

Structural characterization of a novel monotreme-specific protein with antimicrobial activity from the milk of the platypus

Janet Newman,^{a*} Julie A. Sharp,^b Ashwantha Kumar Enjapoori,^c John Bentley,^a Kevin R. Nicholas,^b Timothy E. Adams^a and Thomas S. Peat^a

Received 5 December 2017

Accepted 12 December 2017

Edited by C. S. Bond, University of Western Australia, Crawley, Australia

Keywords: novel folds; SAD phasing; monotremes; platypus; antibacterial; monotreme lactation protein.

PDB references: monotreme lactation protein, native, space group *P1*, 4v00; native, space group *C2*, 4v3j; iodine soak, space group *P2*₁, 6b4m

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

^aBiomedical Manufacturing, CSIRO, 343 Royal Parade, Parkville, VIC 3052, Australia, ^bInstitute for Frontier Materials, Deakin University, Geelong, VIC 3217, Australia, and ^cSchool of Medicine, Deakin University, Geelong, VIC 3217, Australia. *Correspondence e-mail: janet.newman@csiro.au

Monotreme lactation protein (MLP) is a recently identified protein with antimicrobial activity. It is present in the milk of monotremes and is unique to this lineage. To characterize MLP and to gain insight into the potential role of this protein in the evolution of lactation, the crystal structure of duck-billed platypus (*Ornithorhynchus anatinus*) MLP was determined at 1.82 Å resolution. This is the first structure to be reported for this novel, mammalian antibacterial protein. MLP was expressed as a FLAG epitope-tagged protein in mammalian cells and crystallized readily, with at least three space groups being observed (*P1*, *C2* and *P2*₁). A 1.82 Å resolution native data set was collected from a crystal in space group *P1*, with unit-cell parameters $a = 51.2$, $b = 59.7$, $c = 63.1$ Å, $\alpha = 80.15$, $\beta = 82.98$, $\gamma = 89.27^\circ$. The structure was solved by SAD phasing using a protein crystal derivatized with mercury in space group *C2*, with unit-cell parameters $a = 92.7$, $b = 73.2$, $c = 56.5$ Å, $\beta = 90.28^\circ$. MLP comprises a monomer of 12 helices and two short β -strands, with much of the N-terminus composed of loop regions. The crystal structure of MLP reveals no three-dimensional similarity to any known structures and reveals a heretofore unseen fold, supporting the idea that monotremes may be a rich source for the identification of novel proteins. It is hypothesized that MLP in monotreme milk has evolved to specifically support the unusual lactation strategy of this lineage and may have played a central role in the evolution of these mammals.

1. Introduction

The ability to lactate is a feature that is only found among mammalian lineages and involves a facet of maternal care where mothers secrete a nutrient-rich milk which is delivered to the young *via* the mammary gland. The evolution of lactation is the most efficient and adaptable means of post-natal nutrient provision that has arisen among vertebrates (Blackburn, 1993). Lactation is beneficial for the young because it safeguards against fluctuations in the quantity, quality and toxicity of the adult food supply (Pond, 1977). In addition, the consumption of maternal milk augments protection against infection of the offspring, particularly in monotremes and marsupials. Various hypotheses have been presented in an attempt to elucidate the evolutionary history of mammary-gland development (Blackburn, 1991). Secretions of ancestral mammary glands may have had antimicrobial properties that protected either eggs or hatchlings, and organic components that supplemented offspring nutrition (Oftedal, 2002). The extant monotremes *Ornithorhynchus anatinus* (duck-billed platypus) and *Tachyglossus aculeatus* and *Zaglossus bruijnii* (echidna genera) are regarded as the most ancestral mammals and have a mammary-gland structure

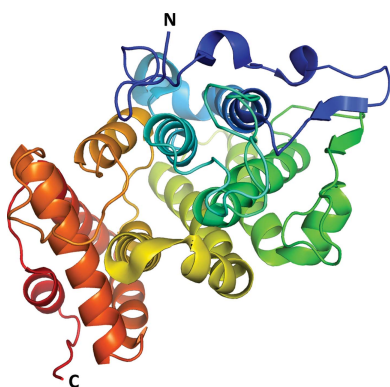


Table 1
Macromolecule-production information.

Source organism	Platypus (<i>O. anatinus</i>)
Expression vector	pTarget (Promega)
Expression host	Human embryonic kidney (HEK) Free-Style 293 cells
Complete amino-acid sequence of the construct produced	MALSCLVFLFTLASVVSQGHVAHPSLGRGDGF PFLWDNAASTLDQLNGTDTTIIILNGFNY LDRLSMFKTVLEGTRKYFDSFAPNNTAN IYWGFIIYLNWILATGRSADPTGHTTCG LAHGDPMLAEESWNCIKYNPAIAFF AAKAGIFGDVTKTIVLAKPKAANSFYC SSEECQAAYPDVMATYLDYFEYLSLE KTGESIDMDKAQQLLWKAHVTSMENSTIA VCKPRLKYNIIERQLDRDYLLSLLYFA ATNFPNFIKFKVADMPHRQLRFGDI APFIIDMDMKNNLLVVLHGFYTVHSL GGSSLTHWRNLMESVPSREARMVNLI LAGTPVEVQVELAKLGIPTVVDYKDDDD K

