Ovine β -lactoglobulin at atomic resolution**George Kontopidis,[‡] Anna
Nordle Gilliver[§] and Lindsay
Sawyer***Structural Biochemistry Group, Institute of Cell
and Molecular Biology, The University of
Edinburgh, Swann Building, King's Buildings,
Mayfield Road, Edinburgh EH10 3BF, Scotland[‡] Present address: Department of Veterinary
Medicine, University of Thessaly, Karditsa,
Greece.[§] Present address: Novo Nordisk A/S,
Bagsvaerd, Denmark.

Correspondence e-mail: l.sawyer@ed.ac.uk

Received 26 March 2014

Accepted 19 September 2014

PDB reference: β -lactoglobulin, 4ck4

The crystal structure of the triclinic form of the milk protein β -lactoglobulin from sheep (*Ovis aries*) at 1.1 Å resolution is described together with a comparison of the triclinic structures of the low-pH bovine and high-pH ovine proteins. All three structures are remarkably similar, despite the well known pH-dependent conformational transition described for the bovine and porcine proteins that occurs in solution. The high resolution of the present structure determination has allowed a more accurate description of the protein than has hitherto been possible, but it is still not clear whether flexibility changes in the external loops can compensate for the presence of a significant void in the unliganded interior of the structure.

1. Introduction

Ovine β -lactoglobulin (β -Lg), like its more commonly studied relative from cow, is a small acid-stable protein that normally exists as a dimer of subunit molecular weight around 18 200. The bovine polypeptide chain comprises 162 amino acids, with one free cysteine and two disulfide bridges and, as a typical lipocalin (Ganfornina *et al.*, 2006), the structure contains a β -barrel with eight antiparallel β -strands with (+1)₈ topology and a three-turn α -helix on the outer surface. A ninth β -strand is used in dimer formation. β -Lg is readily isolated from milk. Ovine (*Ovis aries*) and bovine (*Bos taurus*) β -Lg have 96% sequence identity and the amino-acid changes are largely conservative (Ali & Clark, 1988). Both the cow and the sheep proteins have a number of genetic variants.

The structures of several crystal forms of the cow protein and those from reindeer (*Rangifer tarandus*; Oksanen *et al.*, 2006) and pig (*Sus scrofa domestica*; Hoedemaeker *et al.*, 2002) have been determined at resolutions from 1.6 to 3.0 Å and all show the characteristic lipocalin fold noted above. The pig protein dimerizes in a distinct fashion relative to the other species, as reflected in the pH profile: while the ruminant protein is a dimer at physiological pH and becomes a monomer at low pH, that from pig is a dimer that dissociates as the pH is raised. A common feature of the structures is flexibility of the external loops, often leading to difficulty of interpretation of what is often weak and/or discontinuous electron density (Brownlow *et al.*, 1997). In particular, the residues around 33–34, 86–88 and 112–114 have proved particularly troublesome during refinement, with main-chain conformations often appearing in unfavourable regions of the Ramachandran plot (Kontopidis *et al.*, 2004). An explanation of the flexibility of the external loops has been provided by Jameson *et al.* (2002), who noted that the entropy decrease that occurs when a hydrophobic ligand is bound to the protein can be offset by an increase in the flexibility of the external loops. Further, an internal void will also be stabilized by loop flexibility. The short tripeptide Thr-Asp-Tyr around position 99 forms a γ -turn which places Tyr99 in or very close to an unfavourable region, but generally the density is clear and unambiguous, and indeed where the other lipocalins share this tripeptide sequence they too have a γ -turn.

By chance, we found that triclinic crystals obtained from sheep β -lactoglobulin at pH 5.4 diffracted to a resolution better than 1 Å, although because of the physical restrictions of the station the data were limited to about 1.1 Å. The structure has been determined from

these data collected at around 100 K in the hope that clear density might emerge for the less well defined regions in other electron-density maps of β -Lg. Loch *et al.* (2014) have recently published two high-pH structures of the sheep protein at significantly lower resolution than that presented here. A brief comparison of the triclinic form (PDB entry 4nlj) will be made here. After the present article had been submitted we also became aware of the deposited coordinates (PDB entry 4tlj; Crowther *et al.*, 2014) of the 1.17 Å resolution structure of recombinant goat β -Lg (*cf.* the sheep A variant) crystallized at pH 6.8 from an alcohol-PEG solution.

