NMR Derived Solution Structure of an EF-Hand Calcium-Binding Protein from *Entamoeba Histolytica*†

H. S. Atreya,‡ S. C. Sahu,‡ A. Bhattacharya,§ K. V. R. Chary,*,‡ and Girjesh Govil‡

Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai-400005, India, and School of Life Sciences, Jawaharlal Nehru University, New Delhi-110067, India

Received July 18, 2001; Revised Manuscript Received September 10, 2001

ABSTRACT: We present the three-dimensional (3D) solution structure of a calcium-binding protein from *Entamoeba histolytica* (*Eh*CaBP), an etiologic agent of amoebiasis affecting millions worldwide. *Eh*CaBP is a 14.7 kDa (134 residues) monomeric protein thought to play a role in the pathogenesis of amoebiasis. The 3D structure of Ca^{2+}-bound *Eh*CaBP has been derived using multidimensional nuclear magnetic resonance (NMR) spectroscopic techniques. The study reveals the presence of two globular domains connected by a flexible linker region spanning 8 amino acid residues. Each domain consists of a pair of helix-loop-helix motifs similar to the canonical EF-hand motif of calcium-binding proteins. *Eh*CaBP binds to four Ca^{2+} with high affinity (two in each domain), and it is structurally related to calmodulin (CaM) and troponin C (TnC) despite its low sequence homology (∼29%) with these proteins. NMR-derived structures of *Eh*CaBP converge within each domain with low RMSDs and angular order-parameters for backbone torsion angles close to 1.0. However, the presence of a highly flexible central linker region results in an ill-defined orientation of the two domains relative to one other. These findings are supported by backbone ^{15}N relaxation rate measurements and deuterium exchange studies, which reveal low structural order parameters for residues in the central linker region. Earlier, biochemical studies showed that *Eh*CaBP is involved in a novel signal transduction mechanism, distinct from CaM. A possible reason for such a functional diversity is revealed by a detailed comparison of the 3D structure of *Eh*CaBP with that of CaM and TnC. The studies indicate a more open C-terminal domain for *Eh*CaBP with larger water exposed total hydrophobic surface area as compared to CaM and TnC. Further dissimilarities between the structures include the presence of two Gly residues (G63 and G67) in the central linker region of *Eh*CaBP, which seem to impart it a greater flexibility compared to CaM and TnC and also play crucial role in its biological function. Thus, unlike in CaM and TnC, wherein the length and/or composition of the central linker have been found to be crucial for their function, in *Eh*CaBP, both flexibility as well as amino acid composition is required for the function of the protein.

Entamoeba histolytica, a protozoan parasite, is the causative agent of amoebiasis and amoebic dysentry. It infects nearly 50 million people worldwide, resulting in about 40000−100000 deaths every year (*1*). Though ubiquitous in distribution, this parasite is more prevalent in tropical and subtropical regions. It can invade extraintestinal tissues such as liver and brain and result in the formation of abscesses which could be life threatening (*1*). Although the biology of the parasite has been studied extensively, the mechanism governing its pathogenesis remains unclear. *E. histolytica* generally does not invade, and it is not clear as to why only a few infected individuals get invasive disease. A number of studies suggest that calcium (Ca^{2+}) may be involved in the pathogenetic mechanisms of amoebiasis (*2*−*4*). Ravdin et al. (*2*) have observed a transient increase in Ca^{2+} levels in the target cells during the process of cell killing. In another study, Carbajal et al. (*4*) have reported the effect of cytosolic free Ca^{2+} on cell adhesion of *E. histolytica* trophozoites. More recently, Meza (*5*) reviewed on various possible signaling pathways in *E. histolytica* leading to pathogenesis.

To understand the mechanism by which Ca^{2+} effects virulence and also to elucidate the role of Ca^{2+}-binding proteins (hereafter abbreviated as CaBP) in host-parasite relationships, a gene encoding a novel CaBP has been isolated from *E. histolytica* (*Eh*CaBP) (*6*). A number of experiments on this protein suggests that it is involved in a novel signal transduction pathway involving cellular proliferation, distinct from the ubiquitous Ca^{2+} signal-transducing molecule calmodulin (CaM) (*7*, *8*). Further, analysis of different species of *Entamoeba* indicates that this protein is not present in nonpathogenic species, namely, *Entamoeba invadens* and *Entamoeba moshkovskii*, thus suggesting its possible role in the pathogenesis of amoebiasis (*8*).

*Eh*CaBP is a 14.7 kDa (134 amino acid residues) monomeric protein. An analysis of its primary sequence reveals the presence of four Ca^{2+}-binding sites (Scheme 1), similar to those observed in other EF-hand CaBPs (*9*, *10*). The specific sites labeled as X, Y, Z, −Y, -X, -Z in Scheme 1, refer to 1st, 3rd, 5th, 7th, 9th, and 12th positions, respectively,

† A.B. gratefully acknowledges support from DST.

* To whom the correspondence should be addressed. E-mail: chary@tifr.res.in. Fax: 0091-22-2152110.

‡ Department of Chemical Sciences.

§ School of Life Sciences.

10.1021/bi0114978 CCC: $20.00 © 2001 American Chemical Society
Published on Web 11/08/2001

Scheme 1

Protein	Ca^{2+}-binding loop position												
	Sites	X		Y		Z		-Y		-X		-Z	
		1	2	3	4	5	6	7	8	9	10	11	12
*Eh*CaBP	I	**D**	V	**N**	G	**D**	**G**	A	V	S	Y	E	**E**
	II	**D**	A	**D**	G	**N**	**G**	E	I	D	Q	N	**E**
	III	**D**	V	**D**	G	**D**	**G**	K	L	T	K	E	**E**
	IV	**D**	A	**N**	G	**D**	**G**	Y	I	T	L	E	**E**
Calmodulin	I	**D**	K	**D**	G	**D**	**G**	E	V	S	F	E	**E**
	II	**D**	K	**D**	G	**N**	**G**	T	I	T	T	K	**E**
	III	**D**	K	**D**	G	**N**	**G**	Y	I	S	A	A	**E**
	IV	**D**	I	**D**	G	**D**	**G**	E	V	N	Y	E	**E**
Troponin C	I	**D**	A	**D**	G	**G**	**G**	D	I	S	T	K	**E**
	II	**D**	E	**D**	G	**S**	**G**	T	I	D	F	E	**E**
	III	**D**	K	**N**	A	**N**	**G**	F	I	D	I	E	**E**
	IV	**D**	K	**N**	N	**N**	**G**	R	I	D	F	D	**E**

in the Ca^{2+}-binding loops of the protein, which coordinate to Ca^{2+} in a pentagonal bipyramidal geometry (*9*).

Extensive analyses of molecular databases including human genome show that *Eh*CaBP has a limited overall sequence similarity with known CaBPs. The maximum identity at the amino acid level is 29% with calmodulins. In general, CaM belongs to a highly conserved protein family showing more than 65% sequence identity among members, irrespective of the species. Similarity in the protein sequences of *Eh*CaBP and CaM is restricted mainly to EF-hand Ca^{2+}-binding loops (see Scheme 1) with a negligible similarity in the functional central linker region.

As a prelude to the characterization of structure−function relationship in *Eh*CaBP, we have initiated its NMR study in solution in the Ca^{2+}-bound form. Backbone and side-chain ^1H, ^{13}C, and ^{15}N NMR assignments and secondary structure characterization of this protein have been reported earlier (*11*−*13*). These studies reveal that, despite its low overall sequence homology, the structural topology of the protein resembles that of intracellular CaBPs such as CaM (*14*−*17*) and troponin C (TnC) (*18*−*21*), in terms of two globular domains (the N-terminal and the C-terminal) with a pair of EF-hands each and connected by a flexible linker region. Despite such marked structural similarities, these proteins have evolved to be functionally diverse (*7*, *8*). Thus, subtle and important differences in structural properties may exist between *Eh*CaBP and CaM/TnC, which may provide clues to its function.

Unlike CaM and TnC, which have been extensively studied using X-ray crystallography, *Eh*CaBP crystals yielded poor X-ray diffraction limiting the resolution to 3.5 Å and thereby eluding a good crystal structure (*22*). This has been attributed to the presence of highly flexible linker region in *Eh*CaBP, connecting the two individual globular domains. In this paper, we present the complete 3D structure of the *Eh*CaBP in solution as derived by NMR data. The structural morphology of the protein is compared with CaM and TnC in order to highlight the similarities and differences, which exist between these proteins.

MATERIALS AND METHODS

*Sample Preparation. Eh*CaBP was overexpressed in *Escherichia coli* BL21(DE3) strain containing a pET-3c expression system. Unlabeled, isotopically ^{13}C or ^{15}N labeled, doubly labeled (both ^{13}C and ^{15}N), and amino acid selectively unlabeled−^{13}C fractionally (10%) labeled *Eh*CaBP, were purified using the protocol described earlier (*11*, *12*, *23*). NMR experiments were performed at 35 °C on 0.6 mL of 3 mM protein samples in 30 mM $CaCl_2$ and 50 mM deuterated Tris buffer (pH 6.0), either in 99.9% ^2H_2O or in a mixed solvent of 90% H_2O and 10% $2H_2O$.

