

**STRUCTURAL STUDIES OF CALCIUM BINDING
PROTEINS AND CYSTEINE SYNTHASE**



**Thesis submitted to
Jawaharlal Nehru University
For the award of
DOCTOR OF PHILOSOPHY
By
MANISH KUMAR**



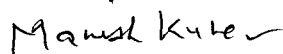
**School of Life Sciences
Jawaharlal Nehru University
New Delhi-110067
INDIA
2009**

**SCHOOL OF LIFE SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI- 110 067
INDIA**

CERTIFICATE

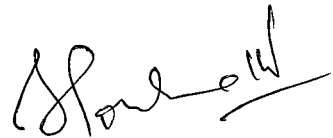
The research work embodied in the thesis entitled "***Structural Studies of Calcium Binding Proteins and Cysteine Synthase***" has been carried out in School of Life Sciences, Jawaharlal Nehru University, New Delhi.

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.



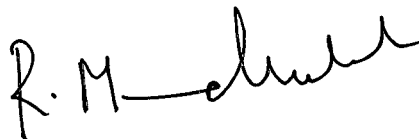
MANISH KUMAR

(Candidate)



Dr. SAMUDRALA GOURINATH

(Supervisor)



Prof. R. Madhubala

(Dean)

Acknowledgement

I would like to start thanking my mentor and supervisor Dr. S. Gourinath, for his support, patience and wise supervision during the preparation of this thesis. I owe him my profound admiration and respect. His help in designing the experiment is highly appreciated. He helped me with his knowledge to unravel some of the complicated theories of Biophysics. I also would like to thank him from my bottom of heart for extending a helping hand at any point of time whenever required. The immense amount of help, you provided to me during my PhD can't be described in this small piece of paper.

I am grateful to Dr. Neelima Alam for his guidance, intellectual support and encouragement during the initial stages of this work. Her consistent analysis of the results always gave rise to new ideas and prompted me to design new experiments.

I am thankful to Prof. Alok Bhattacharya for giving clone of EhCaBP2 and letting me work on this in the initial years of my PhD.

I am thankful to present Dean Prof. Madhubala and former Dean Prof. P. K. Yadava for making all the experimental facilities available to me.

I highly appreciate the support provided by my lab members Shivesh, Tara, Arif, Sudhir, Nitesh, Isha, Faisal, Gokula, Sanjeev and Mohit. Numerous help provided by my lab members made me accomplish my PhD.

I also acknowledge support provided by Karan, Sharma j and Ved..

A special thanks goes to Krishna and Pankaj for their innumerable help they provided me during my PhD.

I am also thankful to all my friends, without their support it would not have been possible to complete my PhD.

I am also thankful to SLS office and CIF staff for their help. I acknowledge the financial support from CSIR.

My everlasting gratitude to my family members for unrelenting support they provided during the entire tenure of my PhD.

Manish

Dedicated to Almighty

Contents

1 An Overview of <i>Entamoeba histolytica</i> and Calcium Binding Proteins	1
1.1 <i>Entamoeba</i> : General features	1
1.1.1 Historical Background	2
1.1.2 Epidemiology	2
1.1.3 Life cycle of the parasite	3
1.1.4 Pathogenesis	7
1.1.5 Genome of <i>E. histolytica</i>	7
1.1.6 Symptoms	8
1.1.7 Treatment	8
1.2 Calcium and its importance	9
1.2.1 Calcium Signaling	9
1.2.2 Calcium as a biological messenger	9
1.2.3 Uniqueness of Ca ²⁺	10
1.2.4 Calcium homeostasis	10
1.3 Extracellular Calcium Binding Proteins	13
1.3.1 Intracellular Calcium Binding Proteins	13
1.3.2 Calcium binding motifs	14
1.4 The EF-hand: A Ca²⁺-Binding Unit	16
1.4.1 The Canonical EF-loop	17
1.4.2 The Non-Canonical EF-loop	18
1.4.3 Calmodulin: a prototypical calcium sensor	19

1.5 Calcium in Protozoan Biology	21
1.5.1 Role of Calcium in <i>Entamoeba</i>	22
1.5.2 EhCaBP1 and EhCaBP2 : Calcium binding proteins in <i>Entamoeba Histolytica</i>	22
2 Expression, Purification, Crystallization and Preliminary Crystallographic Analysis of EhCaBP2 and its Complexes	23
2.1 Introduction	23
2.2 Materials and Methods	25
2.2.1 Sources of Materials	25
2.2.2 Organisms and Growth Conditions	25
2.2.3 Culture Media	25
2.2.3.1 Luria Broth (LB)	25
2.2.3.2 LB Agar	26
2.2.4 Plasmid DNA Isolation from <i>E. coli</i> Transformants	26
2.2.4.1 Preparation of competent cells and transformation	26
2.2.4.2 Mini-preparation of plasmid DNA	26
2.2.4.3 Agarose gel electrophoresis	27
2.2.5 SDS-Polyacrylamide gel electrophoresis	27
2.2.6 Protein estimation	28
2.2.7 Expression and purification of recombinant EhCaBP2 in <i>Escherichia coli</i>	28
2.2.8 Preparation of Sr–EhCaBP2 complex	29

2.2.9	Preparation of IQ1–EhCaBP2 complex	29
2.2.10	Crystallization	30
2.2.11	Data collection and processing	30
2.3	Results and Discussion	31
3	Homology Modeling of Seven EhCaBPs	36
3.1	Introduction	36
3.2	Methods	38
3.2.1	Homology modeling	38
3.2.2	Model validation	39
3.2.3	Phylogenetic analysis	39
3.3	Results and Discussion	40
3.4.1	CaBP family (<i>E. histolytica</i>) and Sequence Conservation	59
3.4.2	Structural description of Seven EhCaBPs models	60
3.4.3	EF hands	62
3.4.4	Validation of EhCaBP protein models	63
3.4.5	Electrostatic Representation	65
3.4.6	Amino Acid Composition And Solvent Accessibilty	66
3.4.7	Comparison of models with their respective templates	67
3.4.8	Alternative model for EhCaBP7 and EhCaBP9	70
3.4.9	Phylogenetic Studies	72
3.4.10	Conclusion	72

4 Calcium Binding Constants Database of Calcium Binding	
Proteins	74
4.1 Description about the database	74
4.2 Results and Discussion	76
4.2.1 Determinants of Ca ²⁺ affinity	77
4.2.2 Determinants of intrinsic Ca ²⁺ affinity	77
4.2.3 Entropic contributions to affinity	78
5 An Overview of OASS	80
5.1 General description of OASS	80
5.1.1 <i>Entamoeba histolytica</i> OASS	81
5.1.2 <i>T. vaginalis</i> OASS	84
5.1.3 Bacterial OASS	84
5.1.4 Plant OASS	85
5.2 Regulation of OASS	87
6 Expression, Purification, Crystallization & Structural Analysis of EhOASS and in Complex with Cysteine	90
6.1 Introduction	90
6.2 Materials and Methods	92
6.2.1 Overexpression and purification	92
6.2.2 Crystallization	93

6.2.3	Co-Crystallization of EhOASS with Cysteine	93
6.2.4	Collection and processing of diffraction data (Eh OASS)	93
6.2.5	Data collection and Processing Cys-EhOASS	94
6.2.6	Structure determination and refinement of EhOASS	95
6.2.7	Structure determination and refinement of EhOASS in complex with Cysteine	95
6.3	Results and Discussion	97
6.5.1	Overall structure	104
6.5.2	Comparison with other structures	104
6.5.3	Dimer interface and Extended N-terminal region	106
6.5.4	C-terminal region	107
6.5.5	Active site	108
6.5.6	SAT binding site	109
6.5.7	Chloride binding site	110
6.5.8	EhOASS in complex with Cysteine	110
6.5.9	Conclusion	112
7	Spectral and Kinetic Studies of EhOASS	113
7.1	Introduction	113
7.1.1	Cysteine Biosynthesis	113
7.1.2	Active site changes in OASS during cysteine biosynthesis	113
7.1.3	Kinetic mechanism of OASS	114
7.1.4	OASS inhibition	114

7.2 Materials and Methods	115
7.2.1 Source of materials	115
7.2.2 Overexpression and purification of EhOASS	115
7.2.3 Ligand Binding and Spectrum Studies	115
7.2.4 Kinetics and Ammonium Sulphate inhibition studies	116
7.2.5 Initial Velocity Studies	117
7.3 Results and Discussion	118
7.4.1 Kinetics mechanism of EhOASS	124
7.4.1.1 Initial Velocity Patterns	126
7.4.2 Spectrum studies of EhOASS with different ligands	126
7.4.3 Kinetics of Eh OASS with Ammonium sulphate as an Inhibitor	128
7.4.4 Kinetics of Cysteine with EhOASS	131
7.4.5 Conclusion	131
References	132
Appendices	147

ABBREVIATIONS AND SYMBOLS

γ	gamma
λ	lambda
μ	micro
$^{\circ}\text{C}$	Degree centigrade
Abs.	Absorbance
APS	Ammonium per Sulfate
β Me	β Mercaptaethanol
CaBP	Calcium Binding Protein
Cam	Calmodulin
CCP4	Collaborative Computational Project 4
CNS	Crystallography and NMR system
Conc.	Concentration
Cys	Cysteine
CS	Cysteine Synthase
DEAE	Di Ethyl Amino Ethyl
DTNB	5, 5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol
EDTA	Ethylene diamine tetraacetate
EGTA	Ethylene Glycol Tetraacetic Acid
EhCaBP	<i>Entamoeba histolytica</i> Calcium Binding Protein
EhCS	<i>Entamoeba histolytica</i> Cysteine Synthase
EhOASS	<i>Entamoeba histolytica</i> O-acetylserine
kDa	Kilo Dalton
LB	Luria broth
M	Marker
Meth	Methionine
MPD	Methyl Pentanediol
mM	Milli molar

Mg	Magnesium
mg	Milligram
min	Minute
mM	Milli molar
ml	Milli liter
NaOH	Sodium hydroxide
Nine EhCaBPs	EhCaBP1, EhCaBP2, EhCaBP3, EhCaBP4, EhCaBP5, EhCaBP6, EhCaBP7, EhCaBP9, EhCaBP10
nm	Nanometer
OAS	O- acetyl- L- Serine
OD.	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
PDB	Protein Data Bank
PEG	Polyethylene Glycol
PMSF	Phenyl Methyl sulphonyl flouride
PLP	Pyridoxal Phosphate
rpm	Revolutions per minute
SAT	Serine O-acetyltransferase
SDS	Sodium Dodecyl Sulfate
Sec	Second
Seven EhCaBPs	EhCaBP3, EhCaBP4, EhCaBP5, EhCaBP6, EhCaBP7, EhCaBP9, EhCaBP10
T₁₀E₁	Tris-EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
TNB	2-nitro-5-thiobenzoic acid
Tris	Tris (hydroxymethyl) amino ethane
U	Unit
v/v	volume/volume
W	Watt
w/v	weight/volume

CHAPTER 1

An Overview of *Entamoeba histolytica* and Calcium Binding Proteins

1.1 *Entamoeba*: General features

Entamoeba is an unicellular protozoan organism which exists either as internal parasites or commensals of animals. In nature there are several species such as *Entamoeba histolytica*, *Entamoeba dispar*, *Entamoeba coli*, *Entamoeba gingivalis*, *Entamoeba invadens*, *Entamoeba moshkovskii*. Among all these species *E. histolytica* is the pathogen which is attributed for amoebiasis (this involves amoebic dysentery and amoebic liver abscesses) *E. histolytica* also causes diseases in animals like cattle etc. *E. coli* and *E. dispar* are nonparasitic form of *Entamoeba* genus. These are commensal species that does not cause disease and reside in the gut of humans and other animals along with other gut floras. One species of *Entamoeba* genus which infects the oral cavity of humans and animals belongs to the species *gingavilis*. The main food for the *Entamoeba gingivalis* is oral cavity fauna. The protozoan *Entamoeba* is not only limited to humans but also affects variety of other animals such as *Entamoeba invadens* which resides in reptiles like snake and lizards. Some of the species of *Entamoeba* live as an independent organism, they are non parasitic in nature. This independent group belongs to the *E. moshkovskii* and it is aquatic in nature. *E. terrapinae* is a commensal species found in turtles.

Taxonomic Classification

Kingdom:	Protista
Subkingdom:	Protozoa
Phylum:	Sarcomastigophora
Sub phylum:	Sarcodina
Class:	Lobosea
Order:	Amoebida
Family:	Entamoebidae
Genus:	<i>Entamoeba</i>
Species:	<i>histolytica</i>

1.1.1 Historical Background

The history of *E. histolytica* is traced to 1875 when Feder Losch observed it in a patient suffering from diarrhoea (Losch, F, A 1875). The very first name attributed by Feder Losch to this organism was “*Amoeba coli*”. In 1901 William Councilman further investigated this organism and suggested the term “amoebic dysentery” and “amebic abscess of the liver” to explain the pathology of this organism. In 1903 Fritz Schaudinn gave the taxonomic classification of the amoeba and named it *Entamoeba histolytica* (Schaudinn F, 1903). In the later years that are in 1919, Clifford had insight in the number of the nuclei in the cyst and grouped this organism as per the nuclei in the cyst. In the context of human beings in 1925, Emile Brumpt said that *Entamoeba histolytica* causes virulence in humans where as *Entamoeba dispar* is non infective (Brumpt, 1925). However, not much credence was attached to the above mentioned theory on the basis of the fact that the two species has very similar morphology. The first proof in favour of Brumpt hypothesis came in 1973 when difference in the agglutination properties of the parasite was observed in the patients suffering from amoebic dysentery and asymptomatic individuals (Martinez-Palomo et al., 1973a; Martinez-Palomo et al., 1973b). After this a lot of experimental evidences kept in coming and tried to differentiate on the basis of antigenic differences (Strachan et al., 1988) and genetic markers (Tannich et al., 1989).

Finally in 1993, *E. histolytica* was finally resolved from *E. dispar* (Diamond and Clark, 1993), validating Brumpt's hypothesis after almost seventy years.

1.1.2 Epidemiology

The name *E. histolytica* encompasses the mode of action of *Entamoeba histolytica* (histo-lytic = tissue destroying), and as per its name it does pathogenesis by destroying the intestinal linings and liver. And this leads to amoebic dysentery or amoebic liver abscess. Amoebiasis is an endemic disease, more or less confined to developing countries, inflicts a high level of morbidity and mortality worldwide. After *Plasmodium falciparum*, *E. histolytica* is the second biggest cause of death in a year as per the WHO estimate (WHO, 1998). Humans and other primates are the only known reservoirs (PHAC, 2001). Amoebiasis is most rampant in the developing countries like India because of poor health hygiene and sanitary conditions. Dissemination of amoebic spores happen mainly through contaminated water and food. In ~90% cases of infection there is no manifestation of symptoms. Only 10% of infection show the classical symptoms of amoebiasis and in these cases *E. histolytica* develop invasive disease (Gathiram and Jackson, 1987). Generally the parasite harbouring place is the intestinal lumen but when it acquires the invasive form, it bores hole in the gut and reaches the blood stream from where it goes to the vital organs, mainly the liver and sometimes the brain. This causes abscesses which proved to be detrimental if not treated well. Males are much more frequented with dreaded form of amoebiasis that is invasive amoebiasis than females. However, amoebic liver abscess is equally common in both sexes among prepubertal children. It is also documented that oral-anal sexual contact can be also one of the mode of propagation of *E. histolytica* infection (Hung et al., 2008; Keystone et al., 1980).

1.1.3 Life cycle of the parasite

The life cycle of *E. histolytica* has two distinct stages: infective and non-infective. Non-infective stage is marked by motility, activeness, fragileness and this stage is known as trophozoite [fig. 1.6.1 (d)]. Infective stage is marked by dormant, hardy, non motile and this stage is known as cyst [fig. 1.6.1 (a)]. Human infection begins with the ingestion of

faecally contaminated food or water. The trophozoites are very active, represent diverse form of morphology and ranges from 10–50 μm in diameter and most of the trophozoites have a nucleus with a central karyosome. The cysts in comparison to the trophozoite is very small and measure 10–15 μm in diameter and are surrounded by a refractile wall made of chitin (Clark, 2000). All classical organelles of a typical eukaryotic cell like mitochondria, Golgi apparatus, rough endoplasmic reticulum, centrioles and microtubules are absent. Nucleus as usual is spherical in shape, 4-7 μm in diameter and has an outer covering called nuclear envelope. Inside the nuclear membrane aggregated chromatin is quite evident and evenly distributed (Clark, 2000). In humans the first stage of infection is oral cavity which happens through the ingestion of cyst [fig. 1.6.1 (b)]. Once the cyst enters the mouth, it traverses down and reaches the gastro-intestinal tract and then into the ileo-cecal region, where it leaves the cyst stage to produce eight trophozoites per cyst (Katz, M et al., 1989). During excystation, there is somehow movement of the amoeba within the cyst wall [fig. 1.6.1 (c)]. During this whole exercise there is depletion in the thickness of the cyst wall and then through any point on the cyst wall amoeba comes out with the help of pseudopodia. The multi-nucleated trophozoite undergoes cytokinesis and nuclear division to give rise to eight daughter amoebae (Marinets et al., 1997). After the birth of the daughter amoeba the trophozoites fix themselves to intestinal epithelial cells and colonic mucus layer. In general most of the individuals (~90%) who become infected with *E. histolytica* does not show any symptom (Gathiram and Jackson, 1987). The next phase of *Entamoeba* trophozoite starts with re-encystation of the trophozoites within the lumen of the colon [fig. 1.6.1 (f)]. After the formation of cyst in the gut lumen it passes along faeces and there is further continuation of the life cycle. There is also a chance that trophozoites reinvade the colonic epithelium, and cause further degeneration of intestinal lumen. In contrast to the cysts that are much tougher and remain viable in a moistened environment and capable enough to spread infection, trophozoites which are very fragile and short-lived outside the body, cannot sustain their journey through the human gut.

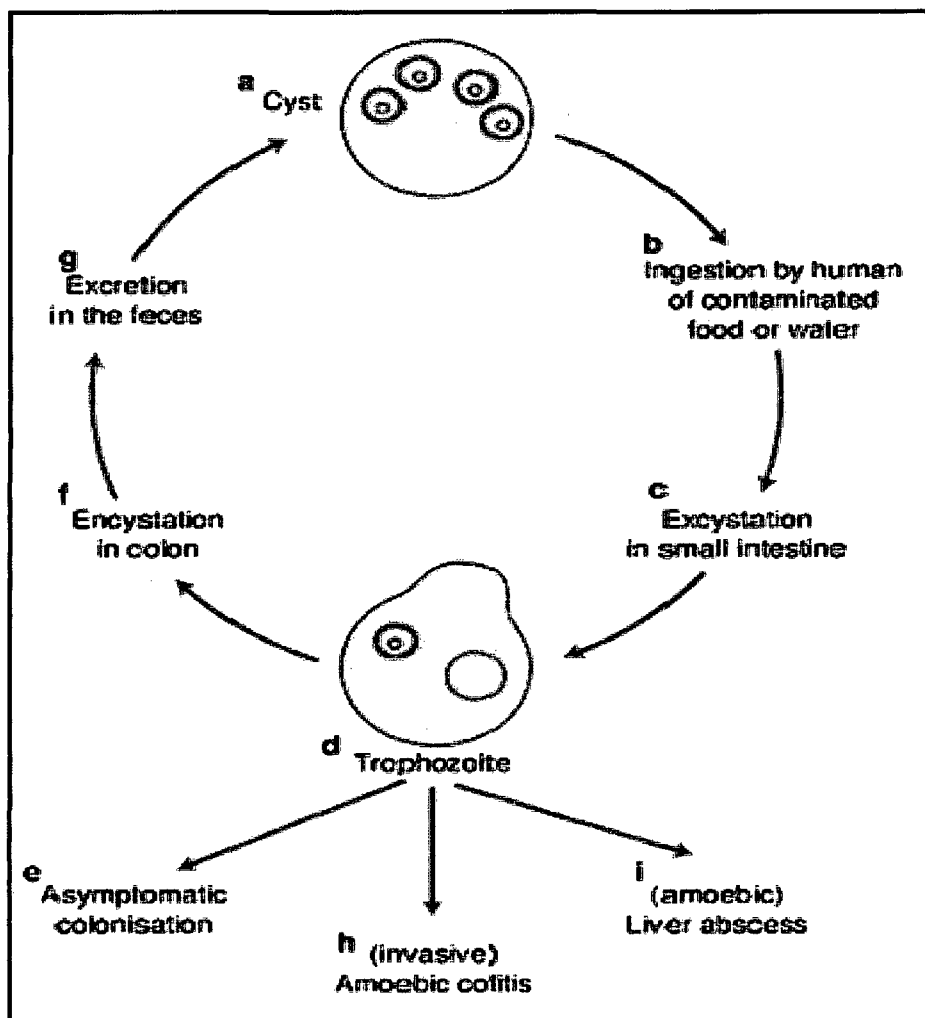


Fig. 1.6.1 Life cycle of *Entamoeba histolytica* and the clinical manifestations of infection in humans. Life cycle consists two stages: cysts and trophozoites. (a) Cysts measure 10–15 μm in diameter and typically contain four nuclei; (b) They spread via the ingestion of faecally contaminated food or water. (c) During excystation within the lumen of the small intestine, nuclear division is followed by cytoplasmic division, giving rise to eight trophozoites. (d) Trophozoites, which measure 10–50 μm in diameter and contain a single nucleus with a central karyosome, reside in the lumen of the caecum and large intestine, where they adhere to the colonic mucus and epithelial layers. (e) Approximately 90% of individuals infected with *E. histolytica* are asymptotically colonised; (f) Re-encystation of the trophozoites occurs within the lumen of the colon, resulting in (g) the excretion of cysts in the faeces and continuation of the life cycle. (h) Alternatively, the trophozoites can invade the colonic epithelium, causing amoebic colitis (in ~10% of infected people). Amoebic dysentery usually occurs gradually, with symptoms [such as abdominal pain and tenderness, and painful sudden bowel evacuation (tenesmus) and diarrhoea] developing over a period of one to several weeks, often followed by weight loss. *E. histolytica* can spread in the bloodstream (haematogenously) after it has penetrated the colonic epithelium (not shown) and can establish persistent extraintestinal infections, most commonly (i) amoebic liver abscess. Liver abscess is overwhelmingly the most common extraintestinal manifestation of amoebiasis. Adapted from (Huston et al., 1999).

1.1.4 Pathogenesis

In the pathogenesis of amoebiasis there is large interplay of many factors. It involves large number of steps and variety of agents. The food for *E. histolytica* ranges from human intestinal epithelium, erythrocytes, neutrophils to lymphocytes (Burchard and Bilke, 1992; Ravdin and Guerrant, 1981). Several virulence factors/molecules have been identified that are responsible for the pathogenesis of amoebiasis, however the prime molecules responsible for unleashing of the amoebiasis is still not evident. The whole process virulence can be divided into four major categories on the basis of the generation of experimental evidences. Following are the categories –

- (i) Cross talk with the intestinal flora,
- (ii) Destruction of the cells upon attachment ,
- (iii) Destruction of target cell by release of toxins and
- (iv) Engulfment of the attached cells

E. histolytica main mode of pathogenicity whereby it kills the victim's cells is contact dependent. Cytolysis of the host cells starts with the attachment of the target cells via galactose/N-acetyl D-galactosamine-inhibitable (Gal/GalNAc) lectin (Ravdin and Guerrant, 1982). There is instantaneous killing of target cell that is within 5–15 min and after this there is phagocytosis. The Gal/GalNAc lectin is a membrane complex comprised of a heavy subunit (Hgl, 170 kDa) and light subunit (Lgl, 30–35 kDa). These two chains are associated with the help of disulphide bonds, and a non-covalently associated intermediate subunit (Igl, 150 kDa) (Cheng et al., 2001; Mann, 2002). Lectin not only plays role in the adherence to the host cells but also offers resistance to complement lysis. The genome has more than one genes encoding subunits of Gal/GalNAc lectin.