which lacks a nipple. They are the only mammals that begin early developmental stages in a small (~15 mm) external egg covered by a soft leathery shell that is incubated outside the mother’s body. The hatchling consumes milk from a mammary pad located on the abdomen of the mother, and the milk composition is appropriate for the development of the young and protection from pathogenic infection in the nursing environment (Griffiths, 1978). Monotremes diverged from the mammalian lineage 166 million years ago (MYA), and by 148 MYA mammals had diverged into three distinct groups: the eutherian mammals, the marsupial mammals and the monotremes (Warren *et al.*, 2008; Fig. 1). The mammary gland in marsupials and eutherians evolved with a teat that allows the young to suckle directly from the mammary gland, potentially reducing exposure to microbial infection.

Studies describing the composition and function of monotreme milk components have received limited attention compared with similar studies involving marsupials and eutherians. We previously reported that a highly expressed monotreme lactation protein (MLP) was a major protein present in the milk of the echidna and platypus throughout lactation and showed that this protein had antibacterial properties. Cell-based studies showed that MLP is regulated by lactogenic hormones: insulin, dexamethasone and prolactin (Enjapoori *et al.*, 2017). MLP is unique to the monotreme lineage and we hypothesize that the mammary glands secrete

Table 2
Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	SD-2
Temperature (K)	281
Protein concentration (mg ml ⁻¹)	4
Buffer composition of protein solution	100 mM Tris pH 8.0, 150 mM NaCl
Composition of reservoir solution	0.068 M ammonium acetate, 25.1% (w/v) PEG 8K, 0.1 M sodium cacodylate buffer pH 6.6
Volume and ratio of drop	400 nl + 400 nl
Volume of reservoir (µl)	50

this protein into milk to protect the hatchlings from microbes as they ingest the milk from the areolar skin surface of the pouch (in echidna) or the incubatorium (in platypus). As this protein is not present in other mammalian species, we speculate that it may be required to protect the hatchling from bacterial infections in the absence of a teat. MLP therefore may have played a central role in the evolution of lactation to support a reproductive strategy that utilized a mammary gland with a teat. We present here the crystal structure of MLP and show it has a unique fold that may reveal a unique biological function.

2. Experimental

2.1. Plasmid construction

The assembly of a mammalian expression vector encoding a C-terminally FLAG-tagged MLP has been described previously (Enjapoori *et al.*, 2014).

2.2. Protein purification

The C-terminally FLAG-tagged recombinant MLP was produced following transient transfection of scaled-up cultures of suspension-adapted FreeStyle 293 cells (Life Technologies) grown in the presence of 5 µM kifunensine (Cayman Chemical; Ren *et al.*, 2016) and was purified as described previously (Enjapoori *et al.*, 2014). The protein appeared as a single band at ~40 kDa on a Coomassie-stained gel and was concentrated to 9.5 mg ml⁻¹ in Tris-buffered saline solution (TBS; 100 mM Tris pH 8.0, 150 mM NaCl) for initial crystallization trials. The protein was tested for stability using a standard differential scanning fluorimetry (DSF) assay

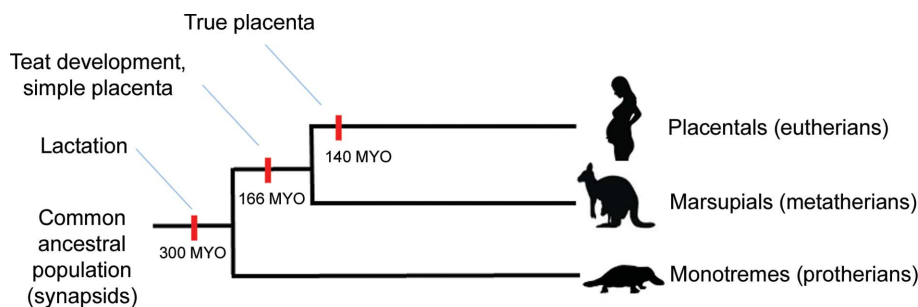


Figure 1

Evolution of lactation and teats. Lactation evolved in the synapsids, while teat development occurred in the therians after the split of the protherian and metatherian lineages. Of the features shown in this phylogeny, only lactation is a characteristic that is common to the entire group.

Table 3

X-ray data statistics.

Values in parentheses are for the highest resolution shell.

