distinct peak in the output of the translation search, having a correlation coefficient of 42.2% and an R factor of 50.1%. In the subsequent rigid-body refinement, the correlation coefficient increased to 46.1% and the R factor decreased to 48.7%. The solution was applied to the model coordinates using *LSQKAB* (Collaborative Computational Project, Number 4,

1994) and the transformed coordinates were used for the refinement.

After molecular replacement, two rounds of refinement using *X-PLOR* (Brünger *et al.*, 1987) followed by manual rebuilding using the program *O* (Jones *et al.*, 1991) were then performed on the protein structure. The refinement protocol in *X-PLOR* consisted of rigid-body refinement (40 cycles using all the data) followed by simulated annealing, a conventional positional refinement (50 cycles) and a restrained group *B*-factor refinement (20 cycles). The simulated-annealing runs involved heating to 3000 K and then cooling in increments of 25 K with 50 steps of molecular-dynamics simulations for 0.5 fs at each temperature. At this stage, the *R* factor and $R_{\rm free}$ (the $R_{\rm free}$ reflections, 5% of the data set, were selected using the *X-PLOR* program *FREERFLAG* and were excluded from the refinement) fell to 0.26 and 0.38, respectively. The N lobe (residues 1–344) and the C lobe (residues 345–689) were refined as two separate rigid bodies. Further energy minimization by 40 steps of positional (*xyz*) refinement followed by ten steps of group *B*-factor refinement (two *B*-factor groups for each residue: one for backbone and one for side-chain atoms) led to a drop in the *R* factor to 0.20 for all data and a fall in $R_{\rm free}$ to 0.32. The final refinement statistics are given in Table 2.

3. Results and discussion

3.1. Final model

The final coordinate set contains 5281 protein atoms from 689 amino-acid residues. The protein structure has a geometry close to ideal, with r.m.s. deviations of 0.016 Å and 2.4° from standard values of bond lengths and angles, respectively. The final *R* factor was 0.20 for all reflections to 3.8 Å resolution. A section of representative electron density from a final $(2F_o - F_c)$ map is shown in Fig. 1. The main-chain torsion angles are shown in a Ramachandran plot (Ramachandran & Sasisekharan, 1968) with various allowed regions as defined by Laskowski *et al.* (1993) in the program *PROCHECK*. The most favoured and additionally allowed regions account for more than 90% of the residues (Fig. 2). Only two residues, Leu299 and Leu640, were present in the disallowed regions. These are parts of γ-turns with torsion angles around 70.0, −50.0°. These γ-turns are conserved in all transferrins (Smith *et al.*, 1995).

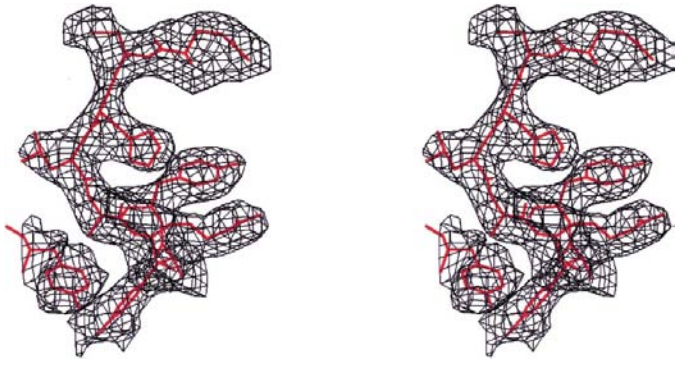

Figure 1
A section of the final electron-density $(2F_o - F_c)$ map of MALT. The contour levels are drawn at 1.5σ. The residues are Tyr92, Tyr93, Tyr192, Glu244, Cys245, His246, Leu247, Ala248 and Arg249.

Table 2
Refinement statistics for the final structure.

