Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Cloning, overexpression, purification, crystallization and preliminary X-ray analysis of a female-specific lipocalin (FLP) expressed in the lacrimal glands of Syrian hamsters

Ved Prakash Dubey,‡ Biswajit Pal,‡ Subramanya Srikantan,§ Sambhavi Pottabathini, Prabir Kumar De and Rajan Sankaranarayanan*

Centre for Cellular and Molecular Biology, Council of Scientific and Industrial Research, Uppal Road, Hyderabad 500 007, India

‡ These authors contributed equally to this work.
§ Present address: Laboratory of Cellular and Molecular Biology, NIA-IRP, NIH, Baltimore, MD 21224, USA.

Correspondence e-mail: sankar@ccmb.res.in

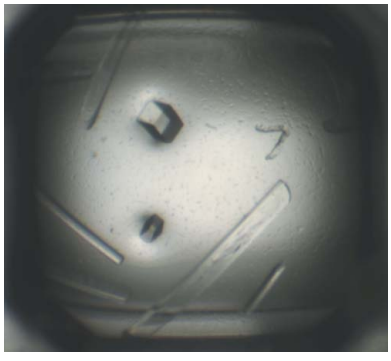
Received 11 January 2010
Accepted 3 March 2010

Proteins belonging to the lipocalin superfamily are usually secretory proteins of molecular mass ~20 kDa with a hydrophobic pocket for the binding and transport of diverse small ligands. Various lipocalins have been associated with many biological processes, *e.g.* immunomodulation, odorant transport, pheromonal activity, retinoid transport, cancer-cell interactions *etc.* However, the exact functions of many lipocalins and the ligands bound by them are unclear. Previously, the cDNA of a 20 kDa lipocalin (FLP) which is female-specifically expressed in the lacrimal glands of Syrian (golden) hamsters and secreted in the tears of females has been identified and cloned. His-tagged recombinant FLP (rFLP) has now been cloned, overexpressed in *Escherichia coli* as a soluble protein and purified to homogeneity using Ni-affinity followed by size-exclusion chromatography. Purified rFLP was crystallized using the sitting-drop vapour-diffusion method. The crystals tested belonged to space group $P2_12_12_1$ and diffracted to beyond 1.86 Å resolution. Solvent-content analysis indicated the presence of one monomer in the asymmetric unit.

1. Introduction

The lipocalin protein superfamily consists of abundant secretory and mostly acidic pI proteins of molecular mass ~20 kDa, which usually have low sequence similarity but contain two or three small conserved motifs and have similar three-dimensional architectures (Flower *et al.*, 2000). The lipocalin fold is a highly symmetrical all-β structure dominated by a single eight-stranded antiparallel β-sheet that is closed back on itself to form a continuously hydrogen-bonded β-barrel which encloses a ligand-binding site. Lipocalins bind small hydrophobic molecules, *e.g.* retinoids, steroids, bilins, siderophores, lipids, odorants, drugs *etc.* (Åkerstrom *et al.*, 2000). Many functions have been proposed for lipocalins, such as the binding and transport of volatile odorants/pheromones, retinoid transport, prostaglandin synthesis, immunomodulatory effects, bacteriostatic effects, anti-inflammatory activity and even allergenic properties (Cavaggioni & Mucignat-Caretta, 2000; Tegoni *et al.*, 2000; Goetz *et al.*, 2002; Lögdberg & Wester, 2000; Mäntyjärvi *et al.*, 2000). However, the exact physiological functions of many lipocalins and their natural ligands are still unknown.

