**Sanjeev Kumar,^{a,b} Rana Zaidi^b
and Samudrala Gourinath^{a*}**^aSchool of Life Sciences, Jawaharlal Nehru
University, New Delhi 110 067, India, and^bDepartment of Biochemistry, Jamia Hamdard
University, New Delhi, IndiaCorrespondence e-mail:
samudralag@yahoo.com,
sgourinath@mail.jnu.ac.in

Received 18 September 2012

Accepted 28 October 2012

Cloning, purification, crystallization and preliminary crystallographic study of calcium-binding protein 5 from *Entamoeba histolytica*

Entamoeba histolytica is the causative agent of human amoebiasis. Phagocytosis is the major route of food intake by this parasite and is responsible for its virulence. Calcium and calcium-binding proteins play major roles in its phagocytosis. Calcium-binding protein 5 from *E. histolytica* (EhCaBP5) is a cytoplasmic protein; its expression is very sensitive to serum starvation and it seems to be involved in binding to myosin I. In this study, EhCaBP5 was cloned, expressed in *Escherichia coli* and purified using affinity and size-exclusion chromatography. The purified protein crystallized in space group C222 and the crystals diffracted to 2 Å resolution. The Matthews coefficient indicated the presence of one molecule in the asymmetric unit, with a V_M of 2.35 Å³ Da⁻¹ and a solvent content of 47.7%.

1. Introduction

The protozoan parasite *Entamoeba histolytica* is the causative agent of amoebiasis, a disease with worldwide occurrence and a high degree of morbidity and mortality, particularly in developing countries. The genomic sequence of this organism has been reported (Loftus *et al.*, 2005), but no metabolite transporter has been identified to date. The main source of its food intake is through phagocytosis. Calcium plays a major role in the progression and completion of phagocytosis, growth and host–parasite interaction resulting in pathogenesis by this organism (Jain *et al.*, 2008).

A genome-sequence analysis showed the presence of 27 calcium-binding proteins in *E. histolytica* with multiple EF-hand motifs (Bhattacharya *et al.*, 2006). It has been shown that calcium-binding protein 1 from *E. histolytica* (EhCaBP1) is involved in cytoskeletal dynamics and that blocking the expression of EhCaBP1 affects the survival of the organism (Sahoo *et al.*, 2004). Furthermore, it has been shown that EhCaBP1 helps in phagocytic cup formation (Jain *et al.*, 2008). A crystal structure of EhCaBP1 reported by our laboratory showed the trimeric N-terminal domain, but the C-terminal half could not be traced owing to a lack of electron density (Kumar *et al.*, 2007). The crystal structure of the N-terminal domain showed the same trimeric arrangement as observed in the structure of full-length EhCaBP1 (Kumar *et al.*, 2010). It has also been shown that the EF-hand motifs are flexible and can accommodate several heavy atoms (Kumar *et al.*, 2012). A mutagenesis study showed that the N-terminal domain performs most of the functions of the full-length protein (Jain *et al.*, 2009).

As indicated above, *E. histolytica* contains many calcium-binding proteins, but structural and functional details are only available for EhCaBP1. Preliminary crystallographic data of EhCaBP2 have been reported previously (Gourinath *et al.*, 2005). NMR models of some other calcium-binding proteins from *E. histolytica* are available, such as EhCaBP2 (PDB entry 2jnx; S. M. Mustafi, R. B. Mutalik & K. V. Chary, unpublished work) and EhCaBP3 (Rout *et al.*, 2010).

In order to understand the calcium-signalling pathways and phagocytic cup formation, structural and functional investigations of various calcium-binding proteins from *E. histolytica* are being carried out. Here, we report the cloning, overexpression, purification, crystallization and preliminary crystallographic analysis of calcium-binding protein 5 from *E. histolytica* (EhCaBP5). Our preliminary results show that EhCaBP5 is a 16.3 kDa calmodulin-like cytoplasmic

© 2012 International Union of Crystallography
All rights reserved

protein and that its expression is very sensitive to serum starvation (unpublished data). Compared with known structures, it has a highest protein sequence identity of only 29% to potato calmodulin (Yun *et al.*, 2004).

2. Experimental procedures

2.1. Cloning of EhCaBP5

The calcium-binding protein 5 gene was amplified by PCR from genomic DNA of *E. histolytica* strain HM1:IMSS. The amplified product was cloned in a pET-28b expression vector between *Nco*I and *Xho*I restriction-enzyme sites with a C-terminal His₆ tag. The following forward and reverse primers were used: CaBP5_Fp, 5'-CATGCCATGGCAATGCAAAAACACAATGAAGAC-3', and CaBP5_Rp, 5'-CGGCTCGAGTTGATTAATAAAATCATCAACTTTTA-3', respectively.

