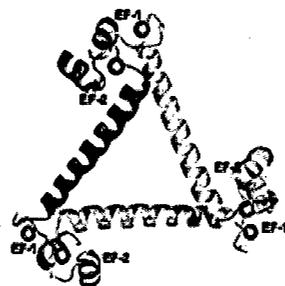


**STRUCTURAL STUDIES OF CALCIUM BINDING  
PROTEIN-1 FROM *Entamoeba histolytica***



**Thesis submitted to  
Jawaharlal Nehru University  
For the award of the degree of  
DOCTOR OF PHILOSOPHY**

**SHIVESH KUMAR**



**School of Life Sciences  
Jawaharlal Nehru University  
New Delhi - 110067  
INDIA  
2009**



**SCHOOL OF LIFE SCIENCES  
JAWAHARLAL NEHRU UNIVERSITY  
NEW DELHI - 110067  
INDIA**

**CERTIFICATE**

The research work embodied in the thesis entitled “**Structural Studies of Calcium Binding Protein-1 from *Entamoeba histolytica***” has been carried out in School of Life Sciences, Jawaharlal Nehru University, New Delhi.

This work is original and has not been submitted so far in part or full, for award of any degree or diploma of any university.

*Shivesh Kumar*

**SHIVESH KUMAR**

*(Candidate)*

**Dr. SAMUDRALA GOURINATH**

*(Supervisor)*

**Prof. R. MADHUBALA**

*(Dean)*

*Dedicated to...*

*My Parents*

## *Acknowledgement*

*It gives me immense pleasure to express my gratitude to the innumerable persons for providing refreshing insight, enthusiasm and criticisms, which have been helpful to me in the due course of completion of my PhD thesis. Without their generosity, interesting feedbacks and valuable suggestions which exhilarated me constantly, this thesis would have been incomplete.*

*I whole heartedly owe great gratitude to my mentor and supervisor Dr. S. Gourinath for his guidance and constant support throughout the thesis. His constant encouragement and precious contributions as well as his solidarity introduced me to minute methodological approaches which helped me in exploration to this research area. He helped me understand the complicated issues of this research area within the philosophy of science. His truly scientist intuition has developed him into a repository of ideas and passions in science, which exceptionally inspired and enriched my growth as a student, a researcher and a scientist want to be. Sir, your thoughts and input were always meaningful due to which I had an opportunity to learn and further developed a penchant to this interesting field of Biophysics. Collecting data is invariably a trying experience. Thanks owed to you for providing me with such a wonderful experience. I can just express my gratitude in few words i.e., Sir, I saw you as a light at the end of the tunnel.*

*Special thanks to Dr. Neelima Alam for her supervision, and crucial contribution during my initial years of research. Her involvement with her originality has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come.*

*My heartfelt thanks to Prof. Alok Bhattacharya, School of Life Sciences, J.N.U., for providing the clones of EhCaBP1 which scintillated my initial and exhaustive research on calcium-binding proteins.*

*I am much indebted to Prof. Sudha Mahajan Cowsik for her initial support and guidance and also for the valuable advice in various discussions and her unflinching encouragement and support in various ways.*

*My heartfelt and sincere thanks to Prof. R. Madhubala, Dean, School of Life Sciences and the former Dean, Prof. P. K. Yadava for providing me a platform to work in a scientifically rich environment.*

*I am very grateful to Dr. Dinakar Salunke, National Institute of Immunology, New Delhi and Dr. Amit Sharma, International Centre for Genetic Engineering and Biotechnology, New Delhi for providing me with the facility of XRD.*

*I gratefully acknowledge Prof. T.P. Singh and Dr. Savita Yadav, All India Institute of Medical Sciences for access to Spectroscatter 201 Dynamic Light scattering equipment*

*It is a pleasure to acknowledge Dr. Sharma, Mr. Mishra, Mr. Rajender and all CIF and CCIF members, Meenu mam and all staffs of the school for their help.*

*Many thanks go to the Director, Advanced Instrumentation Facility (A.I.F.), J.N.U., New Delhi to allow me using the X-ray source for sets of data collection.*

*I would also like to thank Dr. Mannicam Yogavel, I.C.G.E.B., Vineet and Tarique, N.I.I. for their help in data collection.*

*I gratefully acknowledge Narendra Padhan, Ruchi Jain, Anjali Dike for providing me unflinching encouragement and support in various ways.*

*Without camaraderie of my labmates Manish, Tara, Arif, Sudhir, Isha, Faisal, Gokula, Sanjeev, Nitesh and Mohit, whose presence somehow perpetually refreshed and created a pleasant working atmosphere. Thanks for the scientific discussions and sharing various thoughts during tea meetings. Thanks for such an exhilarating time we all spent together. I want to convey special acknowledgement to Arif for devoting his precious time to proof-read the thesis. I am also thankful to Sanjeev for the indispensable help, who is always ready to lend a hand.*

*I have been fortunate to come across many funny and good friends, without whom life would be bleak. My heartfelt thanks to Anupam, Siteshu, Rashmi, Ankush, Jyoti, Anita, Hina, Khyati, Babita, Ratnakaran, Hemant, Ashraf, Tarique, Raghu, Chanchal for their ever-present help and support in various ways.*

*I would like to thank Karan, Sharma Jee and Ved for laboratory assistance.*

*Special thanks go to Hema Adhikari, for her constant help and support since the beginning of my Ph.D. work,*

*I convey special acknowledgement to Department of Science and Technology, Govt. of India and Elettra Synchrotron source, Trieste, Italy for providing me the opportunity for the Data collection experiments with full travel allowance.*

*I gratefully acknowledge University Grant Commission for the financial support during the course of my Ph.D.*

*My Parents deserve special mention for their inseparable support and prayers. My father, Dr. Bijoy Mall, in the first place is the person who put the fundamental to my learning character, showing me the joy of intellectual pursuit ever since I was a child. My mother, Smt. Laxmi Devi, is the one who sincerely raised me with her caring and gentle love. My younger brother Jitesh, my elder sisters Sujata Di and Mukta Di and my brother-in-laws Ajay Kumar and Dev Kumar, thanks for being supportive and caring for the entire years.*

*Last but not the least, praise and thanks goes to 'The Almighty' for the many blessing undeservingly bestowed upon me.*

*This thesis arose in part out of years of research that has been done by me with the enormous and constant help of all the people during this entire study and made a mark in my life at the end. At the end I would like to quote the words of Ralph Waldo Emerson, which inspired me in due course of my Ph.D.*

*"Do not go where the path may lead, go instead where there is no path and leave a trail."*

*Shivesh Kumar*

## Abbreviations and Symbols

$\gamma$	gamma
$\lambda$	lambda
$\mu$	micro
Å	Angstrom
AIDS	Acquired Immune Deficiency Syndrome
APS	Ammonium Per Sulfate
Ba	Barium
bp	base pair
BSA	Bovine Serum Albumin
Ca	Calcium
CaBPs	Calcium Binding Proteins
CaM	Calmodulin
CBB	Coomassie Brilliant Blue
CCP4	Collaborative Computational Project 4
CDPK	Calcium Dependent Protein Kinase
CHCl <sub>3</sub>	Chloroform
Cl	Chloride
CNS	Crystallography and NMR System
COOT	Crystallographic Object-Oriented Toolkit
DEAE	Di Ethyl Amino Ethyl
DLS	Dynamic Light Scattering
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylene Diamine Tetra Acetate
EGTA	Ethylene Glycol Tetraacetic Acid
EhCaBP1	<i>Entamoeba histolytica</i> Calcium Binding Protein1
EtOH	Ethanol
GalNAc	N-Acetylglucosamine
GLB	Gel Loading Buffer
hr	hour
HSF	Heat Shock Factor
Hsp	Heat shock protein

IEC	Ion Exchange Chromatography
Ig	Immunoglobulin
Kb	Kilobase
kDa	kilo Dalton
LB	Luria Broth
LGT	Lateral Gene Transfer
M	Marker
mA	milli Ampere
MALDI	Matrix Assisted Laser Desorption Ionization
mg	milligram
Mg	Magnesium
min	minute
ml	millilitre
mM	milliMolar
MME	Mono Methyl Ether
MPD	2-methyl-2,4-pentandiol
MS	Mass Spectrometry
NaOH	Sodium hydroxide
NK	Natural Killer
nm	nanometer
NMR	Nuclear Magnetic Resonance
°C	degree centigrade
PAF	Polymerase Associated Factors
PAGE	Polyacrylamide Gel Electrophoresis
PAHO	Pan American Health Organization
Pb	Lead
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene Glycol
PFGE	Pulsed Field Gel Electrophoresis
pm	picomole
PMSF	Phenyl Methyl Sulphonyl Fluoride
RTK	Receptor Tyrosine Kinases
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid

Ins(1,4,5)P <sub>3</sub>	Inositol-1,4,5-trisphosphate
cADPR	cyclic ADP ribose
NAD	Nicotinamide-Adenine Dinucleotide
NAADP	Nicotinic Acid Dinucleotide Phosphate
PLC	Phospholipase C
rpm	revolutions per minute
SAD	Single-wavelength Anomalous Dispersion
SDS	Sodium Dodecyl Sulphate
sec	second
Sr	Strontium
SSC	Standard Saline Citrate
T <sub>10</sub> E <sub>1</sub>	Tris-EDTA
TAFs	TATA binding protein Associated Factors
TBE	Tris-Borate EDTA
TEMED	N,N,N',N', Tetramethylethylenediamine
T <sub>m</sub>	Annealing temperature
TnC	Troponin C
Tris	Tris (hydroxymethyl) amino ethane
tRNA	transfer- Ribo Nucleic Acid
U	Unit
UNESCO	United Nations Educational, Scientific and Cultural Organization
v/v	volume/volume
W	Watt
w/v	weight/volume
WHO	World Health Organization
β-ME	β-Mercaptoethanol

# TABLE OF CONTENTS

## ACKNOWLEDGEMENT

## ABBREVIATIONS AND SYMBOLS

i-iii

Page No.

## INTRODUCTION

1-33

<b>1.1</b>	<b><i>Entamoeba histolytica</i>: A pathogenic enteric protozoan.....</b>	<b>1</b>
<b>1.1.1</b>	<b>Historical background.....</b>	<b>1</b>
<b>1.1.2</b>	<b>Epidemiology.....</b>	<b>2</b>
<b>1.1.3</b>	<b>Life cycle of the parasite.....</b>	<b>3</b>
<b>1.1.4</b>	<b>Pathogenesis.....</b>	<b>4</b>
<b>1.1.5</b>	<b>Metabolism.....</b>	<b>6</b>
<b>1.1.6</b>	<b>Genome organization of <i>E. histolytica</i>.....</b>	<b>7</b>
<b>1.2</b>	<b>Ca<sup>2+</sup> as second messenger.....</b>	<b>8</b>
<b>1.2.1</b>	<b>The Ca<sup>2+</sup>-signalling toolkit.....</b>	<b>11</b>
<b>1.2.2</b>	<b>Calcium Binding Motifs.....</b>	<b>13</b>
<b>1.3</b>	<b>The EF-Hand: A Ca<sup>2+</sup>-Binding Unit.....</b>	<b>14</b>
<b>1.3.1</b>	<b>The canonical EF-loop.....</b>	<b>16</b>
<b>1.3.2</b>	<b>Non-canonical EF-loops.....</b>	<b>19</b>
<b>1.3.3</b>	<b>The EF-hand pair.....</b>	<b>22</b>
<b>1.3.4</b>	<b>EF-hand domain organization.....</b>	<b>23</b>
<b>1.4</b>	<b>Consequences of Ca<sup>2+</sup>-Binding.....</b>	<b>25</b>
<b>1.4.1</b>	<b>Ca<sup>2+</sup>-binding and induced structure formation.....</b>	<b>25</b>

1.4.2	Ca <sup>2+</sup> binding and conformational change.....	26
1.5	Calcium Binding Proteins of <i>Entamoeba histolytica</i> .....	27
1.6	Target binding mode of Calmodulin: A prototypic Calcium Binding Protein.....	30
1.7	Metals in Proteins and plasticity of calcium binding loops.....	32
<b>AIMS AND OBJECTIVES</b>		<b>34</b>
<b>MATERIALS AND METHODS</b>		<b>35-55</b>
2.1	Sources of materials.....	35
2.2	Organisms and growth conditions.....	35
2.3	Culture media.....	35
2.3.1	Luria Broth.....	35
2.3.2	LB Agar.....	36
2.4	Plasmid DNA isolation.....	36
2.4.1	Boiling Method (mini-prep).....	36
2.4.2	Alkaline lysis method (midi-prep).....	37
2.5	Transformation of <i>E. coli</i> cells.....	37
2.5.1	Preparation of competent cells.....	37
2.5.2	Transformation of competent cells.....	38
2.6	Clones of full-length EhCaBP1 and N-terminal domains of EhCaBP1.....	38
2.7	Agarose gel electrophoresis.....	38
2.8	Protein purification.....	38
2.8.1	Over-expression and purification of full-length EhCaBP1.....	38
2.8.2	Over-expression and purification of N-terminal domain of EhCaBP1.....	39
2.8.3	Packing and Equilibration of column.....	39

2.8.4	Binding and Elution of EhCaBP1.....	40
2.8.5	Binding and Elution of N-terminal domain of EhCaBP1.....	40
2.9	Desalting and equilibration of protein samples.....	40
2.10	SDS-Polyacrylamide gel electrophoresis.....	40
2.11	Protein concentration.....	41
2.11.1	Centricon concentration method.....	41
2.11.2	Lyophilization.....	41
2.12	Protein estimation.....	41
2.13	Dynamic light scattering.....	42
2.14	Size exclusion chromatography.....	42
2.15	Crystallization of EhCaBP1.....	43
2.16	Mass spectroscopy of EhCaBP1 crystals.....	43
2.17	Data collection and processing structure determination of EhCaBP1 crystals.....	43
2.18	Interhelical Angle determination.....	45
2.19	Crystallization of EhCaBP1-Phe complex.....	45
2.20	Data collection and processing and structure determination of EhCaBP1-Phe complex.....	46
2.21	Mass spectroscopy of N-terminal domain of EhCaBP1.....	47
2.22	Crystallization of N-terminal domain of EhCaBP1.....	48
2.23	Data collection, processing and structure determination of N-terminal domain of EhCaBP1.....	48

2.24	Preparation of Pb <sup>2+</sup> -EhCaBP1 complex, Crystallization and Data collection.....	49
2.25	Preparation of Ba <sup>2+</sup> -EhCaBP1 complex, Crystallization and Data collection.....	51
2.26	Preparation of Sr <sup>2+</sup> -EhCaBP1 complex, Crystallization and Data collection.....	53
2.27	Structure determination of EhCaBP1 –heavy metal complexes.....	54
2.28	Anomalous signal analysis.....	55
2.29	Phasing with Anomalous signal.....	55

## CHAPTER

I.	<b>Crystal structure of Calcium Binding Protein-1 from <i>Entamoeba histolytica</i></b>	<b>56-78</b>
3.1	Abstract.....	56
3.2	Introduction.....	57
3.3	Results and Discussion.....	62
3.3.1	EhCaBP1 purification and crystallization.....	63
3.3.2	Overall crystal structure of EhCaBP1.....	64
3.3.3	Sequence comparison of EhCaBP1.....	66
3.3.4	Solvent content.....	67
3.3.5	Calcium Binding loops of EhCaBP1.....	67
3.3.6	Mass spectrometry of EhCaBP1 crystals .....	68
3.3.7	Symmetry interactions in domain swapped trimeric form of EhCaBP1.....	69
3.3.8	Trimeric nature of EhCaBP1.....	71
3.3.8.1	Size-exclusion chromatography.....	71
3.3.8.2	Dynamic Light Scattering.....	72

3.3.9	Hydrophobic Core.....	73
3.3.10	Comparison with calmodulin, Troponin C, and NMR structure.....	73
3.4	Conclusion.....	76
II.	<b>Crystal structure of EhCaBP1 in complex with Phenylalanine: Visualization of mode of target binding</b>	<b>79-85</b>
4.1	Abstract.....	79
4.2	Introduction.....	80
4.2.1	Calmodulin: Target recognition and activation mechanism.....	80
4.3	Results and Discussion.....	82
4.3.1	Structural analysis of EhCaBP1- Phenylalanine complex.....	82
4.4	Conclusion.....	84
III.	<b>Crystal structure of N-terminal domain of EhCaBP1</b>	<b>86-97</b>
5.1	Abstract.....	86
5.2	Introduction.....	87
5.2.1	Schematic representation of independent domains of EhCaBP1.....	87
5.2.2	Functional properties of N-terminal and C-terminal domains.....	88
5.3	Results and Discussion.....	89
5.3.1	Purification of N-terminal domain.....	89
5.3.2	Mass spectroscopy of N-terminal domain.....	90
5.3.3	Crystal structure of N-terminal domain of EhCaBP1.....	91

5.3.4	Overall structure of N-terminal domain.....	92
5.3.5	Calcium-binding loop of N-terminal domain.....	93
5.3.6	Symmetry-related trimeric structure of N-terminal domain.....	94
5.4	Conclusion.....	95

**IV. Flexibility and plasticity of EF-hand motifs: Structure of Calcium Binding Protein-1 from *Entamoeba histolytica* in complex with  $Pb^{2+}$ ,  $Ba^{2+}$ , and  $Sr^{2+}$  98-113**

6.1	Abstract.....	98
6.2	Introduction.....	99
6.3	Results and Discussion.....	104
6.3.1	$Pb^{2+}$ -EhCaBP1 complex.....	104
6.3.2	Overall structure of $Pb^{2+}$ -EhCaBP1 complex.....	105
6.3.3	$Ba^{2+}$ -EhCaBP1 complex.....	106
6.3.4	Overall structure of $Ba^{2+}$ -EhCaBP1 complex.....	107
6.3.5	$Sr^{2+}$ -EhCaBP1 complex.....	109
6.3.6	Overall structure of $Sr^{2+}$ -EhCaBP1 complex.....	109
6.3.7	Anomalous signal analysis and verification of heavy atom binding.....	111
6.3.8	Phasing with anomalous signal.....	111
6.4	Conclusion.....	111

<b>SUMMARY</b>	<b>114-118</b>
<b>REFERENCES</b>	<b>119-155</b>
<b>APPENDIX</b>	<b>156-161</b>
LIST OF TABLES.....	156
LIST OF FIGURES.....	157
LIST OF PROTEIN SEQUENCES.....	160
LIST OF PUBLICATIONS.....	161

# INTRODUCTION

---

## 1.1 *Entamoeba histolytica*: A pathogenic enteric protozoan

The enteric protozoan parasite *Entamoeba histolytica* is the causative agent of amoebiasis, an endemic disease in developing countries causing morbidity and mortality among large number of individuals (WHO/PAHO/UNESCO report, 1997; Stanley, 2003). It belongs to Phylum- Sarcomastigophora, Class- Labosea and Order- Amoebida. Unlike many protozoan parasites, *E. histolytica* has a simple life-cycle, existing as either the infectious cyst form or the amoeboid trophozoite stage. Human beings and some non-human primates are the only natural hosts. *E. histolytica* cysts are round, four nucleated usually 10-15µm in diameter and are surrounded by a refractile wall made of chitin (Clark, 2000). Trophozoites are highly motile with a pleuromorphic shape, 10-50 µm in diameter and mostly uninucleated. Typical mitochondria, Golgi apparatus, rough endoplasmic reticulum, centrioles and microtubules are absent. Nucleus is spherical, 4-7 µm in diameter and covered by a bilayered nuclear envelope. Chromatin clumps are usually uniform in size and evenly distributed inside the nuclear membrane (Clark, 2000).

### 1.1.1 Historical background

Amoebic colitis and amoebic liver abscess were known to the ancients; Dysentery (Greek dys-alteration, enteron-bowl) was described by Celso and Hippocrates (460 to 377 AC) as "belly flow" and later by Huang Ti (140-87 AC) as dysentery. Lambal (1850) showed the presence of protozoan in the fecal matter of a dysentery patient. However, more than 2000 years would elapse before amoebas were identified as a cause of dysentery when, in 1875, the St. Petersburg physician Fedor Aleksandrovich Löschi described amoebic trophozoites in the stool and colonic ulcerations of a farmer with a fatal case of dysentery and named it *Amoeba coli* (Losch, 1875). Koch (1886) in Egypt demonstrated the presence of the parasite in liver injury. In 20th century, this parasite received different names as *Amoeba coli*, *Amoeba dysenteria*, *Entamoeba dysenteria* etc. In 1901, William Councilman and H. A. Lafleur described the pathology of amoebiasis and introduced the terms "amoebic dysentery" and "amoebic abscess of the liver". Later in 1903, Fritz Schaudinn documented the taxonomic description and named as *Entamoeba histolytica* (Schaudinn, 1903). In 1919, Clifford classified this organism into several species on the basis of number of nuclei present

in the cyst. In 1925, Emile Brumpt proposed that there were actually two species of *Entamoeba*, of which one was infective to human whereas the other was not. The former was *Entamoeba histolytica* and the later he called *Entamoeba dispar* (Brumpt, 1925). However, as the two species were morphologically indistinguishable, Brumpt's hypothesis was ignored for 50 years till the biochemical evidence in favor of subgroups within *E. histolytica* was reported in 1973, based on differences in lectin agglutination properties of parasites isolated from amebic patients and asymptomatic individuals (Martinez-Palomo *et al.*, 1973). Over the next fifteen years evidence for the existence of two groups continued to accumulate and came to include antigenic differences (Strachan *et al.*, 1988) and DNA evidence (sequencing of ribosomal and other highly conserved genes) (Burch, Clark) (Tannich *et al.*, 1989) in addition to isoenzymes. Finally in 1993, *E. histolytica* was formally redescribed to separate it from *E. dispar*, (Diamond *et al.*, 1993; Clark, 2000), validating Brumpt's hypothesis after almost seventy years.

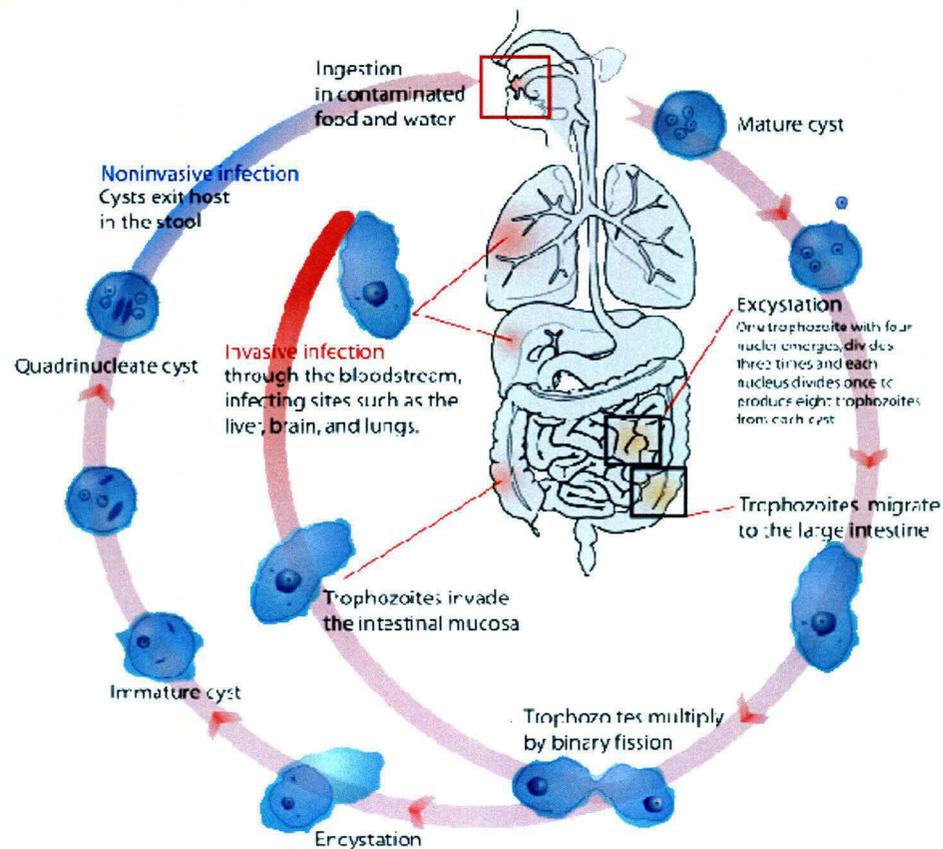
### 1.1.2 Epidemiology

*E. histolytica* is distributed throughout the world and is a substantial health risk in almost all countries where the barriers between human faeces and food and water are inadequate. It is the third leading cause of morbidity and mortality due to parasitic disease in humans after malaria and schistosomiasis (Anand, 1997) and is estimated to be responsible for between 50,000 and 100,000 deaths worldwide every year (Stanley, 2003). It is a serious health hazard, particularly in developing countries. Many of the estimated 500 million individuals infected with *Entamoeba* are colonised by *E. dispar*; this organism is not a pathogen, and causes no signs of disease or mucosal invasion even in patients with AIDS (Allanson-Jones *et al.*, 1986; Reeds *et al.*, 1991). *E. histolytica* is the causative agent of amoebic colitis and all forms of extraintestinal amoebiasis, but is also present in many asymptomatic individuals. Data from epidemiological surveys with techniques that can differentiate between *E. dispar* and *E. histolytica* in stool samples suggest that most asymptomatic individuals infected with *Entamoeba* are colonized with *E. dispar*; however, in some regions, the prevalence of *E. histolytica* both in asymptomatic individuals as well as in patients with diarrhoea is very high (Adb-Alla *et al.*, 2002).

Amoebic colitis has generally been regarded as an equal-opportunity disease affecting children and adults of both sexes although, this notion has been challenged by data indicating a significant male predominance (Acuna-Soto *et al.*, 2000). However, amoebic liver-abscess mainly affects men between the ages of 18 and 50, in whom rates are 3-20 times higher than females (Acuna-Soto *et al.*, 2000; Thompson *et al.*, 1985; Abuabara *et al.*, 1982; Shandera *et al.*, 1998; Barnes *et al.*, 1987; DeBakey *et al.*, 1951; Katzenstein *et al.*, 1982). The parasite usually thrives in the intestinal lumen but during the invasive form of the disease it penetrates the intestinal wall and enters the blood stream from where it goes to vital organs, mainly the liver and sometimes the brain, causing abscesses, which are fatal if not diagnosed, timely.

### 1.1.3 Life cycle of the parasite

The dimorphic parasite has two stages in its life cycle, infective, non-motile and dormant stage known as cyst and invasive, motile and dividing stage known as trophozoite. Infection starts with ingestion of food and water contaminated with fecal material (Fig.1.1). The cysts get excysted in the ileo-cecal region and give rise to eight trophozoites each. The trophozoites migrate to the colon and continue to divide further and colonize the host's tissues, causing dysentery. The process of excystation has been studied both *in vitro* (Marinets *et al.*, 1997) and in monkey with almost identical results (Espinosa-Cantellano *et al.*, 1991). First the amoeba moves within the cyst wall, later at one point thinning of wall occurs through which amoeba emerges, usually after numerous extension and retraction of the pseudopodia leaving behind an empty cyst wall. The multi-nucleated trophozoite undergoes cytokinesis and nuclear division to give rise to eight daughter amoebae (Marinets *et al.*, 1997). The trophozoites multiply by binary fission and re-encyst and are finally passed through the faeces, thus completing the life cycle. Although cysts may remain viable in a humid environment and stay infective for several days, trophozoites are short lived outside the body and do not survive during the passage through the upper gastrointestinal tract. Tissue invasion is not a part of the life cycle of the organism as those



**Figure-1.1** Life cycle of *Entamoeba histolytica*. (Adapted from the website [http://commons.wikimedia.org/wiki/File:Entamoeba\\_histolytica\\_life\\_cycle-en.svg](http://commons.wikimedia.org/wiki/File:Entamoeba_histolytica_life_cycle-en.svg)).

organisms that pass through the mucosa are no longer capable of giving rise to new infections. The invasive disease is therefore viewed as an aberrant behaviour on the part of the organism.

#### 1.1.4 Pathogenesis

The pathogenesis of infection by *E. histolytica* is mostly governed at three levels-

1. Adherence of trophozoite to the target cell
2. Lysis of target cell and
3. Phagocytosis of target cell

Disease begins after adherence of trophozoites to colonic epithelial cells, probably through the galactose/N-acetylglucosamine specific lectin (Gal/GalNAc), a complex of a disulphide-linked 170 kDa and 35/31 kDa subunits and an associated 150 kDa protein (Mann *et al.*, 2002). Mammalian cells without N-terminal galactose or N-acetylgalactosamine residues are resistant to adherence by amoebic trophozoites and to killing by *E. histolytica*, consistent with a physiological role for the lectin adhesion and the known requirement for cell-cell contact in killing process (Ravdin *et al.*, 1989). Analysis of the genome reveals redundancy in the genes encoding the subunits of Gal/GalNAc lectin.

The cytolytic capabilities of *E. histolytica* were first documented decades ago. Human cells touched by amoebic trophozoites become immobile within minutes and lose their cytoplasmic granules, structures and eventually their nucleus (Ravdin *et al.*, 1980). A family of pore forming peptides named amoebapores are attributed to the cytolytic effect and are structurally and functionally related to granulolysins and (natural killer) NK-lysins produced by mammalian T cells (Leippe, 1997). In addition to three amoebapores, a homologue of haemolysin III was identified in recent genome sequencing and suggested haemolysins may have a role in host cell lysis in addition to amoebapores (Loftus *et al.*, 2005).

*E. histolytica* trophozoites can also kill mammalian cells by induction of programmed cell death (apoptosis), determined by the host cell DNA ladder formation and over-expression of Bcl-2 (Ragland *et al.*, 1994). Soluble amoebic proteins activate the heat shock transcription factor-1 to form trimers from its inactive monomers. Activated HSF-1 induces the expression of Hsp27 and Hsp72, which perform different functions. Phospho-Hsp27 associates with IKK (Inhibitor of NF- $\kappa$ B-kinase) and inhibits its activity, thereby suppressing NF- $\kappa$ B activation induced by the pro-inflammatory cytokines released in response to amoebic infection. Suppression of NF- $\kappa$ B favors increased apoptosis. However, Hsp72 that apparently does not play a role in NF- $\kappa$ B signaling in intestinal epithelial cells (IEC) might function to promote cell survival through its potent anti-apoptotic activity. Together, the stress response induced by amoebic proteins functions to inhibit intestinal inflammation and promote cell survival (Kammanadiminti *et al.*, 2006).

Amoebic invasion through the mucosa and into the sub-mucosal tissue is the hallmark of amoebic colitis. Cysteine proteinases are a key virulence factor of *E. histolytica* and play a role in intestinal invasion by degrading the extra cellular matrix and circumventing the host immune response through cleavage of secretory immunoglobulin A (sIgA), IgG, and activation of complement. So far, seven distinct genes encoding pre-pro forms of papain family proteases have been identified in *E. histolytica* out of which many are absent in the nonpathogenic sibling *E. dispar* (Que *et al.*, 2000). Proteases are secreted by *E. histolytica* trophozoites and large quantity of cysteine proteases are seen extracellularly in amoebic liver abscess in animals. Genetically engineered, protease deficient *E. histolytica* trophozoites have reduced virulence in a model of amoebic liver abscess in rodents (Ankri *et al.*, 1999).

### 1.1.5 Metabolism

The metabolism of *E. histolytica* seems to have been shaped by secondary gene loss and lateral gene transfer (LGT), primarily from bacterial lineages (Loftus *et al.*, 2005). *E. histolytica* is an obligate fermenter, using bacterial-like fermentation enzymes and lacking proteins of the tricarboxylic acid cycle and mitochondrial electron transport chain. An atrophic, mitochondrion-derived organelle has been identified in *E. histolytica* (Leon-Avila *et al.*, 2004), and the genome data support the absence of a mitochondrial genome. Glucose is the main energy source; however, in place of the typical eukaryotic glucose transporters, those of *E. histolytica* are related to the prokaryote glucose/ribose transporter family, with the amino- and carboxy-terminal domains switched relative to their prokaryotic counterparts.

Most pathways for amino acid biosynthesis have been eliminated, except those for serine and cysteine, which are probably retained for the production of cysteine, the major intracellular thiol. The high levels of cysteine in *E. histolytica* may compensate for the lack of glutathione and its associated enzymes, a major component of oxidative stress resistance in many organisms (Fahey *et al.*, 1984). *E. histolytica* lacks de novo purine, pyrimidine and thymidylate synthesis and must rely on salvage pathways, similar to *G. lamblia* and *T. vaginalis* (Abrahamsen *et al.*, 2004). In addition, *E. histolytica* appears to lack ribonucleotide reductase, a characteristic that it shares with *G. lamblia* (Baum *et al.*, 1989). *E. histolytica* is unable to synthesize fatty acids but retains the ability to synthesize a variety of phospholipids. The absence of

identifiable pathways for the synthesis of isoprenoids and the sphingolipid head group aminoethylphosphonate suggests the existence of novel pathways. These pathways, once characterized, might represent attractive drug targets. Two unusual enzymes of fatty acid elongation are shared between *E. histolytica* and *G. lamblia*, including a predicted acetyl-CoA carboxylase with two carboxyltransferase domains (Jordan *et al.*, 2003). This enzyme removes a carboxyl group from oxaloacetate and transfers it to acetyl-CoA to form malonyl-CoA and pyruvate. *E. histolytica* also has five members of a fatty acid elongase family, previously identified only in plants, green algae and *G. lamblia* (Lawrence, 2003; Qvarnstrom *et al.*, 2005). Folate is a cofactor essential for thymidylate synthesis, methionine recycling and is also required for organelle protein synthesis in mitochondria and chloroplast in plants. Loss of mitochondrial genome may have paved for the loss of folate dependent functions.

#### **1.1.6 Genome organization of *E. histolytica***

The genome of *E. histolytica* is 24 Mb in size, encodes around 9938 genes of average size 1.17 Kb, which comprise about 49% of the total genome. Only 25% of genes contain introns and only 6% contain multiple introns (Loftus *et al.*, 2005). The genome is highly A+T rich (67% in the coding region and 72% in the intergenic region) and it is estimated that each trophozoite contains about 0.24 picogram of DNA (Dvorak *et al.*, 1995). From the total predicted genes 31.8 % of proteins do not have any homologue. The *E. histolytica* chromosomes do not condense during cell cycle. Also the lengths of the homologous chromosome are variable. From Pulsed Field Gel Electrophoresis (PFGE) it is established that the genome of strain HM-1: IMSS has 31-35 chromosomal bands of sizes ranging from 300Kb to 2200Kb. The size distribution however changes with the isolation conditions and separation conditions (Dhar *et al.*, 1996; Bagchi *et al.*, 1999; Willhoeft *et al.*, 1999). The chromosome size variation observed may be due to expansion and contraction of subtelomeric repeats, as in other protists and it has been suggested that in *E. histolytica* these regions may consist of tRNA-containing arrays. Molecular analysis showed that there are 14 linkage groups (Willhoeft *et al.*, 1999).

A number of extra-chromosomal circles of different size varying from 5kb to 50kb are present in the genome (Dhar *et al.*, 1996). Among them, the 24.5 Kb EhR1 circle encoding the rRNA genes is the most abundant with around 200 copies per

haploid genome (Mittal *et al.*, 1994). The organization and role of the other circular molecules is not clear.

Multicopy protein coding gene families are also abundant in *E. histolytica* genome. These include virulence factor genes like Gal/GalNAc lectin intermediate subunit (30 copies), Cysteine proteinase family (at least 20 different members), the protein kinase family (271 members across seven super families), calcium binding proteins (CaBPs) and phosphatases (about 100 members) (Loftus *et al.*, 2005).

## 1.2 $\text{Ca}^{2+}$ as second messenger

Protein function is governed by shape and charge.  $\text{Ca}^{2+}$  binding triggers changes in protein shape and charge. The abilities of  $\text{Ca}^{2+}$  and phosphate ions to alter local electrostatic fields and protein conformations are the two universal tools of signal transduction. The calcium ion is used as a major signaling molecule in a diverse range of eukaryotic cells including several human parasitic protozoa, such as *Trypanosoma cruzi*, *Trypanosoma brucei*, *Leishmania sp.*, *Plasmodium sp.*, *Toxoplasma gondii*, *Cryptosporidium parvum*, *Entamoeba histolytica*, *Giardia lamblia* and *Trichomonas vaginalis*.  $\text{Ca}^{2+}$  is critical for invasion of intracellular parasites, and its cytosolic concentration is regulated by the concerted operation of several transporters present in the plasma membrane, endoplasmic reticulum, mitochondria and acidocalcisomes. The ubiquitous second messenger  $\text{Ca}^{2+}$  is responsible for regulating a wide range of cellular processes (Clapham *et al.*, 1995). Intracellular calcium plays a crucial role as a second messenger for the control of a variety of cell functions in eukaryotes, including muscle contraction, secretion, exocytosis, energy metabolism, cell division and differentiation, sodium and potassium permeability, chemotaxis and synaptic plasticity during learning and memory.  $\text{Ca}^{2+}$  is an unlikely candidate to perform this role of a universal messenger because prolonged elevations of  $[\text{Ca}^{2+}]$  result in irreversible damage as occurs during cardiac or cerebral ischaemia (Trump *et al.*, 1995). Because of its cytotoxicity, the intracellular level of  $\text{Ca}^{2+}$  in resting cells is normally held within a narrow range of 20–100 nmol  $\text{l}^{-1}$ . The signaling functions of  $\text{Ca}^{2+}$  have to be performed against this background of a tightly controlled  $\text{Ca}^{2+}$  homeostasis. Another consequence of this rigid homeostatic control over  $\text{Ca}^{2+}$  is that this messenger has a very low diffusibility in cytoplasm. An extensive array of  $\text{Ca}^{2+}$  pumps are distributed throughout the

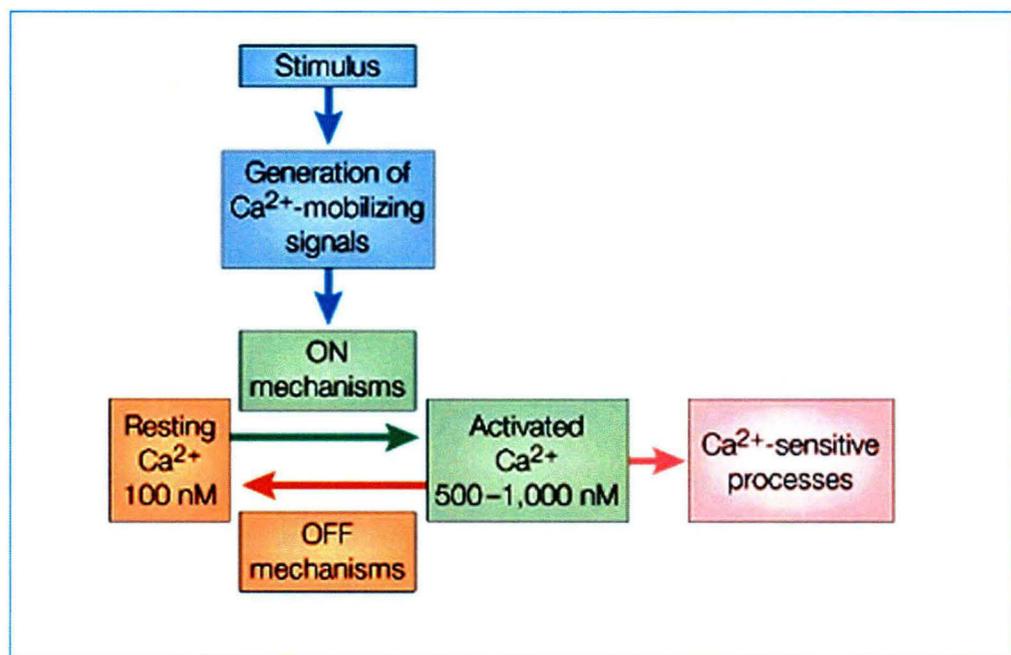
cytoplasm (Carafoli *et al.*, 1994) which rapidly sequester  $\text{Ca}^{2+}$ , thus restricting its diffusion. In order to overcome the twin problems of an inherent cytotoxicity and low diffusibility, cells have evolved an ingenious mechanism of signaling based on presenting  $\text{Ca}^{2+}$  as brief spikes often organized as regenerative waves (Cheek *et al.*, 1991; Berridge *et al.*, 1993; Clapham *et al.*, 1995).  $\text{Ca}^{2+}$  enters from the outside through a variety of channels such as the voltage-operated channels (VOCs), receptor-operated channels (ROCs) or store-operated channels (SOCs). Also, it can be released from internal stores (Fig.1.2). Utilization of these sources varies from cell to cell. In most cells, it is the internal stores which provide most of the signal  $\text{Ca}^{2+}$ , so attention has focused on the intracellular  $\text{Ca}^{2+}$  channels, which are of two main types (Berridge *et al.*, 1993; Clapham *et al.*, 1995). First, there is the ryanodine receptor (RYR) family comprising three members: RYR1 found in skeletal muscle and certain neurons (e.g. Purkinje cells), RYR2 found in cardiac muscle, brain and some other cells, and RYR3 found in smooth muscle, brain and other cells (Bennett *et al.*, 1996; Giannini *et al.*, 1995). Second, the inositol 1,4,5-trisphosphate receptor (InsP3R) family has a number of members (Furuichi *et al.*, 1995; Taylor *et al.*, 1995; Bezprozvanny *et al.*, 1995). There are four InsP3R genes, and further diversity results from alternative splicing. These two receptor families must have evolved from a common ancestor since they display considerable sequence homology which is matched by a number of physiological similarities, particularly with regard to the control of channel opening (Taylor *et al.*, 1995). Cytosolic  $\text{Ca}^{2+}$  homeostasis in resting cells is achieved by balancing the leak of  $\text{Ca}^{2+}$  (entering from the outside or from the stores) by the constant removal of  $\text{Ca}^{2+}$  using pumps either on the plasma membrane or on the internal stores (Fig.1.2). These pumps ensure that cytoplasmic  $[\text{Ca}^{2+}]$  remains low and that the stores are loaded with signal  $\text{Ca}^{2+}$ . The brief burst of  $\text{Ca}^{2+}$  responsible for cell activation is usually produced by the coordinated opening of either the RYRs or the InsP3Rs. Perhaps their most important property is their sensitivity to  $\text{Ca}^{2+}$  *i.e.*, they display the phenomenon of  $\text{Ca}^{2+}$ -induced  $\text{Ca}^{2+}$  release (CICR) which is of major significance for the generation of complex signals.  $\text{Ca}^{2+}$  has a biphasic effect on the RYRs and InsP3Rs: as its concentration is increased, it initially exerts a positive feedback effect by enhancing the opening of the channels (*i.e.*, CICR), but as soon as the concentration reaches a certain level the feedback switches from positive to negative and  $\text{Ca}^{2+}$  then inhibits the channel (Bezprozvanny *et al.*, 1995). This negative feedback effect ensures that just enough  $\text{Ca}^{2+}$  is released to give a meaningful signal,

thus avoiding the cytoplasm from being swamped with this potentially cytotoxic agent. The fact that  $\text{Ca}^{2+}$  release is regenerative has important implications for signaling because it provides one of the mechanisms for coordinating the activity of individual receptors *i.e.*, they can communicate with each other using  $\text{Ca}^{2+}$  as a messenger (Bootman *et al.*, 1995). A specific region within the cell usually functions as an initiation site where the release of calcium takes place, which then diffuses outwards to excite neighboring receptors, thereby setting up a  $\text{Ca}^{2+}$  wave. A global  $\text{Ca}^{2+}$  signal is created by coordinating release from all the receptors using  $\text{Ca}^{2+}$  as the messenger. A more specialized mechanism of coordination is found in skeletal and cardiac muscle, where the opening of the RYRs is tightly coupled to the action potential sweeping over the plasma membrane (Cannell *et al.*, 1995; Lopez-Lopez *et al.*, 1995). Using a regenerative process is inherently dangerous because it is liable to be triggered by the stochastic opening of a single channel. To avoid such random triggering of regenerative  $\text{Ca}^{2+}$  waves, cells have developed mechanisms for regulating the excitability of these intracellular receptors such that they are turned off in resting cells but become increasingly excitable when  $\text{Ca}^{2+}$  signals are being generated. In the case of the InsP3Rs, excitability is regulated by the agonist-dependent generation of InsP3 by cell surface receptors. This InsP3 binds to the InsP3Rs, greatly enhancing their sensitivity to the stimulatory action of  $\text{Ca}^{2+}$ . In effect, the InsP3R is under the dual regulation of two agonists – InsP3 and  $\text{Ca}^{2+}$ . The primary function of the former is to increase the  $\text{Ca}^{2+}$  sensitivity of the InsP3R. Similarly, the RYR may also be under dual regulation, at least in some cell types (Lee *et al.*, 1994; Galione *et al.*, 1994). The putative second messenger cyclic ADP ribose (cADPR) is able to enhance the  $\text{Ca}^{2+}$  sensitivity of the RYRs. In summary, through the ability of InsP3 or cADPR to enhance the sensitivities of the InsP3Rs and RYRs respectively, these messengers convert the quiescent cytoplasm into an excitable medium in which these intracellular channels can communicate with each other to generate global  $\text{Ca}^{2+}$  signals. Cells invest much of their energy to effect changes in  $\text{Ca}^{2+}$  concentration. Underlying the speed and effectiveness of  $\text{Ca}^{2+}$  is the 20,000-fold gradient maintained by cells between their intracellular (~100 nM free) and extracellular (mM) concentrations. In contrast, the concentration of  $\text{Ca}^{2+}$ 's cousin,  $\text{Mg}^{2+}$ , barely differs across the plasma membrane. This versatility is exploited to regulate diverse cellular responses.

