Acta Cryst. (1998). D**54**, 1442–1445

Crystallization and preliminary X-ray studies of a recombinant calcium-binding protein from *Entamoeba histolytica*

B. Gopal,[a] R. Suma,[a] M. R. N. Murthy,[a]* A. Bhattacharya[b] and S. Bhattacharya[c] at [a]*Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560 012, India,* [b]*School of life Sciences, Jawaharlal Nehru University, N. Delhi-110 067, India, and* [c]*School of Environmental Sciences, Jawaharlal Nehru University, N. Delhi-110 067, India. E-mail: mrn@mbu.iisc.ernet.in*

(*Received 27 May 1997; accepted 29 January 1998*)

Abstract

A calcium-binding protein (CaBP) of *Entamoeba histolytica* was purified from an *E. coli* recombinant clone carrying the *CaBP* gene in a pET-3c expression vector using anion-exchange and size-exclusion chromatography. Examination of the amino-acid sequence of the recombinant protein suggested that it has four independent EF-hand motifs. The protein dissolved in cacodylate buffer was crystallized using the hanging-drop method with 2-methylpentane-2,4-diol (MPD) as the precipitant. X-ray diffraction data have been collected on these crystals using a MAR Research imaging-plate detector system attached to a Rigaku RU200 rotating-anode X-ray generator. The crystals belong to the hexagonal space group $P6_122$ with unit-cell dimensions of $a = b = 96.21$, $c = 65.48$ Å. Preliminary molecular-replacement computations suggest that the structure of this protein is likely to be similar to that of calmodulin (CAM).

1. Introduction

The role of calcium as a secondary messenger has been fairly well established in a variety of eucaryotic systems. A number of studies have shown the involvement of calcium in various metabolic processes of protozoa. It has been suggested that calcium regulates the cellular pathogenesis of *E. histolytica*, based on the ability of calcium antagonists to block the cytopathic function of amoeba and release of secretary collagenase (Ravdin *et al.*, 1985; Munoz *et al.*, 1991). To elucidate the mechanism of calcium signaling and the role of calcium and CaBP in host–parasite relationships, a gene encoding CaBP from *E. histolytica* was cloned into the *E. coli* expression vector pET-3c and expressed in the *E. coli* strain BL21 (DE3)PLys-S (Prasad *et al.*, 1993). An analysis of the amino-acid sequence reveals the presence of four calcium-binding sites, the canonical EF-hand motifs (Table 1). However, the sequence identity of this CaBP with other CaBP's is fairly low (30.2% with a calmodulin from *Drosophila melanogaster*). The ion-binding parameters of this CaBP have been shown to be different to those of other CaBP's (Gopal *et al.*, 1997). Analysis of the three-dimensional structure of the protein could reveal the possible basis for these changes.

CaBP exists in solution as a monomer of molecular mass 14.7 kDa. The holo (calcium-saturated) form of CaBP exhibits a far-UV CD spectrum typical of a protein with a high α-helical content. The apo form of this protein, however, does not seem to have a well defined structure (Gopal *et al.*, 1997), a feature shared by most of the calmodulins. The physiological function of this protein is unknown. However, there is some evidence to suggest that CaBP activates a kinase/kinases and functions as a trigger in the calcium-signalling pathway (Yadava *et al.*, 1997). This paper describes the crystallization and preliminary X-ray diffraction studies on crystals of the recombinant calcium-bound CaBP.

2. Experimental

2.1. *Crystallization*

The purification of recombinant CaBP involves application of the crude cell lysate to a DE52 anion exchanger, followed by size-exclusion chromatography using a Pharmacia Superose HR 10/30 using a Pharmacia FPLC system (Gopal *et al.*, 1997).