The small molecules responsible for the lysis of host cells are a family of small proteins, amoebapores which are present in the lysosome-like granular vesicles of *E. histolytica*. Three amoebapore isoforms (amoebapore A, B, and C) have been reported and described in details. The structural aspect of these molecules matches with granulysins and NK-

lysins produced by mammalian T cells and has capacity of forming pores in lipid bilayers (Leippe, 1997). These proteins are 70–80 amino acids long and has typical compact alpha helical, disulphide-bonded structure known as the saposin-like fold. All the SAPLIP (putative saposinlike proteins) proteins have great affinity for the lipid and this is instrumental in killing of target cells. These variant forms of amoebapore are present in virulent and non virulent forms that is *E. histolytica* and *E. dispar*. Non-pathogenic *E. dispar* possess low amount of these peptides (Nickel et al., 1999).

The *E. histolytica* cysteine proteinases has been proven in invivo and invitro studies that these proteins has direct role to play in the overall pathogenicity that is invasion and inflammation of the gut (Ankri et al., 1999; Li, E., et al., 1995; Stanley et al., 1995). The very first entry of *E. histolytica* in the intestinal tissue (submucosal tissues) is with the help of cysteine proteinases that digest extracellular matrix proteins and facilitate the entry of the pathogen in submucosal layer of the instestine (Que and Reed, 2000).

Proteophosphoglycans (PPG) are rampant on the outer surface of *Entamoeba histolytica*. These proteins are hydrophobic and have a lot of glycosylation. It has two parts – one is of 110 kDa (PPG1) and other is of 45 kDa (PPG2) (Bhattacharya et al., 1992). These molecules are hallmarks of pathogenic form of *Entamoeba* and totally absent in non-pathogenic form that is *E. dispar* (Bhattacharya et al., 2000).

1.1.5 Genome of *E. histolytica*

The genome of *E. histolytica* has been sequenced and it has been found to be 23.7 Mbp in size and encodes for around 8264 (www.pathema.org) genes (Loftus et al., 2005). The *E. histolytica* genome (24 Mbp) (Loftus et al., 2005) is similar in size in comparison to another protozoan parasite genomes such as *P. falciparum* (23 Mbp) (Gardner et al., 2002), *Trypanosoma brucei* (26 Mbp) (Berriman et al., 2005) and the free-living amoeba *Dictyostelium discoideum* (34 Mbp) (Eichinger et al., 2005). The genome has a very high content of A+T (67 % in the coding region and 72 % in the intergenic region) and in each trophozoite has around 0.24 picogram of DNA (Dvorak et al., 1995). From the total estimated genes 31.8 % are totally unique and do not have any homologue. There are 14 chromosomes ranging in sizes from 300 to 2200 Kb (Willhoeft and Tannich, 1999). As

per analysis of the genome it has been found to be tetraploid (Willhoeft and Tannich, 1999).

In the *E. histolytica* genome there are many genes which have more than one copy. These genes belong to the multicopy gene family. These are virulence factor genes such as the Gal/GalNAc lectin intermediate subunit (30 copies), the cysteine proteinase family (atleast 20 different members), the protein kinase family (271 members across seven superfamilies) calcium binding proteins (CaBPs) and phosphatases (about 100 members) (Loftus et al., 2005).

1.1.6 Symptoms

In patients affected by amoebiasis there is manifestation of variety of symptoms and it varies from dysentery (Wilihoeft et al., 2001) with bloody mucus containing stools and constitutional symptoms (amoebic dysentery), to mild abdominal discomfort with diarrhea containing blood or mucus, alternating with periods of constipation or remission. Apart from this other things also appear such as prolonged pain in abdomen and tenderness, painful sudden bowel evacuation (tenesmus), and diarrhoea.

1.1.7 Treatment

A number ofazole derivative more specifically Nitroimidazole derivatives like metronidazole, tinidazole, ornidazole are the first line of treatment against amoebic infection. Metronidazole is right now the drug of choice for treatment of amoebiasis, but the other drugs are also used, for example tinidazole. As with any other drug these drugs also has some side effects like metronidazole produce nausea, vomiting, headache, and abdominal discomfort (Pehrson and Bengtsson, 1984). Metronidazole is not very effective against the luminal form of the disease. Amoebiasis has two lines of treatment first amoebic colitis is treated by metronidazole, second line is done by a luminal agent (paromomycin, iodoquinol, or diloxanide furoate) that eradicates colonization (Pehrson and Bengtsson, 1984; Powell et al., 1973). Chloroquine, an anti-malarial drug, also comes handy against the *E. histolytica* trophozoites, but is not used for the routine treatment, it is

only recommended for patients with large and multiple amebic abscesses. Availability of vaccine against *Entamoeba histolytica* will be a great boon for poor patient as it will prevent the occurrence of amoebiasis. But right now there is no vaccine against amoebiasis available.

1.2 Calcium and its importance

Calcium is soft alkaline and grey metal and is the fifth most abundant element by mass in the earth's crust after oxygen, silicon, aluminum, and iron. Even in sea it is fifth most abundant. As far as biological processes are concerned it has a big role to play. For e.g. calcium along with phosphate and hydroxide ion forms hydroxyapatite and this is a major ingredient of tooth and enamel. In bone almost all the calcium is in the bound form that is it is immobilized. In other organisms also calcium is very important, it exists as a calcium carbonate in egg shells, exoskeleton of mollusks, barnacles etc.

1.2.1 Calcium Signaling

All living cells are required to have a cross talk with the environment where it resides. And this communication between the two is facilitated by a lot of factors. So any change in the environment elicits a response from the living cell. The process of communication between a cell and its external environment through stimulus derived intracellular messengers is called signal transduction. In all living organisms from unicellular to multicellular various small compounds are used in response to the environment, these molecules are termed as second messengers. Intracellular signal transduction is largely carried out by these second messengers. The nature of these molecules has a lot of variation. Basically three kinds of compounds are used in the cell to communicate with outer world. These are hydrophobic (e.g., DAG, IP3 and phosphatidylinositols), hydrophilic (e.g., cAMP, cGMP, Ca^{2+} and Mg^{2+}) and gases (e.g., NO and CO) (Berridge et al., 2003; Pattni and Banting, 2004). After the receipt of information by these

molecules it is passed on to next downstream molecules which in turn lead to change in expression profile of the cell.

1.2.2 Calcium as a biological messenger

The credit for discovering the role calcium plays as a participant in the regulation of cellular activity goes to British physiologist Ringer (Medvedev et al., 2005). In 1883 he found that any deficiency of calcium impairs the functioning of heart muscle. In the later years a lot of work was done and a role of second messenger was assigned to calcium by Howard Rasmussen (Rasmussen et al., 1976). After sensing the stimuli from the environment there is increase in the concentration of calcium inside the cell. These calcium ions bind to the target proteins and bring about a change in shape and charge of the protein. This in turn led to cascading effect inside the cell which culminates in some kind of cellular response to the stimuli. Calcium control diverse functions like contraction, secretion, gene transcription, metabolism, proliferation and cell cycle, learning and memory and finally apoptosis and, in the case of dysregulation, cancer.

1.2.3 Uniqueness of Ca^{2+}

There are some unique properties of calcium that gives the calcium a special status. Calcium has greater affinity for the oxygen atom which is readily available by carbonyl and carboxyl group of the proteins. Though magnesium ion is also good competitor for oxygen yet calcium has upper hand because there are important differences between calcium and magnesium. Magnesium likes water more than calcium. The loss of hydration shell from magnesium is much slower. Calcium settles down along with phosphate which is the energy currency of the cell. That's why calcium is maintained at lower concentration. Ca^{2+} ion can not be metabolized, unlike other second messenger molecules.

1.2.4 Calcium homeostasis

The dynamic equilibrium of calcium inside the cell is delicately balanced by a number of processes. There is a continuous exchange of calcium between cytosol and other cell organelles (nucleus, endoplasmic reticulum and or related organelles) (fig. 1.6.2) (Berridge et al., 2003). In any cell there is always difference in the concentration of the calcium in the different compartment of the cellular organelles. Free calcium is kept in nanomolar range concentration the cytosol whereas in the ER , SR and extracellular fluid it is in the millimolar range (Clapham, 1995). There are some dedicated pumps (Ca^{2+} -ATPase) in the cell to maintain difference in the concentration of calcium across the different compartments of the cell. These pumps act as per requirement of the calcium inside the cytosol. These pumps including various ion exchangers ($\text{Na}^+/\text{Ca}^{2+}$ exchangers), mitochondrial exchangers ($\text{Na}^+/\text{Ca}^{2+}$ exchanger, Ca^{2+} uniporter, $\text{H}^+/\text{Ca}^{2+}$ exchanger), Golgi pumps (SPCA1 and SPCA2) either pump the calcium in the cytosol or back to the different store house that is ER or SR (Berridge et al., 2003). At any given point of time, a dynamic equilibrium is achieved between the cytosol and different cellular organelles by the pumping in and pumping out mechanism. (fig. 1.6.2)

Broadly the Ca^{2+} signalling network can be divided into four functional units

- I. On the receipt of the stimulus, Ca^{2+} mobilizing signals are elicited.
- II. Next step leads to the pumping of calcium in the cytosol.
- III. Now calcium acts as a second messenger and downstream processing of the signal is done
- IV. Then at the end of the process calcium is pumped back into the store house.

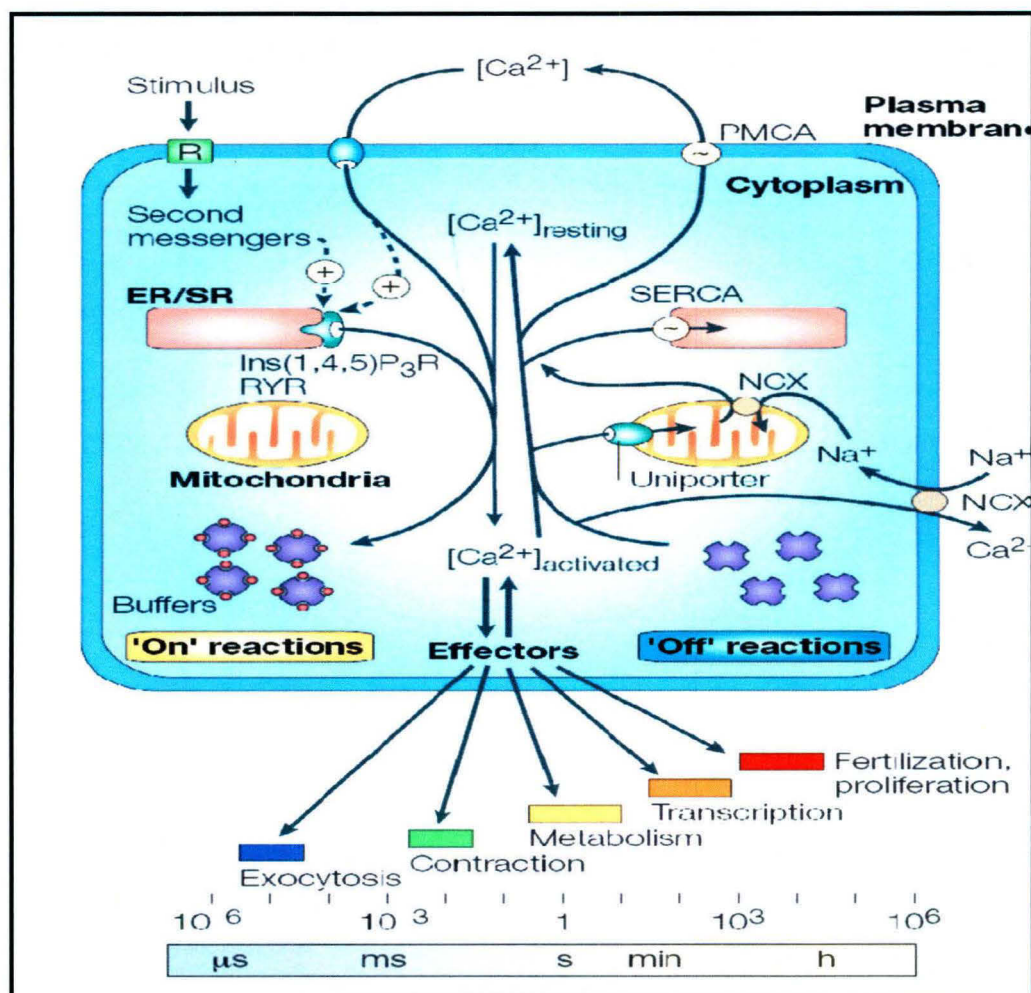


Fig. 1.6.2 Mechanisms responsible for calcium homeostasis in the cell. Examples of processes controlled by calcium-dependent effectors are shown with their respective time scales of occurrence. ER/SR: Endoplasmic reticulum/Sarcoplasmic reticulum; PMCA: Plasma membrane Ca²⁺ ATPase; R: receptor; SERCA: sarco/endoplasmic reticulum Ca²⁺ ATPase; Ins(1,4,5)P₃R: Inositol 1,4,5 tris phosphate receptor; RYR: Ryanodine receptor; NCX: sodium calcium exchanger. Adapted from (Berridge et al., 2003).

For increasing the concentration of calcium inside the cytosol different ion channels like Voltage Operated channel (VOCs), second messenger-operated channels (SMOC), store-operated channels (SOCs), receptor-operated channels (ROCs) and different exchangers are activated.

There are also certain receptors for calcium present in the cytosol itself like IP₃ receptor which acts a calcium channel on ER membrane. Sphingosine-1-phosphate controls efflux from ER through a sphingolipid Ca²⁺ release-mediating protein of endoplasmic reticulum (SCaMPER) (Mao et al., 1996). Another class of proteins known as cytosolic buffer proteins such as calbindin-D_{9k}, Calretinin and parvalbumin sequester almost all of the calcium thereby leaving very small fraction of free calcium (Berridge et al., 2003). These cytosolic buffers are the mainstay of the regulation of amplitude and duration of Ca²⁺ signals and in confining the spreading of local Ca²⁺ signals in nearby places. All Once the desired response is generated by the cell, all the calcium is sent back by various Ca²⁺ ATPase pumps (Pozzan et al., 1994), exchangers (Blaustein and Lederer, 1999) and a mitochondrial calcium uniporter (Favaron and Bernardi, 1985) to make the cytosol almost calcium free (Berridge et al., 2003). In this way, there are many mediators in the cell to execute the calcium response and subsequently remove the calcium from the cytosol.

1.3 Extracellular Calcium Binding Proteins

Extracellular Ca²⁺ plays a great role in higher organism that is multicellular, here it involved in one of the many vital functions such as blood clotting. In blood clotting it binds to modified γ -carboxyglutamate (Gla) residues, in the Gla domains (this domain is present the enzymes of blood clotting processes) (Maurer et al., 1996; Selander-Sunnerhagen et al., 1992). Ca²⁺ ion plays a structural roles in other extracellular calcium binding proteins such as C-type lysozymes, α -lactalbumin (McKenzie and White, 1991), pancreatic enzyme deoxyribonuclease I (Pan and Lazarus, 1999) etc.

1.3.1 Intracellular Calcium Binding Proteins

The first step of the calcium signal is the binding to specific Ca²⁺ receptors or Ca²⁺-binding proteins (CaBPs). The Ca²⁺ binding proteins has been grouped into three classes: inducer or sensor proteins (e.g., calmodulin, troponin C) (Berridge et al., 2000), buffer proteins (e.g., S100G and parvalbumin) (Schroder et al., 1996), or Ca²⁺-stabilized proteins (e.g., thermolysin) (Buchanan et al., 1986). This grouping has been done on the

basis of their work. The most common calcium binding motif in the biological system is the EF - hand motif. Till today more than 66 sub families has been included in the family of EF - hand proteins (Kawasaki et al., 1998). A list of different Ca²⁺ binding proteins is shown in table 1.7

Table 1.7 : List of Ca²⁺ binding proteins

Table 1. Calcium-binding proteins of <i>Entamoeba</i>						
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	Homologues in other species of <i>Entamoeba</i>
EhCaBP1	CaBP 1	134	EAL48959	4	Yes	<i>E. terrapinae</i> , <i>E. moshkovskii</i>
EhCaBP2	CaBP 2	134	EAL51694	4	Yes	<i>E. terrapinae</i> , <i>E. moshkovskii</i>
CaM	CaBP 3	151	EAL46322	3	Yes	ND
CaM par	CaBP 4	146	EAL51814, EAL46978	3	ND	ND
CaM	CaBP 5	144	EAL46660	2	ND	ND
CaM	CaBP 6	150	EAL50371, EAL50341	2	ND	ND
CaM-like	CaBP 7	155	EAL43751	4	ND	ND
CaBP	CaBP 8	263	EAL50453	2	ND	ND
CaBP	CaBP 9	157	EAL52004	4	ND	<i>E. moshkovskii</i>
CaBP	CaBP 10	153	EAL50237	3	ND	ND
CaBP	CaBP 11	627	EAL50040	2	ND	ND
CaBP	CaBP 12	145	EAL51761	4	ND	ND
CaBP	CaBP 13	336	EAL48778	2	Yes	ND
CaBP	CaBP 14	264	EAL46305	3	Yes	<i>E. dispar</i>
Granin 1	CaBP 15	215	EAL47654	3	ND	ND
Granin 1	CaBP 16	215	EAL44934, EAL44709, EAL44971	3	Yes	<i>E. terrapinae</i> , <i>E. dispar</i> , <i>E. moshkovskii</i>
Granin 2	CaBP 17	213	EAL44955, EAL44970	3	Yes	ND
Granin 2	CaBP 18*	630	EAL49043, EAL42990	3	Yes	ND
CaBP/URE3BP	CaBP 19	220	EAL42646	3	Yes	ND
Myosin light chain	CaBP 20	146	EAL52163, EAL50546	3	ND	<i>E. dispar</i>
Calcinurin B-subunit	CaBP 21	179	EAL45078	2	Yes	<i>E. terrapinae</i> , <i>E. dispar</i> , <i>E. invadens</i> , <i>E. moshkovskii</i>
Hypothetical	CaBP 22	317	EAL45535	3	ND	ND
Hypothetical	CaBP 23	169	EAL43738	2	ND	ND
Hypothetical	CaBP 24	318	EAL49541	6	ND	ND
Hypothetical	CaBP 25	307	EAL45702	6	ND	ND
Hypothetical	CaBP 26	673	EAL51827	6	ND	ND
Hypothetical	CaBP 27	658	EAL51029	5	ND	ND

Adapted from (Bhattacharya et al., 2004).

1.3.2 Calcium binding motifs

Broadly four kinds of calcium binding motifs have been outlined keeping in view some of the typical characteristic of each group.

- 1) Annexin domain (Gerke et al., 2005)
- 2) C2 domain (Sutton and Sprang, 1998)

- 3) C-lectin (Weis et al., 1998) and
- 4) EF-hand domain (Kretsinger and Nockolds, 1973).

The way the calcium is bound to the proteins is typical to the the particular group of calcium binding motif. The separate group of calcium binding proteins that is EF hand family of proteins has 12 residues in the calcium binding loop of which 6 residues (negatively charged) are actually coordinated with the calcium in pentagonal bipyramidal geometry. Whereas, in other structural motifs the actual number residues (acidic) taking part in the calcium coordination is 4-5 which are far apart but come closer in the tertiary structure. Annexin, C2 domain proteins goes to the cell membranes in response to calcium signaling [fig. 1.6.3 (a) (b)]. And the bound Ca^{2+} is a linker between proteins (Annexin proteins and C2 domains) and the membrane surface [fig. 1.6.3 (a) (b)]. In the above mentioned case calcium interacts with the acidic residue of the calcium binding protein and by a phosphoryl head group of the cell membrane (Nalefski and Falke, 1996; Seaton and Dedman, 1998). Calcium binding to the C2 domain of synaptotagmin brings about a change in the protein which makes it suitable for the binding to the plasma membrane induces little conformational change (Sutton et al., 1995). Calcium binding to C - lectins at the membrane acts as a identifier in endocytosis and cell - cell interactions (Weis et al., 1998). Another group of EF hand proteins in the neuronal calcium sensing family, such as frequenin (NCS-1), visinin-like proteins (VILIPs), Kv-channel interacting proteins (KChIPs), recoverin and guanylate cyclase activating proteins (GCAPs) has the affinity for the membrane and attachment is done through myristoylated N-terminus (Burgoyne and Weiss, 2001; R, 2004). The majority of the calcium binding proteins involved in cell signaling are EF hand proteins.

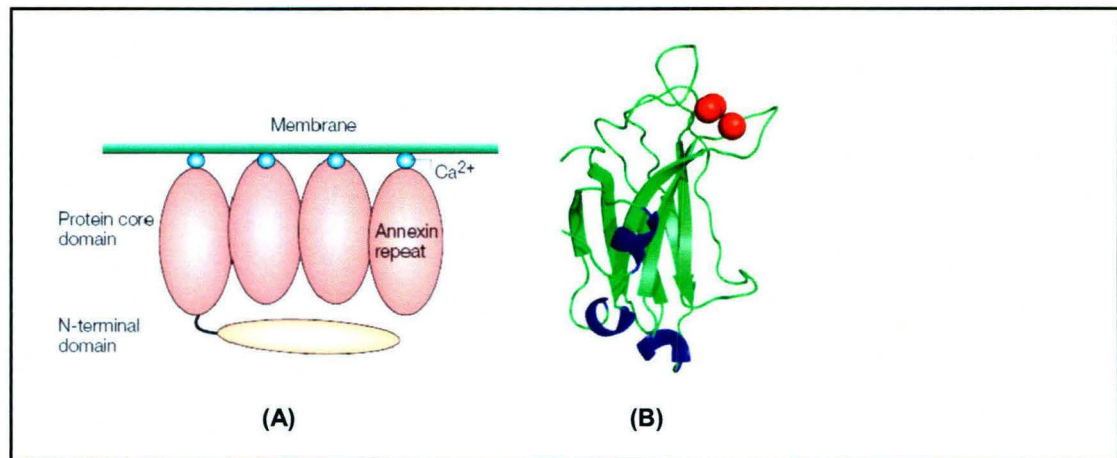


Fig. 1.6.3 Different calcium binding motifs (A) Schematic drawing of an annexin that is peripherally attached to a membrane surface through bound Ca^{2+} ions (blue), adapted from (Gerke et al., 2005), (B) Structure of C2 domain protein kinase C consisting of two, four stranded β -sheets creating three loops [loops 1 and 3 binds Ca^{2+} (red sphere)] at the top of the domain and four at the bottom. Adapted from (Sutton and Sprang, 1998).

1.4 The EF-hand: A Ca^{2+} -Binding Unit

In the EF hand family of proteins the functional unit is a pair of EF-hand motifs (Kretsinger and Nockolds, 1973). In an EF hand motif there are three parts namely - helix, loop and helix (fig. 1.6.4). The first helix which is N terminal and known as E consist of 10-12 residues. Next to helix E is loop which is calcium coordination loop and has 12 residues (fig. 1.6.4). And the last part is helix F which is at C terminal and consists of 10-12 residues (fig. 1.6.4). Structurally the two helices (alpha) E and F are 90 degree to each other. The name EF hand derives from the the three-dimensional arrangement of helix , loop and helix which gives an impression of the thumb, index and middle fingers of a hand (Kretsinger and Nockolds, 1973). In the majority of the cases the EF-hand motif occur in pairs and there is a small anti-parallel β -sheet present in between the paired EF-loops, which is responsible for cooperative Ca^{2+} -binding (Osawa et al., 1999). Broadly the EF-hand has been classified into major classes the canonical EF-hands (present in calmodulin (CaM) and other CaM like proteins) and the other group is non-canonical EF-hands found in the N-termini of S100 and S100-like proteins.