NMR Experiments. NMR experiments were carried out on a Varian Unity+ 600 MHz NMR spectrometer equipped with pulsed field gradient unit and triple resonance probe with actively shielded *Z*-gradients, operating at a ^1H frequency of 600.051 MHz.

The two-dimensional (2D) experiments with unlabeled *Eh*CaBP in ^2H_2O include two quantum-filtered correlation spectroscopy (2QF−COSY) (*24*), three quantum-filtered correlation spectroscopy (3QF−COSY) (*25*), clean total correlation spectroscopy (clean-TOCSY) (*26*) with a mixing time (τ_m) of 80 ms and NOESY (*27*) with a τ_m of 100 ms. Spectra in 90% H_2O + 10% ^2H_2O include WGRAF−NOESY (*28*) with a τ_m of 100 ms and clean-TOCSY with a τ_m of 100 ms.

Experiments with uniformly ^{15}N-labeled or ^{13}C-labeled *Eh*CaBP include sensitivity enhanced 2D (^{15}N-^1H) HSQC and 2D (^{13}C-^1H) HSQC (*29*) in both constant time and nonconstant time options, 3D TOCSY−(^{13}C/^{15}N-^1H)−HSQC (τ_m = 80 ms) (*30*), 3D NOESY−(^{13}C/^{15}N-^1H)−HSQC (τ_m = 100 ms) (*31*) and 3D HNHA (*32*). Further, using the doubly labeled *Eh*CaBP, the following 3D triple resonance experiments were recorded: HNCA (*33*), HN(CO)CA (*34*), HNCO (*35*), CBCANH (*36*), CBCA(CO)NH (*37*), and HN(CA)CO (*38*). Data transformation and processing were done on a Silicon Graphics workstation (R10000 based Indigo II Solid Impact Graphics) using the Felix 97 software (MSI). ^1H chemical shifts were calibrated relative to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at 308 K (0.0 ppm). ^{13}C chemical shifts were calibrated indirectly relative to DSS. ^{15}N chemical shifts were calibrated with respect to an external standard of $^{15}NH_4Cl$ (2.9 M in 1 M HCl).

Distance Constraints for Structure Calculations: (a) nOe-Derived Distance Constraints. Cross-peaks from 3D (^1H-^1H) NOESY−(^{15}N-^1H)−HSQC and 2D ^1H NOESY spectra were integrated for obtaining distance restraints. The calibration of cross-peaks were done using the CALIBA macro of DYANA (*39*) with the minimum distance set to 2.4 Å and the maximum distance set to 6.0 Å. For cross-peaks in the NOESY spectrum which could not be integrated accurately, distances were classified based on their intensities as 1.8−2.4 Å (strong), 1.8−3.5 Å (medium), 1.8−5.0 Å (weak), or 1.8−6.0 Å (very weak). Hydrogen bond constraints were added only for residues which were involved in either a α-helix or a β-sheet as characterized by chemical shift indices, $^3J(H^N-H^\alpha)$ values and deuterium exchange rates (*12*). A lower limit of 2.0 Å was used for H−O distance in all the hydrogen bonds. Stereospecific assignments of prochiral methyl groups of Val and Leu residues were obtained using selectively unlabeled−^{13}C fractionally labeled samples as described in ref *23*. Standard pseudo-atom

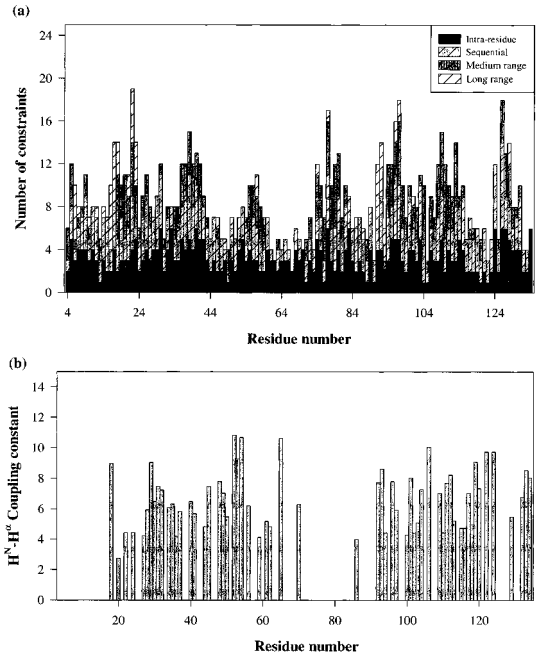

FIGURE 1: (a) Number of nOes for each amino acid residue used in 3D structure calculation of *Eh*CaBP, plotted as a function of the residue number (shown for residues 4−134). (b) Values of H^N−H^α coupling constants in *Eh*CaBP obtained using the 3D HNHA experiment, as a function of residue number. Vertical bars indicate values for only those residues for which coupling constants could be measured unambiguously.

distance corrections were incorporated to account for center averaging (*40*) for all methyl protons and nonstereospecifically assigned methylene protons. A total of 1265 distance constraints (with an average of about 10/residue) were used for structure calculations, which include 430 intraresidue, 389 interresidue (sequential), 335 medium range, and 111 long-range distance constraints. Only intraresidue and sequential nOes were observed for residues belonging to the linker region, while a few short-range nOes were seen for residues in segments joining the two EF-hands in a domain. The number of distance constraints used in the structure calculation for each amino acid residue is shown in Figure 1a. The N-terminal M1, A2, and E3 residues do not show up correlations in any of the 2D and 3D NMR spectra. These residues were therefore excluded from structure calculations.

(b) Ca^{2+}-Ligand Distance Constraints. Though Ca^{2+} ions were not included in the initial stages of structure calculations, they were incorporated in the final stages of structure refinement. These were based on two independent experimental observations of substantial downfield shifts of (i) ^{1}HN spins belonging to the individual Gly residues at the sixth position in each of the four Ca^{2+}-binding loops, namely, G15, G51, G90, and G122 (at 10.34, 10.75, 10.51, and 9.92 ppm, respectively), and (ii) backbone ^{15}N spins belonging to the individual hydrophobic residues at the eighth position in each of the four Ca^{2+}-binding loops, namely, V17, I53, L92, and I124 (at 127.8, 130.6, 128.7, and 128.1 ppm, respectively). The downfield shift observed for Gly residues is due to the hydrogen bonding of their HN with one of the two side-chain oxygens of an invariant Asp at the first position in the Ca^{2+}-binding loop (*41*), and those observed for residues at the eighth position in the loop are due to the coordination of backbone C′ of its previous residue to Ca^{2+} (*42*). Thus, these downfield shifts serve as signatures of Ca^{2+} binding to all the four sites in the protein.

The ligands chosen for Ca^{2+} coordination included residues at positions X, Y, Z, −Y, and −Z in the respective Ca^{2+}-binding loops (Scheme 1) (*9*). The bridging water molecule, hydrogen bonding with side-chain oxygen atoms of residues at Y and −X position and simultaneously coordinating to Ca^{2+}, was not included in the structure calculations. If the residue at position 1, 3, or 5 in the loop happened to be an Asp, then the distance between its Cγ and Ca^{2+} was restrained to an upper limit of 4.0 Å. This is because only one of the side-chain carboxyl oxygen coordinates to Ca^{2+}. On the other hand, if the residue at these positions happened to be an Asn, a direct upper limit O−Ca^{2+} distance constraint of 2.8 Å was used. Similarly, for residues at positions 7 and 12 in the loop, which coordinate to Ca^{2+} directly via their backbone C′ and both the side-chain carboxyl oxygens, respectively, an O−Ca^{2+} distance constraint of 2.8 Å was used. Thus, a total of six Ca^{2+}-ligand distance constraints per Ca^{2+}-binding loop were incorporated, with the lower bound set to 2.0 Å in all cases.

Torsion Angle Constraints. Torsion angle restraints for ϕ angles were derived from $^3J(H^N-H^\alpha)$ values measured in 3D HNHA experiment (*32*). Coupling constants thus obtained for different amino acid residues are given in Figure 1b. For residues which were characterized by ^{1}H, ^{13}Cα, and ^{13}C′ chemical shift indices (CSI) as belonging to a α-helix and showed a $^3J(H^N-H^\alpha) < 5$ Hz, ϕ was restrained to $-60 \pm 30°$ (α-helix). On the other hand, residues which were simultaneously characterized by the same CSI as belonging to β-sheet structure and which showed $^3J(H^N-H^\alpha) > 7$ Hz, ϕ was restrained to $-120 \pm 30°$ (β-sheet). Backbone ψ angle constraints were included only for those residues within a regular α-helical ($-45 \pm 30°$) or β-sheet ($135 \pm 30°$) conformation. A total of 200 ϕ and ψ torsion angle constraints were used in structure calculations with 176 constraints for residues in various α-helices and 24 constraints for residues in β-sheets.