The conditions listed by Zeelan *et al.* (1991) were used in the initial crystallization trials. The conditions which appeared promising were further refined by variations in the concentrations of MPD and protein, additives, pH and temperature. The presence of $CaCl_2$ was essential for crystallization. The FPLC elution profile showed a shoulder following the peak corresponding to CaBP. Fractions of the protein corresponding to the peak alone yielded crystals of better quality when compared with the pooled protein. However, the crystals which appeared after a period of 10 d were of size 0.05 × 0.02 × 0.02 mm and did not grow further even after two months. A couple of serial seedings in 63% MPD, 8 mg ml^{-1} protein, 15 m*M* $CaCl_2$ and 50 m*M* cacodylate buffer pH 4.3 led to crystals of dimensions 0.8 × 0.2 × 0.2 mm in 3–4 weeks (Fig. 1).

2.2. *Data collection*

Diffraction data collection was carried out using an X-ray source (Cu $K\alpha$), provided by a Rigaku RU200 rotating-anode generator equipped with a 0.2 mm focusing cup and operating at 40 kV, 58 mA. Data were recorded at 277 K on a MAR Research imaging-plate system at a crystal-to-film distance of 146 mm. Each frame was exposed for 15 min. The data set was processed using the *DENZO*/*SCALEPACK* (Otwinowski, 1993; Minor, 1993) suite of programs. Equivalence of reflection intensities and systematic absences suggested that the space group is either $P6_122$ or $P6_522$. The refined cell parameters and the intensity statistics are listed in Table 2.

2.3. *Molecular-replacement computations*

The presence of four appropriately placed canonical calcium-binding EF-hand motifs in the amino-acid sequence of *E. histolytica* CaBP suggested that its structure is likely to be similar to that of the calmodulins. Alignment of the amino-acid sequence of CaBP with calmodulins from other sources indicated that it is most similar to *D. melanogaster* calmodulin. Hence, the coordinates of *D. melanogaster* calmodulin (PDB code 4CLN) available in the 1996 release of the Protein Data Bank (Bernstein *et al.*, 1977) were used to obtain a tentative

Table 1. *The sequences of the calcium-binding sites of CaBP compared with the EF loops of calmodulin, troponin C and parvalbumin*

Loop position		1	2	3	4	5	6	7	8	9	10	11	12
CaBP (*E. hist*)	Loop 1	D	V	N	G	D	G	A	V	S	Y	E	E
	Loop 2	D	A	D	G	N	G	E	I	D	Q	N	E
	Loop 3	D	V	D	G	D	G	K	L	T	K	E	E
	Loop 4	D	A	N	G	D	G	Y	I	T	L	E	E
Calmodulin	Loop 1	D	K	D	G	D	G	E	V	S	F	E	E
	Loop 2	D	K	D	G	N	G	T	I	T	T	K	E
	Loop 3	D	K	D	G	N	G	Y	I	S	A	A	E
	Loop 4	N	I	D	G	D	G	E	V	N	Y	E	E
Troponin C	Loop 1	D	A	D	G	G	G	D	I	S	T	K	E
	Loop 2	D	E	D	G	S	G	T	I	D	F	E	E
	Loop 3	D	K	N	A	N	G	F	I	D	I	E	E
	Loop 4	D	K	N	N	N	G	R	I	D	F	D	E
Parvalbumin	Loop 1	D	Q	D	K	S	G	F	I	E	E	D	E
	Loop 2	D	S	D	G	D	G	K	I	G	V	D	E

structure of CaBP by molecular-replacement methods. The automated molecular-replacement program *AMoRe* (Navaza, 1994) was used. The solutions obtained had low correlation coefficients and high *R* factors. Further rigid-body refinement of the best solution using *X-PLOR* (Brünger, 1992) failed to reduce the *R* factor significantly.