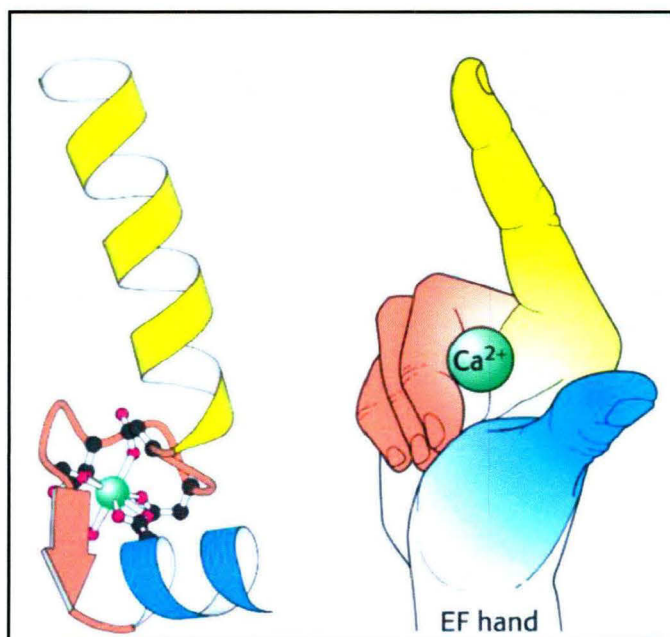


Fig. 1.6.4 EF hand Ca^{2+} binding motif. The helix (blue) loop (light brown) helix (yellow) motif is represented as thumb, middle and index fingers respectively of a hand. Adapted from (www.biochem.arizona.edu/classes/bioc462/462).

1.4.1 The Canonical EF-loop

The canonical EF-hand motifs which has been most extensively studied is highly conserved in the calcium coordination loop at position 1, 3, 5, 7, 9, and 12 and oxygen is offered by these residues to the calcium for binding [fig. 1.6.5 (a)]. At the position 12 of the calcium coordination loop generally glutamate is found and it offers both its side-chain oxygens for calcium coordination. Out of the seven ligands used for the coordination with the calcium five are given by the nine-residue loop (loop sequence positions 1, 3, 5, 7 and 9) [fig. 1.6.5 (b)]. The rest of the ligands that is three or four ligands are contributed by the side-chain carboxy groups of negatively charged amino acids and one from a backbone carbonyl group [fig. 1.6.5 (b)]. There is a lot of variation in the calcium binding constant in different calcium binding proteins and might be having some coorelation with the sequence position at 1, 3, 5, 7, and 9.

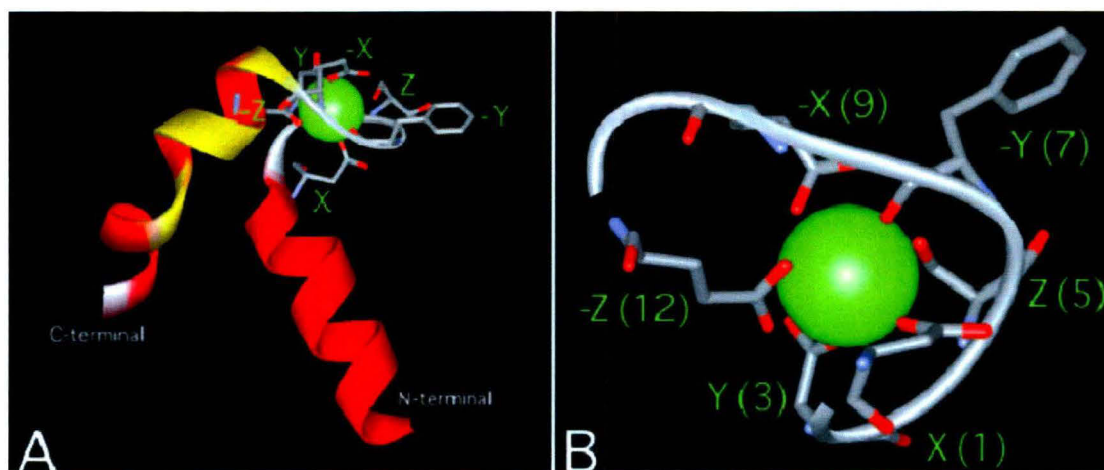


Fig. 1.6.5 Three dimensional structure of an EF-hand motif from parvalbumin. It consists of (A) two perpendicular α -helices (red) with a Ca^{2+} -binding loop of 12 amino acids (grey). The amino acids in the loop that provide oxygen ligands to the Ca^{2+} ion (green) are indicated with X (No. 1), Y (No. 3), Z (No. 5), 3Y (No. 7), 3X (No. 9) and 3Z (No. 12), which are Asp, Asp, Ser, Phe, Glu and Glu in the case of parvalbumin. (B) Close view of the six amino acids in the loop that provide oxygen ligands to the Ca^{2+} ion in a pentagonal bipyramidal fashion. Adapted from (Honore and Vorum, 2000).

1.4.2 The Non-Canonical EF-loop

So the mechanism mentioned above explains most of the calcium binding to the EF hand. And there is a lot of diversity in terms of length and composition of EF hand. The different non canonical EF loop can be grouped into four groups.

- I. The length of the EF hand in this group is identical to the canonical loop (12 residues). The variation lies in the coordination of calcium to the ligands. At the 12th position of canonical loop glutamic residue is present but in case of non canonical loop aspartic residue is found (e.g. EF3 from CIB (calcium- and integrin-binding protein) (Gifford et al., 2007). The second variation is binding of the calcium to the coordination loop which is primarily by main-chain carbonyl groups such as EF4 of AtCBL2 (*Arabidopsis thaliana* calcineurin B-like protein) (Nagae et al., 2003).
- II. In the second variety length is different and it is grouped into those, where length of EF-loop is longer (14 residue) than canonical EF-loop due to

insertions at N-terminal part of the loop. These are called as pseudo EF-loop forming an EF hand variant known as the ψ hand. This pseudo EF-hand loop chelates Ca^{2+} mainly via backbone carbonyls (positions 1, 4, 6, 9). EF1 of S100 and S100-like proteins including calbindin $\text{D}_{9\text{K}}$ are typical examples of this group.

- III. Members from this class have EF hand sequence shorter than canonical sequence. It has 11 residues only. Examples of this group are EF1 in most members of the penta-EF-hand subfamily (Gifford et al., 2007).
- IV. This group is totally different from the rest of the three groups with respect to configuration that it attains in the calcium coordination. This group attains octahedral configuration whereas the other three members attain pentagonal bipyramidal. It possesses two additional residues at the C-terminal part of loop. Examples are EF5 of the apoptosis-linked protein ALG-2 (apoptosis-linked gene-2). It belongs to the penta-EF-hand subfamily and has 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).

1.4.3 Calmodulin: a prototypical calcium sensor

Calmodulin (CaM) (an abbreviation for **CAL**cium **MODUL**ated **proTEIN**) is a calcium-binding protein present across all the all eukaryotic cells (including plants, yeast and animals) (Vogel, 1994; Zhang and Yuan, 1998). CaM, the most investigated protein is small acidic protein of 148 amino acids (~17.0 kDa). The X-ray crystal structure of Ca^{2+} -saturated CaM shows that it has dumbbell shape with two lobes (globular domains), each containing a pair of EF-hand motifs, separated by a long, central α -helical linker (Chattopadhyaya et al., 1992) [fig. 1.6.6 (a)]. These two symmetrical globular domains are very much similar (75% homology) (Chou et al., 2001; Fallon and Quioco, 2003) and structurally similar in the presence and absence of Ca^{2+} ions (Kuboniwa et al., 1995; Zhang et al., 1995). However, biochemically these two domains are distinct and calcium

affinity of CaM-C is 10-fold stronger than CaM-N (Linse et al., 1995; Bhattacharya et al., 2004). CaM-C has more acidic residues than CaM-N whose acidic patches has been partly neutralized by the basic residues (binding pocket). The central linker is highly flexible which allows it to interact with variety of targets (Barbato et al., 1992). Cam binds to one calcium in each calcium binding motif and exhibit conformational change upon calcium binding. In response to elevated intracellular Ca^{2+} levels, CaM can bind four Ca^{2+} ions, one in each EF-hand calcium binding motifs. Conformational change leads to the exposure of hydrophobic residue (methionine rich crevices), these hydrophobic residues interact with Basic Amphiphilic Helices (BAA helices) on the target proteins [fig. 1.6.6 (b)]. These helices contain complementary hydrophobic regions. Calcium bound calmodulin targets lot of Cam kinases (Hoeflich and Ikura, 2002), adenylylate cyclase (Kessin and Franke, 1986), phosphodiesterase (Lin et al., 1974) and calcineurin (Klee et al., 1979) and modulates their activity, thereby affecting the downstream targets that regulate cellular processes such as motility, secretion, cell differentiation, apoptosis and gene expression.

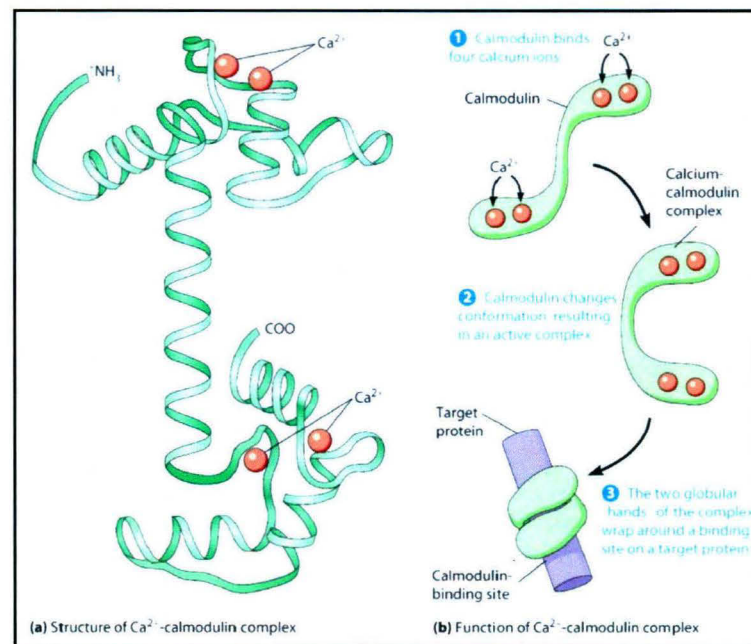


Fig 1.6.6 Schematic representation of calmodulin (A) showing the four Ca^{2+} ions in light red color and (B) calcium-bound calmodulin complexed with target protein. Adapted from (The World of the Cell, 7th edition Becker, Kleinsmith, Hardin & Bertoni).

1.5 Calcium in Protozoan Biology

In protozoa, intracellular calcium ions are involved in variety of activities including ciliary, flagellar, and ameboidal motion (Prusch and Hannafin, 1979), exocytosis (Gilligan and Satir, 1983), and the release of surface coat proteins (Voorheis et al., 1982). The intracellular calcium ion (Ca^{2+}) is used as a major signaling molecule in case of a several human parasitic protozoa, such as *Trypanosoma cruzi*, *Trypanosoma brucei*, *Leishmania* spp, *Plasmodium* spp, *Toxoplasma gondii*, *Cryptosporidium parvum*, *Entamoeba histolytica*, *Giardia lamblia* and *Trichomonas vaginalis*.

In *T. cruzi* calcium helps in the invasion of non-phagocytic host cells which occurs by the fusion of host cell lysosome with plasma membrane at the entry (Berriman et al., 2005; Tardieux et al., 1992).

In *Toxoplasma gondii*, the invasion of host cells is calcium dependent. And the invasion process involves motility, cytoskeletal rearrangements, attachment and secretion (Lovett and Sibley, 2003). *Plasmodium falciparum* is known to require Ca^{2+} for the regulation of its cell cycle and for its long term survival. In *Plasmodium falciparum* calcium-dependent protein kinase 1 (PfCDPK1) is present during schizogony of blood stages and in the infectious sporozoite stage. It is involved in active membrane biogenesis.

Centrin a calcium binding cytoskeletal protein needed for centrosome duplication overcomes the G2/M check point, thus calcium responsible for *Leishmania* growth (Selvapandiyar et al., 2001).

In *Giardia lamblia*, CaM crucial for the starting of infection is involved in its excystation process (Bernal et al., 1998).

TH-16258

CaM has been found and characterized in many protozoa, e.g. *Tetrahymena*, *Leishmania*, *Giardia* and *plasmodium*. Different intracellular organelles such as the ER, mitochondria, and acidocalcisomes are controlling the Ca^{2+} regulation in many protozoan parasites. In some protozoa where these organelles are absent calcium regulation is carried out by homologous organelles such as the mitosome/cryptome of *E. histolytica* or the hydrogenosome of trichomonads. *Giardia lamblia* do not have mitochondria.



572-67
M3144
St.

Acidocalcisomes are acidic calcium stores of Ca^{2+} in many trypanosomatid and apicomplexan protozoa (Docampo and Moreno, 2001). These are absent in mammals. Many CaM controlled enzymes such as Ca^{2+} -ATPases and adenylyl cyclase are CaM independent in protozoa in in vitro conditions in contrast to eukaryotes.

1.5.1 Role of Calcium in *Entamoeba*

Ca^{2+} signalling is crucial to the pathogenesis of many protozoan parasites, including *E. histolytica* (Ravdin et al., 1985). Calcium is released in the target cells which lead to cytolysis by *E. histolytica* (Ravdin et al., 1988). Calcium and phorbol esters are also playing a lead role in the release of many digestive enzymes, such as proteinases and lipases, some of the molecules thought to be involved in pathogenesis (Meza, 2000; Weikel et al., 1988). In general, functions of calcium ions are mediated by a set of calcium binding proteins (CaBPs) and in case of *E. histolytica* it has been shown that a repertoire of CaBPs (27 in number) exist (Bhattacharya et al., 2006).

1.5.2 EhCaBP1 and EhCaBP2: Calcium binding proteins in *Entamoeba histolytica*

Two novel Ca^{2+} -binding proteins (EhCaBP1 and EhCaBP2) from *E. histolytica* has been characterized well (Chakrabarty et al., 2004; Prasad et al., 1993). Both are 134 amino (14.7 kDa) acid residues and have four canonical EF-hand Ca^{2+} binding domains. In comparison to CaM they are similar in negative charge, heat stability and number calcium binding domains. Unlike CaM, EhCaBP1 and EhCaBP2 do not stimulate c-AMP phosphodiesterase (a diagnostic test for CaM). It has been shown that EhCaBP1 is responsible for the actin remodeling pseudopod formation and phagocytosis (Sahoo et al., 2004). Both EhCaBP1 and EhCaBP2 vary in their central linker, upstream and downstream sequences. At the nucleotide sequence level both are 79 % identical but in the central linker region identity is at 40 %. Both EhCaBP1 and EhCaBP2 are different with respect to activation of kinases and recognition of different targets (Chakrabarty et al., 2004).

CHAPTER 2

Expression, Purification, Crystallization and Preliminary Crystallographic Analysis of EhCaBP2 and its Complexes

2.1 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 developing countries. There are a number of studies that show the involvement of Ca^{2+} and its binding proteins in amoebic pathogenesis (Ravdin et al., 1988). Previously, a novel Ca^{2+} -binding protein from *E. histolytica* (EhCaBP1) has been characterized and its three dimensional structure has been derived using multidimensional nuclear magnetic resonance (NMR) spectroscopic techniques in the apo form as well as in a complexed form with Ca^{2+} (Atreya et al., 2001; Yadava et al., 1997). EhCaBP1 is a 14.7 kDa (134 amino-acid residues) protein that has been shown to participate in cytoskeletal dynamics (Sahoo et al., 2004). The study reveals the presence of two globular domains connected by a flexible linker region spanning eight amino-acid residues. EhCaBP1 binds to four Ca^{2+} ions with high affinity (two in each domain) and it is structurally related to calmodulin (CaM) and troponin C (TnC), despite having low sequence homology with these proteins. The NMR structure shows a more open C-terminal domain for EhCaBP1 with a larger water-exposed total hydrophobic surface area compared with CaM and TnC (Atreya et al., 2001). Further dissimilarities between the structures include the presence of two Gly residues (Gly 63 and Gly 67) in the central linker region in EhCaBP1, which seem to impart a greater flexibility compared with CaM and TnC and may also play a crucial role in its biological function. The major differences in the structure of EhCaBP1 with respect to those of CaM and TnC are in the Ca^{2+} binding loops, interhelical angles and exposed hydrophobic surface. These structural features make EhCaBP1 functionally distinct from other CaM like Ca^{2+} binding proteins. In a recent study, a paralogous isoform of EhCaBP1, EhCaBP2, was identified and partially characterized (Chakrabarty

et al., 2004). The two isoforms are encoded by genes of the same size (402 bp). Comparison between the two genes showed an overall identity of 79 % at the nucleotide-sequence level (78 % at the protein level). This identity dropped to 40 % in the 75-nucleotide central region (56 % at the protein level) between the second and third Ca²⁺ binding domains. An array of biochemical studies indicated that despite their structural similarities, the two EhCaBPs are functionally distinct.

EhCaBP2, a CaM-related protein involved in cytoskeletal dynamics, may bind to the IQ motifs of myosin. Myosins, neuronal growth proteins, voltage-gated channels and certain signaling molecules contain IQ motifs (Bahler & Rhoads, 2002) that can bind to either CaM or CaM-related proteins. These motifs are of about ~25 amino acids in length and conform to the consensus sequence (I,L,V)QxxxRxxxx(R,K) (reviewed by Bahler & Rhoads, 2002).

2.2 Materials and Methods

2.2.1 Sources of Materials

E. coli strain DH5 α was obtained from Bethesda Research Labs (B.R.L., U.S.A). DEAE & Phenyl Sepharose resin was purchased from Amersham Pharmacia (U.S.A.), Promega (U.S.A.), Sigma (U.S.A.) and other reagents from Sigma Aldrich (U.S.A.) and Qualigens (India). *E. coli* media components were from DIFCO (U.S.A.). 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).

2.2.2 Organisms and Growth Conditions

E. coli DH5- α has the genotype: *SupE44 lacU169* (ϕ 80 *lacZ* M15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*. Cells from an agar stab or frozen glycerol stock was 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 225 rpm at 37°C. The cells were grown overnight, were used as inoculum for further growth by diluting 100 fold in fresh LB medium and grown with aeration at 37°C for 3-4 h to obtain log phase cultures.

2.2.3 Culture Media

2.2.3.1 Luria Broth (LB)

Bacterial cells were grown in Luria Broth (LB). It was prepared by dissolving 25 gms of LB powder (Amersham) in 1 liter of distilled water and pH adjusted to 7.0 using 2 N NaOH. The medium was sterilized by autoclaving.

2.2.3.2 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 (when required) after cooling the LB agar to around 55°C and plates were poured.

2.2.4 Plasmid DNA Isolation from *E. coli* Transformants

2.2.4.1 Preparation of competent cells and transformation (Hanahan, 1983)

Single colony of *E. coli* [strain DH5 α , BL21(DE3), SG130069] was grown overnight in 5ml LB medium and 1% inoculum was added to 50 ml LB medium in 500 ml flask. The cells were grown at 37°C to an OD₆₀₀ of 0.38-0.42. The cells were vigorously shaken on ice water for 15 min and were thereafter collected by centrifugation at 5000 rpm for 5 min at 4°C and resuspended in 25 ml ice-cold filter sterilized 0.1 M CaCl₂. Cells were incubated for 1 h on ice with occasional shaking; the cells were collected again by centrifugation at 5000 rpm for 5 min at 4°C. The halo shaped pellet was finally resuspended in 2 ml of ice-cold 0.1 M CaCl₂. The competent cells were stored in 15% glycerol stocks in 100 µl aliquots at -70°C.

Competent cells were thawed on ice and to 100 µl cell suspension, 5-10 ng of plasmid DNA was added. The cells were incubated on ice for 30 min. Cells were then given a heat shock at 42°C for 90 s and incubated on ice for 5 min. 0.9 ml of LB was added to the cells and the cells were grown at 37°C for 1 h at 200 rpm. Transformants were plated on LB agar plates with appropriate antibiotic and incubated at 37°C for 14-16 h.

2.2.4.2 Mini-preparation of plasmid DNA (Alkaline lysis method) (Birnboim and Doly, 1979)

A single colony harboring the desired plasmid was inoculated in 2 ml of LB medium containing appropriate antibiotic and grown overnight at 37°C. The cells were pelleted at 6,000 rpm for 5 min and soup was aspirated out. The pellet was suspended in 100 µl of solution I (50 mM glucose, 25 mM Tris-Cl pH 7.5, 10 mM EDTA pH 8.0). To the tube

200 μ l of freshly prepared solution II (0.2 N NaOH and 1% SDS) was added, mixed gently by inverting and incubated on ice for 5 min, 150 μ l of chilled solution III (3 M potassium acetate, pH 5.2) was then added and the contents were mixed gently by inverting the tube and kept on ice for 10 min. The mixture was centrifuged at 14,000 rpm for 10 min at 4°C. The soup was transferred to the fresh tube and 0.7 volumes of isopropanol was added and centrifuged at 14000 rpm for 10 min. The pellet was washed with 70 % ethanol by centrifugation at 14000 rpm for 5 min at RT. The soup was discarded and pellet was air dried. The dried pellet was suspended in T₁₀E₁-RNase or autoclaved Milli-Q.

2.2.4.3 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 TBE [45 mM Tris- borate and 1mM EDTA, pH 8.0] by heating and was cooled to about 50⁰C before adding 0.5 μ g/ml of ethidium bromide. The molten agarose was poured in a tray and allowed to gel. 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 UV trans-illuminator.

2.2.5 SDS-Polyacrylamide gel electrophoresis (Laemmli, 1970)

SDS-PAGE was carried out under reducing conditions. The separating gels (10-14 % 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 2 min. After electrophoresis, proteins were fixed in the gel by incubating in fixing solution (50 %

methanol, 7.5 % acetic acid) and detected by Coomassie Brilliant Blue (0.25 % CBB R-250 in fixing solution) staining for 1 h. The gels were destained in the fixing solution and dried.

2.2.6 Protein estimation (BCA assay)

The amount of protein in a sample was estimated by the bicinchoninic acid assay using BSA as the standard (Smith et al, 1985). The working solution was prepared by mixing bicinchoninic acid (Sigma, U.S.A.) and 4 % copper sulphate in a ratio of 50:1. To 20 μ l of protein (appropriate dilutions) was added 180 μ l of the working solution in a microtitre plate and incubated at 37°C for 30 min. The absorbance was taken at 562 nm using a microtiter plate reader (Bio-Rad, U.S.A.).

2.2.7 Expression and purification of recombinant EhCaBP2 in *Escherichia coli*.

A 5ml culture (LB + 60 μ g/ml kanamycin) of a single well-isolated EhCaBP2 colony was inoculated in a 100 ml conical flask (Borosil) and incubated overnight at 37°C at 220 rpm. A 200 ml culture (LB + 60 μ g/ml kanamycin) in 1L of conical flask was inoculated with 1% of the above culture and grown at 37°C/220 rpm till the OD₆₀₀ 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 1/25th 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 1min 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.