	Hg soak	Native, $P1$	Native, $C2$	Iodine soak, $P2_1$
Data collection				
Space group	$C2$	$P1$	$C2$	$P2_1$
a, b, c (Å)	92.7, 73.2, 56.5	57.2, 59.7, 63.1	92.4, 73.0, 56.7	56.9, 79.3, 98.4
α, β, γ (°)	90.0, 90.3, 90.0	80.2, 83.0, 89.3	90.0, 90.1, 90.0	90.0, 92.3, 90.0
Resolution (Å)	2.20 (2.32–2.20)	1.82 (1.86–1.82)	1.97 (2.08–1.97)	2.50 (2.64–2.50)
R_{merge}	0.101 (0.636)	0.089 (0.814)	0.075 (0.737)	0.116 (0.731)
$R_{\text{p.i.m.}}$	0.022 (0.183)	0.054 (0.533)	0.023 (0.263)	0.045 (0.292)
$\langle I/\sigma(I) \rangle$	26.0 (4.3)	15.1 (2.2)	22.2 (2.7)	14.4 (2.8)
$CC_{1/2}$	0.999 (0.886)	0.998 (0.711)	0.999 (0.807)	0.997 (0.839)
Completeness (%)	95.4 (72.3)	99.5 (91.2)	99.1 (97.8)	94.6 (73.5)
Multiplicity	21.2 (12.6)	7.4 (6.2)	11.2 (7.7)	7.6 (7.1)
Estimated maximum resolution, anomalous (Å)	2.80 [$CC_{1/2} > 0.15$]			
Refinement				
No. of protomers		2	1	2
Resolution (Å)		61.7–1.82	57.3–1.97	48.4–2.50
Unique reflections		69095	24961	27276
$R_{\text{work}}/R_{\text{free}}$ (%)		16.6/18.8	17.2/21.1	18.4/22.0
No. of atoms				
Total		6267	2957	5644
Protein		5801	2818	5541
Water		448	139	87
B factors (Å ²)				
Overall		21.0	37.7	37.8
Protein		20.8	37.9	38.5
Water		27.4	34.5	30.9
R.m.s. deviations				
Bond lengths (Å)		0.009	0.017	0.014
Bond angles (°)		1.162	1.576	1.497
PDB code		4v00	4v3j	6b4m

(Seabrook & Newman, 2013). Briefly, the protein was diluted 60-fold into 13 different buffer/pH solutions at two concentrations of NaCl, SYPRO Orange dye (Sigma, catalogue No. S5692) was added and the fluorescence was measured as the temperature was ramped up from room temperature. Data are summarized in Table 1.

2.3. Crystallization

Initial crystallization trials were set up in four screens (PACT, JCSG, PSgradient and Morpheus_C3; the contents of these screens can be found at <http://c6.csiro.au>) at 281 K and one screen (JCSG) at 293 K. The droplets consisted of 150 nl protein solution and 150 nl reservoir solution and were set up in SD-2 plates (Molecular Dimensions, England). Crystals grew overnight and suggested that the protein crystallizes with high-molecular-weight polyethylene glycols at 281 K. After several rounds of optimization, which included fine screening with PEG 6K, 8K and 10K and various buffers at neutral or basic pH, microseeding, drop-size and drop-ratio variation, and *in situ* treatment with PNGaseF enzyme (produced in-house), a crystal was produced that diffracted X-rays to beyond 2 Å resolution. The crystallization condition consisted of 0.068 M ammonium acetate, 25.1% (w/v) PEG 8K, 0.1 M sodium cacodylate buffer pH 6.6 (Table 2). The crystallization droplet consisted of 400 nl protein solution and 400 nl crystallant solution in an SD-2 sitting-drop plate. Crystals appeared after two weeks and grew to full size within a further two weeks. The crystals used for data collection were

stabilized by the addition of 1 µl 32% (w/v) PEG 8K with 20% glycerol before being harvested and flash-cooled in liquid nitrogen.

2.4. Data collection and structure solution

Phase information was obtained from a mercury derivative that was obtained from a 24 h soak of a crystal of PNGaseF-treated MLP protein grown in 0.389 M sodium chloride, 25.5% (w/v) PEG 8K, 0.1 M sodium HEPES pH 7.9. The soaking solution was prepared by adding a small amount of solid ethylmercury phosphate to 20 µl 32% (w/v) PEG 8K, and 1 µl of this solution was layered over the droplet containing the crystal. The mercury crystal was briefly back-soaked by transferring it into 1 µl 32% (w/v) PEG 8K with 20% glycerol.

The crystals were quite variable, with some crystals giving no diffraction at all and some diffracting to beyond 2 Å resolution. At least three space groups were observed ($P1$, $C2$ and $P2_1$). The 1.82 Å resolution native data set was indexed in space group $P1$, with unit-cell parameters $a = 51.2$, $b = 59.7$, $c = 63.1$ Å, $\alpha = 80.15$, $\beta = 82.98$, $\gamma = 89.27^\circ$. The mercury-derivative crystal adopted space group $C2$, with unit-cell parameters $a = 92.7$, $b = 73.2$, $c = 56.5$ Å, $\beta = 90.28^\circ$. A second crystal in space group $C2$ diffracted further (1.97 Å resolution) and was used to fully refine the structure in this space group (see Table 3). The $P2_1$ crystal diffracted to 2.50 Å resolution and had unit-cell parameters $a = 56.9$, $b = 79.3$, $c = 98.4$ Å, $\beta = 92.3^\circ$.

