Title

The structure of bovine β-lactoglobulin in crystals grown at pH 3.8 exhibiting novel threefold twinning.

Journal

Acta Crystallographica Section F: Structural Biology Communications, 75(Pt 10)

Authors

Yeates, Todd

Mcpherson, Alexander

Publication Date

2019-10-01

DOI

10.1107/S2053230X1901224X

Peer reviewed

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

The structure of bovine β-lactoglobulin in crystals grown at pH 3.8 exhibiting novel threefold twinning

Todd O. Yeates[a] and Alexander McPherson[b]*

[a]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095-1569, USA, and [b]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697-3900, USA. *Correspondence e-mail: amcphers@uci.edu

Received 16 January 2019
Accepted 3 September 2019

Edited by G. G. Privé, University of Toronto, Canada

Keywords: β-lactoglobulin; milk; tritohedral twinning; symmetry; whey proteins; space group determination.

PDB reference: β-lactoglobulin, 6nkq

Supporting information: this article has supporting information at journals.iucr.org/f

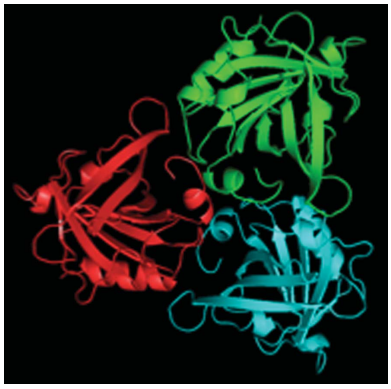
Bovine β-lactoglobulin was crystallized from 3 *M* NaCl buffered at pH 3.8 with sodium citrate as thick hexagonal prisms of greater than 1 mm in edge length. Analyses of the X-ray diffraction intensities using three different current algorithms were unanimous in specifying the space group to be $P6_322$, with unit-cell dimensions $a = b = 75.47$, $c = 140.79$ Å. No progress could be made, however, towards an acceptable solution by molecular replacement using this symmetry. In the end, it was found that the true space group was $C222_1$, a subgroup of $P6_322$, with $a = 65.89$, $b = 114.12$, $c = 140.51$ Å, with the apparent 622 symmetry arising from an unusual threefold or tritohedral twinning. An assembly based on a model of the protein in another crystal form (PDB entry 1beb) containing three molecules in the asymmetric unit was refined to 2.3 Å resolution with a final *R* factor of 0.23 and R_{free} of 0.26. NCS restraints were maintained throughout. For the most part, the molecules found in this crystal form are virtually the same as in PDB entry 1beb, although there are numerous local variations, particularly in loop elements, rotamer conformation differences and some alterations, including additions, at the termini.

1. Introduction

The two principal proteins in the milk of cows and sheep are casein, which is not a whey protein but constitutes the major component of dairy milk, and β-lactoglobulin (βLG), which is the predominant protein of whey. Whey (Boland, 2011) is defined as the liquid remaining after the making of cheese. Given that France alone has over 350 varieties of cheese, it is perhaps not surprising that the term, although widely used, is somewhat imprecise. βLG is estimated as constituting from 45% to 60% of the total whey protein and from 0.2% to 0.4% of the total weight of skimmed milk (Chatterton *et al.*, 2006; Jost, 1993). An unfortunate property of βLG is that it and casein are the most allergenic components of dairy products (Juarez & Ramos, 2003; Mills & Tatham, 2003).

βLG is a protein of 162 amino acids after the removal of a 17-amino-acid signal peptide from its gene transcript (UniProtKB P02754 [LACB_BOVIN]). It has a molecular weight of 18 300 Da (Kontopidis *et al.*, 2004). The molecule exists as a twofold-symmetric dimer over the pH range 5–8, but has been reported to form tetramers, octamers and higher oligomers below pH 5 (Kontopidis *et al.*, 2004; Jost, 1993; Qin *et al.*, 1998). At pH values below 3 the dimer dissociates into monomers. The protein is well known for undergoing a measurable, and extensively investigated, pH-induced conformational change (Qin *et al.*, 1998).