Examination of the aligned amino-acid sequences of *E. histolytica* CaBP and *D. melanogaster* calmodulin (Fig. 2) revealed that two additional residues occur in the putative linker region of CaBP when compared to *Drosophila* CAM. These residues might result in an extension of the central helix and hence alter both the distance and the relative orientation of the amino- and carboxy-terminal domains. To accommodate the changes in structure that might have resulted from the variation in the length of the linker segment, three additional molecular models were constructed as follows. Residues 66–91 of 4CLN (corresponding to residues 55–80 of CaBP as shown in Fig. 2) constitute the linker helix. The transformation that superposes residues 66–(91 − *i*) onto (66 + *i*)–91 of *Drosophila* CAM was derived using the program *O* (Jones *et al.*, 1991). Residues 85–C-terminal end were transformed using this matrix. After transformation, residue 85 will nearly coincide with residue (85 + *i*). The new coordinates were appended to residues 1 (85 + *i* − 1) of the original 4CLN, resulting in a model *i* residues longer than 4CLN. These synthetic models were used in the molecular-replacement calculations in both $P6_122$ and $P6_522$. The calculations with the model two residues longer (model 2) than 4CLN provided the best results. The *R* factors and correlation coefficients for this and the other models used in the molecular-replacement computations are shown in Table 3. The model with two additional residues not only gave better figures of merit but also had a large difference in the quality of the best and second-best solutions. The transformation required to transfer model 2 to the CaBP cell is as follows,

$$\begin{matrix} X \\ Y \\ Z \end{matrix} = \begin{matrix} 0.65922 & -0.61932 & +0.42647 & +80.28 \\ -0.69245 & -0.72109 & +0.02319 & +50.70 \\ 0.29316 & -0.31060 & -0.90420 & +35.87 \end{matrix}$$

Using this transformation, model 2 was placed in the CaBP cell and all the other molecules in the unit cell were generated using crystal symmetry. Examination of the packing diagram indicated that there were no severe short contacts between

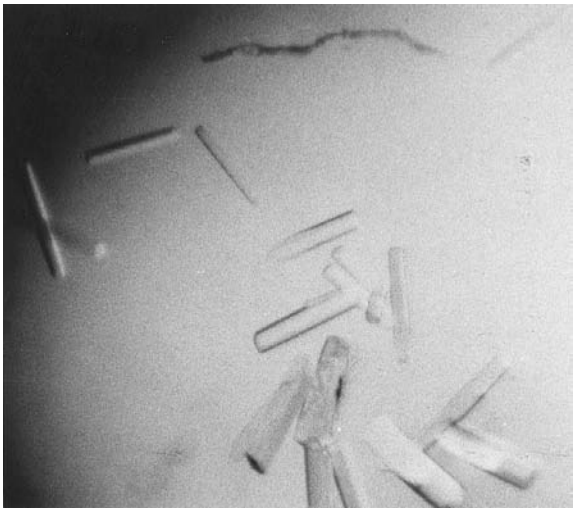

Fig. 1. Crystals of CaBP grown in 64% MPD. Average size of the crystals is 0.6 × 0.2 × 0.2 mm.

```
                                                        10
E. hist CaBP :                          M A E A L F K E I D V N G
D. melo CAM  : A D Q L T E E Q I A E F K E A F S L F D K D G

                        20                      30
E. hist CaBP    D G A V S Y E E V K A F V S K K R A I K N E
D. melo. CAM    D G T I T T K E L G T V M R S L G Q N P T E

                    40                      50
E. hist CaBP    Q L L Q L I F K S I D A D G N G E I D Q N E
D. melo. CAM    A E L Q D M I N E V D A D G N G T I D F P E

                  60                      70
E. hist CaBP    F A K F Y G S I Q G Q D L S D D K I G L K V
D. melo. CAM    F L T M M A R K M K D T D S E E E I - - R E

                80                      90
E. hist CaBP    L Y K L M D V D G D G K L T K E E - - - - V
D. melo. CAM    A F R V F D K D G N G F I S A A E L R H V M

               100                     110
E. hist CaBP    T S F F K K H G I E K V A E Q V M K A D A N
D. melo CAM     T N L G E K L T D E E V D E M I R E A D I D

               120                     130
E. hist CaBP    G D G Y I E L E E F L E F S L
D. melo CAM     G D G Q V N Y E E F V T M M T S K
```

Fig. 2. Sequence alignment of *Entamoeba histolytica* CaBP with the model *Drosophila melanogaster* calmodulin (30.2% identity for the 129 aligned amino acids).