The purification of EhCaBP2 was performed on DEAE anion exchange chromatography. The column was equilibrated with 20 bed-volumes of 50 mM of Tris-Cl (pH 7.5) + 2 mM

EGTA (pH 7.5). The sample was loaded to the column 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 CaCl₂ and the OD₂₈₀ was taken. This protein was further purified using Phenyl Sepharose affinity chromatography. The protein was dialyzed in 5 mM CaCl₂, 50 mM HEPES buffer pH 7.5 and passed through the Phenyl Sepharose column, which had been pre-equilibrated with the same buffer. This column was washed with 500 mM NaCl with 50 mM HEPES buffer pH 7.5, 5 mM CaCl₂ to remove non specific bound proteins. The column was further washed with 5 mM CaCl₂, 50 mM Hepes buffer pH 7.5. The EhCaBP2 was eluted with 5 mM EGTA, 10 mM EDTA in 25 mM cacodylate buffer pH 5.0. The sample was 99% pure as estimated by SDS-PAGE [fig. (2.3.1)]. The CaBP2 was concentrated to 15 mg/ml [estimated by BCA assay (Smith et al., 1985)] using Centricon microconcentration devices (Amicon Inc. Beverly, MA, USA) and used for crystallization.

2.2.8 Preparation of Sr-EhCaBP2 complex

The purified EhCaBP2 was dialyzed against 6 M urea, 10 mM EDTA to remove the bound Ca²⁺ from the EhCaBP2. Subsequently, this apo-form EhCaBP2 was dialyzed against 10 mM cacodylate buffer with 5 mM SrCl₂ with two buffer changes to form the Sr-EhCaBP2 complex. The Sr-EhCaBP2 complex was concentrated to 15 mg/ml [estimated by BCA assay (Smith et al., 1985)] using Centricon microconcentration devices (Amicon Inc. Beverly, MA, USA) and used for crystallization.

2.2.9 Preparation of IQ1-EhCaBP2 complex

EhCaBP2 was mixed with IQ1 peptide motif (generously donated by R. Dominguez, Boston Biomedical Research Institute, Boston, MA, USA) in a 1:2 molar ratio. Initially, the IQ1 peptide was dissolved in DMSO and then added to EhCaBP2 which had been pre equilibrated with 50 mM acetate buffer pH 3.5. This mixture of IQ1 and EhCaBP2 was dialyzed against 50 mM cacodylate pH 5.0, 5 mM CaCl₂. When a similar experiment was performed without calcium, more precipitate was observed in the dialysis membrane.

However, when the pH was raised to 5.0 some precipitate was observed in the dialysis membrane. IQ1 was bound to EhCaBP2 and the excess IQ1 was precipitated in the presence of calcium, while all the IQ1 precipitated and none was bound to EhCaBP2 in the absence of calcium. It is noted that EhCaBP2 activates kinase in a calcium-dependent manner (Chakrabarty et al., 2004). The IQ1–EhCaBP2 complex was concentrated to 15 mg/ml [estimated by BCA assay (Smith et al., 1985)] using Centricon microconcentration devices (Amicon Inc. Beverly, MA, USA) and used for crystallization.

2.2.10 Crystallization

Initial crystallization experiments were performed at both 289 K and cold-room (277 K) temperature for EhCaBP2 and its complex with IQ1 using the hanging-drop vapor-diffusion method. The crystallization trials were performed using different ratios of protein and precipitant solution (1:1, 1.5:1, 2:1, and 3:1) and then equilibrated by vapor diffusion with same precipitant. The details of crystallization conditions for EhCaBP2 and its complexes are summarized in Table 2.3.3.

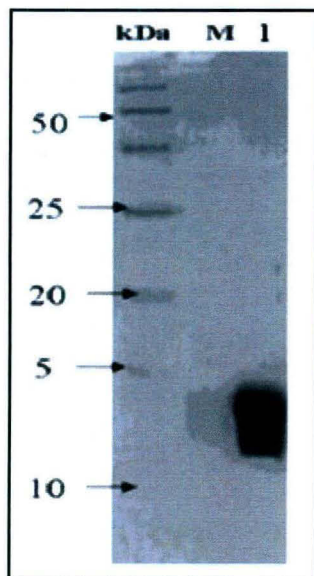
2.2.11 Data collection and processing

The X-ray diffraction experiments were performed at 100 K using crystals mounted in cryoloops at cryogenic conditions and flash frozen in liquid nitrogen. All data sets were collected at 100 K. The X-ray diffraction data from EhCaBP2 crystals were collected at BNL (Brookhaven National Laboratory) using beam line X9. The data from Sr–EhCaBP2 crystals were collected at CHESS using the A1 beam line. The data from IQ1–EhCaBP2 crystals were collected at SSRL beam line 11-1. All the data sets were indexed and scaled with the programs DENZO and SCALEPACK (Otwinowski & Minor, 1997).

2.3 Results and Discussion

Table 2.3.3

Crystallization of EhCaBP2 and its complexes.



EhCaBP2	45–60 % (w/v) MPD, 5 mM CaCl ₂ , 50 mM acetate buffer, pH 4.6
St–EhCaBP2	40–50 % (w/v) MPD, 50 mM acetate buffer pH 4.6, 5 mM SrCl ₂
IQ1–EhCaBP2	30 % (w/v) ethanol, 40–50 % (w/v) MPD, 50 mM cacodylate buffer, pH 6.5

Fig. 2.3.1 14 % SDS PAGE showing purified EhCaBP2 in lane 1.

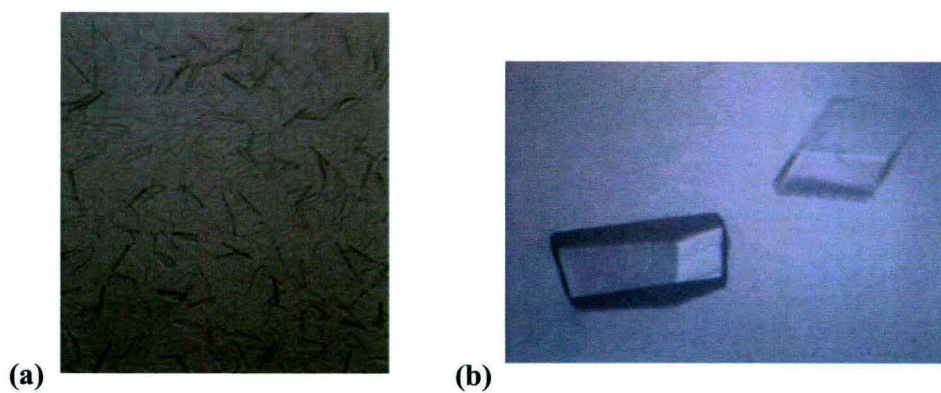
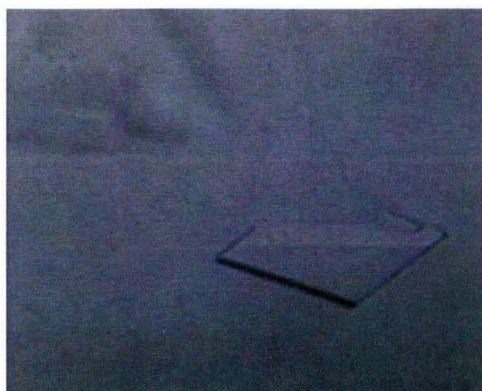


Fig. 2.3.2 (a), (b) EhCaBP2 crystallization condition 45–60 % (w/v) MPD, 5 mM CaCl₂, 50 mM acetate.



(c)

Fig. 2.3.2 (c) Sr-EhCaBP2 crystallization condition 40–50 % (w/v) MPD, 50 mM acetate buffer pH 4.6, 5 mM SrCl₂.



(d)

Fig. 2.3.2 (d) IQ1-EhCaBP2 crystallization condition 30 % (w/v) ethanol, 40–50 % (w/v) MPD, 50 mM cacodylate buffer pH 6.5.

Table 2.3.4 (values in parentheses are for the last resolution shell)

Data Collection Statistics		
	EhCaBP2	Sr-EhCaBP2
X ray source	BNL X9	Chess A1
Wavelength (Å)	1.25473	0.9764
Space Group	P2 ₁	P2 ₁
Unit - cell parameters		
a (Å)	111.74	69.2
b (Å)	68.8	112
c (Å)	113.25	93.4
β(°)	116.7	92.8
Resolution range	30.0-2.5	20.0-2.68
R _{sym} (%)	5.2	8.5(36.0)
Completeness (%)	87.9	88.1(66.8)
Total number of observations	565291	447423
No of unique observations	47410	37740
Redundancy	11.9	5.9
Average I/σ(I)	22.5	10.2(3.3)
Crystal mosaicity (°)	0.4	0.44

Table 2.3.5 (values in parentheses are for the last resolution shell)

Data Collection Statistics	
	IQ1-EhCaBP2
X ray source	SSRL 11-1
Wavelength (Å)	0.9764
Space Group	P2 ₁
Unit - cell parameters	
a (Å)	60.5
b (Å)	69.86
c (Å)	86.5
β(°)	97.9
Resolution range	50.0-3.11
R _{sym} (%)	4.5(32.5)
Completeness (%)	94.8(95.7)
Total number of observations	497664
No of unique observations	13758
Redundancy	36.2
Average I/σ(I)	19.9(3.8)
Crystal mosaicity (°)	1.5

EhCaBP2 purity was assayed on SDS – PAGE (fig. 2.3.1) and found it to be 99 % pure. Only these fractions containing 99 % pure EhCaBP2 were taken for crystallization trials. The initial trials of EhCaBP2 crystallization gave high- mosaicity poorly diffracting crystals in 45–65 % methylenepentanediol (MPD) with 5 mM CaCl₂ and 50 mM acetate buffer pH 4.6 using the hanging drop vapor - diffusion method at 289 K. In a subsequent crystallization experiment, EhCaBP2 was crystallized under similar conditions in the presence of 1 mM strontium (Table 2.3.3) as an additive [(fig. 2.3.2 (b)]. These crystals diffracted to 2.6 Å resolution and belonged to space group P21, with unit-cell parameters $a = 111.74$, $b = 68.83$ $c = 113.25$ Å, $\beta = 116.7^\circ$.

Prior to crystallization, the Sr–EhCaBP2 complex was concentrated to 15 mg/ ml and crystallized with 40–50 % MPD as precipitant in 50 mM acetate buffer pH 4.6 and 5 mM SrCl₂ using the hanging drop vapour - diffusion method at 289 K. At higher concentrations of MPD, fast-growing plate-like poly-microcrystals of the Sr–EhCaBP2 complex were formed. These crystals were used for microseeding. The best diffracting crystals appeared in 10–15 d at 42 % MPD when the protein and precipitant were mixed in a 2.5:1 ratio; after microseeding, the concentration of precipitant was increased to 60 % [fig. 2.3.2 (c)]. The crystals grew to approximately 0.5 x 0.5 x 0.1 mm in about 7–10 d. The crystals were frozen in 65 % MPD, 10 mM cacodylate, 25 mM acetate pH 4.6, 5 mM SrCl₂. The crystals diffracted to 2.55 Å and belonged to space group P21, with unit-cell parameters $a = 69.18$, $b = 112.03$, $c = 93.42$ Å, $\beta = 92.8^\circ$. The data-collection statistics are summarized in (Table 2.3.4). Based on a Matthews coefficient calculation (Matthews, 1968), each asymmetric unit in this cell could contain 7–11 molecules (V_M is in the range 3.4–2.2 Å³ Da⁻¹), with a solvent content ranging from 64 to 43.4 %, respectively. The C-terminal domain of calmodulin from paramecium (PDB code 1exr; Wilson & Brunger, 2000) was used as model for molecular replacement and seven peaks were obtained using MOLREP (Vagin & Teplyakov, 1997).

This IQ1–EhCaBP2 complex was concentrated to 15 mg/ml and crystallized using 50 mM cacodylate buffer pH 6.4 with 30 % MPD, 30–45 % ethanol as precipitants. All these crystals were obtained at a temperature of 289 K [fig. 2.3.2 (d)]. These crystals were of

poor quality with layers and cracks; very few crystals were good-looking. The better looking crystals were transferred into cryoprotectant solutions containing mother liquor with an MPD concentration increased by 5 % from the crystallization condition prior to data collection. These crystals diffracted to 3.0 \AA and belonged to space group P21 with unit-cell parameters $a = 60.5$, $b = 69.86$, $c = 86.5 \text{ \AA}$, $\beta = 97.9^\circ$. The data collection statistics are summarized in (Table 2.3.5). Based on a Matthews coefficient calculation (Matthews, 1968), each asymmetric unit in this cell can contain three to five molecules (V_M in the range $3.8\text{--}2.3 \text{ \AA}^3 \text{ Da}^{-1}$), with a solvent content ranging from 67 to 45 % respectively.

CHAPTER 3

Homology Modeling of Seven EhCaBPs

3.1 Introduction

Ca^{2+} is a ubiquitous second messenger involved in signal transduction processes in eukaryotes (Berridge et al., 2000). It has also been documented that Ca^{2+} signaling plays a crucial role in the pathogenesis of many protozoan parasites including *Trypanosoma cruzi* (Yakubu et al., 1994), *Leishmania amazonensis* (Lu et al., 1997) and *Toxoplasma gondii* (Vieira et al., 2000). As early as 1992 Scheibel et al. did some profound work related to calcium binding proteins and identified a number of CaBPs in *E. histolytica*. Among these are two related EF-hand-containing proteins, granin 1 and granin 2, which were characterized and found to be localized in intracellular granules (Nickel et al., 2000). And then some propositions for these proteins were put forward that they might be involved in phagocytosis, control of endocytotic pathways and 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^{2+} dependent DNA binding properties (Gilchrist et al., 2003). On extensive analysis of *Entamoeba histolytica* genome sequence it was revealed that there are 27 CaBPs, which have two or more EF hand motifs [(Bhattacharya et al., 2006), for the table of EhCaBPs refer to chapter 1, An Overview of *Entamoeba histolytica* Calcium Binding Protein)]. CaBP1 is well characterized protein among them, and about nine CaBPs are of similar length, which have about four EF hand motifs.

CaBP1 (EhCaBP1) from *E. histolytica* had been previously characterized (Prasad et al., 1993; Sahoo et al., 2004). EhCaBP1 has 134 amino acid residues with four canonical EF-hand Ca^{2+} binding domains. Though this protein has some structural similarity with CaM, it is functionally quite distinct (Yadava et al., 1997). Inducible expression of EhCaBP1 antisense RNA demonstrated its role in cellular proliferation (Sahoo et al., 2004). According to NMR structure EhCaBP1 and EhCaBP2 are structurally related to CaM and troponin C (TnC) in spite of low sequence homology with these proteins. (Atreya et al.,

2001). While crystal structure showed complete different arrangement of EF hand motifs in EhCaBP1 (Kumar et al., 2007). Among the nine CaBPs, which have similar size and putative EF hand motifs, two structures have been reported. It becomes interesting to unravel the structure of the seven EhCaBPs in the light of the EhCaBP1 crystal structure which states a novel rearrangement of EF hand motifs. With this view in mind models for seven EhCaBPs which are having similar size were generated with available homologues models.

3.2 Methods

Database search and sequence alignment was done for identifying templates. The single letter amino acid sequences for the seven EhCaBPs were downloaded from NCBI GenBank (<http://www.ncbi.nlm.nih.gov/sites/entrez>) as indicated in Table 3.5.1. A PSI-BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) (Altschul et al., 1990) search was performed using seven EhCaBPs sequences as query against the PDB database (<http://www.rcsb.org/pdb/Welcome.do>) (Berman et al., 2000). After removing redundant sequences, it was found that each search resulted in significant hit. And for each of CaBP (*Entamoeba histolytica*) a separate template was selected for modeling of seven EhCaBPs. The ClustalX program (www.embl.de/chenna/clustal) was used for sequence alignment (Ramu et al., 2003). The Bioedit program (<http://www.mbio.ncsu.edu/BioEdit>) was used to represent conservation in sequence alignment. Solvent accessibility chart was made with the programme <http://gibk26.bse.kyutech.ac.jp/shandar/netasa/asaview/> with default settings (Shandar et al., 2004). Figures were made with PyMol (Delano, 2002).

3.2.1 Homology modelling

The three-dimensional (3D) structures of seven EhCaBPs were modelled in a stepwise procedure, using MODELLER version 9v1. This software implements homology modelling of proteins by satisfaction of spatial restraints (Sali et al., 1993). 100 models for each of the seven members were generated. Bond and angle values are taken from CHARMM-22 force field. 3D models were generated by molecular probability density function optimisation. The X-ray crystallographic structure of template, was used as the template for modelling seven EhCaBPs. Among all seven EhCaBPs, only couple of members has cysteines. ModLoop (<http://alto.compbio.ucsf.edu/modloop>) (András et al., 2003; Fiser et al., 2000) was used to model the bigger and problematic loops carefully as these were showing maximum root mean square deviation (RMSD).

3.2.2 Model validation

Evaluation of the models of seven EhCaBPs was carried out using the programmes Prosa-web (<https://prosa.services.came.sbg.ac.at/prosa.php>) (Wiederstein et al., 2007; Sippl et al. 1993], PROCHECK (<http://nihserver.mbi.ucla.edu/SAVS/>) (Laskowski et al., 1993) and WHATIF (<http://swift.cmbi.kun.nl/WIWWWI/>) (Vriend et al., 1990;, Hooft et al., 1996). Only those models that showed a satisfactory PROSA, PROCHECK and WHATIF profile were selected. A 3D structural superimposition to compare the structures was done using the program STAMP, which is a part of VMD version 1.8.4 (Humphrey et al., 1996).

3.2.3 Phylogenetic analysis

To analyse EhCaBP protein sequences from an evolutionary perspective the neighbour-joining (NJ) tree was constructed using the programme of <http://www.ebi.ac.uk/Tools/es/cgi-bin/clustalw2> with default settings.

3.3 Results and Discussion

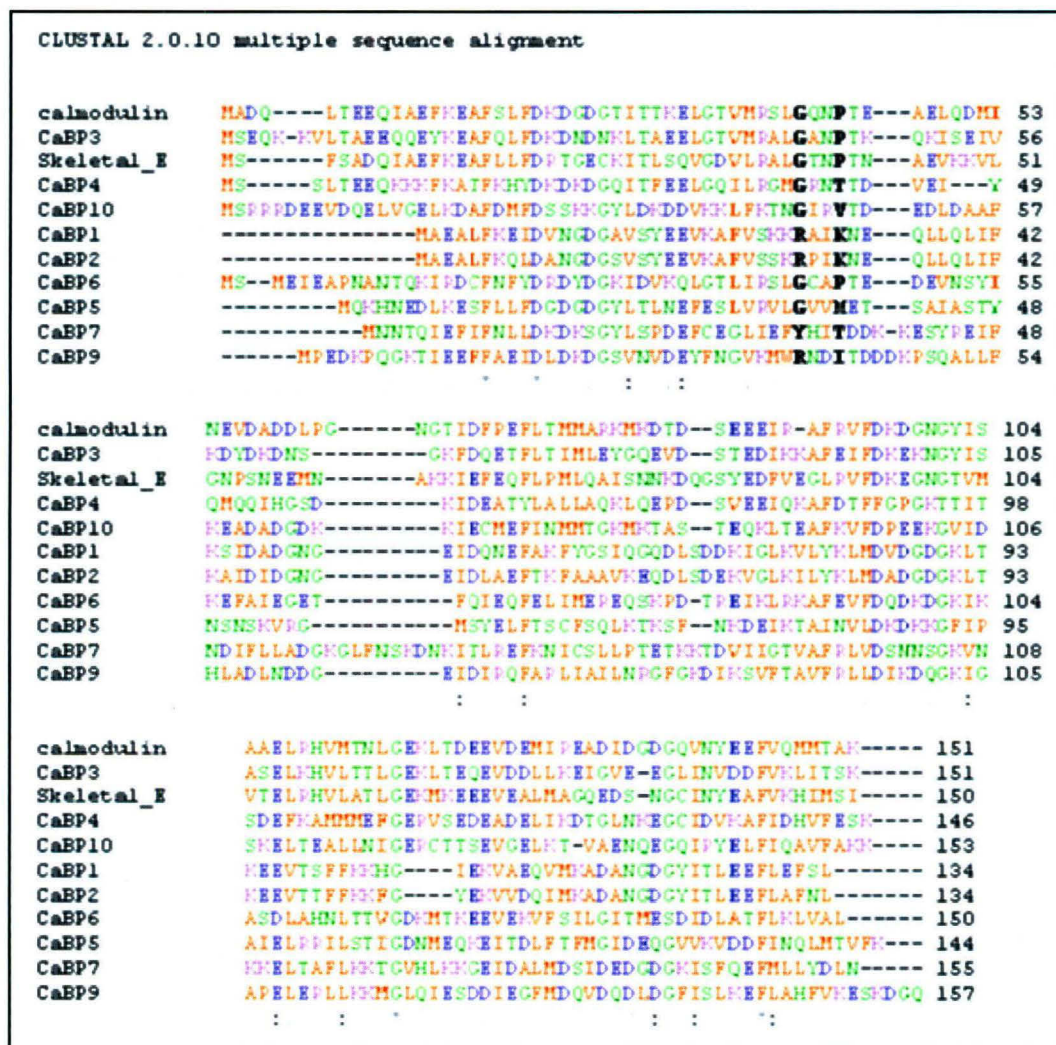
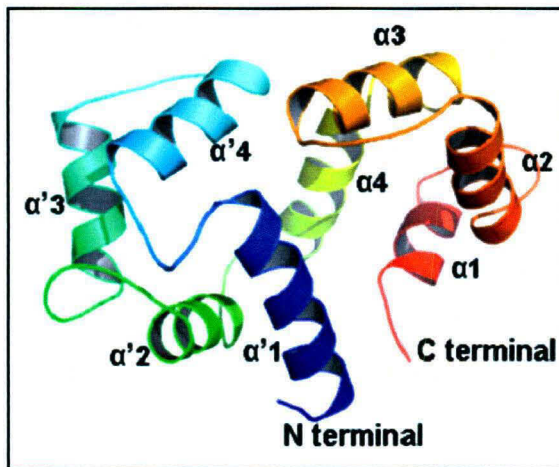
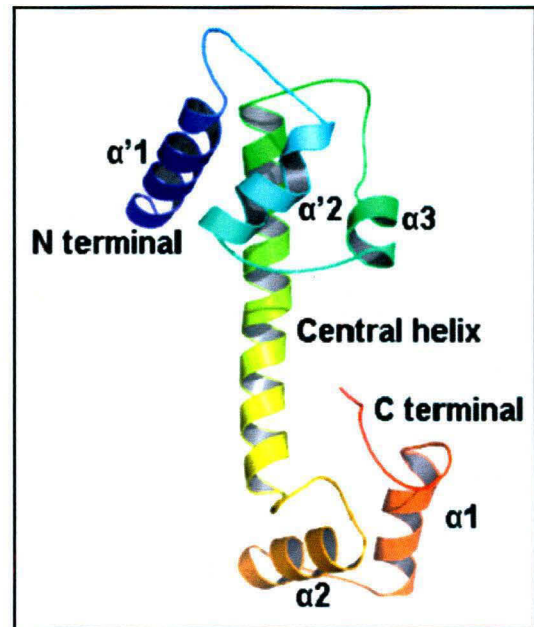


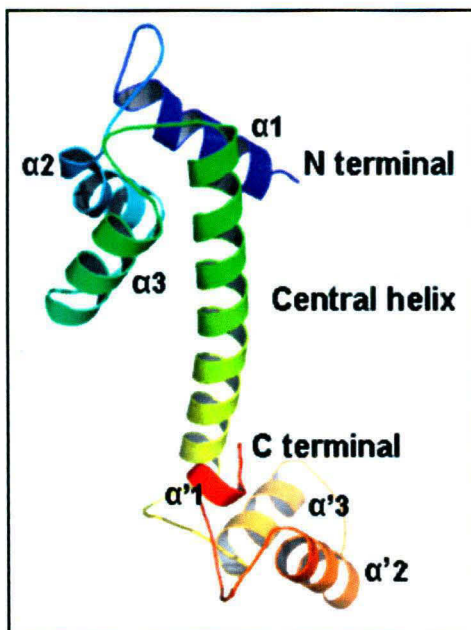
Fig. 3.3.1 Multiple sequence alignment of nine sequences belonging to the EhCaBP family with Calmodulin and Skeletal Essential Light Chain. * represent identical amino acids, : represent Conserved substitutions, . represent semiconserved substitutions.