The architecture of the protein is primarily β-structure and extended polypeptide and loops (43% β-structure, 47% extended chain, β-turns and one three-turn α-helix; Qin *et al.*, 1998; Kontopidis *et al.*, 2004). Nine β-strands form what is



known as a calyx, a goblet-shaped cavity enclosed by the strands. A great variety of conventional lipophilic molecules, including long-chain fatty acids, retinol, vitamin D, cholesterol and a host of other such molecules, are often accommodated in the calyx space. Hence, the protein is referred to as a lipocalin (Sawyer & Kontopidis, 2000). The structure of the protein has been determined in multiple crystal forms to a resolution of as high as 1.8 Å in a triclinic crystal (Brownlow *et al.*, 1997; Kontopidis *et al.*, 2004).

The function of βLG has been a source of controversy, and indeed no function has been definitively assigned. It may in fact have multiple functions, as several physiological responses have been linked to βLG and its oligopeptide cleavage products (Hernández-Ledesma *et al.*, 2008; Jost, 1993; Chatterton *et al.*, 2006; Creamer *et al.*, 2011). Among these are antiviral, anticarcinogenic and hypocholesterolemic effects, and prevention of pathogen adhesion (Chatterton *et al.*, 2006), as well as the regulation of mammary-gland phosphorus metabolism, transport of vitamin D, cholesterol and retinol, transfer of passive immunity to newborns and enhancement of pre-gastric esterase activity (Madureira *et al.*, 2007). Although found in large quantities in milk, a classical nutritional food, its resistance to proteolysis argues against its being strictly a nutritional protein (Morr & Ha, 1993; Mills & Tatham, 2003; Sawyer, 2013). It is nonetheless rich in branched-chain amino acids, making it the richest known source of these nutritional components (Boland, 2011). Its binding of lipophilic molecules suggests that it may principally act as a transport protein, while others contend that stimulation of lipase activity may be its principal function (Sawyer, 2013). The structure, functions and biological properties of βLG have been thoroughly and ably reviewed many times (Sawyer & Kontopidis, 2000; Hambling *et al.*, 1992; Jost, 1993; Creamer *et al.*, 2011) and little more can be added here.

Some years ago, during the course of studies of dye binding in protein crystals (McPherson & Larson, 2018), βLG was crystallized and used as a test sample. The conditions used for its crystallization were somewhat unusual in that the crystals, which were large (0.5–1.0 mm) hexagonal prisms, formed at pH 3.8 in 2.5–3 *M* NaCl. This is well outside the pH range in which other βLG crystals were grown, and indeed the crystals have a unique cell and a novel symmetry. The solution of the structure, in the end, posed a challenging crystallographic problem.

2. Materials and methods

Lyophilized βLG was purchased from Sigma–Aldrich (St Louis, Missouri, USA) and dissolved to a concentration of 30 mg ml^{-1} in distilled water. When the protein solution was combined with an equal volume of 3 *M* NaCl buffered with sodium citrate at pH 3.8 (pH range 3.8–4.2), a heavy white precipitate immediately formed. After 15–30 min the precipitate was centrifuged away, leaving a clear solution with a substantially reduced protein concentration. To effect crystallization, 12 µl drops of the protein–NaCl solution were deployed into Cryschem sitting-drop plates (Hampton Research, Aliso Viejo, California, USA) along with 0.6 ml reservoirs of 3 *M* NaCl buffered with citrate at pH 3.8. Hexagonal prisms appeared after 12 h at temperatures of between 37 and 4°C and grew to full size after three days. Their properties are detailed in Table 1. Because of their large size, reproducibility, abundance and apparent flawlessness, they were easily manipulated for data collection at room temperature and cryotemperature.

Table 1
Crystal properties.

Crystal habit	Thick hexagonal prisms
Apparent unit-cell dimensions (Å)	$a = b = 75.47$, $c = 140.79$
Apparent space group	$P6_322$
True unit-cell dimensions (Å)	$a = 65.89$, $b = 114.12$, $c = 140.51$
True space group	$C222_1$
No. of molecules in unit cell	24
No. of molecules in asymmetric unit	3
V_M (Å^3 Da^{-1})	3.13
Solvent content (%)	60
Twin fractions	0.417, 0.305, 0.278

For the structure analysis and model reported here, X-ray diffraction intensities were collected on beamline 8.3.1 at the Advanced Light Source (ALS), Lawrence Berkeley National Laboratory, Berkeley, California, USA using a Dectris PILATUS 300K detector. After submersion in 3 *M* NaCl pH 3.8 containing 25% glycerol for 30 s to 1 min, the crystals were flash-cooled in the cryostream at 173 K. Room-temperature data were also collected for comparison at University of California Irvine using a Rigaku MicroMax-007 HF generator with Osmic mirrors and a Saturn 944+ CCD as a detector.