### 1.2.1 The $\text{Ca}^{2+}$ signaling toolkit

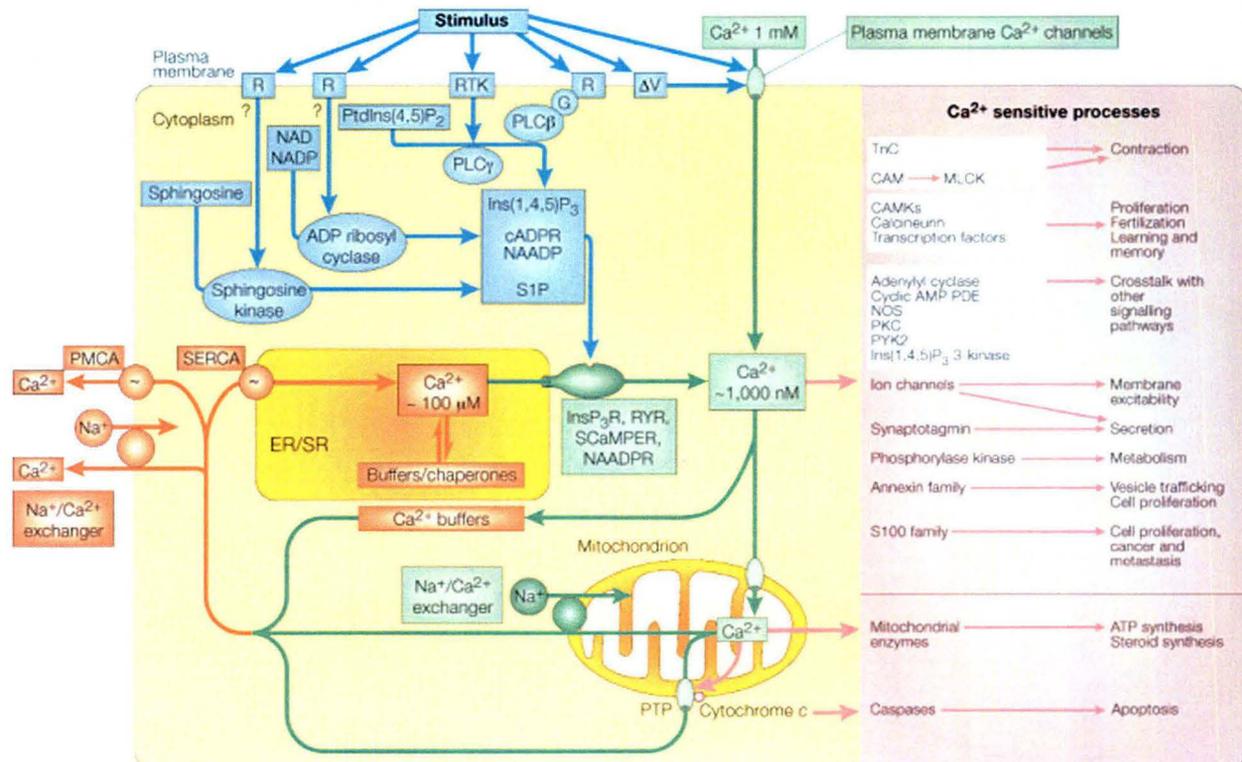
The  $\text{Ca}^{2+}$  signaling network can be divided into four functional units (Fig.1.2):-

- Signaling is triggered by a stimulus that generates various  $\text{Ca}^{2+}$ -mobilizing signals.
- The latter activate the ON mechanisms that feed  $\text{Ca}^{2+}$  into the cytoplasm.
- $\text{Ca}^{2+}$  functions as a messenger to stimulate numerous  $\text{Ca}^{2+}$ -sensitive processes.
- Finally, the OFF mechanisms, composed of pumps and exchangers, remove  $\text{Ca}^{2+}$  from the cytoplasm to restore the resting state.



**Figure-1.2** *The four units of the  $\text{Ca}^{2+}$  signaling network.* Stimuli act by generating  $\text{Ca}^{2+}$ -mobilizing signals that act on various ON mechanisms to trigger an increase in the intracellular concentration of  $\text{Ca}^{2+}$ . The increased level of  $\text{Ca}^{2+}$  stimulates various  $\text{Ca}^{2+}$ -sensitive processes to trigger many different cellular pathways. The response is terminated by OFF mechanisms that restore  $\text{Ca}^{2+}$  to its resting level (Adapted from Berridge et al., 2000).

The functional relationship between these units is illustrated in Fig.1.3, which reveals that the signaling network is composed of many components (the  $\text{Ca}^{2+}$  signaling toolkit). Each specific cell type can exploit this large repertoire to construct versatile  $\text{Ca}^{2+}$  signaling networks.



**Figure-1.3. Elements of the  $\text{Ca}^{2+}$  signaling toolkit.** Cells have an extensive signaling toolkit that can be mixed and matched to create  $\text{Ca}^{2+}$  signals of widely different properties.  $\text{Ca}^{2+}$ -mobilizing signals (blue) are generated by stimuli acting through a variety of cell-surface receptors ( $R$ ), including G-protein ( $G$ )-linked receptors and receptor tyrosine kinases ( $RTK$ ). The signals generated include:  $\text{Ins}(1,4,5)\text{P}_3$ , generated by the hydrolysis of phosphatidylinositol-4,5-bisphosphate ( $\text{PtdIns}(4,5)\text{P}_2$ ) by a family of phospholipase C enzymes ( $\text{PLC}\beta$ ,  $\text{PLC}\gamma$ ); cyclic ADP ribose ( $\text{cADPR}$ ) and nicotinic acid dinucleotide phosphate ( $\text{NAADP}$ ), both generated from nicotinamide-adenine dinucleotide ( $\text{NAD}$ ) and its phosphorylated derivative  $\text{NADP}$  by  $\text{ADP ribosyl cyclase}$ ; and sphingosine 1-phosphate ( $\text{SIP}$ ), generated from sphingosine by a sphingosine kinase. ON mechanisms (green) include plasma membrane  $\text{Ca}^{2+}$  channels, which respond to transmitters or to membrane depolarization ( $\Delta V$ ), and intracellular  $\text{Ca}^{2+}$  channels - the  $\text{Ins}(1,4,5)\text{P}_3$  receptor

(*InsP3R*), ryanodine receptor (*RYR*), NAADP receptor and sphingolipid  $\text{Ca}^{2+}$  release-mediating protein of the ER (*SCaMPER*). The  $\text{Ca}^{2+}$  released into the cytoplasm by these ON mechanisms activates different  $\text{Ca}^{2+}$  sensors (purple), which augment a wide range of  $\text{Ca}^{2+}$ -sensitive processes (purple), depending on cell type and context. OFF mechanisms (red) pump  $\text{Ca}^{2+}$  out of the cytoplasm: the  $\text{Na}^+/\text{Ca}^{2+}$  exchanger and the plasma membrane  $\text{Ca}^{2+}$  ATPase (*PMCA*) pumps  $\text{Ca}^{2+}$  out of the cell and the sarco-endoplasmic reticulum  $\text{Ca}^{2+}$  ATPase (*SERCA*) pumps it back into the ER/SR. (*TnC*, troponin C; *CAM*, calmodulin; *MLCK*, myosin light chain kinase; *CAMK*,  $\text{Ca}^{2+}$ /calmodulin-dependent protein kinase; cyclic AMP PDE, cyclic AMP phosphodiesterase; *NOS*, nitric oxide synthase; *PKC*, protein kinase C; *PYK2*, proline-rich kinase 2; *PTP*, permeability transition pore.) (Adapted from Berridge *et al.*, 2000).

### 1.2.1 Calcium Binding Motifs

Based on cellular location there are two type of calcium binding proteins- extra cellular and intracellular. On the basis of its function the  $\text{Ca}^{2+}$  binding proteins can be divided into three categories: trigger or sensor proteins (e.g., calmodulin, troponin C) (Berridge *et al.*, 2000), buffer proteins (e.g., S100G and parvalbumin) (Schroder *et al.*, 1996, or  $\text{Ca}^{2+}$ -stabilized proteins (e.g., thermolysin) (Buchanan *et al.*, 1986). There are four kinds of calcium binding motifs which have been characterized at the molecular level. They are- Annexin domain (Gerke *et al.*, 2005), C2 domain (Sutton *et al.*, 1998), C-lectin (Weis *et al.*, 1998), and EF-hand domain (Kretsinger *et al.*, 1973). The most common calcium binding motif is the EF-hand motif. It is one of the most frequently used motifs in eukaryotic systems and can be found in each category mentioned above and constitute more than 50% of all well-characterized  $\text{Ca}^{2+}$ -binding proteins.

### 1.3. The EF-Hand: A Ca<sup>2+</sup>-Binding Unit

The Ca<sup>2+</sup> that flows into the cytoplasm during the ‘on’ reaction does not remain free. Instead it becomes bound to a wide variety of CaBPs (Ca<sup>2+</sup>-Binding Proteins), many of which belong to a homologous family that binds this cation using a characteristic helix–loop–helix structural motif termed the ‘EF-hand’ (Fig.1.4) (Kretsinger *et al.*, 1973). This motif was first observed in the crystal structure of parvalbumin, a small Ca<sup>2+</sup>-binding protein isolated from carp muscle (Berridge *et al.*, 2000). The EF-hand motifs were later identified in the amino acid sequence of troponin C (Berridge *et al.*, 2003) the myosin light chains (Grabazek *et al.*, 2006), the ubiquitous calmodulin, (Capozzi *et al.*, 2006; Falke *et al.*, 1994) and in many other Ca<sup>2+</sup>-binding proteins (Linse *et al.*, 1995; Kretsinger *et al.*, 1973).



**Figure-1.4** *A single EF-hand motif from N-terminal domain of calmodulin (PDB code 1EXR) showing the helix-loop-helix structure with bound calcium ion in cyan sphere in the loop region.*

Calcium and the EF-hand Ca<sup>2+</sup>-binding proteins have been recognized as the key players in nearly all aspects of cell functions, starting with a cell’s birth during mitosis and ending with its apoptotic death. (Pidcock *et al.*, 2001; Mcphalen *et al.*, 1991). The

exceptional versatility of the EF-hand motif is clearly reflected in the growing database of three-dimensional structures of EF-hand proteins revealing a great diversity of conformations, domain organization, and structural responses to calcium (Likic *et al.*, 2003; Evanus *et al.*, 2001). EF-hand proteins are categorized into two general classes: the  $\text{Ca}^{2+}$  sensors and the  $\text{Ca}^{2+}$  buffers. The  $\text{Ca}^{2+}$  sensors translate the chemical signal of an increased  $\text{Ca}^{2+}$  concentration into diverse biochemical responses.  $\text{Ca}^{2+}$  binding to the ubiquitous sensor CaM elicits a structural response through which CaM binds its target proteins, in many cases removing enzymatic auto-inhibitory domains, thereby activating these enzymes in a  $\text{Ca}^{2+}$ -dependent manner. In the case of the vision-associated protein recoverin, the binding of  $\text{Ca}^{2+}$  results in the extrusion of a hydrophobic myristoyl group, allowing recoverin to associate with the membrane and subsequently its membrane-bound protein target. The  $\text{Ca}^{2+}$  buffers are a smaller subset of the EF-hand protein family.

Exemplified by calbindin D9K and parvalbumin, these proteins help to modulate the  $\text{Ca}^{2+}$  signal both spatially and temporally as they bind the free  $\text{Ca}^{2+}$  to transmit the signal throughout the cell or to remove the potentially harmful ion from the cytoplasm.

For most protein,  $\text{Ca}^{2+}$ -binding motifs, including the EF-hand, a turn-loop structure provides the bulk of the ligands for the bound cation. Unlike helices and sheets, which can only provide a few appropriate ligands in a given sequence, owing to geometric or spacing constraints, the turn-loop structure is flexible and can readily supply three ligands from a sequence of five amino acids (Pidcock *et al.*, 2001).  $\text{Ca}^{2+}$  is a 'hard' metal ion and likes 'hard' ligands that provide an interaction dominated by ionic forces with little covalency (Falke *et al.*, 1994). As such, oxygen is the coordinating atom of choice and consequently EF-hands, and in particular EF-loops, are rich in the negatively charged amino acids, glutamic acid and aspartic acid (Fig.1.5A). Of the two amino acids,  $\text{Ca}^{2+}$  ligands are more commonly provided by the side chain of aspartic acid residues, a frequency which may reflect the preference of the EF-loop for carboxylate ligands with a less bulky side chain (McPhalen *et al.*, 1991). Consistently, asparagine is present to a greater extent than glutamine. The preferred coordination geometry of the  $\text{Ca}^{2+}$  ion is seven ligands arranged in a pentagonal bipyramidal fashion. This is the geometry of the  $\text{Ca}^{2+}$  ion in water and, in most protein binding

sites, including EF-hands, this geometry is preserved with six or seven of the chelating groups provided by the binding motif (Strynadka *et al.*, 1989).

### 1.3.1 The canonical EF-loop

In the canonical EF-loop, five of  $\text{Ca}^{2+}$ 's seven co-ordinating groups are provided by the nine-residue loop. The remaining two come from a bidentate carboxylate ligand supplied by the side chain of an acidic amino acid located in the exiting helix, three residues removed from the loop's C-terminus. Though this residue is not structurally part of the loop, it is commonly referred to as the EF-loop's twelfth residue. Of the ligands provided by the loop itself, three or four come from the side-chain carboxy groups of negatively charged amino acids and one from a backbone carbonyl group.

Frequently, the coordination sphere of the  $\text{Ca}^{2+}$  ion is completed by a water molecule hydrogen-bonded to one of the side chains of the loop (Fig.1.5B). The chelating residues of the loop are notated in two ways, the first based on linear position and the second on the tertiary geometry imposed by their alignment on the axes of a pentagonal bipyramid: 1(+X), 3(+Y), 5(+Z), 7(-Y), 9(-X), 12(-Z). The Y- and Z-axis pairs align along the vertices of an approximately planar pentagon, whereas the X-axis pair takes up an axial position perpendicular to the Y/Z plane (Strynadka *et al.*, 1989).

In addition to the Ca-O bonds formed through the chelating interactions, the canonical EF-loop also contains an extensive network of hydrogen bonds between chelating and non-chelating residues. These bonds help to stabilize the otherwise unfavourably close proximity of the negatively charged oxygen atoms in the coordination sphere of the bound  $\text{Ca}^{2+}$ . Several of the loop's residues contribute individual and identifiable roles to both the stabilization and fold of this structure. Through numerous intraloop hydrogen bonds, the first residue plays an important part in defining a precise stereochemical arrangement for the loop. The conserved glycine residue found at the sixth position facilitates the unusual main-chain conformation ( $\phi$  and  $\psi \sim 60^\circ$  and  $20^\circ$  respectively) which results in a  $90^\circ$  turn that enables the remaining  $\text{Ca}^{2+}$  ligands to take up co-ordinating positions. The eighth residue of the loop is a highly conserved hydrophobic residue (Godzik *et al.*, 1989), the main chain NH and CO groups of which face away from the  $\text{Ca}^{2+}$ -binding site towards the loop of the paired EF-hand. This forms the short anti-parallel  $\beta$ -sheet with the corresponding

groups of the paired loop's eighth position. As well as chelating the bound  $\text{Ca}^{2+}$  ion either directly or indirectly via a bridging water molecule, the hydrogen-bonding pattern formed by the side chain of loop position 9 initiates the exiting helix, a conformation stabilized through additional hydrogen bonds with the side-chain carboxylate group of position 12. Finally, the bidentate side-chain ligands supplied by the twelfth loop position are critical to both the structure and function of the EF-loop.

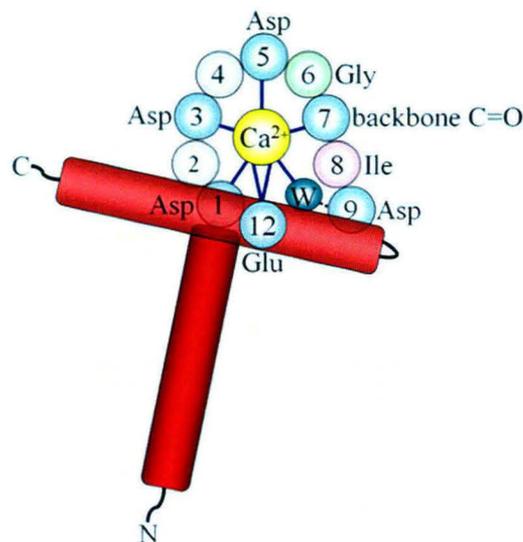
As the above discussion reflects EF-loop structural information predominantly obtained through X-ray crystal structures, it represents a static view. MD (molecular dynamics) simulations corroborate many of these points, but they have shed light on the transitory nature of the  $\text{Ca}^{2+}$  ligands. Although EF1 (EF-hand 1) remained essentially as observed in the crystal structure, in the course of simulations the bound  $\text{Ca}^{2+}$  ion in EF2 of CaM had a tendency to gain an extra water molecule as a ligand, the appearance of which coincided with a decreased contribution from the main-chain carbonyl group of the  $-Y$  ligand (Likic *et al.*, 2003). This ligand exchange results in two water molecules in the coordination sphere, a liganding scheme that increases the flexibility of the loop and perhaps the  $\text{Ca}^{2+}$  off-rate (Evanus *et al.*, 2001).

A second short equilibrium simulation on  $\text{Ca}^{2+}$ -bound EF3 of the C-terminal domain of CaM found that the coordination number of the  $\text{Ca}^{2+}$  ion fluctuated between seven and eight as the Asp95 (+Y ligand) flipped between monodentate and bidentate (Kobayashi, 2006). As this same phenomenon was observed in EF2 of calbindin D9K (Marchand *et al.*, 1998), it appears that the current van der Waals parameters of  $\text{Ca}^{2+}$  that are used in the simulations, coordination numbers of 7 and 8, are in a quasi-equilibrium.

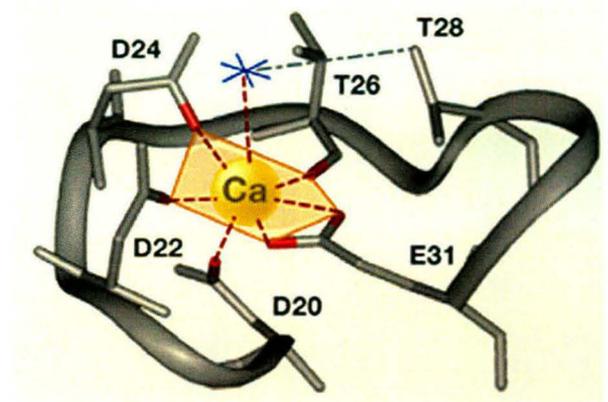
A)

EF-loop position	1	2	3	4	5	6	7	8	9	10	11	12
coordinating ligand	X sc		Y sc		Z sc		-Y bb		-X sc*			-Z sc2
most common	Asp 100%	Lys 29%	Asp 76%	Gly 56%	Asp 52%	Gly 96%	Thr 23%	Ile 68%	Asp 32%	Phe 23%	Glu 29%	Glu 92%
also frequently observed		Ala Gln Thr Val Ile Ser Glu Arg	Asn	Lys Arg Asn	Ser Asn		Phe Lys Gln Tyr Glu Arg	Val Leu	Ser Thr Glu Asn Gly Gln	Tyr Ala Thr Leu Glu Lys	Asp Lys Ala Pro Asn	Asp

B)



C)



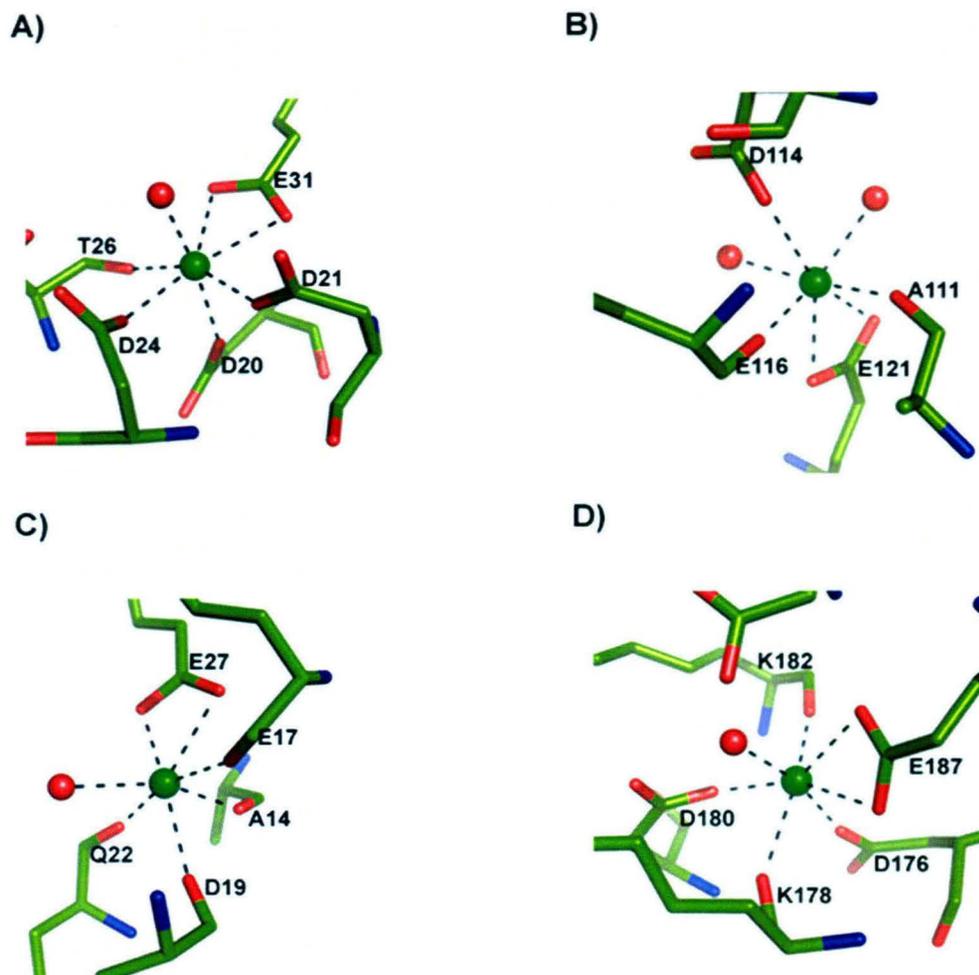
**Figure-1.5 The canonical EF-hand loop.** (A) The sequence preference of the EF-hand loop. The  $\text{Ca}^{2+}$  ligands are indicated by both their position in the EF-loop and in the co-ordinating array with whether or not coordination occurs via the side chain (sc) or through the backbone (bb) indicated below. The asterisk (\*) highlights the ligand typically provided by a water molecule that is hydrogen-bonded to the side chain of the amino acid found at position 9. Also noted in the Figure are the most common amino acids at each position, with their corresponding percentages of occurrence, and those that occur with a frequency greater than 5% in known EF-loops (Falke et al., 1994). (B) A schematic diagram of the  $\text{Ca}^{2+}$  coordination sphere with the entering and exiting helices in red, the co-ordinating protein ligands in blue and co-ordinating water molecule (W) in teal (dark blue). Light green corresponds to

*the conserved glycine residue that provides the bend in the loop. Purple highlights the conserved hydrophobic residue that forms the short  $\beta$ -sheet in the paired EF-hand. Also indicated are the most common amino acids found at the critical positions. (C)  $\text{Ca}^{2+}$  is coordinated by seven oxygen atoms (five in the plane of the orange pentagon and two perpendicular to the plane, thus forming a bipyramidal pentagon).*

*(Adapted from the website [http://structbio.vanderbilt.edu/chazin/cabp\\_database](http://structbio.vanderbilt.edu/chazin/cabp_database)).*

### 1.3.2 The non-canonical EF-loops

Although the  $\text{Ca}^{2+}$  chelation scheme outlined above is employed by the majority of EF-hands, the composition and length of functional binding loops vary significantly in the 'EF-handome' (Haiech *et al.*, 2004) (Fig.1.6). The different non-canonical EF-loops can be classified into four groups, three of which demonstrate different ways whereby the pentagonal bipyramidal coordination is achieved. The first group contains EF-hands that, although they are the same length as the canonical sequence (12 residues), do not use canonical ligands to bind the  $\text{Ca}^{2+}$  ion (Fig.1.5C). This group contains two types of deviation. In approximately 10% of known EF-hands, an aspartic acid residue instead of a glutamic acid residue occupies loop position 12. These 'Asp12' EF-loops tend to be smaller and more compact than the canonical loop (Cook *et al.*, 1993; Vijay-kumar *et al.*, 1992) a characteristic that has effects on  $\text{Ca}^{2+}$  coordination and  $\text{Mg}^{2+}$  selectivity. As seen in the crystal structure of EF3 from CIB (calcium- and integrin-binding protein), the side chain of Asp12 is too short for both oxygen atoms to co-ordinate  $\text{Ca}^{2+}$ , and although one does, the second is replaced by a bridging water molecule (Gentry *et al.*, 2005). The second deviation in this group are EF-loops that bind  $\text{Ca}^{2+}$  through an increased use of main-chain carbonyl groups such as EF4 of AtCBL2 [*Arabidopsis thaliana* (thale cress) calcineurin B-like protein] (Nagae *et al.*, 2003). This deviation is most likely a response to substitution of otherwise disabling amino acids, since, through a shift in the entering



**Figure-1.6** *The canonical and Non-canonical EF-loops. The loop of CaM EF1 (PDB code 1EXR) indicating the canonical arrangement. (B) EF1 of calpain domain VI is shorter than the canonical sequence and has an additional ligand provided by a backbone carbonyl group as well as an additional water molecule (PDB code 1DVI). (C) EF-loops with insertions: the  $\psi$ -hand of calbindin D9K overcomes a two-residue insertion by turning inside out and using the backbone carbonyl groups for coordination (PDB code 3ICB) (D) EF-loops that have the canonical length, but do not chelate the  $\text{Ca}^{2+}$  ion with the canonical ligands: AtCBL2 EF4 has a lysine residue at the +Y position and as such chelates the  $\text{Ca}^{2+}$  ion through the backbone carbonyl group (PDB code 1UHN). In (B)–(D) the  $\text{Ca}^{2+}$  ion is represented by a green sphere, water molecules by red spheres, the chelating groups as red sticks and the side chain and backbone carbonyl groups as light green sticks.*

helix, the main-chain carbonyl oxygen atom substitutes for the missing side-chain carboxy group. The second group of non-canonical EF-loops are those with insertions

(Fig.1.6D). An entire subfamily, the S100s and related proteins including calbindinD9K, contain a non-canonical EF-loop in EF1 termed the 'pseudo-EF-loop' forming an EF-hand variant known as the  $\psi$ -hand ('pseudo-EF-hand'). To accommodate a two-residue insertion in the first part of the loop, this segment is turned inside-out with the dipoles of the main-chain carbonyl groups providing the chelating +X, +Y and +Z ligands. Interestingly, the presence of these unusual liganding groups does not prevent the formation of several favourable hydrogen bonds and, at least in the case of calbindin D9K, high-affinity  $\text{Ca}^{2+}$  binding still occurs. Main-chain co-ordinating groups are also used to overcome a three-residue insertion between the +X and +Y positions and enable pentagonal bipyramidal geometry in EF1 of AtCBL2 (Nagae *et al.*, 2003) and a one-residue insertion between these same positions in EF1 of the extracellular glycoprotein BM40 (also known as SPARC and osteonectin) (Hohenester *et al.*, 1996). Interestingly, the loops of EF1 and the canonical EF2 of BM40 can be superimposed with an rmsd (root mean square deviation) of only 0.25 Å (1 Å=0.1 nm), demonstrating that the one-residue insertion of EF1 results only in a very localized effect. The final non-canonical EF-hand that falls into this subgroup is the single functional EF-hand found in the ELC (essential light chain) of scallop (*Argopecten irradians*) myosin. The EF-loop of its EF1 is unusual in a number of ways:

- (I) although there is a significant number of aspartic acid residues present, owing to their unusual separation, main-chain carbonyl groups provide a number of ligands;
- (II) a substitution of lysine for the C-terminal glutamic acid ligand is compensated for by an insertion of two residues after the first ligand, enabling  $\text{Ca}^{2+}$ -coordination entirely by the N-terminal part of the loop; and
- (III) instead of coming from a glutamic acid residue found in the exiting helix, the final two ligands are provided by two side chains of the entering helix (Houdusse *et al.*, 1996).

TH-16987

Members of the third group of non-canonical EF-hands are rarer. At present there is only one example of a  $\text{Ca}^{2+}$ -chelation loop shorter than the canonical sequence: the eleven-residue loop of EF1 in most members of the penta-EF-hand subfamily (Fig.1.6B). In these EF-loops, as is the case for several of the loops mentioned above,



572-69 159606 St

the lack of prototypic side chains in the X and Y positions is compensated for by both an alternative loop conformation that allows for the oxygen of a main-chain carbonyl group to act as a ligand and by coordination through an additional water molecule (Blanchard *et al.*, 1997). Of course, one way in which a non-canonical EF-loop can try to bind  $\text{Ca}^{2+}$  is to abandon the pentagonal bipyramidal coordination scheme. An octahedral coordination scheme is seen in EF5 of the apoptosis-linked protein ALG-2 (apoptosis-linked gene-2). The loop of this EF-hand in all known members of the penta-EF-hand subfamily contains a two-residue insertion in the C-terminal part of the loop that inactivates the site in all members except ALG-2 (Tarabykina *et al.*, 2000). In this protein the  $\text{Ca}^{2+}$  is co-ordinated only by the N-terminal part of the loop and has an octahedral geometry, as the potential glutamine ligand is too distant and it is replaced by a single water molecule (Jia, *et al.*, 2001). Octahedral coordination is also seen in EF1 of the conventional myosin ELC from the acellular true slime mould *Physarum polycephalum* (Debreczeni *et al.*, 2005). The unusual conformation of this EF-loop prevents the loop's C-terminal ligands from participating in the coordination scheme.

### 1.3.3 The EF-hand pair

The EF-hand motif almost always occurs in pairs, creating an 11 Å distance between the two bound  $\text{Ca}^{2+}$  ions (Biekofsky *et al.*, 1998). Stacked against one another in a face-to-face manner, this pair forms a four-helix bundle with the amphipathic helices packed together to make a hydrophobic core. The structural integrity of the two EF-hand motifs is further stabilized through the short antiparallel  $\beta$ -sheet formed between the pairs EF-loops (Grabazek *et al.*, 2006). Even though the two EF-hands in the domain are related by an approximate 2-fold axis of symmetry that passes through the eighth position of the loop (Strynadka *et al.*, 1989), they are not identical. In fact, an EF-hand's position in a pair is strictly defined and designated as either 'odd' or 'even'. This asymmetry translates into asymmetric dynamic and  $\text{Ca}^{2+}$ -binding properties well documented for CaM (Evanus *et al.*, 1998; Malmendal *et al.*, 1999) and calbindin D9K (Akke *et al.*, 1993; Maler *et al.*, 2000). Two pieces of evidence point to the pervasiveness of this pairing. Firstly, it is maintained even if one of the EF-hands is non-functional, a good example of which is seen in the  $\text{Ca}^{2+}$ -bound forms of the SCPs (sarcoplasmic  $\text{Ca}^{2+}$ -binding proteins) of the invertebrates *Nereis diversicolor* (a ragworm) and *Branchiostoma* (formerly *Amphioxus*; the lancelet). For

*Nereis* SCP, although EF2 is non-functional, owing to an extensive insertion, it is still paired with a functional EF1 and the two-fold symmetry is approximately preserved (Vijay-kumar. *et al.*, 1992). In the case of *Branchiostoma* SCP, the non-functional EF4 is again paired with a Ca<sup>2+</sup>-bound EF3, but in this case the antiparallel  $\beta$ -sheet between the two EF-hands is lost (Cook. *et al.*, 1993). The second piece of evidence is provided by peptide studies on individual EF-hands. The hydrophobic packing of the EF-hand motif provides a strong driving force for peptide dimerization, the extent of which has been demonstrated in experiments using isolated EF-hands from several EF-hand containing proteins, including CaM (Reed *et al.*, 1990), TnC (troponin C) (Shaw *et al.*, 1992), calbindin D9K (Linse *et al.*, 1993) and *Nereis* SCP (Durussel *et al.*, 1993). These peptides are observed to form homodimers in the presence of Ca<sup>2+</sup>, which includes a short  $\beta$ - sheet formed between the loops. Tellingly, when a homodimer is mixed with its natural partner, a heterodimer is preferentially formed at the expense of the homodimer. This suggests a highly specific packing of the core (Godzik *et al.*, 1989), a suggestion corroborated by the 200-fold increased affinity of the heterodimer compared with the homodimer for Ca<sup>2+</sup> (Reid *et al.*, 1981; Shaw *et al.*, 1991). Interestingly, in addition to dimer formation, there is a cation-induced transition from a random-coil structure to a helix-loop-helix conformation, a transition which may have important implications for protein folding (Lopez *et al.*, 2002). Finally, the interdependent folding of a pair of Ca<sup>2+</sup>-binding sites is supported by calorimetric studies showing that the Ca<sup>2+</sup>- saturated forms of TnC and CaM unfold as a 'co-operative block' (Tsalkova *et al.*, 1980). Even though the structure of the EF-hand pair is highly conserved throughout this protein family, diversity is introduced through both structural additions to the N-terminus of the Ca<sup>2+</sup>- binding domain and by modifications to both the length and composition of the linker connecting the two EF-hands. An additional and significant level of diversity is also introduced through the topic of the next subsection – the organization of the EF-hand domains in a protein.

#### **1.3.4 EF-hand domain organization**

The smallest member of this family is calbindin D9K, a 9kDa protein which contains a single EF-hand pair. Although the other members of the S100 subfamily to which this protein belongs also contain a single pair of EF-hands, these members occur as part of either homo- or hetero-dimers, giving each of the S100 proteins four EF-

hands. Most of the proteins of the EF-hand family contain two pairs, each similar to calbindin D9K, totalling four EF-hand motifs. Three different arrangements of the two domains have been observed, depending on the conformation and length of the linker between the domains of paired EF-hands. In the first, two independent globular domains are connected by a flexible linker. As observed for sTnC (skeletal-muscle TnC) (Herzberg *et al.*, 1988), and extensively for CaM (Babu *et al.*, 1985), the plasticity of the linker in this arrangement allows the two EF-hand domains to take up a variety of orientations with respect to each other, enabling an extensive array of target interaction arrangements. In the second conformation, which is seen in the structure of recoverin and other members of the NCS (neuronal  $\text{Ca}^{2+}$ -sensor) subfamily, a U-shaped linker between the domains places the four EF-hands in a compact tandem array on one face of the protein (Ames *et al.*, 2000; Ames *et al.*, 1999). Here, the relative disposition of the domains is fixed by intra- and inter-domain contacts. The final observed arrangement for the four EF-hands is the compact globular fold seen in the structure of the invertebrate SCPs (Cook *et al.*, 1993; Vijaykumar *et al.*, 1992). In this conformation, the two pairs of EF-hands reside on opposite faces of the compact molecule. In fact, the highly apolar surfaces of the two domains exhibit such an affinity and specificity for each other that they preclude recognition and binding to the hydrophobic surface of a target protein (Rabah *et al.*, 2005). Finally, there are those members of the EF-hand superfamily which contain six EF-hands, including the neuroprotective protein calbindin D28K and its homologue calretinin, as well as a sub-family of ER regulatory proteins, including reticulocalbindin. For calbindin D28K the three pairs of EF-hands form a compact and ellipsoid-type conformation similar to that of the NCS subfamily (Kojetin *et al.*, 2006). In these larger EF-hand-containing proteins, rarely are the entire EF-hands active in binding  $\text{Ca}^{2+}$ . Because the EF-hand motif has a strong tendency to be part of a pair, examples of EF-hand proteins that contain an odd number of EF-hands are rare. A single, functional EF-hand is found in the C-terminal domain of the CaV1.2 subunit of the voltage-gated  $\text{Ca}^{2+}$  channel. Despite the non-canonical sequence found in its EF-loop, this motif is believed to bind bivalent cations and may function in  $\text{Ca}^{2+}$  homeostasis (Brunet *et al.*, 2005). For the parvalbumins, a three-EF-hand protein subfamily, the unpaired EF1 lacks several characteristic features of the EF-hand motif. Functionally, this site does not bind  $\text{Ca}^{2+}$  but serves to stabilize the  $\text{Ca}^{2+}$ -binding EF2/3 pair by behaving like an endogenous peptide bound to the C-terminal

domain in a manner reminiscent of Ca-CaM-peptide interactions (Babini *et al.*, 2005). Interestingly, substitution of a canonical loop into this inactivated EF-hand does not restore  $\text{Ca}^{2+}$ -binding and causes an overall destabilization of the structure (Cox *et al.*, 1999). Finally, an entire subfamily contains five EF-hands. However, the function of this fifth hand is to promote self-association to create a ten-EF-hand structural unit. This dimerization property of the EF-hand is particularly interesting, as it suggests that, in addition to being a  $\text{Ca}^{2+}$ -binding unit, it can also serve as a dimerization motif. In structures of penta-EF-hand subfamily members, EF5 of the first monomer is packed against the equivalent section of the second, burying one-fifth [( for example, the homodimer of calpain domain VI formed through interactions between EF5/5' (PDB code 1DVI)] of the total surface of each monomer and forming a tightly associated four-helix bundle (Blanchard *et al.*, 1997; Jia *et al.*, 2001; Ileri *et al.*, 2002; Jia, *et al.*, 2000) (Fig.1.6A). Pairing between fifth EF-hands is also used to mediate heterodimer formation between domains IV and VI of calpain as well as ALG-2 and peflin (Kitaura *et al.*, 2002). Interestingly, the structure of the EF5/5 pair mimics that of the traditional EF-hand pair, including the antiparallel  $\beta$ -sheet. Variation in the amino acid composition of the EF-hand helices leads to different helix packing interactions between the EF-hands as well as between a given EF-hand domain and the rest of the protein.

## 1.4 Consequences of $\text{Ca}^{2+}$ binding

### 1.4.1 $\text{Ca}^{2+}$ -binding and induced structure formation

Traditionally, the role of  $\text{Ca}^{2+}$  binding has been examined through the lens of signal transduction, which focuses on the  $\text{Ca}^{2+}$ -induced conformational changes and their effects on target interactions.  $\text{Ca}^{2+}$  can also play a de novo structural role and, in many cases, the binding of this cation is key to the structural integrity of the protein. EF-hands that serve a structural function have a high  $\text{Ca}^{2+}$ -affinity sufficient enough for  $\text{Ca}^{2+}$  to be bound even in the resting cell. Proteins such as the sarcoplasmic CaBPs (the invertebrate SCPs (Christova *et al.*, 2000; Precheur *et al.*, 1996), the C-terminal domain of CaVP (calcium vector protein) (Theret *et al.*, 2000) and the prokaryotic protein calerythrin (Aitio *et al.*, 2001) all exhibit an unfolded, molten-globule-like state in the absence of  $\text{Ca}^{2+}$  in vitro. Another example is the bilobal TnC, in which the

two EF-hands of the C-terminal domain have an approximately 100-fold higher affinity for  $\text{Ca}^{2+}$  than have the N-terminal EF-hands – an affinity high enough to ensure  $\text{Ca}^{2+}$  saturation in the resting cell. Like the aforementioned proteins, the C-terminal domain of TnC is unfolded in the absence of  $\text{Ca}^{2+}$  (Li *et al.*, 1994). When compared with the N-terminal domain, this lack of structure is thought to be due to the absence of stabilizing hydrophobic interactions that overcome the destabilizing electrostatic interactions, the main source of which are the carboxy ligands of the EF-hand motif (Ingraham *et al.*, 1983). Interestingly, a structural role for the EF-hand has also been suggested for several multidomain signalling proteins, since for both phospholipase  $\text{C}\delta$  (Bairoch *et al.*, 1990) and EH (Eps15-homology) domains (de Beer *et al.*, 1998),  $\text{Ca}^{2+}$  binding has no effect on the function of the protein. In contrast with these proteins, in which  $\text{Ca}^{2+}$ -binding induces structure formation, most EF-hand-containing proteins are structured in the apo state, and the binding of this cation leads to a conformational change through which this protein transduces the message of an increased  $\text{Ca}^{2+}$  concentration.

#### 1.4.2 $\text{Ca}^{2+}$ -binding and conformational change

Initially the  $\text{Ca}^{2+}$ -induced conformational transformation was considered in terms of a closed-to-open domain transition quantified through the change in interhelical angle of vectors defined by the helices that bind  $\text{Ca}^{2+}$ . For the members of the CaM subfamily and several other EF-hand-containing proteins, these helices, antiparallel in the apo structure with interhelical angles of 130–140°, open upon  $\text{Ca}^{2+}$  binding to become roughly perpendicular with an angle that approximates to 90° (Nelson *et al.*, 1998; Yap *et al.*, 1999). Through this change in angle there is a loss of interhelical contacts between the entering and exiting helices of the EF-hand, as the top where  $\text{Ca}^{2+}$  binds is ‘pinched’ together and the bottom moved apart – a conformational change that exposes once-buried hydrophobic side chains and creates an extensive target-protein interaction site. In CaM these surfaces are approximately 10 Å × 12.5 Å in each domain (Zhang *et al.*, 1995). At a molecular level this conformational change appears to be a consequence of  $\text{Ca}^{2+}$  chelation by the EF-loop. Initially the  $\text{Ca}^{2+}$  ion is bound by ligands in the N-terminal part of the loop (1, 3, 5 and 7), as this region is quite flexible and ‘waves around’ to catch the ion (Kobayashi *et al.*, 2006). At this

point the C-terminal ligands, predominantly those provided by the twelfth position, are too far away to chelate the bound  $\text{Ca}^{2+}$  ion directly. In order for the loop to complete the ion's coordination sphere (an energetically favourable process), the exiting helix must be repositioned by  $\sim 2\text{\AA}$ . It is the movement of this helix that causes the conformation change in the EF-hand. The repositioning of Glu12 is stabilized not only through the Ca–O 'bonds' that form, but also through two new hydrogen bonds between this side chain and the backbone at positions 2 and 9. The importance of these hydrogen bonds to the  $\text{Ca}^{2+}$ -induced conformational change has been seen in structural studies in which this glutamic acid residue is replaced with a glutamine residue (Evanus *et al.*, 1998; Evanus *et al.*, 1997; Wimberly *et al.*, 1995). The inability of the glutamine residue's side chain to simultaneously form both of these hydrogen bonds leads to a conformational exchange between two conformations similar to those of the closed and open states. This conformational exchange suggests that, unlike in the wild-type protein, the equilibrium here is not being pushed towards the open structure and that one additional hydrogen bond ( $\sim 12$  kJ/mol) is required to shift the balance (Evanus *et al.*, 1997). This two-step binding pathway suggests an explanation for the observed asymmetry in the EF-hand pair mentioned above. As the linker between EF1 and EF2 provides a strong covalent coupling between helices II and III of the pair, rotation of helix II will cause pulling on helix III, forcing it to follow and re-orient. By contrast, as there is no covalent link between helices I and IV, a rotation of helix IV by itself can occur more readily. The consequence of this smaller conformational cost of  $\text{Ca}^{2+}$ -binding to EF2 compared with EF1 is seen in several EF-hand proteins, including CaM, TnC and CIB, where the second EF-hand of the pairs is filled first, since the binding affinity is higher.

### **1.5 Calcium Binding Proteins of *Entamoeba histolytica***

Calcium plays an essential role in many fundamental processes in almost all eukaryotic cells including protozoan parasite *Entamoeba histolytica* (Moreno *et al.*, 2003). Intracellular calcium plays a crucial role for the control of a variety of cell functions in eukaryotes including cell migration, contraction, secretion, proliferation and differentiation, exocytosis, transcellular ion transport, neurotransmitter release and gap junction regulation (Tsein *et al.*, 1990). The uptake and release of calcium ions ( $\text{Ca}^{2+}$ ) across the plasma membrane and intracellular organelles is orchestrated

by the concerted operation of distinct and number of calcium transporting systems. These processes are essentially mediated by a variety of  $\text{Ca}^{2+}$ -binding proteins (CaBPs) (Berridge *et al.*, 2000.), which are involved in binding  $\text{Ca}^{2+}$  and transducing the signal through down-stream effectors. Many CaBPs can also directly function as effectors, for example, calcium-dependent protein kinases (CDPKs) of plants (Harper *et al.*, 2005). Once inside the cell,  $\text{Ca}^{2+}$  can either interact with so-called soluble CaBPs or become sequestered into intracellular organelles.

One of the well-studied CaBP is calmodulin (CaM), a four EF-hand highly conserved CaBP. It has been found in almost all eukaryotic cells and has been implicated in a large number of cellular processes (Chin *et al.*, 2000).  $\text{Ca}^{2+}$ -signaling also plays a crucial role in the pathogenesis of many protozoan parasites (Scheibel *et al.*, 1992). In Plasmodium, the expression of CaM has been shown to be stage-specific and is involved in erythrocyte invasion as well as schizont maturation (Matsumoto *et al.*, 1987). Chelation of  $\text{Ca}^{2+}$  can prevent hepatocyte invasion by merozoites (Johnson *et al.*, 1981).

A number of CaBPs have been identified in *E. histolytica* (Scheibel *et al.*, 1992). In the protozoan parasite *Entamoeba histolytica*, the causative agent of amebiasis,  $\text{Ca}^{2+}$  is reported to be involved in its pathophysiology by initiating the amebic cytolytic activity (Meza *et al.*, 2000). Amebic cytolytic activity could be blocked by  $\text{Ca}^{2+}$ -channel blockers, or by treatment with EGTA (Ravdin *et al.*, 1988), whereas stimulation of amebic PKC activity with phorbol esters enhanced lysis of target epithelial cells (Weikel *et al.*, 1988). Changes in the  $\text{Ca}^{2+}$  profile were also related to the cell cycle and the developmental stages of the parasite, i.e., the cyst or the trophozoite stage (Ganguly *et al.*, 2001). Moreover, extracellular  $\text{Ca}^{2+}$ , amebic intracellular  $\text{Ca}^{2+}$  flux, bepridil-sensitive  $\text{Ca}^{2+}$  channels and a putative CaM-dependent signal transduction pathway have been implicated in the growth and encystation of *Entamoeba* (Makioka *et al.*, 2001).  $\text{Ca}^{2+}$  is thus speculated to play a direct role in precipitating the cytopathic effects of *E. histolytica*. Among these are two related EF-hand-containing proteins, granin 1 and granin 2, which are localized in intracellular granules (Nickel *et al.*, 2000). They may be involved in phagocytosis, control of endocytotic pathways and  $\text{Ca}^{2+}$ -dependent granular discharge. However, there is no experimental evidence in support of any of the suspected functional involvement of these proteins. Another protein, URE3-BP, was shown to have a transcription

regulatory function with Ca<sup>2+</sup>-dependent DNA binding properties (Gilchrist *et al.*, 2001, 2003). *E. histolytica* genome encodes a large repertoire of CaBPs as revealed by a motif-based search for EF-hand containing proteins suggesting an extensive Ca<sup>2+</sup>-based signaling network in this organism (Table-1.1) (Bhattacharya *et al.*, 2006).