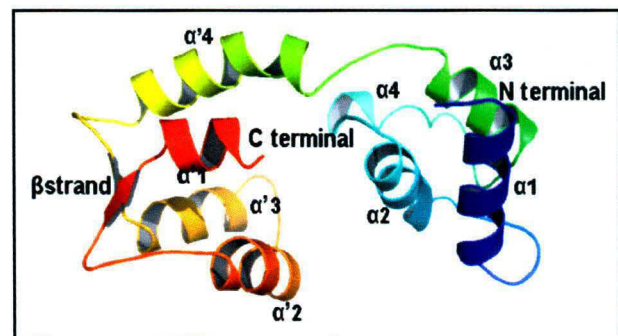
(a) CaBP3



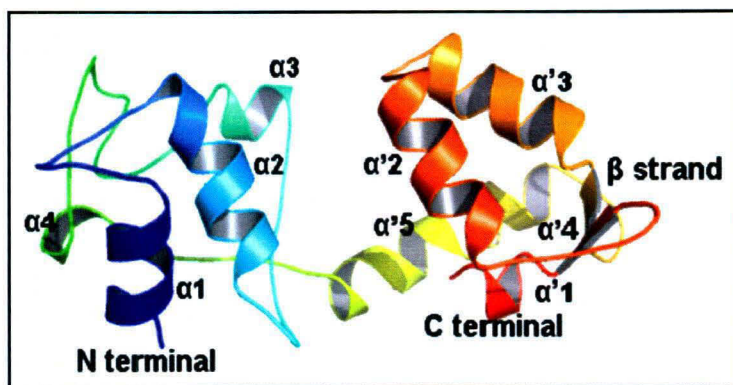
(b) CaBP4



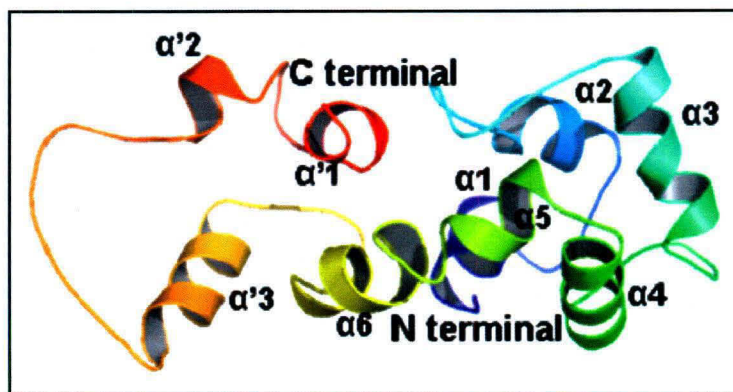
(c) CaBP5



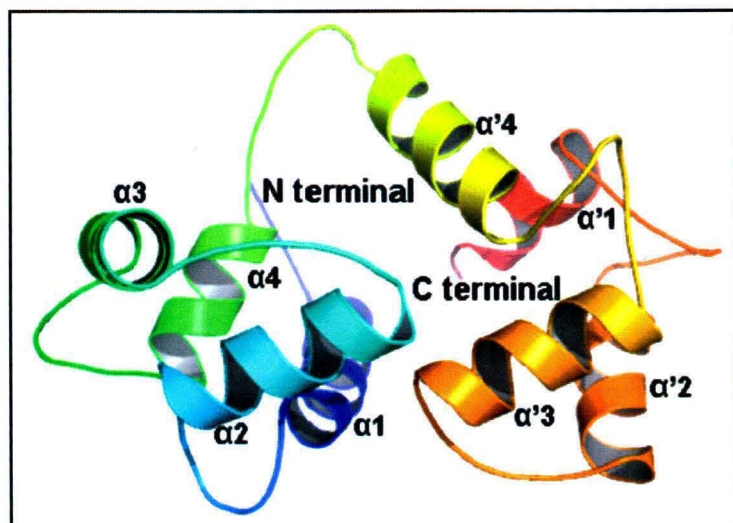
(d) CaBP6



(e) CaBP7

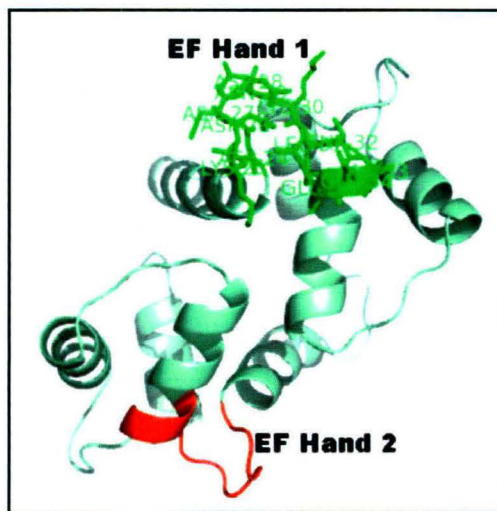


(f) CaBP9

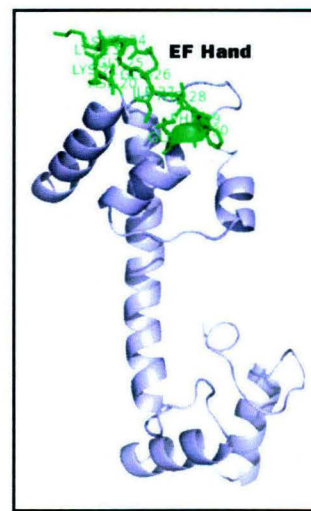


(g) CaBP10

Fig. 3.3.2 Cartoon representations of models of seven EhCaBPs. The entire helices and N& C terminal has been labeled.



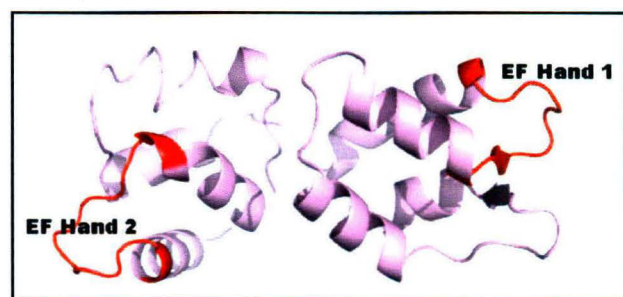
(a) CaBP3



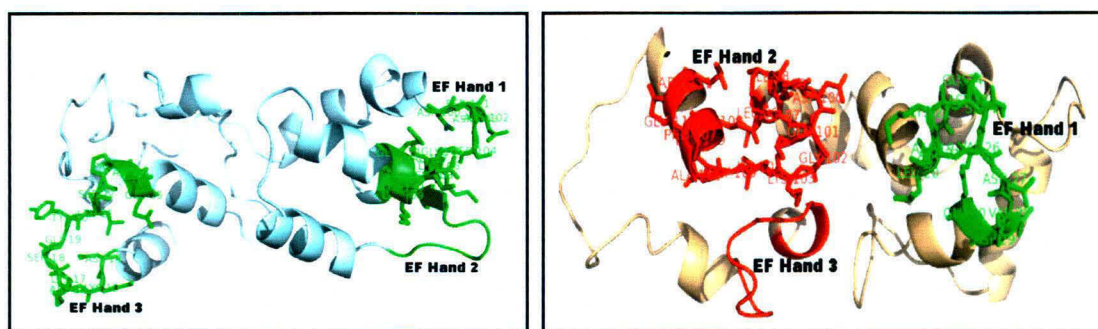
(b) CaBP4



(c) CaBP5

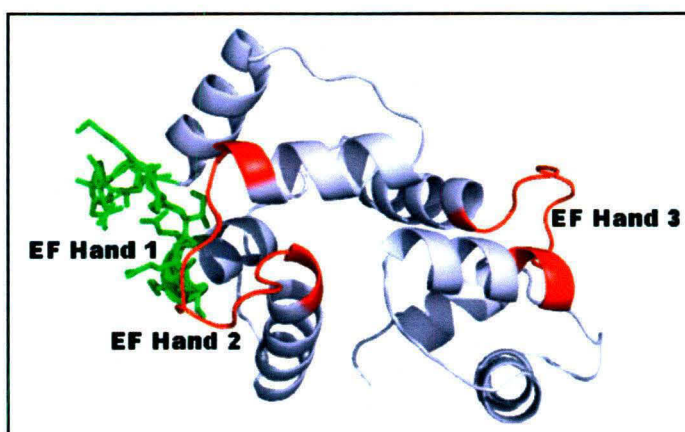


(d) CaBP6



(e) CaBP7

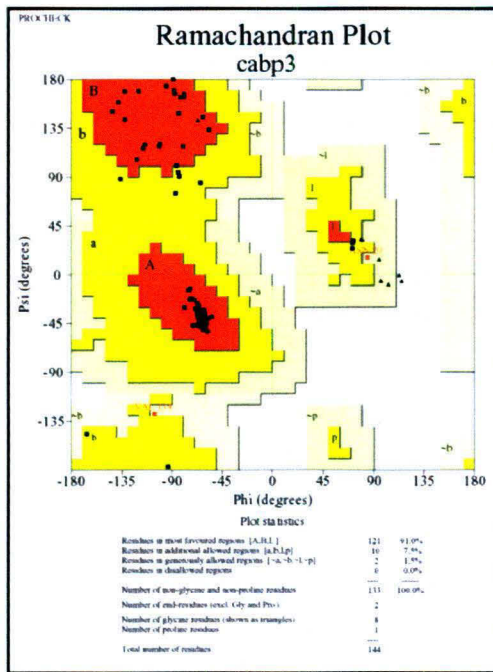
(f) CaBP9



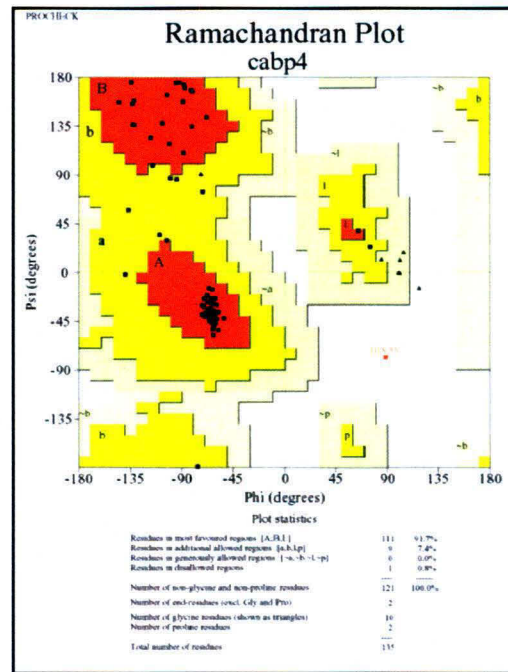
(g) CaBP10

Fig. 3.3.3 Cartoon representations of models of seven EhCaBPs with labelled EF hands (a), (b), (c), (d), (e), (f), (g). All the EhCaBP protein models show labeled EF Hands (with side chain) in green colour and red colour.

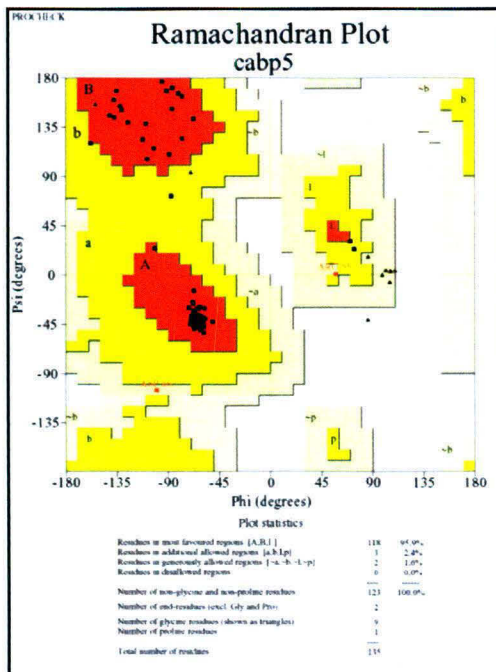
Model validation



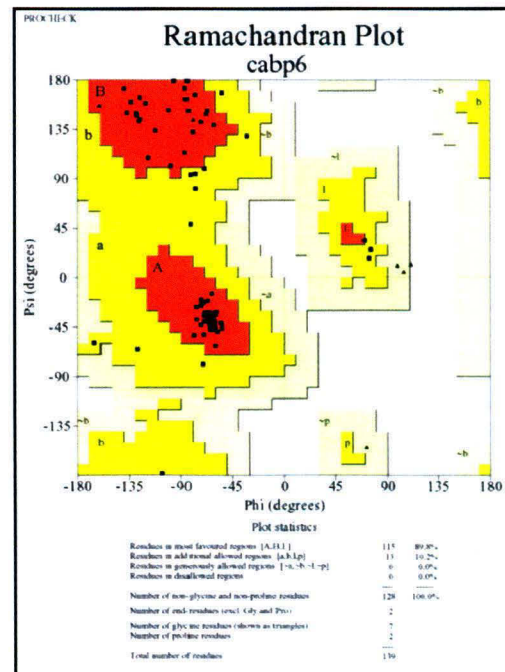
(a)



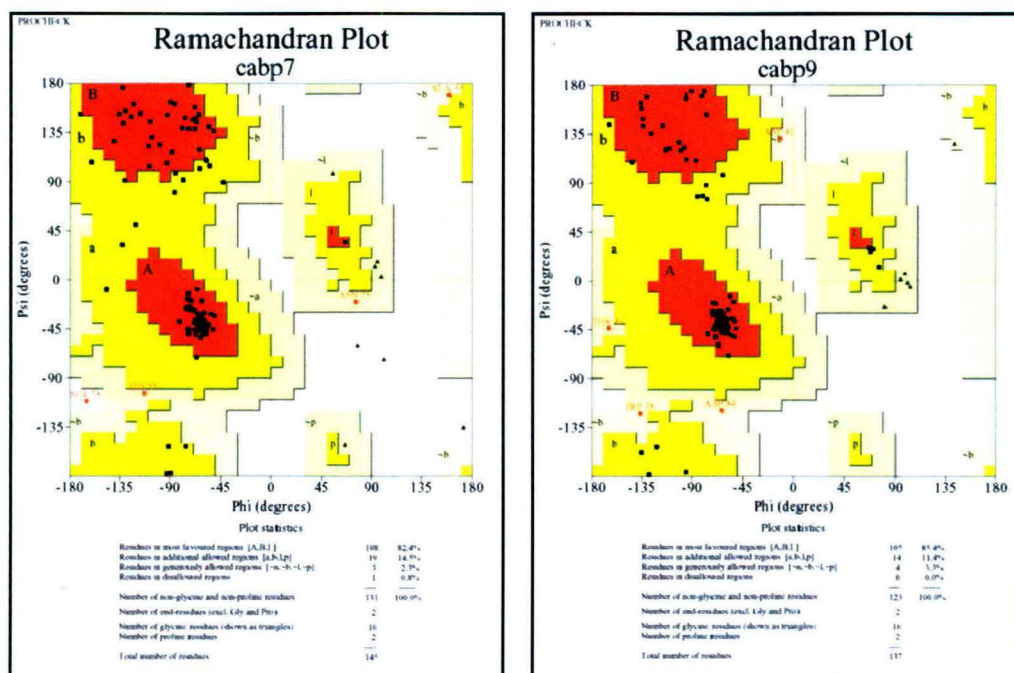
(b)



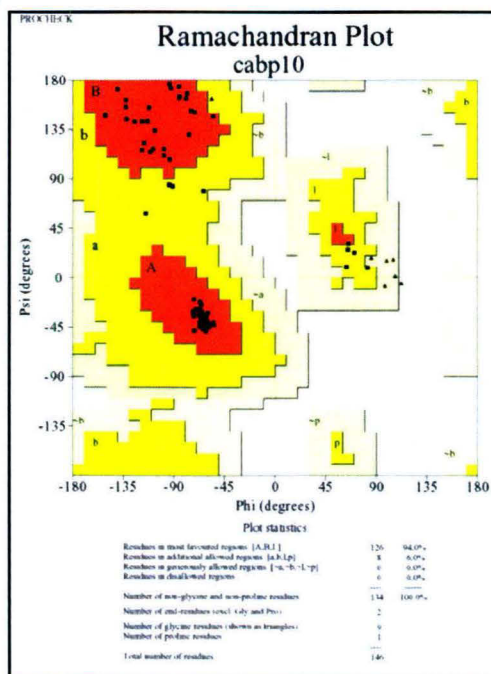
(c)



(d)



(e)



(f)

(g)

Fig. 3.3.4 Ramachandran plots for the models of seven EhCaBPs proteins (a), (b), (c), (d), (e), (f), (g). The percentage of residues in the most favoured, allowed, and generously allowed regions are shown below the Ramachandran plots.

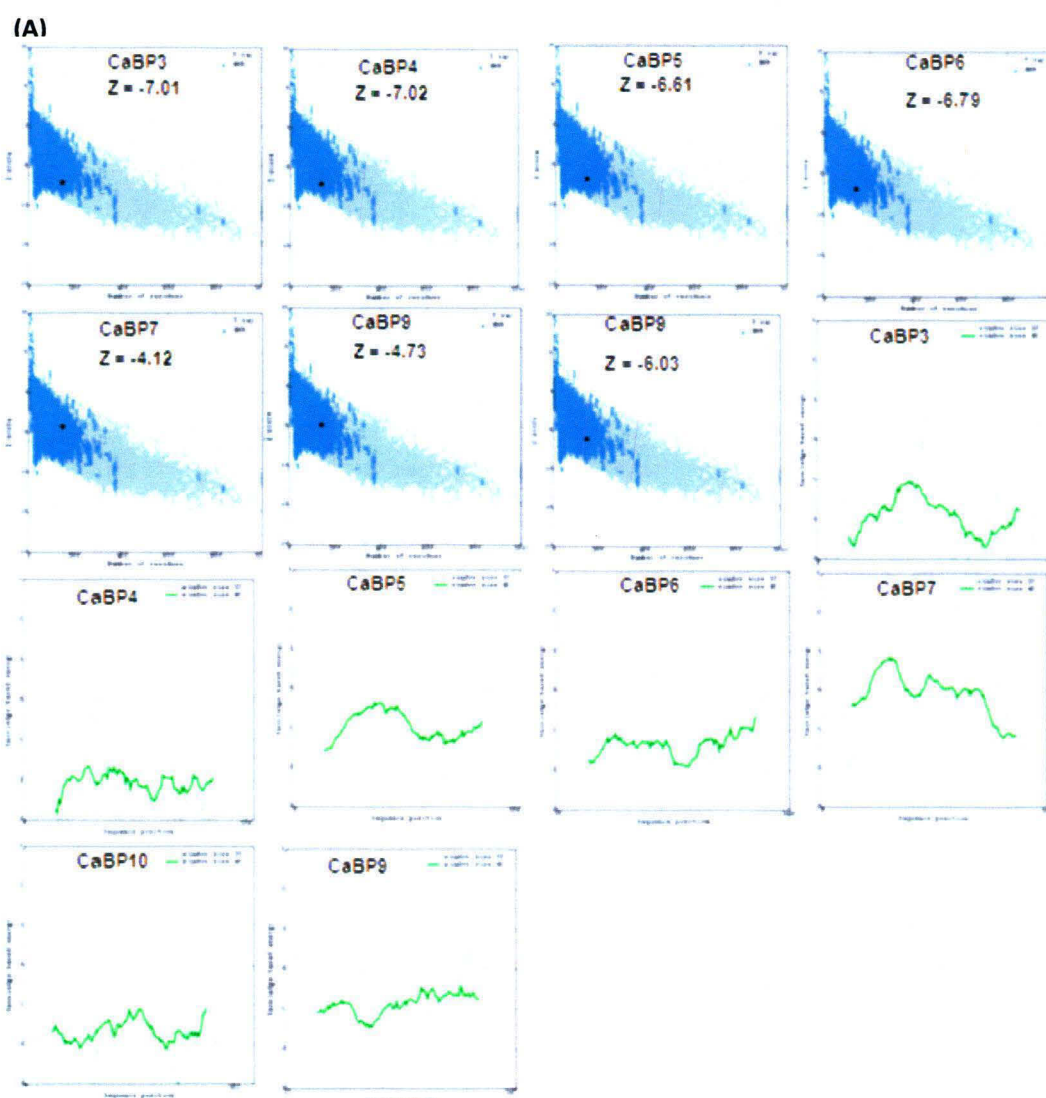
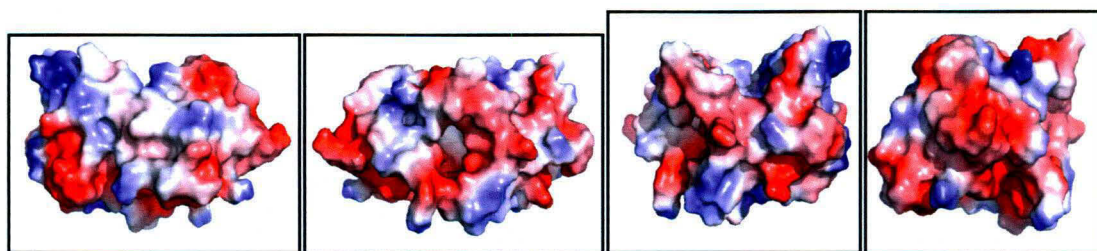
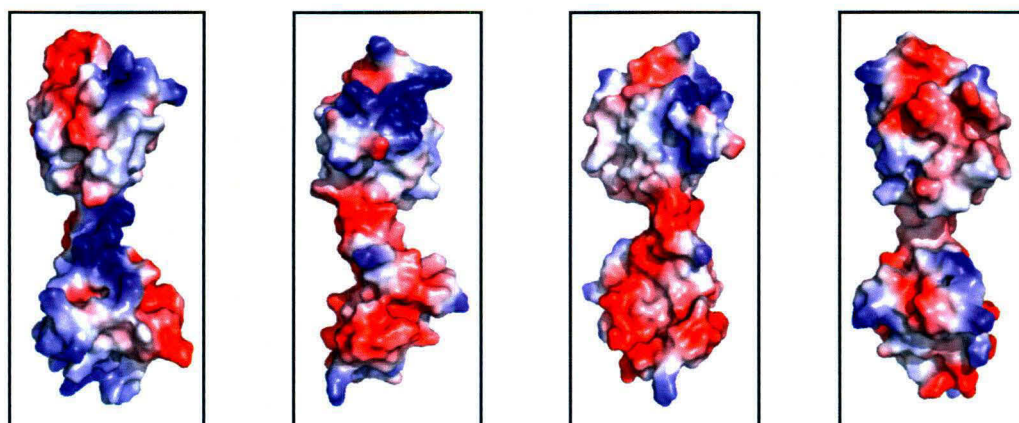


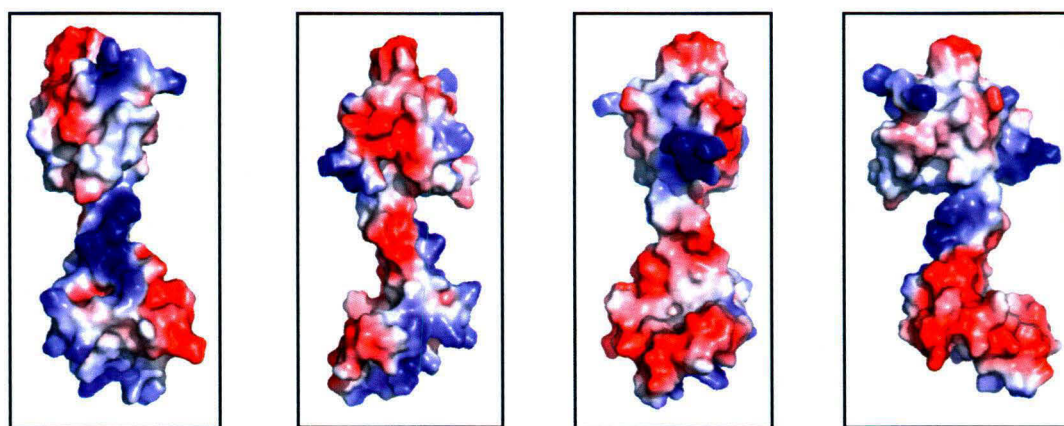
Fig. 3.3.5 ProsaWeb analysis of theoretical 3D models of seven EhCaBPs. (A) Z - Score plot of the EhCaBP obtained; the z-scores of all the seven models lie within the normal range. (B) Energy plots for the seven EhCaBP models; the energy of residues for all the models developed is close to zero or negative indicating the stability of the models.

