Crystal structure of human seminal diferric lactoferrin at 3.4 Å resolution

Janesh Kumar1, Wolfgang Weber2, Sabine Münchau2, Savita Yadav1, S Bhaskar Singh1, K Saravanan1, M Paramasivam1, Sujata Sharma1, Punit Kaur1, A Bhushan1, A Srinivasan1, Christian Betzel3 and T P Singh1*

1Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110029, India
2Universitätsklinikum Hamburg-Eppendorf, IMBM, Martinistrasse 52, D-20246 Hamburg, Germany
3Institute of Medical Biochemistry and Molecular Biology, C/o DESY, Geb 22a Notkestrasse 85, 22603, Hamburg, Germany

Received 29 October 2002; revised 16 January 2003

Lactoferrin was purified from human seminal fluid obtained from the semen bank. The purified samples were saturated with Fe³⁺ and crystallized by microdialysis method. The crystals belong to orthorhombic space group P2₁2₁2₁ with a = 55.9 Å, b = 97.2 Å, c = 156.1 Å and Z = 4. The structure was determined with molecular replacement method and refined to an R factor of 18.7% for all the data to 3.4 Å resolution. The overall structure of seminal lactoferrin is similar to human colostrum lactoferrin. The amino acid sequence of seminal lactoferrin shows that it has one amino acid less than human colostrum lactoferrin and the structure of its N-terminal region is far more ordered than other lactoferrins. The structure of the iron-binding site and its immediate surroundings indicate well defined features.

Materials and Methods

Isolation and purification

Samples of human semen obtained from a laboratory specialized in fertility analyses...
Preparation of iron-saturated lactoferrin

The purified lactoferrin was dissolved in 1 ml of 0.1 M sodium bicarbonate, buffered with sodium citrate at pH 8.0 to obtain a concentration of 1 mM. To this solution, 1.2 ml of ferric chloride reagent (2 mM FeCl₃.6H₂O in 0.1 M sodium bicarbonate buffered with sodium citrate at pH 8.0) was added and equilibrated for 24 hr at 298 K. The excess of ferric chloride reagent was removed by passing the preparation through a Sephadex G-50 column (50 × 2 cm) and eluting with distilled water. The iron saturation of human seminal lactoferrin was estimated to be 92%.

Crystallization

The purified iron-saturated samples of human seminal lactoferrin were used for crystallization. The crystals were obtained by microdialysis method with protein concentration of 70 mg/ml in 0.025 M Tris-HCl at pH 8.0. This solution was dialyzed against the same buffer containing 13% (v/v) ethanol at pH 8.0. All crystallization experiments were carried out at 277 K. Rectangular shaped dark brown crystals grew to the dimensions of 0.3 × 0.16 × 0.12 mm³ in 3 weeks.

Data collection and processing

Owing to the size, the crystals of sLf did not diffract beyond 3.4 Å. Despite low resolution, it was of enough interest to determine the structure of lactoferrin from a secretion other than colostrum/milk. X-ray intensity data were measured at 285 K using a MAR Research imaging-plate scanner having a diameter of 300 mm. Monochromatic CuKα radiation was produced by a graphite-crystal monochromator mounted on a Rigaku RU-200 rotating anode generator operating at 40 kV and 100 mA with a focal
point of $3 \times 0.3 \text{ mm}^2$. The crystals diffracted to 3.4 Å resolution and remained stable in the X-ray beam. Therefore, a complete intensity data set was obtained using one crystal.

The HKL package and MARSCALE were used for the determination of unit cell parameters, data processing, scaling and merging\textsuperscript{24}. The details of crystallographic data collection and processing statistics are given in Table 2.

**Structure determination**

The structure was determined by molecular replacement method employing the AmoRe\textsuperscript{25}, incorporated in the CCP4\textsuperscript{26} package, using the model of diferric human lactoferrin\textsuperscript{5}. Since the sequences sLf and human milk lactoferrin (mLf) is identical, a full model was used for generating the rotation function. The model obtained from molecular replacement was refined using the program REFMAC\textsuperscript{27}.