2. Materials and methods

Pooled whole milk was obtained from the lactating herd of ewes owned by H. J. Errington & Co., Walston Braehead Farm, Lanarkshire, Scotland. The ewes were largely Lacaune with some Friesland and crossed Greyface cross-breeding. All other chemicals were of the best grade available and were supplied by Sigma-Aldrich, Poole, England.

2.1. Purification

The purification was based upon those of Armstrong *et al.* (1967) and Rocha *et al.* (1996). Ammonium sulfate (264 g l⁻¹) was added over a period of 45 min to ovine milk with constant stirring. The mixture was stirred for a further 2 h at room temperature before being filtered through a fluted Whatman 541 filter paper. The filtrate, which is the slightly yellow whey fraction, was clear and was treated with sodium azide (10 mg l⁻¹) as a preservative. The pH was adjusted to 3.2 by careful, slow addition of 1 M HCl (~50 ml l⁻¹) over 45 min with constant stirring. The resulting precipitate of unwanted protein was removed by centrifugation in a Beckman Avanti J-25 centrifuge at 34 000g for 30 min, after which the pH of the supernatant was returned to pH 6.0 with 1 M ammonia solution. Further addition of ammonium sulfate (264 g l⁻¹) precipitated the β -Lg, which was collected after centrifugation at 34 000g for 40 min. The precipitate was dissolved in a few millilitres of 0.16 M acetate buffer pH 5.2 and dialysed for 24 h against several changes of this same buffer. The resulting solution was concentrated in a Vivaspin concentrator to a final volume of 2 ml before being loaded onto a Sephacryl S-200 column (15 × 700 mm) and eluted with 50 mM sodium phosphate pH 6.2. The major peak eluting first was β -Lg followed by α -lactalbumin. The pooled fractions of the major peak were dialysed against 10 mM phosphate buffer pH 7.5 before being run onto a Pharmacia Biotech Mono-Q column (10 × 100 mm) 10 ml at a time. After washing with 100 ml buffer, a linear gradient was applied (35 ml 10 mM phosphate buffer pH 7.5 + 35 ml 10 mM phosphate buffer pH 7.5 containing 1 M KCl). The major peak that emerged around 0.5 M KCl was pooled and concentrated to 20 mg ml⁻¹ in 10 mM phosphate buffer pH 7.5 for use in crystallization trials. SDS-PAGE showed the protein to be greater than 97% pure, while native PAGE at pH 8.6 (Davies, 1974) showed a single band with a different net charge from that of the bovine A or B variants. Tryptic digestion followed by mass spectrometry allowed some 72% of the sequence to be identified, which together with the overall measured mass of 18 146 corresponds well to the calculated value from ovine β -Lg variant B (Swiss-Prot Accession No. P67976) of 18 151. Only the N-terminal Leu-to-Ile substitution could not be confirmed by mass spectrometry.

2.2. Crystallization

Hanging-drop crystallization was carried out using 5 µl of 20 mg ml⁻¹ protein solution and 5 µl of a well solution made up of

0.2 M sodium citrate, 0.1 M sodium potassium tartrate (adjusted to pH 5.4–5.8 with 1 M acetic acid) and 1.9–2.0 M ammonium sulfate. Crystals were obtained both at 283 and 290 K. Streak-seeding was used to improve the size and character of the crystals. It was also found that approximately 1.5 M sodium citrate pH 5.6 alone produced similar crystals.