**Table 1.1 - Calcium Binding Proteins of *Entamoeba histolytica***

Predicted name (in genome database)	Class	Length (amino acids)	Number of copies in genome database (with accession no.)	No. of EF-hands (SMART/Pfams)	EST	Homology in other species of <i>Entamoeba</i>
EhCaBP1	CaBP1	134	EAL48959	4	Yes	<i>E. terrapinae</i> <i>E. moshkovskii</i>
EhCaBP2	CaBP2	134	EAL51694	4	Yes	<i>E. terrapinae</i> <i>E. moshkovskii</i>
CaM	CaBP3	151	EAL46322	3	Yes	ND
CaM-put	CaBP4	146	EAL51814, EAL46978	2	ND	ND
CaM	CaBP5	144	EAL46660	2	ND	ND
CaM	CaBP6	150	EAL50371, EAL50341	4	ND	ND
CaM-like	CaBP7	155	EAL43751	3	ND	ND
CaBP	CaBP8	263	EAL50453	2	ND	ND
CaBP	CaBP9	157	EAL52004	4	ND	<i>E. moshkovskii</i>
CaBP	CaBP10	153	EAL50237	2	ND	ND
CaBP	CaBP11	627	EAL50040	3	ND	ND
CaBP	CaBP12	145	EAL51761	3	ND	ND
CaBP	CaBP13	336	EAL48778	3	Yes	ND
CaBP	CaBP14	264	EAL46305	3	ND	<i>E. dispar</i>
Granin 1	CaBP15	215	EAL47854	3	ND	ND
Granin 1	CaBP16	215	EAL44984, EAL44709, EAL44971	3	Yes	<i>E. terrapinae</i> , <i>E. dispar</i> , <i>E. moshkovskii</i>
Granin 2	CaBP17	213	EAL44985, EAL44970	3	Yes	ND
Granin 2	CaBP18*	600	EAL49043, EAL49043, EAL42900	3	Yes	ND
CaBP/URE3BP	CaBP19	220	EAL42646	3	Yes	ND
Myosin light chain	CaBP20	146	EAL52163, AL50546	3	ND	<i>E. dispar</i>
Calcineurin B-subunit	CaBP21	179	E45078AL	2	Yes	<i>E. terrapinae</i> , <i>E. dispar</i> , <i>E. inadens</i> , <i>E. moshkovskii</i>
Hypothetical	CaBP22	317	EAL45535	3	ND	ND
Hypothetical	CaBP23	169	EAL43738	2	ND	ND
Hypothetical	CaBP24	318	EAL49541	6	ND	ND
Hypothetical	CaBP25	307	EAL45702	6	ND	ND
Hypothetical	CaBP26	673	EAL51827	6	ND	ND
Hypothetical	CaBP27	658	EAL51029	5	ND	ND

ND, not determined, \* likely to be actinin (Adapted from Bhattacharya *et al.*, 2006)

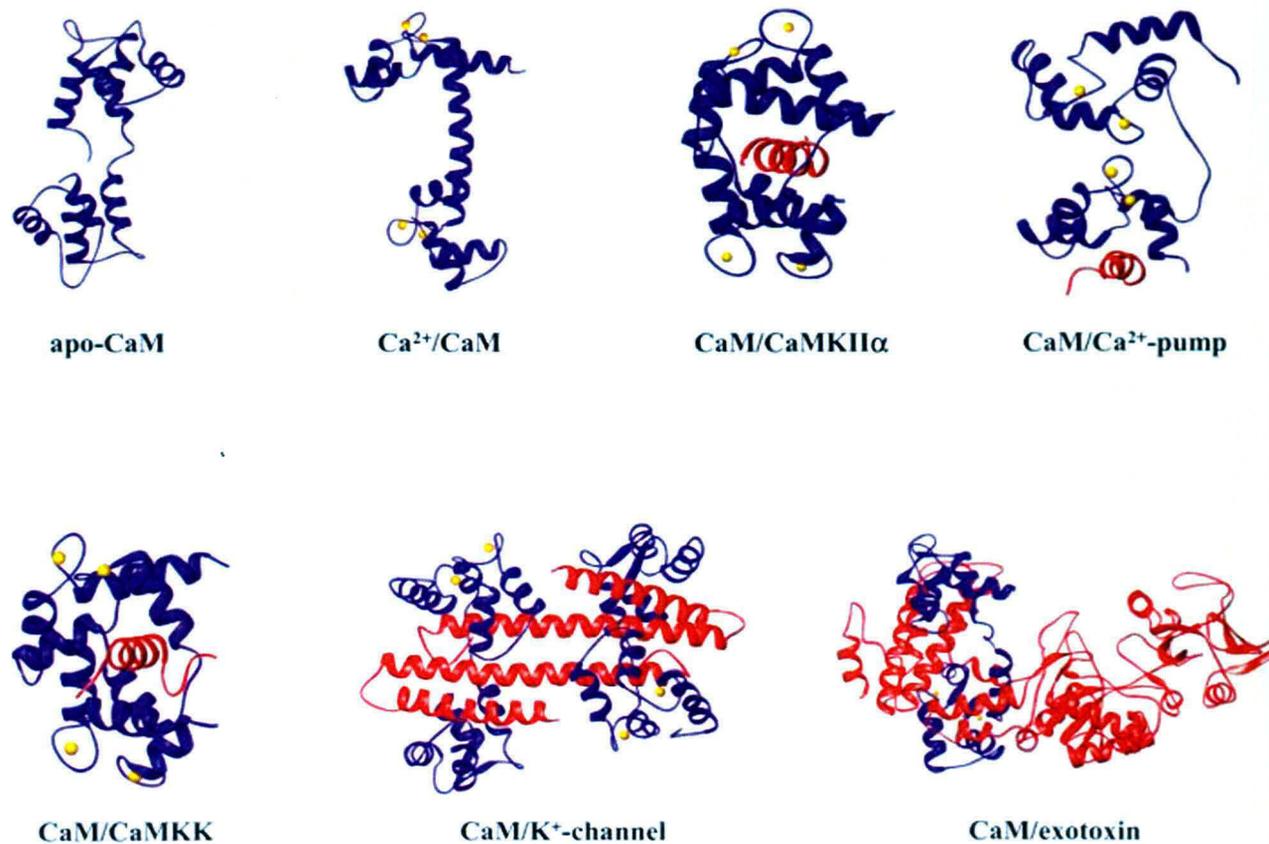
## 1.6 Target binding mode of Calmodulin: A prototypic Calcium Binding Protein

Calmodulin (CaM) is a highly conserved  $\text{Ca}^{2+}$  binding protein, which is ubiquitous and central in translating  $\text{Ca}^{2+}$  levels into physiological signals. For example, CaM-dependent kinases are activated when  $\text{Ca}^{2+}$  CaM binds to the auto-inhibitory region of the target with hydrophobic patches in each of its two lobes to contact two hydrophobic anchor residues located at “1-10,” “1-14,” or “1-16” spacings in the primary sequence of the kinase (Ikura *et al.*, 1992; Meador *et al.*, 1992, 1993; Osawa *et al.*, 1999). This small 17 kDa acidic protein belongs to a family of homologous calcium-binding proteins that bind  $\text{Ca}^{2+}$  through the EF-hand motif (e.g., parvalbumin (Krietsinger *et al.*, 1973) or troponin C (Herzberg *et al.*, 1988)).

CaM interacts with target proteins in a  $\text{Ca}^{2+}$ -dependent manner (Klee, 1988; Crivici *et al.*, 1995) that translates the second messenger calcium into a variety of cellular responses. Calcium binding to CaM exposes hydrophobic clefts in each of the CaM lobes that interact with downstream targets (Crivici *et al.*, 1995). The crystal structures of  $\text{Ca}^{2+}$  bound CaM show that this protein contains two similar domain structures linked by an extended  $\alpha$ -helix (Babu *et al.*, 1985; Babu *et al.*, 1988; Taylor *et al.*, 1991; Chattopadhyaya *et al.*, 1992; Rao *et al.*, 1993; Wilson *et al.*, 2000), each domain containing two helix-loop-helix EF-hand  $\text{Ca}^{2+}$ -binding motifs. The incorporation of  $\text{Ca}^{2+}$  into CaM is essential and leads to a major conformational change which includes the opening of a hydrophobic cavity in each globular domain necessary for target protein recognition (Zhang *et al.*, 1995).

Various modes of target peptide recognition have been thoroughly studied especially by X-ray crystallography (Fig.1.7) (Ikura *et al.*, 1992; Meador *et al.*, 1992; Osawa *et al.*, 1999; Elshorst *et al.*, 1999; Schumacher *et al.*, 2001; Han *et al.*, 2000). A compact, calcium-free, apo form of CaM is converted to an extended dumbbell-shaped form on binding  $\text{Ca}^{2+}$  (Zhang *et al.*, 1995). This extended conformation consists of two structurally similar domains separated by a flexible 28-residue helix. Each domain has two EF-hand motifs with bound  $\text{Ca}^{2+}$ . The calcium-induced extension of CaM exposes two hydrophobic pockets, one per domain, which represent the binding sites for target proteins. Each hydrophobic pocket in the N- and C-lobes of CaM grabs well-conserved bulky hydrophobic residues separated by a specific number of residues (10, 14 or 16) on the target peptide. On binding to the protein

target, the central CaM helix unwinds, and the two hydrophobic pockets wrap around the  $\alpha$ -helix of the protein target. Various target sequences are classified according to this spacing (Rhoads *et al.*, 1997; Yap *et al.*, 2000).



**Figure-1.7** *Ribbon presentations of CaM and CaM in complex with target peptides. CaM is colored blue, the CaM targets are red,  $\text{Ca}^{2+}$  ions are yellow. The N-terminal lobe of CaM is orientated to the top, the C-terminal lobe to the bottom of the figures. Structural data were taken from the Protein Data Bank, accession codes: apo-CaM (1CFD)  $\text{Ca}^{2+}$ /CaM(1CLL); CaM/CaMKIIa (1CM1); CaM/CaMKK (1IQ5); CaM/ $\text{Ca}^{2+}$ -pump (1CFF); CaM/K<sup>+</sup>-channel (1G4Y), CaM/anthrax exotoxin (1K93) (Adapted from Vetter *et al.*, 2003).*

Comparison of the structures of apo-CaM and  $\text{Ca}^{2+}$ /CaM determined by NMR and X-ray crystallography reveals that the overall structure and distribution of

secondary structure elements is very similar in both molecules (Kuboniwa *et al.*, 1995; Finn *et al.*, 1995; Zhang *et al.*, 1995).

The major conformational change induced by the binding of  $\text{Ca}^{2+}$  ions into the EF-hand pairs is a significant alteration of the relative orientation of the helices in each lobe (Ikura *et al.*, 1996; Potter *et al.*, 1983; Chou *et al.*, 2001). As the result of a twist-like motion, the relative angles between the helices change from  $121^\circ$ - $144^\circ$  in apo-CaM to  $86^\circ$ - $116^\circ$  in  $\text{Ca}^{2+}$ /CaM (Zhang *et al.*, 1995). This rearrangement of the helices leads to the exposure of several hydrophobic residues to the solvent, which forms a large hydrophobic, concave patch or channel on the surface of each lobe.

### 1.7 Metals in Proteins and plasticity of calcium binding loops

Metals are important for the biological activity of proteins. Metal ions in proteins can act as structure promoters or can take part in enzymatic reactions. Divalent metal cations such as  $\text{Zn}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Cu}^{2+}$  and  $\text{Ca}^{2+}$  are often associated with the catalytic or regulatory activities of proteins that constitute some of the fundamental chemical life processes (e.g. Nelson *et al.*, 2000; Dudev *et al.*, 2003; Nicholson, 1997; Cowan, 1998). For example, Mg in chlorophyll is important for photosynthesis, Cu (together with Fe) has a role in oxygen-carrying proteins and Zn, Mg and Ca can serve as Lewis acids. Although the coordination geometry of proteins specifically invites essential metal ions, it is susceptible to interact with toxic heavy metals. Heavy metals may accumulate in the body and cause disease states even at low concentrations (Cline *et al.*, 1996; Lidsky *et al.*, 2003). Calcium-binding proteins have received considerable attention due to their diverse roles in biological systems. The vast majority of these sites conform to the EF-hand motif proposed by Kretsinger (Kretsinger *et al.*, 1973), yet their affinities and specificities vary considerably (Levine *et al.*, 1982). Understanding the physical properties of calcium-binding loops has been difficult because in most proteins containing EF-hand type sites, the loops are paired (Kretsinger *et al.*, 1987). Recently, proteins have been considered as a tool for removing heavy metals from the environment in order to reduce pollution (Hennig *et al.*, 1986; Sano *et al.*, 2006a,b). In this respect, knowledge of the preferred coordination surrounding a metal ion could help in choosing proteins that would be

more effective in removing certain metal ions from the environment than those already available.

Calmodulin (CaM) is a prototypic EF-hand calcium-binding protein that acts by sensing calcium levels and binding to target proteins in a regulatory manner. Its importance is highlighted by its 100% sequence conservation across all vertebrates. Although CaM is best characterized by its ability to specifically bind  $\text{Ca}^{2+}$ , a number of studies have indicated that in fact it can also be activated by other metal ions (Habermann *et al.*, 1983; Chao *et al.*, 1984; Richardt *et al.*, 1986; Ouyang *et al.*, 1998; Ozawa *et al.*, 1999). Of these,  $\text{Pb}^{2+}$  is of special interest because of the possible role that  $\text{Pb}^{2+}$ -CaM may play in  $\text{Pb}^{2+}$  toxicity (Goering *et al.*, 1993), which has a high affinity towards CaM and is able to activate it even at low concentrations (Habermann *et al.*, 1983; Richardt *et al.*, 1986; Kern *et al.*, 2000). To a large extent,  $\text{Pb}^{2+}$  seems to be able to substitute for calcium in the regulation of CaM function (Habermann *et al.*, 1983; Fullmer *et al.*, 1985) and many target proteins are indeed activated by Pb-CaM similarly to Ca-CaM (Chao *et al.*, 1995).  $\text{Pb}^{2+}$  binds to all four metal-binding loops in CaM simultaneously and with such high affinity that  $\text{Pb}^{2+}$  can displace bound  $\text{Ca}^{2+}$  from CaM (Aramini *et al.*, 1996; Fullmer *et al.*, 1985). Furthermore,  $\text{Pb}^{2+}$  can functionally substitute for  $\text{Ca}^{2+}$  in some aspects of CaM function, including the ability of  $\text{Pb}^{2+}$ -CaM to bind to the target peptide from myosin light-chain kinase (Chao *et al.*, 1995).  $\text{Ba}^{2+}$  and  $\text{Mg}^{2+}$  have been reported to have significantly lower, if any, affinity towards CaM (Ozawa *et al.*, 1999). Isothermal titration calorimetry of  $\text{Sr}^{2+}$ -binding has been reported recently with Centrin, a highly conserved calcium-binding proteins present in all eukaryotic cells (Charbonnier *et al.*, 2006).

## AIMS AND OBJECTIVES

*Entamoeba histolytica*, a protozoan parasite, is the causative agent of amoebiasis. The genome analysis shows that *E. histolytica* encodes a large number of calcium-binding proteins. A number of studies show that  $\text{Ca}^{2+}$  and its binding proteins are centrally involved in amebic pathogenesis. One of the first identified four EF-hand motifs containing  $\text{Ca}^{2+}$ -binding protein from *E. histolytica* (EhCaBP1) is a 14.7 kDa (134 amino acid residues) protein, with no known homologue in the database. Though this protein has some structural similarity with CaM, it is functionally quite distinct and binds to different set of proteins compared to CaM. Inducible expression of EhCaBP1 antisense RNA categorizes it to be an essential gene and demonstrated its role in cellular proliferation. Also, the confocal microscopic studies suggest the cytoplasmic localization of the protein including pseudopods and beneath the plasma membrane. The enrichment of EhCaBP1 on pseudopods suggests its association with areas rich in actin cytoskeleton and its associated proteins. This suggests the involvement of EhCaBP1 in dynamic changes of the actin cytoskeleton. Recently, it has been reported that the N-terminal domain of EhCaBP1 carry out most of the functions of the full length protein. Based on the previous discussions, the aims of this study have been designed. They are summarized below:-

1. Structural studies of Calcium Binding Protein-1 from *Entamoeba histolytica*.
2. Visualization of mode of target binding of EhCaBP1.
3. Structural studies of N-terminal domain of EhCaBP1.
4. Flexibility and plasticity of calcium binding loops of EhCaBP1.

This study envisages X-ray crystallography and other biophysical techniques to reveal the three-dimensional structure of EhCaBP1, its mode of target binding as well as the plasticity of the calcium binding loops.

# MATERIALS AND METHODS

---

## 2.1 Sources of materials:

Clones of full-length EhCaBP1 and its N-terminal domain in pET-3c vector were kindly provided by Prof. Alok Bhattacharya, School of Life Sciences, Jawaharlal Nehru University, New Delhi, India.

Ion-exchanger resins were obtained from Amersham Pharmacia, (U.S.A.) Chromatographic columns were obtained from Sigma-Aldrich (U.S.A.), Centricons were obtained from Millipore. Membrane filter were purchased from Millipore. Dialysis tubings from Sigma-Aldrich (U.S.A.) and Spectrum, Protein Marker from Fermentas, Crystallization reagents from Hampton (U.S.A.), Jena-Biosciences (Germany), Sigma-Aldrich (U.S.A.) and all the Heavy metals were purchased from Merck.

(All concentrations indicated in percentage are in (w/v) basis unless stated otherwise. All solutions were prepared in double distilled water unless stated otherwise. Autoclaving was done at a pressure of 15 lbs per square inch for 20 min.)

## 2.2 Organisms and growth conditions:

*E. coli* DH5 $\alpha$  has the genotype: *SupE44 lacU169* ( $\phi$ 80 *lacZ* M15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*. Cells from an agar stab or frozen glycerol stock were first streaked on an LB plate (containing the appropriate antibiotic wherever necessary) and allowed to grow overnight at 37°C. Liquid cultures in LB medium were initiated from a single colony and were grown with constant shaking at 220 rpm at 37°C. The cells were grown overnight, were used as inoculums for further growth by diluting 100 fold in fresh LB medium and grown with aeration at 37°C for 3-4 hr to obtain log phase cultures.

## 2.3 Culture media:

### 2.3.1 Luria Broth (LB):

#### Composition per litre

Tryptone	10g
Yeast extract	5g
Sodium Chloride	5g

The components were dissolved in double distilled water and pH adjusted to 7.0 using 2 N NaOH. The medium was sterilized by autoclaving.

### **2.3.2 LB Agar:**

LB agar was prepared by adding 1.5% (W/V) of Bacto-Agar to LB medium and sterilized by autoclaving. Ampicillin was added to a final concentration of 100 µg/ml after cooling the LB agar to around 55°C and plates poured.

## **2.4 Plasmid DNA isolation:**

### **2.4.1 Boiling Method (mini-prep) (Holmes and Quigley, 1981):**

A single bacterial colony containing the plasmid was inoculated in 2 ml LB medium containing the appropriate antibiotic (50 µg/ml in case of Ampicillin) and grown overnight at 37°C with shaking at 220 rpm. The overnight culture was transferred to 1.7 ml microfuge tubes and centrifuged at maximum speed for 30 seconds at 4°C. The medium was completely drained off and the pellet was resuspended in 100 µl of ice-cold Alkaline lysis solution I [50 mM Glucose, 25 mM Tris-Cl (pH-8.0) and 10 mM EDTA (pH - 8.0)] by vigorous vortexing. 200 µl of freshly prepared Alkaline lysis solution II (0.2 N NaOH and 1% SDS) was then added to each bacterial suspension. The tubes were closed and mixed the contents by inverting the tubes rapidly (five times). Then 200 µl of Alkaline lysis solution III [3 mM potassium acetate (pH-5.5)] was added and dispersed through the viscous bacterial lysate by inverting the tube several times. The tubes were stored on ice for 10 minutes and centrifuged at maximum speed for 10 minutes at 4°C. The supernatant was transferred to a fresh microfuge tubes. The DNA was precipitated from the supernatant by adding 0.6 vol of isopropanol and mixed by inverting. The precipitated nucleic acid was collected by centrifugation at maximum speed for 10 minutes at room temperature. The pellet so obtained was washed with 500 µl of 70% (v/v) ethanol, air dried and dissolved in 20 µl of T<sub>10</sub>E<sub>1</sub> (10 mM Tris. Cl, pH 8.0, 1 mM EDTA, pH 8.0) containing 0.1 mg/ml RNase A. The tube was incubated at 37°C for 30 min for RNase digestion and 2 µl was loaded on a 0.8% agarose gel to quantitate the DNA.

#### **2.4.2 Alkaline lysis method (midi-prep) (Birnboim *et al*, 1979):**

Bacterial cells were grown overnight in 50 ml LB medium with the appropriate antibiotic and collected by centrifugation (6,000 rpm at 4°C for 8 minutes). The cell pellet was resuspended in lysis buffer [25 mM Tris.Cl (pH 8.0), 10 mM EDTA (pH 8.0), 15% Sucrose] containing 2 mg/ml lysozyme, and kept on ice for 10 min. To the lysed cells were added 3 ml of denaturing solution (freshly prepared solution containing 0.2 N NaOH and 1% SDS) and mixed gently by inverting. To the tube was added 1.6 ml of 3 M sodium acetate (pH 4.6) and incubated on ice for 20 minutes followed by centrifugation at 12,000 rpm for 20 min at 4°C. RNA was removed by incubating the supernatant with 12-15 µl of RNase A (10 mg/ml) at 37°C for 45 min. The supernatant was extracted twice with phenol:chloroform:isoamyl alcohol [25:24:1 (v/v/v)] and once with equal volume of chloroform:isoamyl alcohol [24:1 (v/v)]. The upper aqueous layer was transferred to an oakridge tube and the DNA precipitated by addition of 2.5 volumes of chilled ethanol and left for overnight incubation at -20 °C or 45 minutes at -80°C. The DNA was precipitated by centrifugation at 12,000 rpm for 20 minutes at 4°C. The pellet was resuspended in 0.4 ml of nuclease free water and to this was added 120 µl of 4M NaCl and 0.5 ml 13% PEG 8000. The tube was incubated in ice for 1h and centrifuged at RT at 12,000 rpm for 15 min. The pellet was washed with 200 µl of 70% ethanol, dried in a 37°C incubator and resuspended in 50 µl of T<sub>10</sub>E<sub>1</sub>.

#### **2.5 Transformation of *E. coli* cells:**

##### **2.5.1 Preparation of competent cells (Hanahan, 1985):**

A fresh single colony of *E. coli* (strain BL21-DE3) was grown overnight at 30°C, 200 rpm into 10 ml 2X LB medium in a 50 ml. flask. 1% of the overnight culture was added into 50 ml 2X LB medium in 500 ml flask. The cells were grown at 30°C, 200 rpm to an OD<sub>600</sub> of 0.40-0.45. The cells were chilled in ice water bath for 2 hr and were thereafter collected by centrifugation at 6,000 rpm for 5 min at 4°C and resuspended by swirling in 25 ml ice-cold filter sterilized acid-salt buffer (100 mM CaCl<sub>2</sub>, 70 mM MnCl<sub>2</sub> and 40 mM CH<sub>3</sub>COONa; pH-5.4). The resuspended cells were incubated in ice for 45 minutes and collected by centrifugation at 5,500 rpm, 8 min,

4°C. The pellet was finally resuspended in 2.5 ml of ice-cold acid-salt buffer; glycerol was then added and mixed gently. Immediately 100 µl aliquots were snap-frozen by immersing the tightly closed tubes in liquid nitrogen and stored at -80°C until needed.

### **2.5.2 Transformation of competent cells (Hanahan, 1985):**

Competent cells were thawed on ice and to 100 µl cells, 5-10 ng of plasmid DNA was added. The cells were incubated on ice for 45 min. Cells were then given a heat shock at 42°C for 90 seconds and incubated on ice for 2 min. 0.9 ml of LB was added to the cells and the cells were grown at 37°C for 1h at 225 rpm. Transformants were plated on LB agar plates with appropriate antibiotic and incubated at 37°C for 12-16 h.

### **2.6 Clones of full-length EhCaBP1 and N-terminal domains of EhCaBP1:**

The full-length EhCaBP1 gene and its N-terminal domain in the pET-3c expression vector and was kindly provided by the laboratory of Prof. Alok Bhattacharya, School of Life Sciences, Jawaharlal Nehru University, New Delhi, India).

### **2.7 Agarose gel electrophoresis:**

The agarose concentrations used in electrophoresis separation were chosen based on the size of the DNA to be resolved. Agarose was melted in 0.5X or 1X TBE [45 mM Tris-borate and 1 mM EDTA, pH 8.0] by heating and was cooled to about 50°C before adding 0.5µg/ml of ethidium bromide. The molten agarose was poured in a tray and allowed to set. After the gel had set, DNA samples were loaded and electrophoresed in 0.5X to 1X TBE in appropriate electric field strength for optimum separation. The DNA was visualized at 302 nm using a transilluminator.

### **2.8 Protein purification**

#### **2.8.1 Over-expression and purification of EhCaBP1:**

A 5ml culture (LB + 100 µg/ml ampicillin) of a single well-isolated EhCaBP1 colony was inoculated in a 100 ml conical flask (Borosil) and incubated overnight at 37°C at 220 rpm. A 200 ml culture (LB + 200 µg/ml ampicillin) in 1L of conical flask was inoculated with 1% of the above culture and grown at 37°C/220 rpm till the OD<sub>600</sub>

reaches between 0.5 to 0.7. The bacterial culture was induced with 1 mM IPTG and incubated for 3-4 hr at 37°C/220 rpm. The induced bacteria was collected at 6000 rpm at 4°C for 10 min and washed once with Wash buffer (50 mM Tris.Cl, pH 7.5 and 100 mM NaCl). The cell pellet was suspended in 1/25<sup>th</sup> volume of the original culture in 50 mM Tris.Cl pH 7.5 and 2 mM EGTA). The cells were lysed by freeze-thawing thrice in liquid nitrogen followed by sonication (3 X 30sec, full burst, with 1 min interval) on ice. The sonicated sample were spun down at 12,000 X g for 30 min at 4°C and the supernatant was loaded on to a packed DEAE sepharose column.

### **2.8.2 Over-expression and purification of N-terminal domain of EhCaBP1:**

A 5ml culture (LB + 100 µg/ml ampicillin) of a single well-isolated EhCaBP1 colony was inoculated in a 100 ml conical flask (Borosil) and incubated overnight at 37°C at 220 rpm. A 200 ml culture (LB + 200 µg/ml ampicillin) in 1L of conical flask was inoculated with 1% of the above culture and grown at 37°C/220 rpm till the OD<sub>600</sub> reaches between 0.5 to 0.7. The bacterial culture was induced with 1mM IPTG and incubated for 3-4 hr at 37°C/220 rpm. The induced bacteria was collected at 6000 rpm at 4°C for 10 min and washed once with Wash buffer (50 mM Tris.Cl, pH 7.5 and 100 mM NaCl). The cell pellet was suspended in 1/25<sup>th</sup> volume of the original culture in 50 mM Tris.Cl pH 7.5 and 2 mM EGTA). The cells were lysed by freeze-thawing thrice in liquid nitrogen followed by sonication (3 X 30sec, full burst, with 1 min interval) on ice. The sonicated sample were spun down at 12,000 X g for 30 min at 4°C and the supernatant was loaded on to a packed DEAE sepharose column.

### **2.8.3 Packing and Equilibration of column:**

The DEAE-Sepharose beads was soaked in autoclaved water (Milli Q) and mixed and the pH was maintained at 7.5 with dilute HCl and kept at 4°C for overnight incubation. Next day, the beads were transferred to the Sigma column and equilibrated with buffer pH 7.5 containing 100 mM Tris-Cl and 2 mM EGTA. This process was repeated 4-5 times and finally the column was packed and kept at 4°C until further use.

#### **2.8.4 Binding and Elution of EhCaBP1:**

The column was equilibrated with 20 bed-volumes of 50 mM of Tris-Cl (pH 7.5) + 2 mM EGTA (pH 7.5) (20 bed volumes). The sample was loaded to the column either using gravity flow method or using peristaltic pump from Amersham Biosciences and the flow through was collected. The column was washed with 20 bed-volumes of 50 mM of Tris-Cl (pH 7.5), 2 mM EGTA (pH 7.5). The protein was eluted in 1ml fractions with 50 mM Tris-Cl (pH 7.5), 10 mM CaCl<sub>2</sub> and the OD<sub>280</sub> was taken.

#### **2.8.5 Binding and Elution of N-terminal domain of EhCaBP1:**

The column was equilibrated with 20 bed-volumes of 50mM of Tris-Cl (pH 7.5) + 2 mM EGTA (pH 7.5) (20 bed volumes). The sample was loaded to the column either using gravity flow method or using peristaltic pump from Amersham Biosciences and the flow through was collected. The column was washed with 10 bed-volumes of 50 mM of Tris-Cl (pH 7.5), 2 mM EGTA (pH 7.5) and 10 bed-volumes of 50 mM of Tris-Cl (pH 7.5), 2 mM EGTA (pH 7.5) and 20 mM of NaCl. The protein was eluted in 1ml fractions with 50 mM Tris-Cl (pH 7.5), 3 mM CaCl<sub>2</sub> and the OD<sub>280</sub> was taken.

#### **2.9 Desalting and equilibration of protein samples:**

Protein samples were desalted by dialysis against the appropriate buffer using a dialysis membrane of 10 kDa/3 kDa molecular weight cut-off. Samples were dialyzed against 100 volumes of desired buffers at 4°C with minimum 5-6 changes.

#### **2.10 SDS-Polyacrylamide gel electrophoresis (Laemmli UK, 1970):**

SDS-PAGE was carried out under reducing conditions. The separating gels (12-14% acrylamide as per need) was prepared using acrylamide (acrylamide:bis-acrylamide=29:1) in 1.5% Tris-Cl pH 8.8, 0.1% (w/v) SDS, 0.04% (w/v) APS and TEMED. After polymerization of separating gel, stacking gel was poured. The stacking gel contained 4% acrylamide in 0.5% Tris-Cl pH 6.8, 0.1% (w/v) SDS, 0.04% (w/v) APS and TEMED. Prewarmed samples and 4X SDS-PAGE loading dye [125 mM Tris-Cl pH 6.8, 4% (w/v) SDS, 10% (w/v) 2-mercaptoethanol, 20% (v/v) glycerol and 0.2% (w/v) bromophenol blue] were mixed to 1X dye concentration and reboiled for 3 min. After electrophoresis, proteins were fixed in the gel by incubating

in fixing solution (45% methanol, 10% acetic acid) and detected by Coomassie Brilliant Blue (0.25% CBB R-250 in fixing solution) staining for 30 minutes to 1 hr. The gels were destained in the fixing solution and scanned using Gel Doc.

### **2.11 Protein concentration:**

The peak fractions of the elutions from the gel filtration column was pooled and concentrated using the following methods-

#### **2.11.1 Centricon concentration method:**

A 10 kDa cut-off centricon from Millipore was used to concentrate the protein and the final concentration of the protein was measured using Bradford assay.

#### **2.11.2 Lyophilization:**

Lyophilization is carried out using a simple principle of physics called sublimation. Sublimation is the transition of a substance from the solid to the vapor state, without first passing through an intermediate liquid phase. To concentrate the protein for crystallization purpose the fractions was pooled and lyophilized in *LABCONCO*. After lyophilization, the dried sample was dissolved in 50 mM of Tris-Cl (pH 7.5), 10 mM of CaCl<sub>2</sub> and the protein concentration was measured using Bradford assay.

### **2.12 Protein Estimation:**

The amount of protein in a sample was estimated by the Bradford Dye-Binding Protein Assay using BSA as the standard (Bradford, 1976). The micro assay was done in Spectramax-M2 and the concentration of protein is determined. This method involves the measurement of the binding of Coomassie Brilliant Blue G-250 dye to proteins and measurement of the absorbance of the dye-protein complex at 595 nm. Its advantages as a protein concentration measuring procedure over other chemical methods as it is fast and only requires one reagent, its color intensity is relatively stable over a period of an hour so more flexibility in reading the absorbance is provided and it has fewer substances that interfere with it 5 µl of protein sample was mixed 250 µl of the working solution in a microtitre plate and incubated at 37°C for

25 min. The absorbance was taken at 595 nm. BSA of known concentration was used as standard.

### **2.13 Dynamic light scattering (DLS):**

Dynamic light scattering (DLS) is a method to measure the size of the particles in solution in a size range not accessible for a light microscope. This method exploits the fact that the particles contained in the liquid undergo Brownian motion. This average velocity of this motion is determined by the particle size, their thermal energy and by the viscosity of the medium. This dependence is described by the Stokes-Einstein equation. It is thus enough to measure the particle velocity. For doing so, a laser beam is directed through the volume containing the particles, so that the light is scattered by the particles. Because the laser light is coherent, the scattered light forms an interference pattern. When the particles are in motion, this pattern changes permanently, so that the particle movement is transferred to an ever-changing light intensity distribution. A detector collects a fraction of the scattered light under a certain angle, thus seeing permanent light intensity fluctuations. Now the fluctuation rate is measure for the particle velocity. The detector signal is used to calculate the so-called autocorrelation function, representing the temporal behavior of the intensity fluctuations in a mathematical form. From this function the particle's diffusion velocity is derived, so that finally an application of the Stokes-Einstein equation delivers the desired particle size. The DLS experiment was performed by using Spectroscatter 201 Dynamic Light scattering equipment (RiNA *NETZWERK RNA-Technologien*, Berlin).

The DLS experiments was carried out at different conditions as follows:-

- A) 50 mM Tris pH 7.5, 10 mM CaCl<sub>2</sub> and 150 mM of NaCl
- B) 50 mM Tris pH 7.5, 10 mM CaCl<sub>2</sub> and 300 mM of NaCl
- C) NMR condition (50 mM Tris pH 6.0, 30 mM CaCl<sub>2</sub>)
- D) 50 mM Tris pH 7.5, 10 mM CaCl<sub>2</sub> and 20% ethanol

### **2.14 Size exclusion chromatography:**

The molecular size and the multimeric form of EhCaBP1 was determined by subjecting the protein to size exclusion chromatography on Pharmacia Sephadex G-75, 16/60 (GE health care) in 150 mM NaCl, 5 mM CaCl<sub>2</sub>, Tris pH 7.5 buffer.

Approximately 2 mg of protein was injected and passed through the column at a flow rate of 0.5 ml/min. The column was calibrated using following molecular weight standards: bovine serum albumin (66 kDa), ovalbumin (45 kDa), chymotrypsinogen (25.6 kDa) and lysozyme (14.3 kDa).

### **2.15 Crystallization of EhCaBP1:**

The purified EhCaBP1 was concentrated to 30 mg/ml in 50 mM Tris buffer (pH 7.5) containing 10 mM CaCl<sub>2</sub> for crystallization trials. Initial crystallization experiments were performed at both 16°C and cold room (4°C) temperatures. It was crystallized in hanging drops by mixing equal volumes (5 µl) of protein with the precipitant solution containing 60% MPD, 5 mM CaCl<sub>2</sub> and 50 mM Acetate buffer pH 4.4. Hexagonal shaped crystals (200 × 200 × 100 µm<sup>3</sup>) of EhCaBP1 appeared at 16°C approximately after ten days.

### **2.16 Mass spectroscopy of EhCaBP1 crystals:**

Mass spectroscopy was done in an in-house source. EhCaBP1 crystals were washed in the mother liquor twice and mixed (1:1) with matrix (sinapinic acid in acetonitrile containing 0.1% TFA) on a target plate and left to dry, before analysis by MALDI-MS. The mass scale of the instrument was externally calibrated using calibration mixture [insulin (bovine), ubiquitin, Cytochrome C, myoglobin)]. Peak at 14.84 kDa (m/z) confirms the full-length protein. A doubly ionized (m/2z) peak is also present.

### **2.17 Data collection, processing and structure determination of EhCaBP1 crystals:**

The X-ray diffraction experiments were performed at 100 K with EhCaBP1 crystals mounted on cryoloops in mother liquor and flash frozen in liquid nitrogen. These crystals diffracted to 3 Å with in-house rotating anode generator (At NII or ICGEB, New Delhi) and same crystals diffracted at a higher resolution of 2.4 Å at synchrotron X-ray source beamline A1 of CHESS (Cornell High Energy Synchrotron Source). They belong to space group P6<sub>3</sub> (Table 2.1) with two molecules per asymmetric unit. The data sets were indexed, processed and scaled with HKL2000 program (Otwinowski *et al.*, 1997).

**Table-2.1 Crystallographic data-collection and refinement statistics of EhCaBP1.**

Data Set	EhCaBP1
<b>Crystallographic data</b>	
X-ray source	CHESS A1
Wavelength	0.977
Space group	P6 <sub>3</sub>
Unit Cell Parameters (Å)	
a = 95.250, b =95.250, c = 64.996	
<b>Resolution range(Å)</b>	50 - 2.4
R <sub>sym</sub> (%)	7.1
Completeness	91.2 (73.9)
Total no. of observations	56718
No. of unique observations	12484
Redundancy	4.5(1.7)
Average I/s(I)	15.5 (1.9)
Crystal mosaicity (°)	0.5
<b>Refinement</b>	
Resolution	50 - 2.4
R_factor (%)	25.7 (34.6)
Free_R factor (%)	28.2 (36.5)
Mean B_factor	66.4
Number of atoms	
Protein/Ca/water/others	1013/4/31/20
<b>RMS deviations</b>	
Bonds (Å)	0.009
Bond angles (°)	2.0
Dihedral angles (°)	21.6
Improper angles (°)	1.0
Cross validated error	0.41

Values in parentheses are for the last resolution shell.

Free R factor was calculated with a subset of 7.5% randomly selected reflections.

The structure was solved by molecular replacement with Phaser program (Storoni *et al.*, 2004) using the N-terminal domain of human calmodulin (1CLL) as the search model. This search molecule corresponded to half of the EhCaBP1 molecule. The structure was refined to 2.4 Å resolution by iterative model building either by “O” graphics package (Jones *et al.*, 1993) or the COOT graphics package (Emsley *et al.*, 2004) combined with conjugate-gradient minimization with bulk solvent correction in CNS (Brunger *et al.*, 1998).

In the model, the N-terminal domain of CaM, comprised two EF hand motifs connected by a short loop. After an initial solution was located, each EF hand motif was treated as rigid body for the initial refinement, followed by manual extension of the model guided by both 2Fo-Fc (better than 1.5 Å) and Fo-Fc (better than 3 Å) electron density maps. The final model refined well with good electron density (Fig.3.3C; Chapter-1, Page No. 65) and crystallographic  $R_{\text{factor}}$  and  $R_{\text{free}}$  (Table 2.1) values that are within the range of average values for structures refined at this given resolution (Kleywegt *et al.*, 1996).

Despite acceptable refinement statistics, electron density for the C-terminal half of the molecule was absent. At this stage, we reduced the sigma cut-off of the 2fo-Fc and Fo-Fc electron density maps in an attempt to build the missing C-terminal part of the molecule without success. The water molecules and acetate molecules were added manually where justified by hydrogen bonds and Fo-Fc electron density at  $\geq 3.0 \sigma$  contour level in the final stages of refinement (Table 2.1). The structure was validated and deposited in the Protein Data Bank (PDB code: 2NXQ).

### **2.18 Interhelical Angle determination:**

The interhelical angle was calculated using the program interhlx (Yap, 1995). The program uses the following convention to calculate the sign of the angle between two helices:

The two helices are taken to be positioned such that helix I is "in front of" helix II. Helix I (from N to C) is used to define a vertical vector. A second vertical vector is defined, with its tail at the C-terminus of helix II. The angle between helices I and II is the rotation required to align the head of the second vector with the N-terminus of helix II. The vector is rotated in the direction (clockwise or counter-clockwise) that produces an angle of less than 180 degrees. If the rotation is clockwise, the sign is positive. If it is counter-clockwise, the sign is negative.

### **2.19 Crystallization of EhCaBP1-Phe complex:**

The purified protein was concentrated to 30 mg/ml in 50mM Tris pH 7.5 buffer containing 10 mM  $\text{CaCl}_2$  and 2 mM phenylalanine. This mixture was kept for crystallization similar to native crystallization condition. The complex was crystallized in hanging drops by mixing equal volumes (3 to 5  $\mu\text{l}$ ) of the complex with

the precipitant solution containing 63 to 65% MPD, 5 mM CaCl<sub>2</sub> and 50 mM Acetate buffer pH 4. Rod shaped crystals (400×675×675 μm<sup>3</sup>) of EhCaBP1-Phe appeared at 16°C within 7-10 days.

## **2.20 Data collection and processing and structure determination of EhCaBP1-Phe complex:**

The X-ray diffraction experiments were done at 100 K with EhCaBP1-Phe crystals mounted on cryoloops in mother liquor and flash frozen in liquid nitrogen. These crystals diffracted to 2.9 Å resolution with an in-house rotating anode generator (Advanced Instrumentation Facility, JNU). They belong to space group P6<sub>3</sub> (Table 2.2) with two molecules per asymmetric unit similar to native structure (Kumar *et al.*, 2007). The data sets were indexed, processed and scaled with Automar program.

The structure was solved by molecular replacement with Phaser program (Storoni *et al.*, 2004) using the native structure of EhCaBP1 (2NXQ) as the search model. The structure was refined to 2.4 Å resolution by iterative model building by the COOT graphics package (Emsley *et al.*, 2004) combined with conjugate-gradient minimization with bulk solvent correction in CNS (Brunger *et al.*, 1998).

The structure looked similar to native EhCaBP1 structure except large Fo-Fc density at the interface of EF1 and EF2 of the assembled domain to accommodate Phenylalanine. The final model refined well with good electron density and bound Phe (Fig.4.1C, Chapter II, Section-4.3.1) and crystallographic R\_factor and R\_free (Table 2.2) values that are within the range of average values for structures refined at this given resolution (Kleywegt *et al.*, 1996). Despite acceptable refinement statistics, electron density for the C-terminal half of the molecule was absent similar to the native structure. The water molecules, acetate molecules and phenylalanine molecules were added manually where Fo-Fc electron density at 3.0σ contour level and justified by hydrogen bonds or hydrophobic interactions in the final stages of refinement (Table 2.2).

**Table-2.2 Crystallographic Data-Statistics of EhCaBP1-Phe complex.**

Data Set	EhCaBP1-Phenylalanine
<b>Crystallographic data</b>	
X-ray source	Microstar
Wavelength	1.5418
Space group	P6 <sub>3</sub>
Unit Cell Parameters (Å)	
a = 95.201, b =95.201, c = 64.287	
<b>Resolution range(Å)</b>	50 - 2.8
R <sub>sym</sub> (%)	6.1 (7.33)
Completeness	97.9 (86.1)
Total no. of observations	15536
No. of unique observations	8289
Redundancy	8.9(5.4)
Average I/s(I)	51.92 (2.07)
Crystal mosaicity (°)	0.5
<b>Refinement</b>	
Resolution	50 - 2.8
R_factor (%)	25.7 (26.4)
Free_R factor (%)	28.5 (28.8)
Mean B_factor	93.4
Number of atoms	
Protein/Ca/water/Acetate/Phe	1024/4/42/2/2
<b>RMS deviations</b>	
Bonds (Å)	0.009
Bond angles (°)	1.5
Dihedral angles (°)	18.8
Improper angles (°)	0.73
Cross validated error	0.5

Values in parentheses are for the last resolution shell.

Free R factor was calculated with a subset of 7.5% randomly selected reflections.

### 2.21 Mass spectroscopy of N-terminal domain of EhCaBP1:

Mass spectroscopy was done in an in-house Bruker Daltonics flexAnalysis instrument of type autoflex TOF-TOF. The purified protein solution was mixed (1:1) with matrix (sinapinic acid in acetonitrile containing 0.1% TFA) on a target plate and left to dry, before analysis by MALDI-MS. The mass scale of the instrument was externally calibrated using calibration mixture [Bradykinin, Angiotensin I, Angiotensin II, Substance P, Bombesin, Renin Substance, ACTH clip and Somatostatin].

### 2.22 Crystallization of N-terminal domain of EhCaBP1:

The purified N-terminal protein was concentrated to 10-15 mg/ml in 50 mM Tris buffer (pH 7.5) containing 3 mM CaCl<sub>2</sub> and 20 mM of NaCl for crystallization trials. Initial crystallization experiments were performed at both 16°C and cold room (4°C) temperatures. It was crystallized in hanging drops by mixing equal volumes (3µl) of protein with the precipitant solution containing 15-18% PEG 400, 5mM CaCl<sub>2</sub>, 10% Isopropanol, 50 mM NaCl and 100 mM Acetate buffer pH 3.6- 4.2. Thin Needle appeared within 2 days. The rod shaped crystals (150×150×100 µm<sup>3</sup>) appeared at 16°C approximately after 5-10 days upon addition of a mixture of glycine, proline, glucose and sucrose from 1M stock to a final concentration of 2 mM and 100 mM of KCl solution.

### 2.23 Data collection, processing and structure determination of N-terminal domain of EhCaBP1:

The X-ray diffraction experiments were performed at 100K with N-terminal crystals mounted on cryo-loops in mother liquor and flash frozen in liquid nitrogen. These crystals diffracted to 2.5 Å with in-house rotating anode generator (Advanced Instrumentation Facility (A.I.F., Jawaharlal Nehru University, New Delhi, India). They belong to space group P3 (Table 2.3) with four molecules per asymmetric unit. The data sets were indexed, processed and scaled with Automar program (Marresearch).

The structure was solved by molecular replacement with Phaser program (Storoni *et al.*, 2004) using the native structure of EhCaBP1 (2NXQ) as the search model. The structure was refined to 2.5 Å resolution by iterative model building by the COOT graphics package (Emsley *et al.*, 2004) combined with conjugate-gradient minimization with bulk solvent correction in CNS (Brunger *et al.*, 1998). The structure (Chapter-III, Fig.5.5; Section 5.5.4) looked similar to native EhCaBP1 structure. The final model refined well with good electron density and crystallographic R\_factor and R\_free (Table 2.3) values that are within the range of average values for structures refined at this given resolution (Kleywegt *et al.*, 1996).

**Table-2.3 Crystallographic Data-Statistics of N-terminal domain of EhCaBP1.**

Data Set	N-terminal Domain
<b>Crystallographic data</b>	
X-ray source	Microstar
Wavelength	1.5418
Space group	P3
Unit Cell Parameters (Å)	
a = 89.589, b =89.589, c = 35.049	
<b>Resolution range(Å)</b>	25 - 2.5
R <sub>sym</sub> (%)	3.01(19.56)
Completeness	99.0 (99.5)
Total no. of observations	39681
No. of unique observations	10746
Redundancy	3.68
Average I/s(I)	10.3 (1.6)
Crystal mosaicity (°)	0.47
<b>Refinement</b>	
Resolution	25 - 2.5
R_factor (%)	23.3 (23.9)
Free_R factor (%)	27.1 (27.5)
Mean B_factor	51.0
Number of atoms	
Protein/Ca/water/	1970/8/68/
<b>RMS deviations</b>	
Bonds (Å)	0.012
Bond angles (°)	2.4
Dihedral angles (°)	20.5
Improper angles (°)	1.21
Cross validated error	0.5

Values in parentheses are for the last resolution shell.

Free R factor was calculated with a subset of 6.5% randomly selected reflections.

The water molecules were added manually where Fo-Fc electron density at 3.0  $\sigma$  contour level and justified by hydrogen bonds or hydrophobic interactions in the final stages of refinement.

#### **2.24 Preparation of Pb<sup>2+</sup>EhCaBP1 complex, Crystallization and Data collection:**

EhCaBP1 was over-expressed and purified as previously described (Kumar *et al.*, 2007). To the purified protein, 1M Pb(NO<sub>3</sub>)<sub>2</sub> solution was added to bring to the final

concentration of 5 mM. Upon addition of  $\text{Pb}^{2+}$ , the solution became turbid and protein was precipitated. The solution was partially clarified upon the addition of 1M sodium acetate pH 4.0 to a final concentration of 100 mM. Finally, the solution became clear with addition of 2-methyl-2, 4-pentandiol (MPD) to a final concentration of 10%. The EhCaBP1-Pb complex was centrifuged to 12,000 rpm to remove any precipitate and the supernatant was used for crystallization.

Hanging-drop vapor-diffusion crystallization trials were performed similar to native protein crystallization condition (Kumar *et al.*, 2007). Crystals suitable for diffraction were obtained by mixing 5  $\mu\text{l}$  of EhCaBP1-Pb complex and 2  $\mu\text{l}$  of the reservoir solution containing 60-65% MPD, in 50mM sodium acetate pH 3.6. Rod shaped crystals ( $250 \times 250 \times 150 \mu\text{m}^3$ ) appeared after 7-10 days of equilibration (Fig.6.2, Chapter-IV; Section-6.3.1).