2.2. Overexpression and purification

The EhCaBP5 construct was transformed into *Escherichia coli* strain BL21 (DE3) for expression. Cells were grown in LB medium supplemented with 50 µg ml⁻¹ kanamycin at 310 K. The culture was induced with 0.6 mM IPTG when the OD reached 0.8. It was then kept at the same temperature for 3 h for further growth. Cells were harvested by centrifugation at 7000 rev min⁻¹ for 10 min. The cell pellet was suspended in suspension buffer (50 mM Tris pH 8.0, 30 mM NaCl, 0.2 mM CaCl₂, 5 mM β-mercaptoethanol). The cells were then lysed by freeze-thaw followed by sonication. Clear supernatant was obtained by centrifugation at 12 000 rev min⁻¹ for 30 min. The protein supernatant was passed through an Ni-NTA column pre-equilibrated with ten bed volumes of suspension buffer for affinity purification. It was then washed with 25–30 ml wash buffer (50 mM Tris, 60 mM NaCl, 5 mM β-mercaptoethanol) to remove nonspecific binding. Finally, the protein was eluted with elution buffer (50 mM Tris, 30 mM NaCl, 0.2 mM CaCl₂, 5 mM β-mercaptoethanol, 40 mM imidazole). The Ni-NTA purified protein was concentrated and was subjected to size-exclusion chromatography on a Superdex 75

HiLoad 26/60 column (GE Healthcare, Piscataway, New Jersey, USA), which also exchanged the protein into crystallization buffer (50 mM Tris pH 8.0, 30 mM NaCl, 0.2 mM CaCl₂, 5 mM β-mercaptoethanol). The purity of the protein was checked using 12% SDS-PAGE (Fig. 1). Even though the protein eluted as a sharp single peak in the gel-filtration profile, it showed two bands very close to each other on the SDS gel. Protein identification using MALDI mass spectrometry (data not shown) suggested that both bands are CaBP5, indicating some degradation of the protein. The degradation must be either at the N-terminal end or at a loop in the middle, as the protein still binds to the Ni-NTA column. The peak fractions had different ratios of intact protein and degraded protein. At the beginning of the peak the higher molecular-weight intact protein was predominant, while at the end of the peak the lower molecular-weight degraded protein was predominant (Fig. 1). The purified protein with six residues of the C-terminal His tag was concentrated to 10 mg ml⁻¹ using a 10 kDa cutoff Centricon prior to crystallization.

2.3. Crystallization

Initial crystallization trials were carried out by the hanging-drop vapour-diffusion method in 96-well plates using a crystallization robot. 200 nl protein solution was mixed with an equal volume of precipitant solution and equilibrated against 100 µl reservoir solution (precipitant). The plates were incubated at 289 K. After 10–15 d of equilibration, tiny crystals or needles appeared in Index condition Nos. 24 (2.8 M sodium acetate pH 7.0) and 72 [0.2 M MgCl₂, 0.1 M bis-tris pH 5.5, 25% (w/v) PEG 3350], Crystal Screen condition No. 23 [0.2 M MgCl₂, 0.1 M Na HEPES pH 7.5, 30% (v/v) PEG 400], SaltRx condition No. 2 (0.1 M bis-tris propane pH 7.0, 2.8 M sodium acetate pH 7.0) and Natrix condition No. 1 (0.01 M MgCl₂, 0.05 M MES pH 5.6, 1.8 M lithium sulfate), all from Hampton Research (Aliso Viejo, California, USA). The conditions were further optimized in 24-well plates by experimenting with different pH values in the range 5.0–7.0 and various concentrations of sodium acetate, lithium sulfate and PEG. Diffraction-quality crystals were only obtained in 2.8 M sodium acetate, 0.1 M bis-tris pH 5.5 (Fig. 2).

Figure 1
Purification of EhCaBP5. A gel-filtration profile of EhCaBP5 was obtained by passage through a HiLoad Superdex 75 16/60 column. Inset (a): molecular mass of EhCaBP5 deduced from a standard plot. The standard plot was obtained using the elution profiles of bovine serum albumin (66 kDa), ovalbumin (45 kDa), chymotrypsinogen (25.6 kDa) and lysozyme (14.3 kDa). Inset (b) shows the SDS-PAGE of the purified protein. Lane M, molecular-mass marker (labelled in kDa); lanes 1 and 2, different gel-filtration peak fractions.

Figure 2
Crystals of calcium-binding protein 5 from *E. histolytica*.

Figure 3
Diffraction image of a typical crystal of EhCaBP5.