Structure Calculations. Structure calculations were performed using the program DYANA (*39*). The standard simulated annealing protocol was used with 10 000 torsion angle dynamics (TAD) steps. Each round of structure calculations started with 200 randomized conformers. Out of all the energy-minimized calculated structures, 20 structures with the lowest residual target function values were chosen for analysis. All atom pairwise-RMSDs were computed using DYANA and MOLMOL (*43*). The quality of these structures was analyzed using PROCHECK (*44*). The corresponding PDB files for an ensemble of 20 structures (PDB code: 1JFJ) and a representative structure of *Eh*CaBP (PDB code: 1JFK) have been deposited.

Backbone Dynamics Studies. Backbone dynamics of *Eh*CaBP was studied using ^{15}N relaxation data (R_1, R_2, and nOe) obtained from 2D [^{15}N-^{1}H] NMR spectroscopy (*45*−*47*). The relaxation data were analyzed using the model free approach (*48*). Models of spectral density functions were selected for each residue using the procedure described by Mandel et al. (*49*), for both isotropic and axially symmetric diffusion tensors. They are essentially $S^2-\tau_m$, $S^2-\tau_m-\tau_e$, $S^2-\tau_m-R_{ex}$, $S^2-\tau_m-\tau_e-R_{ex}$, and a two-time scale model (*50*), where τ_m is the global correlation time of the molecule, S^2 represents the N−H bond order parameter, τ_e is the

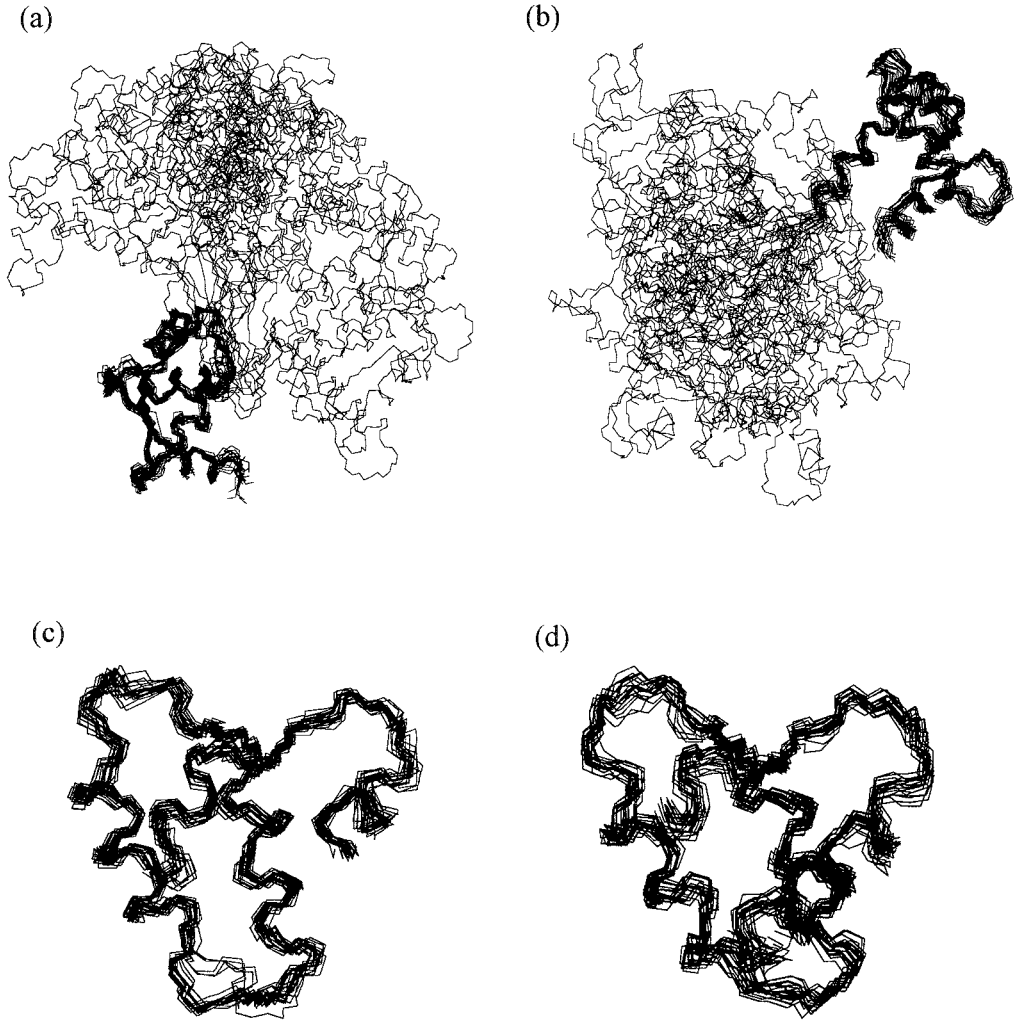

FIGURE 2: (a) Superposition of the 20 energy-minimized conformers of *Eh*CaBP as aligned with respect to the N-terminal domain (residues 4−62). Only backbone atoms are shown. (b) Superposition of the energy-minimized conformers of *Eh*CaBP aligned with respect to the C-terminal domain (residues 71−134). Only backbone atoms are shown. (c) Superposition of the N-terminal domain of the energy-minimized conformers of *Eh*CaBP (residues 4−62). (d) Superposition of the C-terminal domain of the energy-minimized conformers of *Eh*CaBP (residues 71−134).

effective correlation time for internal motions and R_{ex} refers to the conformational exchange rate constant. In the first stage, the best model for individual residue was selected by fitting the experimental data to different models separately. In such a selection, the criterion has always been a model that required a minimum number of parameters to fit the experimental data. After selection of the best model, τ_m was optimized along with other model parameters using the grid search method (*49*). The individual optimizations involved the minimization of χ^2 function using the program Modelfree (4.01) (*49*).

To ascertain whether relaxation data fitted using R_{ex} terms arise from rotational diffusion anisotropy or conformational exchange, a model incorporating anisotropic diffusion was tested for *Eh*CaBP in a fashion similar to the one described by Tjandra et al. (*51*). In such calculations, all residues which either have large amplitude internal motions or potentially undergoing conformational exchange, were excluded. The former was identified by their respective low values of nOe (<0.6) while the latter were identified using the following condition (*51*):

$$(\langle T_2\rangle - T_2,n)/\langle T_2\rangle - (\langle T_1\rangle - T_1,n)/\langle T_1\rangle > 1.5 \times \mathrm{SD}$$

where SD is the standard deviation of the left-hand side of the expression. The data were then fitted to two models of the diffusion tensor, namely, isotropic and axially symmetric. In such calculations, the following error function was optimized to obtain anisotropic diffusion parameters (*51*):

$$E = \sum_n (T_{1,e}/T_{2,e} - T_{1,c}/T_{2,c})^2/\sigma_{T_1/T_2}{}^2$$

where σ_{T_1/T_2} is the estimated error in the experimentally derived T_1/T_2 ratio, and the summation extends over all residues used in the fit. The program, R2R1_1.11 available in Prof. A. G. Palmer's laboratory was used for these calculations. Further, the significance of reduction in the error function was evaluated using a statistical *F*-test (*52*), to assess whether the fit occurred by chance with the inclusion of anisotropic diffusion.

RESULTS

Quality of EhCaBP Structures. A superposition of 20 NMR-derived minimum energy conformers of *Eh*CaBP is shown in Figure 2. Figure 2a shows the superposition of the structures when aligned with respect to the N-terminal

Table 1: Structural Parameter Statistics for the 20 Energy-Minimized Conformers of *Eh*CaBP Calculated Using DYANA[a]

(1) DYANA parameters	
avg target function (Å^2)	0.51 ± 0.08
upper limit violations	
constraints violated in more than 1 structure	0
sum of violations (Å)	4.7 ± 0.5
dihedral angle violations	
constraints violated in more than 1 structure	0
sum of violations (deg)	0.3 ± 0.1
van der Waals violations	
number > 0.2 Å	0
sum of violations (Å)	1.3 ± 0.3
(2) RMSD values (Å)	
N-terminal domain	
residues 4−62	0.68 ± 0.20
EF-hand 1 (residues 4−29)	0.54 ± 0.18
EF-hand 2 (residues 35−62)	0.59 ± 0.14
C-terminal domain	
residues 71−134	0.73 ± 0.19
EF-hand 3 (residues 71−102)	0.55 ± 0.18
EF-hand 4 (residues 107−134)	0.60 ± 0.20
(3) Ramachandran plot statistics	
residues in most favorable regions (%)	82.2
residues in allowed regions (%)	16.4
residues in generously allowed regions (%)	1.2
residues in disallowed regions (%)	0.2

[a] RMSDs were evaluated using MOLMOL and statistics for the Ramachandran plot were obtained using PROCHECK.