Rounds of automated refinement were interspersed with manual rebuilding into 2Fo-Fc and Fo-Fc omit maps using the program O\textsuperscript{28}. A bulk solvent correction was also applied in the final rounds of refinement and map calculations. The refined model of sLf has an R factor of 0.187 and the free R factor of 0.242 with a good geometry (Table 3).

**Results and Discussion**

**Sequence analysis**

The sequence of sLf (Table 1) is almost identical to that of human milk lactoferrin (mLf; X5396118). The only differences exist in the N-terminal region, where the first two residues are absent in sLf. Two more differences correspond to the residues at positions 10 and 28 (mLf structure numbering), where alanine and lysine in mLf have changed to threonine and arginine in sLf.

**Overall structure**

Since the resolution is limited to 3.4 Å, the structure was refined with stringent constraints. The final backbone electron density is continuous and is generally well defined for the side chains as well (Fig. 1). The electron density for the iron-binding ligands including Fe$^{3+}$ ions is also clearly defined (Figs. 2a and b).

The molecule is folded into two globular lobes, the N-lobe, comprising the N-terminal half of the polypeptide chain (residue 2-333), C-lobe comprising C-terminal half of the polypeptide chain (residues 345-691) and a connecting helical peptide residues (334-344). Each lobe is further sub-divided into two domains of approximately 160 residues each (Fig. 3).

**Metal and anion-binding sites**

The two iron-binding sites situated one in each lobe at the end of each inter-domain cleft are 45 Å apart. The geometry of iron-binding coordination is a distorted octahedron in both lobes with four protein ligands occupying the four positions about the iron atoms (Figs. 2a and b). The composition of metal coordination distances involving protein ligands in

---

**Table 2—Crystallographic data on seminal lactoferrin**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>$P2_12_1$</td>
</tr>
<tr>
<td>$a$ (Å)</td>
<td>55.9</td>
</tr>
<tr>
<td>$b$ (Å)</td>
<td>97.2</td>
</tr>
<tr>
<td>$c$ (Å)</td>
<td>156.1</td>
</tr>
<tr>
<td>$Z$</td>
<td>4</td>
</tr>
<tr>
<td>Matthew's coefficient $\bar{V}$/Da</td>
<td>2.72</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>55</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.0-3.4</td>
</tr>
<tr>
<td>No. of observed reflections</td>
<td>85,368</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>10,547</td>
</tr>
<tr>
<td>Overall completeness (20.0-3.4 Å)%</td>
<td>94.3</td>
</tr>
<tr>
<td>Completeness in the highest resolution shell (3.6-3.4Å)%</td>
<td>89.1</td>
</tr>
<tr>
<td>$R_{\text{sym}}$ for all data (20.0-3.4 Å)%</td>
<td>14.7</td>
</tr>
<tr>
<td>$R_{\text{free}}$ in the highest resolution shell (3.6-3.4 Å)%</td>
<td>39.6</td>
</tr>
<tr>
<td>Average I/σ(I) in the highest resolution shell (3.6-3.4)</td>
<td>6.4</td>
</tr>
</tbody>
</table>

**Table 3—Refinement statistics**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB code</td>
<td>1N76</td>
</tr>
<tr>
<td>Space group</td>
<td>$P2_12_1$</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.0-3.4</td>
</tr>
<tr>
<td>$R_{\text{cryst}}$ (for all data)%</td>
<td>18.7</td>
</tr>
<tr>
<td>$R_{\text{free}}$ (for 5% randomly selected data)</td>
<td>24.2</td>
</tr>
<tr>
<td>R.M.S. deviations from ideal geometry bond distances (Å)</td>
<td>0.023</td>
</tr>
<tr>
<td>Bond Angles (°)</td>
<td>1.8</td>
</tr>
<tr>
<td>Torsion angles (°)</td>
<td>20.0</td>
</tr>
<tr>
<td>Overall average B factor (Å$^2$)</td>
<td>52.8</td>
</tr>
<tr>
<td>G-factor</td>
<td>−0.18</td>
</tr>
<tr>
<td>Residues in the most allowed regions (%)</td>
<td>82</td>
</tr>
</tbody>
</table>
Fig. 1—Stereoview of a section of final (2Fo-Fc) electron density map contoured at 1σ level [The residue numbers are indicated]

Fig. 2—View of the Fe$^{3+}$ binding site showing difference electron density (Fo-Fc) map for iron and iron-binding protein ligands including CO$_3^{2-}$ ion [(a): N-lobe; (b): C-lobe. The contours are drawn at 2σ level]
two lobes shows an essentially similar binding environment (Table 4).