The X-ray diffraction experiments were done at 100 K with EhCaBP1-Pb crystals mounted on cryoloops in mother liquor and flash frozen in liquid nitrogen. These crystals diffracted to 2.8 Å with in-house Bruker Microstar rotating anode generator and MAR imaging plate (Advanced Instrumentation Facility, Jawaharlal Nehru University, New Delhi). The crystals belong to  $P6_3$  space group (Table-2.4) with two molecules per asymmetric unit similar to native structure (Kumar *et al.*, 2007). The data sets were indexed, processed and scaled with Automar program (Marresearch).

**Table-2.4 Crystallographic data-collection and refinement statistics of Pb<sup>2+</sup>-EhCaBP1 complex**

Data Set	Pb <sup>2+</sup> -EhCaBP1complex
<b>Crystallographic data</b>	
X-ray source	Microstar
Wavelength	1.5418
Space group	P6 <sub>3</sub>
Unit Cell Parameters (Å)	
a = 95.02, b =95.02, c = 64.27	
<b>Resolution range(Å)</b>	50 - 2.8
R <sub>sym</sub> (%)	0.035 (0.047)
Completeness	99.8 (100.0)
Total no. of observations	56365
No. of unique observations	8180
Redundancy	6.8(6.9)
Average I/s(I)	8.7 (0.7)
Crystal mosaicity (°)	0.5
<b>Refinement</b>	
Resolution	50 - 2.8
R_factor (%)	26.54(26.64)
Free_R factor (%)	28.87 (28.93)
Mean B_factor	89.9
Number of atoms	
Protein/Pb/water	1024/4/43
<b>RMS deviations</b>	
Bonds (Å)	0.009
Bond angles (°)	1.7
Dihedral angles (°)	18.8
Improper angles (°)	0.73
Cross validated error	0.49

Values in parentheses are for the last resolution shell.

Free R factor was calculated with a subset of 7.5% randomly selected reflections.

### 2.25 Preparation of Ba<sup>2+</sup>-EhCaBP1 complex, Crystallization and data collection:

Unlike lead, which can easily replace calcium from the calcium-binding loops of CaM, barium ions cannot replace it, in all EF hand motifs. High concentration of barium is required for the activation of CaM compared to lead (Yamazaki *et al.*, 1996). This is basically due to low binding affinity of barium compared to calcium. To replace all the bound calcium ions from the calcium binding loops, the protein has to be unfolded and the calcium ions has to be removed through dialysis and then the protein has to be refolded in presence of excess of barium ions.

The EhCaBP1 protein solution was denatured with 8M urea to remove the tightly bound calcium ions and further dialyzed against 10mM Barium chloride in Tris buffer (pH 7.5) several times to remove urea. This protein was concentrated to 20-25 mg/ml and used for crystallization. Hanging-drop vapor-diffusion crystallization trials were performed similar to native condition. Crystals suitable for diffraction were obtained by mixing 5  $\mu$ l of EhCaBP1-Ba complex and 5  $\mu$ l of the reservoir solution (60-65% MPD, 50 mM sodium acetate, pH 3.8). Rod shaped crystals with oiling was appeared after 15-20 days of equilibration (Fig.6.4, Chapter-IV; Section No. 6.3.3, Page No. 106).

**Table-2.5 Crystallographic data-collection and refinement statistics of Ba<sup>2+</sup>-EhCaBP1 complex**

Data Set	Ba <sup>2+</sup> -EhCaBP1complex
<b>Crystallographic data</b>	
X-ray source	Microstar
Wavelength	1.5418
Space group	P6 <sub>3</sub>
Unit Cell Parameters (Å)	
a = 94.19, b =94.201, c = 64.947	
<b>Resolution range(Å)</b>	50 - 3.2
R <sub>sym</sub> (%)	0.082 (0.686)
Completeness	99.2 (98.1)
Total no. of observations	56365
No. of unique observations	8180
Redundancy	5.0(5.0)
Average I/s(I)	9.62 (1.43)
Crystal mosaicity (°)	0.5
<b>Refinement</b>	
Resolution	50 - 3.2
R_factor (%)	26.1 (26.4)
Free_R factor (%)	28.1 (28.8)
Mean B_factor	89.9
Number of atoms	
Protein/Ba/water	1024/6/42
<b>RMS deviations</b>	
Bonds (Å)	0.009
Bond angles (°)	1.7
Dihedral angles (°)	189.9
Improper angles (°)	0.83
Cross validated error	0.49

Values in parentheses are for the last resolution shell.

Free R factor was calculated with a subset of 7.5% randomly selected reflections.

The X-ray diffraction experiments were done at 100 K with EhCaBP1-Ba crystals mounted on cryoloops in mother liquor and flash frozen in liquid nitrogen stream. These crystals diffracted to 3.2 Å with in-house rotating anode generator (Advanced Instrumentation Facility, Jawaharlal Nehru University, New Delhi). They belong to space group  $P6_3$  with two molecules per asymmetric unit similar to native structure (Kumar *et al.*, 2007). The data sets were indexed, processed and scaled with Automar program (Marresearch). Table 2.5 represents the crystallographic data statistics for  $Ba^{2+}$ -EhCaBP1 complex.

### **2.26 Preparation of $Sr^{2+}$ -EhCaBP1 complex, Crystallization and data collection:**

Strontium has lower affinity than calcium for the calcium-binding site and it cannot replace the bound calcium ions. EhCaBP1-Sr complex was prepared following similar protocol used to prepare EhCaBP1-Ba complex. The EhCaBP1-Sr complex was concentrated to 20-25 mg/ml and taken for crystallization trials.

Hanging-drop vapor-diffusion crystallization trials were performed similar to native condition. Crystals suitable for diffraction were obtained by mixing 3 µl of EhCaBP1-Sr complex and 3µl of the reservoir solution (60-65 % MPD, 50mM sodium acetate pH 3.6). Hexagonal crystals appeared after 15-20 days of equilibration (Fig.6.6 Chapter-IV; Section-6.3.5). The X-ray diffraction experiments were done at 100K with EhCaBP1-Sr complex crystals mounted on cryoloops in mother liquor and flash frozen in liquid nitrogen stream. These crystals diffracted to 3.0 Å at beamline BL 5.2 R, XRD1, Elettra synchrotron source, Trieste, Italy. They belong to space group  $P6_3$  (Table 2.6) with two molecules per asymmetric unit similar to native structure (Kumar *et al.*, 2007). The data sets were indexed, processed and scaled with HKL2000 program (Otwinowski *et al.*, 1997).

**Table-2.6 Crystallographic data-collection and refinement statistics of Sr<sup>2+</sup>-EhCaBP1 complex.**

Data Set	Sr <sup>2+</sup> -EhCaBP1complex
<b>Crystallographic data</b>	
X-ray source	Elettra
Wavelength	1.5418
Space group	P6 <sub>3</sub>
Unit Cell Parameters (Å)	
a = 95.43, b =95.43, c = 63.97	
<b>Resolution range(Å)</b>	50 - 3.0
R <sub>sym</sub> (%)	0.068 (0.045)
Completeness	94.9 (96.8)
Total no. of observations	11558
No. of unique observations	6442
Redundancy	4.0 (4.0)
Average I/s(I)	18.8 (2.81)
Crystal mosaicity (°)	0.5
<b>Refinement</b>	
Resolution	50 - 3.0
R_factor (%)	25.9 (26.4)
Free_R factor (%)	28.6 (28.8)
Mean B_factor	88.9
Number of atoms	
Protein/Sr/water	1024/4/49
<b>RMS deviations</b>	
Bonds (Å)	0.009
Bond angles (°)	1.4
Dihedral angles (°)	18.4
Improper angles (°)	0.82
Cross validated error	0.5

Values in parentheses are for the last resolution shell.

Free R factor was calculated with a subset of 6.5% randomly selected reflections.

### 2.27 Structure determination of EhCaBP1-heavy metal complexes:

All the three complex structures were solved by molecular replacement with Phaser program (Storoni *et al.*, 2004) using the native structure of EhCaBP1 (2NXQ) as the search model. The structure was refined by iterative model building by the COOT graphics package (Emsley *et al.*, 2004) combined with conjugate-gradient minimization with bulk solvent correction in CNS (Brunger *et al.*, 1998). Calcium atom binding sites shown high difference fourier map in all the three complexed

structures. The barium bound structure had higher difference fourier map ( $6\sigma$ ) compared to other complex structures, where as strontium bound structure had low difference fourier map with 3.5 sigma level, following the trend of molecular weight of heavy atoms bound. The R-factors did not converge till the replacement of respective heavy atoms in the place of calcium. The structure looked similar to native EhCaBP1 structure. The final model refined well with good electron density and crystallographic  $R_{\text{factor}}$  and  $R_{\text{free}}$  values that are within the range of average values for structures refined at this given resolution (Kleywegt *et al*, 1996). Despite acceptable refinement statistics, electron density for the C-terminal half of the molecule was absent similar to the native structure. Very few water molecules were added manually where Fo-Fc electron density at  $\geq 3.0 \sigma$  contour level and justified by hydrogen bonds in the final stages of refinement.

### **2.28 Anomalous signal analysis:**

To confirm the presence of heavy atoms at the EF hand motifs, the anomalous signal was checked. Barium and lead atoms have  $f''$  about 9 electrons at Cu  $K\alpha$  wavelength of 1.54 Å, where as strontium has  $F''$  of 1.845 electrons at 1.2 Å. The anomalous signal was calculated using phenix.xtriage (Zwart *et al.*, 2005) for all data sets.

### **2.29 Phasing with Anomalous signal:**

The structure of EhCaBP1 was attempted to solve using SAD technique exploiting the anomalous signal of barium ions and lead ions independently. Both the structures were tried to solve using the AutoSol and AutoBuild protocol of PHENIX (Adams *et al.*, 2002) and to check the strength of phasing power for solving structure by automated direct method using home X-ray source.

# CHAPTER-I

---

**Crystal structure of Calcium Binding Protein-1 from  
*Entamoeba histolytica***

### 3.1 Abstract

Calcium plays a pivotal role in the pathogenesis of amoebiasis, a major disease caused by *Entamoeba histolytica*. Several EF-hand containing calcium-binding proteins (CaBPs) have been identified from *E. histolytica*. Even though these proteins have very high sequence similarity, they bind to different target proteins in a  $\text{Ca}^{2+}$  dependent manner, leading to different functional pathways (Bhattacharya *et al.*, 2006; Chakrabarty *et al.*, 2004). The crystal structure of the *Entamoeba histolytica* Calcium Binding Protein-1 (EhCaBP1) has been determined at 2.4 Å resolution. The crystals were grown using MPD as precipitant and they belong to  $P6_3$  space group with unit cell parameters of  $a = 95.25 \text{ \AA}$ ,  $b = 95.25 \text{ \AA}$ ,  $c = 64.99 \text{ \AA}$  (Materials and Methods, Table 2.1, Section-2.17, Page no. 44). Only two out of the four expected EF-hand motifs could be modelled into the electron density map. The presence of three glycine residues in the central linker region connecting the two domains of the protein imparts flexibility and is the major cause of the disorder of the C-terminal half of the protein. The final model refined to R\_factor of 25.6% and Free\_R of 28.2% (Materials and Methods, Table 2.1, Section-2.17, Page no. 44). Unlike CaM, the first two EF-hand motifs in EhCaBP1 are connected by a long helix and form a dumbbell shaped structure. Owing to domain swapping oligomerization, three EhCaBP1 molecules interact in a head-to-tail manner to form a triangular trimer. The mass spectroscopic analysis of the EhCaBP1 crystals confirms the full-length protein crystals. Also, the trimeric nature of the protein is further confirmed by gel permeation and dynamic light scattering experiments. This trimeric arrangement allows the EF-hand motif of one molecule to interact with that of an adjacent molecule to form a two EF-hand domain similar to that seen in the N-terminal domain of the NMR structure of CaBP1, calmodulin and troponin C. The oligomeric state of EhCaBP1 results in reduced flexibility between domains and may be responsible for the more limited set of targets recognized by EhCaBP1.

### 3.2 Introduction

*Entamoeba histolytica* is the etiological agent for human amoebic colitis and liver abscess and causes a high level of morbidity and mortality worldwide particularly in the developing countries. Calcium is a ubiquitous intracellular signal responsible for controlling numerous cellular processes in wide spectrum of organisms. Cells respond to an extra-cellular stimulus by a transient change in intracellular  $\text{Ca}^{2+}$  - concentration, which in turn is sensed by calcium binding proteins (Berridge *et al.*, 2000). A number of studies show that  $\text{Ca}^{2+}$  and its binding proteins are centrally involved in amebic pathogenesis (Ravdin *et al.*, 1988). Pathogenesis involves penetration into human tissues, attachment of *E. histolytica* to the host cells followed by their cytolysis and phagocytosis. *E. histolytica* not only phagocytoses epithelial cells, but also red blood cells (Tsutsumi *et al.*, 1992), bacteria (Bracha *et al.*, 1982) and cells from the immune system (Ravdin *et al.*, 1988). Motility and phagocytosis of the parasite require dynamic cytoskeletal organization based on the activity of microfilaments and actin-binding proteins (Voigt *et al.*, 1999). Like eukaryotic cells, regulation of the cytoskeleton in protozoans is triggered by calcium and calcium binding proteins that are critical for the progression and completion of phagocytosis, growth and host-parasite interaction (Camacho, 2003; Burleigh *et al.*, 1998). In *Histoplasma capsulatum*, CBP, a calcium binding protein aids the yeast's survival when they are in a low-calcium environment, such as the phagolysosomal compartment within macrophages (Batanghari *et al.*, 1998). Calmodulin (CaM) is a highly conserved  $\text{Ca}^{2+}$  binding protein, which is ubiquitous and central in translating  $\text{Ca}^{2+}$  levels into physiological signals. CaM has been shown to be involved in regulation of cell growth in *Trypanosoma cruzi* and deletion of the CaM gene slows the procyclic parasite growth by about 50% (Eid *et al.*, 1991). In *Toxoplasma gondii*, depletion of calcium, addition of channel blockers and use of CaM inhibitors blocked the cell invasion process. CaM was found to concentrate at the apical end of *Toxoplasma* and may be involved in cytoskeletal rearrangements for cell entry (Pezzella *et al.*, 1997). Centrin, a  $\text{Ca}^{2+}$ -dependent cytoskeletal protein is essential for overcoming the G2/M check point, thus implicating calcium in *Leishmania* growth (Selvapandiyan *et al.*, 2001). Free calcium concentration is acutely elevated upon the engagement of phagocytic receptors of human neutrophils (Stendahl *et al.*, 1994). The calcium/CaM

signal is known to be involved in vacuole fusion, trafficking and exocytosis (Peters *et al.*, 1998; Colombo *et al.*, 1997).

Calcium activation events have been less studied in *E. histolytica*. Fibronectin-mediated adhesion in *E. histolytica* can modify cytosolic calcium concentration. This induces the formation of actin adhesion plates and focal contacts, which is a link between calcium signaling and cytoskeletal structures (Carbajal *et al.*, 1996). Also it has been shown that protein kinase C relocates from the cytosol to the membrane fraction in actively phagocytosing trophozoites (De Meester *et al.*, 1990). The ubiquitous calcium binding protein calmodulin has been shown biochemically to be present in *E. histolytica* (Munoz *et al.*, 1991). CaM has been implicated in many functions, such as channel activation, electron dense granule release, cell proliferation and pathogenic activity of *E. histolytica* (Ravdin *et al.*, 1982; Makioka *et al.*, 2001). However, the mechanism of CaM action is not known, as the corresponding gene has not yet been characterized. *E. histolytica* has an extensive signaling system unlike many other protozoan parasites. The extent of the signaling network can be gauged by the fact that this microorganism has a large repertoire of novel CaBPs and transmembrane kinases (Loftus *et al.*, 2005; Bhattacharya *et al.*, 2006) many of which possess four EF-hand motifs similar to Calmodulin (CaM), although none of them exhibit high sequence identity with CaM. CaMs are highly conserved; therefore it appears that there is no CaM gene in the *E. histolytica* genome.

To elucidate the mechanisms of Ca<sup>2+</sup> signalling in *E. histolytica*, various calcium-binding proteins have been cloned, expressed, and biochemically characterized. (Chakrabarty *et al.*, 2004; Prasad *et al.*, 1993; Gopal *et al.*, 1998; Nickel *et al.*, 2000; Sahoo *et al.*, 2004 ). One of the first identified four EF-hand motif containing Ca<sup>2+</sup>-binding protein from *E. histolytica* (EhCaBP1) is a 14.7 kDa (134 amino acid residues) protein, with no known homologue in the database (Prasad *et al.*, 1992). Though this protein has some structural similarity with CaM, it is functionally quite distinct and binds to different set of proteins compared to CaM (Yadava *et al.*, 1997)

Inhibition of the expression of the EhCaBP1 gene by regulatable antisense RNA expression results in loss of cell growth, suggesting that the gene is essential for *E. histolytica* (Sahoo *et al.*, 2003). Inducible expression of EhCaBP1 antisense RNA

demonstrated its role in cellular proliferation (Sahoo *et al.*, 2003). EhCaBP1 is involved in dynamic changes of the actin cytoskeleton (Sahoo *et al.*, 2004). EhCaBP1 directly interacts with F-actin and co-localize in phagocytic cups and in pseudopods with no significant change in the kinetics of *in vitro* polymerization of actin, indicating that EhCaBP1 does not affect filament treadmilling (Sahoo *et al.*, 2004). In addition, using atomic force microscopy; it was found that filaments of F- actin, polymerized in presence of EhCaBP1, were thinner (Sahoo *et al.*, 2004). These results indicate that EhCaBP1 may be involved in dynamic membrane restructuring at the time of cell pseudopod formation, phagocytosis and endocytosis, in a process mediated by direct binding of EhCaBP1 to actin, affecting the bundling of actin filaments (Sahoo, N. *et al.*, 2004).

Confocal microscopic studies suggest the cytoplasmic localization of the protein including pseudopods and beneath the plasma membrane (Sahoo, N. *et al.*, 2004). The enrichment of EhCaBP1 on pseudopods suggests its association with areas rich in actin cytoskeleton and its associated proteins. However, this enrichment pattern was not seen over the portion of the membrane not undergoing any ruffle. Therefore, it appears that the presence of actin in some of the areas underneath the plasma membrane ensures recruitment of EhCaBP1 or vice versa. When amoebic cells are incubated with RBCs, phagocytic cups can be seen prior to engulfment of the RBCs. EhCaBP1 is observed to be present around the cup along with actin. It is probable that EhCaBP1 is recruited early in the initiation of phagocytosis and is lost after phagocytic vesicles are formed, although association of actin with the vesicles continues. This suggests that EhCaBP1 may not play a direct role in actin filament production but rather in the higher organization of these filaments (Sahoo *et al.*, 2004). A number of calcium binding proteins are known to exert influence on the functioning of cytoskeleton. DAP-kinase, a calcium-regulated death-promoting kinase is known to bind actin filaments. One of the substrates of DAPk was identified as myosin light chain (MLC), the phosphorylation of which mediates membrane blebbing (Shohat *et al.*, 2002). The neuronal calcium sensors are a family of EF-hand-containing  $\text{Ca}^{2+}$ -binding proteins. Neurocalcin, a member of this family is an N-myristoylated calcium-binding protein that directly interacts with actin in a calcium-dependent manner (Mornet *et al.*, 2001). There is a possibility that neurocalcin delta may be involved in the control of clathrin-coated vesicle traffic (Ivings *et al.*, 2002).

CaM is known to participate in endocytosis and actin cytoskeleton organization (Geli *et al.*, 1998; Ohya *et al.*, 1994). Regulation of cytoskeletal organization by CaM is through modulation of PtdIns(4,5)P<sub>2</sub> levels and subsequently phospholipase D activity (Desrivieres *et al.*, 2002). The cytoplasm of *E. histolytica* has numerous granules. Calcium binding proteins have been identified in these granules that may have important role in granule discharge during cell killing (Nickel *et al.*, 2000). Therefore it is probable that EhCaBP1 modulates actin-mediated processes through the participation of other proteins in a complex network. It is also possible that EhCaBP1 may be able to pull the membrane-attached actin filaments away from the membrane or denucleate these into shorter filaments, thereby allowing ingestion of RBC as phagosomes, similar to some of the actin depolymerization factors, such as ADF/cofilins.

EhCaBP1 shares 29% sequence identity with the ubiquitous CaBP, Calmodulin (CaM). However, this protein is functionally distinct from CaM (Yadava *et al.*, 1997). EhCaBP1 is an essential protein, as down regulation of its expression blocks proliferation of the parasite (Sahoo *et al.*, 2003). A phagocytosis deficient *E. histolytica* mutant, L6, showed reduced expression of EhCaBP1, further confirming its involvement in phagocytosis (Hirata *et al.*, 2007). Detailed analysis showed the involvement of EhCaBP1 in different forms of endocytosis, such as pinocytosis and erythrophagocytosis (Sahoo *et al.*, 2004). EhCaBP1 is likely to participate in the initiation step of endocytosis as it associated transiently with phagocytic cups and was not found in phagosomes (Jain *et al.*, 2008). Interestingly, the recruitment of EhCaBP1 to the phagocytic cups was not dependent on its ability to bind calcium. The mechanism by which EhCaBP1 is recruited to the phagocytic cups is not yet clear, although its ability to bind both F- and G- actin directly has been demonstrated (Sahoo *et al.*, 2004).

The three dimensional structure of EhCaBP1 was determined using multidimensional nuclear magnetic resonance (NMR) spectroscopic techniques in the Ca<sup>2+</sup> bound form (Atreya *et al.*, 2001) and a recent preliminary crystallographic analysis of another calcium binding protein (EhCaBP2) was reported from our group (Gourinath *et al.*, 2004). Despite low sequence identity between EhCaBP1 and calmodulin (CaM), the NMR structure of EhCaBP1 indicated structural similarity to CaM and troponin C (TnC). Structural studies show that CaM molecule has two lobes

separated by eight turn central helix, each lobe consisting of two EF-hand motifs (Babu *et al.*, 1985; Wilson *et al.*, 2000).  $\text{Ca}^{2+}$  binding induces an opening of each lobe and exposure of hydrophobic pockets, which are the target binding sites. Additionally, CaM also undergoes a large structural transition upon binding to targets, by bringing both EF-hand domains into close proximity and collapsing around an amphiphilic  $\alpha$ -helix in the target. Melting or unwinding of central helix allows this large conformational rearrangement of lobes, and extensive protein flexibility accounts for the ability to interact with various targets in a sequence dependent manner (Meadow *et al.*, 1993; Hoeflich *et al.*, 2002).

### $\text{Ca}^{2+}$ -binding loop position

Protein	Sites	X		Y		Z		-Y		-X		-Z	
		1	2	3	4	5	6	7	8	9	10	11	12
<b>EhCaBP1</b>	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
<b>Calmodulin</b>	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
<b>Troponin C</b>	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

**Scheme 1.** *EF-hand motif sequence alignment from EhCaBP1, Calmodulin and Troponin C. The calcium-binding loop positions has been indicated according to their consensus sequence.*

The previous biochemical studies have clearly shown that the calcium binding affinity of the EF3 and EF4 are much higher than that of the EF1 and EF2 (Gopal *et al.*, 1997). This strongly indicates that only EF1 and EF2 are affected by the  $\text{Ca}^{2+}$  concentration fluctuations around it and C-terminal domain (EF3 and EF4) should be

rigid and may not be influenced by  $\text{Ca}^{2+}$  concentration changes. An analysis of the EhCaBP1 primary sequence reveals the presence of four  $\text{Ca}^{2+}$ -binding sites, similar to those observed in other EF-hand CaBPs (Strynadka *et al.*, 1989; Nelson *et al.*, 1998). The specific sites labelled as X,Y, Z, -Y, -X, -Z in Scheme 1, refer to 1<sup>st</sup>, 3<sup>rd</sup>, 5<sup>th</sup>, 7<sup>th</sup>, 9<sup>th</sup> and 12<sup>th</sup> positions, respectively in the calcium binding loops of the protein, which coordinate to  $\text{Ca}^{2+}$  in a pentagonal bipyramidal geometry (Strynadka *et al.*, 1989).

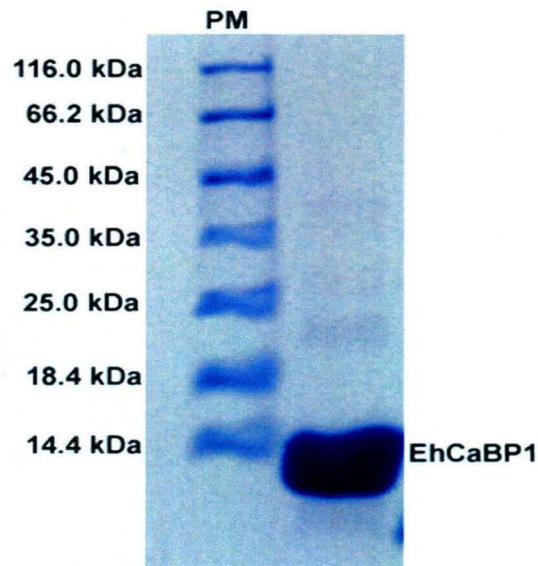
Extensive analyses of molecular databases including human genome show that EhCaBP1 has a limited overall sequence similarity with known CaBPs. The maximum identity at the amino acid level is 29% with calmodulins. In general, CaM belongs to a highly conserved protein family showing more than 65% sequence identity among members, irrespective of the species. Similarity in the protein sequences of EhCaBP and CaM is restricted mainly to EF-hand  $\text{Ca}^{2+}$ -binding loops with a negligible similarity in the functional central linker region.

To understand its flexibility and differential target recognition the structure of EhCaBP1 was determined by X-ray crystallography at 2.4 Å resolution.

### **3.3 Results and Discussion**

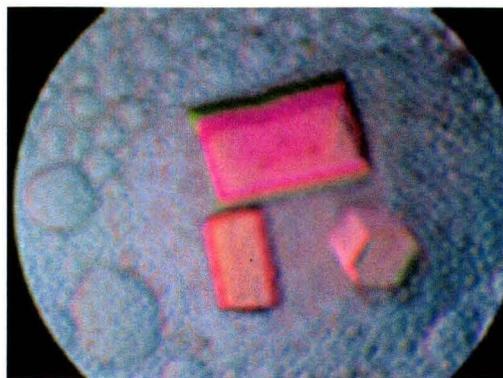
#### **3.3.1 EhCaBP1 purification and crystallization**

The protein was over-expressed and purified using ion-exchange chromatography as (Materials and Methods, Section No. 2.8.1, Page No. 38). The purity of protein was analysed on 12% SDS-PAGE gel (Fig.3.1). The band observed was corresponding to the size of full-length EhCaBP1 protein (i.e., 14.7 kDa). The purified protein was concentrated using centricon and was used for crystallization trials. Hanging drop method was used for crystallization and the crystals were obtained at 16°C using MPD as a precipitant.



**Figure-3.1** *Over-expression and purification of full-length EhCaBP1. SDS-PAGE (12%) analysis of the purification of EhCaBP1. A total of 10  $\mu$ l of the protein marker and 5  $\mu$ l of the purified and concentrated EhCaBP1 protein was loaded in the respective lane. The gel was stained with coomassie blue. The mobility of molecular-weight markers were as indicated. PM indicates the Protein Marker.*

The crystals of EhCaBP1 were obtained in MPD and were hexagonal in shape (Fig.3.2) (Materials and Methods, Section No. 2.15, Page No. 43).

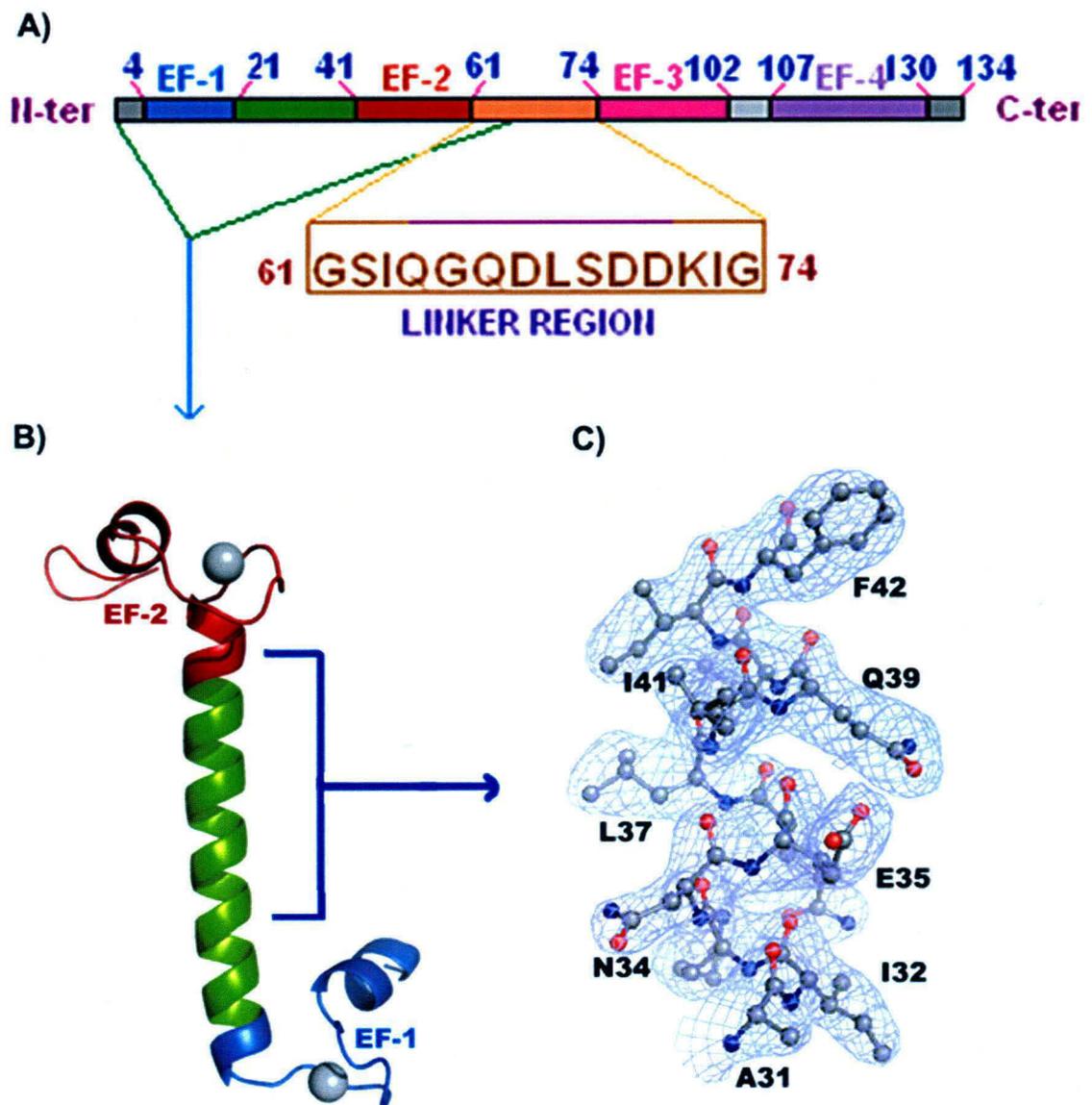


**Figure-3.2.** *Hexagonal crystals of EhCaBP1.*

The initial molecular replacement model for EhCaBP1 was the CaM N-terminal domain, where two EF-hand motifs are connected by a short loop and form a lobe. Further model building and refinement of EhCaBP1 reveals that these two interacting EF-hand motifs belong to two molecules related by crystallographic symmetry.

### 3.3.2 Overall crystal structure of EhCaBP1

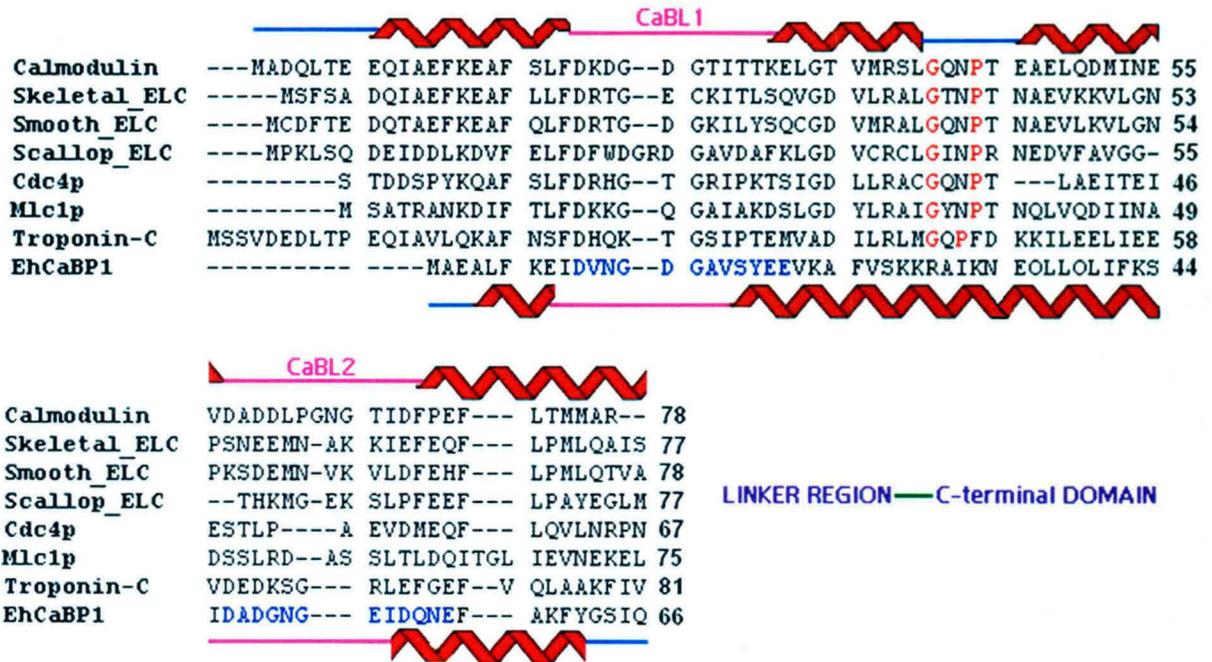
The final model reveals that the two EF-hand motifs in EhCaBP1 are connected by a straight helix (Fig.3.3), separating the two EF-hands by 44 Å and forming a dumbbell shaped structure, which is different from any known CaM-like proteins. The two EF-hand motifs belong to the N-terminal half of EhCaBP1, and each of them is bound to a calcium ion. The calcium coordination in both EF-hand motifs is very similar to that of CaM, with six ligands coming from protein and one ligand from water. The C-terminal half of the molecule, which has two more EF-hand motifs and is connected by a linker to the N-terminal part of the molecule, could not be modelled because the electron density for this region of the protein was absent. The connecting linker has three glycines and these may be responsible for the disorder in this region (Fig.3.3). Extensive flexibility and disorder of the C-terminal domain was observed in the NMR structure of EhCaBP1, although both structures differ in the relative arrangement of EF-hand motifs. The large-scale domain disorder observed for EhCaBP1 is not unique and several structures have been reported in the literature, including the CaM-Pb<sup>2+</sup> complex structure (Wilson *et al.*, 2000) that contain flexible and large domains (up to 55% of the molecule) which are missing in the final structures (Van Raaij *et al.*, 2001; Thomassen *et al.*, 2003). The final model of the EhCaBP1 molecule contains 66 out of 134 residues, two tightly bound calcium atoms, two acetate molecules, and 15 water molecules. In EhCaBP1, two EF-hand motifs are connected by a long helix. In contrast, the corresponding EF motifs in CaM are connected by a short loop, thus bringing these two EF-hand motifs into close proximity and forming a two EF-hand domain.



**Figure-3.3** Crystal structure of Calcium binding protein 1 from *E. histolytica* (*EhCaBP1*). A) Schematic representation of CaBP1, showing calcium binding loops (CaBL) with residue numbers (CaBL1 in blue and CaBL2 in red). The central linker region of the protein is displayed in orange, which has three glycines imparting flexibility to the molecule. (B) Structure of N-terminal half of the CaBP1 traced in the crystal structure is displayed as a ribbon diagram generated by PyMOL, (DeLano, 2000) calcium's (grey) are shown as spheres in two EF-hand motifs. Two EF-hand motifs are connected by straight helix (green). (C) Connecting helix between two EF-hand motifs is well defined. Part of the helix is shown with 2Fo-Fc electron density map at  $2\sigma$ .

### 3.3.3 Sequence comparison of EhCaBP1

Sequence comparison of EhCaBP1 with other Ca<sup>2+</sup>-binding proteins of known structure indicates the probable explanation for this structural variation. In CaM and the essential light chains (ELC) of the myosin light chain family, the region between EF1 and EF2 contains



**Figure-3.4** Sequence homology of N-terminal half of EhCaBP1 with CaM and other homologous essential light chains (ELC) indicating secondary structures. Gly and Pro (indicated in bold) are conserved in between calcium binding loop1 (CaBL1) and calcium binding loop2 (CaBL2) in CaM, ELC's and troponin C. These residues are changed to Arg and Lys in EhCaBP1.

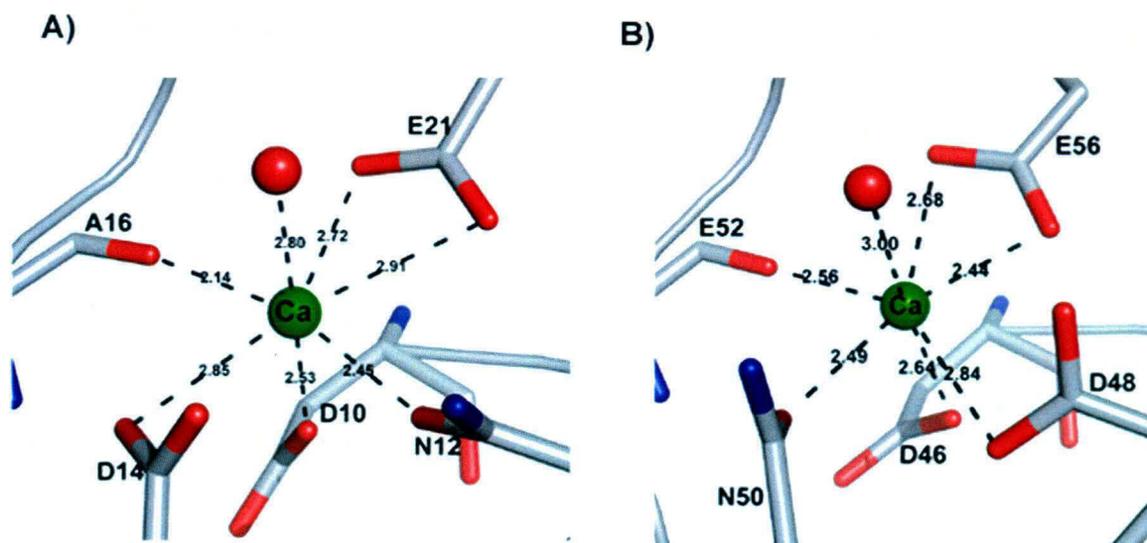
conserved Gly and Pro residues (Fig.3.4) that induce a bend in this region. These residues, which are helix breakers, are replaced by Arg and Lys in EhCaBP1, thereby causing the stabilization and straightening of the helix. This provides the probable explanation for the differences in the arrangement of EF-hand motifs in EhCaBP1 and CaM.

### 3.3.4 Solvent content

EhCaBP1 consists of 134 amino acids and was crystallized in  $P_{63}$  symmetry with two molecules in one asymmetric unit. With this molecular content, the unit cell has a solvent content of 57.2% and  $V_m$  of 2.9. If the molecule crystallized was only 66 residues long, as observed in the crystal structure, the unit cell would have 78.2% solvent content and a  $V_m$  of 5.7, which is unlikely given the typical solvent content of protein crystals (Kantardjieff *et al.*, 2003; Matthews, 1968) Even though there are few structures reported in the literature with more than 80% solvent content, the probability of crystallizing with such a high solvent content is low. To confirm that the crystals contain the full EhCaBP1 molecule, the crystals were washed with mother liquor, dissolved, and the molecular weight was determined using SDS-PAGE and MALDI mass spectrometry (Fig.3.6). The solvent content and the supporting biochemical data confirm that the crystals contain full-length EhCaBP1 with 134 amino acids, only 66 residues of which are traceable in electron density maps. The refinement statistics are also within the average values of structures at this resolution, (Kleywegt *et al.*, 1996) suggesting that the disordered residues do not contribute significantly to the diffraction data.

### 3.3.5 Calcium Binding loops of EhCaBP1

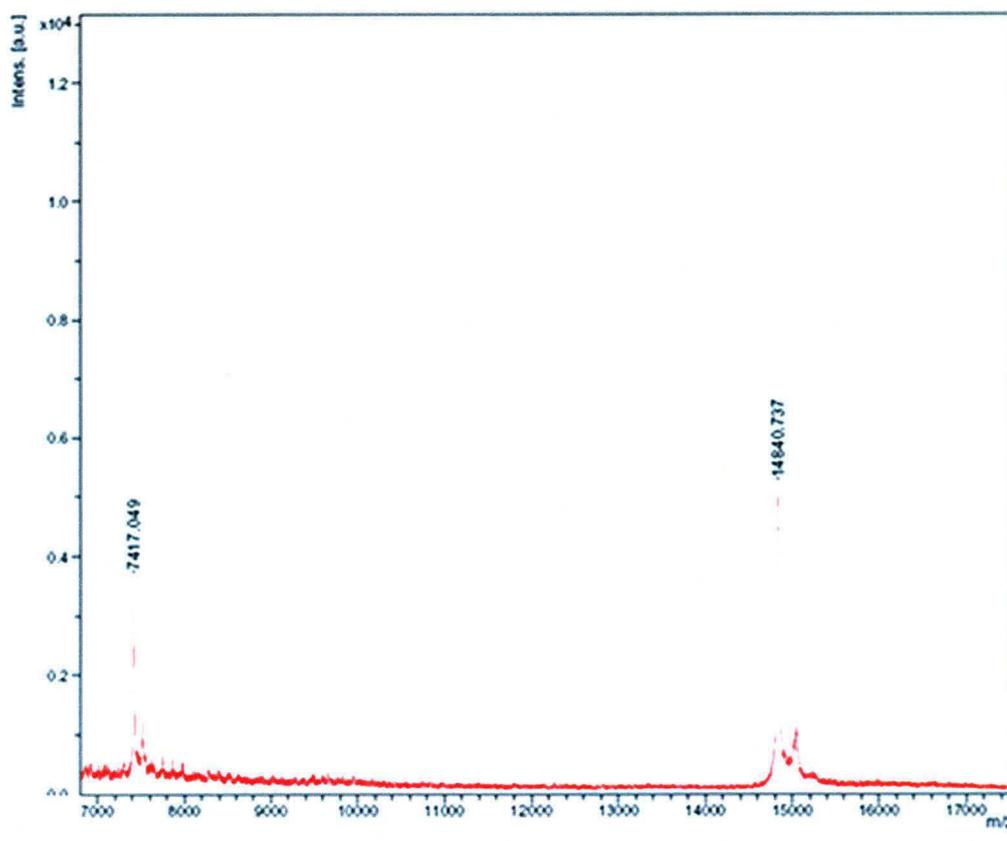
The nature of the co-ordination geometry of the calcium ion in EF-hand motifs are pentagonal bipyramidal. In each of the EF-hand motifs, one water molecule is also taking part in the interaction. The calcium ligands in EF-hand 1 motif are D10, D14, N12, A16 and E21 and the calcium ligands for EF-2 hand motif are D46, D48, N50, E52 and E57 respectively. The co-ordination distances have been shown for all the interacting partners (Fig.3.5). The average co-ordination distance for  $Ca^{2+}$ - ligand interaction in EF-hand 1 and EF-hand 2 motifs are 2.64 Å and 2.60 Å, respectively. The amino-acid side chains have been represented as sticks and calcium ion and water molecule as green and red sphere, respectively.



**Figure-3.5 Calcium coordination in EF-hand motifs of EhCaBP1.** The calcium coordination of the EF-hand 1 and EF-hand 2 motifs are shown with co-ordination distances. The calcium ion is interacting with seven oxygen ligands. In both the EF-hand motifs, one water molecule is also taking part to satisfy the coordination geometry. The calcium ion and the water molecule are represented by green and red spheres, respectively.

### 3.3.6 Mass spectrometry of EhCaBP1 crystals

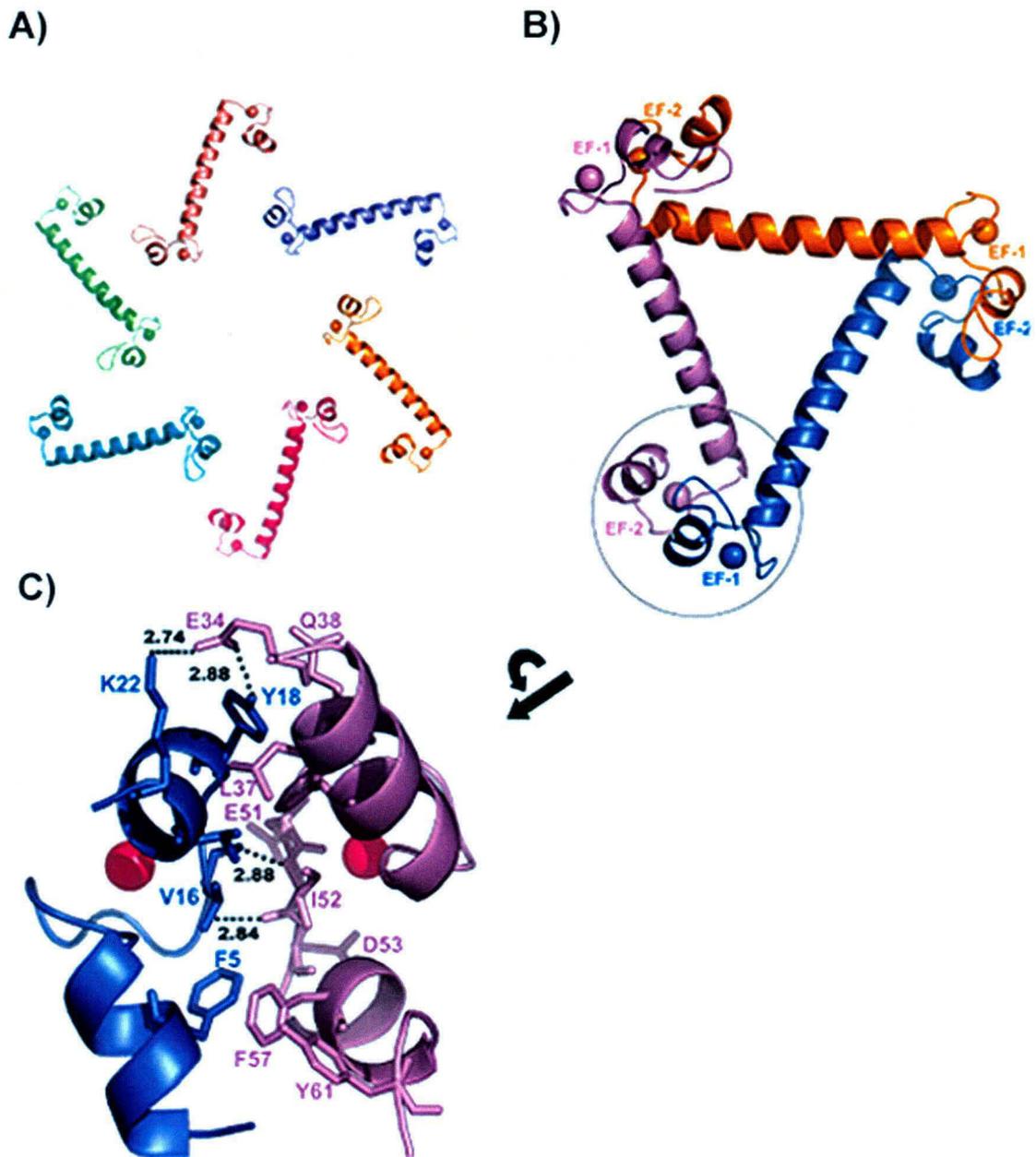
To confirm whether the protein molecule has not been cleaved from the flexible central linker region, the intact mass analysis of the EhCaBP1 crystals was done using mass spectrometry. The MALDI-TOF results of the EhCaBP1 crystals have clearly shown the peak at 14840 Dalton which corresponds to the molecular weight of the full length protein with bound calcium ions. This result further confirmed that the crystal was of the full length protein (Materials and Methods, Section- 2.16., Page No. 43)



**Figure-3.6** *Mass Spectroscopy of EhCaBP1 crystal. Peak at 14840 Dalton corresponds to the molecular wt. of full length protein bound with four calcium ions. This confirms that the protein was intact after crystallization.*

### 3.3.7 Symmetry interactions in domain swapped trimeric form of EhCaBP1

The N-terminal half of EhCaBP1 makes extensive contacts with symmetry-related molecules in the crystal. The unit cell contains six molecules that are positioned in such a way that the C-terminal region is pointing towards a void near the center of the cell [Fig.3.7 (A)] that resembles the N-terminal domain of CaM, TnC, and NMR structure of the same protein [Fig.3.7 (C)]. The EF1 motif of one molecule makes extensive hydrophobic and couple of hydrogen bond contacts with the EF2 motif of the symmetry-related molecule, thereby forming a domain. Three molecules of EhCaBP1 interact in a head-to-tail manner in the crystal to compose a triangular trimer [Fig.3.7 (B)]. The EF-hand motifs from three molecules interact at the three corners of the triangle to form domains containing two EF-hands.