domain (residues 4−62), while Figure 2b shows the superposition of structures aligned with respect to the C-terminal domain (residues 71−134). Structural statistics for the family of 20 *Eh*CaBP structures is presented in Table 1. Although, each domain exhibits structural heterogeneity when aligned with respect to the latter (Figure 2a and 2b), residues within each domain converge with low RMSDs (Table 1). This is depicted in Figures 2c and 2d, which show the superposition of the 20 conformers of individual N- and C-terminal domains (residues 4−62 and 71−134). Further, as shown below using relaxation data, the central linker region (residues 63−70) exhibits high structural flexibility (in a subnanosecond time scale) giving rise to spatial heterogeneity of the two domains about this region. Poor convergence of structures in this region is thus attributed to a small number of nOes observed for these residues due to conformational averaging. A lack of secondary structure in this region is also reflected in $^1H^\alpha$, $^{13}C^\alpha$, and $^{13}C'$ chemical shift indices (*12*). On the other hand, it is seen that NMR structures within a given domain show good convergence and covalent geometry with a well-defined secondary structure. In all the 20 structures, the absolute mean deviation of χ^1 angles is 20°, while the absolute deviation of ω angles from the mean remained below 0.2°. Further, individual EF-hands within a given domain show higher convergence when compared to the domain as a whole, as indicated by their low RMSD from the mean (Table 1). This is attributed to be due to segments connecting the two individual EF-hands within each domain (residues 30−34 in N-terminal domain and 103−107 in C-terminal domain), which show slightly higher disorder than the rest of the polypeptide. These inter-EF-hand loop regions also undergo higher mobility as compared to the other structurally ordered segments in the protein (see below). Even in the Ramachandran plot of 20 structures (statistics shown in Table 1), residues that were observed in the disallowed regions (0.2% in total) belonged mostly to either the linker region or to the segments connecting the two EF-hands in individual domains. In very few cases (0.05%), residues belonging to the Ca^{2+}-binding loops were seen in the disallowed regions.

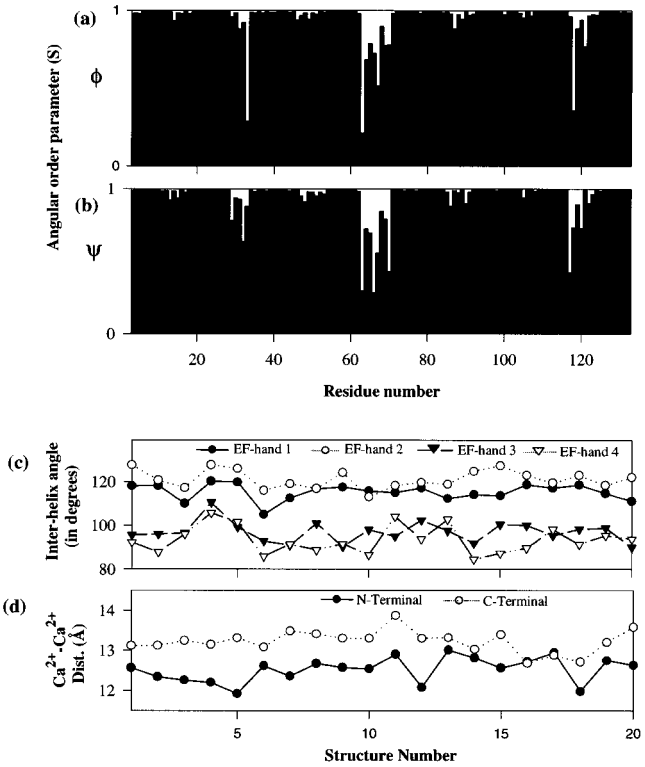

FIGURE 3: (a) Backbone ϕ and (b) ψ torsion angular order-parameters for the 20 structures of *Eh*CaBP as a function of residue number. (b) Interhelical angle within an EF-hand in the N- and C-terminal domains of *Eh*CaBP plotted for the energy-minimized structures. (c) $Ca^{2+}-Ca^{2+}$ distance in the N- and C-terminal domain of *Eh*CaBP plotted for the energy-minimized structures.

The family of *Eh*CaBP conformers was further analyzed with respect to angular order parameters for ϕ and ψ torsion angles using the method of Hyberts et al. (*53*). This is plotted in Figure 3a and 3b, as a function of residue number. The order parameter is related to the standard deviation of torsion angles in a logarithmic manner and a value close to 1.0 indicates a good convergence (*53*). As evident from Figure 3, *Eh*CaBP structures exhibit good convergence (except for residues 63−70) in terms of backbone torsion angles with mean values of 0.95 ± 0.12 and 0.96 ± 0.13 for ϕ and ψ torsion angular order parameters, respectively. Residues 63−70, which form an extended linker region, show poor convergence. Other residues, which show low angular order parameters, are the ones belonging to inter-EF-hand loop regions and a few residues in the Ca^{2+}-binding loops.

The interhelix angle between two helices within individual EF-hands was computed using MOLMOL and is shown in Figure 3c for all the individual structures. As evident from this figure, the interhelical angles are conserved in all the structures with an average (over 20 conformers) of 115 ± 3.75° for EF-hand I, 120 ± 4.2° for EF-hand II, 96 ± 4.8° for EF-hand III, and 93 ± 6.0° for EF-hand IV. These analyses reveal that, in its Ca^{2+}-bound state, the C-terminal domain of *Eh*CaBP shows a more open conformation compared to the N-terminal domain, as evidenced from the interhelical angles. Further, to see the effect of this open

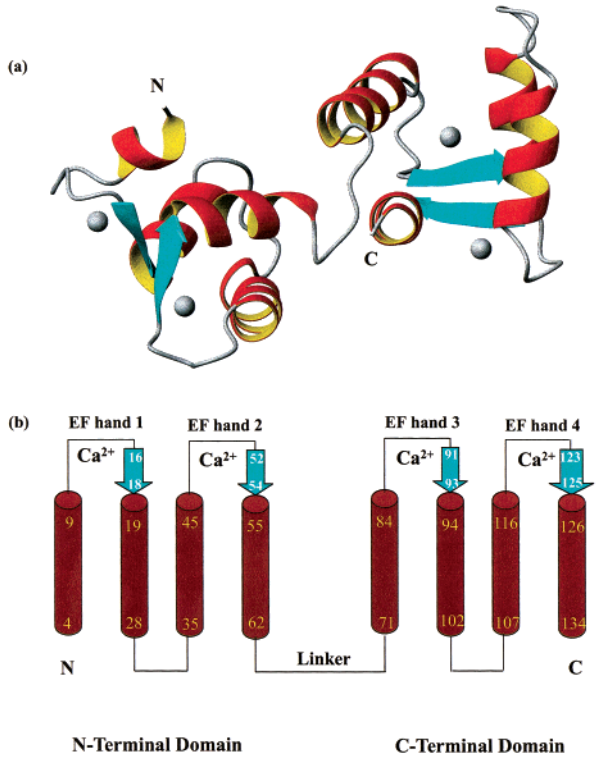

FIGURE 4: (a) A representative structure of *Eh*CaBP having the least residual target function value among the 20 conformers generated using DYANA. Residues 4−134 are shown in the figure. Ca^{2+} atoms in the structure are shown in gray. (b) A topology diagram of the same molecule (panel a), depicting the various secondary structure elements in the protein. Start and end residue numbers in α-helices (brown) and *β*-sheet (blue) structures along the polypeptide chain of the protein are indicated, along with the numbering of EF-hands.

structure on Ca^{2+}−Ca^{2+} distances in both the domains, inter-Ca^{2+} distance in each domain was computed for all the 20 structures and shown in Figure 3d. The Ca^{2+}−Ca^{2+} distance is conserved in all these structures with an average value of 12.6 ± 0.3 Å in N-terminal domain and 13.2 ± 0.28 Å in the C-terminal domain. A longer Ca^{2+}−Ca^{2+} distance in the C-terminal domain as compared to the N-terminal domain can be attributed to a more open structure of the former.

Description of the Structure. A minimum energy structure of *Eh*CaBP, having the least residual target function value among the 20 conformers was chosen as a representative structure and is shown in Figure 4a. A topology diagram of the same molecule is shown in Figure 4b, depicting the various secondary structure elements in the protein. The structure of *Eh*CaBP reveals two globular domains (N-terminal and C-terminal), each containing a pair of canonical EF-hands. The two domains are connected by an eight amino acid residue linker, which lacks a secondary structure.

(a) EF-Hands. Each of the four EF-hands in *Eh*CaBP consists of two amphipathic helices flanking the Ca^{2+}-binding loop in a typical helix-loop-helix motif (*9*, *10*). There are in all, eight α-helices [residues A4−I9 (α1), Y19−K28 (α2), E35−A45 (α3), Q55−Y62 (α4), D71−M84 (α5), E94−K102 (α6), E107−A116 (α7), L126−L134 (α8)] and four antiparallel *β*-strand segments, each formed by three amino acid residues [V16−S18 (*β*1), E52−D54 (*β*2), K91−T93 (*β*3), and Y123−T125 (*β*4)]. Helices α1 and α2 form the first EF-hand in the N-terminal domain followed by helices α3 and α4 forming the second EF-hand. In the C-terminal domain, the two respective EF-hands are formed by helices α5 and α6 and α7 and α8 (Figure 4b). All α-helices and *β*-strands were characterized by characteristic nOe patterns, $^3J(H^N-H^\alpha)$ values and deuterium exchange studies (*12*).