At beamline 8.3.1, $\Delta\omega$ rotation sectors of 0.1° were collected with collection times of 0.1, 0.2 and 0.4 s. The prismatic crystals, which exhibited a mosaicity of 0.80°, proved to be rather resistant to X-ray exposure and were so large that five to six different volumes, with 360° of data per site, could be recorded from a single crystal. The specific data used in this analysis were collected from five different locations on a single crystal, and this yielded a total of 2 241 005 intensities that resulted in an average multiplicity in space group $C222_1$ of 36. On the Rigaku rotating-anode home source, the crystals were conventionally mounted (McPherson, 1982) in quartz capillaries. $\Delta\omega$ rotations of 0.30° were recorded with collection times of 30 s. Data-collection statistics are given in Table 2.

The X-ray diffraction patterns were not, unfortunately, as flawless as the appearance of the crystals might have suggested. The reflections were frequently mosaic, with spreads of up to 1° common. Importantly, the patterns were consistently highly anisotropic, with the diffraction strength declining with scattering angle much more rapidly in the **hk** plane of reciprocal space than along the **l** direction. While $CC_{1/2}$ for the data fell below 0.30 at 2.2 Å overall and at 2.0 Å along **l**, this occurred at 2.6 Å in the **hk** plane. This suggested that some form of disorder in the **a** × **b** plane was likely to be present in the crystals.

At ALS, the data were processed and reduced as they were recorded using *XDS* (Kabsch, 1988, 1993, 2010*a*,*b*) but only to unaveraged intensities in space group *P*1. This allowed subsequent averaging in space groups of lower symmetry than

Table 2
Data-collection, processing and scaling statistics.

Values in parentheses are for the outer shell.

X-ray source	ALS beamline 8.3.1
Detector	Dectris PILATUS 300K
Mosaicity (°)	0.8
Resolution (Å)	58.0–2.3 (2.38–2.30)
No. of observations	2241005
No. of unique reflections	23730 (1056)
$CC_{1/2}$	0.999 (0.976)
R_{merge}	0.220
R_{meas}	0.225
$R_{p.i.m.}$	0.031
Completeness (%)	99.46 (99.0)
Multiplicity	36.6 (38.7)
$\langle I/\sigma(I)\rangle$	10.6 (2.2)
No. of batches	5
No. of crystals	1

the predicted space group $P6_322$. The crystal data collected at room temperature were processed by the software supplied by Rigaku and reduced using *iMosflm* (Battye *et al.*, 2011). Data collected at both sources and temperatures were scaled, reduced to an asymmetric unit and converted to structure amplitudes using *POINTLESS* and *AIMLESS* (Evans, 2006, 2011; Evans & Murshudov, 2013) from the *CCP*4 suite (Winn *et al.*, 2011). Molecular-replacement searches, based primarily on PDB entry 1beb (Brownlow *et al.*, 1997), were carried out using *Phaser* (Read, 2001; Storoni *et al.*, 2004; McCoy *et al.*, 2005, 2007), also from the *CCP*4 suite. Refinement utilized *REFMAC*5 (Murshudov *et al.*, 1997, 2011) from the *CCP*4 suite. No TLS refinement (Winn *et al.*, 2001) was used and NCS restraints were imposed throughout. Graphic analysis and water addition was performed and most figures were generated by *Coot* (Emsley *et al.*, 2010) and *PyMOL* (DeLano, 2002). All computing operations were conducted on an Apple MacBook Air or a 64-bit Intel Xeon CPU.