2.3. X-ray data

A crystal of 0.2–0.3 mm in size obtained at pH 5.6 using 2.0 M ammonium sulfate was collected in a 0.4 mm CryoLoop (Hampton Research), dipped briefly in immersion oil (Type B, Cargille) and cooled by plunging into liquid N₂. X-ray data were then collected from the cooled crystal at 100 K (Cryostream; Oxford Cryosystems). Initial data were collected in-house on a 300 mm MAR Research imaging-plate system mounted upon an Enraf-Nonius FR571 rotating-anode generator (40 kV, 80 mA, Cu K α radiation, graphite monochromator). These X-ray data were to 1.88 Å resolution and were processed using *DENZO* and scaled with *SCALEPACK* (Otwinowski & Minor, 1997). The space group was determined to be *P*1, with unit-cell parameters $a = 37.40$, $b = 49.36$, $c = 49.42$ Å, $\alpha = 69.66$, $\beta = 68.98$, $\gamma = 77.67^\circ$. The unit-cell parameters and the space group of the ovine β -Lg crystals were distinct from those found previously for either the bovine protein (Brownlow *et al.*, 1997) or the ovine protein (Rocha *et al.*, 1996; Loch *et al.*, 2014) and could not be transformed into a higher symmetry space group.

For a monomer molecular weight of 18 150 and a unit-cell volume of 77 867 Å³, a dimer gives $V_M = 2.15$ Å³ Da⁻¹, which corresponds to a solvent content of 42.7%, which is in good agreement with the various crystal forms of β -Lg (Matthews, 1968; Sawyer, 2013; Hoedemaeker *et al.*, 2002; Oksanen *et al.*, 2006). The unit-cell parameters are similar to, but distinct from, those of the lattice X form of cow β -Lg, the unit-cell parameters of which ($a = 37.8$, $b = 49.5$, $c = 56.6$ Å, $\alpha = 123.4$, $\beta = 97.3$, $\gamma = 103.7^\circ$) were chosen on the basis of the crystal habit. They do not represent the reduced cell (Ladd & Palmer, 2003), which has unit-cell parameters $a = 37.8$, $b = 49.5$, $c = 50.7$ Å, $\alpha = 100.7$, $\beta = 111.9$, $\gamma = 103.7^\circ$ and $V = 81\,460$ Å³, with the slight difference in volume from the ovine enzyme reflecting the different crystallization conditions and the slight change in surface charge.

The self-rotation function determined by *MOLREP* (Vagin & Teplyakov, 2010) showed the expected single peak 0.43 times the origin peak on the $\chi = 180^\circ$ section at $\theta = 63.0^\circ$ and $\varphi = 48.0^\circ$, with the next highest peak being 0.25 of the origin. The structure was solved using the molecular-replacement protocol from *MOLREP* with data between 10.0 and 1.9 Å resolution. The search model consisted of a monomer of bovine β -Lg from the trigonal lattice Z crystals at pH 7.4 (Kontopidis *et al.*, 2002). The first two solutions were related by a rotation of almost 180°, which was a further indication that the two monomers form a dimer. Since the space group is *P*1, the model of the monomer fixed in the orientation of the first solution of the rotation function (RF) was used together with rotation-function solutions 2–5 for a translation search to locate the second molecule in the asymmetric unit. The translation function (TF) showed a clear translation solution corresponding to the second rotation-function solution. The combined coordinates of the first solution from the rotation function ($\alpha = 33.0$, $\beta = 153.0$, $\gamma = 164$, RF/ $\sigma = 11.4$, next highest 7.4) and the first solution of the translation function ($\alpha = 213.0$, $\beta = 28.7$, $\gamma = 190$, $x = 0.593$, $y = 0.220$, $z = 0.168$; TF/ $\sigma = 24.9$, next highest 8.5, correlation 0.62) formed the starting model for the refinement of the dimer.

A second crystal obtained at pH 5.4 using 1.8 M ammonium sulfate was cooled in a similar manner to the first crystal. X-ray data were then collected to 1.12 Å resolution from this cooled crystal at 100 K

Table 1
Crystallographic data for ovine β -lactoglobulin.
Values in parentheses are for the outer shell.