Resolution limits (Å)	32–3.8
Present *R* factor	0.20
$R_{\rm free}$	0.320
Number of reflections used	7887
Number of protein atoms	5281
Average *B* value (Å^2)	44.5
R.m.s. deviations from ideal values	
Bond lengths (Å)	0.016
Bond angles (°)	2.4
Dihedral angles (°)	24.6
Improper angles (°)	2.11

3.2. Structure of the protein

A very unexpected result was found in the structure of MALT (Fig. 3): both the N and the C lobes were found to be in the closed conformation. The polypeptide-chain folding of the protein was very similar to that of diferric lactoferrin. The only conformational differences between the two structures were found in the iron-binding sites and surface loops. Superposition of the MALT structure onto the mare diferric lactoferrin (MDLT) structure gave a root-mean-square deviation of 1.6 Å (687 C^{α} atoms). If the C lobe of MALT is superimposed on the C lobe of MDLT, a rotation of 4.4° is required to bring

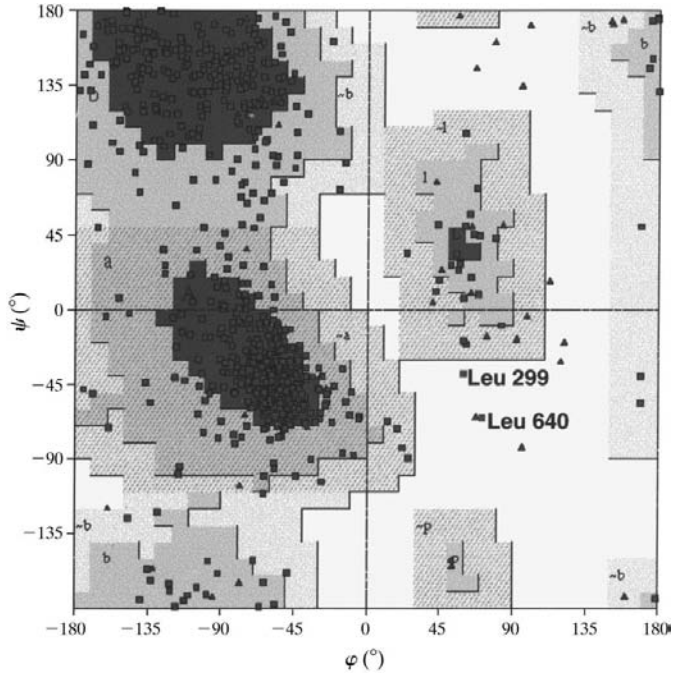

Figure 2
Ramachandran plot of the main-chain dihedral angles of MALT with 90% of the non-glycine and non-proline residues in the energetically favourable core regions. Residues Leu299 and Leu640 are part of the conserved γ-turns in lactoferrins. Squares correspond to non-glycine residues, triangles correspond to glycine residues.

Table 3
Superimposition of mare apolactoferrin and mare diferric lactoferrin.

(*a*) R.m.s. deviations of mare apolactoferrin when superimposed onto mare diferric lactoferrin.

	R.m.s. deviation (Å)
Full molecule	1.6
N lobe	0.8
C lobe	0.9

(*b*) Relative orientation of lobes and domains of mare apolactoferrin with respect to mare diferric lactoferrin.

	Relative orientation (°)
N lobe (C lobe is superimposed)	4.4
N2 domain (N1 domain is superimposed)	0.8
C2 domain (C1 domain is superimposed)	0.7

the N lobe of apolactoferrin into correspondence with that of diferric lactoferrin (Fig. 4). For individual lobes of MALT and MDLT, the r.m.s. deviations after superposition of the N lobes were 0.8 and 0.9 Å, respectively, for the C lobes. If the N1 domain of MALT was superimposed onto the N1 domain of MDLT, a rotation of 0.8° was required to bring the N2 domains into correspondence. Similarly, a rotation of 0.7° was required to bring the C2 domains of two forms into correspondence if the C1 domains were superimposed. These values are summarized in Table 3. These data clearly indicate that the two structures are essentially similar.

The major differences in the polypeptide folding of MALT and MDLT occurred in the loops present on the surface of the protein and the antiparallel β-strands present at the back of the iron-binding site. In the N lobe, these included loop 83–88, strand 88–99, loop 99–105, loop 230–237 and strand 237–253, and in the C lobe, loops 418–424 and 573–578. The r.m.s. deviations in these regions were in the range of 2.0 Å.