3. Results and discussion

What initially appeared to be a relatively straightforward crystal structure determination, anisotropic disorder aside, was not. The Rigaku system, on which data were first collected at room temperature, indicated the space group to be $P6_322$ with unit-cell parameters $a = b = 75.47$, $c = 140.79$ Å with a high level of probability. The unit cell yielded a V_M of 3.13 Å^3 Da^{-1} (Matthews, 1968), which implied a solvent volume of about 60% and one molecule of βLG in the asymmetric unit. Processing at the time of data collection at ALS by *XDS* similarly predicted the space group to be $P6_322$ with the same unit-cell dimensions, again with confidence. Finally, the combined *POINTLESS* and *AIMLESS* programs indicated the space group to be $P6_322$, again with near-certainty (89% probability). Scaling in *AIMLESS*, however, was worse than expected using point-group symmetry 622. Scaling at moderate resolution was acceptable, for example with a merging R around 0.12 to 3 Å resolution, but the R value rose to about 0.30 overall when processing to 2.3 Å resolution. The divergence from perfect 622 symmetry provided an early warning sign of a possible space-group problem.

The unit cell in space group $P6_322$, which had one molecule of βLG composing the asymmetric unit, suggested the presence of the common βLG dimer in the crystal with its dyad axis coincident with a crystallographic twofold axis. However, using a 1.8 Å resolution structure of a monomer model from PDB entry 1beb as a probe, no molecular-replacement solution could be found by *Phaser*. Alternative models from the PDB resulted in additional failures. It became evident that the space group of the crystals must be other than $P6_322$.

The inability to find a structure solution in $P6_322$ prompted an investigation into possible twinning. Initial statistical tests on the intensity distribution gave a relatively weak indication of twinning when diffraction data to the 2.3 Å resolution limit were included in the analysis. The value for the local twin statistic was $\langle|L|\rangle = 0.419$ (compared with expected values of 0.5 for untwinned data and 0.375 for perfectly twinned data; Padilla & Yeates, 2003). The corresponding estimate of only 0.14 for the twin fraction did not give confidence that a lower symmetry space group had given rise to near-622 symmetry by twinning. The situation was substantially clarified by repeating the statistical analysis to lower resolution limits in order to exclude more weakly recorded intensities. For data between 20 and 3 Å the value of $\langle|L|\rangle$ was 0.377, nearly matching the value of 0.375 expected for perfect hemihedral (twofold) twinning. When considering only data in a relatively thin shell at moderate resolution (3.5–3.0 Å, encompassing more than 14 000 reflections), the value was $\langle|L|\rangle = 0.341$. Remarkably, this value falls slightly beyond (*i.e.* below) the value for perfect twofold twinning, hinting at some more unusual, more severe form of twinning. The much clearer indications of twinning obtained when the analysis was confined to more moderate resolution limits highlights the potential masking effects that can occur when intensity statistics are analyzed to the outer limits of resolution, where the signal-to-noise ratio may be low.

Given the inference of twinning, molecular replacement was re-examined in several space groups representing lower subgroups of $P6_322$. Searches in the possible twinned space group $P6_3$ with two molecules in the asymmetric unit initially appeared to be successful, but no model could be formulated that yielded satisfactory refinement, even when twinning was assumed. The space group was then considered to be $C222_1$, which is also a subgroup of $P6_322$, where now three monomers of βLG would fill the asymmetric unit. In the $C222_1$ unit cell *Phaser* was able to locate all three of the βLG molecules in the asymmetric unit with very high LLG and TFZ. Two of the molecules were related by an NCS twofold axis as in the structure with PDB code 1beb, while the third was related to another by a crystallographic twofold axis (Fig. 1).

In order for a solution in an orthorhombic (222 symmetry) space group to give rise to near-622 diffraction intensities, a threefold-type twinning would be required; 622 point symmetry is a group product of symmetries 222 and 3. Refinement of the $C222_1$ model against the near-622 observed data was accomplished in *REFMAC*5 with the amplitude