	High resolution	Atomic resolution	Merged data
Data collection	In-house	SRFC Daresbury 14.2	
Wavelength ($\text{\AA}$)	1.542	0.932	
Unit-cell parameters			
a ($\text{\AA}$)	37.40	37.148	37.148
b ($\text{\AA}$)	49.36	49.031	49.031
c ($\text{\AA}$)	49.42	49.053	49.053
α ($^\circ$)	69.66	69.754	69.754
β ($^\circ$)	68.98	69.076	69.076
γ ($^\circ$)	77.67	77.564	77.564
Resolution range ($\text{\AA}$)	15–1.88	10–1.12	15–1.12
Temperature (K)	100	100	
Oscillation range ($^\circ$)	80×1.5	100×2	
Exposure time	30 min	30 and 3 s	
No. of reflections			
Measured	62013	419802	470467
Unique	25013	107037	116059
Mean $I/\sigma(I)$	9.8 (1.3)	15.1 (4.9)	7.2 (3.7)
Completeness (%)	80.1 (73.0)	89.3 (89.6)	92.3 (89.4)
Multiplicity	2.5	3.6	4.0
R_{merge} (%)	2.5 (24.2)	4.4 (12.1)	9.4 (12.3)
Final R factor (all data) (%)			13.93
Free R factor (5% of data) (%)			16.94
Solvent sites			483
R.m.s.d. from ideality			
Bond lengths ($\text{\AA}$)			0.010
Bond angles ($^\circ$)			1.93
Mean B factor ($\text{\AA}^2$)			
Protein atoms			15.97
Solvent atoms			29.87

(Cryostream; Oxford Cryosystems) on a 345 mm MAR Research imaging-plate system at Daresbury SRS station 14.2, $\lambda = 0.932 \text{ \AA}$. The processing and refinement statistics are also shown in Table 1. The crystal diffracted to beyond $1.1 \text{ \AA}$ resolution but the detector geometry limited recording of these higher resolution reflections. Two passes ($100 \times 2^\circ$ images, 30 or 3 s exposure) produced a data set of 89.3% completeness, which was reduced by the limited φ -rotation range and ice rings. The R_{merge} was 4.4%. By merging in the home-source data set, the completeness increased to 92.3% but at the expense of the R_{merge} , which increased to 9.4%. However, it was these combined data (Table 1) that were used for the refinement of ovine β -Lg.

The initial refinement used *CNS* (Brünger *et al.*, 1998) with the merged data and the dimer model of the bovine β -Lg found by

Figure 1
Electron-density difference map ($2F_o - 1F_c$, contoured at 2.8σ) around Tyr99 of chain A. Although the particular conformation fits exactly into the electron-density difference maps, this particular residue is listed as unfavourable in the Ramachandran plot. This and all the other figures were prepared using *Coot* (Emsley & Cowtan, 2004).

MOLREP as described above. Rigid-body refinement with simulated annealing using data from 15.0 to $2.2 \text{ \AA}$ resolution lowered the R factor from 35.2 to 32.8%, whereupon further refinement was carried out with *SHELXL* (Sheldrick, 2008) using alternate rounds of positional and B -factor refinement using all data in the resolution range 10– $1.8 \text{ \AA}$. The final model was the result of anisotropic refinement with *REFMAC* (Murshudov *et al.*, 2011) using all data to $1.12 \text{ \AA}$ resolution, interspersed with model-building using *Coot* (Emsley & Cowtan, 2004). Structure validation used, *inter alia*, *MolProbity* (Chen *et al.*, 2010), with a final *MolProbity* score of 1.66, which is in the middle of the range expected for a structure at this resolution. The coordinate and structure-factor data for ovine β -lactoglobulin have been deposited in the Protein Data Bank (Berman *et al.*, 2000) as entry 4ck4.

3. Results and discussion

The results of the refinement are shown in Table 1. The main chain is mostly well defined, with 96% of the amino acids lying within the most favourable regions of the Ramachandran plot, and with the well documented γ -turn at Tyr99 (Brownlow *et al.*, 1997; Sawyer, 2013) and Ala34 (conformation 1) being the only two amino acids in less favourable regions. In fact, Tyr99 of chain A can be brought into a favourable region by rotation of ψ by 1° (Fig. 1). Ala34 is on the long flexible loop between strands A and B which was refined by Loch *et al.* (2011) with the preceding peptide bond flipped (conformation 2), thereby allowing the chain to exist in a more favourable conformation. We have examined this region critically in the ovine structure,