One of the main features of the construction of the iron sites is that the four protein ligating residues are widely spaced along the polypeptide chain and belong to distinctly different parts of the protein structure. As seen from Figs. 4a and b, Asp 60 belongs to the main body of domain I and Tyr 192 to domain II, while Tyr 92 and His 253 come from each of the two backbone strands, crossing between the two domains at the back of the iron site. This arrangement might be useful for a protein whose function is the binding and release of metal ions, with an associated large-scale conformational change. It is also clearly different from many other metalloproteins, where two, three or more of the ligands are contributed by a small loop of the polypeptide chain. It is indeed noteworthy that the carboxylate groups (of Asp 60 and Asp 395) appear to fill a particularly important structural role in each iron-binding site. While one oxygen, Oδ1, is bound to iron, the other is acceptor in two hydrogen bonds, from the main-chain NH groups of residues 62 and 122. Both NH groups are at the N-termini of helices 3 and 5. The carboxylate group thus has a dual role, binding to iron and forming a strong link between the two domains through interaction with the opposing helix N-termini. In contrast, the other iron ligands do not seem to be much constrained by the surrounding protein structure.

**Interactions between the two lobes**

The two lobes are covalently connected by a peptide segment 334-344, which forms a regular α-helix. It may be noted that the connecting peptide is on the outside and rather separate from the main area of contact between the two lobes. A major contribution to the interlobe regions comes from the C-terminal helix, residues 676-691, which runs across the surface of the C-lobe and finishes up against the surface of the N-lobe. This helix is in contact with residues 310-314 in the N-lobe and three other helices, numbers (315 to 321) and (321 to 332) in the N-lobe, and number (376 to 388) in the C-lobe.

**Interactions between domains**

In the N-lobe of lactoferrin, Glu 211 forms a salt bridge with Arg 89 in the other domain and is hydrogen bonded to the main chain of one crossover strand. In the C-lobe, Asp547 is hydrogen bonded to the crossover strand via a water molecule. Lys 301 forms an ion pair across the N-lobe domain interface with Glu216. An equivalent interaction does not occur in C-lobe.
Conclusions
The sequences of sLf and mLf differ by four residues. The structure of sLf is identical to that of mLf. The structure of sLf was superimposed with the structure of mLf in order to minimize the rms shift of the main chain atoms and was found to be 0.4 Å. The observed specific mutations in the sLf may pertain to some of its roles that are not required by the milk protein. The most striking differences between sLf and mLf are the absence of two residues at the N-terminus. The iron-binding geometry and its environment are also identical in the two proteins. There are two sites of carbohydrate attachment on sLf, one in each lobe, but in neither case the electron density was observed. It may be due to several factors such as the low resolution, flexibility, disorder or the heterogeneity in the carbohydrate itself. The total carbohydrate content in sLf and mLf was reported to be 6.1% (ref. 29) and 6.4% (ref. 17), respectively. The regions of the lactoferrin structure that may be involved in cell receptor binding are difficult to identify.

Acknowledgement
Financial support from the Department of Biotechnology, New Delhi is gratefully acknowledged. Janesh Kumar thanks the Council of Scientific and Industrial Research (CSIR), New Delhi for the award of a fellowship. Skillful technical assistance in protein purification by Sigrid Himple is gratefully acknowledged.

References

Fig. 4—(a): Showing the locations of iron-binding ligands in N-lobe; (b): Showing the locations of iron-binding ligands in C-lobe