Proc Indian Natn Sci Acad, **70, A,** No. 4, July 2004, pp. 615–625
© Printed in India.

NMR OF EF-HAND CALCIUM BINDING PROTEINS

SULAKSHANA MUKHERJEE, SOURAJIT MUSTAFI, AJAY A SHETH,
HANUDATTA S ATREYA AND KANDALA V R CHARY*

*Department of Chemical Sciences, Tata Institute of Fundamental Research,
Homi Bhabha Road, Mumbai 400 005 (India)*

(Received 31 March 2003 ; Accepted 21 April 2003)

EF-hand proteins belong to a growing family of calcium binding proteins (CaBPs) that are known to undergo large conformational changes from a Ca^{2+}-free (apo) to the Ca^{2+}-bound (holo) state. In this review, we first describe the characteristics of EF-Hand motives and follow it up with a review on the Ca^{+2} affinity, cooperativity, selectivity and displacement in these proteins. Further, the focus is on the recent NMR structural studies on various forms of EF-hand CaBPs in both apo and holo states.

Key Words: EF-Hand Protein; Calmodulin; *Eh*CaBP; Pseudocontact Shifts; Chemical Shifts; Ytterbium

Introduction

Calcium is one among the most commonly used ions, in a multitude of biological processes, which range from bone mineralization to cell signaling[1]. It triggers new life at fertilization. It controls several developmental processes and once cells have differentiated, it even functions to control diverse cellular processes such as metabolism, proliferation, secretion, contraction, learning and memory. It is even involved in cell death[2]. In all these processes, Ca^{2+} interacts with large number of proteins so called as calcium binding proteins (CaBPs)[3-5]. EF-hand Ca^{2+} binding proteins (hereafter referred to as EF-CaBP) belong to a growing super family of Ca^{2+} binding proteins. Ever since the recognition of the EF-hand motif in the parvalbumin structure in 1973[6], more than 1000 distinct primary sequences in this class of proteins are known and catalogued into 66 different sub-families[4,5]. They bind co-operatively to Ca^{2+} in sub-nanomolar to millimolar range and function as signal transducers or modulators. Some of them have also been realized to be involved in extracellular functions such as cell migration, differentiation and association[7]. These proteins have been a subject of great interest for structural biologists, resulting in the availability of more than 300 three-dimensional (3D) structures, as of today[3-5,8], both by NMR spectroscopy and X-ray crystallography (Table I). This can be primarily attributed to the development of multinuclear and multidimensional NMR techniques for solving the solution structures of large biomolecules, during the last one decade. Interestingly, most of the known 3D structures of Ca^{2+} free (Apo-form) EF-CaBPs are revealed by high resolution NMR spectroscopy. Besides, NMR has been extensively used in understanding not just the 3D structures and dynamics of these proteins, but also the Ca^{2+} affinity, cooperativity, selectivity and displacement among the various EF- hand units present in a given protein.

In the following paragraphs, we first describe the characteristics of EF-Hand motifs and then embark on to review the state-of-the-art in our understanding of various EF-CaBPs and their functional properties. The current review highlights recent NMR studies and limits to correlating these with their biological functions. In addition, wherever appropriate, biophysical studies other than X-ray are discussed.

EF-Hand Ca^{+2} Binding Loops

The canonical Ca^{2+}-binding motif (hereafter referred to as the Ca^{2+}-binding loop) in these proteins consists of a contiguous 12-residue loop flanked by 2 helices, forming the so-called "EF-hand motif"[6] (see Fig. 1A). Within the Ca^{2+}-binding loop, residues that are involved in co-ordination with Ca^{2+} are 1, 3, 5, 7, 9 and 12, forming a pentagonal bi-pyramidal geometry[3,4] as shown below (highly conserved residues are shown and ±X, ±Y, ±Z refer to the vertices of the pentagonal bi-pyramid) (see Fig. 1B). The numbering is relative to the Ca^{2+}-binding loop[3,4].

1	2	3	4	5	6	7	8	9	10	11	12
D	·	(D/N)	·	(D/N)	·	·	·	·	·	·	E
X		Y		Z		-Y		-X			-Z

*Author for Correspondence:
E-mail; chary@tifr.res.in; *Tel. :* 0091-22-2215 2971 Ext. 2489; *Fax :* 0091-22-2215 2110

Table I

NMR Structures of EF-Hand Calcium-Binding Proteins Deposited in PDB as on Janaury 6, 2003

Protein	PDB CODE	APO/HOLO STATE	REFERENCE
Calmodulin			
Vertebrate caltractin (centrin)	1CFC	Apo	R1
	1DMO	Apo	R2
Vertebrate	1AK8	Holo	R3
N domain vertebrate	1CMF	Apo	R4
C domain	1CMG	Holo	R4
Vertebrate complexed with peptide	2BBN	Holo	R5
Troponin C			
Skeletal calmodulin	1TNW	Holo	R6
N domain	1TRF	Apo	R7
	1TNP	Apo	R8
	1SMG	Holo	R9
Cardiac N domain	1SPY	Apo	R10
	1AP4	Holo	R10
Cardiac C domain	3CTN	Holo	R11
Parvalbumin			
Alpha lineage	1PAS	Holo	R12
Calbindin D_{9k}			
bovine	1CLB	Apo	R13
	2BCB	Holo	R14
	1CDN	Holo	R15
	1BOD	Holo	R16
porcine	1CB1	Holo	R17
Recoverin			
bovine	1IKU	Apo	R18
bovine with hippocalcin myristoylated	1JSA	Holo	R19
S100B			
bovine	1CFP	Apo	R20
human	1UWO	Holo	R21
rat	1SYN	Apo	R22
	1QLK	Holo	R23
Calcyclin			
rabbit	1CNP	Apo	R24
	1A03	Holo	R25
EhCaBP	1JFK	Holo	R26

R1 1CFC H Kuboniwa, N Tjandra, S Grzesiek, H Ren, C B Klee and A Bax *Nature Struct Biol* 2 (1995) 768