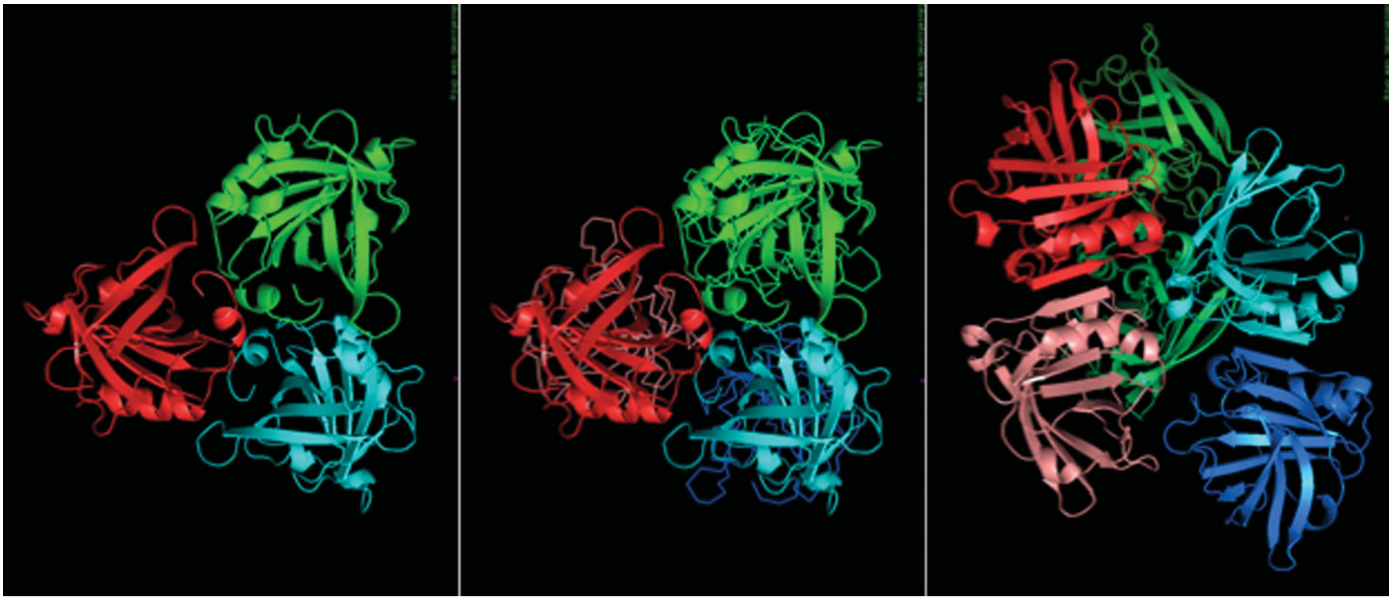
Table 3
Refinement and model statistics.

No. of reflections, working set	22481
No. of reflections, test set	1070 (4.84%)
R_{work}	0.237
R_{free}	0.262
Mean B factor, overall (Å^2)	93.0
R.m.s.d., bond lengths (Å)	0.005
R.m.s.d., angles (°)	1.19
R.m.s.d., chiral volumes. (Å^3)	0.034
Ramachandran outliers	11
Rotamer outliers	0
NCS restraints	Yes
TLS applied	No
B factors	Isotropic
No. of waters	83

twinning option enabled, with the required twin operators being automatically determined [(h, k, l), ($h/2 + k/2$, $3h/2 - k/2$, $-l$), ($-h/2 + k/2$, $3h/2 + k/2$, $-l$)]. The final refined twin fractions were 0.417, 0.305, 0.278. The deviation from perfect (*i.e.* 1/3) twin fractions is consistent with the deviation of the observed intensities from exact 622 symmetry.

The type of twinning observed here is only possible owing to a centered orthorhombic lattice with fortuitous unit-cell values nearly conforming to hexagonal geometry. Specifically, in order for a *C*-centered orthorhombic unit cell to support a pseudo-threefold twinning operation (along the c axis), the a and b unit-cell lengths must be related by a ratio of the square root of 3. In our case, a = 75.5 Å and b = 140.8 Å satisfies this condition. Note also that the reciprocal orthorhombic lattice would obey the same length ratio of the square root of 3 for a^* versus b^*, and so as a result the reciprocal lattice is also necessarily hexagonal in shape. To the degree that the required relationship between the a and b axis lengths is strictly true, then reflections that are not crystallographically related in the orthorhombic space group are superimposed by a 120° rotation about the c axis. Under a growth pathology with crystal subdomains rotated in such a fashion, three crystallographically independent reflections overlap and their separate intensities sum to give the observed intensity. This behavior amounts to pseudo-merohedral twinning. As far as we are aware, threefold pseudo-merohedral twinning has not previously been reported, at least not in macromolecular crystallography. According to the system of prefixes used for other established forms of twinning, *i.e.* hemihedral (twofold) and tetartohedral (fourfold), this newly observed type would be referred to as tritohedral twinning. Because its existence was not anticipated, expected values for the local L statistic for threefold twinning were not established in earlier work (Padilla & Yeates, 2003; Yeates & Tsai, 2012). We rectify this omission here. For perfect tritohedral twinning (twin fractions all equal to 1/3), the expected values are $\langle|L|\rangle = 5/16 = 0.3125$ and $\langle L^2\rangle = 1/7 \simeq 0.1429$. These values conform to the previously observed pattern of rational fractions as a function of twin order [for twin order n, $\langle L^2\rangle = 1/(2n + 1)$]. The values above were further verified here by numerical simulations.