Table 2. *Data collection details*

Cell parameters (Å)	$a = b = 96.21$ $c = 65.48$
Space group	$P6_122$ or $P6_522$
Crystal-to-detector distance (mm)	146
Exposure per frame (s)	900
Oscillation angle (°)	1
No. of frames collected	90
Maximum resolution for processing (Å)	3.5
Processing software	*DENZO/ SCALEPACK*
Total No. of reflections processed	93727
Average No. of times each reflection was measured	11.6
Total No. of unique reflections with $I > 2\sigma(I)$	2324
Completeness of the data (20.0–3.5 Å) (%)	94.1
Completeness of the data (3.6–3.5 Å) (%)	84.0
$R_{\rm merge}$† (%)	12.1
$R_{\rm meas}$‡ (%)	11.7

† $R_{\rm merge}$ is defined as $\sum|I_h - \langle I\rangle|/\sum I_h \times 100$, where I_h is the intensity of the reflection h and $\langle I\rangle$ is the average of measurements for reflection h. ‡ $R_{\rm meas}$ (Diederichs & Karplus, 1997) is defined as $\sum_h (n_h/n_h - 1)^{1/2} \sum_i |\langle I_h\rangle - I_{i,h}|/\sum_h \sum_I I_{h,i}$ with $\langle I_h\rangle = \sum_I I_{h,i}/n_h$.

neighbouring molecules. The few short contacts that were observed were confined to residues 106 and 107. Examination of the sequence alignment (Fig. 2) shows that these two residues correspond to a site of a four-residue deletion in CaBP. Extensive van der Waals contacts are observed between the closest neighbours related by the crystal twofold. The regions that contribute to stabilizing interactions between these molecules include the polypeptide segments in the helices *C*, *G* and the central linker (Fig. 3). When subjected to refinement using *X-PLOR*, the transformed model 2 showed a 10% drop in the *R* factor and free *R* (Brünger, 1993), a feature which none of the other models exhibited. These observations provide strong support to the validity of the molecular-replacement solution.

2.4. *Preliminary electron-density map calculation*

CaBP reflections were phased on the basis of a polyalanine model of 4CLN with two additional residues in the linker region (model 2) suitably placed in the CaBP cell. The resulting phases were used to compute $2F_o - F_c$ and $F_o - F_c$ maps using *X-PLOR*. Examination of these maps revealed good density for a large fraction of the side chains in the N-terminal region and the linker helix. Representative examples are shown in Fig. 4. However, the density corresponding to the carboxy-terminal domain was comparatively more difficult to interpret as there were occasional breaks in the main-chain density and lack of substantial density for most of the side chains.

Table 3. *Comparison of the quality of molecular-replacement solutions with choice of models*

Model type	Solution No.	Correlation coefficient (%)	*R* factor (%)
Native (4 CLN)	1	48.2	49.3
	2	48.8	49.3
Model 1	1	52.2	46.0
(1-residue insertion)	2	50.6	47.7
Model 2	1	56.2	45.2
(2-residue insertion)	2	52.1	48.1
Model 3	1	48.5	48.5
(3-residue insertion)	2	48.0	48.4

3. Results and discussion

Crystals of the CaBP (Fig. 1) were obtained by vapour diffusion using the hanging-drop method (McPherson, 1992). The protein crystallized in about 10 d at room temperature, yielding microcrystals. However, crystals suitable for diffraction studies could be obtained by serial seeding techniques. The crystals belong to the hexagonal space group $P6_122$ with unit-cell dimensions $a = b = 96.21$, $c = 65.48$ Å. The volume of the asymmetric unit (524 903.82 Å^3) is consistent with one CaBP molecule (Matthews number 2.98 Å^3 Da^{-1}).