Mammalian lipocalins displaying stringent sex-specific expression in nonreproductive tissues have only been identified in the lacrimal gland of female hamsters (female-specific lacrimal protein; FLP), the submandibular gland of male hamsters (male-specific submandibular protein; MSP) and in the submandibular gland of boars and are secreted in tears or saliva (Srikantan *et al.*, 2005; Thavathiru *et al.*, 1999; Spinelli *et al.*, 2002). Interestingly, in both sexes of humans a lipocalin of obscure function is abundantly expressed and secreted by the lacrimal and lingual salivary glands into tears and saliva (Breustedt *et al.*, 2005). On the other hand, male-predominant lipocalins, synthesized in the liver, are present in rat and mouse urine (Cavaggioni & Mucignat-Caretta, 2000). The crystal structures of the boar salivary and mouse urinary lipocalins have been solved; they bind pheromonal ligands and function in intraspecies chemical communication (Spinelli *et al.*, 2002; Cavaggioni & Mucignat-Caretta, 2000). The structure of human tear lipocalin, which binds ligands



of diverse chemical structure, has also been solved (Breustedt *et al.*, 2005). However, the structures of hamster FLP and MSP, their natural ligands and their functions (if any) in hamster chemical communication are not known.

FLP and MSP are abundantly expressed hamster lipocalins that have been identified in our laboratory. We have purified and characterized the lipocalins, cloned their cDNAs and studied their unusual sex-hormone-repressed regulation in the lacrimal and submandibular glands (Ranganathan & De, 1995; De, 1996; Ranganathan *et al.*, 1999; Thavathiru *et al.*, 1999; Srikantan *et al.*, 2005, 2007). The precursor forms of both FLP and MSP are 172 amino acids long, with 85% sequence identity, and have signal peptides of 16 and 15 residues, respectively. In female hamster lacrimal extracts mature FLP (156 residues) has a molecular mass of 20 kDa on SDS–PAGE and is nonglycosylated (the sequence has no N-glycosylation site), while in male submandibular extracts mature MSP (157 residues with a single N-glycosylation site) is detectable as a 24 kDa (N-glycosylated) and a 20.5 kDa (nonglycosylated) species. Both FLP and MSP have maximum identity to rat odorant-binding lipocalin (∼61%; PDB code 3fiq; White *et al.*, 2009), which is expressed in nasal glands in both sexes; the next highest identity is to aphrodisin (∼39%; PDB code 1e5p; Vincent *et al.*, 2001), which is another hamster lipocalin secreted in vaginal discharge that binds odorants/pheromones and has aphrodisiac activity on male hamsters. Homology-search results, their sex-specific and tissue-specific expression patterns and their presence in tears and saliva suggest that FLP and MSP may have different sex-specific roles in hamster chemical communication that are distinct from that of aphrodisin.

We have initiated a study to determine the three-dimensional structures of FLP and MSP. Here, we report the bacterial overexpression, purification, crystallization and preliminary X-ray analysis of recombinant FLP (rFLP).

2. Experimental methods

2.1. Preparation of overexpression constructs for recombinant FLP

Using the full-length cDNA of FLP (Srikantan *et al.*, 2005; GenBank AF345648.1; gi:13430368) as template and FOE-F (5′-ATCAAT**CATATG**CATTATCAGAATCTTGAAGTC-3′) and FOE-R (5′-ACT**CTCGAG**TTGATCACAATTGTCTGTGGG-3′) primers containing anchored *Nde*I and *Xho*I restriction sites, respectively (indicated in bold), the partial FLP cDNA encoding the complete mature FLP polypeptide (156 residues) was amplified. This cDNA product was cloned within the *Nde*I and *Xho*I sites of pET21a vector and transformed into ultracompetent DH5α *Escherichia coli* cells. The insert within the purified plasmid was verified by sequencing.

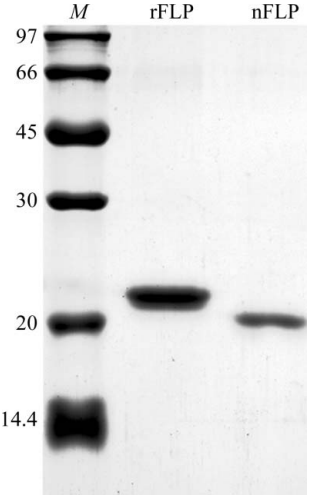


Figure 1
Protein separation in 12% SDS–PAGE under reducing conditions stained using Coomassie Blue. Purified recombinant FLP (rFLP) and natural FLP (nFLP) purified from female hamster lacrimal gland were run separately alongside molecular-mass markers (lane *M*).