Figure 2
Difference electron-density maps around Ala34 of chains A and B produced after several cycles of refinement with the region omitted from the calculation. A single conformation observed for Ala34 (1.6σ) in chain A (a), although Ala34 (1.3σ) in chain B seems to partially adopt (30%) an alternative conformation (b).

refining it exhaustively in both conformations 1 and 2. While we see a minor conformer with an occupancy of about 0.3 consistent with conformation 2 in chain *B*, there is little evidence for this conformation in chain *A*. Fig. 2 shows Ala34 in chain *B*. The two polypeptide chains in the asymmetric unit show small structural differences mostly due to crystal packing. The most significant difference is that the rigid and well determined N-terminal residues in the *B* chain are not matched by a similar clarity in the *A* chain, where residues 3–4 appear disordered with little obvious density for Ile1 and Ile2. The electron density for Asn63 in the *B* chain is indistinct and this residue has been omitted. The EF-loop is in the closed position as expected for a structure determined at a pH below the Tanford transition, the pH-dependent conformational transition triggered by deprotonation of Glu89 around pH 7 and involving a significant movement of the EF-loop that permits access to the internal ligand-binding site (Tanford *et al.*, 1959; Qin *et al.*, 1998; Sakurai *et al.*, 2009; Sawyer, 2013). The hybrid form described by Vijayalakshmi *et al.* (2008) is not observed in our structure.

Recently, Loch *et al.* (2014) have published the structures of two crystal forms of ovine β -Lg grown at high pH: a trigonal form that is identical to that reported by Rocha *et al.* (1996) and a triclinic form (PDB entry 4nlj) that is distinct from both the bovine triclinic form (Brownlow *et al.*, 1997) and that reported here. Interestingly, their high-pH forms both have the EF-loop in the closed position. As pointed out by Lewiński and coworkers (Loch *et al.*, 2014), the same buried Glu89 is present in the ovine protein, so that it is likely that a similar pH-dependent transition to the transition reported by Tanford and coworkers (Tanford *et al.*, 1959; Sakurai *et al.*, 2009) for the bovine protein will occur in the ovine protein. To the best of our knowledge, no such solution studies have been carried out for the ovine protein as they have for the porcine protein (Ragona *et al.*, 2003), in which the transition is shifted to significantly higher pH. However, pH titration of the goat protein, essentially the ovine *A* variant, by Ghose *et al.* (1968) shows the presence of at least one abnormal carboxyl group, consistent with a Tanford-like transition.

The r.m.s.d. between corresponding atoms of chains *A* and *B* in our structure is 1.18 Å. The r.m.s.d. between backbone atoms of 152 (7–159) residues in the ovine and bovine (lattice X; PDB entry 1beb; Brownlow *et al.*, 1997) proteins is 0.549 Å. Comparison of the two triclinic ovine structures (PDB entry 4ck4 and 4nlj) in *Coot*, superimposing the high-pH coordinates for chain *A* with those reported

here, shows that the structures are remarkably similar. The 132 C α atoms of the *A* chains which were matched (alternative conformations led to the omission of some 30 residues) had an r.m.s.d. of 1.110 Å, while the *B*-chain r.m.s.d. was 1.396 Å for 130 main-chain C α atoms. It is clear that there is a different orientation of one subunit relative to the other in the two structures, but it is impossible to say whether this is real or, more likely, a feature of the different crystal packing. Strikingly, the main chains of the EF loops in both the *A* and *B* chains are almost indistinguishable, with some variation in the side chains. However, the buried Glu89 side chain adopts identical conformations in all four instances. The variability of the 33–34 peptide group referred to above is present in both the *A* and *B* chains of the high-pH structure, perhaps also reflecting the different crystal contacts. A preliminary comparison of the coordinates 4tlj of the recombinant goat protein shows a very similar arrangement of main-chain and side-chain atoms. The same regions of increased flexibility observed here are also present, but there is a significantly larger change in the intersubunit arrangement mentioned above. The *A*-chain C α superposition with 4ck4 has an r.m.s.d. of 0.652 Å, the *B*-chain superposition has an r.m.s.d. of 0.857 Å and the angular difference of the intramolecular superposition axis of 4tlj from that of 4ck4 is 4.9° compared with 2.0° for the ovine high-pH structure 4nlj. As the crystallization medium used for the goat structure is significantly different from that used in our work, it is possible that the intersubunit rearrangement is a crystallization artefact.