(b) Ca^{2+}-Binding Loops. The four Ca^{2+}-binding loops in *Eh*CaBP, as evidenced from sequence homology (Scheme 1) and supported by NMR data consist of 12 contiguous residues with an invariant Asp, Gly, and Glu at the 1st, 6th, and the 12th positions, respectively. Residues D10−E21 and D46−E57 form Ca^{2+}-binding loops in the N-terminal domain, while those belonging to the C-terminal domain are formed by residues D85−E96 and D117−E128.

(c) Hydrophobic Pockets in the Protein. Figure 5a illustrates the hydrophobic surface area of the N- and C-terminal domains (residues 4−62 and 71−134, respectively) in *Eh*CaBP. The hydrophobic residues shown (in red) are Ala, Val, Leu, Ile, Met, and Phe. The solvent (water) exposed surface area was computed using the software, NACCESS (*54*). Using this software, relative solvent exposed surface area of each residue is calculated as percentage accessibility compared to the accessibility of that residue type in an extended ALA-X-ALA tripeptide (*55*). Relative accessibility helps in evaluating the extent of accessible surface area of a given residue type in different regions of the protein. Thus, the C-terminal domain, which shows a more open conformation in *Eh*CaBP exhibits a slightly higher exposed hydrophobic surface of 37.3% (total solvent accessible area = 1483 Å) compared to 33% (total solvent accessible area = 1264 Å^2) in the N-terminal domain. The number of hydrophobic residues in N-terminal and C-terminal domains is 26 and 24, respectively.

Hydrophobic pockets in the protein were characterized based on nOes observed between apolar residues. Residues, which formed the hydrophobic cluster with a relative solvent accessible area of less than 10%, are shown in Figure 5b for the two domains. The core of the hydrophobic pocket arises mainly from residues in the amphipathic helices of the EF-hands. Phe residues at 6, 25, 42, 58, and 61 comprise the buried aromatic residues in the N-terminal domain, while those in the C-terminal domain are comprised of Phe residues at 101, 129, and Tyr 81 (Figure 7b). The other Phe residues in the C-terminal domain are relatively exposed to solvent with a relative solvent accessible surface area of 32% (F100) and 58% (F130). Other hydrophobic residues which have more than 30% of their surface area accessible to solvent include A2, L5, V11, A16, A24, A31, I32, L37, L40, A47, and A59 in the N-terminal domain and L77, V79, L80, L83, M84, V86, I106, and A110 in the C-terminal domain. There are few polar residues, mainly belonging to the Ca^{2+}-binding loops, that are also buried away from solvent.

Dynamic Properties of EhCaBP. The different flexible and rigid structural elements in the protein have been analyzed using the backbone ^{15}N relaxation data, following the procedure outlined in the Materials and Methods. Order parameters (S^2) reflect the mobility of individual N−H bond vectors on the nanosecond to picosecond time scales (*48*). An estimate for τ_c of 5.95 ns was obtained from R_2/R_1 ratio using an isotropic diffusion model. The average S^2 value for residues 10−62 (belonging to the N-terminal domain) is found to be 0.89 ± 0.07, while it is 0.88 ± 0.05 for residues 71−134 (belonging to the C-terminal domain). Of 112

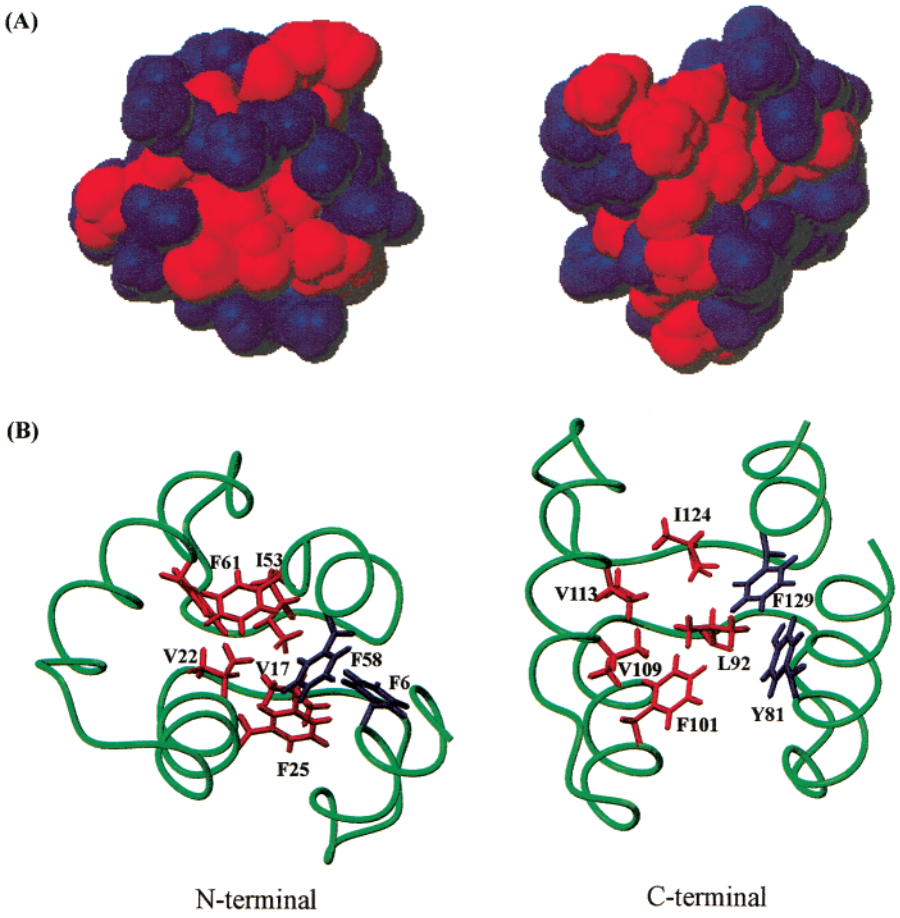

FIGURE 5: (a) Exposed hydrophobic surface of N- and C-terminal domain of *Eh*CaBP. The surface area of hydrophobic residues (Ala, Val, Leu, Ile, Phe, and Met) is shown in red, while that of nonhydrophobic residues are shown in blue for each domain. (b) Side chains of hydrophobic residues (shown in red and blue) in the N- and C-terminal domain of *Eh*CaBP, having relative solvent accessible area of less than 10% (see text). The backbone polypeptide chain of the protein is shown in green.

amino acid residues for which relaxation data were unambiguous, 60 could be fitted to $S^2-\tau_m$, 15 to $S^2-\tau_m-\tau_e$, 7, to $S^2-\tau_m-R_{ex}$, and 4 to $S^2-\tau_m-\tau_e-R_{ex}$ model, assuming isotropic diffusion.

To verify whether the relaxation data fitted using R_{ex} terms could arise from rotational diffusion anisotropy as against conformational exchange, a model incorporating axially symmetric diffusion tensor was tested for *Eh*CaBP. Ninety residues that satisfied the conditions described in Materials and Methods were included in the analyses. The relaxation data (R_2/R_1 ratio) were found to fit well with an axially symmetric model over the isotropic model of diffusion, as evidenced by a slight reduction in the error function from 1.80 to 1.44. Values obtained for $D_\parallel/D_\perp$, θ, and ϕ were 0.90 (oblate), 10°, and −22°, respectively. As mentioned earlier, reduction in error function was tested using the statistical *F*-test (*52*). Such a test ruled out the possibility of reduction in the error function by chance, with the inclusion of additional parameters, as the calculated probability turned out to be 4.9×10^{-4} ($F = 6.54$), thereby justifying the anisotropic diffusion model for *Eh*CaBP. Order parameters and τ_m were optimized in the same manner as for the isotropic model. However, no significant differences in τ_m (5.97 ns) and average S^2 were noticed in the results obtained from the two models. Of the seven residues, which required an inclusion of the R_{ex} term in the isotropic model, two did not require R_{ex} and fitted well in the $S^2-\tau_e$ model under axially symmetric diffusion. These results imply that, in the case of *Eh*CaBP, although anisotropy is required for selection of an appropriate model, its inclusion is not necessary for evaluating the backbone order parameters.