Attempts to model the structure of CaBP using the coordinates of *D. melanogaster* were not successful. However, a model constructed using the coordinates of *D. melanogaster* CAM with two additional residues in the linker helix provided significantly better molecular-replacement results. This is consistant with the length of the putative central linker in

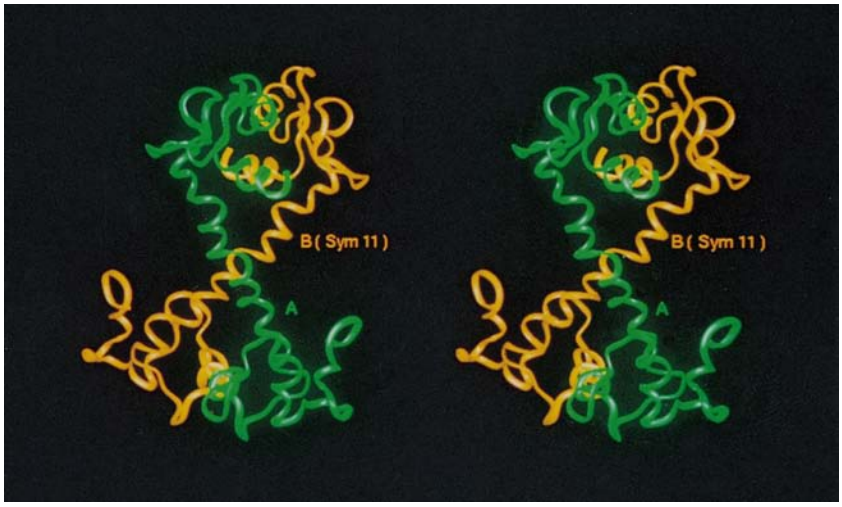

Fig. 3. Arrangement of the twofold-related neighbours. The monomers shown in green and yellow are related by the symmetry operation $(x, x - y, 1/6 - z)$.

CaBP. The molecular-replacement results led to a good packing model of CaBP in the $P6_122$ cell without severe intermolecular short contacts. The closest neighbours in the crystal, shown in Fig. 4, are related by the crystal twofold parallel to a^* at $z = 1/12$. The axes of the linker helices of the CaBP molecules shown in Fig. 3 are at an angle of 80°. The residues that approach closest on the linker are the serines at position 74 in either molecule. Further stabilization is provided by van der Waals contacts between the respective N- and C-terminal domains of the two molecules. These results demonstrate that *E. histolytica* CaBP has a structure similar to that of calmodulin. The additional residues in the putative linker region increase the length of the linker helix and alter the relative orientation between the N- and the C-terminal domains. A similar model has recently been proposed for the structure of the *Arabidopsis* calcium-binding protein (Khan *et al.*, 1997) using molecular modelling followed by energy minimization.