R2 1DMO M Zhang, T Tanaka and M Ikura *Nature Struct Biol* **2** (1995) 758

R3 1AK8 D Bentrop, I Bertini, M A Cremonini, S Forsén, C Luchinat and A Malmendal *Biochemistry* **36** (1997) 11605

R4 1CMF, 1CMG B E Finn, J Evënas, T Crakenberg, J P Waltho, E Thulin and S Forsén *Nature Struct Biol* **2** (1995) 777

R5 2BBN M Ikura, G M Clore, A M Gronenborn, G Zhu, C B Klee and A Bax *Science* **256** (1992) 632

R6 1TNW C M Slupsky and B D Sykes *Biochemistry* **34** (1995) 15953

R7 1TRF W A Findlay, F D Sönnichsen and B D Sykes *J Biol Chem* **223** (1994) 601

R8 1TNP S Gagné, S Tsuda, M Li, L Smillie and B Sykes *Nature Struct Biol* **2** (1995) 784

R9 1SMG S M Gagné, M X Li and B D Sykes *Biochemistry* **36** (1997) 4386

R10 1SPY, 1AP4 L Spyracopoulos, M X Li, S K Sia, S M Gangne, M Chandra, R J Solaro, Sykes *Biochemistry* **36** (1997) 12138

R11 3CTN S K Sia, M X Li, L Spyracopoulos, S M Gagné, W Liu, J A Putkey, B D Sykes *J Biol Chem* **272** (1997) 18216

R12 1PAS A Padilla, A Cavé, J Parello *J Mol Biol* **204** (1998) 995

R13 CLB N J Skelton, J Kördel and W J Chazin *J Mol Biol* **249** (1995) 441

R14 2BCB J Kördel, N Skelton, M Akke and W Chazin *J Mol Biol* **231** (1993) 711

R15 1CDN M Akke, S Forsén and W J Chazin *J Mol Biol* **252** (1995) 102

R16 1BOD C Johansson, M Ullner and T Drakenberg *Biochemistry* **32** (1993) 8429

R17 1CB1 M Akke, T Drakenberg and W J Chazin *Biochemistry* **31** (1992) 1011

R18 1IKU T Tanaka, J B Ames, T S Harvey, L Stryer and M Ikura *Nature* **376** (1995) 444

R19 1JSA J B Ames, R Ishima, T Tanaka, J I Gordon, L Stryer and M Ikura *Nature* **389** (1997) 198

R20 1CFP P M Kilby, L J Van Eldik and G C K Roberts *Structure* **4** (1996) 1041

R21 1UWO S P Smith and G S Shaw *Structure* **6** (1998) 211

R22 1SYN A C Drohat, J C Amburgey, F Abildgaard, M R Starich, D Baldisseri and D J Weber *Biochemistry* **35** (1996) 11577

R23 1QLK A C Drohat, D M Baldisseri, R R Rustandi and D J Weber *Biochemistry* **37** (1998) 2729

R24 1CNP B Potts, J Smith, M Akke, T Macke, K Okazaki, H Hidaka, D Case and W Chazin *Nature Struct Biol* **2** (1995) 790

R25 1AO3 M Sastry, R R Ketchem, O Crescenzi, C Weber, M J Lubienski, H Hidaka and W J Chazin *Structure* **6** (1998) 223

R26 1JFK H S Atreya, S C Sahu, A Bhattacharya, K V R Chary and G Govil *Biochemistry* **40** (2001) 14392

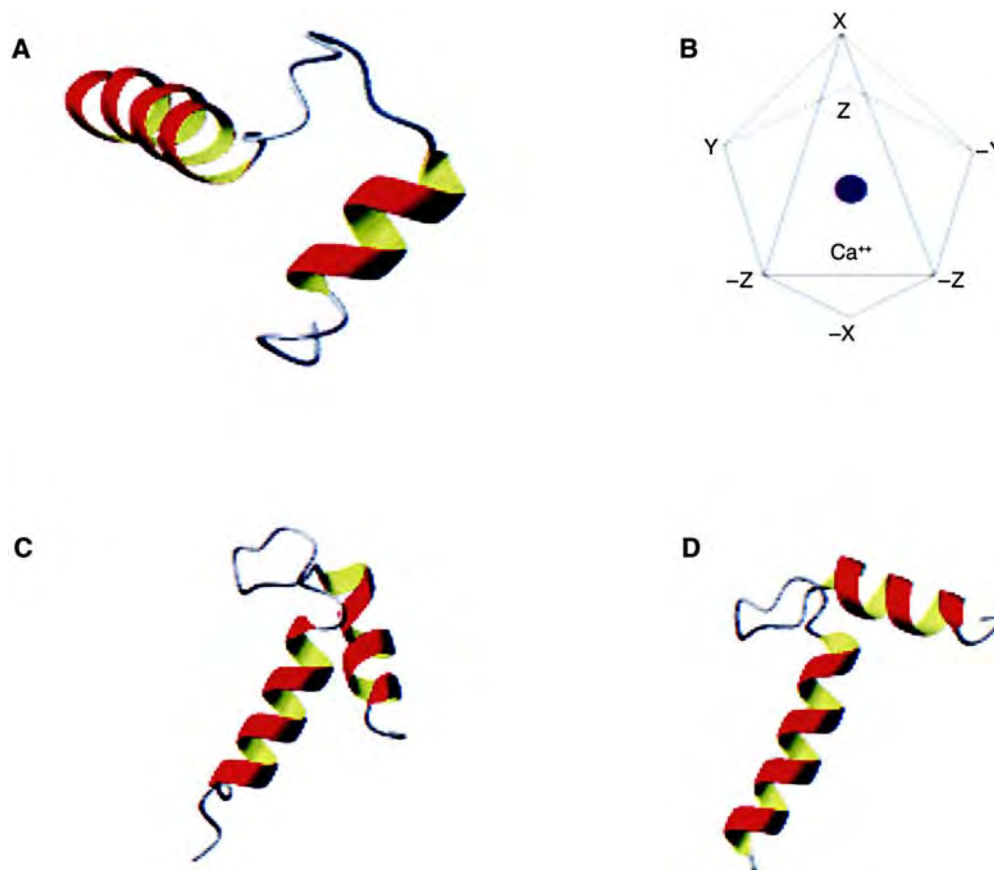


Fig. 1 (A) Helix-loop-helix motif, (B) Pentagonal-bipyramidal geometry. Illustrative examples of (C) Ca^{2+} -free EF-hand motif, where in the E and F helices are almost antiparallel and (D) Ca^{2+} -bound EF-hand motif, where in the E and F helices are almost perpendicular.