During refinement, it became obvious that at 1.1 Å resolution it is possible to detect alternative conformations with occupancies less than 0.3. This provided an excellent opportunity to observe traces of lipophilic ligands such as fatty acids in the calyx. Although the EF-loop was in the closed position, it was possible that a small fatty acid bound at physiological pH might still be present in our crystal structure. An overlay of the structure of β -Lg bound to palmitic acid (PDB entry 1gxa; Kontopidis *et al.*, 2002) revealed no corresponding electron density (Fig. 3): the calyx remained empty. No water molecules were present inside either, although some of the void is occupied by alternative side-chain conformations, for example Met107, located in the centre of the calyx.

There are seven amino-acid changes in the sheep B genetic variant compared with the bovine B genetic variant, four of which lead to a change in charge (Asp53Asn, Asp130Asn, Glu158Gly and, perhaps the most interesting, Tyr20His; bovine amino acid first) in a structurally conserved region of the lipocalins. These changes cause a net change of +3 at pH values above the pI, a feature clearly indicated in the native gel which was run at pH 8.6 (data not shown). Only a single band was observed from the preparation used for crystallization, in comparison with the distinct bands seen in bovine AB from Sigma (data not shown). However, a distinct blob of positive electron density was observed too close to His20 to be a water molecule (Fig. 4a), together with some negative density beside the imidazole ring. As the Cys and Met residues are all well defined, it seems improbable that the extra density represents a radiation-damage-induced modification (Weik *et al.*, 2000). A reassessment of the mass-spectrometric data showed that peaks at both ($M + H$)⁺ of 2681.38 and 2707.32 were observed in the spectrum which corresponded to the expected ($M + H$)⁺ of 2681.36 and 2707.38 for peptides containing His and Tyr, respectively, although no quantitative estimate of the ratio was possible. Further, refinement of a tyrosine at this position left the *R* factor unaltered and led to Fig. 4(b), which shows a possible but not entirely convincing side chain (compare the side chain of Tyr99 shown in Fig. 1). Trial-and-error refinement of both His and Tyr at position 20 converged with an occupancy ratio of His to Tyr in both chains of 0.4 to 0.6 (Fig. 4c). In parallel, a few cycles of *SHELXL* with

Figure 3
Overlay of β -Lg ovine (green) and the bovine structure (PDB entry 1gxa, cyan) with palmitic acid bound in the central lipophilic calyx of the protein. Difference electron-density maps ($1F_o - 1F_c$; positive, green; negative, red; 3.7σ) around the central calyx of the ovine β -Lg structure show that there are no traces of ligand present in the central calyx. While the central calyx is empty, some residues, such as Met107 and Leu58, adopt more than one conformation, thus partially filling the void.

the occupancies of both amino acids at position 20 refined together produced broadly similar values for the R factor but a somewhat lower ratio of His to Tyr. It is probable therefore that both the A and B genetic variants are present in roughly equal amounts. Although any individual animal can be homozygous or heterozygous in the A or B genes, milk pooled from a flock will generally exhibit a non-integral ratio. The other ovine A/B variation is an Asp/Asn substitution at position 130. The density is clear in both the A and B chains, although alternative conformations of χ_2 appear to be likely in chain B. As both residues are in contact only with the solvent, it is not possible to distinguish between the amide and carboxyl groups especially, as already seen, when both appear to be present. Published ratios for the Lacaune and Friesland breeds give an indication that the A variant is slightly more common than the B variant (Erhardt,

1989; Felipe & Law, 1997; Barillet *et al.*, 2005), which is in keeping with the crystallographic result, although individual flocks do show significant variation. The C variant is much rarer (Barillet *et al.*, 2005) and can be discounted here as there is no evidence of the substitution at Arg148Gln.