The binding clefts in both the N and C lobes of MALT are closed. The metal ligands in the binding sites of MALT have shifted and have acquired slightly different orientations

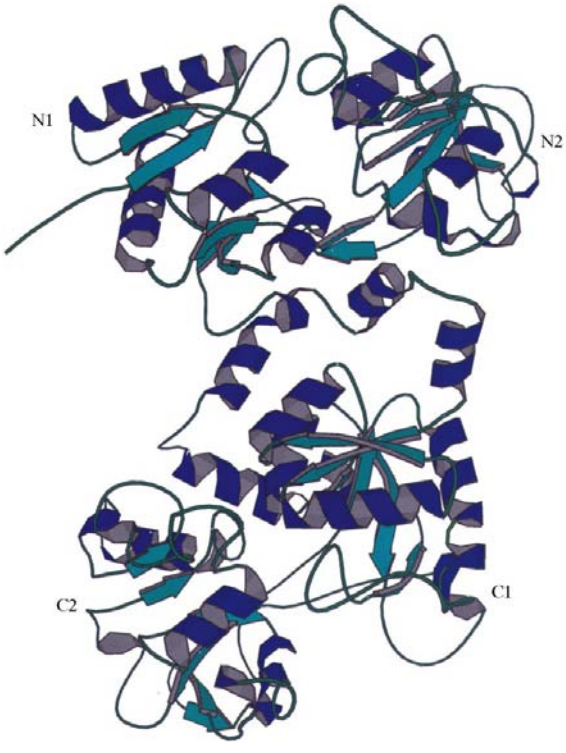

Figure 3
Structure of MALT. The molecule is folded into two globular lobes representing the N-terminal (upper) and the C-terminal (lower) halves. The binding cleft in each lobe is between two domains, N1 (upper left) and N2 (upper right) in the N lobe and C1 (lower right) and C2 (lower left) in the C lobe. The figure was drawn using *MOLSCRIPT* (Kraulis, 1991).

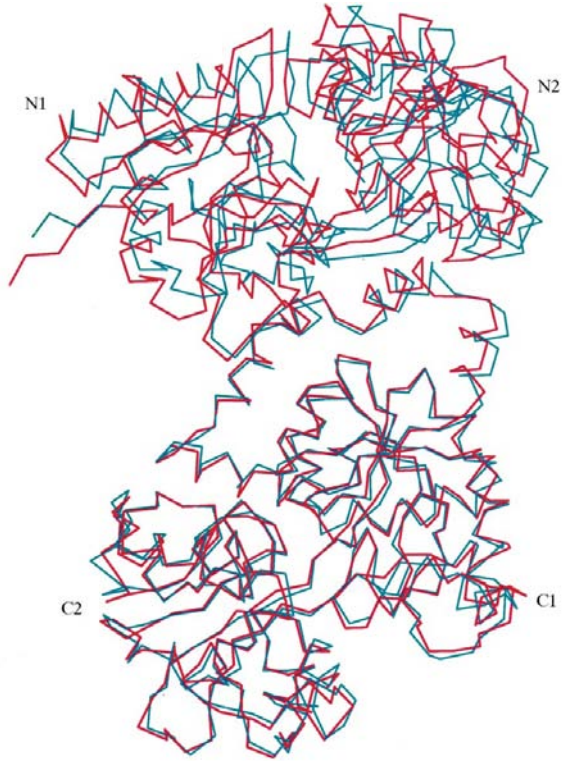

Figure 4
View of the conformations of MALT (blue) and MDLT (red) after least-squares fitting of the C^{α} atoms (superposition is based on the C lobe only). An additional rotation of 4.4° is required to bring the N lobes of the two forms into correspondence.

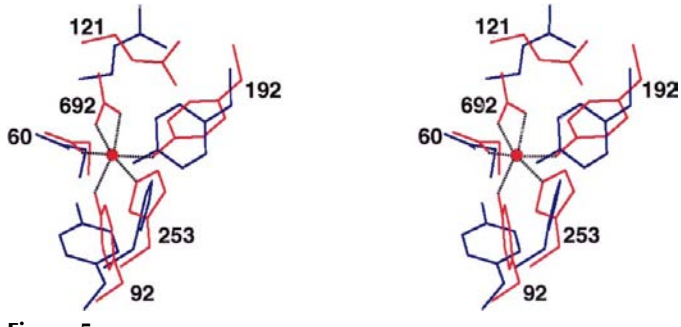

Figure 5
Superposition of the N-lobe binding site of the diferric form (red) onto the apo form (blue) of mare lactoferrin. Arg121 has a different orientation in the apo form. Residues Tyr92 and His253 move away from the binding site.