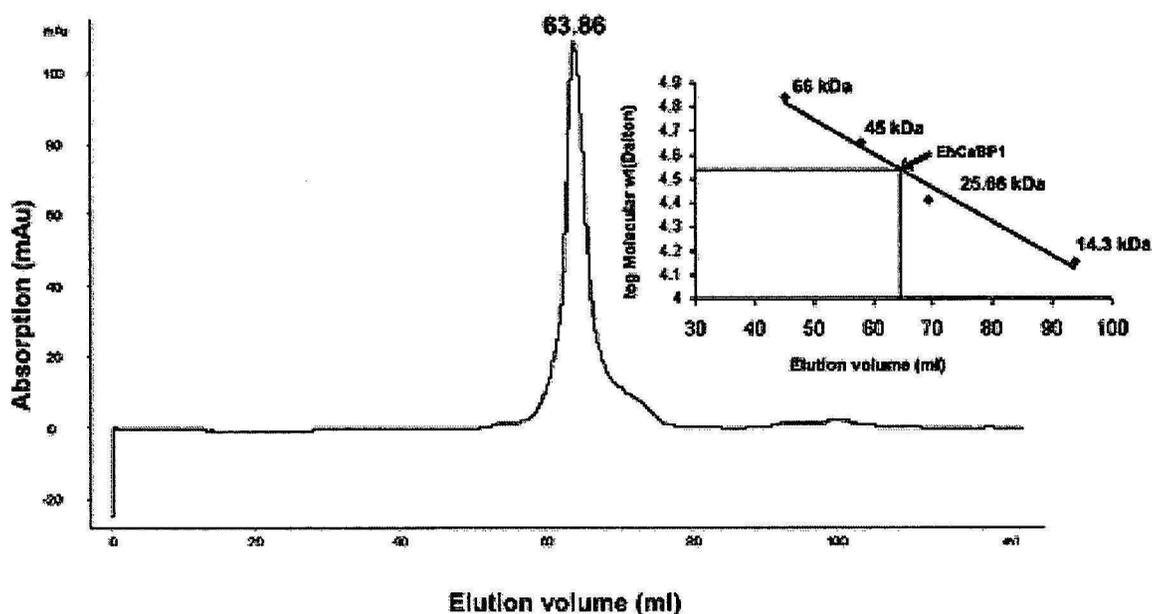
**Figure-3.7 Symmetry interactions and domain formation.** (A) The arrangement of six molecules of EhCaBP1 in the unit cell. (B) Interaction between three symmetry related molecules. Three molecules (in blue, purple and gold) are shown interacting with each other in head to tail manner. EF1 hand motif of one molecule makes extensive interactions with EF2 hand motif of symmetry related molecule to form a lobe/domain. (C) Magnified view of the assembled domain displaying the extensive interactions that stabilize the domain formation. The hydrophobic pocket formed at the interface of two molecules can also be seen.

### 3.3.8 Trimeric nature of EhCaBP1

To verify that the trimer formation and three domain arrangement is not an artefact of the crystal lattice, the multimeric form of the EhCaBP1 was confirmed using both size exclusion chromatography and DLS (Dynamic Light Scattering) experiments.

#### 3.3.8.1 Size exclusion chromatography

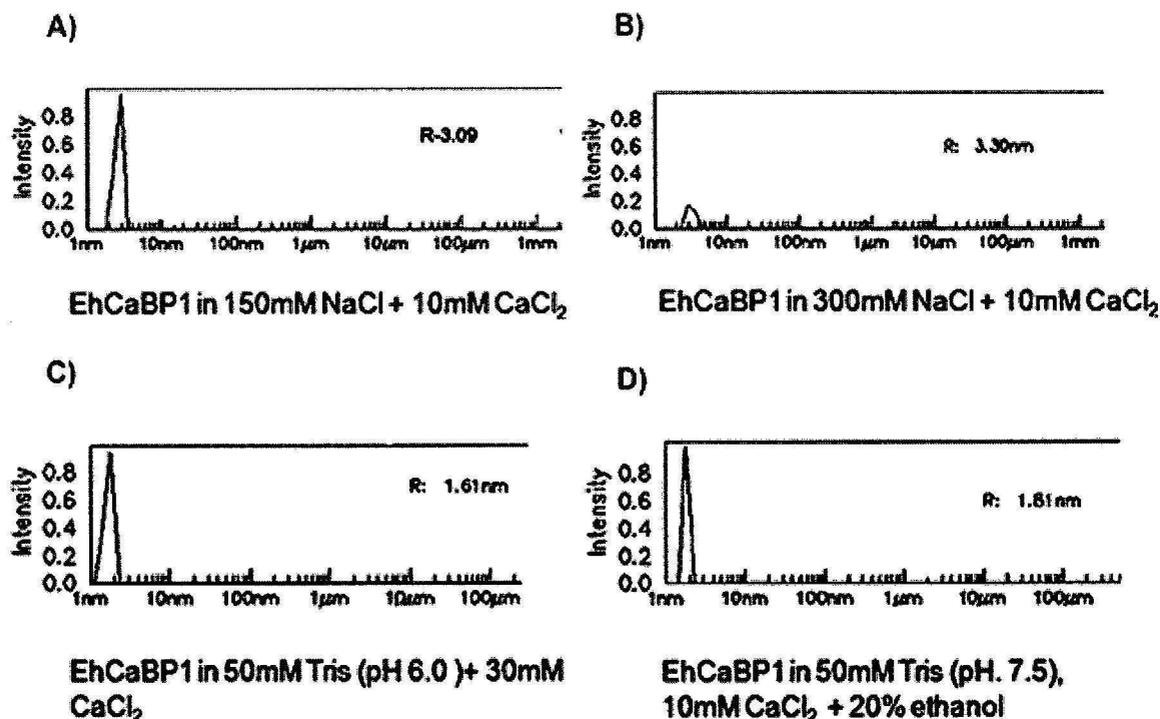
The multimeric form of this protein was further confirmed at physiological salt concentration of 150 mM NaCl by gel filtration. The peak of EhCaBP1 on SephadexG-75 was eluted at the 63.8mL volume corresponding to 34 kDa, (Fig.3.8) which corresponds to the molecular weight in between dimer and trimer.



**Figure-3.8** *Size exclusion chromatography of EhCaBP1.* The protein was passed through a Sephadex G-75, 16/60 column and 1 mL fractions were collected in each tube. The elution volume as well as the elution pattern of the protein is shown. The elution pattern of the gel filtration standards is marked in the inset as followed: bovine serum albumin (66 kDa), ovalbumin (45 kDa), chymotrypsinogen (25.6 kDa) and lysozyme (14.3 kDa). Molecular masses of the standards were plotted in logarithmic scale against the elution volume. Molecular mass of the eluted form of EhCaBP1 was deduced from the plot.

### 3.3.8.2 Dynamic Light Scattering

The molecular mass of EhCaBP1 obtained from DLS data was 44.2 kDa consistent with trimeric form of EhCaBP1 in Tris buffer (pH 7.5), 5 mM CaCl<sub>2</sub> containing 150 mM NaCl and 300 mM NaCl.



**Figure-3.9** *Dynamic Light Scattering experiment of EhCaBP1 in various conditions.* A) At physiological condition, a peak of radius 3.09 nm is obtained which corresponds to the trimeric form of the protein. B) The addition of excess of salt (300 mM NaCl) maintained the the trimeric form of protein C) The NMR experimental conditions has shown the monomeric form of the protein with a peak at 1.61 nM. D) The presence of 20% EtOH in the buffer resulted in the dissociation of the trimer form of the protein to its monomer with the peak at 1.81 nM.

While in low ionic strength buffer like Tris buffer pH 6.0, 30 mM CaCl<sub>2</sub> (NMR experimental condition) (Atreya *et al.*, 2001) or in the presence of 20% ethanol it exhibited a molecular mass of 14.1 kDa indicating the monomeric form of the protein.

### 3.3.9 Hydrophobic Core

The trimeric oligomerization of EhCaBP1 consequence of one EF1 motif of one molecule interact with the EF2 motif of another molecule to form a domain, which is similar to N-terminal or C-terminal domain of CaM. These two EF-hand domains form a hydrophobic core similar to the hydrophobic regions that are present in each of the two domains of CaM. In CaM, these hydrophobic regions are the primary binding surfaces for CaM targets. L4, F5, I8, V16, Y18, and V21 from EF1 were in close proximity to F57, I52, F41, and L37 from the EF2 of a symmetry-related molecule and form a hydrophobic core [Fig.3.7(C)]. V16 of EF1 and I52 of EF2 interact through backbone hydrogen bonds in such a way that they form an anti-parallel  $\beta$ -sheet. E34 of EF2 also forms a salt bridge with K22 and a hydrogen bond with Y18 of EF1. Overall, the two EF-hand motifs from two molecules of EhCaBP1 interact extensively and form a distinct domain that is reminiscent of the two domains of CaM and may be involved in target binding by EhCaBP1.

### 3.3.10 Comparison with calmodulin, Troponin C, and NMR structure

The individual EF-hand motifs of EhCaBP1 show a high structural similarity with their counterparts in CaM, TnC, and NMR structure of EhCaBP1. EhCaBP1 EF motifs are most similar to the EF-hands in CaM and least similar with the EF-hand motifs in the NMR structure of EhCaBP1 (Table 3.1).

The comparison of interhelical angles indicate that both the NMR structure and crystal structure of EhCaBP1 have a less open hydrophobic pocket as compared to the  $\text{Ca}^{2+}$ -bound CaM and TnC structures. The NMR structure of EhCaBP1, although, has a relatively less open hydrophobic pocket than the crystal structure (Table 3.1B). Notably, the overall arrangement of these EF-hands is completely different in the crystal structure of EhCaBP1 compared with all these structures. The NMR structure of EhCaBP1 is similar to the CaM structure, in that it has two globular domains connected by a flexible linker helix, and each globular domain comprises of two adjacent EF-hand motifs. Unlike CaM, the flexible linker does not form a helix in the NMR structure of EhCaBP1 but instead it is so flexible that the relative positions of the globular domains are ill defined (Atreya et al., 2001). In the crystal structure of EhCaBP1 reported here, the C-terminal region, which corresponds to 50% of the

molecule, is not visible in the electron density map. Moreover, the N-terminal half of EhCaBP1 forms a dumbbell shaped structure with the EF1 and EF2 motifs separated by an eight turn helix.

A) RMSD of individual EF-hand motifs and Assembled domain

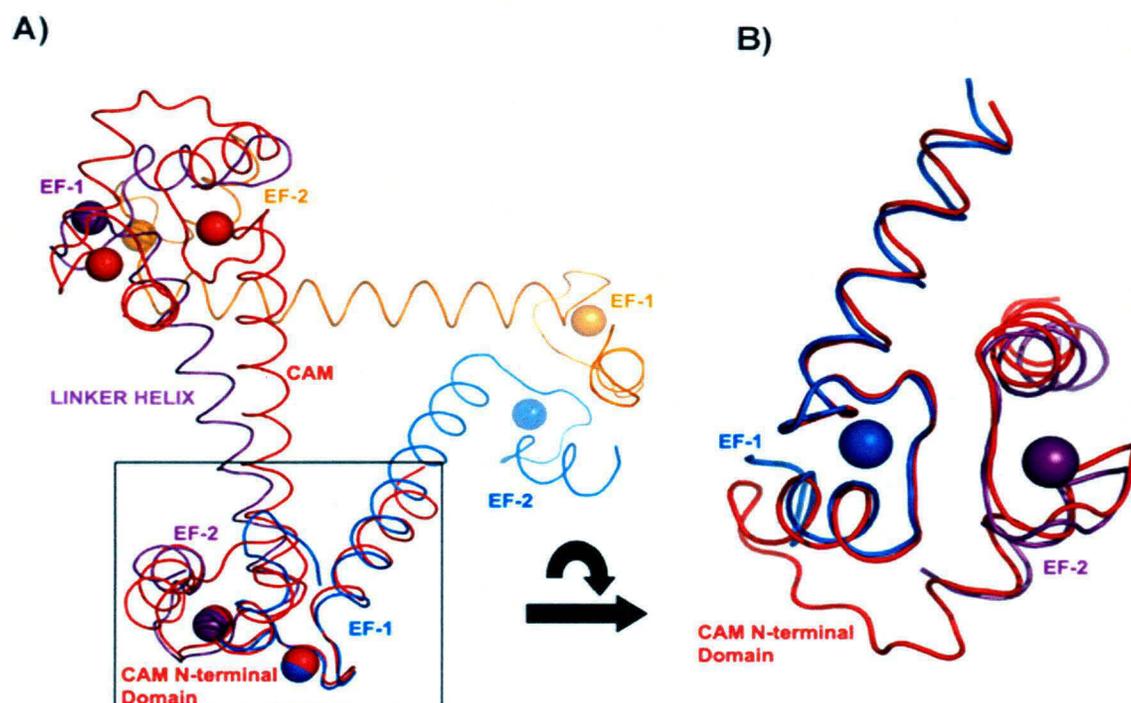
Molecule	EF1 (Å)	EF2 (Å)	Assembled domain
NMR (1JFK)	1.19	1.67	1.76 (42)
CaM (1CLL)	0.46	0.78	1.15 (50)
TnC (1TNX)	0.93	1.34	1.14 (36)

B) Comparison of Inter-helical Angles of EF-hand motifs

Molecule structure	I and II	III and IV	II and III	I and IV
Present structure	97	106	129 <sup>a</sup>	124 <sup>a</sup>
NMR (1JFK)	117	122	133	117
CaM (1CLL)	87	86	112	107
TnC (1TNX)	82	76	117	116

**Table 3.1 Comparison of the present crystal structure with N-terminal domain of NMR structure, CaM, and TnC.** C $\alpha$  RMS deviations for independent EF-hand motifs and with assembled domain are listed in the table. Inter helical angles (in degrees) were calculated using the algorithm of Kuboniwa et al., 1995 as implemented in the program “interhlx” by K. Yap. Angles are stated as “pening angle” onto the hydrophobic pocket, higher angle represents less open hydrophobic pocket. The angle between II, III, and I, IV are calculated in the assembled domain of EhCaBP1 crystal structure.

In contrast, CaM and other structurally related EF-hand proteins have these motifs closer, connected by a short loop to form a domain. However, the trimeric assembly of EhCaBP1 brings EF1 from one molecule and EF2 from another into close proximity in a domain-swapped manner to form a domain that is similar to the CaM N-terminal domain [Fig.3.10 (B)].



**Figure-3.10** *A superposition of CaM on trimer of EhCaBP1 that reveals the similarity of the assembled domain to CaM N-terminal domain. (A) The CaM N-terminal domain (red) was superposed (by least square fit) on assembled domain of EF1 of one molecule (light blue) and EF2 of neighbouring molecule (purple) highlights the domain similarity. The distance between N-terminal and C-terminal domains in CaM is approximately equal to the distance between the two assembled domains of EhCaBP1 trimer. But the linker helix in CaM bends about 158 and thus C-terminal domain of CaM do not overlap with the next assembled domain. (B). Magnified view of CaM N-terminal domain (red) superposed on assembled domain of EhCaBP1 trimer (EF1 in blue, EF2 in purple).*

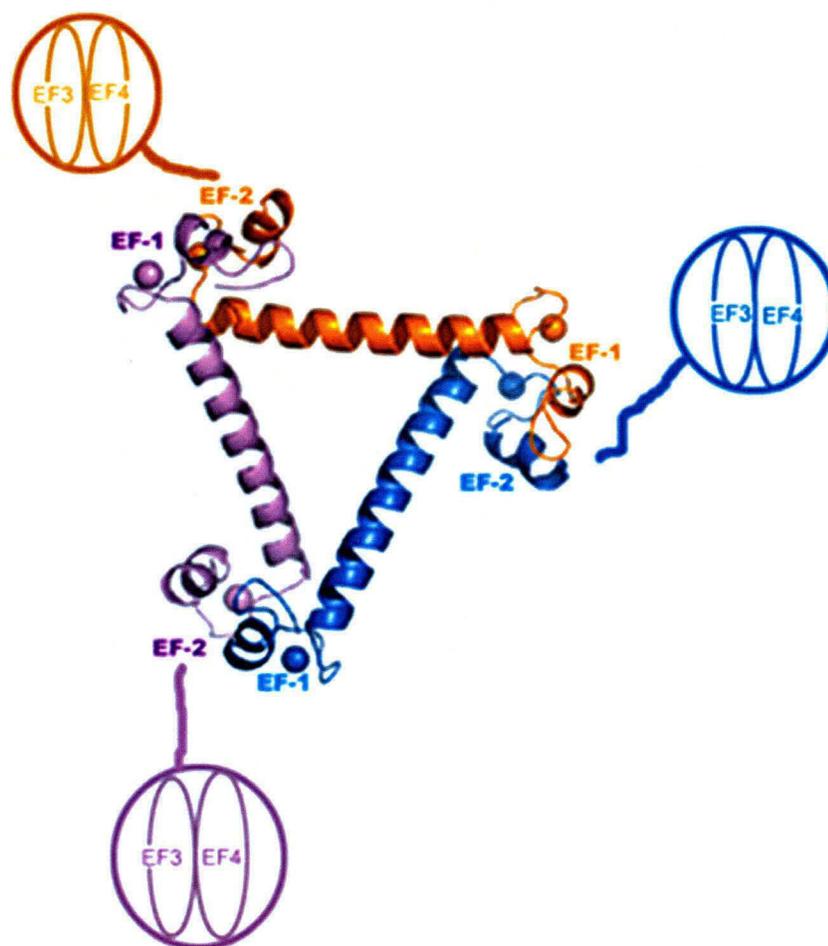
The assembled domain has RMSD of 1.15 Å (50 residues) with the N-terminal domain of CaM, 1.14 Å (36 residues) with the N-terminal domain of TnC and 1.76 Å (42 residues) with the N-terminal domain of NMR structure of EhCaBP1 (Table 3.1A).

The assembled domain closely resembles the N-terminal domain of CaM in its overall structure and hydrophobic core, suggesting that this domain may be interacting with other cellular targets at hydrophobic residues in a fashion that is

comparable to the CaM N-terminal domain. The different target binding specificities between CaM and EhCaBP1 may be due to the differences in their respective EF-hand linker regions. In CaM, the linker between C-terminal domain and N-terminal domain is highly flexible and can accommodate structurally diverse target regions which have hydrophobic residues separated by 10, 14, or 16 residues (Hoeflich *et al.*, 2002). In contrast, the trimeric quaternary arrangement of EhCaBP1, places severe constraints on the mobility and flexibility of these EF-hand domains and may therefore limit the range of targets that EhCaBP1 can bind. The distance between the assembled domains in the trimer and the two domains in the CaM is almost the same. The two assembled domains in trimer are separated by about 44 Å (distance between two calcium atoms in EF-hand motifs) whereas the two domains in CaM extended structure are 43 Å apart. In spite of this, interestingly when the assembled domain was superposed with the CaM N-terminal domain, the C-terminal domain of CaM deviates from the next assembled domain, because of the 158° bend in the central helix of CaM [Fig.3.10 (A)], whereas in EhCaBP1 trimer the helices connecting the domains are straight.

### 3.4 Conclusion

In summary, the 2.4 Å resolution crystal structure displays an unusual molecular organization and EF-hand domain formation in EhCaBP1. We have found that EhCaBP1 exists as trimer both in solution and in the crystal, where three molecules interact with each other in head-to-tail manner. Three molecules of EhCaBP1 N-terminal domains participate in domain swapping to form trimers led by couple of critical residues difference in between EF1 and EF2 motifs in comparison to CaM and ELC's (Fig.3.4). This further confirms the previous observations that one or a few critical amino acid replacements can lead to 3D domain-swapping (Rousseau F, *et al.*, 2003). Similar kind of domain swapping phenomenon was also observed in apo-CaM structure, where CaM forms dimer (Schumacher *et al.*, 2004) but the domain swapping has never been seen before in any calcium bound CaM or CaM related proteins.



**Figure-3.11** *Model of EhCaBP1 crystal structure.* The N-terminal domain forms a domain-swapped trimer where EF-1 of one molecule interacts with EF-2 of the symmetry-related molecule in a head-to-tail manner. The circle represents the C-terminal domain, which was not traced in the crystal structure, is supposed to be a globular domain containing EF-3 and EF-4 motifs. The connection of the flexible central linker has not been shown.

Variations in the buffer conditions like salt concentration and presence of a polar solvent seems to influence the molecular state (monomeric/trimeric) of EhCaBP1. DLS experiments clearly show that at around physiological conditions (150 mM NaCl and pH 7.5) this protein exists in trimeric form.

Thus it clearly shows that the trimeric state observed in the crystal structure, is the most probable functional and natural state of the protein compared with the monomeric form found in the NMR structure. In the trimeric form hydrophobic

pockets that may be involved in target binding are formed at the interface of two molecules. The hydrophobic pockets in EhCaBP1 are very similar to the ones seen in CaM and the distance between them in trimer of EhCaBP1 is almost equal to the distance between the hydrophobic pockets in the CaM extended structure. The functional difference between EhCaBP1 and CaM can be explained by the difference in the flexibility of the helix that connects the EF-hand domains. In EhCaBP1, three domains are connected by three helices, and these helices are constrained in triangular arrangement by oligomerization. In CaM, the two EF-hand domains are connected by a flexible linker helix that permits greater conformational flexibility and is largely responsible for CaM's ability to bind to a large array of targets. Taken together, these crystallographic and biochemical results expand the range of structural diversity observed for the EF-hand family and suggest that oligomerization of EhCaBP1 is both essential for target binding and may restrict the structural diversity of the target proteins to which EhCaBP1 can bind. These results represent fundamental advances in our understanding of EF-hand comprising proteins and their structure-function relationship. This opens up a new area to better understand the possibility of unconventional EF-hand motif arrangement in CaM like proteins and its functional implications.

# CHAPTER-II

---

**Crystal structure of EhCaBP1 in complex with phenylalanine: Visualization of mode of target binding**

#### 4.1 Abstract

The crystal structure of EhCaBP1 shows only the N-terminal domain whereas the C-terminal domain has not been traced due to the presence of flexible linker region connecting the two domains. Unlike the structure of CaM, the arrangement of EF-hand motifs of the N-terminal domain is unusual where the EF-hand 1 of one molecule interacts with EF-hand 2 of the symmetry related molecule in a domain swapped manner. The assembled-domain closely resembles the N-terminal domain of CaM in its overall structure and hydrophobic core, suggesting that this domain may be interacting with other cellular targets at hydrophobic residues in a fashion that is comparable to the CaM N-terminal domain. The trimeric quaternary arrangement of EhCaBP1 places severe constraints on the mobility and flexibility of these EF-hand domains and may therefore limit the range of targets that EhCaBP1 can bind. The functional studies of independent domains clearly show that N-terminal domain is the functional domain of EhCaBP1 in activating kinase and localization in phagocytic cup (Jain *et al.*, 2009). To understand the mode of target binding, EhCaBP1 was co-crystallized with Phenylalanine and the structural studies show that the Phe binds to the hydrophobic pocket of the assembled-domain. The crystal structure of the EhCaBP1-phenylalanine complex has been determined at 2.8 Å resolution. The crystals were grown using MPD as precipitant and they belong to P6<sub>3</sub> space group with unit cell parameters of a = 95.201 Å, b = 95.201 Å, c = 64.287 Å. The final model refined to R\_factor of 25.6% and Free\_R of 28% (Materials and Methods, Table 2.2, Section-2.20, Page No. 47). The Phe-bound assembled-domain was superimposed on the N-terminal domain of CaM, bound to hydrophobic IQ motif of cardiac Ca<sup>2+</sup> channel. Both structures superimposed very well with an RMS deviation of 1.23 Å. The hydrophobic residue Phe of the peptide (bound to CaM N-terminal domain) and Phe bound to EhCaBP1 assembled-domain are located at the similar regions.

## 4.2 Introduction

Crystal structure of EhCaBP1 showed an unusual arrangement of the domains of EhCaBP1 (Kumar, S., *et al.*, 2007). The N-terminal domains of three molecules of EhCaBP1 interact in a head to tail manner to form a domain-swapped trimer. In the trimeric form, hydrophobic pockets are formed at each interface, and inter-pocket distance is almost equal to the distance between the hydrophobic pockets in the extended structure of CaM. The region connecting EF hands I and II was found to be less flexible with extended conformation. On the other hand, the three glycine residues (G63, G67 and G76) present in the central linker region makes it more flexible as compared to CaM. Due to the flexibility of the central linker region, the C-terminal domain structure could not be traced.

The N-terminal domain of EhCaBP1 undergoes a major conformational change upon binding to calcium, similar to the full length protein (Jain R., *et al.*, 2009). EhCaBP1 is known to activate endogenous kinase(s) in a  $\text{Ca}^{2+}$  dependent manner (Yadava *et al.*, 1997). The endogenous kinase activation and the erythrophagocytic studies show that the N-terminal domain behaves like a full-length protein (Jain *et al.*, 2009). Further, the fluorescence microscopic studies show the presence of the N-terminal protein at the phagocytic cup and its complete co-localization with F-actin (Jain *et al.*, 2009). These studies clearly indicate that N-terminal domain is the functional domain of EhCaBP1 which performs most of the functions of the full-length protein.

### 4.2.1 Calmodulin: Target recognition and activation mechanism

CaM interacts with target proteins in a  $\text{Ca}^{2+}$ -dependent manner (Klee, 1988; Crivici *et al.*, 1995) that translates the second messenger calcium into a variety of cellular responses. Calcium binding to CaM exposes hydrophobic clefts in each of the CaM lobes that interact with downstream targets (Crivici *et al.*, 1995). This small 17-kDa acidic protein belongs to a family of homologous calcium-binding proteins that bind  $\text{Ca}^{2+}$  through the EF-hand motif (e.g., parvalbumin (Krietsinger *et al.*, 1973) or troponin C (Herzberg *et al.*, 1988)). Calmodulin (CaM) is a highly conserved  $\text{Ca}^{2+}$  binding protein, which is ubiquitous and central in translating  $\text{Ca}^{2+}$  levels into physiological signals. The crystal structures of  $\text{Ca}^{2+}$  bound CaM show that this protein

contains two similar domain structures linked by an extended  $\alpha$ -helix (Babu *et al.*, 1985; Babu *et al.*, 1988; Taylor *et al.*, 1991; Chattopadhyaya *et al.*, 1992; Rao *et al.*, 1993; Wilson *et al.*, 2000;), each domain containing two helix-loop-helix EF-hand  $\text{Ca}^{2+}$  binding motifs. The incorporation of  $\text{Ca}^{2+}$  into CaM is essential and leads to a major conformational change which includes the opening of a hydrophobic cavity in each globular domain necessary for target protein recognition (Zhang *et al.*, 1995). Various modes of target peptide recognition have been thoroughly studied especially by X-ray crystallography (Ikura *et al.*, 1992; Meador *et al.*, 1992; Osawa *et al.*, 1999; Elshorst *et al.*, 1999; Schumacher *et al.*, 2001; Han *et al.*, 2000). A compact, calcium-free, apo form of CaM is converted to an extended dumbbell-shaped form on binding  $\text{Ca}^{2+}$  (Zhang *et al.*, 1995). This extended conformation consists of two structurally similar domains separated by a flexible 28-residue helix. Each domain has two EF-hand motifs with bound  $\text{Ca}^{2+}$ . The calcium-induced extension of CaM exposes two hydrophobic pockets, one per domain, which represent the binding sites for target proteins. Each hydrophobic pocket in the N- and C-lobes of CaM grabs well-conserved bulky hydrophobic residues separated by a specific number of residues (10, 14 or 16) on the target peptide. On binding to the target, the central CaM helix unwinds, and the two hydrophobic pockets wrap around the  $\alpha$ -helix of the protein target. Various target sequences are classified according to this spacing (Rhoads *et al.*, 1997; Yap *et al.*, 2000).

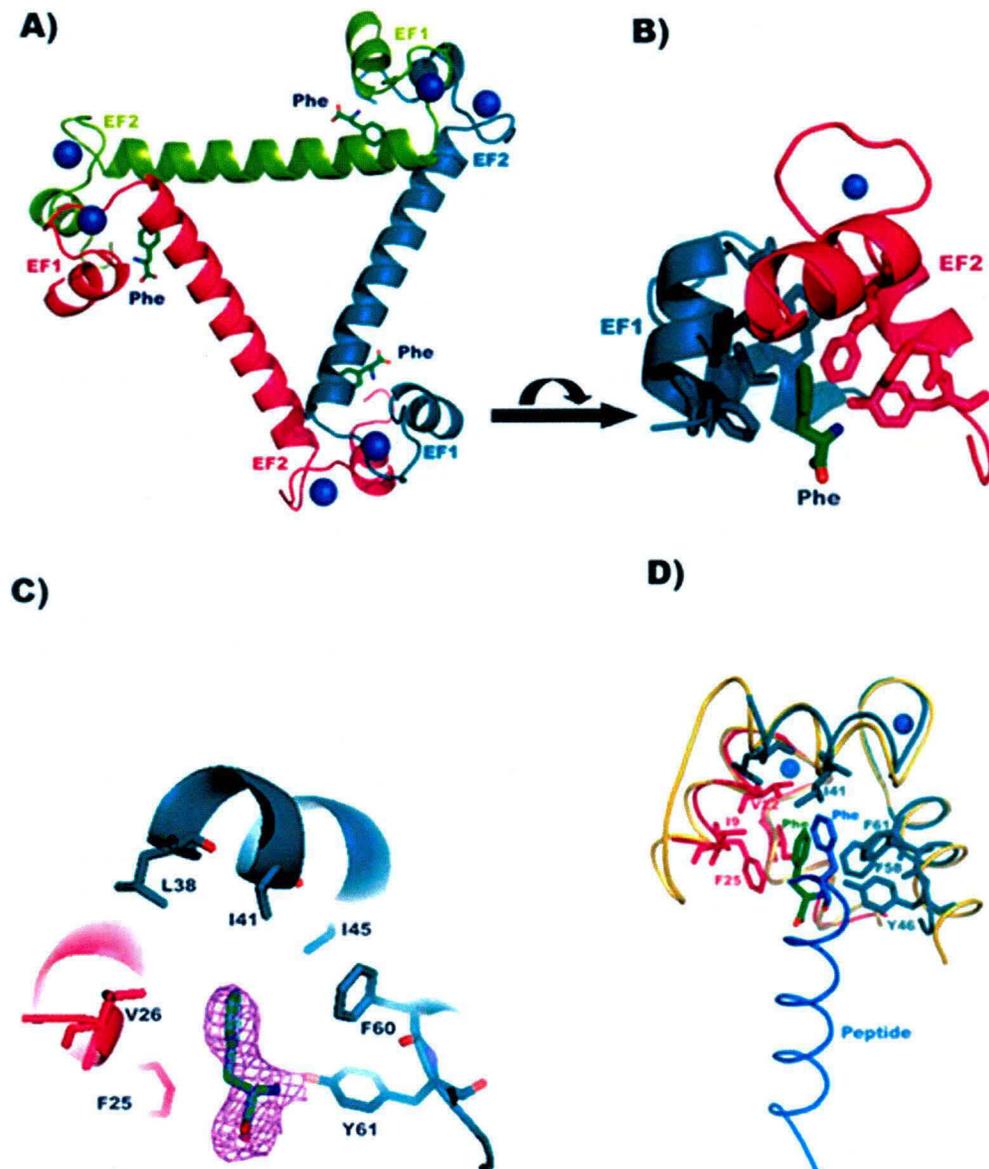
The two EF hand motifs belonging to N-terminal domain of EhCaBP1 are separated by long helix. In contrast, the corresponding EF motifs in CaM are connected by a short loop, thus bringing these two EF hand motifs into close proximity and forming a two EF-hand domain. The N-terminal domain of three molecules of EhCaBP1 participates in domain swapping to form trimers (Chapter-I, Fig.3.7 (B), Section-3.3.7) (Kumar *et al.*, 2007). This allows the EF1-hand motif of one molecule to interact with EF2-hand motif of an adjacent molecule to form a two EF-hand domain. This assembled-domain is similar to that of the two EF hand domains of CaM and TnC. This is essentially facilitated by a couple of critical residues in the linker that separate EF1 and EF2 motifs in comparison to CaM and ELC's (Kumar *et al.*, 2007). CaM and CaM-like proteins bind to their targets by anchoring hydrophobic residue of the target. Two EF hand motifs of each domain of

CaM bind to one hydrophobic residue of the target. To understand the target binding mode, EhCaBP1 was co-crystallized with Phenylalanine (Phe).

### 4.3 Results and Discussion

#### 4.3.1 Structural analysis of EhCaBP1- Phenylalanine complex

Crystal structure of EhCaBP1 with Phe showed that the hydrophobic pocket formed at the interface between EF1 and EF2 in the assembled-domain is bound to Phe [Fig.4.1 (B)] with good electron density [Fig.1 (C)]. The Phe forms several hydrophobic interactions with I8, F 24, V21 and V25 residues of EF hand motif 1 of one molecule and Y61, F60, F57, I40 and L 37 residues of EF hand motif 2 of another molecule (Fig.4.1 (C)). The Phe-bound assembled-domain was superimposed on the N-terminal domain of CaM, bound to hydrophobic IQ motif of cardiac Ca<sup>2+</sup> channel (Fallon *et al.*, 2005) [(Fig. 4.1 (D))]. Both structures superimposed very well with an RMS deviation of 1.23 Å. The hydrophobic residue Phe of the peptide (bound to CaM N-terminal domain) and Phe bound to EhCaBP1 assembled-domain are located at similar regions. Both ligands were bound at their respective hydrophobic pockets of the target protein. The distance between the two assembled-domains is approximately the same as the N and C-terminal domains of CaM, but the assembled-domain cannot change its structure after binding to target. This rigidity in the trimeric structure may be responsible for differential recognition of the targets.



**Figure-4.1** *Structural characterization of EhCaBP1 bound to Phenylalanine molecule.* (A) Structure of trimeric EhCaBP1 showing the phenylalanine molecule bound in the hydrophobic pocket formed by the assembled-domain. Three molecules interact with each other by head-tail manner forming an assembled-domain at the interface. (B) Close view of assembled-domain showing phenylalanine in the hydrophobic pocket surrounded by several hydrophobic residues. (C) 2Fo-Fc electron density map of phenylalanine at 2 $\sigma$  level. (D) The N-terminal domain of CaM-peptide complex is superposed on the assembled-domain of EhCaBP1-phenylalanine complex. The hydrophobic hanking residue of the peptide in CaM-peptide complex is bound to similar location as Phe in the assemble-domain of EF1 of one molecule (hot pink) and EF2 of neighbouring molecule (deep teal).

The binding affinity of EF hand motifs of EhCaBP1 (Gopal *et al.*, 1997), strongly indicates that only EF1 and EF2 are affected by the  $\text{Ca}^{2+}$  concentration fluctuations around it and C-terminal domain (EF3 and EF4) should be rigid and may not be influenced by  $\text{Ca}^{2+}$  concentration changes. This is also evident from the results obtained from crystallization studies presented here that the assembled-domain in trimer bind to the hydrophobic amino acid revealing the mode of target binding. Therefore the evolution of the CaBPs, such as EhCaBP1 may have been designed to offer both functional and structural diversity suitable for a pathogen to modulate host-pathogen relationship.

#### 4.4 Conclusion

Three molecules of EhCaBP1 N-terminal domains participate in domain swapping to form trimers led by couple of critical residues difference in between EF1 and EF2 motifs in comparison to CaM and ELC's. This allows the EF1-hand motif of one molecule to interact with EF2 of an adjacent molecule to form a two EF-hand domain (assembled-domain) similar to that of calmodulin and troponin C. The hydrophobic pocket formed at the interface between EF1 and EF2 in the assembled-domain is bound to phenylalanine [(Fig.4.1 (A)]. The phenylalanine forms several hydrophobic interactions with Ile 9, Phe 25, Val 22 and Val26 residues of EF-hand motif 1 of one molecule and Tyr 62, Phe 61, Phe 58, Ile 41 residues of EF-hand motif 2 of another molecule. The Interacting distances of the bound Phenylalanine in the hydrophobic pocket of the assembled-domain with the surrounding residues in EhCaBP1-Phe complex is shown in Table 4.1.

The assembled-domain bound to Phe was superimposed with the N-terminal domain Calmodulin bound to hydrophobic cardiac IQ motif of calcium channel (Cardiac calcium channel) [Fig.4.1 (D)]. Both structures superimposed very well with an rms deviation of 1.23 Å and the hydrophobic residue Phe of the peptide bound to Cam N-terminal domain and Phe bound to EhCaBP1 assembled-domain are located at similar location. Both the ligands bind their targets in similar manner at the hydrophobic pockets of respective domains.

The distance between two assembled-domains is approximately same as the distance between calmodulin N and C-terminal domains. But the assembled-domain can't change its structure after binding to target as calmodulin wraps around its target.

This rigidity in the trimeric structure may be responsible for differential recognition of the targets.

**Table-4.1 Interacting distances of the bound Phenylalanine in the hydrophobic pocket of the assembled-domain with the surrounding residues in EhCaBP1-Phe complex.**

Interacting residues	Phenylalanine	Distances ( Å)
OH /62 Tyr/B	N/ Phe	2.79
CE1/25 Phe/B'	CB/ Phe	3.53
CD1/45 Phe/B	CZ/ Phe	4.27
CG2/41 Ile/ B	CZ/ Phe	4.47
CG2/26 Val/B'	CZ/ Phe	4.78
CG1/22 Val/B'	CZ/ Phe	3.89
CE1 /58 Phe/B	CZ/ Phe	4.74

It is evident from the biochemical data and the structure that only N-terminal domain is involved in target binding and activating kinase. It is also noted from earlier biochemical data that the calcium binding affinity of the EF3 and EF4 are much higher than the EF1 and EF2 (Gopal *et al.*, 1997), which makes clear that only EF1 and EF2 are affected by the calcium concentration fluctuations around it and C-terminal domain, which consists of EF3 and EF4 motifs, should be rigid and may not be effected by calcium concentration changes.

# CHAPTER-III

---

**Crystal structure of N-terminal domain of EhCaBP1**

## 5.1 Abstract

The crystal structure of EhCaBP1 traced only the N-terminal region of the full-length protein with a novel arrangement of EF-hand motifs. The presence of the long helix joining the two EF-hand motifs results in the extended conformation. Three symmetry-related molecules are interacting in a head-to-tail manner to form a trimeric structure which limits the movement and hence the targets of this protein are different from CaM, which is highly flexible in nature. The functional studies on the independent domains show that the N-terminal domain undergoes a major conformation change upon binding calcium and it activates the endogenous kinase more efficiently than the full-length EhCaBP1 (Jain *et al.*, 2009). Also, the absence of dominant negative effect in N-terminal-GFP cells suggests that this domain is likely to behave like the full length EhCaBP1 protein. Fluorescence microscopy studies show the presence of the N-terminal domain at the phagocytic cup and its complete co-localization with F-actin. Furthermore, majority of N-terminal molecules were found around the phagocytic cups and not much in cytoplasm. The full-length EhCaBP1 protein molecules are found around phagocytic cups as well as in the cytoplasm. Distribution of C-terminal protein is quite different. Most of the molecules are localized in the cytoplasm with no specific relation with F-actin at the site of attachment of RBC (Jain *et al.*, 2009). These studies clearly demonstrate that the N-terminal domain has most of the functions in comparison to the full length EhCaBP1. As the full length EhCaBP1 structure revealed only the N-terminal domain and half of the structure, i.e., the C-terminal domain is missing, raised several questions. To validate the full-length structure, the initiatives have been taken to determine the structures of the individual domains. The crystal structure of N-terminal domain has been determined at 2.5 Å resolution. The crystal were grown using PEG 400 as a precipitant and they belong to P3 space group with unit cell parameters of  $a = 89.589$  Å,  $b = 89.589$  Å,  $c = 35.049$  Å. The final model refined to R\_factor of 23.3 % and Free\_R of 27.1%, respectively (Materials and Methods, Table 2.3, Section-2.23, Page No. 49). The structure is similar to the N-terminal structure of the full-length protein with the extended conformation of the EF-hand motifs connected with a long central helix. Unlike CaM, the overall structure of N-terminal domain is dumbbell shaped. Owing to domain swapping oligomerization, three EhCaBP1 molecules interact in a head-to-tail manner to form a triangular trimer.

## 5.2 Introduction

Crystal structure of EhCaBP1 showed an unusual arrangement of the EF-hand motifs of N-terminal domain (Kumar *et al.*, 2007). Also, the three glycine residues (G63, G67 and G76) present in the central linker region makes it more flexible as compared to CaM. Due to the flexibility of this central linker region, the C-terminal domain structure could not be traced. N-terminal domain contains two EF-hand motifs, each binds to one calcium ion. In CaM, the presence of glycine and proline residues in the linker region connecting EF-hand 1 and EF-hand 2 motifs results in loop-formation and close-contacts of the two motifs forming the N-terminal domain. On the other hand, the linker region connecting EF-hand 1 and EF-hand 2 motifs of EhCaBP1 has arginine and lysine residues resulting in the extended conformation of the N-terminal domain, where EF-1 and EF-2 motifs are far apart from each other (Kumar *et al.*, 2007). Three molecules of EhCaBP1 interact in a head-to-tail manner to form a domain-swapped trimer.

Taking leads from the full-length structure, it was proposed that only N-terminal domain can bind to targets and both N-terminal and C-terminal domains may have different functions. The calcium binding affinities of EF-hand motifs also suggests that N-terminal domain may sense calcium levels in the cell, whereas the C-terminal domain may be rigid. To study the functional properties of the independent domains of EhCaBP1, it was separately cloned, and characterized (Jain *et al.*, 2009).

## 5.3 Schematic representation of independent domains of EhCaBP1

The N-terminal domain extends from 1-66 amino acid residues and the C-terminal domain extends from 67-134 amino acid residues, respectively (Fig.5.1). Both the domains contain two EF-hand motifs and each EF-hand motifs bind one calcium ion.

A)



B)



**Figure-5.1 Schematic representation of independent domains of EhCaBP1.**

A) Schematic representation of N-terminal domain containing 1-66 residues with EF-1 and EF-2 motifs extending from 10-21 and 46-57 amino acids residues, respectively. B) Schematic representation of C-terminal domain containing 67-134 residues with EF-3 and EF-4 motifs extending from 85-96 and 117-128 amino acids residues, respectively. One calcium ion binds to each EF-hand motif.

#### 5.4 Functional properties of N-terminal and C-terminal domains

The CD-spectroscopic studies show that both the domains bind  $\text{Ca}^{2+}$ , but the consequence of binding is not the same. N-terminal domain undergoes a major conformation change as compared to the C-terminal domain. In this respect, N-terminal domain behaved like a full-length EhCaBP1 (Jain *et al.*, 2009). EhCaBP1 is known to activate endogenous kinase(s) in a  $\text{Ca}^{2+}$  dependent manner (Yadava *et al.*, 1997). N-terminal domain activates the endogenous kinase more efficiently than the complete EhCaBP1 while, the C-terminal domain shows a marked reduction in activity of about 50–60% of the full-length EhCaBP1 (Jain *et al.*, 2009).

It has been shown that EhCaBP1 has a crucial role in the initiation of erythrophagocytosis and erythrophagocytosis has been linked to the pathogenesis in amebiasis (Orozco *et al.*, 1983). Over-expression of full length EhCaBP1 also did not change significantly the level of erythrophagocytosis (Sahoo *et al.*, 2004). The erythrophagocytic study with the independent domains shows that the over-expression of C-terminal domain results in a dominant negative phenotype with respect to

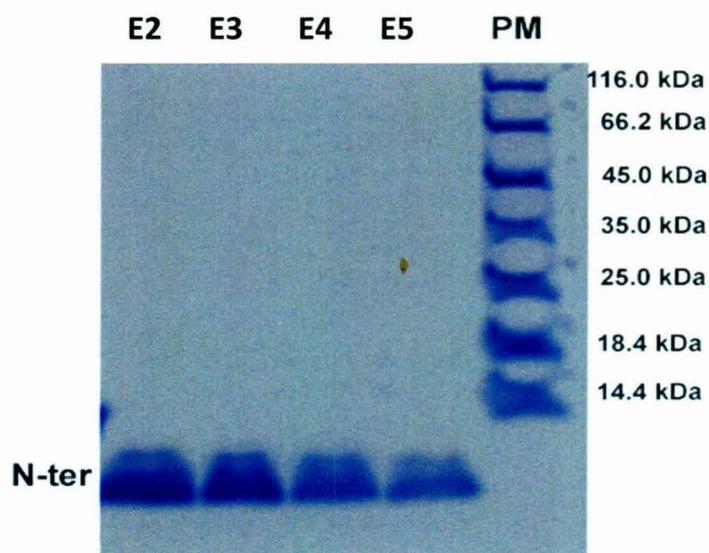
erythrophagocytosis, whereas the N-terminal-GFP cells suggests that this domain is likely to behave like the full length EhCaBP1 protein (Jain *et al.*, 2009). Fluorescence microscopy studies show the presence of the N-terminal domain at the phagocytic cup and its complete co-localization with F-actin. Furthermore, majority of N-terminal molecules were found around the phagocytic cups and not much in cytoplasm. The full-length EhCaBP1 protein molecules are found around phagocytic cups as well as in the cytoplasm. Distribution of C-terminal protein is quite different. Most of the molecules are localized in the cytoplasm, with no specific relation with F-actin at the site of attachment of RBC (Jain *et al.*, 2009).