The dimer interface comprises the antiparallel association of strands I together with the Asp33–Arg40 ion pair as described by Brownlow *et al.* (1997), Qin *et al.* (1998) and Sakurai & Goto (2002). The hydrogen bonding between the main chain of Ala34 and the Asp/Arg ion pair provides a plausible explanation of the unfavourable conformation observed for Ala34. There are a number of crystal contacts involving different regions of the molecule such that a possible distortion in one subunit is absent in the other. This is particularly apparent in the N-termini, where the A chain is disordered relative to the B chain. In the A chain, the side chains of residues 3–4 are close to the C-terminal of residues 162–159 of the A chain of the molecule at $x - 1, y, z$. In the B chain, the O atom of Thr4 makes a good hydrogen bond to the ND2 atom of Asn90 of the molecule at $x + 1, y, z - 1$. There are some 26 intermolecular hydrogen-bond donor–acceptor contacts of less than 3.2 Å within the crystal, involving predominantly Glu, Asn, Lys and Thr side chains. The bovine lattice X is also triclinic but the crystal contacts are quite distinct, with only three of the same residue pairs involved. There are also only 20 intermolecular contacts as defined above. This may reflect the significantly lower resolution of the structure determination.

The solvent structure is made up of 483 molecules, including two sulfate, 16 ammonium, four acetate and three chloride ions, assigned on the basis of their electron density when refined as O atoms. In the case of the NH_4^+ and SO_4^{2-} ions, the local coordination was also considered. A number of water molecules occur with partial occupancy either associated with alternative side-chain conformations or as alternative water positions. As noted above, the central calyx or binding cavity had no unexplained electron density and we conclude that the cavity is empty.

4. Conclusion

Of all the β -Lg structures in the PDB (<http://www.rcsb.org>), the current structure has (just) the highest resolution [1.12 Å, with the next being the 1.17 Å PDB entry 4tlj and then the 1.64 Å PDB entry 4gny (Gutierrez-Magdaleno *et al.*, 2013)] and a good R factor of 0.139 ($R = 0.132$ for PDB entry 4tlj, with the next best R factor being 0.181 for PDB entry 1beb) and $R_{\text{free}} = 0.169$ ($R_{\text{free}} = 0.159$ for PDB entry 4tlj, with the next best R_{free} being 0.218 for PDB entry 4nlj). Despite this high resolution, it is clear that there is still some mobility in the structure at 100 K. The current structure appears to have the highest number of alternative conformations of all β -Lg structures in the PDB, reflecting the real mobility of the structure. It is also the first structure where the presence and ratio of both the A and B genetic variants has been determined crystallographically. Without determining the structure of the ligand-bound protein, however, it is not possible to add further to Jameson's thesis that flexibility changes in the external loops of the lipocalins are able to offset the requirement to maintain a large void in the centre of the molecule.

We are grateful to Humphrey Errington of Errington Cheese Ltd for providing us with the ewe's milk, to Stella Bury, Atlanta Cook, Helen Denton, Carl Holt and Iain McNae for helpful discussions and to Andy Cronshaw for the mass spectrometry. We thank the staff of Station 14.2 at SRS Daresbury, and the financial support from the

Figure 4

The genetic variation at residue 20. (a) His residue refined in position 20. The large positive electron density in the $(1F_o - 1F_c, 3.2\sigma)$ map (green) next to the His ring is too close to the imidazole to be a water molecule. (b) A Tyr residue refined in position 20. The negative electron density in the $(1F_o - 1F_c, 3.4\sigma)$ map (magenta) indicates that Tyr cannot occupy 100% of this position. (c) His (40% occupancy) and Tyr (60%) refined together at position 20. All positive and most negative electron density has disappeared.

BBSRC and a European Union Concerted Action is gratefully acknowledged.