compared with those observed in MDLT (Fig. 5). Asp60 (395) has moved away from its original position in MDLT and interacts strongly with Ser122 (464). Tyr92 (433) interacts with Asp60 (395) and Tyr192 (526) points towards Tyr92 (433), while His253 (595) has moved back from its position in MDLT and does not interact with its surroundings. In the iron-binding cleft of the N lobe, the side chain of Arg121, which is involved in anion binding, has a different orientation compared with that in MDLT (Fig. 5). The guanidinium group has shifted by more than 2 Å from its position in MDLT. On the other hand, in the C lobe, Arg463 does not move at all from its position in MDLT. The strong interactions generated by Lys210 at the back of the iron site in the N lobe of MDLT are absent in MALT because of the changed orientation of its side chain.

There are minor differences in the interdomain and interlobe interactions between diferric lactoferrin and apolactoferrin. The solvent content of apolactoferrin crystals is 51%, while that of diferric form is 55%. The molecules are packed slightly more tightly in apolactoferrin crystals than in the diferric form crystals. There are 40 intermolecular contacts ($d < 3.5$ Å) in the apolactoferrin crystals compared with 32 in the diferric form crystals. In apolactoferrin, the N-lobe residues involved in intermolecular contacts are Lys21, Arg24, Val29, Arg30, Gly101, Lys197, His272, Arg273 and Ala285. The residues involved from the C lobe are Asn366, Lys416, Gln418, Asn421, Gln498, Asn509, Glu510, Asn511, Gly622, Gln652 and Arg674. The common residues involved in intermolecular interactions in the apo and diferric forms are Lys21,