Electrostatic Representation

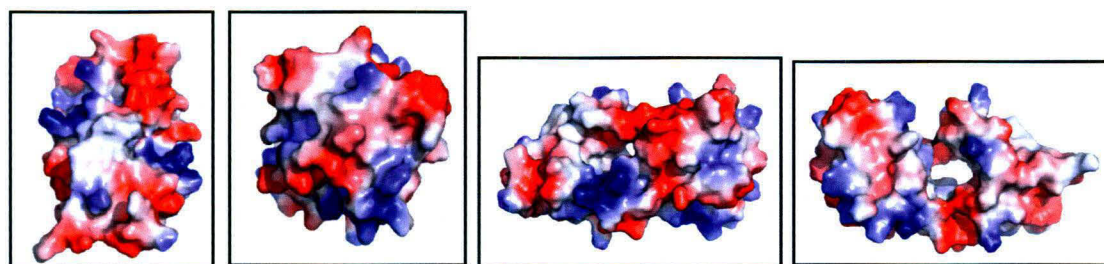
(a) CaBP3 dorsal side CaBP3 ventral side CaBP3 3 side CaBP3 4 side



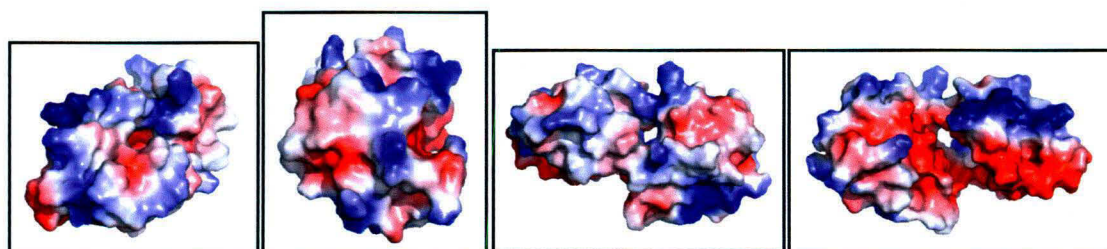
(b) CaBP4 dorsal side CaBP4 ventral side CaBP4 3 side CaBP4 4 side



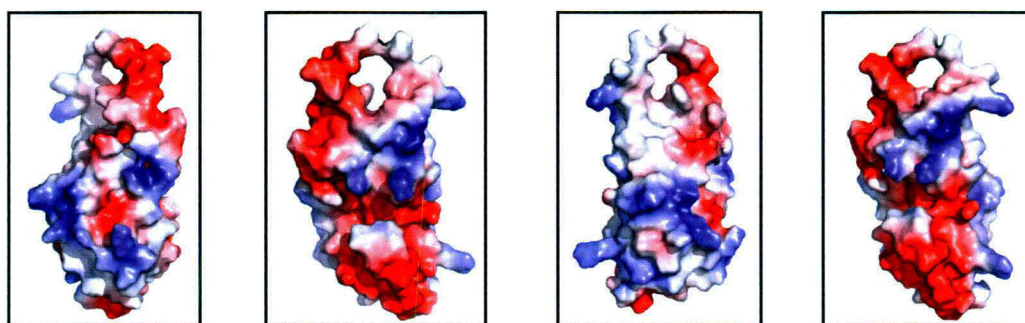
(c) CaBP5 dorsal side CaBP5 ventral side CaBP5 3 side CaBP5 4 side



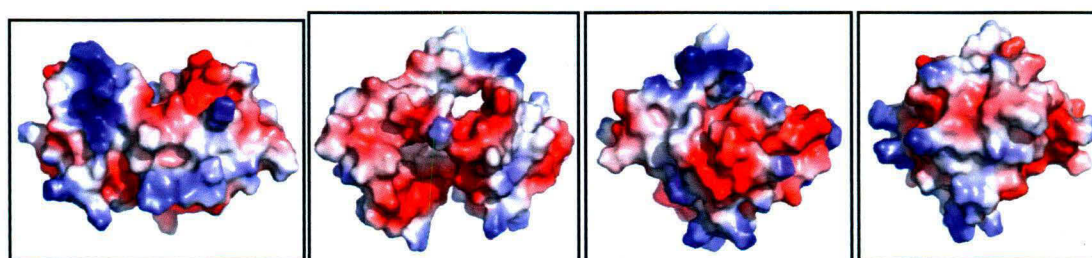
(d) CaBP6 dorsal side CaBP6 ventral side CaBP6 3 side CaBP6 4 side



(e) CaBP7 dorsal side CaBP7 ventral side CaBP7 3 side CaBP7 4 side

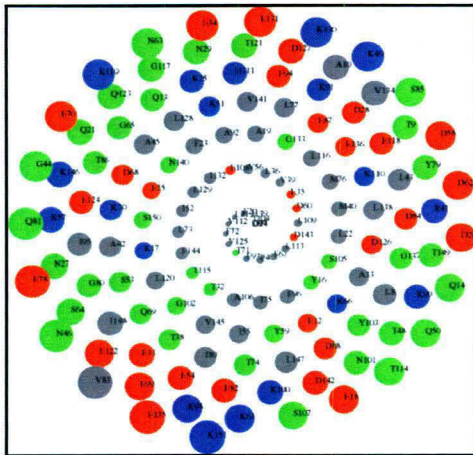


(f) CaBP9 dorsal side CaBP9 ventral side CaBP9 3 side CaBP9 4 side

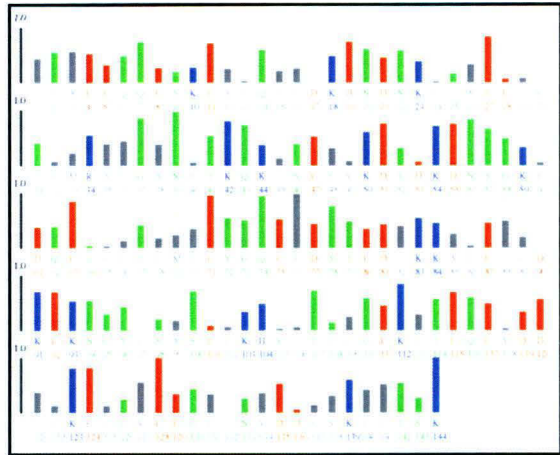


(g) CaBP10 dorsal side CaBP10 ventral side CaBP10 3 side CaBP10 4 side

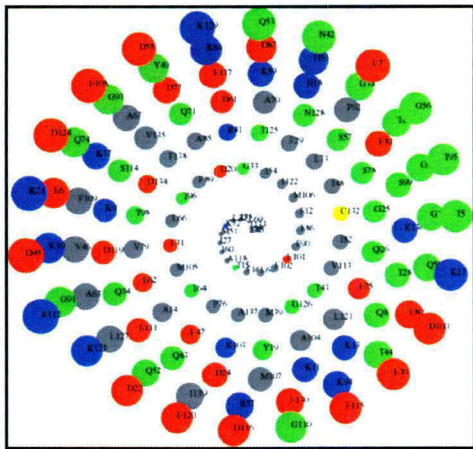
Fig. 3.3.6 Electrostatic molecular surface representation of the models of seven EhCaBPs. – first one - Front view of EhCaBP proteins, the second one - the opposite face after 180° rotation, third and fourth - view of each edge of the molecule.



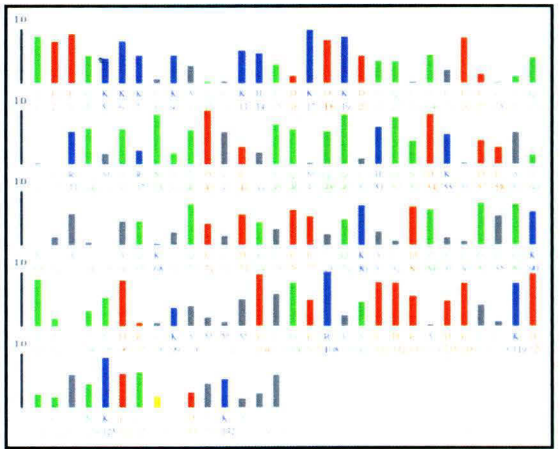
(a) (i) CaBP3



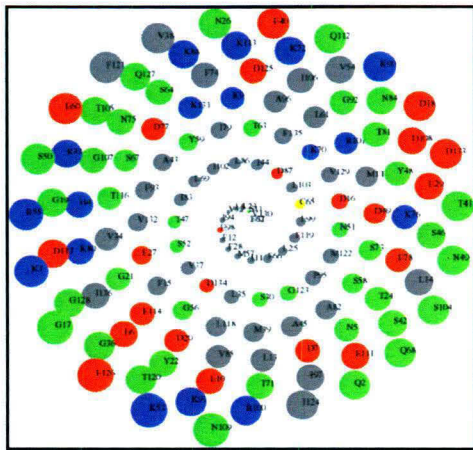
(ii) CaBP3



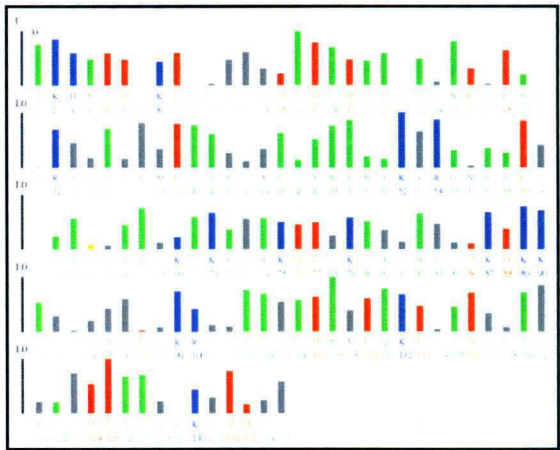
(b) (i) CaBP4



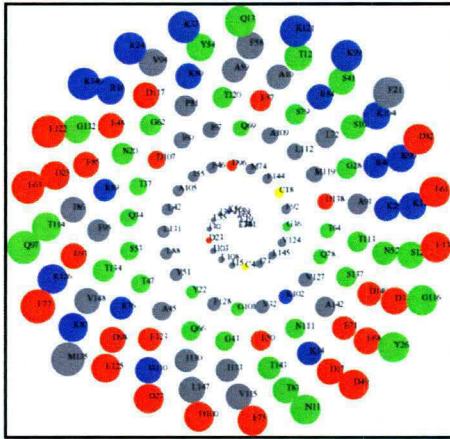
(ii) CaBP4



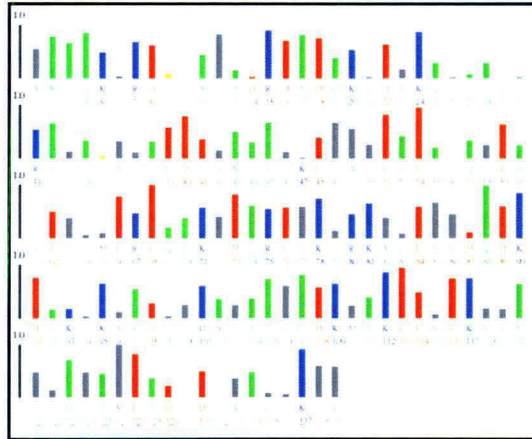
(c) (i) CaBP5



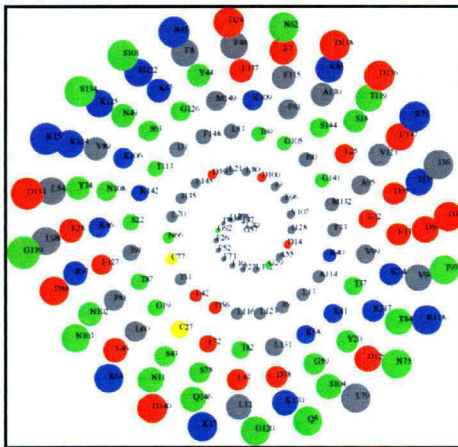
(ii) CaBP5



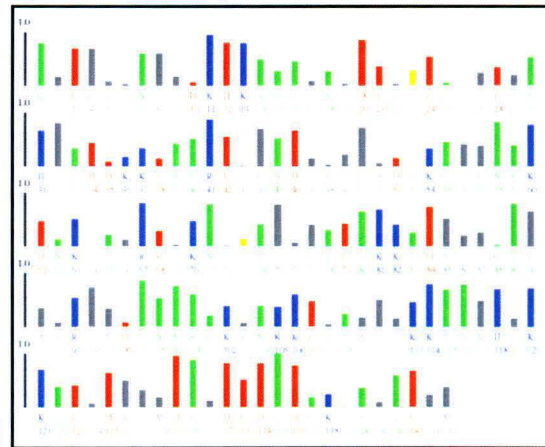
(d) (i) CaBP6



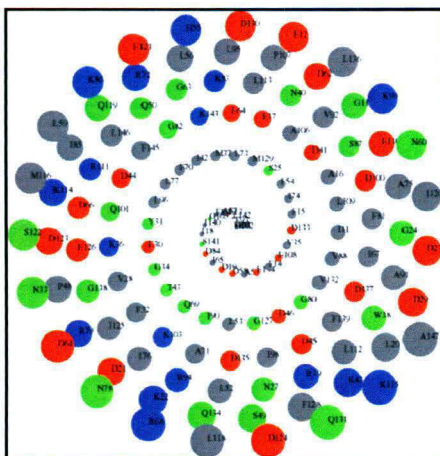
(ii) CaBP6



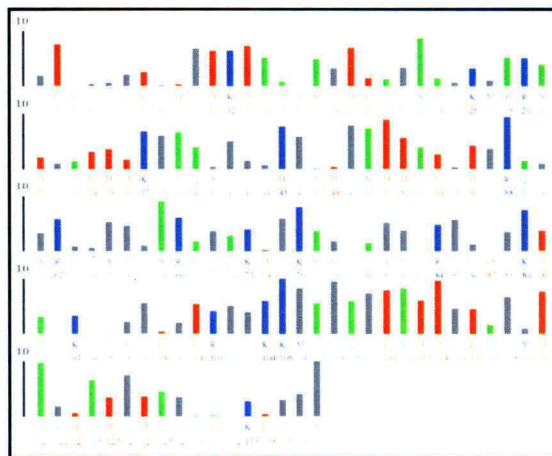
(e) (i) CaBP7



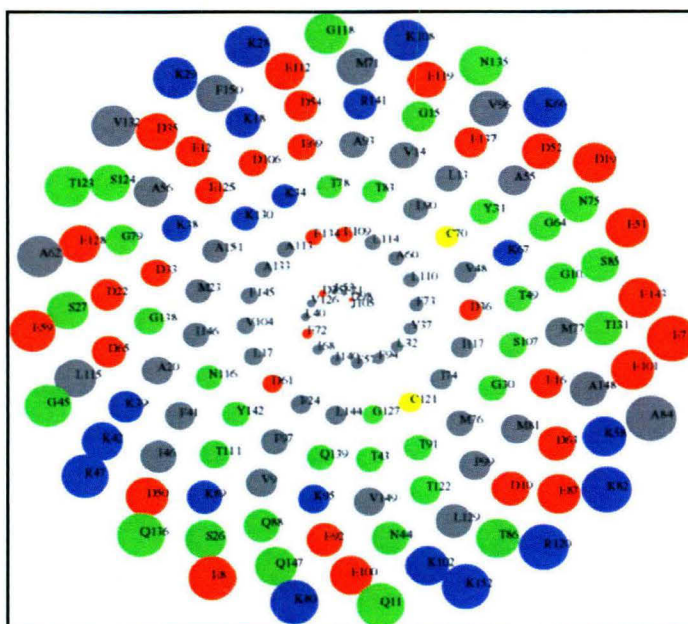
(ii) CaBP7



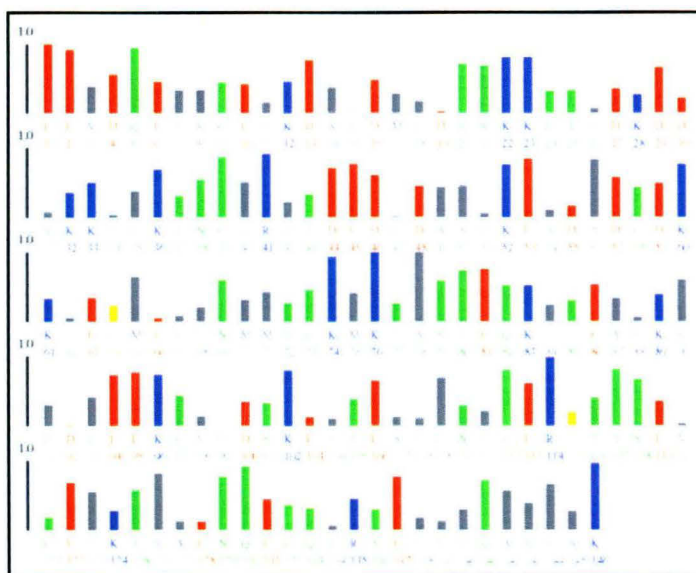
(f) (i) CaBP9



(ii) CaBP9



(g) (i) CaBP10



(ii) CaBP10

Fig. 3.3.7 Graphical representation of solvent accessibility of seven EhCaBPs (a), (b), (c), (d), (e), (f), (g). (i) The spiral plot is a plot of solvent accessibility of residues, arranged in the order of their relative accessible surface area. Radii of the spheres represent relative solvent accessibility of the residue. **(ii)** Length of the bars represents relative solvent accessibility of the residue.

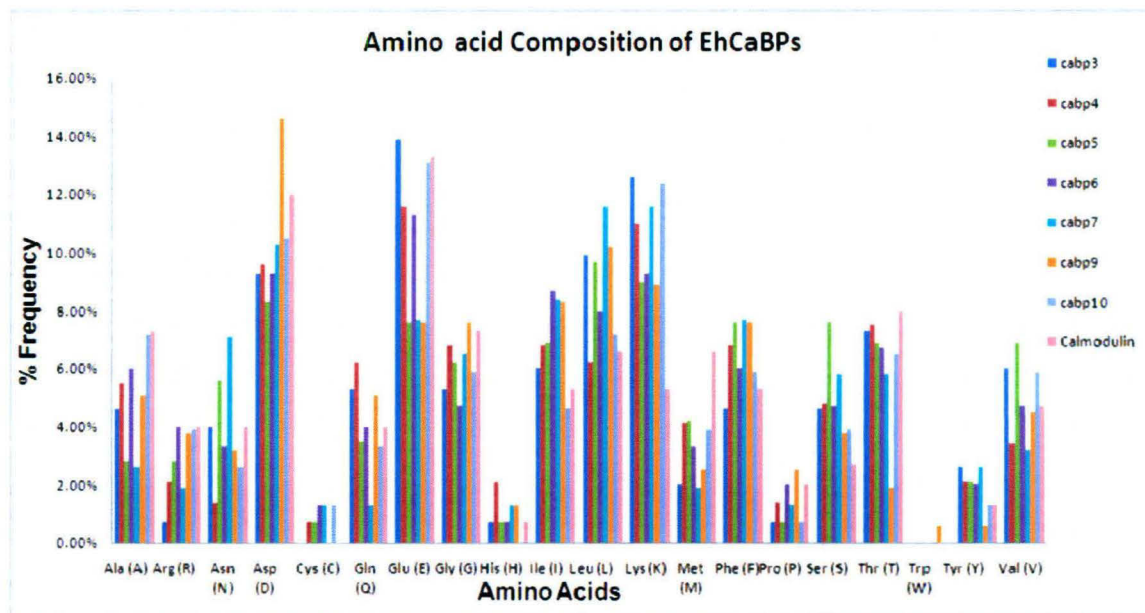


Fig. 3.3.8 (a) Composition of amino acids in the seven EhCaBPs and Calmodulin.

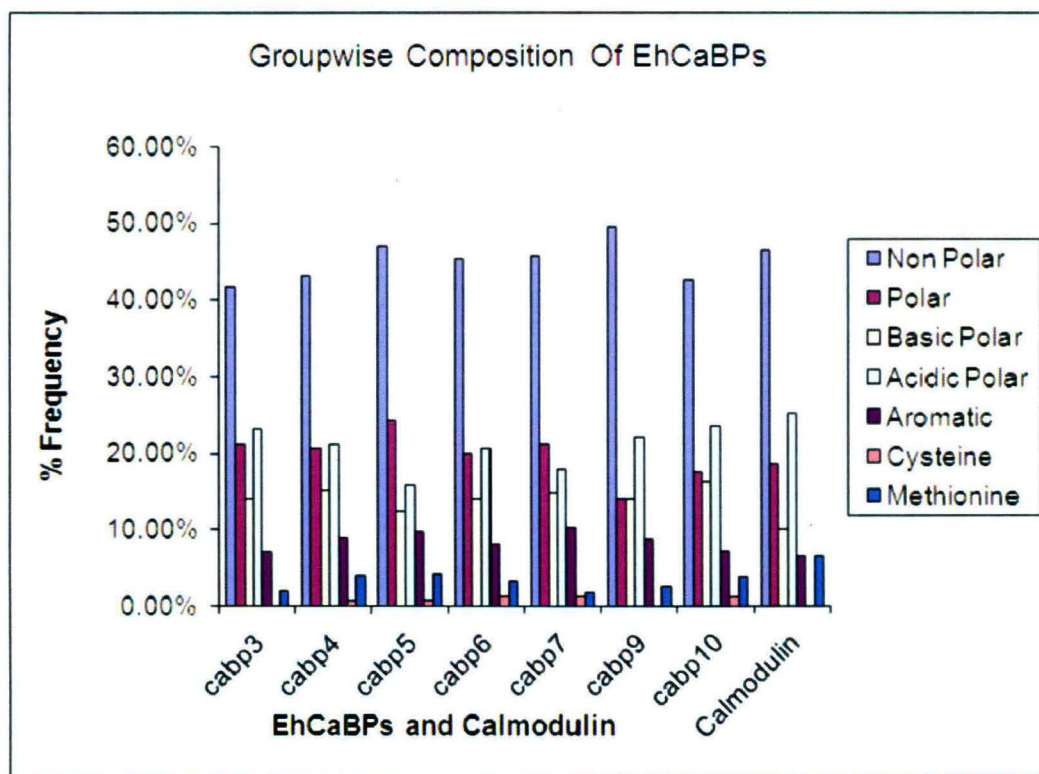
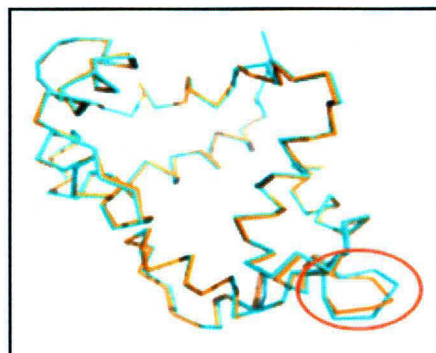


Fig. 3.3.8 (b) Groupwise Composition of amino acids in the seven EhCaBPs & Cam.