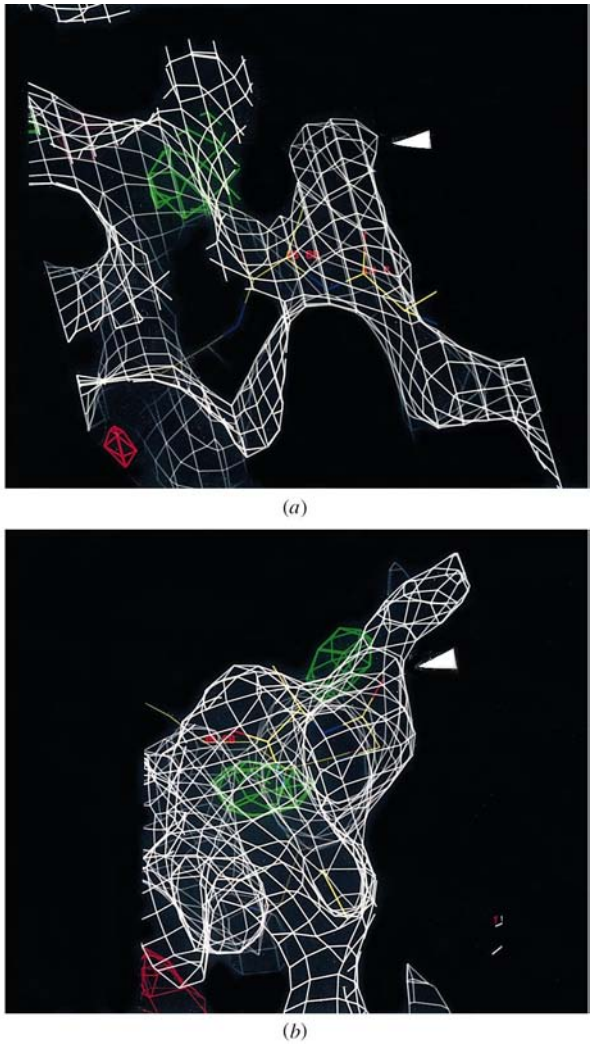

Fig. 4. $2F_o - F_c$ electron-density map at 3.5 Å resolution calculated with phases based on a polyalanine model (model 2) for two representative residues in the N-terminal domain and central linker. The map has clear density for the side chains. (*a*) Leu15, (*b*) Arg40.

The diffraction data obtainable from CaBP crystals are limited to 3.5 Å resolution (Table 2). Hence, the present data are not adequate to obtain detailed structural information on CaBP. The linker region of CaBP contains three glycine residues (Fig. 2). It is likely that the occurrence of glycine residues induces flexibility leading to the poor diffraction quality. Further attempts to improve crystal quality by the addition of reagents that induce helicity and also by replacement of the glycine residues by alanines are in progress.

We wish to thank the National Area Detector Facility, Mr R. Ravishankar, Mr P. F. Babu and Mr James Paul for help during data collection. BG and MRNM wish to thank Professor T. N. Gururow for providing the cryocooler to enable data collection at controlled temperature. This project was supported by grants from the Council of Scientific and Industrial Research and the Department of Science and Technology, Government of India, to MRNM.

References

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Brünger, A. T. (1992). *X-PLOR*. Version 3.1. *A System for X-ray Crystallography & NMR*. New Haven: Yale University Press.

Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.

Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.

Gopal, B., Swaminathan, C. P., Bhattacharya, S., Bhattacharya, A., Murthy, M. R. N. & Surolia, A. (1997). *Biochemistry*, **36**, 10910–10917.

Jones, T. A., Zou, J-Y., Cowan, S. W. & Keldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Khan, A. R., Johnson, K. A., Braam, J. & James, M. N. G. (1997). *Proteins Struct. Funct. Genet.* **27**, 144–153.

McPherson, A. (1992). *J. Cryst. Growth*, **112**, 161–167.

Minor, W. (1993). *XDISPLAYF Program*. Purdue University, West Lafayette, IN, USA.

Munoz, M., De, L., Moreno, M. A. & Hernandez, V. I. (1991) *Mol. Microbiol.* **5**, 1701–1714.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. (1993). *Oscillation Data Reduction Program*, in *Proceedings of the CCP4 Study Weekend*, *Data Collection and Processing*, *29–30 January 1993*, edited by L. Sawer, N. Issacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Prasad, J., Bhattacharya, S. & Bhattacharya, A. (1993). *Cell Mol. Biol. Res.* **39**, 167–175.

Ravdin, J. I., Murphy, C. F., Guerrant, R. L. & Long-Kruz, S. A. (1985). *J. Infect. Dis.* **152**, 542–549.

Yadava, N., Chandok, M. R., Prasad, J., Bhattacharya, S., Sopory, S. K. & Bhattacharya, A. (1997). *Mol. Biochem. Parasitol.* **84**, 69–82.

Zeelen, J. Ph., Hiltunen, J. K., Ceska, T. A. & Wierenga, R. K. (1994). *Acta Cryst.* D**50**, 443–447.