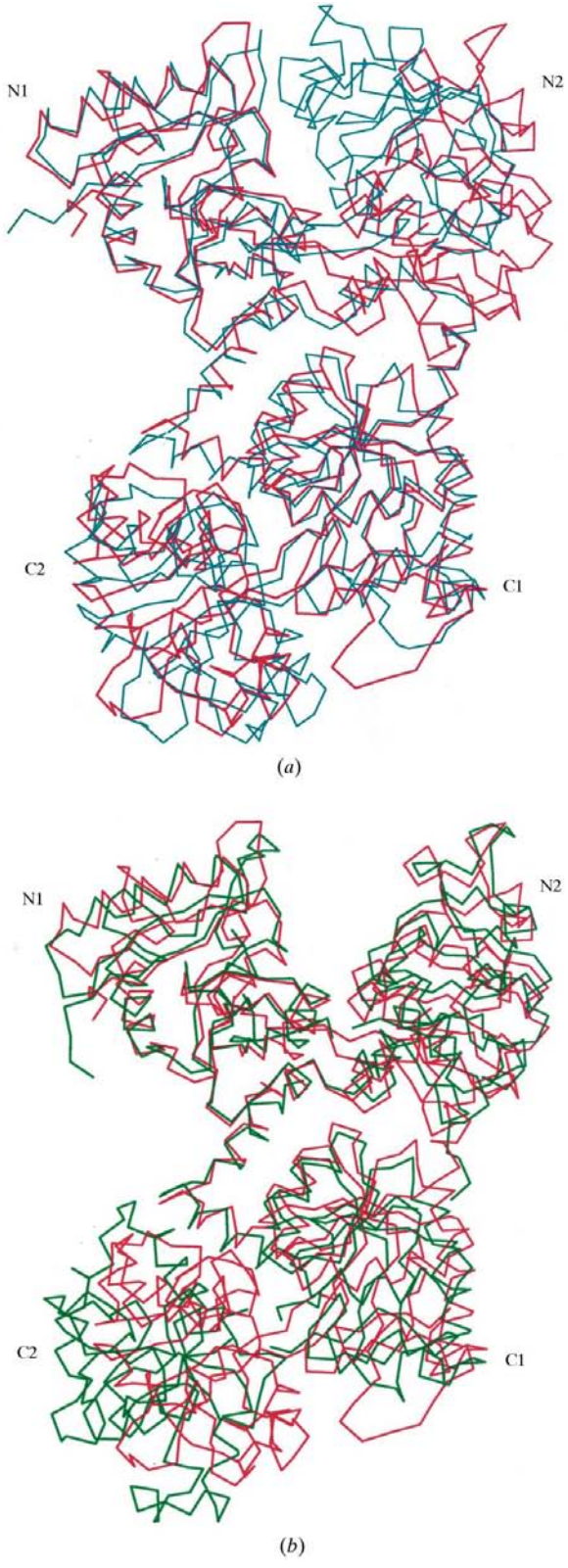

Figure 6
(*a*) C^{α} plot of mare apolactoferrin (MALT) in blue superimposed on human apolactoferrin (HALT) in red. The two structures have been refined to 3.8 and 2.8 Å, respectively. (*b*) Superposition of C^{α} chains of HALT in red and duck apo-ovotransferrin (DAOT) in green. The DAOT has been refined at 4.0 Å resolution. (*c*) The superposition of C^{α} chains of MALT in blue and DAOT in green.

Arg24, Arg30, His272, Gln418, Asn421, Glu510, Asn621 and Gly622. The residues Ala1 and Arg3, which are involved in intermolecular interactions in diferric lactoferrin crystals, do not make any intermolecular contacts in apolactoferrin crystals.

The structure of MALT, with both the N and the C lobes closed, is very different from that of HALT, which has the N lobe open and the C lobe closed (Norris *et al.*, 1991), and that of DAOT (Rawas *et al.*, 1997), which has both lobes open. An examination of the interdomain interactions in the N lobes of MDLT and HDLT clearly indicates the presence of stronger attractive interactions in the former. Furthermore, the interdomain interactions in the C lobe of HDLT (Haridas *et al.*, 1995) are stronger than those in duck diferric ovotransferrin (DDOT). Thus, it is clearly suggested that the presence of strong interdomain interactions helps in keeping the two domains together, even if the Fe^{3+} is removed.

In all the diferric proteins of the transferrin family, both lobes are in the closed conformations. However, the apoproteins show remarkable differences in their domain arrangements. In duck apo-ovotransferrin both lobes are in the open form, in human apolactoferrin the N lobe is in the open conformation while the C lobe is in the closed form, whereas in mare apolactoferrin both lobes are in the closed form. The domain associations in these proteins are compared in Fig. 6.

It is significant to note that separate superpositions of the N and C lobes of MDLT onto the N and C lobes of HDLT show a higher degree of matching between the C lobes (r.m.s. deviation 0.6 Å) than the N lobes (1.0 Å). Similarly, the N lobes of HDLT and DDOT match much more closely (r.m.s. deviation 1.0 Å) than the C lobes (r.m.s. deviation 1.2 Å). This observation is consistent with the similarities of the conformations in the apo forms. In a similar way, both the N and the C lobes of MDLT show similar mismatches with the N and C lobes of DDOT. Though the different forms of the apoproteins appear to be important structural requirements, their functional significance is not yet clear. As lactoferrin is involved in many functions, various structural elements need to be organized to produce functional variations. However, the lactoferrin structures converge to a similar folding when iron is bound to these proteins. The iron-bound forms of these proteins therefore reflect a structural convergence caused by functional convergence.

The authors acknowledge the Department of Biotechnology, New Delhi, for financial support and the EMBL Outstation, Hamburg, Germany, for beam time. They thank Dr Ch. Betzel and S. Karthikeyan for discussions and Dr Paul Tucker for help in data collection.

References

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.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

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

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

Karthikeyan, S., Paramasivam, M., Yadav, S., Srinivasan, A. & Singh, T. P. (1999). In the press.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

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

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

Law, B. A. & Reiter, B. (1977). *J. Dairy Res.* **44**, 595–599.

Mason, P. L., Heremans, J. F. & Ferin, J. (1968). *Fertil. Steril.* **19**, 679–689.

Matthews, B. (1968). *J. Mol. Biol.* **33**, 491–497.

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

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Norris, G. E., Anderson, B. F. & Baker, E. N. (1991). *Acta Cryst.* B**47**, 988–1004.

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. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.

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

Rawas, A., Muirhead, H. & Williams, J. (1997). *Acta Cryst.* D**53**, 464–468.

Sharma, A. K., Paramsivam, M., Srinivasan, A., Yadav, M. P. & Singh, T. P. (1999). In the press.

Smith, C. A., Anderson, B. F., Baker, H. M. & Baker, E. N. (1995). *Acta Cryst.* D**51**, 302–316.

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