Data for all crystal forms and derivatives were collected on the MX2 beamline at the Australian Synchrotron from crystals cryocooled to 100 K. The native data were collected at a wavelength of 1.4586 Å (8500 eV) and $2 \times 360^\circ$ of data were collected from the same crystal by moving up the thin rod. As the space group was $P1$, this gave 99.5% complete data with more than sevenfold multiplicity to 1.82 Å resolution. For the mercury derivative we collected $3 \times 360^\circ$ of data at a wavelength of 0.9919 Å (12 500 eV), again by moving up the rod-like crystal. This crystal was in space group $C2$ and resulted in a 95.4% complete data set with greater than 21-fold multiplicity (95.1% completeness and 10.8-fold multiplicity for the anomalous data) to a resolution of 2.20 Å. Another crystal in space group $C2$ with similar unit-cell parameters diffracted to beyond 2.0 Å resolution and was used to refine the structure in this space group. A third crystal that had been soaked with sodium iodide was found to belong to space group $P2_1$ and diffracted to 2.5 Å resolution using a wavelength of 1.4586 Å. *XDS* (Kabsch, 2010) was used to index the reflections and *AIMLESS* (Evans & Murshudov, 2013) was used for space-group determination and data reduction for all crystals. Six Hg sites were found for the mercury derivative using the anomalous data and were used to solve the structure with *Auto-Rickshaw* (Panjikar *et al.*, 2005). A partial structure was built by *Buccaneer* (Cowtan, 2006) and this was used as a model for *Phaser* (McCoy *et al.*, 2007) to determine the native structure to 1.82 Å resolution. *Coot* (Emsley *et al.*, 2010) was used to manually rebuild the model and this was followed by refinement with *REFMAC* (Murshudov *et al.*, 2011). Final statistics for the structure are given in Table 3. The structure has no Ramachandran outliers, no poor backbone angles or bonds and only six poor rotamers for the 729 side chains (two protomers in the asymmetric unit) in the $P1$ structure (according to the *MolProbity* server; Chen *et al.*, 2010). The structures in the other two space groups were determined by molecular replacement using *Phaser*, manually rebuilt using *Coot* and refined using *REFMAC*. Images were created using the freeware version of *PyMOL* (Schrödinger). The structures and structure factors have been deposited with PDB codes 4v00 for the native $P1$ structure, 4v3j for the $C2$ structure and 6b4m for the $P2_1$ structure.

2.5. Antibacterial assay

To investigate the antibacterial activity of recombinant MLP, we performed an antibacterial assay as described previously (Enjapoori *et al.*, 2014) but with three different concentrations of protein. In brief, *Staphylococcus aureus* ATCC 25923 and *Enterococcus faecalis* ATCC 10100 strains were cultured in Iso-Sensitest broth (Oxoid) from overnight cultures to mid-exponential phase (OD_{600} of 0.6) and were treated with recombinant MLP in a dose-dependent manner at concentrations of 1, 10 and $50 \mu\text{g ml}^{-1}$. Statistical significance was calculated using one-way ANOVA and a *post hoc* Tukey HSD multiple comparison.

3. Results and discussion

Recombinant MLP was purified from an 0.8 l culture supernatant of transiently transfected FreeStyle 293 cells (Thermo Fisher) grown in the presence of kifunensine; approximately 10 mg of protein was recovered, with the majority of the protein migrating as a monomer when analyzed by gel-filtration chromatography, together with a small amount of noncovalently linked dimer (data not shown). The functional status of the protein was assessed using previously described bacterial growth-inhibition assays (Enjapoori *et al.*, 2014). Higher concentrations of MLP ($50 \mu\text{g ml}^{-1}$) were effective at inhibiting the growth of two susceptible bacterial strains, *S. aureus* and *E. faecalis* (Fig. 2). DSF analysis of the MLP protein showed that the protein is stable, with a melting temperature (T_m) of over 331 K in TBS (Fig. 3). Recent work