2.4. X-ray data collection

Crystals were soaked in cryoprotectant solution consisting of 2.8 *M* sodium acetate, 0.1 *M* bis-tris pH 5.5, 20% glycerol. Single crystals were picked up in cryoloops and flash-cooled in liquid nitrogen. X-ray diffraction data were collected on BM14 at the ESRF synchrotron source; the crystals diffracted to 2 Å resolution (Fig. 3). Diffraction data were processed and scaled using *HKL-2000* (Otwinowski & Minor, 1997). The crystals belonged to space group *C222*, with unit-cell parameters $a = 47.5$, $b = 133.6$, $c = 44.5$ Å. Data statistics are reported in Table 1.

3. Results and discussion

Calcium-binding protein 5 from *E. histolytica* was successfully cloned and overexpressed in *E. coli* BL21 (DE3), and the protein was purified to its maximum homogeneity by Ni-NTA chromatography followed by gel-filtration chromatography. The gel-filtration profile indicated the monomeric nature of the protein, with a molecular mass of about 20 kDa, and SDS-PAGE indicated the expected molecular mass of 16.5 kDa. The protein crystallized in various conditions, but diffraction-quality crystals were obtained using 2.8 *M* sodium acetate, 0.1 *M* bis-tris pH 5.5. The EhCaBP5 crystals diffracted to 2 Å resolution and belonged to space group *C222*, with unit-cell parameters $a = 47.5$, $b = 133.6$, $c = 44.5$ Å. The Matthews coefficient V_M is $2.35 \text{ Å}^3 \text{ Da}^{-1}$, indicating the presence of one molecule in the asymmetric unit, with a solvent content of 47.7% (Matthews, 1968). Potato

Table 1

Data-collection statistics.

Values in parentheses are for the outermost resolution shell.

X-ray source	BM14, ESRF
Wavelength (Å)	1.0
Space group	<i>C222</i>
Unit-cell parameters (Å)	
a	47.5
b	133.6
c	44.5
Matthews volume V_M ($\text{Å}^3 \text{ Da}^{-1}$)	2.35
Resolution range (Å)	50–2.0 (2.07–2.00)
$R_{\text{merge}}^{\dagger}$ (%)	9.7 (56)
Completeness (%)	95.1 (93.4)
Total No. of observations	94592
No. of unique observations	9613 (921)
Multiplicity	9.9 (9.0)
Average $I/\sigma(I)$	44.5 (3.5)
Crystal mosaicity (°)	3.5
B value (Wilson plot) (Å^2)	35.4

$$^{\dagger} R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

calmodulin (Yun *et al.*, 2004; PDB entry 1rfj), which shares 29% sequence identity to EhCaBP5 and is the most similar among known structures, was used as a template for molecular replacement, but did not yield a significant solution. Selenomethionyl protein preparation is in progress for heavy-atom phasing and structure solution.

We thank the staff of BM14 at the ESRF and the Department of Biotechnology, Government of India, for access to the beamline. SG is grateful for the Innovative Young Biotechnologist Award and thanks the Department of Biotechnology, Government of India, for funding. SK thanks CSIR for a fellowship. We thank Professor Alok Bhattacharya for providing genomic DNA of *E. histolytica*.

References

- Bhattacharya, A., Padhan, N., Jain, R. & Bhattacharya, S. (2006). *Arch. Med. Res.* **37**, 221–225.
- Gourinath, S., Padhan, N., Alam, N. & Bhattacharya, A. (2005). *Acta Cryst.* **F61**, 417–420.
- Jain, R., Kumar, S., Gourinath, S., Bhattacharya, S. & Bhattacharya, A. (2009). *PLoS One*, **4**, e5269.
- Jain, R., Santi-Rocca, J., Padhan, N., Bhattacharya, S., Guillen, N. & Bhattacharya, A. (2008). *Cell. Microbiol.* **10**, 1373–1389.
- Kumar, S., Ahmed, E., Kumar, S., Khan, R. H. & Gourinath, S. (2012). *BMC Biophys.* **5**, 15.
- Kumar, S., Ahmad, E., Mansuri, M. S., Kumar, S., Jain, R., Khan, R. H. & Gourinath, S. (2010). *Biophys. J.* **98**, 2933–2942.
- Kumar, S., Padhan, N., Alam, N. & Gourinath, S. (2007). *Proteins*, **68**, 990–998.
- Loftus, B. *et al.* (2005). *Nature (London)*, **433**, 865–868.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Rout, A. K., Padhan, N., Barnwal, R. P., Bhattacharya, A. & Chary, K. V. (2010). *Biochemistry*, **50**, 181–193.
- Sahoo, N., Labruyère, E., Bhattacharya, S., Sen, P., Guillén, N. & Bhattacharya, A. (2004). *J. Cell Sci.* **117**, 3625–3634.
- Yun, C., Bai, J., Sun, D., Cui, D., Chang, W. & Liang, D. (2004). *Acta Cryst.* **D60**, 1214–1219.