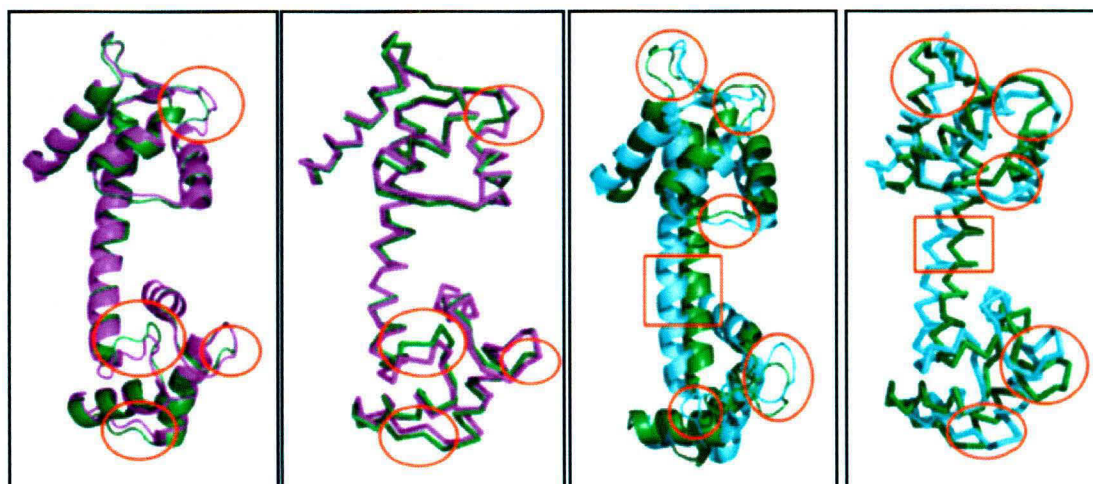


(a) (i) CaBP3

(ii) CaBP3

Template colour - Cyan

Model colour - Orange



(b) (i) CaBP4

(ii) CaBP4

(c) (i) CaBP5

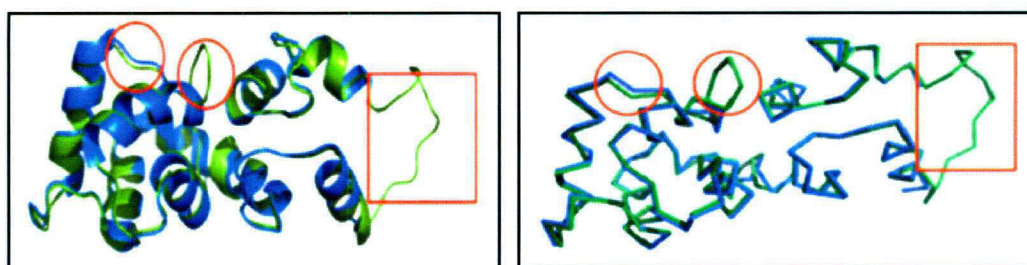
(ii) CaBP5

Template colour - Magenta

Model colour - Forest green

Template colour - Cyan

Model colour - Forest green

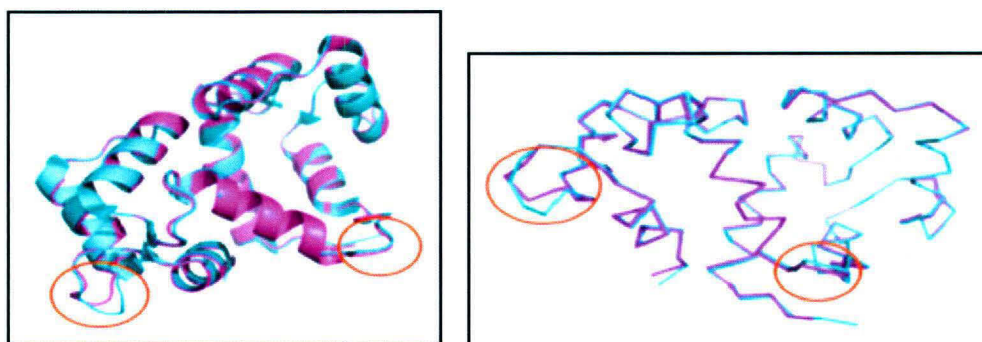


(d) (i) CaBP9

(d) (ii) CaBP9

Template colour - Marine blue

Model colour - Forest green



(e) (i) CaBP 10

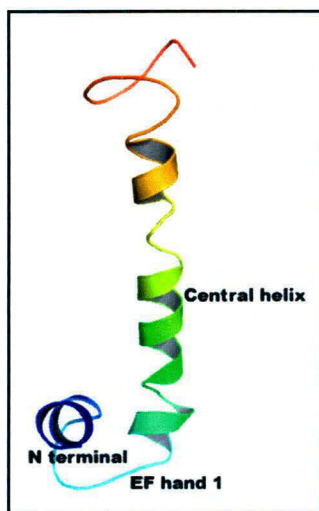
(e) (ii) CaBP 10

Template colour - Cyan

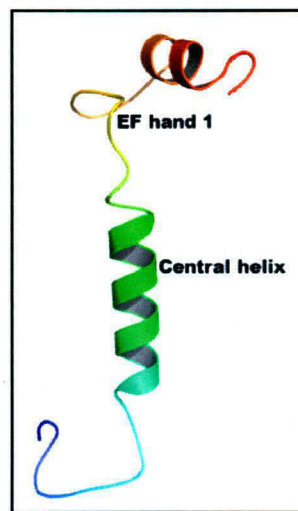
Model colour - Magneta

Fig. 3.3.9 Superimposed structures of theoretical 3D models of the five EhCaBP proteins with (C-alpha atoms only) on their respective templates in (i) cartoon representation (ii) ribbon representation (a), (b), (c), (d), (e). The encircled area shows the difference between the template and the model. Template colour and model colour is given below each figure.

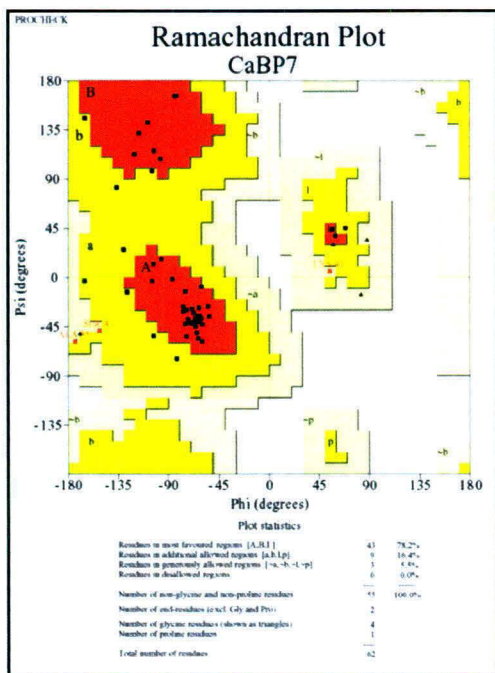
Alternative Models for CaBP7 and CaBP9



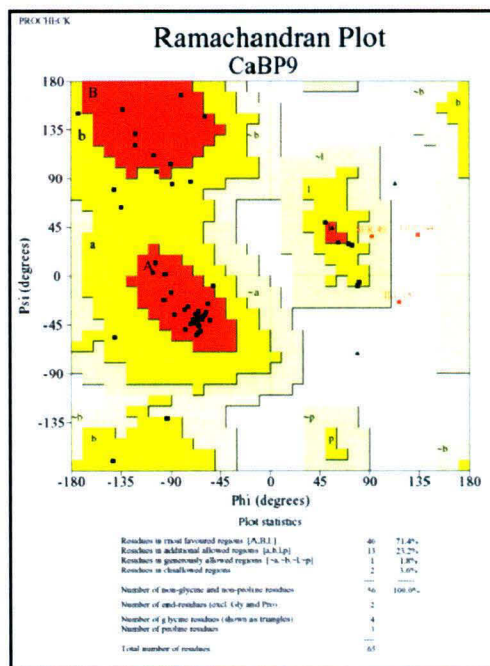
(a) CaBP7



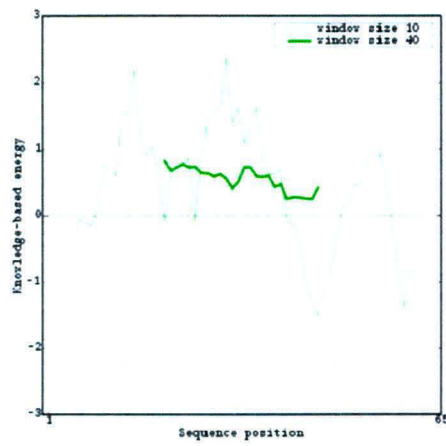
(b) CaBP9



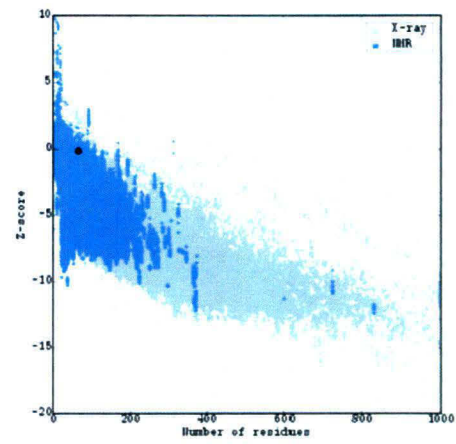
(c)



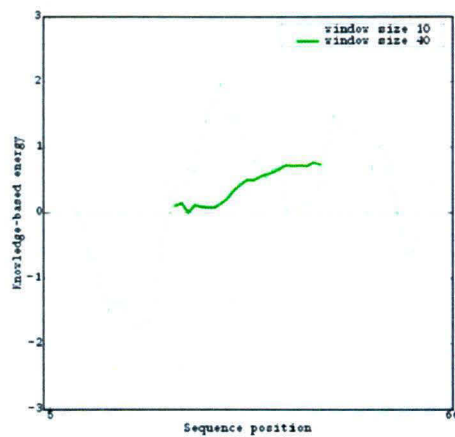
(d)



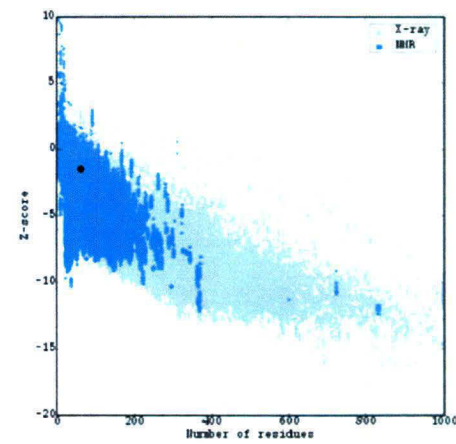
(e) CaBP9



(f) CaBP9 Z = -0.09



(g) CaBP7



(h) CaBP7 Z = -1.51

Fig. 3.3.10 An alternative model showing structural part (a) CaBP7 (b) CaBP9 , (c) & (d) – Ramachandron plot ; (e), (g), (f), (h) – Z score and energy plots respectively.

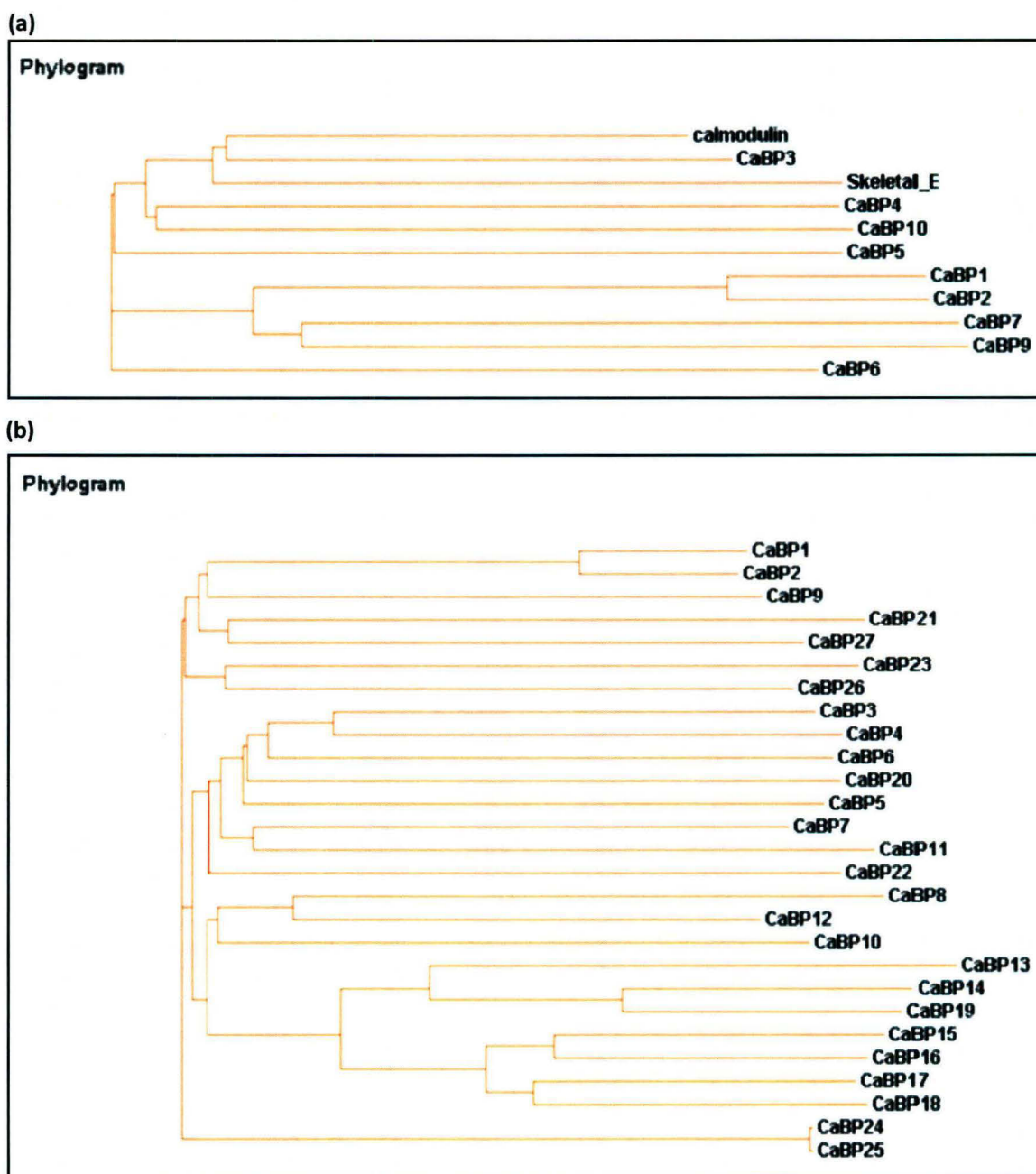


Fig. 3.3.11 Phylogenetic relationship among (a) nine EhCaBPs of *Entamoeba histolytica* and Calmodulin and Skeletal Essential Light Chain. (b) 27 EhCaBPs of *Entamoeba histolytica*. This neighbour joining (NJ) tree was constructed using the programme of <http://www.ebi.ac.uk/Tools/es/cgi-bin/clustalw2> with default settings.

3.4.1 CaBP family (*E. histolytica*) and Sequence Conservation

On the multiple sequence alignment of all 27 CaBPs of *Entamoeba histolytica*, it was found that there is not a single amino acid that has been conserved throughout the evolution of CaBPs in *Entamoeba histolytica*. Reasons for this could be attributed to the fact that location of the EF hand and the length of the protein vary a lot across the 27 EhCaBPs. So when small number and similar length of EhCaBPs were aligned the conservation of amino acid could be seen. On aligning nine EhCaBPs of *Entamoeba histolytica*, which are smaller and contain around 4 putative EF hand motifs, with well characterized calmodulin and Skeletal essential light chain (fig.3.3.1), a few conservations of amino acids is seen, these residues have been highlighted with asterisk mark. The residue positions are Phe21st, Asp25th, Phe145th, Gly117th (equivalent position in the alignment) are conserved (fig.3.3.1). Asp25th is prominent feature of almost all calcium binding EF hands loops. Since aspartate takes part in binding with the calcium ion so it is generally conserved in almost all calcium binding proteins (Jessica et al., 2007). The other residues which are conserved are Phe21st and Phe145th. These aromatic residues are located in the N-terminal and C-terminal halves. In almost all the nine EhCaBPs (the N-terminal Phe21st is surrounded by other adjacent Phe and in the same way the situation is replicated for the C-terminal phenylalanine (Phe145th). This situation is somewhat similar to the Calmodulin (Babu et al., 1988) where there is a formation of cluster of aromatic amino acids in both N- terminal and C-terminal halves. The other conserved residue is Glycine (Gly117th). In Calmodulin (Babu et al., 1988) out of the 10 glycine residues, 8 residues are part of the four calcium binding loops and it offers flexibility. In the same way in case of seven members of EhCaBPs glycine could be offering flexibility to the calcium binding loops. Even at position (equivalent in the alignment) 32nd, 104th there is either isoleucine or valine, which also fall in EF hand motif is seen even in the Calmodulin and Skeletal Essential light chain also (fig.3.3.1). Both these amino acids are non polar in nature. There could be a possibility that these non polar residues might be interacting with the targets. Overall there is good amount of divergence in the sequence is seen among the nine EhCaBPs except for the some of the critical places like calcium binding domain and other residues responsible for the interaction with the target.

3.4.2 Structural description of Seven EhCaBPs models

EhCaBPs Models

In the table 3.5.1, details of the template and CaBP model have been provided. For each of CaBP model separate template was selected, according to the sequence identity. Though some of the CaBP models have been modeled on the template which has less sequence identity with the CaBP model sequence yet it stands valid as observed in other cases for example *Salmonella* OASS (Burkhard et al., 1998) which is less than 20% identical to tryptophan synthase β but overall fold of both of the proteins is totally identical.

Overall Structure

(a) CaBP3

The model of CaBP3 consists of 8 alpha helices but there is no beta sheet formation. It is composed of N-terminal and C-terminal halves which are shown in blue and red colour [fig. 3.3.2 (a)] respectively. The C-terminal is composed of α_1 , α_2 , α_3 , α_4 and N-terminal is of $\alpha'1$, $\alpha'2$, $\alpha'3$, $\alpha'4$. Both N-terminal half and C-terminal half are linked through a helix. There are two EF hands. EF hand 1 is located in the N-terminal half and the EF hand 2 is in C-terminal half.

(b) CaBP4

The model of CaBP 4 is dumbbell in shape but is composed of 6 alpha helices and there is no beta strand or beta sheet formation. It is composed of N-terminal and C-terminal halves which are in blue and red colour [fig. 3.3.2 (b)] respectively. The C-terminal is composed of α_1 , α_2 and N-terminal is of $\alpha'1$, $\alpha'2$, $\alpha'3$. Both N-terminal half and C-terminal half are linked through a central linker alpha helix. There is only one EF hand. EF hand is located in the N-terminal half. The central linker helix is quite flexible in nature, very similar to Calmodulin (Babu et al., 1988) and it might be helpful in the conferring flexibility to CaBP4 in recognizing diverse downstream targets.

(c) CaBP5

The model of CaBP 4 is also having dumbbell shape and is composed of 7 alpha helices and there is no beta strand or beta sheet formation. It is composed of N-terminal and C-terminal halves which are in blue and red color [fig. 3.3.2 (c)] respectively. The N-terminal is composed of α_1 , α_2 , α_3 and C-terminal is of α'_1 , α'_2 , α'_3 . Both N-terminal half and C-terminal half are linked through a central linker alpha helix. There are two EF hands. EF hand 1 is located in the N-terminal half and the EF hand 2 is in C-terminal half. The central linker helix is quiet flexible in nature, very similar to Calmodulin (Babu et al., 1988) and just like CaBP4 it might be helpful in the conferring flexibility to CaBP5 in recognizing diverse downstream targets.

(d) CaBP6

The model of CaBP6 consists of 8 alpha helices and there are two beta strands (antiparallel). It is composed of N-terminal and C-terminal halves which are in blue and red colour [fig. 3.3.2 (d)] respectively. The N-terminal is composed of α_1 , α_2 , α_3 , α_4 and C-terminal is of α'_1 , α'_2 , α'_3 , α'_4 . Both N-terminal half and C-terminal half are linked through a loop region. There are two EF hands. EF hand 1 is located in the N-terminal half and the second EF hand is situated in C-terminal half.

(e) CaBP7

The model of CaBP7 consists of 9 alpha helices and there are two beta strands (antiparallel). It is composed of N-terminal and C-terminal halves which are in blue and red colour [fig. 3.3.2 (e)] respectively. The N-terminal is composed of α_1 , α_2 , α_3 , α_4 and C-terminal is of α'_1 , α'_2 , α'_3 , α'_4 , α'_5 . Both N-terminal half and C-terminal half are linked through a loop region. There are three EF hands. EF hand 1 and EF hand 2 are located in the N-terminal half and the third EF hand is situated in C-terminal half.

(f) CaBP9

The model of CaBP9 consists of 9 alpha helices and there no beta strands formation. It is composed of N-terminal and C-terminal halves which are in blue and red colour [fig. 3.3.2 (f)] respectively. The N-terminal is composed of α_1 , α_2 , α_3 , α_4 , α_5 , α_6 and C-terminal is of $\alpha'1$, $\alpha'2$, $\alpha'3$. Both N-terminal half and C-terminal half are linked through a loop region. There are three EF hands. EF hand 1 is located in the N-terminal half and the second EF hand and third one are situated in C-terminal half.

(g) CaBP10

The model of CaBP10 consists of 8 alpha helices but there is no beta strand formation. It is composed of N-terminal and C-terminal halves which are in blue and red colour [fig. 3.3.2 (g)] respectively. The N-terminal is composed of α_1 , α_2 , α_3 , α_4 and C-terminal is of $\alpha'1$, $\alpha'2$, $\alpha'3$, $\alpha'4$. Both N-terminal half and C-terminal half are linked through a loop region. There are three EF hands. EF hand 1 and 2 are located in the C-terminal half and the third EF hand is situated in N-terminal half.

3.4.3 EF hands

EF hands prediction is done on the basis of two things. First it should satisfy the structure condition of helix – loop - helix formation and the loop sequence (conserved 12 amino acid sequence) which would coordinate with the calcium. Second is - In the prosite scan whichever part of the sequence that was having score of more than 10 for the calcium binding domain was considered. And also some of the experimental values (unpublished data) of calcium binding constants for some of the EhCaBPs were also considered. Those EF hands which satisfy all the conditions of EF hand is shown in the green colour with labeled side chain and those which don't have been shown in green colour in the fig. 3.3.3 (a), (b), (c), (d), (e), (f), (g).