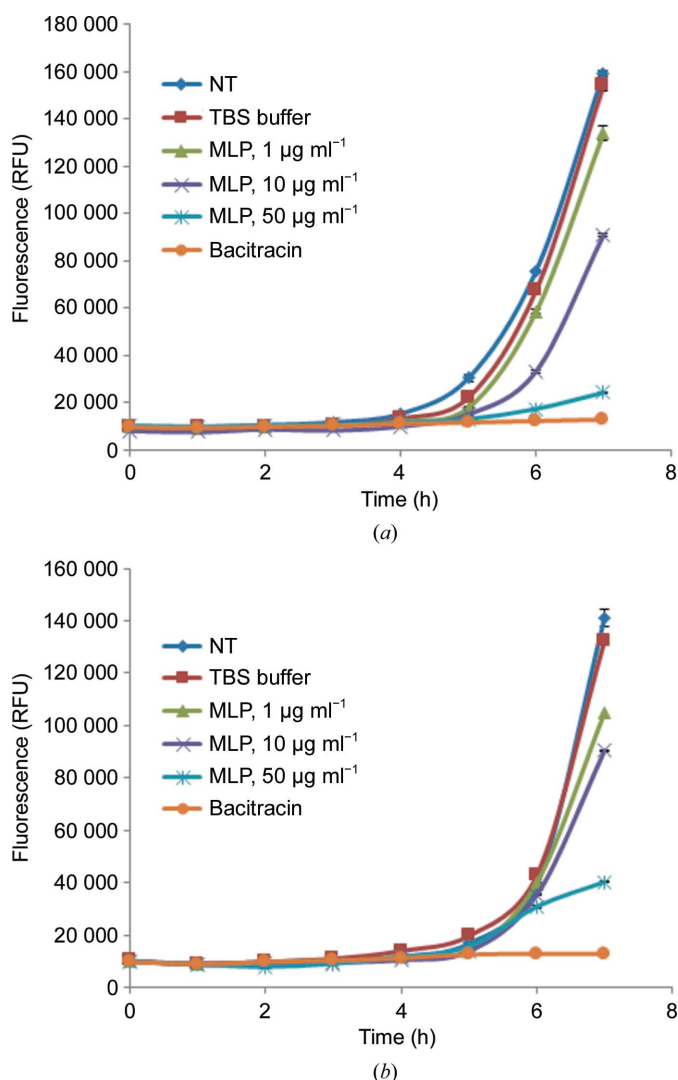


Figure 2 Antibacterial activity of recombinant MLP protein. Bacteriostatic assay using purified MLP protein (1, 10 and $50 \mu\text{g ml}^{-1}$) compared with no treatment (NT) and buffer only (TBS) in the presence of (a) *S. aureus* and (b) *E. faecalis*. The orange curves in both (a) and (b) are for the positive control bacitracin. Statistically significant * $p < 0.05$ or ** $p < 0.01$ compared with NT and TBS at 7 h. Each treatment was performed in triplicate. Standard error bars are shown.

by Dupeux and coworkers suggest that for those proteins that show a clean melt transition by DSF the average T_m is close to 324 K (Dupeux *et al.*, 2011); thus MLP shows a greater stability than average for this technique. The stability

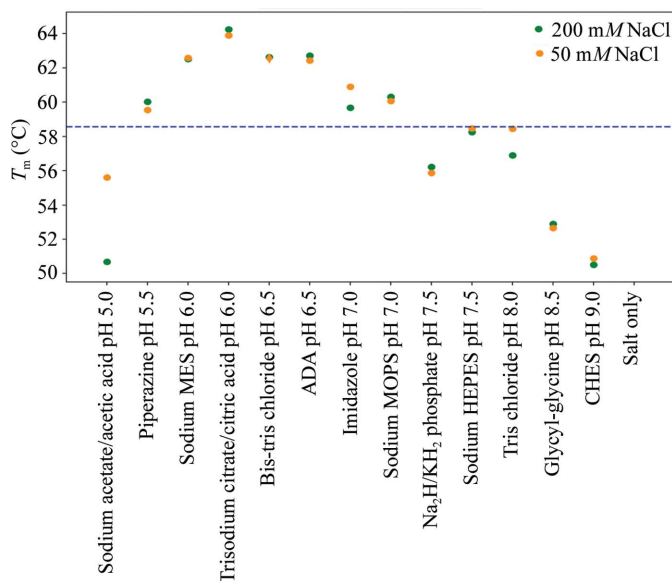


Figure 3

Summary of the results of a systematic buffer screening using differential scanning fluorimetry. Each pH/buffer was tested in triplicate, with either high salt (200 mM NaCl, green spots) or low salt (50 mM NaCl, yellow spots). The protein in the purification buffer (TBS) gave a T_m of $58.55 \pm 0.33^\circ\text{C}$. The highest T_m of $64.25 \pm 0.26^\circ\text{C}$ was seen with 50 mM sodium citrate buffer pH 6 with 200 mM NaCl. The analysis of the melting curves was performed with the program *Meltdown* (Rosa *et al.*, 2015).

maximum (337.5 K) of the MLP protein was at pH 6.5, which matches the pH of mammalian milks (6.4–6.7; Park & Haenlein, 2008).

MLP crystals were thin rods, varying in length from 50 to 300 μm , but rarely more than 20 μm in the other two dimensions. Crystals were often intergrown or split. The native data set came from a crystal (dimensions $90 \times 20 \times 20 \mu\text{m}$) grown from *in situ* PNGase-treated MLP protein at 4 mg ml^{-1} in TBS (Supplementary Fig. S1).