Table 1
Crystallographic statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 40.2, b = 60.2, c = 63.5
Unit-cell volume (Å^3)	153672.5
Resolution (Å)	25.0–1.86 (1.93–1.86)
Observations	52592
Unique reflections	13451 (1296)
Completeness (%)	99.7 (97.4)
Redundancy	3.9 (3.7)
$\langle I/\sigma(I)\rangle$	33.4 (9.9)
V_M (Å^3 Da^{-1})	2.0
R_{merge}† (%)	3.9 (10.4)
$R_{r.i.m.}$‡ (%)	4.7 (12.3)
$R_{p.i.m.}$‡ (%)	2.3 (6.3)
Solvent content (%)	38.8
Molecules per asymmetric unit	1
Overall B factor from Wilson plot (Å^2)	17.7

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity and $I_i(hkl)$ is the intensity of an individual measurement of reflection $I(hkl)$. ‡ $R_{r.i.m.} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where N is the number of times a given reflection hkl was observed, and $R_{p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

2.2. Overexpression and purification

The overexpression construct in pET21a was prepared such that only a methionine residue was added at the N-terminal end of the mature polypeptide of FLP and eight extraneous amino acids (leucine and glutamic acid encoded by the *Xho*I site followed by a six-histidine tag) were added at the C-terminal end. The above plasmid construct for bacterial expression was transformed into competent *E. coli* BL21 (DE3) cells. Transformants from a single colony were inoculated into 10 ml LB medium containing ampicillin (100 µg ml^{-1}) and incubated with shaking at 310 K overnight. This preculture was then used to inoculate 1 l LB medium containing ampicillin (100 µg ml^{-1}), which was incubated with shaking at 310 K. After an OD of 0.6 at 600 nm was reached, protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 m*M* and incubation continued for an additional 4–5 h. The bacterial cells were harvested by centrifugation at 6000*g* for 30 min at 277 K and the cell pellet was resuspended in ice-cold lysis buffer (50 m*M* NaH_2PO_4, 300 m*M* NaCl pH 8.0) containing 5 m*M* imidazole, 1 m*M* PMSF and protease-inhibitor cocktail (Roche). Cells were lysed by sonication on ice and the lysate was centrifuged at 25 000*g* and 277 K for 30 min to remove cell debris. The supernatant was passed through an Ni–NTA column (Qiagen) at 277 K pre-equilibrated with lysis buffer containing 10 m*M* imidazole. The column was then washed with lysis buffer containing 20 m*M* imidazole. Elution was performed with 250 m*M* imidazole in lysis buffer. The eluted fractions containing rFLP were pooled, concentrated using a centrifugal filter (Ultracel-5k; Millipore), loaded onto a calibrated gel-filtration column (1.6 × 95 cm; Sephadex G-75 Superfine; GE Healthcare) and eluted with 20 m*M* Tris–HCl pH 7.4 containing 150 m*M* NaCl at 277 K. The eluted fractions were monitored by measurement of the OD at 280 nm and checked for purity by SDS–PAGE. Pure rFLP eluted as a symmetrical peak with an esti-

mated molecular mass of ∼20 kDa. Fractions containing pure rFLP were pooled and concentrated using an Ultracel-5k centrifugal filter. Purified rFLP at a concentration of 25 mg ml^{-1} in 20 m*M* Tris–HCl, 30 m*M* NaCl pH 7.4 was used in crystallization setups.

2.3. Crystallization

Crystal Screen HT (Hampton Research, USA) was used to screen for initial crystallization conditions. The drops were set up using an Oryx 4 crystallization robot (Douglas Instruments) by mixing equal volumes (1 µl) of protein and reservoir solution at 298 K. Either flat-bottomed Greiner or round-bottomed MRC multi-subwell plates were used to set up the sitting-drop vapour-diffusion method. Three different protein concentrations were used in the initial trials. The plates were incubated at 277 K in a Rigaku Roboincubator and imaged using a Minstrel III imager. Well diffracting crystals were obtained within 24 h in several Crystal Screen HT conditions. Crystals from conditions B3, B10 and D5 diffracted well and belonged to the same space group. A crystal from condition B10 [0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 8.5, 30%(*w*/*v*) polyethylene glycol 4000] was directly used for data collection.