The residues at position 1 and 12 are invariant Asp and Glu, respectively, while the ones that are at positions 3 and 5 can be either Asp or Asn. These residues co-ordinate directly to the Ca^{2+} via their side-chain carboxylate groups. Residue at position 7 co-ordinates to Ca^{2+} via its backbone carbonyl group, while residue at position 9 coordinates to Ca^{2+} indirectly through an intervening water molecule[3]. Conventionally, such a Ca^{2+} binding sequence is termed as an EF-'loop' or a Ca^{2+} binding loop, although the last three positions in it initiate the second helix (or F-helix). The flanking helices (E and F) of the Ca^{2+} binding loop are amphiphilic and possesses 4 conserved hydrophobic residues at positions –1, -4, -5 and –8 in the first helix (E-helix), and positions 13, 16, 17 and 20 in the second helix (F-helix)[8].

A variant of this canonical EF-hand motif, the second most common one, termed as pseudo EF-Hand is seen in the most recently determined structures, belonging to the S100 protein family[3,4]. This has a 14 residue loop, in which most of the calcium ligands are backbone carbonyls. The binding of Ca^{+2} however conforms to the classical pentagonal bipyramidal coordination[3,4].

In both the above mentioned motifs, the final ligand of the binding group is always highly homologous bidentate Glu or Asp, which provides two side-chain carboxylate oxygens. Such a bidentate ligand is expected to provide favourable interaction energy from two oxygens, though it pays only the same entropic penalty for correctly orienting the side-chain as would a monodentate ligand (example: Asn). Mutating this residue with Asn, which is able to provide only one ligand, is expectedly found to reduce the calcium affinity[9].

EF-Hand Motif Mostly Occur in Pairs

In EF-CaBPs, the repeating unit, the "EF-hand", mostly occurs in pairs. Thus, most of the CaBPs have even number of such repeat units [See Table I and II & Fig. 2]. Such pairing is stabilized by a short (3 residues long) anti-parallel β-sheet formation between the two loops and an extensive hydrophobic interaction between the two EF-hand motifs. Such a unit forms a functional domain, which enhances affinity of each of the

Table II

Domain Structures in EF-Hand Calcium Binding Proteins

CaBP	No of EF-hands	No of functional EF-hands	No of independent domains.
Calbindin D_{9k}	2	2	Single
Parvalbumin	3	2	——
CaM	4	4	Two
Troponin C	4	4	Two
Recoverin	4	4	Single(packed together)
Sarcoplasmic	4	4	Single(packed together)
S100	2	2	Single(Homodimerize or heterodimerize)
Calpain domain VI	5	5	Single(Homodimerize)

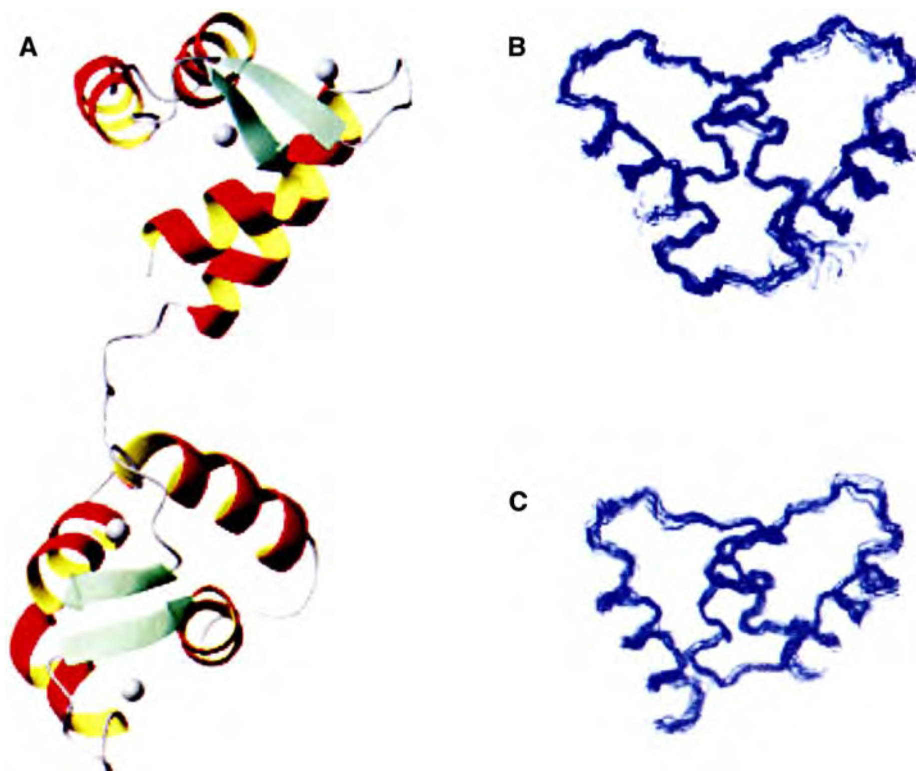


Fig. 2 (A) 3DNMR solution of Ca^{2+}–bound *Eh* CaBP (PDB Code 1JFK)[5]; (B) 25 superimposed structures of N terminal domain (4-62) of *Eh* CaBP; (c) 25 superimposed structures of c-terminal domains (70-134) of *Eh* CaBP.

participating EF-hands for Ca^{2+}. It is also thought to provide a means for positive cooperativity in binding Ca^{+2} to the individual loops. The smallest member of the family is calbindin D_{9k} with two EF-hand units, while the largest known member is the Calpain domain VI with 5 EF-hand units per monomer [Table I and II]. The unpaired EF-hand in the Calpain domain VI induces dimerization, resulting in a 10 EF-hand unit long EF-CaBP.