The refined map for MLP has interpretable density at 1.5σ for 345 residues which form ten major (long) helices, two short helical sections (12 helical regions in total) and two short β -strands (residues 49–53 and 156–160), with much of the N-terminus composed of loop regions (see Fig. 4). One long helix (starting at Asp205) has a decided kink at residue 230 owing to the inclusion of a proline residue at this position. Most of the amino-acid sequence is very well defined, with an average B factor of 21 \AA^2 . In the *P1* structure there is one turn in one of the two protomers (*B* chain), at residues 198–202 that connect helices 5 and 6, which is clearly more mobile than most of the protein, and residues 198–202 have B factors that range from 50 to 70 \AA^2 . The C-terminus of the *A* chain also has some weak density for the last eight residues. There are three disulfide bonds between residues Cys113–Cys122, Cys131–Cys228 and Cys170–Cys176. Although this structure has two protomers in the asymmetric unit, the buried surface area is not sufficient to suggest a biologically important dimer interface (the largest buried surface area is about 510 \AA^2 from *PISA*; Krissinel & Henrick, 2007). We have modelled a single saccharide moiety on Asn82 in both protomers; although there

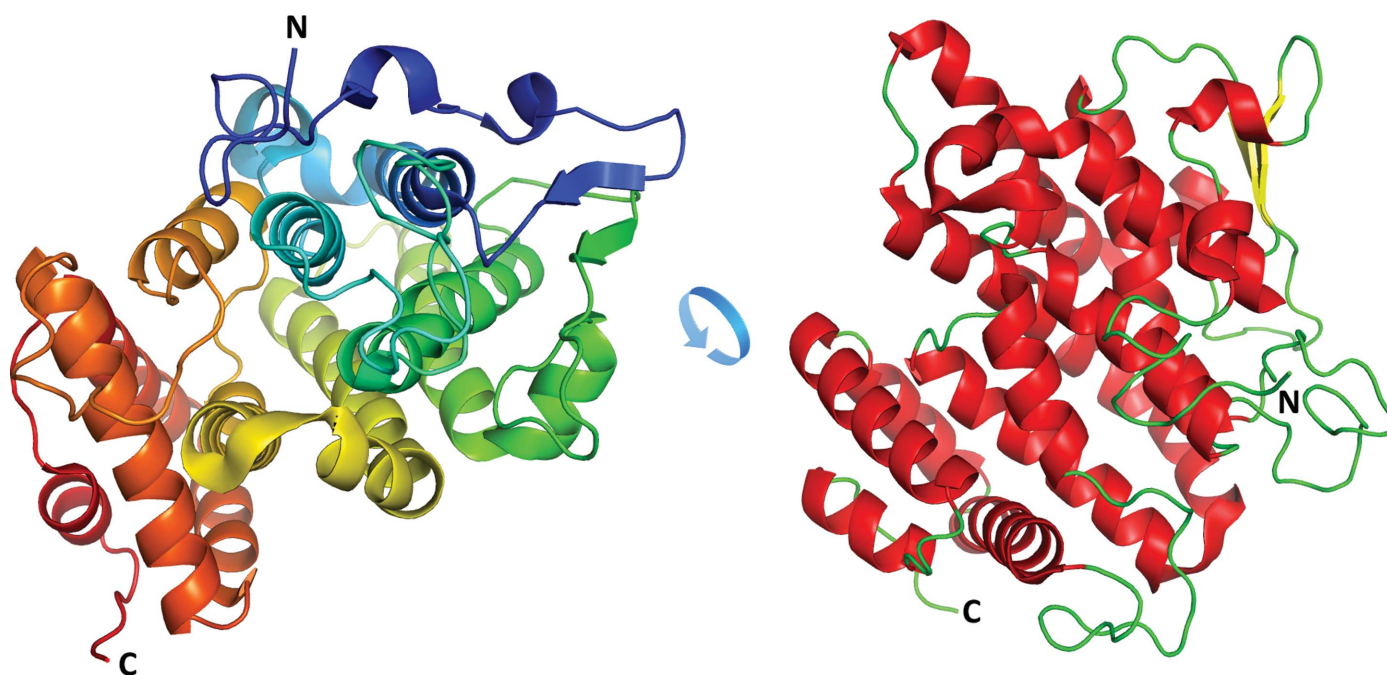


Figure 4

The left image shows a rainbow ribbon representation of MLP showing how the protein folds from the N-terminus (N) to the C-terminus (C). The image on the right is a representation based on secondary structure with helices in red, β -sheet in yellow and loops in green. The right-hand image is rotated approximately 90° around a 45° rotational axis to give a clear view of the N-terminal region, which is largely formed from loops, whereas the rest of the protein is predominantly helical. The secondary-structure assignment was automatically generated in *PyMOL*.

is some extra density beyond this saccharide, there was not enough to fit another full saccharide moiety.