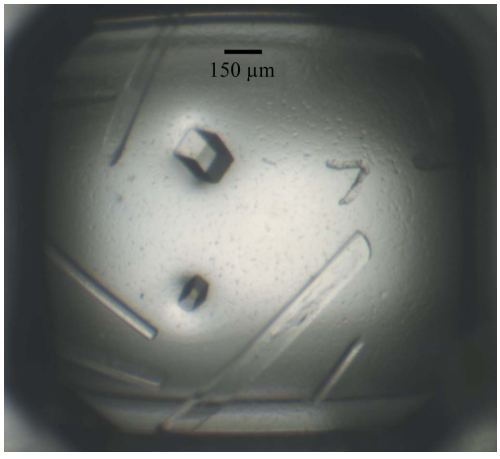


Figure 2
Single crystals of rFLP as imaged by the Minstrel III imager.

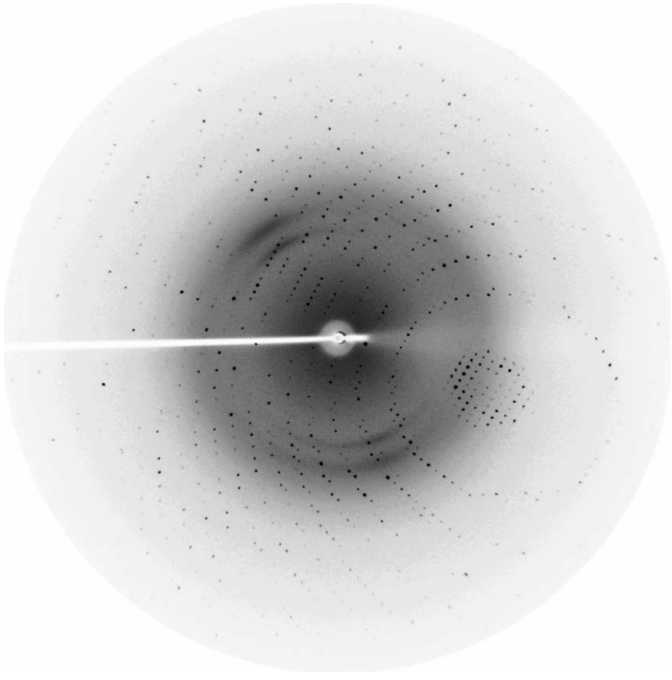


Figure 3
A diffraction image from a rFLP crystal. The edge corresponds to a resolution of 1.86 Å.

2.4. Data collection and processing

An in-house MAR Research MAR345dtb image-plate detector and Cu $K\alpha$ X-rays of wavelength 1.54 Å generated by a Rigaku RU-H3R rotating-anode generator were used to collect diffraction data. The crystal was mounted in a nylon loop and flash-cooled in a nitrogen-gas stream at 100 K using an Oxford Cryostream system. Data were collected using an oscillation angle of 1° and an exposure time of 480 s for each image, with a crystal-to-detector distance of 150 mm. A total of 102° of oscillation data were collected. Indexing, scaling and merging of the data were performed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

3.1. Overexpression, purification and yield of rFLP

FLP was overexpressed in *E. coli* as a soluble protein and purified to homogeneity with a yield of ∼30 mg rFLP per litre of culture. Pure recombinant His-tagged rFLP was obtained by nickel-affinity chromatography of the soluble fraction of the bacterial lysate followed by gel filtration in a pre-calibrated column, in which the rFLP peak had an elution volume that corresponded to an estimated native molecular mass of ∼20 kDa, indicating that it was a monomer. Purified rFLP was found to have a mass of 18 929.9 Da (by mass spectroscopy), which was very similar to the theoretical mass (18 929.8 Da) calculated on the basis of its sequence. In SDS–PAGE (Fig. 1) purified rFLP had a mobility corresponding to an estimated molecular mass of ∼22 kDa, whereas natural FLP purified from female hamster lacrimal gland (theoretical mass 17 737.5 Da) had a mobility corresponding to 20 kDa. This difference must arise from the extraneous amino-acid residues tagged to the C-terminal end of rFLP which are absent in natural FLP (see §2.2).