It is interesting to note that calcium induces association of synthetic peptides with the sequence of an EF-hand motif, to form an intact protein domain[10]. On the other hand, peptides representing two different EF-hands were found to preferentially pair to form natural heterodimers than the formation of homodimers[11].

Ca^{+2} Affinity, Cooperativity and Selectivity

The most important Ca^{+2} binding properties of a given EF-hand pair include the calcium affinity, the cooperativity of Ca^{2+}, and the selectivity of the loops for Ca^{2+} over other metal ions and finally the displacement of Ca^{2+} by other metal ions, such as tripositive lanthanide metal ions (Ln^{3+}). These binding properties essentially correlate with respective functions[12]. The question commonly posed is the one

that addresses as to whether the primary sequence of EF-hand units have any signatures, which can be directly correlated to these properties? In recent years, there has been a growing interest in understanding the factors that govern Ca^{2+}-binding affinities in EF-hand proteins[12]. This is primarily because a wide-range of Ca^{2+} binding affinities/ specificities has been found among these proteins, despite the strong conservation in their primary sequence of Ca^{2+}-binding loops. Several studies[13-16] have concentrated on understanding Ca^{2+} (which is present in micromolar concentrations) binding affinity/ specificity in these proteins, against Mg^{2+}, the only other competitor for Ca^{2+} inside the living cell, which is present in millimolar levels. Models based on the primary sequence of the Ca^{2+}-binding loop have been proposed to explain the origin of differences in Ca^{2+} binding affinities/specificities[14,17,18]. For example, in one of the studies, the presence of neighbouring Asp residues at the +X and +Y, the +Y and +Z, or the +X, +Y and +Z metal ion coordinating positions have been shown to lower metal binding affinity due to dentate-dentate repulsion[17]. In another study, the side-chain of the ligating residue at position +Y in the Ca^{2+}-binding loop has been shown to be important[19]. In yet another reported work, EF-helices that flank the Ca^{2+}-binding loops have also been realized to modulate the Ca^{2+}-binding affinity in this class of proteins[20]. In addition, it has been shown that mutation in non–ligating residues in the binding loop can cause a marked change in the Ca^{+2} affinity, owing primarily to the differences that arise in the electrostatic profile of the loop. The final ligand of the binding loop Glu is found to be a critical residue for Ca^{+2} affinity as mentioned above[9].

The cooperativity of the Ca^{+2} binding to the binding loops found in a functional unit is very important for biological functions. Binding of Ca^{2+} to a pair of sites is cooperative if the binding of Ca^{+2} ion to one site influences the affinity of the 2nd site for Ca^{2+}. If the Ca^{+2} binding to one site enhances the affinity of the 2nd ion, the cooperativity is understood to be positive. Most of the EF-hands are found to exhibit such positive cooperativity[21,22]. Signal transducers remain entirely in 'off' state until Ca^{+2} levels in the cell reach the level corresponding to a calcium signal at which point they must switch entirely to their 'on' state (Ca^{+2} loaded). Positive cooperativity in CaBP provides the means for such a tight signaling system. Now the question to be addressed relates to the specification of core residues, which are most important for such cooperativity. The only knowledge we have so far is from the structural studies on Calbindin D_{9k}[22]. Even mutation of amino acid residues, which are at a distance from the calcium-binding loop, has been found to either increase or decrease the cooperativity[22].

Ca^{+2} Displacement

During the last few years, there have been several studies dealing with the displacement of Ca^{2+} in a Ca^{2+}-saturated (holo) protein by other metal ions. Such studies yield information on the relative specificity and strength with which Ca^{2+} is bound to the protein. Tri-positive lanthanide ions (Ln^{3+}) serve as ideal candidates for the displacement of Ca^{2+} in a protein owing to their similar ionic radii as Ca^{2+}[23]. Such a study yields valuable information on the relative binding affinities of Ca^{2+} and Ln^{3+} for the protein under investigation. Lanthanides such as Tb^{3+} and Eu^{3+} possess optical properties and facilitate structural studies in substituted proteins using fluorescence and UV-Visible absorption spectroscopy[18,21]. Another favourable property of paramagnetic Ln^{3+} is their large anisotropic magnetic susceptibility, which has been harnessed in NMR studies of Ln^{3+}-substituted proteins[24]. The large pseudocontact shifts, observed for residues as far as 40 Å from the metal centre, have been used for 3D structure refinement from NMR data. On the other hand, residual dipolar couplings that arise due to the partial alignment of proteins containing Ln^{3+} in high magnetic field have been used as refinement tools.[25] These also assess the domain alignment in multi-domain EF-hand proteins such as calmodulin (CaM)[26].

Structural Basis for Sequential Displacement of Ca^{2+} by Yb^{3+} in a EF-CaBP

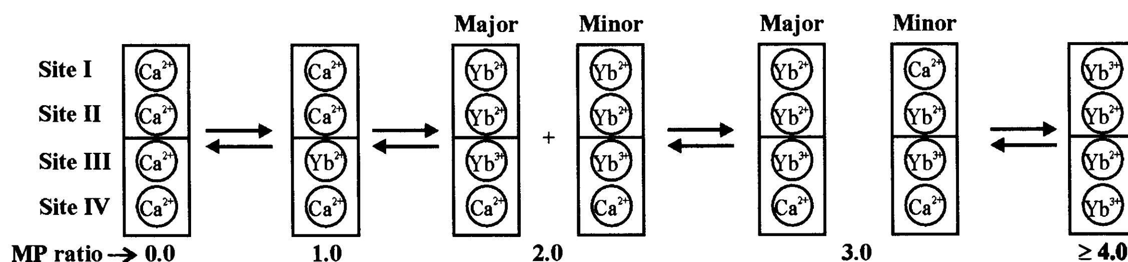
In order to characterize the factors that govern Ln^{3+} binding to EF-hand proteins and harness their favourable magnetic properties for structural studies by NMR, in our laboratory, we carried out Ln^{3+}-substitution studies in an EF-hand Ca^{2+}-binding protein from *Entamoeba histolytica* (hereafter referred to as *Eh*CaBP) by NMR and thermodynamics. Yb^{3+} has been chosen as the paramagnetic probe owing to its large pseudocontact shift to Curie line-broadening ratio[25]. We have used isothermal titration calorimetry (ITC) to obtain binding affinities and free energy changes upon substitution of Ca^{2+} by Yb^{3+}. We have demonstrated, for the first time, the utility of NMR and thermodynamics interwoven to obtain an insight on the