After modelling and refinement the largest difference map peaks are found in a hydrophobic pocket of the *P1* structure that is lined primarily with aromatic amino acids: Phe91, Phe253, Phe258, Phe268, Tyr94, Tyr246 and Trp214. Some additional amino acids are found lining the pocket and include Leu250, Leu296, Leu300, Met272, His217, Thr260, Ser265 and Gly90. It is unclear from the density seen in the pocket what the compound or compounds might be; the density is strongest in the *A* protomer and is left unmodelled, whereas water has been added to the *B* protomer. Whatever resides in the pocket is likely to be hydrophobic in nature as the pocket is essentially lined with hydrophobic residues. There is no obvious channel in the structure to allow access/egress from this pocket to the bulk solvent.

The two protomers in the *P1* structure can be aligned with an r.m.s.d. of 0.27 Å for the C^α atoms. The protomer found in the *C2* crystal form can be superposed with the *A* and *B* protomers of the *P1* structure with r.m.s.d.s of 0.33 and 0.27 Å, respectively. Superposition with the protomers in the *P2*₁ structure also leads to an r.m.s.d. of about 0.3 Å in each case. The packing seen in each crystal form is different, although some interfaces are shared between all three forms. For example, the residues around the C-terminus including Ser313 are always found as crystal contacts, as are residues Pro343 and Met291, whereas other residues (for example Arg278 and His274) are only sometimes seen in the crystal-packing interfaces.

In the *C2* structure both the turn between helices 5 and 6 (residues 198–202) and the two β-strand regions (residues 45–52 and residues 153–159) have weaker density and higher *B* factors than average for this structure. Similar to the *P1* structure, there is density for a single saccharide molecule attached to Asn82. One difference between the two structures is that in the *C2* structure there is extra density around Asn295 which is part of a crystal contact in the *C2* space group and cannot be glycosylation as there is not sufficient space to accommodate a sugar moiety in this position.

The *P2*₁ structure is at lower resolution (2.5 Å) and has two protomers in the asymmetric unit. Like the other two structures, there is some density for a saccharide moiety on residue Asn82. This density is stronger for the *B* chain and there is some additional density beyond the first saccharide to suggest that there is at least one more saccharide in the chain. The density for residues 198–203 in the *A* chain is weaker than for most of the rest of the protein, suggesting mobility in this hinge from helix 5 to helix 6. This crystal had been soaked with sodium iodide in an attempt to obtain phase information to solve the structure. Unfortunately, little anomalous diffraction was obtained from this crystal at the X-ray wavelength used (1.4586 Å). Several high-*B*-factor I atoms were modelled into the structure as the density was too significant to be modelled as water molecules.

Although there is some extra density in the same pocket for the *C2* structure, it is not nearly as extensive. There is little excess density in this pocket for the *P2*₁ structure and water

molecules have been modelled in this region. Fig. 5 shows a surface representation of the MLP structure coloured by *B*-factor value and shows that there are no obvious regions of mobility, although there seems to be a ‘girdle’ around the bottom of the protein which is particularly stable and has a shallow pocket associated with it.

Manual inspection and analysis by the PDBe *PISA* (Krisinel & Henrick, 2007) server suggests that the protein is a monomer. Gel-filtration analysis found one major peak with a shoulder, suggesting that the protein is a monomer in solution but may form a dimer at high concentrations. In addition, the protein appears to be a novel fold as both the *PDBeFold* (Krisinel & Henrick, 2004) and *DALI* (Holm & Laakso, 2016) servers returned no other known protein structure that matched the fold of MLP. Furthermore, this protein was submitted for structure prediction to the 11th CASP challenge, where no groups were successful in predicting the structure (Kryshtafovych *et al.*, 2015). Secondary-structure predictions correctly predicted that the protein was mostly α-helical, but the three-dimensional modellers found it to be extremely difficult to obtain the correct fold for this sequence. The models were of poor quality and were found to be below the level acceptable for practical usability (GDT_TS scores of 17 or lower).

Limited studies of antimicrobial proteins in monotremes have revealed the presence of several antimicrobials in milk that are common to all mammals (Lefèvre *et al.*, 2009). These include lactoferrin, lactalbumin, transferrin, WAP four-disulfide core domain protein 2 and lysozyme, while two antimicrobial proteins, MLP and echAMP (Bisana *et al.*,

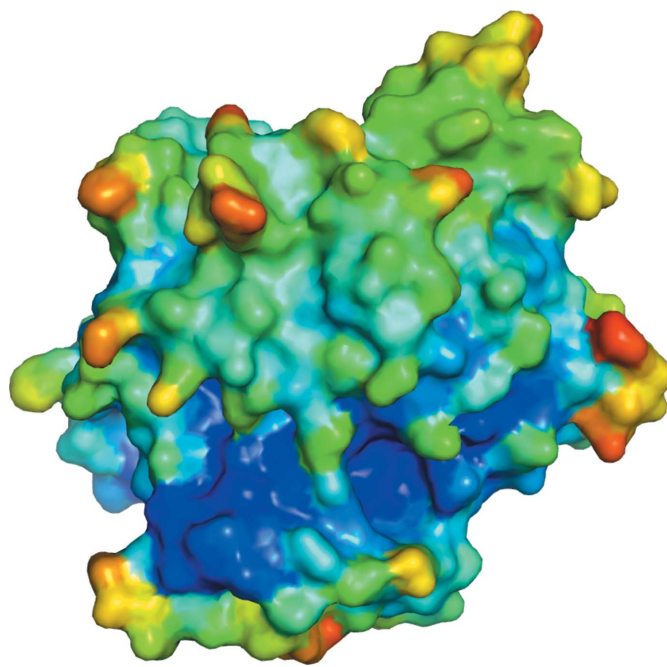


Figure 5 Surface representation of MLP coloured by *B* factor: the lowest *B* factors are in dark blue with the highest *B* factors in red. There is a girdle of low-*B*-factor residues situated at the bottom half of the protein and this corresponds to a small pocket in the surface of the protein.