References

- Ali, S. & Clark, A. J. (1988). *J. Mol. Biol.* **199**, 415–426.
- Armstrong, J. M., McKenzie, H. A. & Sawyer, W. H. (1967). *Biochim. Biophys. Acta*, **147**, 60–72.
- Barillet, F., Arranz, J.-J. & Carta, A. (2005). *Genet. Sel. Evol.* **37**, S109–S123.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Brownlow, S., Morais Cabral, J. H., Cooper, R., Flower, D. R., Yewdall, S. J., Polikarpov, I., North, A. C. T. & Sawyer, L. (1997). *Structure*, **5**, 481–495.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Crowther, J. M., Suzuki, H., Kessans, S. A., Lasse, M., Loo, T. S., Hodgkinson, A. J., Jameson, G. B. & Dobson, R. C. J. (2014). *FEBS Lett.* **588**, 3816–3822.
- Davies, D. T. (1974). *J. Dairy Res.* **41**, 217–228.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Erhardt, G. (1989). *Anim. Genet.* **20**, 197–204.
- Felipe, X. & Law, A. J. R. (1997). *J. Dairy Res.* **64**, 459–464.
- Ganforina, M. D., Sanchez, D., Greene, L. H. & Flower, D. R. (2006). *Lipocalins*, edited by B. Åkerström, N. Borregaard, D. R. Flower & J.-P. Salier, pp. 17–27. Georgetown: Landes Bioscience.
- Ghose, A. C., Chaudhuri, S. & Sen, A. (1968). *Arch. Biochem. Biophys.* **126**, 232–243.
- Gutierrez-Magdaleno, G., Bello, M., Portillo-Tellez, M. C., Rodriguez-Romero, A. & Garcia-Hernandez, E. (2013). *J. Mol. Recognit.* **26**, 67–75.
- Hoedemaeker, F. J., Visschers, R. W., Alting, A. C., de Kruif, K. G., Kuil, M. E. & Abrahams, J. P. (2002). *Acta Cryst.* **D58**, 480–486.
- Jameson, G. B., Adams, J. J. & Creamer, L. K. (2002). *Int. Dairy J.* **12**, 319–329.
- Kontopidis, G., Holt, C. & Sawyer, L. (2002). *J. Mol. Biol.* **318**, 1043–1055.
- Kontopidis, G., Holt, C. & Sawyer, L. (2004). *J. Dairy Sci.* **87**, 785–796.
- Ladd, M. F. C. & Palmer, R. A. (2003). *Structure Determination by X-ray Crystallography*, 4th ed., pp. 583–584. New York: Kluwer.
- Loch, J. L., Molenda, M., Kopeć, M., Świętek, S. & Lewiński, K. (2014). *Biopolymers*, **101**, 886–894.
- Loch, J., Polit, A., Górecki, A., Bonarek, P., Kurpiewska, K., Dziedzicka-Wasylewska, M. & Lewiński, K. (2011). *J. Mol. Recognit.* **24**, 341–349.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Oksanen, E., Jaakola, V.-P., Tolonen, T., Valkonen, K., Åkerström, B., Kalkkinen, N., Virtanen, V. & Goldman, A. (2006). *Acta Cryst.* **D62**, 1369–1374.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M., Baker, E. N. & Jameson, G. B. (1998). *Biochemistry*, **37**, 14014–14023.
- Ragona, L., Fogolari, F., Catalano, M., Ugolini, R., Zetta, L. & Molinari, H. (2003). *J. Biol. Chem.* **278**, 38840–38846.
- Rocha, T., Brownlow, S., Saddler, K. N., Fothergill-Gilmore, L. A. & Sawyer, L. (1996). *J. Dairy Res.* **53**, 575–584.
- Sakurai, K. & Goto, Y. (2002). *J. Biol. Chem.* **277**, 25735–25740.
- Sakurai, K., Konuma, T., Yagi, M. & Goto, Y. (2009). *Biochim. Biophys. Acta*, **1790**, 527–537.
- Sawyer, L. (2013). *Advanced Dairy Chemistry*, 4th ed., Vol. 1A, edited by P. L. H. McSweeney & P. F. Fox, pp. 319–386. New York: Kluwer Academic/Plenum.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Tanford, C., Bunville, L. G. & Nozaki, Y. (1959). *J. Am. Chem. Soc.* **81**, 4032–4036.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Vijayalakshmi, L., Krishna, R., Sankaranarayanan, R. & Vijayan, M. (2008). *Proteins*, **71**, 241–249.
- Weik, M., Ravelli, R. B. G., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 623–628.