relative binding specificities/affinities between Ca^{2+} and Yb^{3+} [27]. As revealed by the NMR titration experiments, Yb^{3+} are found to displace Ca^{2+} from the four metal binding sites present in *Eh*CaBP in a sequential manner. The lanthanides preferentially bind first to Site III in the C-terminal domain and then to sites II and I in the N-terminal domain. They are found to occupy Site IV only to an extent of 70% of the original population. On the other hand, the ITC measurements yielded four macroscopic binding constants, which could be assigned to the Ca^{2+} displacement by Yb^{3+} from respective Ca^{2+}-binding sites in *Eh*CaBP. Based on the NMR data, the highest binding constant obtained in ITC has been ascribed to the displacement of Ca^{2+} by Yb^{3+} in site III, while the lowest to the similar process in site IV. The remaining two macroscopic binding constants, obtained using ITC could then be assigned to the Ca^{2+} displacement by Yb^{3+} in sites I and II in the N-terminal domain. During the course of titration wherein the metal protein ratio was changed from 1 to 3, NMR studies indicated that Yb^{3+} displaces Ca^{2+} preferentially in Site II. Thus, out of the remaining two binding constants, the smaller one could be assigned to Ca^{2+} displacement by Yb^{3+} in site II and the other one to the equilibrium constant of Ca^{2+} displacement by Yb^{3+} in site I, when Yb^{2+} is already bound in its neighbouring site II. In order to verify whether the above assignments of thermodynamic constants to the various reactions are correct, plots of normalized volumes of cross peaks were back calculated by Atreya *et al.*[27], using the equilibrium constants obtained from ITC. As an illustrative example, the experimental and simulated plots of intensity profiles are shown in Figs. 3a and 3b for the original and pseudocontact shifted cross peaks of G15 and G122. As evident from the Figure, simulated plots (Fig. 3b) obtained using the thermodynamic constants are almost identical to the intensity profiles obtained for various species based on NMR titration data (Fig. 3a). Thus, using the NMR and ITC data in concert Atreya *et al.*[27], could arrive at the following scheme of reaction for Ca^{2+} displacement by Yb^{3+} in *Eh*CaBP that is most consistent with both the experimental data. In Scheme 1, white colours indicate metal-binding sites completely filled with Ca^{2+} whereas those colored in gray indicate Yb^{3+} filling.

Further, the study provided structural basis for such a sequential Ca^{2+} displacement by Yb^{3+} in *Eh*CaBP[27]. The highest preference of Yb^{3+} for Site III has been attributed to: (i) the presence of a large number of negatively charged residues that co-ordinate with the metal ion and (ii) the presence of a Tyr at –4 position with respect to the loop, instead of a highly conserved Phe, which results in a relatively more open conformation of the loop. Thermodynamics studies reveal a 10-100 fold higher equilibrium constant for Ca^{2+} displacement by Yb^{3+} at Site III as compared to the rest of the sites[27]. Further, the presence of charged residues in X, Y and Z positions in the Ca^{2+}-binding loop II which co-ordinate to the metal, leads to a dentate-dentate repulsion and hence lowers the affinity for Ca^{2+} as compared to Site I. This results in preferential Ca^{2+} displacement by Yb^{3+} in Site II compared to Site I. Unlike in the past, where Ca^{2+}-binding specificity/affinity has been explained on the basis of either the Ca^{2+}-binding loop sequence or the helices flanking the loop, in this study, both factors have been proposed to influence the specificity of the protein for Ca^{2+} [27].

Signal Transducers Undergo Ca^{2+}-Induced Conformational Change

Calmodulin (CaM) and Troponin C (TnC) are the most studied calcium signal transducers (see Table I and II). They function by undergoing a Ca^{2+}-induced conformational change, which exposes a concave hydrophobic surface in each domain of the protein. These surfaces then interact with target proteins, thereby transducing the calcium signal. The exposure of the hydrophobic surface is expectedly unfavourable



Scheme I

but functionally unavoidable, as this part of the protein interacts with its target. Now, the focus would move on to the energy balance for the exposure of hydrophobic surface. This energy is supposed to come from the large increase in the solvent entropy upon Ca^{2+} extraction from the solvent by the protein and this energetic cost seems to be 'paid for' by a reduction in the calcium affinity.

Ca^{2+} induced conformational change in these signal transducers were reported first by Herzberg *et al.*[28], by modelling the conformation of Ca^{2+}-free domains. This was based on the structural information they had on the N-terminal domain of TnC. According to this model, which is popularly known as the HMJ (named after the three authors of the model; namely, O Herjberg, J Moult and M N G James model[28]), the helices with in an EF-hand unit are anti-parallel in its apo-state and are almost perpendicular in the holo-state (Fig. 1C). The HMJ model remained the only view of the Ca^{2+}-induced conformational changes until 1995, when NMR solution structures of apo and Ca^{2+}-loaded CaM and TnC allowed the first direct determination of these conformational changes in regulatory EF-hand domains[11,12,29,30]. Comparison of the apo and Ca^{2+}-loaded states of CaM threw light upon the calcium-induced conformational changes. In the apo domains, all inter helical interfaces are well packed without any interaction between helices I and III. In the calcium-loaded domains, the two *intra*-EF-hand interfaces (I/II and III/IV) 'swing open' into an inverted-V shape, with the Ca^{2+}- binding loop at their vertex[4] [see Figs. 1C and 1D]. Most importantly, the central helix, connecting the two domains of CaM, as seen in the crystal structure, turned out to be a flexible linker in both the apo and holo states of the protein in solution, as established by the NMR studies[31,32]. This was further substantiated by two independent NMR studies on the structures of the Ca^{2+} loaded CaM-target peptide complexes[33,34]. In both these studies, the CaM is found to be compact as the two domains wrap around the target peptide without much change in the conformation of the individual domains as seen in the peptide free holo from of CaM.