2013), are specific to the monotreme lineage. The adaption of the teat in therians (*i.e.* marsupials and eutherians) may have led to the loss of some of the antimicrobial bioactivity in milk, including these two proteins. Owing to the unique structure of the MLP protein, we propose that it evolved only in the monotreme lineage owing to the lack of nipples, which placed the young in an environment where they were more vulnerable to bacterial infection. Subsequently, therian lineages evolved to adopt teats for milk delivery, and eutherians later adopted a longer gestation period supported by intrauterine development and placental support, allowing the birth of more mature, immune-competent young (Fig. 1). Milk proteins such as MLP, which are specific to the prototherian lineage, are perhaps an adaption of the monotreme reproductive strategy for the survival of the young and played a central role in the evolution of lactation in monotremes.

4. Conclusions

We have determined the crystal structure of MLP to a resolution of 1.82 Å and found the structure to have a novel fold. There is one loop, residues 198–202, that connects helices 5 and 6, which is clearly more mobile than most of the protein, and residues 198–202 have high *B* factors. The novel fold may reflect an ancient activity in this protein which has not been conserved in other species and may be specific to the monotreme lactation strategy.

The high concentration of MLP in milk is unusual for an antibacterial protein and therefore this novel structure may either indicate a different mechanism of action for protection or suggest other roles in development of the suckled young. This may offer a novel drug composition which could potentially be used as a new therapeutic.

Acknowledgements

We thank the C3 Collaborative Crystallisation Centre for crystals and DSF experiments. We also thank the Australian Synchrotron and beamline staff for beamtime and help with data collection.

References

- Bisana, S., Kumar, S., Rismiller, P., Nicol, S. C., Lefèvre, C., Nicholas, K. R. & Sharp, J. A. (2013). *PLoS One*, **8**, e53686.
- Blackburn, D. G. (1991). *Mammal Rev.* **21**, 81–96.
- Blackburn, D. G. (1993). *J. Dairy Sci.* **76**, 3195–3212.
- 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.
- Cowtan, K. (2006). *Acta Cryst.* **D62**, 1002–1011.
- Dupeux, F., Röwer, M., Seroul, G., Blot, D. & Márquez, J. A. (2011). *Acta Cryst.* **D67**, 915–919.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Enjapoori, A. K., Grant, T. R., Nicol, S. C., Lefèvre, C. M., Nicholas, K. R. & Sharp, J. A. (2014). *Genome Biol.* **6**, 2754–2773.
- Enjapoori, A. K., Lefèvre, C. M., Nicholas, K. R. & Sharp, J. A. (2017). *Gen. Comp. Endocrinol.* **242**, 38–48.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- Griffiths, M. (1978). *The Biology of the Monotremes*. New York: Academic Press.
- Holm, L. & Laakso, L. M. (2016). *Nucleic Acids Res.* **44**, W351–W355.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* **D60**, 2256–2268.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Kryshatovych, A. *et al.* (2015). *Proteins*, **84**, Suppl. 1, 34–50.
- Lefèvre, C. M., Sharp, J. A. & Nicholas, K. R. (2009). *Reprod. Fertil. Dev.* **21**, 1015–1027.
- 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.
- 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.
- Oftedal, O. T. (2002). *J. Mammary Gland Biol. Neoplasia*, **7**, 253–266.
- Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* **D61**, 449–457.
- Park, Y. W. & Haenlein, G. F. W. (2008). *Handbook of Milk of Non-Bovine Mammals*. Ames: Blackwell.
- Pond, C. M. (1977). *Evolution*, **31**, 177–199.
- Ren, B., McKinstry, W. J., Pham, T., Newman, J., Layton, D. S., Bean, A. G., Chen, Z., Laurie, K. L., Borg, K., Barr, I. G. & Adams, T. E. (2016). *Dev. Comp. Immunol.* **55**, 32–38.
- Rosa, N., Ristic, M., Seabrook, S. A., Lovell, D., Lucent, D. & Newman, J. (2015). *J. Biomol. Screen.* **20**, 898–905.
- Seabrook, S. A. & Newman, J. (2013). *ACS Comb. Sci.* **15**, 387–392.
- Warren, W. C. *et al.* (2008). *Nature (London)*, **453**, 175–183.