The crystal structure of EhCaBP1 has shown that N-terminal region forms a domain-swapped trimer in which EF-hand 1 of one molecule interacts with EF-hand 2 of the symmetry-related molecule in a head-to-tail manner and the assembled domain formed has similar hydrophobic pocket as that of CaM (Kumar *et al.*, 2007). The functional studies on the independent domains have shown that the N-terminal region of the protein is the functional domain and is sufficient to carry out most of the function of the full-length protein (Jain *et al.*, 2009). Keeping the importance of the N-terminal domain in mind and to remove the model bias of the C-terminal region to get a well-refined structure and to further confirm the trimeric nature of the N-terminal domain, it was crystallized and the structure has been determined.

## **5.5 Results and Discussion**

### **5.5.1 Purification of N-terminal domain**

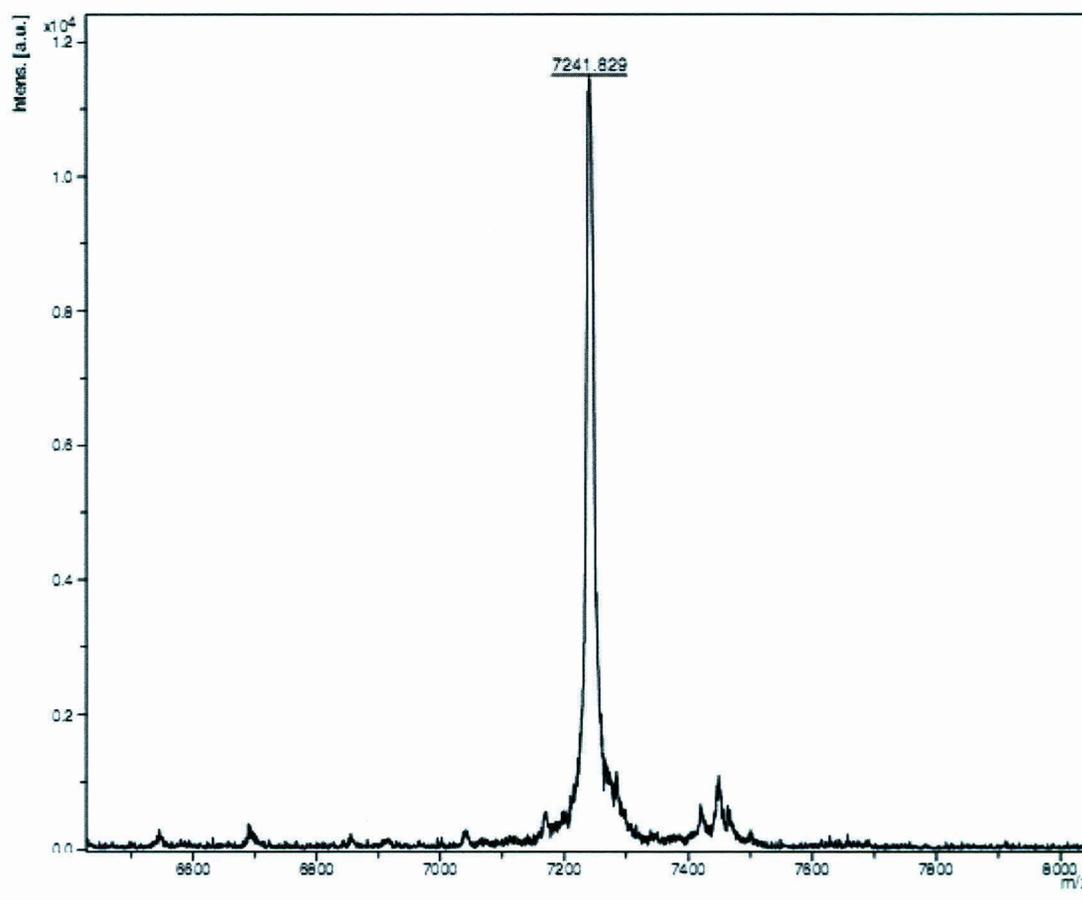
The protein was purified using ion exchange chromatography and was analysed on 14% SDS-PAGE gel (Fig.5.2). The band observed was corresponding to the size of N-terminal domain of EhCaBP1 protein (i.e., 7.2 kDa).



**Figure-5.2** *Over-expression and purification of N-terminal domain of EhCaBP1.* SDS-PAGE (14%) analysis of the purification of N-terminal domain of EhCaBP1. A total of 12  $\mu$ l of the protein marker and 15  $\mu$ l of the purified N-terminal protein was loaded in the respective lane. The gel was stained with coomassie blue. The mobility of molecular-weight markers were as indicated. E2-E5 indicates the elution fractions; PM indicates the Protein Marker.

### 5.5.2 Mass spectroscopy of N-terminal domain

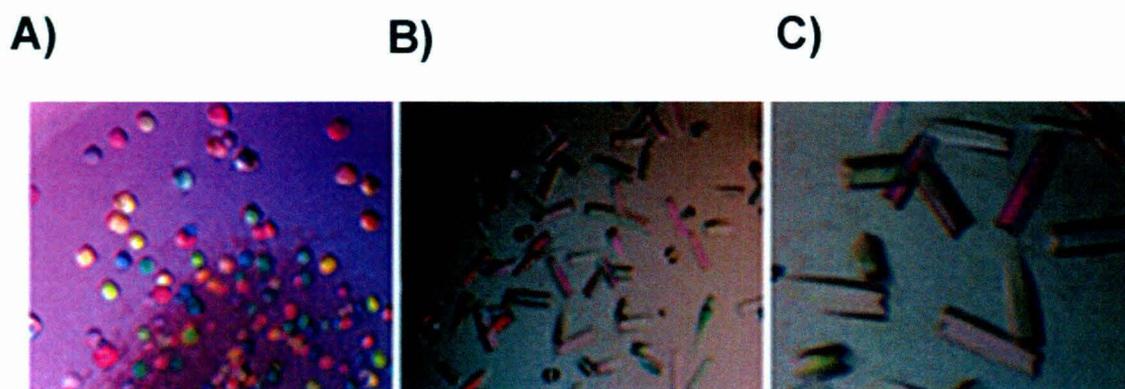
The mass-spectroscopic results confirm the intact mass as well as the purity of the N-terminal protein (Fig.5.3) Mass spectroscopy was done in an in-house Bruker Daltonics flexAnalysis instrument of type autoflex TOF-TOF. (Materials and methods, Section No. 2.21, Page No. 47).



**Figure-5.3** *Mass Spectroscopy of N-terminal domain of EhCaBP1. The peak at 7241 dalton corresponds to the molecular weight of N-terminal domain of EhCaBP1.*

### 5.5.3 Crystal structure of N-terminal domain

The crystals of N-terminal domain were obtained in PEG 400 and PEG 500-MME with different morphology (Fig.5.5) and the crystals in PEG 400 were hexagonal in shape and were used for the data collection using in-house X-ray source, whereas full-length protein crystals were only obtained when MPD was used for crystallization.

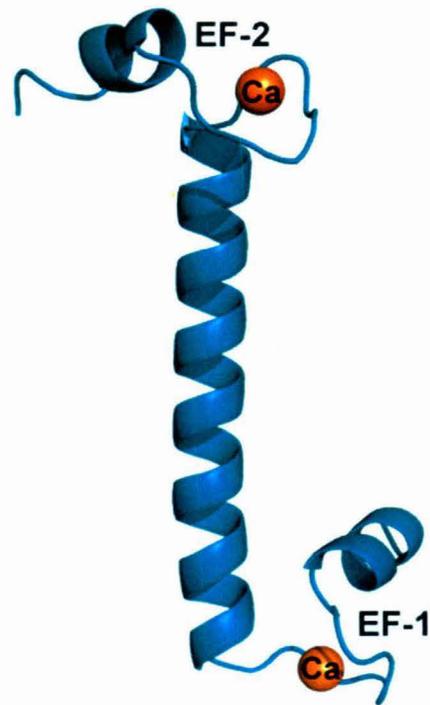


**Figure-5.4** *Crystals of N-terminal domain in different crystallization conditions.*

(A) 18% PEG 500 MME, Acetate buffer pH 3.6.-3.8 (B) 15% PEG 400, Acetate buffer pH 3.6-3.8 C) 15% PEG 400, Acetate buffer pH 3.6 with additive mixture (Materials and Methods, Section No. 2.22, Page No. 48).

#### **5.5.4 Overall structure of N-terminal domain of EhCaBP1**

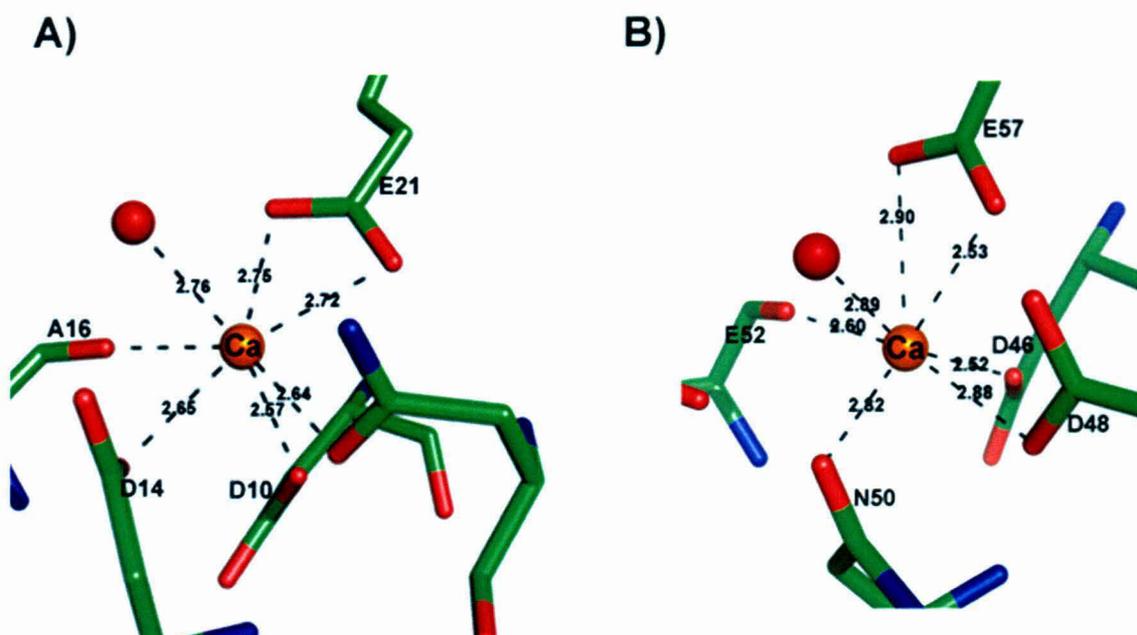
The crystal structure of N-terminal domain forms an extended conformation similar to that of the full-length crystal structure in which only N-terminal region has been traced. Each EF-hand motifs binds one calcium ion. It belongs to P3 space group with four molecules in the asymmetric unit. Also, the N-terminal domain forms domain-swapped trimer similar to the reported crystal structure of full-length protein (Kumar *et al.*, 2007).



**Figure-5.5** *Crystal structure of N-terminal domain of EhCaBP1. Crystal structure of N-terminal domain shows the extended conformation, where EF-hand 1 and EF-hand 2 motifs are far apart and each motif binds one calcium ion. The linker region between the two motifs forms a long helix. The overall structure is similar to the N-terminal domain of the full-length protein. The calcium ion has shown in orange sphere.*

### 5.5.5 Calcium-binding loop of N-terminal domain

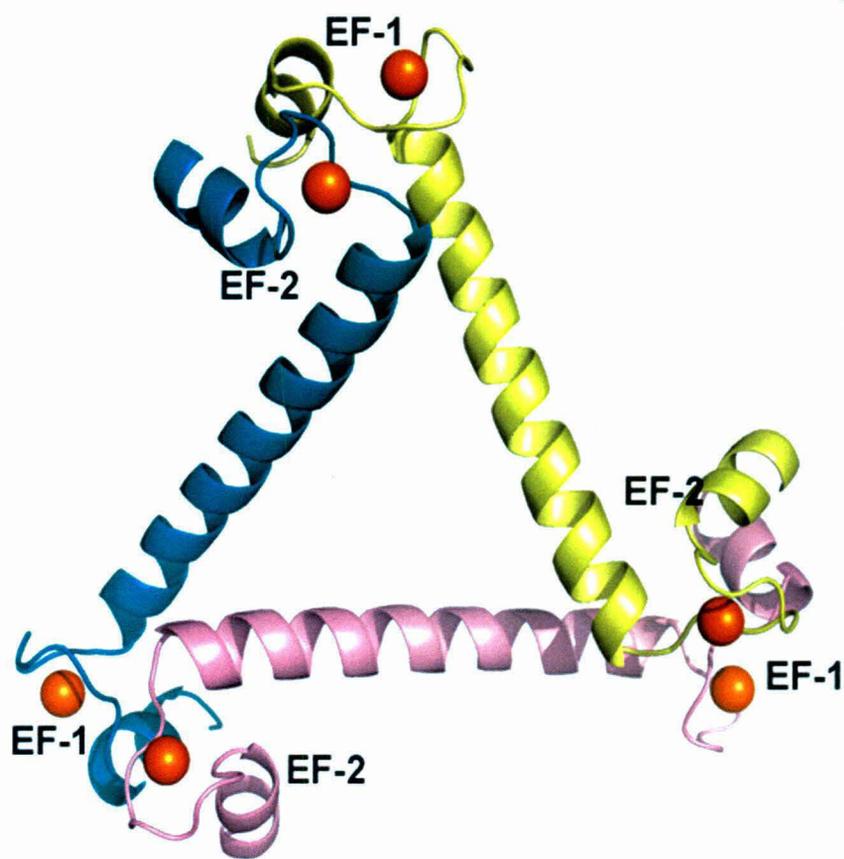
The N-terminal domain contains two EF-hand motifs with bound calcium ions. The coordination geometry of both the calcium-binding loops has shown in Fig.5.6. The calcium ion is coordinating with seven of its ligands. One water molecule is also taking part in satisfying the coordination geometry.



**Figure-5.6**  $Ca^{2+}$ -coordination of EF-hand motifs of N-terminal domain. A)  $Ca^{2+}$ -coordination in calcium-binding loop 1. B)  $Ca^{2+}$ -coordination in calcium-binding loop 2. The coordination distances are shown in Å. Calcium and water molecules are shown as orange and red spheres, respectively.

### 5.5.6 Symmetry-related trimeric structure of N-terminal domain

The extended conformation of N-terminal domain forms a domain-swapped trimer where three symmetry-related molecules are interacting in a head-to-tail manner. EF-hand 1 of one molecule interacts with EF-hand 2 of the symmetry-related molecule and forms a trimer (Fig.5.7). The assembled domains are similar to CaM N-terminal domain.



**Figure-5.7** *Symmetry-related trimer of N-terminal domain.* EF-hand 1 of one molecule is interacting in head-to-tail manner with EF-hand 2 of the symmetry-related molecule in a domain-swapped manner. Three N-terminal molecules are shown in three different colors and the bound calcium ions are shown as orange sphere. EF-hand motifs are labelled.

## 5.6 Conclusion

The high-resolution crystal structure of EhCaBP1 shows an unusual arrangement of EF-hand motifs in which EF-hand 1 of one molecule is interacting with EF-hand 2 of the symmetry-related molecule and forms a trimer. The structure was intriguing and prompted us to carry out structural studies of the individual domains of EhCaBP1. The crystal structure of N-terminal domain is solved at 2.5 Å resolution. It belongs to space group P3 with four molecules in the asymmetric unit. The overall structure of

the N-terminal domain is more refined as compared to the full-length structure. The crystallographic R\_factor and R\_free of N-terminal domain (R\_factor = 23.3%, R\_Free = 27.1%) are comparatively lower than that of the full-length crystal structure (R\_factor = 25.6%, R\_Free = 28.2%).

The crystal structure of N-terminal domain of EhCaBP1 is similar to that of the full-length protein in which only the N-terminal region has been traced. High resolution crystal structure showed an unusual arrangement, with three molecules of N-terminal domain interacting in a head-to-tail manner to form a symmetry-related trimer. This arrangement allows the N-terminal EF-hand motif of one molecule to interact with that of an adjacent molecule to form two EF-hand containing domains, similar to that seen in CaM and TnC. The observation indicating absence of dominant negative phenotype on over-expression, also supports the view that the N-terminal half of the protein is capable of carrying out most of the functions of the full molecule. This may be due to expressed N-terminal domains forming trimeric complexes with each other as well as endogenous N-terminal domains of EhCaBP1 molecules. However, not all functions can be carried out by the N-terminal half. For example, it could not bind to G-actin. The differential behavior suggests that the two halves of the molecule have different functions. The behaviour of C-terminal cells can also be due to alteration in the property of C-terminal due to its fusion with GFP, the latter being much larger than the former. It has been demonstrated that both domains of EhCaBP1 have distinct folding features (Mohan *et al.*, 2008; Mukherjee *et al.*, 2007). This is similar to TnC and CaM where N and C-terminal domains were found to be structurally independent and likely to bind different targets (Babu *et al.*, 1985; Herzberg *et al.*, 1985; Sunderalingam *et al.*, 1985). TnC interacts with only two proteins, troponin I and troponin T. The N-terminal domain functions as the Ca<sup>2+</sup> - specific regulatory switch, while the C-terminal domain plays mainly the structure stabilizing role (Takeda *et al.*, 2003; Vinogradova *et al.*, 2003). On the other hand, domain independence is the key to high level of versatility of CaM (Yamniuk *et al.*, 2004). A genetic screen in Paramecium has also revealed that the domains of CaM have separable physiological roles (Kung *et al.*, 1992).

Previous biochemical studies have clearly shown that the calcium binding affinity of the EF3 and EF4 are much higher than that of the EF1 and EF2 (Gopal *et al.*, 1997). This strongly indicates that only EF1 and EF2 are affected by the Ca<sup>2+</sup>

concentration fluctuations around it and C-terminal domain (EF3 and EF4) should be rigid and may not be influenced by  $\text{Ca}^{2+}$  concentration changes. This is also evident from the results obtained from crystallization studies that N-terminal is the functional domain and plays an important role in target recognition. Therefore, the evolution of the CaBPs, such as EhCaBP1 may have been designed to offer both functional and structural diversity suitable for a pathogen to modulate host-pathogen relationship.

# CHAPTER-IV

---

**Flexibility and plasticity of calcium binding loops of  
EhCaBP1: Crystal structure of EhCaBP1 in complex with  
 $\text{Pb}^{2+}$ ,  $\text{Ba}^{2+}$  and  $\text{Sr}^{2+}$**

## 6.1 Abstract

Heavy metals may accumulate in the body and cause disease states even at low concentrations. These bind to various proteins which ultimately influences the signal pathways and leading to abnormal responses. Although CaM is best characterized by its ability to specifically bind  $\text{Ca}^{2+}$ , a number of studies have indicated that in fact it can also be activated by other metal ions. The most studied of the toxic metal is  $\text{Pb}^{2+}$ , which has a high affinity towards CaM and is able to activate it even at low concentrations. Some other metal ions like  $\text{Mg}^{2+}$ ,  $\text{Ba}^{2+}$ ,  $\text{Sr}^{2+}$ ,  $\text{Hg}^{2+}$ ,  $\text{Cd}^{2+}$  and most lanthanides also show affinity with natural and engineered Calcium Binding Proteins (CaBPs) including CaM. In this work, the structural investigation into flexibility and plasticity of the EF-hand motifs is examined, with the structure of Calcium Binding Protein-1 from *Entamoeba histolytica* (EhCaBP1) complexed with Pb, Ba and Sr. The overall crystal structures of EhCaBP1- heavy metal complexes are found to be similar to native protein with minor deviations in the metal-ion coordination and ligands involved. In  $\text{Pb}^{2+}$ -EhCaBP1, the lead ion coordinates with six ligands in the metal binding loop with minor differences in the coordination distances. The coordination distances of barium ion in  $\text{Ba}^{2+}$ -EhCaBP1 complex has found to be maximum with one extra-bound barium ion outside the EF-hand 2 motif which represents a state where an ion is just about to be bound/released. In  $\text{Sr}^{2+}$ -EhCaBP1 complex, the 3<sup>rd</sup> aspartate residue of the metal binding loop is compensating the loss of water coordinating by donating both of its oxygen atoms. These structures reveal that the EF-hand motifs can accommodate several heavy atoms with different binding affinities.  $\text{Ba}^{2+}$ -EhCaBP1 complex data collected at Cu  $K\alpha$ , showed very strong anomalous signal which could be used for SAD phasing. This clearly indicates that the barium incorporation and data collection using home source could be an easy approach towards the structure solution of Calcium Binding Proteins.

## 6.2 Introduction

Complex formation between inorganic metals and proteins is essential to numerous biological processes, where binding of physiologically-relevant metals (e.g. Ca, Mg, Zn, Fe) effects conformational changes that confer new functions. Nearly 40% of all proteins are known to bind metals (Ibers *et al.*, 1980; Holm *et al.*, 1996; Tainer *et al.*, 1992; Glusker *et al.*, 1991; Silva *et al.*, 1991; Lippard *et al.*, 1994). Metal-binding sites in general are characterized by a central shell of hydrophilic chelating ligands, with a surrounding shell of hydrophobic residues (Yamashita *et al.*, 1990; Bagley *et al.*, 1995), and the binding geometry, coordination number, and ligand preference associated with the binding of essential metal ions to natural proteins has been studied extensively (Glusker *et al.*, 1991; Babor *et al.*, 2005; Dudev *et al.*, 2003; Dudev *et al.*, 2003; Einspahr *et al.*, 1983; Harding *et al.*, 1999; Harding *et al.*, 2000).  $\text{Ca}^{2+}$ -Binding Proteins (CaBPs) represent a large family of these metalloproteins.

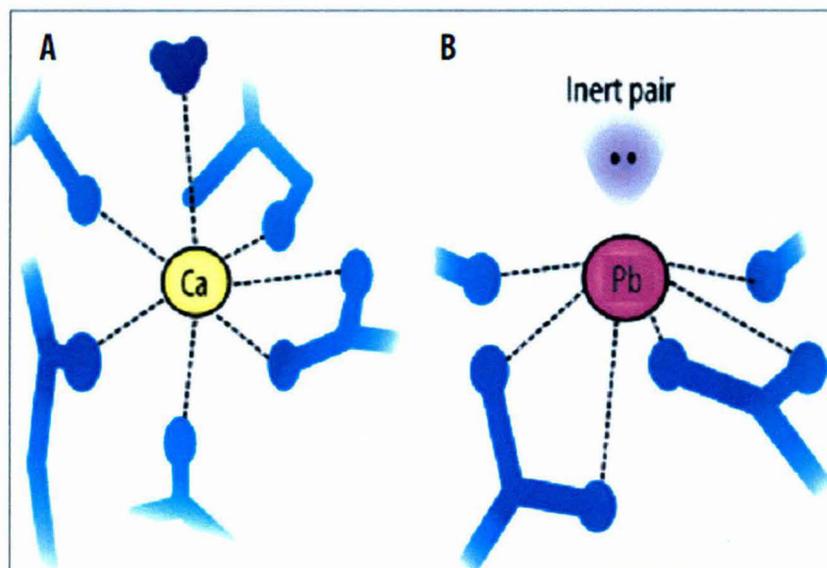
Calmodulin (CaM) is a prototypic EF-hand calcium-binding protein that acts by sensing calcium levels and binding to target proteins in a regulatory manner. Its importance is highlighted by its 100% sequence conservation across all vertebrates. The resulting  $\text{Ca}^{2+}$ -CaM complex mediates biological processes related to inflammation, muscle contraction, memory, nerve growth and the immune response, and may act as a sensor/signal transducer when it binds with proteins unable to bind  $\text{Ca}^{2+}$  themselves (Herzberg *et al.*, 1986; Holmes *et al.*, 1990).  $\text{Ca}^{2+}$  may also interact with proteins to regulate functions related to cell division, apoptosis and intracellular signaling (Carafoli *et al.*, 1994; Santella *et al.*, 1997; Martin *et al.*, 1984; Nelson *et al.*, 1998), as well as to enhance protein stability (Tajima *et al.*, 1986). Although CaM is best characterized by its ability to specifically bind  $\text{Ca}^{2+}$ , a number of studies have indicated that in fact it can also be activated by other metal ions (Habermann *et al.*, 1983; Chao *et al.*, 1984; Richardt *et al.*, 1986; Ouyang *et al.*, 1998; Ozawa *et al.*, 1999). In addition to  $\text{Pb}^{2+}$ ,  $\text{Sr}^{2+}$ ,  $\text{Hg}^{2+}$  and  $\text{Cd}^{2+}$ , most lanthanides have been found to bind with both natural and engineered CaBPs (Forsen *et al.*, 1979; Mills *et al.*, 1985; Wang *et al.*, 1984; Pidcock *et al.*, 1984; Yang *et al.*, 2005), and many of these can activate and then inhibit CaM in response to changes in metal concentration, thus altering the activity of downstream CaM-mediated functions (Chao *et al.*, 1984; Ferguson *et al.*, 2000; Chao *et al.*, 1995; Suzuki *et al.*, 1985; Habermann *et al.*, 1983). Heavy metals may accumulate in the body and cause disease states even at low

concentrations (Cline *et al.*, 1996; Lidsky *et al.*, 2003). Additionally, several recent studies have reported bacterium capable of encoding sensor or chelate proteins for toxic metals (Busenlehner *et al.*, 2006; Fahmy *et al.*, 2006; Ye *et al.*, 2005), including *Ralstonia metallidurans*, which possesses the first identified bacterial resistance determinant found to be specific for  $\text{Pb}^{2+}$  (Borremans *et al.*, 2001). Despite rigorous studies aimed at understanding the structural implications of metal-binding for physiologically-relevant metals, the binding characteristics of toxic metals like lead ( $\text{Pb}^{2+}$ ) are less clearly defined.

$\text{Pb}^{2+}$  is a particularly persistent anthropogenic toxicant whose bioavailability has increased as a result industry. Physiological effects of  $\text{Pb}^{2+}$  uptake include neurological disorders, anemia, kidney damage, hypertension and male fertility decrease (Hernandez-Ochoa *et al.*, 2005; Apostoli *et al.*, 2005; Moore *et al.*, 1987; Chao *et al.*, 1990). In the nervous system, lead affects among others, CaM-regulated processes, neurotransmitter release and protein kinase C (Simons *et al.*, 1993; Marchetti *et al.*, 2003). The effects of lead have also been reported to present concentration-dependence, such that the activation of CaM by  $\text{Pb}^{2+}$  peaks in the 30–500  $\mu\text{M}$  range and higher  $\text{Pb}^{2+}$  concentrations again activate CaM less effectively (Sandhir *et al.*, 1994; Weaver *et al.*, 2002). At the molecular level,  $\text{Pb}^{2+}$  has been shown to specifically target voltage-gated calcium channels (Atchison *et al.*, 2003), activate skeletal muscle troponin C (TnC) (Chao *et al.*, 1990), and displace both  $\text{Ca}^{2+}$  and  $\text{Zn}^{2+}$  in proteins, including CaM, protein kinase C (Markovac *et al.*, 1988) and synaptotagmin (Bouton *et al.*, 2001).  $\text{Pb}^{2+}$  is the most studied toxic metal, which has a high affinity towards CaM and is able to activate it even at low concentrations (Habermann *et al.*, 1983; Richardt *et al.*, 1986; Kern *et al.*, 2000). Low-level lead exposure also affects neuronal growth (Cline *et al.*, 1996). One specific target for lead poisoning can be CaM;  $\text{Pb}^{2+}$  binding to CaM may falsely activate it (Verity *et al.*, 1990; Goering *et al.*, 1993; Ouyang & Vogel, 1998; Kern *et al.*, 2000), leading to an abnormal response. To a large extent, lead seems to be able to substitute for calcium in the regulation of CaM function (Habermann *et al.*, 1983; Fullmer *et al.*, 1985) and many target proteins are indeed activated by Pb-CaM similarly to Ca-CaM (Chao *et al.*, 1995).

The ionic characteristic of lead allows it to bind with greater affinity than calcium and zinc ions to protein binding-sites. Because of its larger ionic radius (119

pm) compared to zinc (74 pm) and calcium (100 pm)], and greater electronegativity (in the Pauling scale, 2.33 versus 1.0 for calcium and 1.65 for zinc), lead establishes very favourable interactions with the coordinating groups of the protein. In contrast to zinc and calcium, where the spherical distribution of electric charge around the ion produces a relatively regular arrangement of the groups that coordinate to them in the protein, for lead the charge distribution is irregular because of the presence of an inert electron pair in its electronic cloud (Fig.6.1).



**Figure-6.1** *Schematic representation of the coordinating groups for calcium and lead. (A) Scheme of calcium coordination in an EF motif by oxygen atoms. A water molecule (in black) also participates in the coordination. The coordination geometry is holodirectional (pentagonal bipyramid), with a more regular distribution of the coordinating groups. (B) In contrast, the hemidirectional disposition adopted by the groups that form the coordination sphere for lead in a compound with coordination number of six. This is caused by the presence of an inert electron pair in the ion. (Adapted from Garza et al., 2006).*

Some other metal ions like  $Mg^{2+}$ ,  $Ba^{2+}$ ,  $Sr^{2+}$ ,  $Hg^{2+}$ ,  $Cd^{2+}$  and most lanthanides also show affinity with natural and engineered Calcium Binding Proteins (CaBPs) including CaM (Forson *et al.*, 1979; Wang, 1984; Pidcock, 2001; Yang, 2005).

However, the maximum activation decreased in the order  $\text{Pb}^{2+} > \text{Ca}^{2+} > \text{Sr}^{2+} > \text{Ba}^{2+} > \text{Cd}^{2+}$  as demonstrated previously in EF-hand motif of D-galactose binding protein (Vyas *et al.*, 1989).

The activation and inhibition depends on the concentration of the metal ion, which further alter the activity of downstream CaM-mediated functions. As lead exhibits highest affinity, it can activate CaM at very low concentrations and can replace all bound Calcium ions (Aramini *et al.*, 1996) leading to abnormal activity and diseased state. While ionic displacement of  $\text{Ca}^{2+}$  by a competing metal ion may represent one mechanism of metal toxicity, an additional opportunistic binding mechanism, resulting from metal–protein interactions in regions lacking an established binding site and related to electrostatic potential interactions, may represent an additional avenue for metal toxicity (Kirberger *et al.*, 2008). Both mechanisms may offer partial explanations for the activation/inhibition of CaM activity reported in related studies (Chao *et al.*, 1984; Ferguson *et al.*, 2000; Chao *et al.*, 1995; Suzuki *et al.*, 1985; Habermann *et al.*, 1983; Kern *et al.*, 2000).  $\text{Ba}^{2+}$  and  $\text{Mg}^{2+}$  have been reported to have significantly lower, if any, affinity towards CaM (Ozawa *et al.*, 1999). For example, the CaM-regulated activation of cerebellar nitric oxide synthase requires an over 200-fold higher concentration of  $\text{Ba}^{2+}$  than of  $\text{Ca}^{2+}$  (Yamazaki *et al.*, 1996) and the interaction between CaM and caldesmon is weakened by the exchange of  $\text{Ca}^{2+}$  for  $\text{Ba}^{2+}$  (Huber *et al.*, 1996). On the other hand,  $\text{Mg}^{2+}$  levels in the cellular environment are much higher than those of  $\text{Ca}^{2+}$  and CaM must be able to function specifically in the presence of an excess of  $\text{Mg}^{2+}$  ions (Malmendal *et al.*, 1999).

The strontium and the calcium ions, members of the alkaline earth series (Group IIB of the Periodic Table), have many properties in common, both having a valency of 2+, similar ionic radii, and the ability to form complexes and chelates of various solubilities and various binding strengths. The relative binding affinity of  $\text{Ca}^{2+}$  and  $\text{Sr}^{2+}$  differs among anionic compounds, some preferring  $\text{Ca}^{2+}$  and others  $\text{Sr}^{2+}$ . The binding of  $\text{Sr}^{2+}$  to alginates exceeds that of  $\text{Ca}^{2+}$  by factors of 1.5 to 4.3, whereas calcium is preferred in other interactions, e.g., binding to G-actin, blocking negative charges on membranes, and binding to collagen (Wasserman *et al.*, 1997). In most biological systems, preference in general is given to  $\text{Ca}^{2+}$  over  $\text{Sr}^{2+}$ , although a marine organism, an *Acantharia*, constructs its internal skeleton of strontium sulfate (da Silva

*et al.*, 1991). In higher vertebrates, the bony skeleton is mostly composed of the highly insoluble calcium phosphate complex, hydroxyapatite. Strontium present in bone is usually considered a contaminant, with no physiological function at the trace values present therein. However, the usefulness of strontium as a potential therapeutic agent for osteoporosis has experimental support. For example, in recent studies, low doses of strontium are reported to have beneficial effect on bone formation in osteoporotic patients and, in some studies, strontium has been given as a salt, designated S 12911 (Brandi *et al.*, 1993; Grynepas *et al.*, 1996). The mechanism of discrimination between  $\text{Ca}^{2+}$  and  $\text{Sr}^{2+}$  is certainly a consequence of interactions that occur during their transfer from one compartment or site to another. For intestinal absorption, the preferential absorption of  $\text{Ca}^{2+}$  over  $\text{Sr}^{2+}$  can be explained, at least in part, by current views of the way that calcium is absorbed from the ingesta. Two factors in the transcellular path of absorption that are involved in the uphill transport of calcium can be implicated as part of the discriminatory process, these being the high-affinity calcium-binding protein, calbindin, that presumably serves to increase the intracellular diffusion of calcium, and the plasma membrane ATP-dependent calcium-pump that actively extrudes calcium from the enterocyte (Wasserman *et al.*, 1997). Binding of  $\text{Ca}^{2+}$  to each of these proteins is considerably greater than that of  $\text{Sr}^{2+}$  (Ingersoll *et al.*, 1971; Pfelger *et al.*, 1975). In addition to active transport, calcium also moves across the intestinal membrane by a paracellular, diffusional-type process, and that path also seems to give preference to the movement of  $\text{Ca}^{2+}$  over  $\text{Sr}^{2+}$ . Therefore, small changes in the reabsorption processes as a result of imposed variables or disease states will result in large effects on the relative urinary excretion of these alkaline earth cations. It seems apparent that strontium will continue to be used, and increasingly so, as a pseudo-calcium in clinical and physiological experimentation. Studies like that of Vezzoli *et al.* (Vezzoli *et al.*, 1998) will help define the advantages and disadvantages of strontium in this role.

The structural and binding affinity investigations on different calcium binding proteins to understand the flexibility and plasticity of EF hand motifs and how they can interact with various heavy metal ions is very limited. Among the EF-hand containing proteins, so far CaM structure has been determined in complex with  $\text{Pb}^{2+}$  (Wilson & Brunger., 2003; Kursula & Majava., 2007) and partly bound  $\text{Ba}^{2+}$  (Kursula

*et al.*, 2007). In CaM-Ba complex, only EF-hand motif 2 was bound with the barium ion (Kursula *et al.*, 2007).

The crystal structure of EhCaBP1 has been solved at 2.4Å resolution in which only N-terminal half has been traced and C-terminal domain is missing due to the presence of the flexible linker region between these two domains (Kumar *et al.*, 2007). The overall structure is trimeric in which EF-hand 1 of one molecule interacts with EF-hand 2 of the other molecule in a domain swapped manner to form an assembled domain similar to that of calmodulin N-terminal domain.

In this study the specificity and flexibility of EF-hand motifs of EhCaBP1 has been investigated in detail using three heavy atoms  $Pb^{2+}$ ,  $Ba^{2+}$  and  $Sr^{2+}$ . The calcium ions in EhCaBP1 has been successfully replaced with the  $Pb^{2+}$ ,  $Ba^{2+}$  and  $Sr^{2+}$  and the crystal structures of the complexes have been determined.

### 6.3 Results and Discussion

#### 6.3.1 $Pb^{2+}$ -EhCaBP1 complex

To study the plasticity of the calcium binding loop of EhCaBP1, it was co-crystallized with  $Pb^{2+}$  (Fig. 6.2) and the structure was determined (Materials and Methods, Section No. 2.24, Page No. 49).

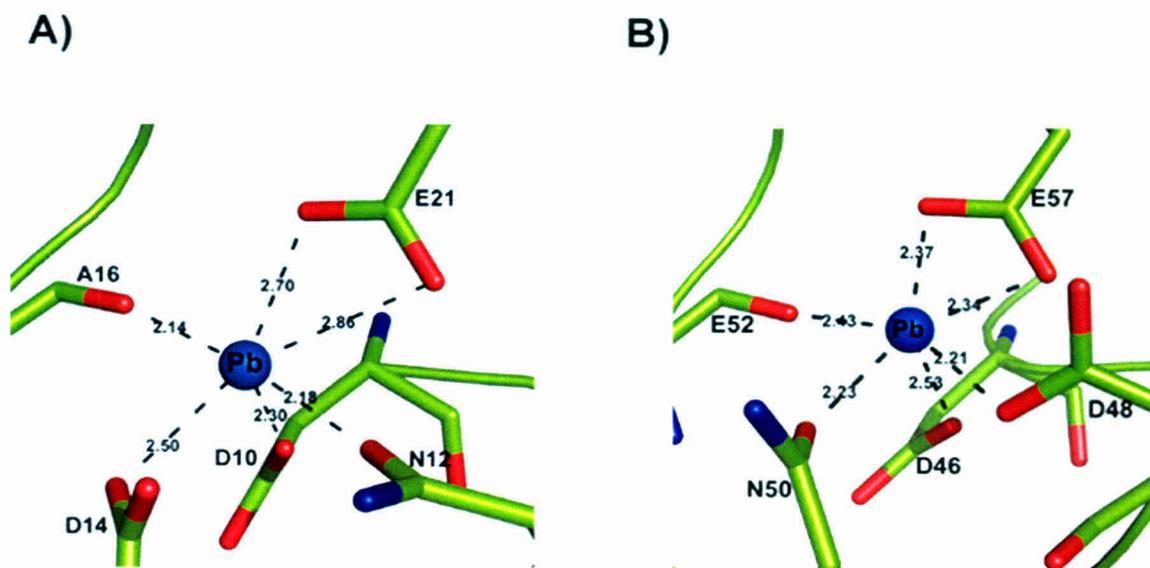


**Figure-6.2** Hexagonal crystal of  $Pb^{2+}$ -EhCaBP1 complex.

Similar to native protein structure, EF hand motifs were connected by a long helix and in both the EF-hands the calcium ions are replaced by lead.

### 6.3.2 Overall structure of $Pb^{2+}$ -EhCaBP1 complex

The final model refined to R\_factor of 26.5 % and Free\_R of 28.8%, respectively (Materials and Methods, Table 2.4, Section No. 2.24, Page No. 51). The overall conformation and metal-coordination geometry of  $Pb^{2+}$  is similar to that of  $Ca^{2+}$  in EhCaBP1 with minor differences. As the lead has one lone pair of electrons in the outermost shell, it interacts with six ligands compared to calcium, making it hemidirectional pentagonal pyramidal shape and  $Pb^{2+}$  interacts with one water molecule less compared to calcium bound structure (Fig.6.3).

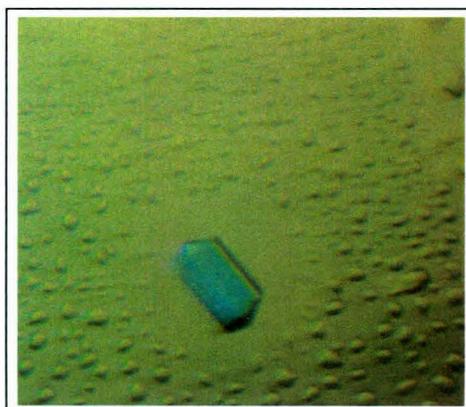


**Figure-6.3** *Lead coordination in EF-hand 1 and EF-hand 2 motifs of EhCaBP1. Lead co-ordinates with six of its ligands in the EF-hand motifs. The coordination of  $Pb^{2+}$  is similar to that of  $Ca^{2+}$  coordination. The coordination distances between the  $Pb^{2+}$  and oxygen atom for EF-hand 1 motif ranges between 2.18 Å and 2.86 Å and for EF-hand 2 motif ranges between 2.21 Å and 2.58 Å.*

The coordination distances vary little as compared to calcium bound structures but the average coordination distance of both calcium and lead is almost same in the EF-hand 1 and EF-hand 2 motifs. Also, due to the presence of flexible linker region between the N-terminal and C-terminal domain, the C-terminal region could not be traced in this complex structure similar to the native structure in which only the N-terminal region has been traced. No significant structural rearrangement occurs upon replacement and the coordination geometry is highly similar suggesting that Pb can act like Ca in binding and activating EhCaBP1. Similar coordination pattern was also seen earlier in Calmodulin-Pb complex structure (Kursula *et al.*, 2007).

### 6.3.3 Ba<sup>2+</sup>-EhCaBP1 Complex

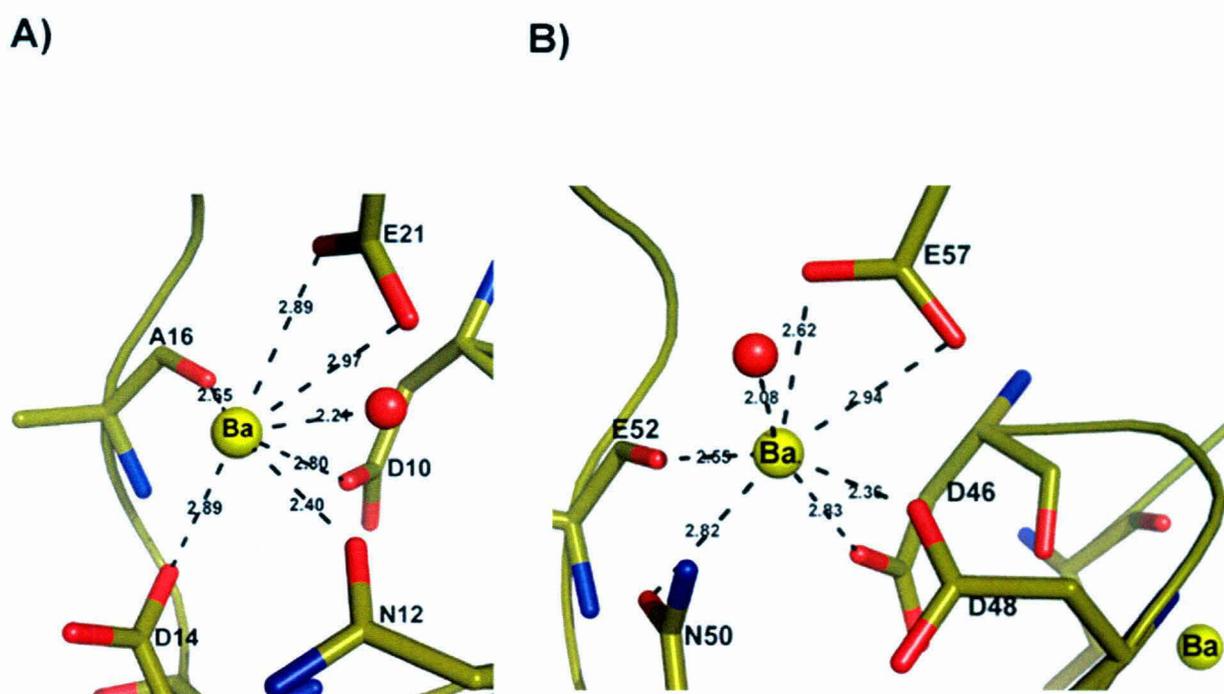
Barium is a small stable divalent ion (ionic radius 1.35 Å; Sharpe, 1992). The barium absorption edges (K edge 37.4 keV, L edges 5.2-6.0 keV) lie beyond the easily accessible wavelength range on commonly available synchrotron beamlines. However, barium has a significant anomalous signal component in the energy range 7–13 keV (wavelength range ~1.7–1.0 Å) which can be exploited for the phasing of protein structures. Barium has low binding affinity to EF hand motifs compared to calcium (Vyas *et al.*, 1989) and it cannot replace the calcium at low concentrations. For the replacement of calcium with barium in the calcium binding loop of EhCaBP1, the protein was denatured and refolded in barium containing buffer and crystallized using MPD as a precipitant (Fig.6.4) (Materials and methods, Section No. 2.25 , Page No. 51).



**Figure-6.4** Hexagonal crystal of Ba<sup>2+</sup>-EhCaBP1 complex surrounded with oiling.

### 6.3.4 Overall structure of Ba<sup>2+</sup>-EhCaBP1 complex

The final model refined to R\_factor of 26.1% and Free\_R of 28.1%, respectively (Materials and Methods, Table 2.5, Section No. 2.25, Page No. 52). The overall structure of Ba<sup>2+</sup>-EhCaBP1 complex structure was similar to native structure with minor differences. The coordination distances in Ba<sup>2+</sup>-EhCaBP1 are slightly higher compared to calcium bound structure with similar geometry. The average of barium coordination distances is about decisive 0.1 Å more than calcium coordinated distances (Fig.6.5).



**Figure-6.5 Barium coordination in EF-hand 1 and EF-hand 2 motifs of EhCaBP1.** Barium co-ordinates with seven of its ligands in the EF-hand motifs. The coordination of Ba<sup>2+</sup> is similar to that of Ca<sup>2+</sup>. Two barium ions are bound at the EF-hand 2 motif, one at the coordination sphere and another barium ion in the close vicinity of EF-hand 2 motif with lower occupancy co-ordinated with K43, D46 and D48 residues, respectively. The coordination distances between the ion and oxygen atom for EF-hand 1 motif ranges between 2.24 Å and 2.97 Å and for EF-hand 2 motif ranges between 2.08 Å and 2.94 Å.

But in EF2 motif, two barium sites have been observed; one at the same site as calcium and another barium atom was detected in the close vicinity of EF-hand 2 loop with lower occupancy coordinated by K43, D46 and D48 side-chain residues. This second barium seems to be partially occupied, which has weaker electron density and high temperature factor compared to the barium ion bound in the centre of EF hand loop. These two sites were confirmed by anomalous map and high intensity difference Fourier map before keeping barium. The fully occupied barium ions have difference Fourier electron density at about  $10\sigma$  and partially occupied barium ion has about  $5\sigma$  level. Anomalous signal evaluation of  $Ba^{2+}$ -EhCaBP1 complex and  $Pb^{2+}$ -EhCaBP1 complex have been compared in Table-5.1. Phenix AutoSol and AutoBuild output is compared for both the complexes.