It is worth mentioning to hear about the recent work of Chou *et al.*[35] who showed for the first time, how RDC's can be used for structure determination in proteins, with special reference to a CaBP. Their work on calmodulin (CaM) showed that the N-and C-terminal domains align with the magnetic field in different degrees, which indicates that the relative orientation of the two domains is modulated dynamically in solution. This study further reveals that N-and C-terminal domains are more closed than their counterparts seen in crystal structures. It is believed that Ca^{2+} binding switches the domain from closed to open form upon binding of calcium. The flexible linker and orientations of the domains are essential for determination of the function of any CaBP, especially it's binding with the other proteins involved in signal transduction.

The solution structure of holo form of *Eh*CaBP revealed the CaM paradigm of two independent globular domains connected by a flexible linker. It also consists of four canonical EF-hands, a pair each in the N- and C-terminal domains [Fig. 2A]. The Ca^{2+}-binding loops in the protein are canonical loops 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[3,4]. E*h*CaBP is structurally related, most closely to CaM[29,30,36] and TnC[12,37,38] in spite of low sequence homology with these proteins. Further, *Eh*CaBP has been thought to be involved in a novel signal transduction pathway, distinct from CaM.

The NMR derived solution structure of *Eh*CaBP[5,39] provided clues on the putative function of the protein. Although EhCaBP, 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 in the case of both CaM and TnC. In *Eh*CaBP, the flexibility in the linker region, imparted by Gly residues, is supposed to be responsible for specificity of the protein towards its interaction with target molecules. Specificity in *Eh*CaBP may also arise from more exposed hydrophobic residues as compared to CaM and TnC. In the case of *Eh*CaBP, the presence of Tyr (Y81) in the C-terminal domain instead of a more conserved Phe in that position is supposed 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. The studies revealed that *Eh*CaBP operates in a distinct manner compared to CaM and TnC, an observation that is supported by biochemical studies[39]. It is expected that the 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, should provide more insight into the dynamics, interactions and Ca^{2+} induced structural changes in *Eh*CaBP.

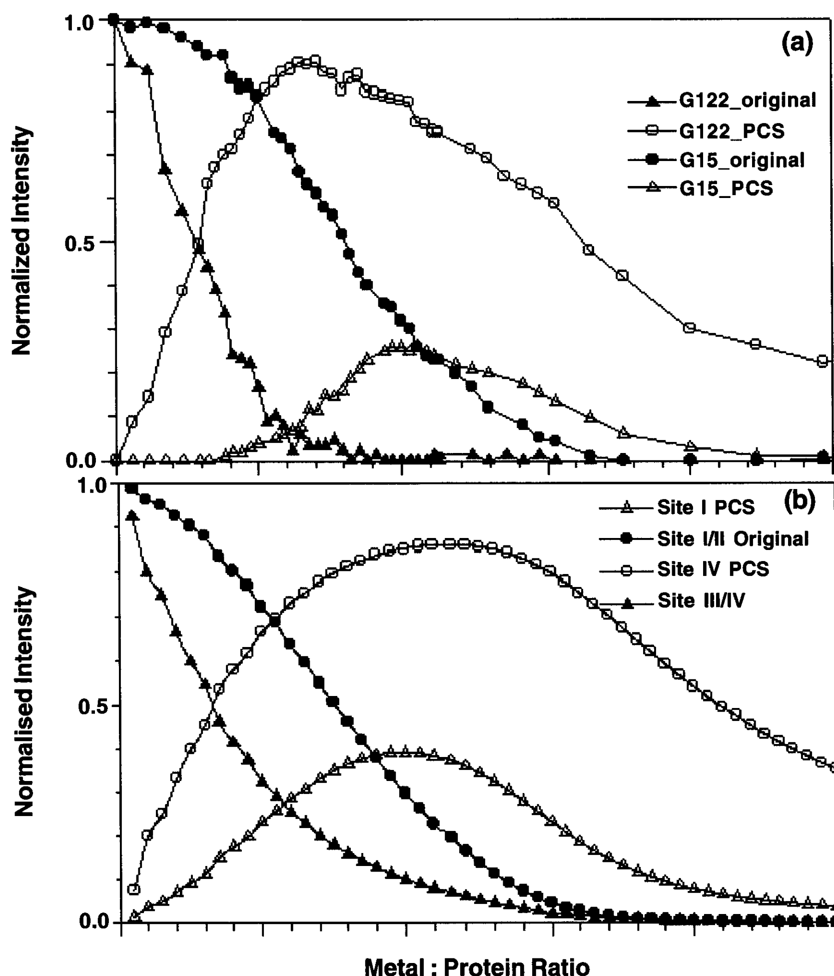


Fig. 3 (a) Experimental and (b) simulated plots of intensity profiles of the original and pseudocontact shifted cross peaks of G15 and G122. The thermodynamic constants derived by ITC measurements have been used in simulations[27]

Signal Modulators do not Undergo Ca^{2+}-Induced Conformational Change

Parvalbumin and calbindin D_{9k} are the most studied calcium signal modultors (see Table I and II). They do not undergo large Ca^{2+}-induced conformational change, as is usually observed in the case of signal transducers. Parvalbumin is supposed to be involved in quenching the Ca^{2+} signal in muscle cells. It is supposed to function as a Ca^{2+} buffer. Such buffering action, at times, is known to deactivate signal transduction. Parvalbumin was the first known structure of an EF-hand protein. It has 3 EF-hand units (Table II), out of which the N-terminal EF-hand is non-functional, with its E and F helices anti-parallel. On the other hand the other two EF-hand units (EF2 and EF3) adopt conformation close to the ones observed in Ca^{2+}-loaded CaM. Further, the unpaired EF1 pack into the hydrophobic surface created by EF2 and EF3, resulting in an increase in their Ca^{2+} affinity, as it prevents exposure of hydrophobic surface to a large extent. It has also been shown that, deletion of EF1 reduces the Ca^{2+} affinity by about 50 fold. The lost affinity could be restored by adding EF1 to EF2-EF3 fragment[40].