3.2. Crystallization, data collection and analysis of diffraction data

For crystallization, three different protein concentrations (15, 20 and 25 mg ml^{-1}) were initially tried; each concentration resulted in crystals in several conditions. The 15 mg ml^{-1} protein concentration was used to reproduce results at 277 K. Crystals of different shapes appeared within 24 h (Fig. 2). Several crystals from a wide range of conditions were tested for diffraction. All crystals tested belonged to the same space group with similar unit-cell parameters. Finally, a crystal grown in 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 8.5, 30%(*w*/*v*) polyethylene glycol 4000 was used for data collection. Diffraction data were collected to a resolution of 1.86 Å (Fig. 3). The crystal belonged to space group $P2_12_12_1$ and the mosaicity of the crystal was 0.45°. The unit-cell parameters were found to be a = 40.2, b = 60. 2, c = 63.5 Å. The $R_{\rm merge}$ of the data set was 3.9% (Table 1). Solvent-content analysis indicated the presence of one monomer in the asymmetric unit. Structure solution using the molecular-replacement method is currently in progress.

VPD and SS thank CSIR (India) for research fellowships. RS is the recipient of a Swarnajayanti Fellowship from the Department of Science and Technology (DST), India. PKD thanks DST (India) for a research grant.

References

Åkerstrom, B., Flower, D. R. & Salier, J. P. (2000). *Biochim. Biophys. Acta*, **1482**, 1–8.
Breustedt, D. A., Korndörfer, I. P., Redl, B. & Skerra, A. (2005). *J. Biol. Chem.* **280**, 484–493.
Cavaggioni, A. & Mucignat-Caretta, C. (2000). *Biochim. Biophys. Acta*, **1482**, 218–228.
De, P. K. (1996). *J. Steroid Biochem. Mol. Biol.* **58**, 183–187.
Flower, D. R., North, A. C. & Sansom, C. E. (2000). *Biochim. Biophys. Acta*, **1482**, 9–24.
Goetz, D. H., Holmes, M. A. & Borregaard, N. (2002). *Mol. Cell*, **10**, 1033–1043.
Lögdberg, L. & Wester, L. (2000). *Biochim. Biophys. Acta*, **1482**, 284–297.
Mäntyjärvi, R., Rautiainen, J. & Virtanen, T. (2000). *Biochim. Biophys. Acta*, **1482**, 308–317.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Ranganathan, V. & De, P. K. (1995). *Biochem. Biophys. Res. Commun.* **208**, 412–417.
Ranganathan, V., Jana, N. R. & De, P. K. (1999). *J. Steroid Biochem. Mol. Biol.* **70**, 151–158.
Spinelli, S., Vincent, F., Pelosi, P., Tegoni, M. & Cambillau, C. (2002). *Eur. J. Biochem.* **269**, 2449–2456.
Srikantan, S., Paliwal, A., Quintanar-Stephano, A. & De, P. K. (2007). *Gen. Comp. Endocrinol.* **151**, 172–179.
Srikantan, S., Parekh, V. & De, P. K. (2005). *Biochim. Biophys. Acta*, **1729**, 154–165.
Tegoni, M., Pelosi, P., Vincent, F., Spinelli, S., Campanacci, V., Grolli, S., Ramoni, R. & Cambillau, C. (2000). *Biochim. Biophys. Acta*, **1482**, 229–240.
Thavathiru, E., Jana, N. R. & De, P. K. (1999). *Eur. J. Biochem.* **266**, 467–476.
Vincent, F., Löbel, D., Brown, K., Spinelli, S., Grote, P., Breer, H., Cambillau, C. & Tegoni, M. (2001). *J. Mol. Biol.* **305**, 459–469.
White, S. A., Briand, L., Scott, D. J. & Borysik, A. J. (2009). *Acta Cryst.* D**65**, 403–410.