The solution for the three βLG monomers obtained from *Phaser* was refined and, taking into account the twinning described above, produced an initial R of 0.29 after rigid-body refinement of the three molecules. Subsequent refinement using *REFMAC*5, after the addition of 83 water molecules, led to an eventual R of 0.237 and an R_{free} of 0.262 at 2.3 Å resolution. TLS refinement (Winn *et al.*, 2001) failed to improve the model or the R values and was not used in refining the structure. NCS restraints were maintained throughout refinement. Statistics are presented in Table 3. The final r.m.s. deviations were 0.005 Å for bond lengths, 1.19° for angles and 0.034 Å^3 for chiral volumes. If the NCS restraints were relaxed



(*a*) (*b*) (*c*)

Figure 1
Arrangement of molecules in the asymmetric unit. (*a*) Three molecules are shown in cartoon format as viewed along the c axis of the $C222_1$ unit cell, which would correspond to a threefold axis along c were the unit cell truly $P6_322$. (*b*) shows the same set of molecules in the same view, but their twofold-related dimer mate is shown in the background so that all dimers are present. In (*c*) the pseudo-threefold axis is vertical.

or dropped there was little change in the residuals or the geometry; thus, the NCS restraints were maintained in the reported refinement and the final model.

According to *Coot*, there are 11 Ramachandran outliers for the three molecules, or about four per molecule. Two of these, Ala34 and Tyr99, were also present in PDB entry 1beb and were justified there by local structural constraints (Brownlow *et al.*, 1997). Although they may have been carried over by model bias, they are also well supported by the electron density and are probably correct. According to *Coot*, there are no side-chain rotamer outliers.

The polypeptide chain structures of the βLG molecules in these crystals, grown at pH 3.8, are not very different from those in the previous 1.8 Å resolution structure with PDB code 1beb, which suggests that neither the pH nor the high NaCl concentration has a profound effect on the overall structure. The three molecules in the asymmetric unit are very similar, as shown in Supplementary Table S1. There are many changes from PDB entry 1beb in rotamer conformations, however, particularly for surface residues and among the three molecules in the asymmetric unit. These are likely to be owing to differences in the water arrangements, and especially to packing interactions. Indeed, inspection of the three molecules comprising the asymmetric unit shows that the lattice environment of each molecule is indeed quite individual and seems to be sufficient to explain the differences in rotamer and loop conformations among the molecules.

The model with PDB code 1beb, determined from a triclinic crystal diffracting to 1.8 Å resolution (Brownlow *et al.*, 1997), began at amino acid Gln5 (amino acids 1–4 were not seen) and ended at the disulfide bond containing Cys160 (amino acids 161–162 were not seen). In the $C222_1$ crystals, amino acid Thr4 is present in the electron density for molecule *A*. For molecules *B* and *C*, amino acids are present from Leu1 and Val3, respectively. Molecules *A*, *B* and *C* also have electron density for His161.

Supplementary Table S1 shows the results of sequential least-squares superposition of each molecule in the asymmetric unit upon every other. The mean deviations are all in the range 0.17–0.21 Å for the main-chain atoms, with some maximum deviations, principally in loops, of as much as 7.8 Å. Examination of the NCS differences among the six molecules indicates that the segments of greatest variation are residues 45–57, 74–80, 124–145 and 151–162.