The various solvent-exposed and buried regions in *Eh*CaBP have been analyzed using deuterium exchange studies. Figure 6 shows a plot of (a) the local backbone RMSD, calculated as an average from the mean structure, (b) the order parameter (S^2) values (under isotropic diffusion model), (c) calculated solvent accessible areas of each residue in the protein, and (d) the deuterium exchange rates. In the case of deuterium exchange rates, only those residues for which exchange rates could be measured (i.e., which exchanged in medium to slow time scales) are plotted. The arrow marks indicate a complete stretch of residues that underwent fast deuterium exchange and, hence, the exchange rates could not be measured. As is evident from the dynamics data, residues in the four Ca^{2+}-binding loops, the loops connecting the EF-hands and the linker region exhibit a high degree of mobility in the increasing order. Residues 67−70, belonging to the central linker, exhibit maximum disorder in the protein (S^2 values of 0.7 ± 0.06). Only a small number of nOes were observed for the residues in this region due to significant line broadening, suggesting that the amino acid residues in this region may experience motions in microsecond to millisecond time scales. This lack of sufficient nOes is reflected in high local backbone RMSDs for residues

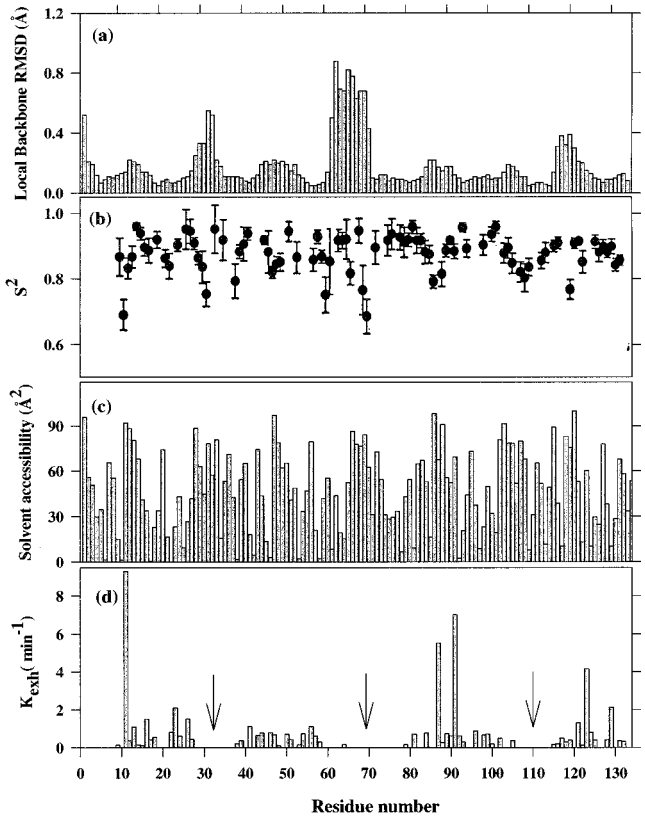

FIGURE 6: (a) Backbone local RMSD from the mean for the 20 energy-minimized conformers of *Eh*CaBP plotted as a function of residue number. (b) Order-parameters (S^2) computed using the model free approach from backbone ^{15}N T_1, T_2 and ^{15}N-^1H heteronuclear nOe data, plotted as a function of residue number. (c) Relative solvent accessible surface area of amino acid residues in *Eh*CaBP plotted as a function of residue number. These area were calculated using the program NACCESS (see text). The structure shown in Figure 4a was used in the calculation of solvent accessible areas. (d) Deuterium exchange rates in *Eh*CaBP plotted as a function of residue number. Vertical bars indicate values for only those residues for which exchange rates could be measured. The arrow marks indicate a stretch of amino acid residues, for which deuterium exchange rates could not be measured due to fast exchange. These residues belong to the central linker region and loop regions connecting two EF-hands in the N- and C-terminal domain of *Eh*CaBP.

in this region (Figure 6a). Amino acid residues in the linker region are also exposed to the solvent, leading to high deuterium exchange rates (Figure 6d). Apart from the N- and C-terminal residues and the linker region, other regions in the protein that have slightly higher mobility are loops connecting the two EF-hands in both the domains (residues 29−33 and residues 103−106) and few residues comprising the Ca^{2+}-binding loops, which also explains high local backbone RMSDs for residues in these segments.

DISCUSSION

The solution structure of *Eh*CaBP reveals a CaM paradigm of two independent globular domains connected by a flexible linker. It consists of four canonical EF-hands, a pair each in the N- and C-terminal domains. The Ca^{2+}-binding loops in the protein consist of 12 contiguous residues flanked by two helices which are oriented nearly perpendicular to each other, a feature common with other members of the family of EF-hand proteins (*9*, *10*). *Eh*CaBP is structurally related, most closely to CaM (*14*−*17*) and TnC (*18*−*21*) inspite of low

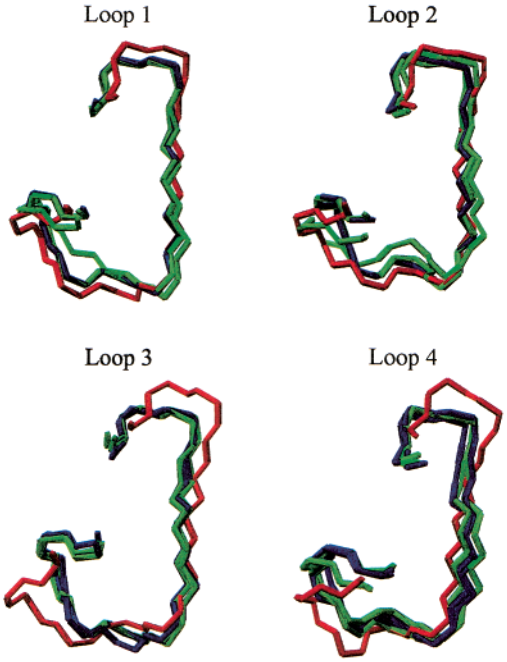

FIGURE 7: Superposition of four Ca^{2+}-binding loops of *Eh*CaBP (shown in red) with the corresponding loops of CaM (blue) and TnC (dark green). Both X-ray derived and NMR-derived structures of CaM and TnC are used in the comparison (see text). The loops are numbered in the order from N-terminus to the C-terminus of the protein. Thus, loops 1 and 2 belong to the N-terminal domain, while loops 3 and 4 belong to the C-terminal domain.

sequence homology with these proteins. Further, *Eh*CaBP has been thought to be involved in a novel signal transduction pathway, distinct from CaM (*7*, *8*). It is of therefore of interest, from the point of view of biological function, to compare the various structural elements/properties in *Eh*-CaBP with the corresponding segments/properties in CaM and TnC.

Comparison of EhCaBP Structure with CaM and TnC. The following 3D structures of CaM and TnC were used for comparison with that of *Eh*CaBP: NMR solution structure of C-terminal CaM (PDB code 1CMG) (*16*), X-ray structure of CaM (PDB code 1CLL) (*17*), NMR structure of TnC (PDB code 1TNX) (*18*) and X-ray structure of TnC (PDB code 1TN4) (*21*). All structures used in such comparison represented the protein in a Ca^{2+} saturated (bound) form and not complexed with any other molecule. The comparison is mostly focused on the conformation of Ca^{2+}-binding loops, interhelical angles within a specific EF-hand, hydrophobic surface area and the interdomain linker region. In the case of NMR derived structures of CaM, a mean structure calculated from the ensemble of deposited structures was used in the analysis.

(a) Ca^{2+}-Binding Loops. Figure 7 shows a superimposition of the four individual Ca^{2+}-binding loops, in the order of primary sequence from N to C-terminus of *Eh*CaBP, with the corresponding loops from CaM and TnC. A backbone RMSD (from the mean) of 0.50, 0.70, 0.73, and 0.75, respectively, were obtained for the four loops. Clearly, the Ca^{2+}-binding loops in the C-terminal domain of *Eh*CaBP do not converge with those of CaM and TnC as compared to the loops in N-terminal domain. The third Ca^{2+}-binding site in particular, shows a maximum divergence in the loop structure. The more open structure of this loop cannot be

Table 2: Primary Sequence Indicating the EF-Hand Ca^{2+}-Binding Loops in Each Domain of *Eh*CaBP, CaM, and TnC[a]

Protein	Domain	Phe-Phe interaction sequence (-4 / 1 ... 13)	Central linker region
*Eh*CaBP	N	^{6}F K E I D V N G D G A V S Y E E . . / . . . D A D G N G E I D Q N E ^{58}F	^{63}G S I Q G Q D L^{70}
	C	^{81}Y K L M D V D G D G K L T K E E . . / . . . D A N G D G Y I T L E E ^{129}F	
Calmodulin	N	^{16}F S L F D K D G D G T I T T K E . . / . . . D A D G N G T I D F P E ^{68}F	^{73}A R K M K D T D^{80}
	C	^{89}F R V F D K D G N G Y I S A A E . . / . . . D I D G D G Q V N Y E E ^{140}F	
Troponin C	N	^{26}F D M F D A D G G G D I S T K E . . / . . . D E D G S G T I D F E E ^{78}F	^{85}E D A K G K S E E E^{94}
	C	^{102}F R I F D K N A D G F I D I E E . . / . . S D K N N D G R I D F D E ^{154}F	

[a] Residues at the −4 position of the first Ca^{2+}-binding loop in a given EF-hand and 13th position of the second Ca^{2+}-binding loop, within the same domain, are highlighted. Their respective positions in the protein primary sequence are indicated as superscripts. The numbering of the Ca^{2+}-binding loops, which are shown underlined, begins with the first residue, an invariant Asp. Last column in the table shows the central linker region in the three proteins, with the respective positions of the start and end residue in the primary sequence indicated by superscripts.

attributed to a more open C-terminal domain (as evidenced from inter-helical angles). This is because, Ca^{2+}-binding loops in the N-terminal domain, which adopt a more closed conformation do not diverge from the corresponding loops of CaM and TnC.