**Table-5.1 Comparison of the anomalous signals and phasing power of barium and lead complexes for automated structure solution using Phenix.**

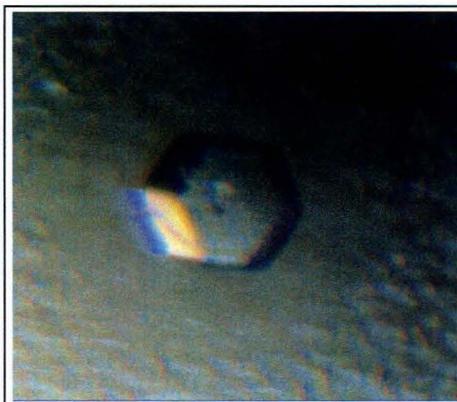
Data set	$Ba^{2+}$ -EhCaBP1 complex	$Pb^{2+}$ -EhCaBP1 complex
Wavelength	1.54179	1.54179
Anomalous $f''$	9.025	8.934
Anomalous signal	0.1014	0.0622
Bijvoet pairs	4987	7281
Lone Bijvoet mates	89	24
<b>AutoSol</b>		
No. of Heavy atom sites	6	4
FOM	0.38	0.18
CC	0.79	0.41
R/Free_R (%)	41.6/43.7	53.0/56.0
<b>AutoBuild</b>		
Residues built (%)	73	65
Map CC	0.72	0.64
R/Free_R	41.6/43.7	47.0/55.0

Similar binding pattern was also observed in EF-hand 2 of barium soaked calmodulin crystals, even though barium was not bound to other EF hand motifs (Kursula *et al.*, 2007). This second barium ion at EF2 motif appears to be in the direction from which

the metal ion can enter or release from its binding pocket. It is possible that this second ion at EF-hand 2 represents a state where an ion is just about to be bound/released.

### 6.3.5 $\text{Sr}^{2+}$ -EhCaBP1 complex

Proteins with the ability to specifically bind strontium would potentially be of great use in the field of nuclear waste management. Unfortunately, no such peptides or proteins are known and it is uncertain whether they exist under natural conditions due to low environmental concentrations of strontium. Strontium has lower binding affinity compared to calcium and lead but it has better binding affinity than barium (Habermann *et al.*, 1983) and also has comparable cation size and hydration energy to calcium (Vyas *et al.*, 1989). Calcium was replaced by strontium using similar process of barium, where the protein was denatured in urea and refolded in strontium containing buffer and crystallized using MPD as a precipitant (Fig.6.6) (Materials and methods, Section No. 2.26, Page No. 53).

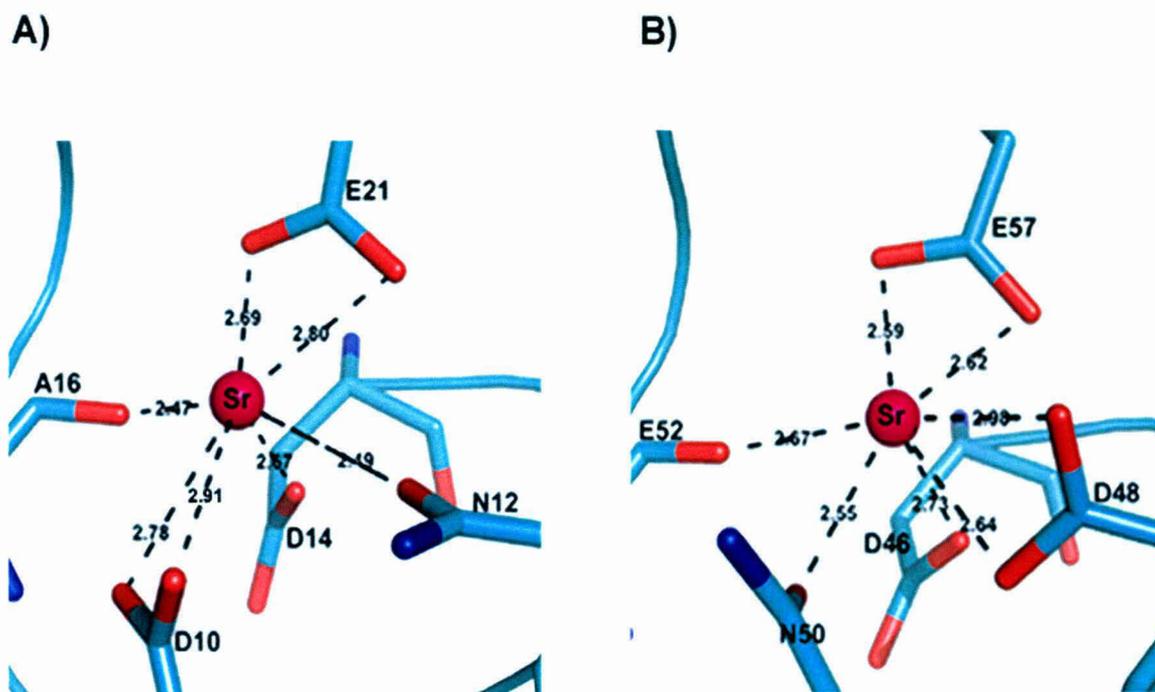


**Figure- 6.6** *Hexagonal crystal of  $\text{Sr}^{2+}$ -EhCaBP1 complex.*

### 6.3.6 Overall structure of $\text{Sr}^{2+}$ -EhCaBP1 complex

The final model refined to R\_factor of 25.9 % and Free\_R of 28.6 %, respectively (Materials and methods, Table No. 2.6, Section No. 2.26, Page No. 53). Overall structure of strontium bound CaBP1 was similar to native calcium bound protein

structure except strontium coordination. Water molecule was not observed in the coordination calcium binding loop of  $\text{Sr}^{2+}$ -EhCaBP1 complex, as it was seen in calcium and barium bound structures (Fig.6.7).



**Figure-6.7** *Strontium coordination in EF-hand 1 and EF-hand 2 motifs of EhCaBP1.* Strontium co-ordinates with seven of its ligands in the EF-hand motifs. The coordination of  $\text{Sr}^{2+}$  is similar to that of  $\text{Ca}^{2+}$  coordination except for the absence of water molecule in the coordination. The 3<sup>rd</sup> aspartate residue is donating both of its oxygen ligands to satisfy the coordination geometry. The coordination distances between the ion and oxygen atom for EF-hand 1 motif ranges between 2.18 Å and 2.86 Å and for EF-hand 2 motif ranges between 2.21 Å and 2.58 Å.

The coordination distances with  $\text{Sr}^{2+}$  are marginally higher compared to  $\text{Ca}^{2+}$  coordination distances. The average coordination distance is about 0.03 Å more than average calcium coordination distances in both EF1 and EF2 hand motifs. The 3<sup>rd</sup> aspartate residue of the calcium binding loop was donating both of its oxygen atoms and hence, compensating for the loss of water coordination.

### 6.3.7 Anomalous signal analysis and verification of heavy atom binding

Barium complex data had good anomalous signal of 0.1014, lead complex had anomalous of 0.0622 while strontium complex data had no anomalous signal at the respective wavelengths, at which the data sets were collected. Calculated anomalous signal followed the pattern of  $f'$  and number of heavy atom binding sites in the protein. As the  $Ba^{2+}$  is bound at three sites (two ions with full occupancy and another ion with partial occupancy), this data showed the highest anomalous signal.

### 6.3.8 Phasing with anomalous signal

The Barium complex data had mean anomalous signal of 0.1034 and the anomalous scatterers were located in *AutoSol* (Adams *et al.*, 2002). A total of six  $Ba^{2+}$  ions were identified in asymmetric unit with a final figure of merit (*FOM*) of 0.38. A preliminary atomic model was built automatically; the partial model obtained from *AutoSol* in *PHENIX* was fed into *AutoBuild* for iterative model building and refinement, resulting in a model with 70% of the residues and  $R_{work}/R_{free}$  of 42/43%. This was achieved without any manual intervention in model building, clearly showing the phasing quality of the barium ions. This clearly indicates that barium incorporation and data collection using Cu  $K\alpha$  X-ray radiation is an easy approach towards the structure solution of Calcium Binding Proteins.

## 6.4 Conclusion

The calcium ions of EhCaBP1 have been successfully replaced with other heavy atoms like lead, barium and strontium, crystallized and the structure has been determined. The difference Fourier electron density and anomalous signal clearly indicate the presence of these heavy metal ions in place of calcium at calcium binding loops in the respective crystal structures. The overall conformation and metal-coordination geometry of  $Pb^{2+}$ ,  $Sr^{2+}$  and  $Ba^{2+}$  in EhCaBP1 are similar to that of  $Ca^{2+}$ -EhCaBP1, thus providing a structural rationale for the ability of  $Pb^{2+}$ ,  $Sr^{2+}$  and  $Ba^{2+}$  to bind and to activate some of the same targets as  $Ca^{2+}$ -EhCaBP1.

Lead has been reported several times to be the element most able to act like calcium in CaM activation. No significant structural rearrangements occur upon

replacement of calcium with lead and the coordination geometry is similar. As the binding affinity of lead is high, it is tightly bound and it is not released from EF hand motif at lower concentrations unlike  $\text{Ca}^{2+}$ , thus it might be keeping calmodulin always in active state. It seems that lead has similar effect on EhCaBP1. As lead is having more affinity towards the calcium binding site and can replace calcium very easily, it can result in the false activation of the downstream signaling pathways with abnormal effects.

The structural analysis of EhCaBP1 in complex with strontium shows that the  $\text{Sr}^{2+}$ -coordination is almost similar to that of native  $\text{Ca}^{2+}$  bound structure. The difference lies in the co-ordinating ligands. There is no water molecule in the coordination and the 3<sup>rd</sup> aspartate residue is donating both of its oxygen ligands for the coordination in a non-canonical fashion. Structure of any calcium-binding protein in complex with strontium is being reported for the first time.

There were several attempts to replace  $\text{Ca}^{2+}$  in EF hand containing proteins, including CaM with  $\text{Ba}^{2+}$ , without much success, except  $\text{Ca}^{2+}$  replacement with  $\text{Ba}^{2+}$  only in EF2 motif of calmodulin (Kursula *et al.*, 2007). In  $\text{Ba}^{2+}$ -EhCaBP1 structure both the EF hand motifs are replaced with  $\text{Ba}^{2+}$  and average coordination distance is about 0.1 Å more than the  $\text{Ca}^{2+}$  bound structure.

The crystallographic data contains detailed information about the molecular flexibility that can be extracted from even moderate-resolution data using modern refinement protocols. Furthermore, it suggests that although protein mobility is limited in the crystalline environment, lattice constraints do not appear to alter the fundamental spatial characteristics of protein flexibility. These results clearly indicate, EhCaBP1 is also capable of binding and being activated by other metal ions and help in understanding the activation mechanisms of EhCaBP1 by different heavy metals and also give more insight into the flexibility of the calcium-binding loop of EhCaBP1.

Also, the structure determination of EhCaBP1 with help of barium derived phases revealed that Barium could produce significant phases using in house X-ray source. This could effectively reduce the structure determination time for calcium binding proteins and other proteins, which can bind barium without changing the

protein structure and can offer significant anomalous signal at a Cu K $\alpha$  X-ray radiation.

The heavy metal complex structures are very similar to native structure with minor differences in EF- hand motifs. The difference lies in the coordination distances between the metals and the ligands, which is maximum in case of barium containing crystals. Ba<sup>2+</sup>-EhCaBP1 complex data collected at Cu K $\alpha$  X-ray source showed very good anomalous signal which could be used for SAD phasing and with the help of direct methods, 70% of the model was traced. This clearly indicates that the barium incorporation or co-crystallization of proteins with barium and data collection using home source is an easy approach towards the structure solution of Calcium Binding Proteins.

These results also have relevance to heavy metal toxicity and giving a structural and biochemical insights into how these metal ions can effectively substitute Ca<sup>2+</sup> in a molecule that is central to several regulatory processes.

# SUMMARY

---

## **Chapter-I**

### **Crystal structure of Calcium Binding Protein-1 from *Entamoeba histolytica*.**

- The crystal structure of EhCaBP1 has been determined at 2.4 Å resolution. It belongs to space group P6<sub>3</sub> with two molecules in an asymmetric unit. The crystal structure displays an unusual molecular organization and EF-hand domain organization.
- EhCaBP1 exists as trimer both in solution and in the crystal, where three molecules interact with each other in head-to-tail manner and forms a domain-swapped trimer.
- The trimeric state observed in the crystal structure, is the most probable functional and natural state of the protein which forms a hydrophobic pockets, very similar to the ones seen in CaM that may be involved in target binding are formed at the interface of two molecules.
- The functional difference between EhCaBP1 and CaM can be explained by the difference in the flexibility of the helix that connects the EF-hand domains.
- In CaM, the two EF-hand domains are connected by a flexible linker helix that permits greater conformational flexibility and is largely responsible for CaM's ability to bind to a large array of targets.
- The crystallographic and biochemical results expand the range of structural diversity observed for the EF-hand family and suggest that oligomerization of EhCaBP1 is both essential for target binding and may restrict the structural diversity of the target proteins. These results represent fundamental advances in our understanding of EF-hand comprising proteins, their structure-function relationship.

## **Chapter-II**

### **Crystal structure of EhCaBP1 in complex with phenylalanine: Visualization of mode of target binding.**

- Domain-swapping and trimerization of EhCaBP1 allows the EF1-hand motif of one molecule to interact with EF2 of an adjacent molecule to form a two EF-hand domain (assembled-domain) similar to that of calmodulin and troponin C.
- The hydrophobic pocket formed at the interface between EF1 and EF2 in the assembled-domain is bound to phenylalanine.
- The phenylalanine forms several hydrophobic interactions with Ile 9, Phe 25, Val 22 and Val26 residues of EF-hand motif 1 of one molecule and Tyr 62, Phe 61, Phe 58, Ile 41 residues of EF-hand motif 2 of another molecule.
- The assembled-domain bound to Phe was superimposed with the N-terminal domain Calmodulin bound to hydrophobic cardiac IQ motif of calcium channel (Cardiac calcium channel). Both structures superimposed very well with an rms deviation of 1.23 Å and the hydrophobic residue Phe of the peptide bound to CaM N-terminal domain and Phe bound to EhCaBP1 assembled-domain are located at similar location. Both the ligands bind their targets in similar manner at the hydrophobic pockets of respective domains.
- The distance between two assembled-domains is approximately same as the distance between calmodulin N and C-terminal domains. But the assembled-domain can't change its structure after binding to target as calmodulin wraps around its target. This rigidity in the trimeric structure may be responsible for differential recognition of the targets.
- It is evident from the biochemical data and the structure that only N-terminal domain is involved in target binding and activating kinase and the calcium binding affinity of the EF3 and EF4 are much higher than the EF1 and EF2 which makes clear that only EF1 and EF2 are affected by the calcium concentration fluctuations around it and helps it in target recognition and binding, whereas the C-terminal domain, which

consists of EF3 and EF4 motifs, should be rigid and may not be effected by calcium concentration changes.

### **Chapter-III**

#### **Crystal structure of N-terminal domain of EhCaBP1.**

- The crystal structure of N-terminal domain has been solved at 2.5 Å resolution. It belongs to space group P3 with four molecules in the asymmetric unit.
- The overall structure of the N-terminal domain is more refined as compared to the full-length structure. (The crystallographic R\_factor and R\_free of N-terminal domain is 23.3%, R\_Free = 27.1%, respectively which is comparatively lower than that of the full-length crystal structure ( R\_factor = 25.6%, R\_Free = 28.2%).
- The crystal structure of N-terminal domain of EhCaBP1 is similar to that of the full-length protein in which only the N-terminal region has been traced.
- The biochemical observations indicate that the N-terminal half of the protein is capable of carrying out most of the functions of the full molecule, but it could not bind to G-actin. The differential behaviour suggests that the two halves of the molecule have different functions.
- The behaviour of C-terminal cells can also be due to alteration in the property of C-terminal due to its fusion with GFP, the latter being much larger than the former.
- This is also evident from the results obtained from crystallization studies that N-terminal is the functional domain and plays an important role in target recognition. Therefore, the evolution of the CaBPs, such as EhCaBP1 may have been designed to offer both functional and structural diversity suitable for a pathogen to modulate host-pathogen relationship.

**Chapter-IV****Flexibility and plasticity of calcium binding loops of EhCaBP1: Crystal structure of EhCaBP1 in complex with  $Pb^{2+}$ ,  $Ba^{2+}$  and  $Sr^{2+}$ .**

- The calcium ions of EhCaBP1 have been successfully replaced with other heavy atoms like lead, barium and strontium, crystallized and the structure has been determined.
- The overall conformation and metal-coordination geometry of  $Pb^{2+}$ ,  $Sr^{2+}$  and  $Ba^{2+}$  in EhCaBP1 are similar to that of  $Ca^{2+}$ -EhCaBP1, thus providing a structural rationale for the ability of  $Pb^{2+}$ ,  $Sr^{2+}$  and  $Ba^{2+}$  to bind and to activate some of the same targets as  $Ca^{2+}$ -EhCaBP1.
- No significant structural rearrangements occur upon replacement of calcium with lead and the coordination geometry is similar except  $Pb^{2+}$  makes six coordination interactions as it has lone pair of electrons in its outermost shell. It seems that lead has similar effect on EhCaBP1.
- The structural analysis of EhCaBP1 in complex with strontium shows that the  $Sr^{2+}$  coordination is almost similar to that of native  $Ca^{2+}$  bound structure. The difference lies in the co-ordinating ligands. There is no water molecule in the coordination and the 3<sup>rd</sup> Aspartate residue is donating both of its oxygen ligands for the coordination in a non-canonical fashion. Structure of any calcium-binding protein in complex with strontium is being reported for the first time.
- There were several attempts to replace  $Ca^{2+}$  in EF hand containing proteins, including CaM with  $Ba^{2+}$ , without much success, except  $Ca^{2+}$  replacement with  $Ba^{2+}$  only in EF2 motif of calmodulin (Kursula *et al.*, 2007). In  $Ba^{2+}$ -EhCaBP1 structure both the EF hand motifs are replaced with  $Ba^{2+}$  and average coordination distance is about 0.1 Å more than the  $Ca^{2+}$  bound structure.
- These results clearly indicate, EhCaBP1 is also capable of binding and being activated by other metal ions and help in understanding the activation mechanisms of EhCaBP1

by different heavy metals and also give more insight into the flexibility of the calcium-binding loop of EhCaBP1.

- Ba<sup>2+</sup>-EhCaBP1 complex data collected at Cu K $\alpha$  X-ray source showed very good anomalous signal which could be used for SAD phasing and with the help of direct methods, 70% of the model was traced. This clearly indicates that the barium incorporation or co-crystallization of proteins with barium and data collection using home source is an easy approach towards the structure solution of Calcium Binding Proteins. This could effectively reduce the structure determination time for calcium binding proteins and other proteins, which can bind barium without changing the protein structure and can offer significant anomalous signal at home source.
  
- The heavy metal complex structures are very similar to native structure with minor differences in EF- hand motifs. The difference lies in the coordination distances between the metals and the ligands, which is maximum in case of barium containing crystals.
  
- These results also have relevance to heavy metal toxicity and giving a structural and biochemical insights into how these metal ions can effectively substitute Ca<sup>2+</sup> in a molecule that is central to several regulatory processes.

# REFERENCES

---

- Abrahamsen, M.S., Templeton, T.J., Enomoto, S., Abrahante, J.E., Zhu, G., Lancto, C.A., Deng, M., Liu, C., Widmer, G., Tzipori, S., Buck, G.A., Xu, P., Bankier, A.T., Dear, P.H., Konfortov, B.A., Spriggs, H.F., Iyer, L., Anantharaman, V., Aravind, L. and Kapur, V. (2004). Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science*. 304, 441-445.
- Abuabara, S.F., Barrett, J.A., Hua, T. and Jonasson, O. (1982). Amebic liver abscess. *Arch. Surg.* 117, 239-244.
- Acuna-Soto, R., Maguire, J.H. and Wirth, D.F. (2000). Gender distribution in asymptomatic and invasive amebiasis. *Am. Gastroenterol.* 95, 1277-1283.
- Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002). PHENIX, building new software for automated crystallographic structure determination. *Acta cryst.* D58, 1948-1954.
- Adb-Alla, M.D. and Ravdin, J.I. (2002). Diagnosis of amoebic colitis by antigen capture ELISA in patients presenting with acute diarrhoea in Cairo, Egypt. *Trop. Med. Int. Health.* 7, 365-370.
- Aitio, H., Laakso, T., Pihlajamaa, T., Torkkeli, M., Kilpelainen, I., Drakenberg, T., Serimaa, R. and Annala, A. (2001). Characterization of apo and partially saturated states of calerythrin, an EF-hand protein from *S. erythraea*, a molten globule when deprived of  $Ca^{2+}$ . *Protein Sci.* 10, 74-82.
- Akke, M., Skelton, N.J., Kordel, J., Palmer, 3rd, A. G. and Chazin, W.J. (1993). Effects of ion binding on the backbone dynamics of calbindin D9k determined by 15N NMR relaxation. *Biochemistry.* 32, 9832-9844.
- Allanson-Jones, E., Mindel, A., Sargeant, P. and Williams, P. (1986). *Entamoeba histolytica* as a commensal intestinal parasite in homosexual men. *N. Engl. J. Med.* 315, 353-356.

- Ames, J.B., Dizhoor, A.M., Ikura, M., Palczewski, K. and Stryer, L. (1999). Three-dimensional structure of guanylyl cyclase activating protein-2, a calcium-sensitive modulator of photoreceptor guanylyl cyclases. *J. Biol. Chem.* 274, 19329–19337.
- Ames, J.B., Hendricks, K.B., Strahl, T., Huttner, I.G., Hamasaki, N. and Thorner, J. (2000). Structure and calcium-binding properties of Frq1, a novel calcium sensor in the yeast *Saccharomyces cerevisiae*. *Biochemistry.* 39, 12149–12161.
- Anand, A.C. (1997). Is it curtains for chronic amebiasis? *Indian. J. Gastroenterol.* 16, 127-129.
- Ankri, S., Bracha, R., Padilla-Vaca, F. and Mirelman, D. (1999). Applying antisense technology to the study of *Entamoeba histolytica* pathogenesis, response. *Trends Microbiol.* 7, 473-474.
- Apostoli, P., Bellini, A., Porru, S. and Bisanti, L. (2000). The effect of lead on male fertility, a time to pregnancy (TTP) study. *Am. J. Ind. Med.* 38, 310–315.
- Aramini, J.M., Hiraoki, T., Grace, M.R., Swaddle, T.W., Chiancone, E. and Vogel, H.J. (1996). NMR and stopped-flow studies of metal ion binding to alpha-lactalbumins. *Biochim. Biophys. Acta.* 1293, 72-82.
- Atchison, W.D. (2003). Effects of toxic environmental contaminants on voltage-gated calcium channel function, from past to present. *J. Bioenergy Biomembr.* 35, 507–532.
- Atreya, H.S., Sahu, S.C., Bhattacharya, A., Chary, K.V. and Govil, G. (2001). NMR derived solution structure of an EF-hand calcium-binding protein from *Entamoeba histolytica*. *Biochemistry.* 40, 14392–14403.

- Babini, E., Bertini, I., Capozzi, F., Luchinat, C., Quattrone, A. and Turano, M. (2005). Principal component analysis of the conformational freedom within the EF-hand superfamily. *J. Proteome Res.* 4, 1961–1971.
- Babor, M., Greenblatt, H.M., Edelman, M. and Sobolev, V. (2005). Flexibility of metal binding sites in proteins on a database scale. *Proteins.* 59, 221–230.
- Babu, Y.S., Sack, J.S., Greenhough, T.J., Bugg, C.E., Means, A.R. and Cook, W.J. (1985). Three-dimensional structure of calmodulin. *Nature.* 315, 37–40.
- Babu, Y.S., Bugg, C.E. and Cook, W.J. (1988). Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.* 204, 191–204.
- Bagchi, A., Bhattacharya, A. and Bhattacharya, S. (1999). Lack of a chromosomal copy of the circular rDNA plasmid of *Entamoeba histolytica*. *Int. J. Parasitol.* 29, 1775–83.
- Bagley, S.C. and Altman, R.B. (1995). Characterizing the microenvironment surrounding protein sites. *Protein Sci.* 4, 622–635.
- Bairoch, A. and Cox, J.A. (1990). EF-hand motifs in inositol phospholipid-specific phospholipase C. *FEBS Lett.* 269, 454–456.
- Barnes, P.F., DeCock, K.M., Reynolds, T.N. and Ralls, P.W. (1987). A comparison of amebic and pyrogenic abscess of the liver. *Medicine. (Baltimore).* 66, 472–83.
- Batanghari, J.W., Deepe, G.S., Jr, di Cera, E. and Goldman, W.E. (1998). Histoplasma acquisition of calcium and expression of CBP1 during intracellular parasitism. *Mol. Microbiol.* 27, 531–539.

- Bennett, D.L., Cheek, T.R., Berridge, M.J., De Smedt, H., Parys, J.B., Missiaen, N.L. and Bootman, M.D. (1996). Expression and function of ryanodine receptors in non-excitable cells. *J. Biol. Chem.* 11, 6356–6362.
- Berridge M.J., Lipp P. and Bootman M.D. (2000). The versatility and universality of calcium signalling. *Nature reviews.* 1, 11-21.
- Berridge, M.J. (1993). Inositol trisphosphate and calcium signalling. *Nature.* 361, 315–325.
- Berridge, M.J., Lipp, P. and Bootman, M.D. (2000). The versatility and universality of calcium signalling. *Nat. Rev. Mol. Cell Biol.* 1, 11-21.
- Berridge, M.J., Bootman, M.D. and Roderick, H.L. (2003). Calcium signalling, dynamics, homeostasis and remodelling. *Nat. Rev. Mol. Cell. Biol.* 4, 517–529.
- Bezprozvanny, I. and Ehrlich, B.E. (1995). The inositol 1,4,5- trisphosphate (InsP3) receptor. *J. Membr. Biol.* 145, 205–216.
- Bhattacharya, A., Padhan, N., Jain, R. and Bhattacharya, S. (2006). Calcium binding proteins of *Entamoeba histolytica*. *Arch. Med. Res.* 37, 221–225.
- Biekofsky, R.R. and Feeney, J. (1998). Co-operative cyclic interactions involved in metal binding to pairs of sites in EF-hand proteins. *FEBS Lett.* 439, 101–106.
- Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7, 1513-23.
- Blanchard, H., Grochulski, P., Li, Y., Arthur, J.S., Davies, P.L., Elce, J.S. and Cygler, M. (1997). Structure of a calpain Ca<sup>2+</sup>-binding domain reveals a novel EF-hand and Ca<sup>2+</sup>-induced conformational changes. *Nat. Struct. Biol.* 4, 532–538.

- Bootman, M.D. and Berridge, M.J. (1995). The elemental principles of  $\text{Ca}^{2+}$  signaling. *Cell*. 83, 675–678.
- Borremans, B., Hobman, J.L., Provoost, A., Brown, N.L. and van Der Lelie, D. (2001). Cloning and Functional Analysis of the pbr Lead Resistance Determinant of *Ralstonia metallidurans* CH34. *J. Bacteriol.* 183, 5651–5658.
- Bouton, C.M., Frelin, L.P., Forde, C.E., Arnold Godwin, H. and Pevsner, J. (2001). Synaptotagmin I is a molecular target for lead. *J. Neurochem.* 76, 1724–1735.
- Bracha, R., Kobilier, D. and Mirelman, D. (1982). Attachment and ingestion of bacteria by trophozoites of *Entamoeba histolytica*. *Infect. Immun.* 36, 396–406.
- Brandi M.L. (1993). New treatment strategies, ipriflavone, strontium, vitamin D metabolites and analogs. *Am. J. Med.* 95, 69S–74S.
- Brumpt, M.E. (1925). Etude sommaire de l' <<*Entamoeba dispar*>> n sp Amibe a kystes quadrinucléés, parasite de l'homme. *Bull.cad. Méd. (Paris)*. 94, 943-52.
- Brunet, S., Scheuer, T., Klevit, R. and Catterall, W.A. (2005). Modulation of  $\text{CaV}1.2$  channels by  $\text{Mg}^{2+}$  acting at an EF-hand motif in the COOH-terminal domain. *J. Gen. Physiol.* 126, 311–323.
- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998). Crystallography and NMR system, A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* 54, 905-921.
- Buchanan, J.D., Corbett, R.J. and Roche, R.S. (1986). The thermodynamics of calcium binding to thermolysin. *Biophys. Chem.* 23, 183-199.

- Burleigh, B.A. and Andrews, N.W. (1998). Signaling and host cell invasion by *Trypanosoma cruzi*. *Curr. Opin. Microbiol.* 1, 461-465.
- Busenlehner, L.S. and Giedroc, D.P. (2006). Kinetics of metal binding by the toxic metal-sensing transcriptional repressor *Staphylococcus aureus* pI258 CadC. *J. Inorg. Biochem.* 100, 1024–1034.
- Camacho, P. (2003). Malaria parasites solve the problem of a low calcium environment. *J. Cell Biol.* 161, 17-19.
- Cannell, M.B., Cheng, H. and Lederer, W.J. (1995). The control of Ca<sup>2+</sup> release in heart muscle. *Science.* 268, 1045–1049.
- Capozzi, F., Casadei, F. and Luchinat, C. (2006). EF-hand protein dynamics and evolution of calcium signal transduction, an NMR view. *J. Biol. Inorg. Chem.* 11, 949–962.
- Carafoli, E. (1994). The signaling function of calcium and its regulation. *J. Hypertens. Suppl.* 12, S47–S56.
- Carafoli, E. (2004). Special issue, calcium signaling and disease. *Biochem. Biophys. Res. Commun.* 322, 1097-1109.
- Carbajal, M.E., Manning-Cela, R., Pina, A., Franco, E. and Meza, I. (1996). Fibronectin-induced intracellular calcium rise in *Entamoeba histolytica* trophozoites: effect on adhesion and the actin cytoskeleton. *Exp. Parasitol.* 82, 11-20.
- Chakrabarty, P., Sethi, D.K., Padhan, N., Kaur, K.J., Salunke, D.M., Bhattacharya, S. and Bhattacharya, A. (2004). Identification and characterization of EhCaBP2 A second member of the calcium-binding protein family of the protozoan parasite *Entamoeba histolytica*. *J. Biol. Chem.* 279, 12898–12908.

- Chao, S.H., Bu, C.H. and Cheung, W.Y. (1995). Stimulation of myosin light-chain kinase by  $\text{Cd}^{2+}$  and  $\text{Pb}^{2+}$ . *Arch. Toxicol.* 69, 197–203.
- Chao, S.H., Suzuki, Y., Zysk, J.R. and Cheung, W.Y. (1984). Activation of calmodulin by various metal cations as a function of ionic radius. *Mol. Pharmacol.* 26, 75–82.
- Chapman, A., Vallejo, V., Mossie, K.G., Ortiz, D., Agabian, N. and Flisser, A. (1995). Isolation and characterization of species-specific DNA probes from *Taenia solium* and *Taenia saginata* and their use in an egg detection assay. *J. Clin. Microbiol.* 33, 1283-8.
- Charbonnier, J.B., Christova, P., Shosheva, A., Stura, E., Du, M.H.L., Blouquit, Y., Duchambon, P., Mironc, S. and Craescuc, C.T. (2006). *Acta Cryst. F62*, 649-651.
- Chattopadhyaya, R., Meador, W.E., Means, A.R. and Quioco, F.A. (1992). Calmodulin structure refined at 1.7 Å resolution. *J. Mol. Biol.* 228, 1177-1192.
- Cheek, T.R. (1991).  $\text{Ca}^{2+}$  regulation and homeostasis. *Curr. Opin. Cell. Biol.* 3, 199–205.
- Chin, D. and Means, A.R. (2000). Calmodulin: A prototypical calcium sensor. *Trends. Cell. Biol.* 10, 322-328.
- Chou, J.J., Li, S., Klee, C.B. and Bax, A. (2001). Solution structure of  $\text{Ca}^{2+}$ -calmodulin reveals flexible hand-like properties of its domains. *Nat. Struct. Biol.* 8, 990–997.

- Christova, P., Cox, J.A. and Craescu, C.T. (2000). Ion-induced conformational and stability changes in Nereis sarcoplasmic calcium binding protein, evidence that the APO state is a molten globule. *Proteins*. 40, 177–184.
- Clapham, D.E. (1995). Calcium signaling. *Cell*. 80, 259–268.
- Clark, C.G. (2000). Cryptic genetic variation in parasitic protozoa. *J. Med. Microbiol.* 49, 489-91.
- Cline, H.T., Witte, S. and Jones, K.W. (1996). Low lead levels stunt neuronal growth in a reversible manner. *Proc. Natl Acad. Sci. USA*. 93, 9915–9920.
- Colombo, M.I., Beron, W. and Stahl, P.D. (1997). Calmodulin regulates endosome fusion. *J. Biol. Chem.* 272, 7707-7712.
- Cook, W.J., Jeffrey, L.C., Cox, J.A. and Vijay-Kumar, S. (1993). Structure of a sarcoplasmic calcium-binding protein from amphioxus refined at 2.4 Å resolution. *J. Mol. Biol.* 229, 461–471.
- Cowan, J.A. (1998). Metal Activation of Enzymes in Nucleic Acid Biochemistry. *Chem. Rev.* 98, 1067–1088.
- Cox, J.A., Durussel, I., Scott, D.J. and Berchtold, M.W. (1999). Remodeling of the AB site of rat parvalbumin and oncomodulin into a canonical EF-hand. *Eur. J. Biochem.* 264, 790–799.
- Crivici, A. & Ikura, M. (1995). Molecular and structural basis of target recognition by calmodulin. *Annu. Rev. Biophys. Biomol. Struct.* 24, 85-115.
- da Silva, J.J.R., Williams, R.J.P. (1991). The biological chemistry of the elements. *Oxford, Oxford University Press*. 1, 478–81.

- de Beer, T., Carter, R.E., Lobel-Rice, K.E., Sorkin, A. and Overduin, M. (1998). Structure and Asn-Pro-Phe binding pocket of the Eps15 homology domain. *Science*. 281, 1357–1360.
- De Meester, F., Mirelman, D., Stolarsky, T. and Lester, D.S. (1990). Identification of protein kinase C and its potential substrate in *Entamoeba histolytica*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 97, 707-711.
- DeBakey, M.E., Ochsner, A., Shandera, W.X., Bollam, P., Hashmey, R.H. Jr., Athey, P.A., Greenberg, S.B. and White, A.C. Jr. (1998). Hepatic amebiasis among patients in a public teaching hospital. *South Med.* 91, 829-37.
- Debreczeni, J.E., Farkas, L., Harmat, V., Hetenyi, C., Hajdu, I., Zavodszky, P., Kohama, K. and Nyitray, L. (2005). Structural evidence for non-canonical binding of Ca<sup>2+</sup> to a canonical EF-hand of a conventional myosin. *J. Biol. Chem.* 280, 41458–41464.
- DeLano WL. (2002). The PyMOL User's Manual. San Carlos, CA: *DeLano Scientific*.
- Desrivieres, S., Cooke, F.T., Morales-Johansson, H., Parker, P.J. and Hall, M.N. (2002). Calmodulin controls organization of the actin cytoskeleton via regulation of phosphatidylinositol (4,5)-bisphosphate synthesis in *Saccharomyces cerevisiae*. *Biochem. J.* 366, 945-951.
- Dhar, S.K., Choudhury, N.R., Mittal, V., Bhattacharya, A. and Bhattacharya, S. (1996). Replication initiates at multiple dispersed sites in the ribosomal DNA plasmid of the protozoan parasite *Entamoeba histolytica*. *Mol. Cell. Biol.* 16, 2314-24.

- Diamond, L.S. and Clark, C.G. (1993). A redescription of *Entamoeba histolytica* Schaudinn, 1903 (Emended Walker, 1911) separating it from *Entamoeba dispar* Brumpt, 1925. *J. Eukaryot. Microbiol.* 40, 340-344.
- Dudev, T. and Lim, C. (2003). Principles governing Mg, Ca, and Zn binding and selectivity in proteins. *Chem. Rev.* 103, 773–788.
- Dudev, T., Lin, Y.L., Dudev, M. and Lim, C. (2003). First-second shell interactions in metal binding sites in proteins, a PDB survey and DFT/CDM calculations. *J. Am. Chem. Soc.* 125, 3168–3180.
- Durussel, I., Luan-Rilliet, Y., Petrova, T., Takagi, T. and Cox, J. A. (1993). Cation binding and conformation of tryptic fragments of Nereis sarcoplasmic calcium-binding protein, calcium-induced homo- and heterodimerization. *Biochemistry.* 32, 2394–2400.
- Dvorak, J.A., Kobayashi, S., Alling, D. and Hallahan, C.W. (1995). Elucidation of the DNA synthetic cycle of *Entamoeba* spp. using flow cytometry and mathematical modeling. *J. Eukaryot. Microbiol.* 42, 610-6.
- Eid, J.E. and Sollner-Webb, B. (1991). Homologous recombination in the tandem calmodulin genes of *Trypanosoma brucei* yields multiple products: compensation for deleterious deletions by gene amplification. *Genes Dev.* 5, 2024-2032.
- Einspahr, H., Bugg, C.E. and Sigel, H. (1984). Calcium and its role in biology, in, M. Dekker (Ed.), *Crystal Structure Studies of Calcium Complexes and Implications for Biological Systems.* New York. 17, 52–97.
- Elshorst, B., Hennig, M., Forsterling, H., Diener, A., Maurer, M., Schulte, P., Schwalbe, H., Griesinger, C., Krebs, J., Schmid, H., Vorherr, T. and Carafoli,

- E. (1999). NMR Solution Structure of a Complex of Calmodulin with a Binding Peptide of the Ca<sup>2+</sup> Pump. *Biochemistry*. 38, 12320-12332.
- Emsley, P. and Cowtan, K. (2004). Coot, model-building tools for molecular graphics. *Acta. Cryst. D60*, 2126-2132.
- Espinosa-Cantellano, M., and Martinez-Palomo, A. (1991). The plasma membrane of *Entamoeba histolytica*, structure and dynamics. *Biol. Cell*. 72, 189-200.
- Evenas, J., Malmendal, A. and Akke, M. (2001). Dynamics of the transition between open and closed conformations in a calmodulin C-terminal domain mutant. *Structure*. 9, 185–195.
- Evenas, J., Malmendal, A., Thulin, E., Carlstrom, G. and Forsen, S. (1998). Ca<sup>2+</sup> binding and conformational changes in a calmodulin domain. *Biochemistry*. 37, 13744–13754.
- Evenas, J., Thulin, E., Malmendal, A., Forsen, S. and Carlstrom, G. (1997). NMR studies of the E140Q mutant of the carboxy-terminal domain of calmodulin reveal global conformational exchange in the Ca<sup>2+</sup>-saturated state. *Biochemistry*. 36, 3448–3457.
- Fahey, R.C., Newton, G.L., Arrick, B., Overdank-Bogart, T. and Aley, S.B. (1984). *Entamoeba histolytica*, a eukaryote without glutathione metabolism. *Science*. 224, 70-72.
- Fahmy, K., Merroun, M., Pollmann, K., Raff, J., Savchuk, O., Hennig, C. and Selenska-Pobell, S. (2006). Secondary Structure and Pd<sup>(II)</sup> Coordination in S-Layer Proteins from *Bacillus sphaericus* Studied by Infrared and X-Ray Absorption Spectroscopy *Biophys. J.* 91, 996–1007.

- Falke, J.J., Drake, S.K., Hazard, A.L. and Peersen, O.B. (1994). Molecular tuning of ion binding to calcium signalling proteins. *Q. Rev. Biophys.* 27, 219–290.
- Fallon, J.L., Halling, D.B., Hamilton, S.L., Quioco, F.A. (2005). Structure of Calmodulin bound to the hydrophobic IQ domain of the cardiac calcium (v) 1.2 calcium channel. *Structure (Cambridge)*. 13: 1881–1886.
- Ferguson, C., Kern, M. and Audesirk, G. (2000). Nanomolar concentrations of inorganic lead increase  $\text{Ca}^{2+}$  efflux and decrease intracellular free  $\text{Ca}^{2+}$  ion concentrations in cultured rat hippocampal neurons by a calmodulin-dependent mechanism. *Neurotoxicology*. 21, 365–378.
- Finn, B.E., Evenas, J., Drakenberg, T., Waltho, J.P., Thulin, E. and Forsen, S. (1995). Calcium-induced structural changes and domain autonomy in calmodulin. *Nat. Struct. Biol.* 2, 777–783.
- Forsen, S., Thulin, E. and Lilja, H. (1979).  $^{113}\text{Cd}$  NMR in the study of calcium binding proteins, troponin C. *FEBS Lett.* 104, 123–126.
- Fullmer, C.S., Edelstein, S. and Wasserman, R.H. (1985). Lead-binding properties of intestinal calcium-binding proteins. *J. Biol. Chem.* 260, 6816–6819.
- Furuichi, T. and Mikoshiba, K. (1995). Inositol 1,4,5-trisphosphate receptor-mediated  $\text{Ca}^{2+}$  signaling in the brain. *J. Neurochem.* 64, 953–960.
- Galione, A. and White, A. (1994).  $\text{Ca}^{2+}$  release induced by cyclic ADP-ribose. *Trends. Cell. Biol.* 4, 431–436.
- Ganguly, A., and Lohia, A. (2001). The cell cycle of *Entamoeba invadens* during vegetative growth and differentiation. *Mol. Biochem. Parasitol.* 112, 277–285.

- Garza, A., Vega, R. and Soto, E. (2006). Cellular mechanism of lead neurotoxicity. *Med. Sci. Monit.* 12, RA57-65.
- Geli, M., Wasp, A. and Reizman, H. (1998). Distinct functions of calmodulin are required for the uptake step of receptor-mediated endocytosis step of yeast: the type I Myo5p is one of the calmodulin targets. *EMBO J.* 17, 635- 647.
- Gentry, H.R., Singer, A.U., Betts, L., Yang, C., Ferrara, J.D., Sondek, J. and Parise, L.V. (2005). Structural and biochemical characterization of CIB1 delineates a new family of EF-hand-containing proteins. *J. Biol. Chem.* 280, 8407–8415.
- Gerke, V., Creutz, C.E. and Moss, S.E. (2005). Annexins, linking Ca<sup>2+</sup> signalling to membrane dynamics. *Nat. Rev. Mol. Cell. Biol.* 6, 449-461.
- Giannini, G., Conti, A., Mammarella, S., Scrobogna, M. and Sorrentino, V. (1995). The ryanodine receptor/Ca<sup>2+</sup> channel genes are widely and differentially expressed in murine brain and peripheral tissues. *J. Cell Biol.* 128, 893–904.
- Gilchrist, C.A., Holm, C.F., Hughes, M.A., Schaenman, J.M., Mann, B.J., and Petri, W.A., Jr., (2001). Control of Ferredoxin and Gal/GalNAc Lectin Gene Expression in *Entamoeba histolytica* by a cis-Acting DNA Sequence. *J. Biol. Chem.* 276, 11838-11843.
- Gilchrist, C.A., Jeo, M., Line, C.G., Mann, B.J. and Petri, W.A., Jr. (2003). Targets of the *Entamoeba histolytica* Transcription Factor URE3-BP *J.Biol.Chem.* 278, 4646-4653.
- Glusker, J.P. (1991). Structural aspects of metal liganding to functional groups in proteins. *Adv. Protein Chem.* 42, 1–76.
- Godzik, A. and Sander, C. (1989). Conservation of residue interactions in a family of Ca-binding proteins. *Protein Eng.* 2, 589–596.

- Goering, P.L. (1993). Lead-protein interactions as a basis for lead toxicity. *Neurotoxicology*. 14, 45–60.
- Gopal, B., Suma, R., Murthy, M.R., Bhattacharya, A. and Bhattacharya, S. (1998). Crystallization and preliminary X-ray studies of a recombinant calcium-binding protein from *Entamoeba histolytica*. *Acta Crystallogr. D Biol. Crystallogr.* 54, 1442–1445.
- Gopal, B., Swaminathan, C.P., Bhattacharya, S., Bhattacharya, A. Murthy, M.R. and Surolia, A. (1997). Thermodynamics of metal ion binding and denaturation of a calcium binding protein from *Entamoeba histolytica*. *Biochemistry*. 36: 10910–10916.
- Gourinath, S., Padhan, N., Alam, N. and Bhattacharya, A. (2005). Crystallization and preliminary crystallographic analysis of calcium-binding protein-2 from *Entamoeba histolytica* and its complexes with strontium and the IQ1 motif of myosin V. *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* 61, 417–420.
- Grabarek, Z. (2006). Structural basis for diversity of the EF-hand calcium-binding proteins. *J. Mol. Biol.* 359, 509–525.
- Grynopas, M.D., Hamilton, E., Cheung, R., Tsouderos, Y., Deloffre, P., Hott, M. and Marie, P.J. (1996). Strontium increases vertebral bone volume in rats at a low dose that does not induce detectable mineralization defect. *Bone*. 18, 253–9.
- Habermann, E., Crowell, K. and Janicki, P. (1983). Lead and other metals can substitute for  $\text{Ca}^{2+}$  in calmodulin. *Arch. Toxicol.* 54, 61–70.
- Haiech, J., Moulhaye, S.B. and Kilhoffer, M.C. (2004). The EF-handome, combining comparative genomic study using FamDBtool, a new bioinformatics tool and the network of expertise of the European Calcium Society. *Biochim. Biophys. Acta*. 1742, 179–183.

- Han, B.-G., Nunomura, W., Takaku wa, Y., Mohandas, N. and Jap, B.K. (2000). Protein 4.1R core domain structure and insights into regulation of cytoskeletal organization. *Nat. Struct. Biol.* 7, 871-875.
- Hanahan, D. (1983). Studies on Transformation of *Escherichia coli* with Plasmids. *J. Mol. Biol.* 166, 557-580, Academic Press, Inc.
- Harding, M.M. (1999). The geometry of metal-ligand interactions relevant to proteins. *Acta Crystallogr. D, Biol. Crystallogr.* 55, 1432-1443.
- Harding, M.M. (2000). The geometry of metal-ligand interactions relevant to proteins. II. Angles at the metal atom, additional weak metal-donor interactions. *Acta Crystallogr. D, Biol. Crystallogr.* 56, 857-867.
- Harper, J.F. and Harman A. (2005). Plants, symbiosis and parasites, a calcium signaling connection. *Nat. Rev. Mol. Cell. Biol.* 6, 555-566.
- Hennig, H.F. (1986). Metal-binding proteins as metal pollution indicators. *Environ. Health Perspect.* 65, 175-187.
- Hernandez-Ochoa, I., Garcia-Vargas, G., Lopez-Carrillo, L., Rubio-Andrade, M., Moran-Martinez, J., Cebrian, M.E. and Quintanilla-Vega, B. (2005). Low lead environmental exposure alters semen quality and sperm chromatin condensation in northern Mexico. *Reprod. Toxicol.* 20, 221-228.
- Herzberg O, James M.N.G. (1988). Refined crystal structure of troponin C from turkey skeletal muscle at 2.0 Å resolution. *J. Mol. Biol.* 203, 761-779.
- Herzberg, O. and James, M.N. (1988). Refined crystal structure of troponin C from turkey skeletal muscle at 2.0 Å resolution. *J. Mol. Biol.* 203, 761-779.