The other signal modulator that has been studied extensively is the Cabindin D_{9k}, which also does not expose its hydrophobic surface upon Ca^{2+} binding. It has two EF-hand units. The N-terminal one is a pseudo EF-hand binding loop that is seen in S100 sub-family, while the C-terminal one is the canonical one. Though both these units have similar binding affinities, the C-terminal EF-hand is found to undergo relatively larger structural change compare to the N-terminal pseudo EF-hand. Cabindin D_{9k} is supposed to be involved in Ca^{2+} uptake and transport. Cabindin D_{9k} was the first EF-hand protein whose structure was determined in both Apo and Holo forms. In both of these forms, it is

well packed and the conformation resembles that of CaM domains in their Apo form[41-43].

Recoverin

Recoverin, which is found in retina, belongs to a completely new subfamily of EF-CaBPs[44]. It has 4 EF hands. When calcium binds to recoverin there, is a large conformational change that causes the flexible loop with the myristoyl post-translational modification to swing out - exposing the myristoyl fatty acid. The hydrophobic fatty acid is then free to insert into membranes. In the holo state, recoverin is supposed to bind to rhodopsin kinase and block the phosphorylation of photoexcited rhodopsin, which in turn prolongs the photo response. This whole process *in vitro* is supposed to be enhanced by myristoylation of recoverin[45].

Recoverin is the lone member of this subfamily for which NMR structures of both apo and holo forms (in the holo form, instead of myristoyl group, a myristate analog was used to improve the solubility of the protein) are known[30,46]. It does not follow the CaM paradigm of two independent domains, which are connected by a flexible linker. Instead, it has four EF-hands that are arranged in one globular domain, with extensive interactions not only within the EF1/EF2 and EF3/EF4 pairs, but also between EF2 and EF3 in both the apo and holo forms of the protein. Besides, it has a longer linker between the EF3 and EF4 and an additional helix at the N-terminal end with an attached myristoyl group. In the apo state of the recoverin, this myristoyl group is completely buried in the core of the protein with close contacts from residues belonging to EF1, EF2 and EF3 units, while in the holo state, the modified myristoyl group is completely exposed to the solvent. The Ca^{2+} induced conformational changes are mainly in the reorganization of the packing between individual EF-hand units and between EF units and the linker loops. The question that is still to be answered is whether such conformational changes have any influence on the interaction of recoverin with its targets.

S100 Proteins

The S100 proteins belong to a unique subfamily of EF-hand CaBPs with 2 EF-hand units, one of which has a pseudo Ca^{2+} binding loop [the N-terminal] discussed above and the other one is the canonical Ca^{2+} binding loop [the C-terminal]. These proteins dimerize and pack together into a 4 EF-hand single globular domain. These are involved in Ca^{2+} transport and regulation of several fundamental cellular processes. NMR structures of both apo and holo forms of two proteins belonging to this sub-family, namely, calcyclin and S100B [See Table I] were deposited in the PDB. There is a high degree of sequence homology among S100 proteins and they all have similar tertiary fold. However, it is surprising to note that there are significant differences in their conformation[34,47]. The 3D NMR structures of the holo form of rat and human S100B proteins are determined[48,49]. The NMR structure of the Ca^{2+} loaded calcylin is also known[50]. The most important inference from these structural studies is that, the 3D structures of S100 proteins and that of Calcyclin do not undergo much conformational change on Ca^{2+} binding.

Interaction studies between S100 and peptides derived from target proteins threw light upon the nature of target recognition[48,51]. A Ca^{2+} induced hydrophobic channel surrounded by several acid residues is supposed to be the target site, based on the 3D structure of S100B[49,52]. Coming years should throw more light upon understanding more about the target recognition in atomic detail, with the availability of the knowledge of 3D structure of many more S100 proteins belonging to this subfamily of EF-hand CaBPs.

Conclusion

Ever since the pioneering work of Kurt Wuthrich in unraveling the protein solution structures by NMR in 1980s, the unraveling of protein structures by NMR has increased exponentially. Till date, the protein data bank *(http://www.rcsb.org/pdb/)* holds 1500 (as on January 27, 2003) NMR structures deposited in it. It is quite interesting to note that the limitations of high resolution NMR, as an analytical tool for macromolecular structure determination are gradually receding with its exponential quality improvement. It is surprising to see the molecular weight limit approaching 1000 kDa[53]. On the other hand, as mentioned above, last decade has provided structures of several EF-hand CaBPs and their complexes with target peptides and other compounds, throwing light upon the intricate details of their structures at atomic resolution and on the mode of recognition and interaction. Yet, several questions remain unresolved about the EF-hand CaBPs as of today. Very few subfamilies of this fascinating field have been the subject of structural investigation though NMR has attracted many researchers. Most of these subfamilies are represented structurally by one or two proteins. For most of the subfamilies the structural information, whether it is by NMR or X-ray, is still not available. With the knowledge of more and

more structures of EF-hand CaBPs and their complexes now amenable with the state-of-the-art NMR techniques, we should soon be able to understand this fascinating molecular mechanism of the transmission of Ca^{2+} triggered signals to the effectors of the message. Thus, in coming years, on one hand there appears great scope for unraveling novel conformations in this protein family, while on the other, deciphering biological functions from such structural information is yet another exciting and challenging task. Therefore, we look forward to stimulating results, which should provide greater insights into the intricate details of this fascinating world of EF-CaBPs structure, dynamics and interaction, which in turn provide insights into their biological functions.

Acknowledgements

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.