The divergence in the third Ca^{2+}-binding loop of *Eh*CaBP is primarily attributed to the presence of a Tyr residue in −4 position (residue 81) instead of a conserved Phe (*56*, *57*) (Table 2). Here, position 1 refers to the start of a canonical EF-hand Ca^{2+}-binding loop. A recent study involving EF-hand proteins has revealed that the highly conserved Phe residue at position −4 of the first EF-hand in a domain interacts with another conserved Phe at the 13th position in the second EF-hand loop within the same domain (*58*). These two Phe residues adopt a perpendicular orientation with respect to each other, bringing their $C^{\alpha}-C^{\alpha}$ distance to 5−8 Å. In both CaM and TnC, these Phe residues are conserved (Table 2). In *Eh*CaBP, the two Phe (in −4th and 13th position) in the N-terminal domain are F6 and F58 (Figure 5b; residues shown in blue), which adopt a perpendicular orientation with a $C^{\alpha}-C^{\alpha}$ distance of 7.3 Å. However, in the C-terminal domain, the aromatic ring of Y81 faces that of its partner, F129, in a nearly parallel orientation (Figure 5b; residues shown in blue) resulting in an average distance $C^{\alpha}-C^{\alpha}$ distance of 10.4 Å. This could be the reason as to why the Ca^{2+}-binding loops in the C-terminal domain of *Eh*CaBP adopt a more open conformation compared to the ones in the N-terminal domain and also diverge from the corresponding loops in CaM and TnC (Figure 7). This hypothesis was further verified by superimposing 27 Ca^{2+}-binding loops from other proteins possessing the canonical EF-hand loops such as parvalbumin, oncomodulin, recoverin, and aequorin along with CaM, TnC, and *Eh*CaBP (Figure 1 in the Supporting Information). It is found that other than *Eh*CaBP, aequorin is the only other protein that possesses a Met residue instead of the highly conserved Phe at the 13th position in its fourth EF-hand. Ca^{2+}-binding loops (3 and 4) of this protein also diverge the most from those of other proteins, akin to the one observed in *Eh*CaBP. The corresponding $C^{\alpha}-C^{\alpha}$ distance between residues at position −4 (Phe) in the third EF-hand and 13 (Met) in the fourth EF-hand of aequorin is found to be 10.5 Å, a value similar to the one in *Eh*CaBP.

(b) Interhelical Angles. Table 3 shows the interhelical angle within an EF-hand for *Eh*CaBP, CaM, and TnC. These angles were calculated using MOLMOL. As evident from interhelical angles, *Eh*CaBP shows a more open C-terminal domain compared to CaM and TnC, while on the other hand, the N-terminal domain of *Eh*CaBP is seen to adopt a slightly more closed state. In both CaM and TnC, the N-terminal domain is the regulatory domain, which interacts with the target molecules while the C-terminal domain acts as an anchoring domain (*59*, *60*). Though such characterization has not been done in the case of *Eh*CaBP, a more open nature of C-terminal domain compared to the N-terminal domain suggests its possible involvement in binding target molecules.

(c) Exposed Hydrophobic Surface. Table 4 shows a percentage of the total surface of hydrophobic residues that is exposed to the solvent (water) for *Eh*CaBP, CaM, and TnC, calculated using the program NACCESS. The percentage values in the table indicate sum total of relative solvent accessibilities of all the hydrophobic residues in a given

Table 3: Interhelical Angles within an EF-Hand in *Eh*CaBP, CaM, and TnC[a]

protein	N-terminal domian		C-terminal domain	
	A/B	C/D	E/F	G/H
*Eh*CaBP	117.8	118.1	92.7	93.4
calmodulin (X-ray)	85.5	87.4	102.4	95.2
calmodulin (NMR)	*b*	*b*	106.6	87.7
troponin C (X-ray)	99.5	96.9	104.6	107.2
troponin C (NMR)	82.5	79.2	91.0	104.8

[a] In the case of CaM and TnC, both X-ray and NMR-derived structures are used for comparison (see text). The helices are labeled A−H as they occur in the primary sequence from N- to C-terminus of the protein. Thus, helices A/B represent the first EF-hand pair in the protein, C/D the second EF-hand pair and so on. Helices A−D belong to the N-terminal domain and helices E−H belong to the C-terminal domain of each protein. [b] NMR structure corresponding to Ca^{2+}-saturated N-terminal CaM (noncomplexed form) was not found in the protein data bank.

Table 4: Total Relative Solvent Accessible Hydrophobic Surface Area of Each Domain in *Eh*CaBP, CaM, and TnC Calculated Using the Program NACCESS (*63*) (see text)[a]

protein	relative solvent accessible surface area (%)			
	N-terminal domain		C-terminal domain	
	no. of hydrophobic residues	surface area	no. of hydrophobic residues	surface area
*Eh*CaBP	26	33.3	24	37.3
calmodulin (X-ray)	25	28.2	24	29.3
calmodulin (NMR)	*b*	*b*	24	23.7
troponin C (X-ray)	33	25.3	24	26.7
troponin C (NMR)	33	33.3	24	34.6

[a] In the case of CaM and TnC, both X-ray and NMR-derived structures were used for comparison. The number of hydrophobic residues used in the calculation for each domain is indicated. The percentage values indicate the extent of total hydrophobic surface area exposed in each protein. [b] NMR structure corresponding to Ca^{2+}-saturated N-terminal CaM (noncomplexed form) was not found in the protein data bank.

domain (the number of hydrophobic residues are indicated in brackets). Residues used in the analysis were Ala, Val, Leu, Ile, Met, and Phe. As evident from the table, the hydrophobic residues in CaM and TnC expose their total surface to solvent to a lesser extent as compared to *Eh*CaBP. Such an observation was also made in thermodynamics studies of binding of peptides to *Eh*CaBP (*61*). In a recent peptide-binding study, it has been reported that even after binding of mellitin, *Eh*CaBP exposes a considerable amount of hydrophobic surface as compared to CaM−mellitin complex (*61*). This is based on the assumption that mellitin interacts in a similar fashion with CaM and *Eh*CaBP. This exposure of hydrophobic surface may be required for other secondary interactions of the protein and also imparts specificity to the protein toward such interactions.

(d) Linker Region. The central linker region, connecting the two globular domains in CaM and TnC has been a region of special interest for biochemists and structural biologists alike. Several site-directed mutageneses have been carried out in these proteins to understand the effect of length and composition of this linker on their individual functions (*62*−*64*). This linker region, which spans 8−10 residues in both CaM and TnC, has been characterized as part of a long helix in X-ray studies (*17*, *20*, *21*) and as an extended flexible region in the solution-state NMR studies (*14*−*16*, *18*, *19*, *65*, *66*). As has been described earlier, in *Eh*CaBP, the linker region spans eight amino acid residues (residues 63−70) (Figure 4). This region exhibits a large structural flexibility (Figure 6b), leading to an ill-defined orientation of the two domains (N- and C-terminal) with respect to each other (Figure 2, panels a and b). The amino acid composition of the linker region in *Eh*CaBP, CaM, and TnC is shown in Table 2 (last column) for comparison. The start and end residues in the linker region are indicated as superscripts. CaM and TnC are composed of a number of charged residues in the central linker region as against only one Asp in *Eh*CaBP. It is interesting to note that the presence of two Gly residues (G63 and G67) may seem to impart a greater flexibility to the central linker region in *Eh*CaBP, as compared to the lone Gly residue present in TnC and none in CaM. Such flexibility has been previously attributed to be the reason as to why X-ray diffraction patterns of *Eh*CaBP crystals are found to be poor in quality (*22*). Further, the presence of G63 and G67 in the linker region has been shown to be important for the function of *Eh*CaBP. In one of the earlier studies on *Eh*CaBP, these two Gly residues were mutated to Ala (*67*). Such mutations were found to impart rigidity to the flexible linker region. Studies involving CD and mass spectrometric techniques revealed that the mutant protein is weakened in its ability to bind Ca^{2+} and its target peptides (*67*), implying a possible role of these glycines in specific interactions with target peptides. In another study, the flexibility in the central linker region was thought to be responsible for the inability of the short peptide, mastoporan, to bind to *Eh*CaBP, although the same peptide is found to bind to CaM with considerable affinity (*61*). Thus, in contrast to CaM and TnC, where length and/or composition of the linker region were found to be crucial for the function, in *Eh*CaBP, both flexibility as well as amino acid composition (with the two Gly residues at position 63 and 67) seem to be necessary for its function.