The source of the pH-induced Tanford transition has been shown to be owing to movement of the tip of the EF-loop (amino acids 85–90) from an open conformation at higher pH to a closed conformation at acidic pH. In all of the molecules in the $C222_1$ crystals the EF-loop is in the closed conformation, as in PDB entry 1beb, and the loop is very well defined by density. This is in contrast to several other segments, in particular those around Pro79, Gln35–Glu36 and residues 8–10 and 66–67 in molecule *A*, where the dispositions are poorly represented or absent in the electron density. Val128 in all molecules bulges out from the main chain and has no density to support it, calling into question its presence in the sequence. There appear to be no ligands in the calyx of any molecule, as might have been expected from the closed conformation of the EF loop, nor were any metal ions found.

Acknowledgements

The authors wish to thank Huiling Lee for his assistance in the room-temperature data collection at University of California Irvine, James Holton and George Meigs at beamline 8.3.1 at ALS for invaluable help in data collection and processing, and Michael Sawaya and Duilio Cascio for assistance with crystallographic calculations. The authors are also indebted to Dale Tronrud for his generous and always appreciated advice and help. The authors also wish to thank the Co-editor Gil Privé, who made important contributions to the analysis and its explication.

References

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Boland, M. (2011). *Handbook of Food Proteins*, edited by G. O. Phillips & P. A. Williams, pp. 30–55. Cambridge: Woodhead.
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.
Chatterton, D. E. W., Smithers, G., Roupas, P. & Brodkorb, A. (2006). *Int. Dairy J.* **16**, 1229–1240.
Creamer, L. K., Loveday, S. M. & Sawyer, L. (2011). *Encyclopedia of Dairy Sciences*, 2nd ed., edited by J. W. Fuquay, P. F. Fox & P. L. H. McSweeney, pp. 787–794. New York: Academic Press.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Hambling, S. G., McAlpine, A. S. & Sawyer, L. (1992). *Advanced Dairy Chemistry I. Proteins*, 2nd ed., edited by P. F. Fox, pp. 141–190. London: Elsevier Applied Science.
Hernández-Ledesma, B., Recio, I. & Amigo, L. (2008). *Amino Acids*, **35**, 257–265.
Jost, R. (1993). *Trends Food Sci. Technol.* **4**, 283–288.
Juarez, M. & Ramos, M. (2003). *Encyclopedia of Food Sciences and Nutrition*, 2nd ed., edited by B. Caballero, L. Trugo & P. M. Finglas, pp. 5198–5205. New York: Academic Press.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 67–72.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.
Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.
Kontopidis, G., Holt, C. & Sawyer, L. (2004). *J. Dairy Sci.* **87**, 785–796.
Madureira, A. R., Pereira, C. I., Gomes, A. M. P., Pintado, M. E. & Malcata, F. X. (2007). *Food. Res. Int.* **40**, 1197–1211.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
McPherson, A. (1982). *The Preparation and Analysis of Protein Crystals*. New York: John Wiley & Sons.
McPherson, A. & Larson, S. B. (2018). *Acta Cryst.* F**74**, 593–602.
Mills, E. N. C. & Tatham, A. S. (2003). *Encyclopedia of Food Sciences and Nutrition*, 2nd ed., edited by B. Caballero, L. Trugo & P. M. Finglas, pp. 143–150. New York: Academic Press.
Morr, C. V. & Ha, E. Y. W. (1993). *Crit. Rev. Food Sci. Nutr.* **33**, 431–476.

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.* D**67**, 355–367.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.
Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M., Baker, E. N. & Jameson, G. B. (1998). *Biochemistry*, **37**, 14014–14023.
Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.
Sawyer, L. (2013). *Advanced Dairy Chemistry*, Vol. 1A, 4th ed., edited by P. L. H. McSweeney & P. F. Fox, pp. 211–259. New York: Springer.
Sawyer, L. & Kontopidis, G. (2000). *Biochim. Biophys. Acta*, **1482**, 136–148.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.
Yeates, T. O. & Tsai, Y. (2012). *International Tables for Crystallography*, Vol. *F*, 2nd online ed., edited by E. Arnold, D. M. Himmel & M. G. Rossmann, pp. 311–316. Chester: IUCr.