- Herzberg, O., Moulton, J. and James, M.N. (1986). A model for the  $\text{Ca}^{2+}$ -induced conformational transition of troponin C. A trigger for muscle contraction. *J. Biol. Chem.* 261, 2638–2644.
- Hirata, K.K., Que, X., Melendez-Lopez, S.G., Debnath, A., Myers, S., Herdman, D. S., Orozco, E., Bhattacharya, A., Mckerrow, J. and Reeds, S. L. (2007). A phagocytosis mutant of *Entamoeba histolytica* is less virulent due to deficient proteinase expression and release. *Exp. Parasitol.* 115, 192–199.
- Hoeflich, K.P. and Ikura, M. (2002). Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell.* 108, 739–742.
- Hohenester, E., Maurer, P., Hohenadl, C., Timpl, R., Jansonius, J.N. and Engel, J. (1996). Structure of a novel extracellular  $\text{Ca}^{2+}$ -binding module in BM-40. *Nat. Struct. Biol.* 3, 67–73.
- Holm, R.H., Kennepohl, P., Solomon, E.I. (1996). Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* 96, 2239–2314.
- Holmes, K.C. Popp, D. Gebhard, W. Kabsch, W. (1990). Atomic model of the actin filament. *Nature.* 347, 44–49.
- Houdusse, A. and Cohen, C. (1996). Structure of the regulatory domain of scallop myosin at 2 Å resolution, implications for regulation. *Structure.* 4, 21–32.
- [http://commons.wikimedia.org/wiki/File:Entamoeba\\_histolytica\\_life\\_cycle-en.svg](http://commons.wikimedia.org/wiki/File:Entamoeba_histolytica_life_cycle-en.svg)
- Huber, P.A., El-Mezgueldi, M., Grabarek, Z., Slatter, D. A., Levine, B. A. and Marston, S. B. (1996). Multiple-sited interaction of caldesmon with  $\text{Ca}^{2+}$  - calmodulin. *Biochem. J.* 316, 413–420.

- Ibers, J.A., Holm, R.H. (1980). Modeling coordination sites in metalloproteins. *Science*. 209, 223–235.
- Ikura, M. (1996). Calcium binding and conformational response in EF-hand proteins. *Trends Biochem. Sci.* 21, 14–17.
- Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. and Bax, A. (1992). Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science*. 256, 632-638.
- Ilari, A., Johnson, K. A., Nastopoulos, V., Verzili, D., Zamparelli, C., Colotti, G., Tsernoglou, D. and Chiancone, E. (2002). The crystal structure of the sorcin calcium binding domain provides a model of  $\text{Ca}^{2+}$ -dependent processes in the full-length protein. *J. Mol. Biol.* 317, 447–458.
- Ingersoll, R.J. and Wasserman, R.H. (1971). Vitamin D-induced calcium-binding protein. Binding characteristics, conformational effects and other properties. *J. Biol.Chem.* 246, 2808–14.
- Ingraham, R.H. and Swenson, C.A. (1983). Stability of the  $\text{Ca}^{2+}$ -specific and  $\text{Ca}^{2+}$ - $\text{Mg}^{2+}$  domains of troponin C. Effect of pH. *Eur. J. Biochem.* 132, 85–88.
- Ivings, L., Pennington, S.R., Jenkins, R., Weiss, J.L. and Burgoyne, R.D. (2002). Identification of  $\text{Ca}^{2+}$ -dependent binding partners for the neuronal calcium sensor protein neurocalcin delta: interaction with actin, clathrin and tubulin. *Biochem. J.* 363, 599-608.
- Jain, R., Kumar, S., Gourinath, S., Bhattacharya, S. and Bhattacharya, A. (2009). N- and C-Terminal Domains of the Calcium Binding Protein EhCaBP1 of the Parasite *Entamoeba histolytica* Display Distinct Functions. *PLoS ONE*. 4(4), e5269.

- Jain, R., Santi-Rocca, J., Padhan, N., Bhattacharya, S., Guillen, N. and Bhattacharya, A. (2008). Calcium binding protein-1 of *Entamoeba histolytica* transiently associates with phagocytic cups in a calcium independent manner. *Cell Microbiol.* 10, 1373–1389.
- Jia, J., Han, Q., Borregaard, N., Lollike, K. and Cygler, M. (2000). Crystal structure of human grancalcin, a member of the penta-EF-hand protein family. *J. Mol. Biol.* 300, 1271–1281.
- Jia, J., Tarabykina, S., Hansen, C., Berchtold, M. and Cygler, M. (2001). Structure of apoptosis-linked protein ALG-2, insights into  $\text{Ca}^{2+}$ -induced changes in penta-EF-hand proteins. *Structure.* 9, 267–275.
- Johnson, J.G., Epstein, N., Shiroishi, T., and Miller, L.H. (1981). Identification of surface proteins on viable *Plasmodium knowlesi* merozoites. *J. Protozool.* 28, 160-164.
- Jordan, I.K., Henze, K., Fedorova N.D., Koonin, E.V. and Galperin, M.Y. (2003). Phylogenomic analysis of the *Giardia intestinalis* transcarboxylase reveals multiple instances of domain fusion and fission in the evolution of biotin-dependent enzymes. *J. Mol. Microbiol. Biotechnol.* 5, 172-89.
- Kammanadiminti, S.J. and Chadee, K. (2006). Suppression of NF-kappaB activation by *Entamoeba histolytica* in intestinal epithelial cells is mediated by heat shock protein 27. *J. Biol. Chem.* 281, 26112-20.
- Kantardjieff, K.A. and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein sci.* 12, 1865–1871.

- Katzenstein, D., Richerson, V. and Braude, A. (1982). New concepts of amebic liver abscess derived from hepatic imaging, serodiagnosis and hepatic enzymes in 67 consecutive cases in San Diego. *Medicine. (Baltimore)*. 61, 237-46.
- Kern, M., Wisniewski, M., Cabell, L. and Audesirk, G. (2000). Inorganic lead and calcium interact positively in activation of Calmodulin. *Neurotoxicology*. 21, 353–363.
- Kitaura, Y., Satoh, H., Takahashi, H., Shibata, H. and Maki, M. (2002). Both ALG-2 and peflin, penta-EF-hand (PEF) proteins, are stabilized by dimerization through their fifth EF-hand regions. *Arch. Biochem. Biophys.* 399, 12–18.
- Klee, C. B. (1988). Interaction of calmodulin with  $\text{Ca}^{++}$  and target proteins. In *Calmodulin* (Cohen, P. & Klee, C. B., eds), pp. 35-56, Elsevier, New York.
- Kleywegt, G.J. and Brunger, A.T. (1996). Checking your imagination: applications of the free R value. *Structure*. 4, 897–904.
- Kobayashi, C. and Takada, S. (2006). Protein grabs a ligand by extending anchor residues, molecular simulation for  $\text{Ca}^{2+}$  binding to calmodulin loop. *Biophys. J.* 90, 3043–3051.
- Kojetin, D.J., Venters, R.A., Kordys, D.R., Thompson, R.J., Kumar, R. and Cavanagh, J. (2006). Structure, binding interface and hydrophobic transitions of  $\text{Ca}^{2+}$ -loaded calbindin-D28K. *Nat. Struct. Mol. Biol.* 13, 641–647.
- Kretsinger R.H. (1980). Structure and evolution of calcium-modulated proteins. *CRC Cr. Rev. Biochem. Mol.* 8, 119–174.
- Kretsinger, R.H. (1987). Cold Spring Harbor Symp. *Quant. Biol.* 52, 499-510.

- Kretsinger, R.H. and Nockolds, C.E. (1973). Carp muscle calcium-binding protein, II. Structure determination and general description. *J. Biol. Chem.* 248, 3313–3326.
- Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C.B. and Bax A. (1995). Solution structure of calcium-free calmodulin. *Nat Struct Biol* 2, 768–776.
- Kumar, S., Padhan, N., Alam, N. and Gourinath, S. (2007). Crystal structure of calcium binding protein-1 from *Entamoeba histolytica*: a novel arrangement of EF hand motifs. *Proteins*. 68(4), 990-998.
- Kung, C., Preston, R.R., Maley, M.E., Ling, K.Y., Kanabrocki, J.A., Seavey, B.R. and Saimi, Y. (1992). In vivo *Paramecium* mutants show that calmodulin orchestrates membrane responses to stimuli. *Cell Calcium*. 13, 413–425.
- Kursula, P. and Majava, V. (2007). A structural insight into lead neurotoxicity and calmodulin activation by heavy metals. *Acta. Cryst. F*63, 653-656.
- Lawrence, J.G. and Hendrickson, H. (2003). Lateral gene transfer, when will adolescence end? *Mol. Microbiol.* 50, 739-49.
- Lee, H.C. (1994). Cyclic ADP-ribose – a new member of a super family of signaling cyclic-nucleotides. *Cell. Signall.* 6, 591–600.
- Leippe, M. (1997). Amoebapores. *Parasitol. Today*. 13, 178-83.
- Leon-Avila, G. and Tovar, J. (2004). Mitosomes of *Entamoeba histolytica* are abundant mitochondrion-related remnant organelles that lack a detectable organellar genome. *Microbiology*. 150, 1245-50.

- Levine, B.A. and Williams, R.J.P. (1982) *Calcium-binding* to proteins. In *Calcium and Cell Function* (Cheung, WV, ed.). Vol 11. *Academic Press, New York*, pp. 1-38.
- Li, M.X., Chandra, M., Pearlstone, J.R., Racher, K.I., Trigo- Gonzalez, G., Borgford, T., Kay, C.M. and Smillie, L.B. (1994). Properties of isolated recombinant N and C domains of chicken troponin C. *Biochemistry*. 33, 917–925.
- Lidsky, T.I. and Schneider, J.S. (2003). Lead neurotoxicity in children, basic mechanisms and clinical correlates. *Brain*. 126, 5–19.
- Likic, V.A., Strehler, E.E. and Gooley, P.R. (2003). Dynamics of Ca<sup>2+</sup>- saturated calmodulin D129N mutant studied by multiple molecular dynamics simulations. *Protein Sci*. 12, 2215–2229.
- Linse, S., Thulin, E. and Sellers, P. (1993). Disulfide bonds in homo- and heterodimers of EF-hand subdomains of calbindin D9k , stability, calcium binding and NMR studies. *Protein Sci*. 2, 985–1000.
- Lippard, S.J. and Berg, J.M. (1994). *Principles of Bioinorganic Chemistry*, *University Science Books, Mill Valley, California*.
- Loftus, B., Anderson, I., Davies, R.U., Alsmark, C.M., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., Nozaki, T., SuhB, Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M.A., Rabbinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guille'n, N., Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N.M., Petri, W.A., Jr., Clark, C.G., Embley, T.M., Barrell, B. and Fraser, C.M., Hall, N. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature*. 433, 865–868.

- Lopez, M.M., Chin, D.H., Baldwin, R.L. and Makhatadze, G.I. (2002). The enthalpy of the alanine peptide helix measured by isothermal titration calorimetry using metal-binding to induce helix formation. *Proc. Natl. Acad. Sci. USA.* 99, 1298–1302.
- Lopez-Lopez, J.R., Shacklock, P.S., Balke, C.W. and Wier, W.G. (1995). Local  $\text{Ca}^{2+}$  transients triggered by single L-type  $\text{Ca}^{2+}$  channel currents in cardiac cells. *Science.* 268, 1042–1045.
- Losch, F.A. (1875). Massive development of amebas in the large intestine. *Am. J. Trop. Med. Hyg.* 24, 383-392.
- Makioka, A., Kumagai, M., Ohtoma, H., Kobayashi, S., and Takeuchi, T. (2001). Effect of calcium antagonists, calcium channel blockers and calmodulin inhibitors on the growth and encystation of *Entamoeba histolytica* and *E. invadens*. *Parasitol. Res.* 87, 833-837.
- Maler, L., Blankenship, J., Rance, M. and Chazin, W.J. (2000). Site–site communication in the EF-hand  $\text{Ca}^{2+}$ -binding protein calbindin D9k . *Nat. Struct. Biol.* 7, 245–250.
- Malmendal, A., Evenas, J., Forsen, S. and Akke, M. (1999). Structural dynamics in the C-terminal domain of calmodulin at low calcium levels. *J. Mol. Biol.* 293, 883–899.
- Malmendal, A., Linse, S., Evenas, J., Forsen, S. & Drakenberg, T. (1999). Battle for the EF-hands: magnesium-calcium interference in calmodulin. *Biochemistry.* 38, 11844–11850.
- Mann, B.J. (2002). Structure and function of the *Entamoeba histolytica* Gal/GalNAc lectin. *Int. Rev. Cytol.* 216, 59-80.

- Marchand, S. and Roux, B. (1998). Molecular dynamics study of calbindin D9k in the apo and singly and doubly calcium-loaded states. *Proteins*. 33, 265–284.
- Marchetti, C. (2003). Molecular Targets of Lead in Brain *Neurotoxicity. Neurotox. Res.* 5, 221–236.
- Marinets, A., Zhang, T., Guillén, N., Gounon, P., Bohle, B., Vollmann, V., Scheiner, O., Wiedermann, G., Stanley, Jr., S.L. and Duchêne, M. (1997). Protection against invasive amebiasis by a single monoclonal antibody directed against a lipophosphoglycan antigen localized on the surface of *Entamoeba histolytica*. *J. Exp. Med.* 186, 1557-65.
- Markovac, J. and Goldstein, G.W. (1988). Picomolar concentrations of lead stimulate brain protein kinase C. *Nature*. 334, 71–73.
- Martin, R. (1984). *Bioinorganic Chemistry of Calcium, Marcel Dekker, New York.*
- Martinez-Palomo, A., Gonzalez-Robles, A. and De la Torre, M. (1973b). Selective agglutination of trophozoites of various strains of *E. histolytica* induced by concanavalin. *Arch. Invest. Med. (Mex), Suppl* 1, 39-48.
- Martinez-Palomo, A., Gonzalez-Robles, A. and De la Torre, M. (1973a). Selective agglutination of pathogenic strains of *Entamoeba histolytica* induced con A. *Nat. New. Biol.* 245, 186-187.
- Matsumoto, Y., Perry, G., Scheibel, L.W. and Aikawa, M. (1987). Role of calmodulin in *Plasmodium falciparum*: implications for erythrocyte invasion by the merozoite. *Eur. J. Cell. Biol.* 45, 36-43.
- McPhalen, C.A., Strynadka, N.C. and James, M.N. (1991). Calcium-binding sites in proteins, a structural perspective. *Adv. Protein Chem.* 42, 77–144.

- Meador, W.E., Means, A.R. and Quioco, F.A. (1992). Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. *Science*. 257, 1251-1255.
- Meador, W.E., Means, A.R. and Quioco, F.A. (1993). Modulation of calmodulin plasticity in molecular recognition on the basis of x-ray structures. *Science*. 262, 1718–1721.
- Meza, I. (2000). Signal transduction pathways in *Entamoeba histolytica*. *Parasitol. Today* 16, 23-28.
- Mills, J.S. Johnson, J.D. (1985). Metal ions as allosteric regulators of calmodulin. *J. Biol. Chem.* 260, 15100–15105.
- Mittal, V., Bhattacharya, A. and Bhattacharya, S. (1994). Isolation and characterization of a species-specific multicopy DNA sequence from *Entamoeba histolytica*. *Parasitology*. 108, 237-44.
- Mohan, P.M., Mukherjee, S. and Chary, K.V. (2008). Differential native state ruggedness of the two Ca<sup>2+</sup>-binding domains in a Ca<sup>2+</sup> sensor protein. *Proteins*. 70, 1147–1153.
- Moore, M.R., Goldberg, A. and Yeung-Laiwah, A.A. (1987). Lead effects on the heme biosynthetic pathway. Relationship to toxicity. *Ann. N.Y. Acad. Sci.* 514, 191–203.
- Moreno, S.N. and Docampo, R. (2003). Calcium regulation in protozoan parasites. *Curr. Opin. Microbiol.* 6, 359–364.
- Mornet, D. and Bonet-Kerrache, A. (2001). Neurocalcin-actin interaction. *Biochim. Biophys. Acta Mol. Cell Res.* 1549, 197-203.

- Mornet, D. and Bonet-Kerrache, A. (2001). Neurocalcin-actin interaction. *Biochim. Biophys. Acta Mol. Cell Res.* 1549, 197-203.
- Mukherjee, S., Mohan, P.M., Kuchroo, K. and Chary, K.V. (2007). Energetics of the native energy landscape of a two-domain sensor protein: distinct folding features of the two domains. *Biochemistry.* 46, 9911–9919.
- Munoz, L., Moreno, M.A., Perez-Garcia, J.N., Tovar, G.R. and Hernandez, V.I. (1991). Possible role of calmodulin in the secretion of *Entamoeba histolytica* electron-dense granules containing collagenase. *Mol. Microbiol.* 5, 1707-1714.
- Nagae, M., Nozawa, A., Koizumi, N., Sano, H., Hashimoto, H., Sato, M. and Shimizu, T. (2003). The crystal structure of the novel calcium-binding protein AtCBL2 from *Arabidopsis thaliana*. *J. Biol. Chem.* 278, 42240–42246.
- Nelson, D.L. and Cox, M.M. (2000). *Lehninger Principles of Biochemistry, 3rd ed.* New York, Worth Publishers.
- Nelson, M.R. and Chazin, W.J. (1998). An interaction-based analysis of calcium-induced conformational changes in Ca<sup>2+</sup> sensor proteins. *Protein Sci.* 7, 270–282.
- Nicholson, A.W. (1997). Ribonucleases, Structures and Functions, edited by G. D'Alessio & J. F. Riordan. *San Diego, Academic Press.*
- Nickel, R., Jacobs, T., Urban, B., Scholze, H., Bruhn, H. and Leippe, M. (2000). Two novel calcium-binding proteins from cytoplasmic granules of the protozoan parasite *Entamoeba histolytica*. *FEBS Lett.* 486, 112–116.
- Ohya, Y. and Botstein, D. (1994). Diverse essential functions revealed by complementing yeast calmodulin mutants. *Science.* 263, 963-966.

- Orozco E, Guarneros G, Martinez-Palomo A and Sanchez T. (1983). *Entamoeba histolytica*: Phagocytosis as a virulence factor. *J. Exp. Med.* 158, 1511–1521.
- Osawa, M., Tokumitsu, H., Swindells, M.B., Kurihara, H., Orita, M., Shibamura, T., Furuya, T. and Ikura, M. (1999). A novel target recognition revealed by calmodulin in complex with Ca<sup>2+</sup>-calmodulin-dependent kinase kinase. *Nat. Struct. Biol.* 6, 819-824.
- Ouyang, H. and Vogel, H.J. (1998). Metal ion binding to calmodulin: NMR and fluorescence studies. *Biometals.* 11, 213–222.
- Ozawa, T., Sasaki, K. and Umezawa, Y. (1999). Metal ion selectivity for formation of the calmodulin–metal–target peptide ternary complex studied by surface plasmon resonance spectroscopy. *Biochim. Biophys. Acta.* 1434, 211–220.
- Peters, C. and Mayer, A. (1998). Ca<sup>2+</sup>/calmodulin signals the completion of docking and triggers a late step of vacuole fusion. *Nature.* 396, 575-580.
- Pezzella, N., Bouchot, A., Bonhomme, A., Pingret, L., Klein, C., Burlet, H., Balossier, G., Bonhomme, P. and Pinon, J.M. (1997). Involvement of calcium and calmodulin in *Toxoplasma gondii* tachyzoite invasion. *Eur. J. Cell Biol.* 74, 92-101.
- Pfleger, H. and Wolf, H.U. (1975). Activation of membrane-bound high-affinity calcium ion- sensitive adenosine triphosphatase of human erythrocytes by bivalent metal ions. *Biochem. J.* 147, 359–61.
- Pidcock, E. and Moore, G.R. (2001). Structural characteristics of protein binding sites for calcium and lanthanide ions. *J. Biol. Inorg. Chem.* 6, 479–489.
- Potter, J.D., Strang-Brown, P., Walker, P.L. and Lida, S. (1983). Ca<sup>2+</sup> binding to Calmodulin. *Methods Enzymol.* 102, 135–143.

- Prasad, J., Bhattacharya, S. and Bhattacharya, A. (1992). Cloning and sequence analysis of a calcium-binding protein gene from a pathogenic strain of *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* 52, 137–140.
- Prasad, J., Bhattacharya, S. and Bhattacharya, A. (1993). The calcium binding protein of *Entamoeba histolytica*: expression in *Escherichia coli* and immunochemical characterization. *Cell Mol. Biol. Res.* 39, 167–175.
- Precheur, B., Cox, J.A., Petrova, T., Mispelter, J. and Craescu, C.T. (1996). *Nereis* sarcoplasmic  $\text{Ca}^{2+}$ -binding protein has a highly unstructured apo state which is switched to the native state upon binding of the first  $\text{Ca}^{2+}$  ion. *FEBS Lett.* 395, 89–94.
- Que, X. and Reed, S.L. (2000). Cysteine proteinases and the pathogenesis of amebiasis. *Clin. Microbiol. Rev.* 13, 196-206.
- Qvarnstrom, Y.C., James, M., Xayavong, B.P., Holloway, G.S., Visvesvara, R., Sriramand A. and da Silva. J. (2005). Comparison of real-time PCR protocols for differential laboratory diagnosis of amebiasis. *J. Clin. Microbiol.* 43, 5491-5497.
- Rabah, G., Popescu, R., Cox, J.A., Engelborghs, Y. and Craescu, C.T. (2005). Solution structure and internal dynamics of NSCP, a compact calcium-binding protein. *FEBS J.* 272, 2022–2036.
- Ragland, B.D., Ashley, L.S., Vaux, D.L. and Petri, W.A. (1994). *Entamoeba histolytica*, target cells killed by trophozoites undergo DNA fragmentation which is not blocked by Bcl-2. *Exp. Parasitol.* 79, 460-7.
- Rao, S.T., Wu, S., Satyshur, K.A., Ling, K.Y., Kung, C. and Sundaralingam, M. (1993). Structure of *Paramecium tetraurella* calmodulin at 1.8 angstrom resolution *Protein Sci.* 2, 436-447.

- Ravdin, J. (1988). Amebiasis. Human infection by *Entamoeba histolytica*. New York, NY: John Wiley and Sons.
- Ravdin, J.I., Croft, B.Y. and Guerrant, R.L. (1980). Cytopathogenic mechanisms of *Entamoeba histolytica*. *J. Exp. Med.* 152, 377-90.
- Ravdin, J.I., Moreau, F., Sullivan, J.A., Petri, W.A. Jr., and Mandell, G.L. (1988). Relationship of free intracellular calcium to the cytolytic activity of *Entamoeba histolytica*. *Infect Immun.* 56, 1505-1512.
- Ravdin, J.I., Sperelakis, N. and Guerrant, R.L. (1982). Effect of ion channel inhibitors on the cytopathogenicity of *Entamoeba histolytica*. *J. Infect. Dis.* 146, 335-340.
- Ravdin, J.I., Stanley, P., Murphy, C.F., and Petri, Jr. W.A. (1989). Characterization of cell surface carbohydrate receptors for *Entamoeba histolytica* adherence lectin. *Infect. Immun.* 57, 2179-2186.
- Reeds, S.L., Wessel, D.W. and Davis, CE. (1991). *Entamoeba histolytica* infection and AIDS. *Am. J. Med.* 90, 269-271.
- Reid, R.E. (1990). Synthetic fragments of calmodulin calcium-binding site III. A test of the acid pair hypothesis. *J. Biol. Chem.* 265, 5971-5976.
- Reid, R.E., Garipey, J., Saund, A.K. and Hodges, R.S. (1981). Calcium-induced protein folding. Structure-affinity relationships in synthetic analogs of the helix-loop-helix calcium binding unit. *J. Biol. Chem.* 256, 2742-2751.
- Rhoads, A.R. and Friedberg, F. (1997). Sequence motifs for calmodulin recognition. *Faseb J.* 11, 331-340.

- Richardt, G., Federolf, G. and Habermann, E. (1986). Affinity of heavy metal ions to intracellular  $\text{Ca}^{2+}$ -binding proteins. *Biochem. Pharmacol.* 35, 1331–1335.
- Rousseau, F., Schymkowitz, J.W.H. and Itzhaki, L.S. (2003). The unfolding story of three-dimensional domain swapping. *Structure.* 11, 243–251.
- Sahoo, N., Bhattacharya, S. and Bhattacharya, A. (2003). Blocking the expression of a calcium binding protein of the protozoan parasite *Entamoeba histolytica* by tetracycline regulatable antisense-RNA. *Mol. Biochem. Parasitol.* 126, 281–284.
- Sahoo, N., Labruyere, E., Bhattacharya, S., Sen, P., Guillen, N. and Bhattacharya, A. (2004). Calcium binding protein 1 of the protozoan parasite *Entamoeba histolytica* interacts with actin and is involved in cytoskeleton dynamics. *J. Cell Sci.* 117, 3625–3634.
- Sandhir, R. and Gill, K.D. (1994). Effect of lead on the biological activity of calmodulin in rat brain. *Exp. Mol. Pathol.* 61, 69–75.
- Sano, D., Myojo, K. and Omura, T. (2006a). Heavy metal-binding proteins from metal-stimulated bacteria as a novel adsorbent for metal removal technology. *Water Sci. Technol.* 53, 221–226.
- Sano, D., Myojo, K. and Omura, T. (2006b). Cloning of a Heavy-Metal-Binding Protein Derived from Activated-Sludge Microorganisms. *Appl. Environ. Microbiol.* 72, 6377–6380.
- Santella, L. Carafoli, E. (1997). Calcium signaling in the cell nucleus. *Faseb J.* 11, 1091–1109.
- Schaudinn, 1903 (Emended Walker, 1911) A Redescription of *Entamoeba Histolytica* separating it from *Entamoeba dispar*. *J. Eukaryot. Microbiol.* 40, 340-344.

- Scheibel, L.W. (1992). Role of calcium/calmodulin-mediated processes in protozoa. *Int. Rev. Cytol.* 134, 165–242.
- Schroder, B., Schlumbohm, C., Kaune, R. and Breves, G. (1996). Role of calbindin-D9k in buffering cytosolic free  $\text{Ca}^{2+}$  ions in pig duodenal enterocytes. *J. Physiol.* 492, 715-722.
- Schumacher, M.A., Crum, M. and Miller, M.C. (2004). Crystal structures of apocalmodulin and an apocalmodulin/SK potassium channel gating domain complex. *Structure.* 12, 849–860.
- Schumacher, M.A., Rivard, A.F., Bachinger, H.P. and Adelman, J.P. (2001). Structure of the gating domain of a  $\text{Ca}^{2+}$ -activated  $\text{K}^+$  channel complexed with  $\text{Ca}^{2+}$ /calmodulin. *Nature.* 410, 1120-1124.
- Selvapandiyan, A., Duncan, R., Debrabant, A., Bertholet, S., Sreenivas, G., Negi, N.S., Salotra, P. and Nakhasi, H.L. (2001). Expression of a mutant form of *Leishmania donovani* centrin reduces the growth of the parasite. *J. Biol. Chem.* 276, 43253-43261.
- Shaw, G.S., Findlay, W.A., Semchuk, P.D., Hodges, R.S. and Sykes, B.D. (1992). Specific formation of a heterodimeric two-site calcium-binding domain from synthetic peptides. *J. Am. Chem. Soc.* 114, 6258–6259.
- Shaw, G.S., Golden, L.F., Hodges, R.S. and Sykes, B.D. (1991). Interactions between paired calcium-binding sites in proteins, NMR determination of the stoichiometry of calcium binding to a synthetic troponin-C peptide. *J. Am. Chem. Soc.* 113, 5557–5563.
- Shohat, G., Shani, G., Eisenstein, M. and Kimchi, A. (2002). The DAP-kinase family of proteins: study of a novel group of calcium-regulated death-promoting kinases. *Biochim. Biophys. Acta Mol. Cell Res.* 1600, 45-50.

- Silva, J.R.R.F.D., Williams, R.J.P. (1991). *The Biological Chemistry of the Elements, The Inorganic Chemistry of Life*, Clarendon Press; Oxford University Press, Oxford [England], New York.
- Simons, T.J. (1993). Lead-calcium interactions in cellular lead toxicity. *Neurotoxicology*. 14, 77–85.
- Stanley, S.L., Jr. (2003). Amoebiasis. *Lancet*. 361, 1025-34.
- Stendahl, O., Krause, K.H., Krischer, J., Jerstrom, P., Theler, J.M., Clark, R.A., Carpentier, J.L. and Lew, D.P. (1994). Redistribution of intracellular  $\text{Ca}^{2+}$  stores during phagocytosis in human neutrophils. *Science*. 265, 1439-1441.
- Storoni, L.C., McCoy, A.L. and Read, R.J. (2004). Likelihood-enhanced fast rotation functions. *Acta. Crystallogr. D*60, 432-438.
- Strachan, W.D., Chiodini, P.L., Spice, W.M., Moody, A.H. and Ackers, J.P. (1988). Immunological differentiation of pathogenic and non-pathogenic isolates of *Entamoeba histolytica*. *Lancet*. 1, 561-563.
- Strynadka, N.C., and James, M.N.G. (1989). Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu. Rev. Biochem.* 58, 951-958.
- Sundaralingam, M., Bergstrom, R., Strasburg, G., Rao, S.T., Roychowdhury, P., Greaser, M. And Wang, B.C. (1985). Molecular structure of troponin C from chicken skeletal muscle at 3 Å resolution. *Science*. 227, 945–948.
- Sutton, R.B. and Sprang, S.R. (1998). Structure of the protein kinase C beta phospholipid-binding C2 domain complexed with  $\text{Ca}^{2+}$ . *Structure*. 6, 1395-1405.

- Suzuki, Y., Chao, S.H., Zysk, J.R. and Cheung, W.Y. (1985). Stimulation of calmodulin by cadmium ion. *Arch. Toxicol.* 57, 205–211.
- Tainer, J.A., Roberts, V.A. and Getzoff, E.D. (1992). Protein metal-binding sites. *Curr. Opin. Biotechnol.* 3(4), 378–387.
- Tajima, M., Urabe, I., Yutani, K. and Okada, H. (1976). Role of calcium ions in the thermostability of thermolysin and *Bacillus subtilis* var. *amylosacchariticus* neutral protease. *Eur. J. Biochem.* 64, 243–247.
- Takeda, S., Yamashita, A., Maeda, K. and Maeda, Y. (2003). Structure of the core domain of human cardiac troponin in the Ca<sup>(2+)</sup>-saturated form. *Nature.* 424, 35-41.
- Tannich, E., Horstmann, R.D., Knobloch, J. and Arnold, H.H. (1989). Genomic DNA differences between pathogenic and nonpathogenic *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. USA.* 86, 5118-5122.
- Tarabykina, S., Moller, A.L., Durussel, I., Cox, J. and Berchtold, M.W. (2000). Two forms of the apoptosis-linked protein ALG-2 with different Ca<sup>2+</sup> affinities and target recognition. *J. Biol. Chem.* 275, 10514–10518.
- Taylor, C.W. and Traynor, D. (1995). Ca<sup>2+</sup> and inositol trisphosphate receptors. *J. Membr. Biol.* 145, 109–118.
- Taylor, D.A., Sack, J.S., Maune, J.F., Beckingham, K. and Quijoch, F.A. (1991). Structure of a recombinant calmodulin from *Drosophila melanogaster* refined at 2.2-Å resolution. *J. Biol. Chem.* 266, 21375-21380.
- Theret, I., Baladi, S., Cox, J.A., Sakamoto, H. and Craescu, C.T. (2000). Sequential calcium binding to the regulatory domain of calcium vector protein reveals

- functional asymmetry and a novel mode of structural rearrangement. *Biochemistry*. 39, 7920–7926.
- Thomassen, E., Gielen, G., Schutz, M., Schoehn, G., Abrahams, J.P., Miller, S. and van Raaij, M.J. (2003). The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. *J. Mol. Biol.* 331, 361–373.
- Thompson, J.E. Jr., Forlenza, S. and Verma, R. (1985). Amebic liver abscess, a therapeutic approach. *Rev. Infect. Dis.* 7, 171-79.
- Trump, B.J. and Berezsky, I.K. (1995).  $Ca^{2+}$ -mediated cell injury and cell-death. *Faseb. J.* 9, 219–228.
- Tsalkova, T.N. and Privalov, P.L. (1980). Stability of troponin C. *Biochim. Biophys. Acta.* 624, 196–204.
- Tsein, R.W. and Tsein, R.Y. (1990). Calcium channels, stores and oscillations. *Annu. Rev. Cell Biol.* 6, 715-760.
- Tsutsumi, V., Ramirez-Rosales, A., Lanz-Mendoza, H., Shibayama, M., Chavez, B., Rangel-Lopez, E. and Martinez-Palomo, A. (1992). *Entamoeba histolytica*: erythrophagocytosis, collagenolysis, and liver abscess production as virulence markers. *Trans. R. Soc. Trop. Med. Hyg.* 86, 170-172.
- Van Raaij, M.J., Schoehn, G., Burda, M.R. and Miller, S. (2001). Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J. Mol. Biol.* 314, 1137–1146.
- Vetter, S.W. and Leclerc, E. (2003). Novel aspects of calmodulin target recognition and activation. *Eur. J. Biochem.* 270, 404-414.

- Vezzoli G, Baragetti I, Zerbi S, Caumo A, Soldati L, Bellinzoni P, Centemero, A., Rubinacci, A., Moro, G.L., Bianchi, G. (1998). Strontium absorption and excretion in normocalciuric subjects, relation to calcium metabolism. *Clin. Chem.* 44, 586–90.
- Vijay-Kumar, S. and Cook, W.J. (1992). Structure of a sarcoplasmic calcium-binding protein from *Nereis diversicolor* refined at 2.0 Å resolution. *J. Mol. Biol.* 224, 413–426.
- Vinogradova, M.V., Stone, D.B., Malanina, G.G., Karatzaferi, C. and Cooke. R., Mendelson, R.A. and Fletterick, R.J. (2005). Ca ( $^{2+}$ ) - regulated structural changes in troponin. *Proc. Natl. Acad. Sci. USA.* 102, 5038–5043.
- Voigt, H. and Guillén, N. (1999). New insights into the role of the cytoskeleton in phagocytosis of *Entamoeba histolytica*. *Cell Microbiol.* 1, 195-203.
- Voigt, H., Olivo, J.C., Sansonetti, P. and Guillén, N. (1999). Myosin IB from *Entamoeba histolytica* is involved in phagocytosis of human erythrocytes. *J. Cell Sci.* 112, 1191-1201.
- Vyas, M.N., Jacobson, B.L. and Quioco, F.A. (1989). The calcium-binding site in the galactose chemoreceptor protein. Crystallographic and metal-binding studies *J. Biol. Chem.* 264, 20817-20821.
- Wang, C.L., Leavis, P.C. and Gergely, J. (1984). Kinetic studies show that Ca $^{2+}$  and Tb $^{3+}$  have different binding preferences toward the four Ca $^{2+}$ -binding sites of calmodulin. *Biochemistry.* 23, 6410–6415.
- Wasserman, R.H. (1997). Vitamin D and the intestinal absorption of calcium and phosphorus. In, Feldman D, Glorieux, F.H., Pike, J.W. eds. Vitamin D. *New York, Academic Press.* Chap. 16.

- Weaver, J., Porasuphatana, S., Tsai, P., Cao, G.L., Budzichowski, T.A., Roman, L.J. and Rosen, G.M. (2002). The Activation of Neuronal Nitric-Oxide Synthase by Various Divalent Cations. *J. Pharmacol. Exp. Ther.* 302, 781–786.
- Weikel, C.S., Murphy, C.F., Orozco, E., and Ravdin, J.I. (1988). Phorbol esters specifically enhance the cytolytic activity of *Entamoeba histolytica*. *Infect. Immun.* 5, 1485-1491.
- Weis, W.I., Taylor, M.E. and Drickamer, K. (1998). The C-type lectin superfamily in the immune system. *Immunol. Rev.* 163, 19-34.
- WHO/PAHO/UNESCO (WHO, Geneva). (1997). Report of a consultation of experts on Amoebiasis. *Weekly epidemiologic record of the WHO.* 72, 97-99.
- WHO/PAHO/UNESCO report (WHO, Mexico). (1997). A consultation with experts on Amoebiasis. *Epidemiol. Bull.* 18, 13-14.
- Willhoeft, U. and Tannich, E. (1999). The electrophoretic karyotype of *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* 99, 41-53.
- Wilson, M.A. and Brunger, A.T. (2000). The 1.0 Å crystal structure of Ca (<sup>2+</sup>) - bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. *J. Mol. Biol.* 301, 1237–1256.
- Wilson, M.A. and Brunger, A.T. (2003). Domain flexibility in the 1.75 Å resolution structure of Pb<sup>2+</sup>-calmodulin. *Acta Crystallogr. D Biol. Crystallogr.* 59, 1782-1792.
- Wimberly, B., Thulin, E. and Chazin, W.J. (1995). Characterization of the N-terminal half-saturated state of calbindin D9k , NMR studies of the N56A mutant. *Protein Sci.* 4, 1045–1055.

- Yadava, N., Chandok, M.R., Prasad, J., Bhattacharya, S., Sopory, S.K. and Bhattacharya, A. (1997). Characterization of EhCaBP, a calcium-binding protein of *Entamoeba histolytica* and its binding proteins. *Mol. Biochem. Parasitol.* 84, 69–82.
- Yamashita, M.M., Wesson, L., Eisenman, G. and Eisenberg, D. (1990). Where metal ions bind in proteins. *Proc. Natl. Acad. Sci. USA.* 87, 5648–5652.
- Yamazaki, J., Urushidani, T. and Nagao, T. (1996). Barium activates rat cerebellar nitric oxide synthase. *Jpn. J. Pharmacol.* 70, 351–354.
- Yamniuk, A.P. and Vogel, H.J. (2004). Calmodulin's flexibility allows for promiscuity in its interactions with target proteins and peptides. *Mol. Biotechnol.* 27, 33–57.
- Yang, W., Wilkins, A.L., Ye, Y., Liu, Z.R., Li, S.Y., Urbauer, J.L., Hellinga, H.W., Kearney, A., van der Merwe, P.A. and Yang, J.J. (2005). Design of a calcium-binding protein with desired structure in a cell adhesion molecule. *J. Am. Chem. Soc.* 127, 2085–2093.
- Yap, K.L., Ames, J.B., Swindells, M.B. and Ikura, M. (1999). Diversity of conformational states and changes within the EF-hand protein superfamily. *Proteins.* 37, 499–507.
- Yap, K.L. (1995). *Interhlx. University of Toronto, Toronto.*
- Yap, K.L., Kim, J., Truong, K., Sherman, M., Yuan, T. and Ikura, M. (2000). Calmodulin target database. *J. Struct. Funct. Genomics.* 1, 8–14.

Ye, J., Kandegedara, A., Martin, P. and Rosen, B.P. (2005). Crystal Structure of the *Staphylococcus aureus* pI258 CadC Cd(II)/Pb(II)/Zn(II)-Responsive Repressor. *J. Bacteriol.* 187, 4214–4221.

Zhang, M., Tanaka, T. and Ikura, M. (1995). Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nat. Struct. Biol.* 2, 758–767.

# APPENDIX

---

## LIST OF TABLES

<b>1.1</b>	<i>Calcium binding proteins of Entamoeba histolytica</i> .....	29
<b>2.1</b>	<i>Crystallographic Data-Statistics and refinement statistics of EhCaBP1</i> .....	44
<b>2.2</b>	<i>Crystallographic Data-Statistics of EhCaBP1-Phe complex</i> .....	47
<b>2.3</b>	<i>Crystallographic Data-Statistics of N-terminal domain of EhCaBP1</i> .....	49
<b>2.4</b>	<i>Crystallographic Data-Statistics of Pb<sup>2+</sup> - EhCaBP1 complex</i> .....	51
<b>2.5</b>	<i>Crystallographic Data-Statistics of Ba<sup>2+</sup> - EhCaBP1 complex</i> .....	52
<b>2.6</b>	<i>Crystallographic Data-Statistics of Sr<sup>2+</sup> - EhCaBP1 complex</i> .....	54
<b>3.1</b>	<i>Comparison of the EhCaBP1 crystal structure with N-terminal domain of NMR structure, CaM, and TnC</i> .....	74
<b>4.1</b>	<i>Interacting distances of the bound Phenylalanine in the hydrophobic pocket of the assembled-domain with the surrounding residues in EhCaBP1-Phe complex</i> .....	85
<b>5.1</b>	<i>Comparison of the anomalous signals and phasing power of barium and lead complexes for automated structure solution using Phenix</i> .....	108

## LIST OF FIGURES

<b>1.1</b>	<i>Life cycle of Entamoeba histolytica</i> .....	4
<b>1.2</b>	<i>The four units of the Ca<sup>2+</sup> signaling network</i> .....	11
<b>1.3</b>	<i>Elements of the Ca<sup>2+</sup> signaling toolkit</i> .....	12
<b>1.4</b>	<i>A single EF-hand motif from N-terminal domain of calmodulin</i> .....	14
<b>1.5</b>	<i>The canonical EF-hand loop</i> .....	18
<b>1.6</b>	<i>The canonical and Non-canonical EF-loops</i> .....	20
<b>1.7</b>	<i>Ribbon presentations of CaM and CaM in complex with target peptides</i> .....	31
<b>3.1</b>	<i>Over-expression and purification of full-length EhCaBP1</i> .....	63
<b>3.2</b>	<i>Hexagonal crystals of EhCaBP1</i> .....	63
<b>3.3</b>	<i>Crystal structure of Calcium binding protein 1 from E. histolytica (EhCaBP1)</i> ....	65
<b>3.4</b>	<i>Sequence homology of N-terminal half of EhCaBP1</i> .....	66
<b>3.5</b>	<i>Calcium-coordination in EF-hand motifs of EhCaBP1</i> .....	68
<b>3.6</b>	<i>Mass Spectroscopy of EhCaBP1 crystal</i> .....	69
<b>3.7</b>	<i>Symmetry interactions and domain formation</i> .....	70

<b>3.8</b>	<i>Size exclusion chromatography of EhCaBP1.....</i>	<b>71</b>
<b>3.9</b>	<i>Dynamic Light Scattering experiment of EhCaBP1 in various conditions.....</i>	<b>72</b>
<b>3.10</b>	<i>A superposition of CaM on trimer of EhCaBP1.....</i>	<b>75</b>
<b>3.11</b>	<i>Model of EhCaBP1 crystal structure.....</i>	<b>77</b>
<b>4.1</b>	<i>Structural characterization of EhCaBP1 bound to Phenylalanine molecule.....</i>	<b>83</b>
<b>5.1</b>	<i>Schematic representation of independent domains of EhCaBP1.....</i>	<b>88</b>
<b>5.2</b>	<i>Over-expression and purification of N-terminal domain of EhCaBP1 .....</i>	<b>90</b>
<b>5.3</b>	<i>Mass-spectroscopy of N-terminal domain of EhCaBP1.....</i>	<b>91</b>
<b>5.4</b>	<i>Crystals pictures of N-terminal domain in different crystallization conditions....</i>	<b>92</b>
<b>5.5</b>	<i>Crystal structure of N-terminal domain of EhCaBP1.....</i>	<b>93</b>
<b>5.6</b>	<i>Ca<sup>2+</sup>-coordination of EF-hand motifs of N-terminal domain.....</i>	<b>94</b>
<b>5.7</b>	<i>Symmetry-related trimer of N-terminal domain.....</i>	<b>95</b>
<b>6.1</b>	<i>Schematic representation of the coordinating groups for calcium and lead.....</i>	<b>101</b>
<b>6.2</b>	<i>Hexagonal crystal of Pb<sup>2+</sup> -EhCaBP1 complex.....</i>	<b>104</b>
<b>6.3</b>	<i>Lead co-ordination in EF-hand 1 and EF-hand 2 motifs of EhCaBP1.....</i>	<b>105</b>

<b>6.4</b>	<i>Hexagonal crystal of Ba<sup>2+</sup>-EhCaBP1 complex surrounded with oiling.....</i>	<b>106</b>
<b>6.5</b>	<i>Barium coordination in EF-hand 1 and EF-hand 2 motifs of EhCaBP1.....</i>	<b>107</b>
<b>6.6</b>	<i>Hexagonal crystal of Sr<sup>2+</sup>-EhCaBP1 complex.....</i>	<b>109</b>
<b>6.7</b>	<i>Strontium co-ordination in EF-hand 1 and EF-hand 2 motifs of EhCaBP1.....</i>	<b>110</b>

**LIST OF PROTEIN SEQUENCES****Protein sequence of full-length EhCaBP1**

MAEALFKEIDVNGDGA VSYEEVKAFVSKKRAIKNEQLLQLIFKSIDADGNGEI  
DQNEFAK FYGSIQGQDLSDDKIGLKVLYKLMDVDGDGKLTKEEVTSFFKKH  
GIEKVAEQVMKADANGDGYITLEEFLEFSL

**Protein sequence of N-terminal domain of EhCaBP1**

MAEALFKEIDVNGDGA VSYEEVKAFVSKKRAIKNEQLLQLIFKSIDADGNGEI  
DQNEFAK FYGSIQ

**Protein sequence of C-terminal domain of EhCaBP1**

GQDLSDDKIGLKVLYKLMDVDGDGKLTKEEVTSFFKKHGIEKVAEQVMKAD  
ANGDGYITLEEFLEFSL

**LIST OF PUBLICATIONS**

**Shivesh kumar**, Narendra Padhan, Neelima Alam, S.Gourinath. (2007) Crystal structure of calcium binding protein-1 from *Entamoeba histolytica*: a novel arrangement of EF-hand motifs. *Proteins: Structure, Function and Bioinformatics*. 68(4):990-8.

Ruchi Jain, **Shivesh Kumar**, S. Gourinath, Sudha Bhattacharya, Alok Bhattacharya. (2009) N- and C-terminal domains of the calcium binding protein EhCaBP1 of the parasite *Entamoeba histolytica* display distinct functions. *PLoS ONE*. 4(4):e5269.

Krishna Chinthlapuri, Manish Kumar, **Shivesh Kumar**, Shefali Jain, Neelima Alam, S. Gourinath. (2008) Crystal structure of native O-acetyl-serine sulfhydrylase from *Entamoeba histolytica* in complex with cysteine: structural evidence for cysteine binding and lack of interactions with Serine acetyl transferase. *Proteins: Structure, Function and Bioinformatics*. 72(4):1222-32.

**Shivesh Kumar** and S.Gourinath. Specificity and Flexibility of calcium binding loops Calcium Binding Protein-1 from *Entamoeba histolytica*. (Communicated)

**Shivesh Kumar** and S.Gourinath. Crystal structure and biophysical characterization of N-terminal domain of EhCaBP1 from *E. histolytica*. (Manuscript under preparation)