Structure−Function Relationship in EhCaBP. The NMR-derived solution structure of *Eh*CaBP presented here provides clues on the putative function of the protein. Although *Eh*CaBP, CaM, and TnC are similar in terms of Ca^{2+} binding affinities and the hydrophobic nature of protein−target interaction, there are dissimilarities in the nature of the central linker region and the extent of the exposed hydrophobic residues. The latter have been shown to be crucial for the function of the protein both in the case of *Eh*CaBP and CaM/TnC. The flexibility in the linker region, imparted by Gly residues, may be responsible for specificity of the protein toward its interaction with target molecules. Specificity may also arise from more exposed hydrophobic residues as compared to CaM and TnC. In the case of *Eh*CaBP, a presence of Tyr (Y81) in the C-terminal domain instead of a more conserved Phe in that position seems to fine-tune the Ca^{2+} binding affinity in that domain. This is revealed in a comparison of Ca^{2+}-binding loop structures of this domain with that of other EF-hand proteins. Thus, we conclude that *Eh*CaBP operates in a distinct manner compared to CaM and TnC, an observation that is supported by biochemical studies (*7*, *8*). Site-directed mutagenesis involving the replacement of Y81 by a more conserved Phe (Y81F), along with structural studies of apo-*Eh*CaBP and *Eh*CaBP−peptide

complexes in solution by NMR, are currently in progress in our laboratory.

ACKNOWLEDGMENT

The facilities provided by the National Facility for High Field NMR, supported by Department of Science and Technology (DST), Department of Biotechnology (DBT), Council of Scientific and Industrial Research (CSIR), and Tata Institute of Fundamental Research, Mumbai, are gratefully acknowledged. We dedicate this paper in the memory of late Prof. G. N. Ramachandran (1922-2001).

SUPPORTING INFORMATION AVAILABLE

Superimposition of canonical EF-hand Ca^{2+}-binding loops from different proteins. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Christopher, D. H., Haque, R., and Petri, W. A., Jr. (1999) *Exp. Rev. Mol. Rev. 22*.
2. Ravdin, J. I., Murphy, C. F., Guerrant, R. L., and Long-Kruz, S. A. (1985) *J. Infect. Dis. 152*, 542−549.
3. Munoz, M. L., O'Shea-Alvarez, M. S., Perez-Garcio, J., Weinbach, E. C., Moreno, M. A., and Tovar, R. (1992) *Comp. Biochem. Physiol. B103*, 517−521.
4. Carbajál, M. E., Manning-Cela, R., Piña, A., Franco, E., and Meza, I. (1996) *Exp. Parasitol. 82*, 11−20.
5. Meza, I. (2000) *Parasitol. Today 16*, 23−28.
6. Prasad, J., Bhattacharya, S., and Bhattacharya, A. (1992) *Mol. Biochem. Parasitol. 52*, 137−140.
7. Yadava, N., Chandok, M. R., Prasad, J., Bhattacharya, S., and Bhattacharya, A. (1997) *Mol. Biochem. Parasitol. 84*, 69−82.
8. Sahoo, N., Chakravarty, P., Yadava, N., Bhattacharya, S., and Bhattacharya, A. (2000) *Arch. Med. Res. 31*, S57−S59.
9. Strynadka, N. C., and James, M. N. G. (1989) *Annu. Rev. Biochem. 58*, 951−958.
10. Nelson, M. R., and Chazin, W. J. (1998) *BioMetals 11*, 297−318.
11. Sahu, S. C., Atreya, H. S., Chauhan, S., Bhattacharya, A., Chary, K. V. R., and Govil, G. (1999) *J. Biomol. NMR 14*, 93−94.
12. Sahu, S. C., Bhattacharya, A., Chary, K. V. R., and Govil, G. (1999) *FEBS Lett. 459*, 51−56.
13. Atreya, H. S., Sahu, S. C., Chary, K. V. R., and Govil, G. (2000) *J. Biomol. NMR 17* (2), 125−136.
14. Zhang, M., Tanaka, T., and Ikura, M. (1995) *Nat. Struct. Biol 2*, 758−767.
15. Kubinowa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C. B., and Bax, A. (1995) *Nat. Struct. Biol. 2*, 768−776.
16. Finn, B. E., Evenas, J., Drakenberg, T., Waltho, P. J., Thulin, E. and ForsÈn, S. (1995) *Nat. Struct. Biol. 2*, 777−783.
17. Chattopadhyaya, R., Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) *J. Mol. Biol. 228*, 1177.
18. Slupsky, C. M., and Sykes, B. D. (1995) *Biochemistry 34*, 15953−15964.
19. Gagné, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) *Nat. Struct. Biol. 2*, 784−789.
20. Herzberg, O., and James, M. N. G. (1988) *J. Mol. Biol. 203*, 751.
21. Houdusse, A., Love, M. L., Dominguez, R., Grabarek Z, and Cohen, C. (1997) *Structure 15*, 1695−711.
22. Gopal, B., Suma, R., Bhattacharya, S., Bhattacharya, A., Murthy, M. R. N., and Surolia, A. (1998) *Acta Crystallogr., Sect. D 54*, 1442−1445.
23. Atreya, H. S., and Chary, K. V. R. (2001) *J. Biomol. NMR 19*, 267−272.
24. Piantini, U., Sörenson, O. W., and Ernst, R. (1982) *J. Am. Chem. Soc. 104*, 6800−6801.
25. Muller, N., Ernst, R. R., and Wüthrich, K. (1996) *J. Am. Chem. Soc. 108*, 6482−6492.
26. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
27. Kumar, A., Wagner, G., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 96*, 1156−1163.
28. Sahu S. C., and Majumdar, A. (1998) *Curr. Sci. 74*, 451−456
29. Kay, L. E., Keifer, P., and Saarinan, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
30. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
31. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T, Bax, A., Gronenborn, A, and Clore, G. M. (1989) *Biochemistry 28*, 6150−6156.
32. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
33. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.
34. Bax, A., and Ikura, M. (1991) *J. Biomol. NMR 1*, 99−104.
35. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.
36. Wittekand, M., and Muller, L. (1993) *J. Magn. Reson. B101*, 201−205.
37. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
38. Clubb, R. T., Thanabal, V., and Wagner, G. (1992) *J. Magn. Reson. 97*, 213−217.
39. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
40. Wüthrich, K., Billeter, M., and Braun W. (1983) *J. Mol. Biol. 169*, 949.
41. Akerfeldt, K. S., Coyne, A. N., Wilk, R. R., Thulin, E., and Linse, S. (1996) *Biochemistry 35*, 3662−3669.
42. Biekofsky, R. R., Martin, S. R., Browne, J. P., Bayley, P. M., and Feeney, J. (1998) *Biochemistry 37*, 7617−7629.
43. Koradi, R., Billeter, M., and Wüthrich K. (1996) *J. Mol. Graph. 14*, 51−55.
44. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thorton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
45. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry 28*, 8972−8979.
46. Skelton, N. J., Palmer, A. G., Akke, M., Kordel, J., Rance, M., and Chazin, W. J. (1993) *J. Magn. Reson. B102*, 253−264.
47. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
48. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559 and 4559−4570.
49. Mandel, A. M., Akke, M., and Palmer, A. G. (1995) *J. Mol. Biol. 246*, 144−163.
50. Clore, G. M, Szabo, A. Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) *J. Am. Chem. Soc. 112*, 4989−4991.
51. Tjandra, N., Feller, S. E., Pastor, R. W., and Bax, A. (1995) *J. Am. Chem. Soc. 117*, 12562−12566.
52. Bevington, P. R., and Robinson, D. K. (1992) in *Data reduction and error analysis for the physical sciences*, McGraw-Hill, New York.
53. Hyberts, S. G., Goldberg, M. S., Havel, T. F., and Wagner, G. (1992) *Science 1*, 736−751.
54. Hubbard, S. J., and Thornton, J. M. (1993) *NACCESS*, Computer Program, Department of Biochemistry and Molecular Biology, University College London.
55. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.
56. Sekharadu, Y. C., and Sundaralingam, M. (1988) *Protein Eng. 2*, 139−146.
57. Falke, J. J., Drake, S. K., Hazard, A. L., and Peersen, O. B. (1994) *Q. Rev. Biophys. 27*, 219−290.

58. Rashidi, H. H, Bauer, M., Patterson, J., and Smith, D. W. (1999) *J. Mol. Microbiol. Biotechnol. 1*, 175−182.
59. Crivici, A., and Ikura, M. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 85−116.
60. Filatov, V. L., Katrukha, A. G., Bulargina, T. V., and Gusev, N. B. (1999) *Biochemistry (Moscow) 64*, 969.
61. Moorthy, A., K., Gopal, B., Satish, P. R., Bhattacharya, S., Bhattacharya, A., Murthy, M. R. N., and Surolia, A. (1999) *FEBS Lett. 461*, 19−24.
62. Putkey, J. A., Ono, T., VanBerkum, M. F., and Means, A. R. (1988) *J. Biol. Chem. 263*, 11242−11249.
63. Fujimori, K., Sörenson, M., Herzberg, O., Moult, J., and Reinach, F. C. (1990) *Nature 345*, 182−184.
64. Dobrowlski, Z., Xu, G. Q., and Hitchcock-DeGregori, S. E. (1991) *J. Biol. Chem. 266*, 5703−5710.
65. Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) *Biochemistry 31*, 5269−5278.
66. Slupsky, C. M., Kay, C. M., Reinach, F. C., Smillie, L. B., and Sykes, B. D. (1995) *Biochemistry 34*, 7365−7375.
67. Moorthy, A. L. (2001) Ph.D. Thesis, Indian Institute of Science, Bangalore, India.

BI0114978