Abbreviations and Symbols

NMR: Nuclear Magnetic Resonance; ITC: Isothermal Calorimetry; HSQC: Heteronuclear Single Quantum Coherence; *Eh*CaBP: *Entamoeba Histolytica* Calcium Binding Protein; CaM: Calmodulin; TnC: Troponin C

References

1 J B Michael, D B Martin and L Peter *Nature* **395** (1998) 645

2 M Allegrozzi, I Bertini, M B L Janik, Y M Lee, G Liu and C Luchinat *J Am Chem Soc* **122** (2000) 4154

3 N C Strynadka and M N James *Annu Rev Biochem* **58** (1989) 951

4 M R Nelson and W J Chazin *Biometals* **114** (1998) 297

5 H S Atreya, S C Sahu, A Bhattacharya, K V Chary and G Govil *Biochemistry* **40** (2001) 14392

6 R H Kretsinger and C E Nockolds *J Biol Chem* **248** (1973) 3313

7 E Hohenester, P Maurer and R Timpl *EMBO J* **16** (1997) 3778

8 C A McPhalen, N C J Strynadka and M N G James *Adv Prot Chem* **42** (1991) 77

9 J F Maune, C B Klee and K Beckingham *J Biol Chem* **267** (1992) 5286

10 G S Shaw, R S Hodges and B D Sykes *Science* **249** (1990) 280

11 B E Finn, J Evenas, T Drakenberg, P J Waltho, E Thulin and S Forsen *Nat Struct Biol* **9** (1995) 777

12 S M Gagne, S Tsuda, M X Li, L B Smillie and B D Sykes *Nat Struct Biol* **2** (1995) 784

13 H J Moeschler, J Schaer and J A Cox *Eur J Biochem* **111** (1980) 73

14 B Gopal, C P Swaminathan, S Bhattacharya, A Bhattacharya, M R N Murthy and A Surolia *Biochemistry* **36** (1997) 10910

15 R Gilli, D Lafitte, C Lopez, M Kilhoffer, A Makarov, C Briand and J Haiech *Biochemistry* **37** (1998) 5450

16 M Imaizumi, M Tanokura and K Yamada *J Biochem (Tokyo)* **107** (1990)

17 B J Mardsen, R S Hodges and B D Sykes *Biochemistry* **27** (1988) 4198

18 X Wu and R E Reid *Biochemistry* **36** (1997) 8649

19 Y Waltersson, S Linse, P Brodin and T Grundstrom *Biochemistry* **31** (1993) 7866

20 Y Sharma, S Chandani, M B Sukhaswami, L Uma, D Balasubramanium and T Fairwell *Eur J Biochem* **243** (1997) 42

21 W D Horrocks Jr and D R Sudnick *Acc Chem Res* **14** (1981) 384

22 B B Kragelund, M Jonsson, G Bifulco, W J Chazin, H Nilsson, B E Finn and S Linse *Biochemistry* **37** (1998) 8926

23 R D Shannon *Acta Cryst* **A32** (1976) 751

24 I Bertini, M B Janik, Y M Lee, C Luchinat and A Rosato *J Am Chem Soc* **123** (2001) 4181

25 I Bertini, C Luchinat and G Parigi *Solution NMR of Paramagnetic Molecules* Elsevier Amsterdam (2001)

26 R R Biekofsky, F W Muskett, J M Schmidt, S R Martin, J P Browne, P M Bayley and J Feeney *FEBS Lett* **460** (1999) 519

27 H S Atreya, S Mukherjee, K V R Chary, Y M Lee and C Luchinat *Protein Science* **12** (2003) 412

28 O Herzberg and M N G James *J Mol Biol* **203** (1988) 761

29 H Kuboniwa, N Tjandra, S Grzesiek, H Ren, C B Klee and A Bax *Nat Struct Biol* **2** (1995) 768

30 M Zhang, T Tanaka and M Ikura *Nat Struct Biol* **2** (1995) 758

31 G Barbato, M Ikura, L E Kay, R W Pastor and A Bax *Biochemistry* **31** (1992) 5269

32 N Tjandra, S E Feller, R W Pastor and A Bax *J Am Chem Soc* **117** (1995) 12562

33 M Ikura, G M Clore, A M Gronenborn, G Zhu, C B Klee and A Bax *Science* **256** (1992) 632

34 P M Kilby, L J Van Eldik and G C K Roberts *Structure* **4** (1996) 1041

35 J J Chou, L Shipeng, B K Clsude and Bax Ad *Nat Struct Biol* **8** (2002) 990

36 R Chattopadhyaya, W E Meador, A R Means and F A Quiocho *J Mol Biol* **228** (1992) 1177

37 C M Slupsky and B D Sykes *Biochemistry* **34** (1995) 15953

38 A Houdusse, M L Love, R Dominguez, Z Grabarek and C Cohen *Structure* **15** (1997) 1695

39 S C Sahu, H S Atreya, S Chauhan, A Bhattacharya, K V R Chary and G Govil *J Biomol NMR* **14** (1999) 93

40 E A Parmyakov, V N Medredkin, Y V Mitin, R H Kretsinger *Biochem Biophy Acta* **1076** (1991) 67

41 J Kordel, N Skelton, M Akke and W Chazin *J Mol Biol* **231** (1993) 711

42 N J Skelton, J K"rdel and W J Chazin *J Mol Biol* **249** (1995) 441

43 D M E Szebenyi and K Moffat *J Biol Chem* **261** (1986) 8761

44 T Tanaka, J B Ames, T S Harvey, L Stryer and M Ikura *Nature 376* **6539** (1995) 444

45 C K Chen, J Inglese, R J Lefkowitz and J B Hurley *J Biol Chem* **270** (1995) 18060

46 J B Ames, R Ishima, T Tanaka, J I Gordon, L Stryer and M Ikura *Nature* **389** (1997) 198

47 A C Drohat, J C Amburgey, F Abildgaard, M R Starich, D Baldisseri and D J Weber *Biochemistry* **35** (1996) 11577

48 A C Drohat, D M Baldisseri, R R Rustandi and D J Weber *Biochemistry* **37** (1998) 2729

49 S P Smith and G S Shaw *Structure* **6** (1998) 211

50 M Sastry, R R Ketchem, O Crescenzi, C Weber, M J Lubienski, H Hidaka and W J Chazin *Structure* **6** (1998) 223

51 B Potts, J Smith, M Akke, T Macke, K Okazaki, H Hidaka, D Case and W Chazin *Nature Struct Biol* **2** (1995) 790

52 H Matsumura, T Shiba, T Inoue, S Harada and Y Kai *Structure* **6** (1998) 233

53 J Fiaux, E B Bertelsen, A L Horwich and K Wuthrich *Nature* **418** (2002) 207