3.4.4 Validation of EhCaBP protein models

Table 3.5.1 Members of the EhCaBP family and the sequences used for modeling

Modelled member of EhCBP	GenBank accession no.	Amino acid range of Eh CaBP model used	Template PDB ID	Percentage (%) Sequence identity with template
CaBP3	EAL46322	8-151	1QTX	48
CaBP4	EAL51814	5-139	2OBH	31
CaBP5	EAL46660	2-136	2OBH	25
CaBP6	EAL50371	10-148	2B10	27
CaBP7	EAL43751	5-149	2B10	32
CaBP9	EAL52004	11-147	2AAO	36
CaBP10	EAL50237	7-152	1QTX	36

Table 3.5.2 Ramachandran plot statistics of EhCaBP proteins studied

Name of EhCaBP	Residues in most favoured regions [A,B,L]	Residues in additional allowed regions [a,b,l,p]	Residues in generously allowed regions [\sim a, \sim b, \sim l, \sim p]	Residues in disallowed regions [\sim a, \sim b, \sim l, \sim p]
CaBP3	91%	7.5%	1.5%	0%
CaBP4	91.7%	7.4%	0.0%	0.8%(HIS55)
CaBP5	95.9%	2.4%	1.6%	0.0%

CaBP6	89.8%	10.2%	0.0%	0.0%
CaBP7	82.4%	14.5%	2.3%	0.8%(SER78)
CaBP9	85.4%	11.4%	3.3%	0.0%
CaBP10	94.0%	6.0%	0.0%	0.0%

Ramachandran plot statistics of the EhCaBP protein models are shown in table 3.5.2. No residue was found in the disallowed region in EhCaBPs modelled, as illustrated in fig. 3.3.4 except for His 55 in CaBP4 and Ser 78 in CaBP7 protein. CaBP6 and CaBP7 has largest deviation from the template and deviation is so large that when there was superposition of these structures on their respective templates then not more than 15 residues could be aligned. But on other accounts of model validation these conform to the criteria of the correct model. ProsaWeb analysis of the six members indicated that z-scores of all the EhCaBPs proteins lie within the standard defined values, and the energy plot indicated that the molecules have overall negative energy and therefore the models are energetically stable (fig. 3.3.5).

Though there could be a lot of sequence diversification among all the seven members EhCaBPs, EF hand formation is quiet constant with variation in number of EF hands. There are 27 EhCaBPs in *Entamoeba histolytica* which indicate that the function assigned to all or most of them could not be just acting as a calcium buffer Structural variation in these EhCaBPs also gives some credence to this idea that the function of all these EhCaBPs could be more than just calcium buffers. The idea of functional variation in these EhCaBPs also stem from the fact that a parasite generally does not keep superfluous things until unless a function is assigned to each one of them.

3.4.5 Electrostatic Representation

Table 3.5.3 Charge of the proteins studied

Protein	Total Charge
CaBP3	-16.00
CaBP4	-11.00
CaBP5	-7.00
CaBP6	-9.00
CaBP7	-6.00
CaBP9	-15.00
CaBP10	-14.00

An electrostatic presentation of different faces of the seven EhCaBP revealed that most of the charged residues are present on the surface of the proteins (fig. 3.3.6). All the members studied carry an overall negative charge. The details of the charge of the seven EhCaBPs proteins studied are given in table 3.5.3. Although among the EhCaBP family members there is some similarity in their structures, the charge distributions on the molecules are quite different from each other. In CaBP6, 7, 9, 10 a groove is suggested. Except CaBP6 all have negative charge on the groove lining. The groove may act as a place for binding with some partner having positive charge on the surface. It can be clearly seen that a large negative patch is present on the ventral side [fig. 3.3.6 (f)] of CaBP9 and 4th side [fig. 3.3.6 (e)] of CaBP7 proteins. A similar-sized negative patch is present on the 3rd side of CaBP4 [fig. 3.3.6 (b)] and 4th side of CaBP5 proteins [fig. 3.3.6 (c)] and this patch surrounds the central groove of the molecule, indicating the affinity of these areas to positively charged molecules. It is also seen that a large positive patch is seen on the dorsal middle segment of CaBP4 and CaBP5 which correspond to the central linker region of the molecule. The central linker region in CaBP is thought to add

flexibility to molecule for binding to the target proteins so this positive patch might be interacting with negatively charged residues of the target proteins. Members of the EhCaBP family show a very different charge distribution when compared to each other, as illustrated in (fig. 3.3.6).

3.4.6 Amino Acid Composition and Solvent Accessibility

To determine whether these proteins have a different amino acid composition in terms of charged residues and other residues, an amino acid composition analysis of positively and negatively charged and other critical residues of the seven EhCaBPs proteins was carried out. The amino acid composition analysis was done along with Calmodulin. The percentage of positively charged amino acids (K, R and H) is almost constant in the EhCaBP proteins except CaBP5 which is having less number of positively charged residues, as shown in fig. 3.3.7 (b), whereas in the case of negatively charged residues (D and E), the amino acid composition was found to vary greatly and there were less negatively charged residues in the CaBP5 and CaBP7 in comparison to other members. In comparison with Calmodulin seven EhCaBPs are less acidic. There is a large percentage of a hydrophobic residue or non polar residues (more than 41%) across the seven EhCaBPs and this comes second to polar residues [fig. 3.3.7 (b)]. The profile for polar and non polar residues also matches with the Calmodulin. Most of these proteins are highly soluble in nature and this is seen in solvent accessible chart (fig 3.3.8). Most of the hydrophobic residues are shielded thereby exposing only charged residues. And this type configuration of residues makes a protein highly soluble in nature. There are lots of loop in some of these EhCaBPs which provide them a lot of flexibility in accommodating greater things.

There is a minimal hydrophobic core in seven members of EhCaBPs molecules with comparatively few residues buried, thus most residues are solvent-accessible. Charged residues are situated mainly on the surface of the molecule (relatively very easily accessible molecular surface) thereby indicating the significance of these residues in the interaction of these proteins with the targets. Methionine percentage in calmodulin is

almost similar to the CaBP4 and CaBP5 [fig. 3.3.7 (a)]. In Calmodulin methionine gets exposed to the surface on the binding of calcium to the calmodulin and methionine interacts with target peptide (Babu et al., 1988). In CaBP4 most of the methionine is buried but in CaBP5 some of them are more exposed in comparison to the CaBP4 [fig. 3.3.8 (b)]. Very few cysteines are present in these CaBPs and in CaBP9 and CaBP3 there is no cysteine [fig. 3.3.7 (a)] which is identical to Calmodulin where there are no cysteine. In terms of aromaticity maximum percentage of aromatic amino acid is present in CaBP7 whereas in other members it is somewhat similar. In comparison to Calmodulin the percentage of aromatic amino acid is more across all the seven EhCaBPs. In CaBP7 some of the Phenyl alanines are solvent accessible [fig. 3.3.8 (c)].

3.4.7 Comparison of models with their respective templates

Table 3.5.4. Comparison of models with their respective templates

Modelled member of EhCBP	Template PDB ID	Number of C _α residues	RMS deviation of C _α of the model with the template
CaBP3	1QTX	125	0.282 Å
CaBP4	2OBH	134	1.438 Å
CaBP5	2OBH	135	4.084 Å
CaBP6	2B10	-	
CaBP7	2B10	-	
CaBP9	2AAO	101	0.452 Å

CaBP10	1QTX	130	0.305 Å
--------	------	-----	---------

In the table 4 structural deviations between the template and model has been mentioned.

(a) CaBP3

Deviation of the CaBP3 model with respect to the template has been mentioned in the table 3.5.4. There is very less deviation of model with respect to the template. The overall structural fold in CaBP3 is almost identical to the template. The structural variation in the model with respect to the template has been encircled [fig. 3.3.9 (a)]. Variation in the model correspond to the 133-140 residue position of the CaBP3 model.

(b) CaBP4

Deviation of the CaBP4 model with respect to the template has been mentioned in the table 3.5.4. In comparison to other deviation mentioned for CaBP5 in the table 3.5.4 it is quiet less. Structural divergence with respect to the template corresponds to the all along the loop both at the C and N-terminal and in terms residue position it is at 90-96, 3-5, 54-56. The structural variation in the model with respect to the template has been encircled [fig. 3.3.9 (b)].

(c) CaBP5

Deviation of the CaBP5 model with respect to the template has been mentioned in the table 3.5.4 and it is quiet large in comparison to other models mentioned in the table 3.5.4. Structural divergence with respect to the template corresponds to the all along the loop both at the C and N-terminal and also at the central helices and in terms residue position it is at 40-45 (helix), 12-127, 90-95, 50-57, 15-18, 68-84 (central helix). The

structural variation in the model with respect to the template has been encircled [fig. 3.3.9 (c)].

(d) CaBP9

Deviation of the CaBP9 model with respect to the template has been mentioned in the table 3.5.4. In comparison to other deviation mentioned for CaBP4 in the table 3.5.4 it is quiet less. Structural divergence with respect to the template corresponds to the 16-126 residue--loop, 134-137--loop, 44-47--loop, 37-40--loop. Missing density for the template between 132-140 residue positions (loop) can also be traced by model's 114-126 residue position. And this can rebuild the structure which is absent in the template. Some minor variation is also seen at 60-62, 2-4 and 35-42 residue position. The structural variation in the model with respect to the template has been encircled [fig. 3.3.9 (d)].

(e) CaBP10

Deviation of the CaBP10 model with respect to the template is very less. In comparison to other deviation mentioned for CaBP3 in the table 3.5.4 it is quiet similar in magnitude. Structural divergence with respect to the template corresponds to the differences along some loop position 133-138. Missing density for 84-89 in the template can also be traced from the model's 86-95 residue positions. And this can rebuild the structure which is absent in the template the structural variation in the model with respect to the template has been encircled [fig. 3.3.9 (e)].

When CaBP6 and CaBP7 were superimposed on their respective templates it was found out there was huge R.M.S. deviation more than 11 Å and that too only for 15 residues. That's why no figure was made for this. In spite of deviations of these models with their respective template, these models stood the test of all the model validations programmes including prosaweb, procheck etc. So even if there was variation of the model with the

template, the model was found up to the mark when the model was tested on other parameters which are used to verify the correctness of the model.

3.4.8 Alternative models for EhCaBP7 and EhCaBP9

Alternative models for EhCaBP7 and EhCaBP9 were also made with N terminal half of EhCaBP1 as a template (PDB id 2NXQ). Template length in case of EhCaBP7 is 5-66 where as in EhCaBP9 it is 1-65. And the length of the template (EhCaBP1) which is same for both of them that is 2-67.

Ramachandran plot of the alternative models of EhCaBP7 and EhCaBP9 reveals that in case of CaBP7 there is no residue in the disallowed region and in CaBP9 there are 2 residues which are in the disallowed region. And also the energy plots for the models are quiet favourable indicating that these are energetically stable. So these two statistics validate the models [fig. 3.3.10 (a), (b), (c), (d), (e), (f), (g)].

Table 3.5.5 Ramachandran plot statistics of alternative model of EhCaBP(7 & 9)

Name of EhCaBP	Residues in most favoured regions [A,B,L]	Residues in additional allowed regions [a,b,l,p]	Residues in generously allowed regions [\sim a, \sim b, \sim l, \sim p]	Residues in disallowed regions [\sim a, \sim b, \sim l, \sim p]
CaBP7	78.2%	16.4%	5.5%	0%
CaBP4	71.4%	23.2%	1.8%	3.6% (Ile18, Glu 64)

Table 3.5.6 Sequence identity of EhCaBPs with EhCaBP1

Sequence 1	Sequence 2	% Identity
CaBP1	CaBP2	78
CaBP1	CaBP3	19
CaBP1	CaBP4	20
CaBP1	CaBP5	11
CaBP1	CaBP6	15
CaBP1	CaBP7	21
CaBP1	CaBP9	26
CaBP1	CaBP10	21

This model reveals that the EF hand motif in EhCaBP7 and EhCaBP9 are separated from the other EF hand of the protein by a straight helix (fig. 3.3.10) which is different from any known CaM-like proteins.

In EhCaBP7 and EhCaBP9 the EF hand is not connected with the other EF hand of the protein by loop but is joined 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. Sequence comparison of EhCaBP7 and EhCaBP9 with other Ca²⁺ 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 conserved Gly and Pro residues which have been highlighted in black colour in fig. 3.3.1 induce a bend in this region. Out of the nine EhCaBP sequence (aligned) shown in the fig. 3.3.1 only in three EhCaBP sequence that is EhCaBP1, EhCaBP7, EhCaBP9 both conserved Gly and Pro residues which are helix breakers, are replaced by Arg and Lys in EhCaBP1, Tyr and Thr in EhCaBP7, Arg and Lys in EhCaBP9 (fig. 3.3.1) thereby causing the stabilization and straightening of the helix (Kumar et al., 2007). These residues have been highlighted

in black colour in fig. 3.3.1. This provides the probable explanation for the differences in the arrangement of EF hand motifs in EhCaBP7 and EhCaBP9 in comparison to CaM.

3.4.9 Phylogenetic Studies

Analysis of the phylogenetic tree for the *Entamoeba histolytica* CaBPs [fig. 3.3.11 (b)], based on the amino acid sequences of 27 EhCaBPs proteins, suggests that the closest relative of well characterized EhCaBP1 and EhCaBP2 (Chakrabarty et al., 2004; Sahoo et al., 2004) is EhCaBP9 also. On the basis of this it could be also said that the role of pseudopodia formation executed by EhCaBP1 could also be shared by EhCaBP9. Among these two least dissimilar proteins are EhCaBP24 and EhCaBP25. Well this may suggest that both of them may be acting complementary to each other in terms of functions in *E. histolytica*. The other peculiarity to EhCaBP24 and EhCaBP25 is that this group has taken a completely a different path to evolve at the outset itself. The dissimilarity in the structure of these homologues is suggestive of divergent evolution of the family.

In another phylogram [fig. 3.3.11 (a)] of nine members of EhCaBPs with Calmodulin and Skeletal Essential Light Chain it was observed that the closest relative of Calmodulin is EhCaBP3 and the other one is EhCaBP4. EhCaBP6 is the one which has completely diverged out in the beginning itself.

Though function for most of them is not known, the closeness in the phylogram indicates some similarity at some level of function. It is corroborated by the fact that all the members that have been investigated so far have EF hands.

3.4.10 Conclusion

The study of seven EhCaBPs has indicated that the structures of these proteins are not similar to each other. This gives credence to the thought that there could be some functionality difference among these proteins. That is these bind to different set of target

proteins and execute different downstream signals. But one thing which is common to all these proteins is the presence of EF hands. Though all of these EhCaBPs has different number of EF hands, this feature is present among all these EhCaBPs. This suggests that the mechanism by which they transfer or receive the signal might be having a common pathway. And at some or the other point they are linked to each other along the signal transduction pathway. The idea of diversity in the function is also corroborated by the fact that there is no conservation of a single amino acid that could be found out when alignment of all 27 EhCaBPs was done. The sequence for EF hand calcium binding loop also vary much thereby giving indication that there could be difference in the calcium binding affinity among all these seven EhCaBPs. In EhCaBP6 there is good number of loops which indicate that there could be a lot of flexibility in the structure of this protein. This greater flexibility in the structure of EhCaBP6 could give it greater role in the maneuvring in the *Entamoeba histolytica* in terms of binding to the target protein. Since most of these proteins are present only in the *Entamoeba histolytica* therefore these proteins can be thought of playing a role in the pathogenesis of amoebiasis. One can also design antibody against these proteins if proven beyond doubt about their role in the pathogenesis.

CHAPTER 4

Calcium Binding Constants Database of Calcium Binding Proteins

4.1 Description about the database

The EF-Hand Calcium-Binding Proteins Data Library (EF-Hand CaBP-DL) is a highly curated collection of sequence of the EF-Hands with calcium binding affinity of EF-Hand superfamily of calcium-binding proteins. It has been conceived, designed, and implemented with a sole aim to find a relationship among the variation in the calcium binding abilities among the different members of EF-hand superfamily across the species.

All information that is not obtained directly from another public database has been published in a peer-reviewed journal. Each data associated with the library has been checked and validated it for inclusion in the database. The data library has the reference associated with every particular piece of information which can be used for checking the methodologies used for the determination of calcium binding constants.

As can be seen in data library, the information in the database is organized around proteins. Each mutant or isoform of a protein is considered a unique protein in the database, and is treated separately. This allows storage of the calcium-binding constants, for instance, of two isoforms of parvalbumin from the same species, and clearly indicates that the two sets of binding constants are for two different chemical entities. It also allows storage of any type of information about a mutant, thereby giving a complete picture about the variation in calcium binding constants among the native and its mutant due to the mutation of EF hand loop.

All of the isoforms and mutants of a given protein associated by a common group are stored in the database. This allows correlating calcium binding constants of isoforms and mutants of a given protein from the evolution perspective.

The EF-Hand CaBP-DL is divided into three main sections: general information (name of the protein), sequence information (EF hand sequence) and Reference.

- **General Information** includes the main entry point for information about a particular protein. The General Information for a protein summarizes name of the protein and to the organism which it belongs.
- **Sequence Information** includes information about the amino acid composition of the binding loops.
- In addition to these informations, there is also a section that includes lists of references.

Data integrity is also ensured by the requirement for the inclusion of the reference from which the information was obtained. Only data that has been published in a peer-reviewed article is included into the database.

4.2 Results and Discussion

Table 4.3

(1) Calcium Binding Constants for Calcium Binding Proteins

Sequence	Ca ²⁺ binding cons.	Loop Name	Reference
DKDGDGTITT KE	1 x 10 ⁷	Loop1 Bovine cam	(1991) <i>J.B.C.</i> 266, 8050-8054.
DADGNGTIDFPE	3.98 x 10 ⁷	Loop2 Bovine cam	(1991) <i>J.B.C.</i> 266, 8050-8054.
DKDGNNGY ISAAE	3.16 x 10 ⁶	Loop3 Bovine cam	(1991) <i>J.B.C.</i> 266, 8050-8054.
DIDGDGQVNYEE	2.5 x 10 ⁶	Loop4 Bovine cam	(1991) <i>J.B.C.</i> 266, 8050-8054.
DTDGSGTIDAKE	8.30 x 10 ⁵	Loop1 caltractin chalmyd	(1994) <i>J.B.C.</i> 269, 15795-15802.
DKDGSSTIDFEE	8.30 x 10 ⁵	Loop2 caltractin chalmyd	(1994) <i>J.B.C.</i> 269, 15795-15802.
DNSGTITI KDLR	6.25 x 10 ³	Loop3 caltractin chalmyd	(1994) <i>J.B.C.</i> 269, 15795-15802.
DKDGSSTIDFEE	6.25 x 10 ³	Loop4 caltractin chalmyd	(1994) <i>J.B.C.</i> 269, 15795-15802.
DKDGDGCITTRE	3.80 x 10 ⁵	Loop1 human cam-like protein	(1992) <i>Bioch.</i> 31, 12826-12832
DRDNGTVDFFPE	1.90 x 10 ⁵	Loop2 human cam-like protein	(1992) <i>Bioch.</i> 31, 12826-12832
DKDGNNGVFSAAE	4.90 x 10 ⁴	Loop3 human cam-like protein	(1992) <i>Bioch.</i> 31, 12826-12832
DTDGDGQVNYE	1.20 x 10 ⁴	Loop4 human cam-like protein	(1992) <i>Bioch.</i> 31, 12826-12832
DQDKSGFIEEDE	2.7 x 10 ⁹	Loop1 parvalbumin cyprinus	(1980) <i>Eur J Bioch.</i> 111, 73-78
DSGDGDKIGVDE	2.7 x 10 ⁹	Loop2 parvalbumin cyprinus	(1980) <i>Eur J Bioch.</i> 111, 73-78
DTLIKRELKQLITKE	2.7 x 10 ⁷	Loop1 S100A12 Sus	(1994) <i>J.B.C.</i> 269, 28929-28936.
DANQDEQVSFKE	6.50 x 10 ⁴	Loop2 S100A12 Sus	(1994) <i>J.B.C.</i> 269, 28929-28936.
DPNQLSKEELKLLQTE	1.6 x 10 ⁸	Loop1 Calbindin D9k Bovine	(1991) <i>Bioch.</i> 30, 154-183
DKNGDGEVSFEE	4 x 10 ⁸	Loop2 Calbindin D9k Bovine	(1991) <i>Bioch.</i> 30, 154-183
DKDKSGTSLVDE	(2.9 ± 0.3) x 10 ⁵	Loop1 S100 like liver lung fish	<i>Eur. J. Biochem.</i> 269, 3433-3441 (2002)
DTNKDGQVSWQE	(6.0 ± 0.7) x 10 ³	Loop2 S100 like liver lung fish	<i>Eur. J. Biochem.</i> 269, 3433-3441 (2002)
ELDTLGEESYKD	5.5 x 10 ⁴	Loop1 GF14 Arabidopsis	<i>The Plant Cell</i> , Vol. 6, 501-510, 1994
VQFDTCHNLDA	4 x 10 ⁴	Loop1 lima bean lectin	<i>Plant Physiol.</i> (1991) 95, 286-290
DSFDTDSKGFITPE	1.6 x 10 ⁶	Loop1 Trop C alpha cray fish	264, 30, <i>JBC</i> , 18240-18246, 1989
DGSGEIEFEFAE	8.1 x 10 ³	Loop2 Trop C alpha cray fish	264, 30, <i>JBC</i> , 18240-18246, 1989
DRCGDGYITQVLRE	1.1 x 10 ³	Loop3 Trop C alpha cray fish	264, 30, <i>JBC</i> , 18240-18246, 1989
DSFDTDSKGFITPE	1.9 x 10 ⁴	Loop1 Trop C gamma cray fish	<i>Eur. J. Biochem.</i> 269, 3433-3441 (2002)
DGSGELEFEFVE	8.1 x 10 ²	Loop2 Trop C gamma cray fish	<i>Eur. J. Biochem.</i> 269, 3433-3441 (2002)
DPDKPGKILLMD	0.014±0.005µMdisCon	Loop1 Rabbit Serum Paraoxonases	<i>Drug Met. and Disp.</i> 26, 7, 653-660, 1998
DEDNIVYLMVVN	5.31±0.94µMdissoCon	Loop2 Rabbit Serum Paraoxonases	<i>Drug Met. and Disp.</i> 26, 7, 653-660, 1998
NPNSPGKILLMD	0.38± 0.09µMdisCons	Loop1 human Serum Paraoxonases	<i>Drug Met. and Disp.</i> 26, 7, 653-660, 1998
DEDNAMYLLVVN	6.2±1.2µMdisConsta	Loop2 human Serum Paraoxonases	<i>Drug Met. and Disp.</i> 26, 7, 653-660, 1998

Ca²⁺ binding cons. - Calcium binding constant

dis. Cons. – Dissociation constant

Loop – Calcium binding loop

Name - Organism name

4.2.1 Determinants of Ca²⁺ affinity

For a clear understanding of the Ca²⁺ binding ability of the various EF-hands, four factors must be considered. The first is the intrinsic Ca²⁺ affinity of each EF-hand (ΔG_{intr}), which mirrors both the ligands presence and the contributions of non coordinating interactions. The second is the ability of a given EF-hand to differentiate between Ca²⁺ and Mg²⁺, a cation chemically similar to Ca²⁺ that can pose significant selectivity problems due to its high cytosolic concentration. The third factor is co-operativity, a phenomenon that enables an EF-hand pair to bind Ca²⁺ as a unit. The final factor for those EF-hand proteins is the presence of target proteins which also somehow affect the calcium affinity. Taken together, the observed Ca²⁺ affinity of an EF-hand can be described as:

$$\Delta G_{\text{obs}} = \Delta G_{\text{intr}} + \Delta \Delta G_{\text{sel}} + \Delta \Delta G_{\text{co-op}} + \Delta \Delta G_{\text{interact}} \quad (1)$$

Where the effects of Mg²⁺ selectivity ($_{\text{sel}}$), co-operativity ($_{\text{co-op}}$) and target interaction ($_{\text{interact}}$) are seen as an energetic coupling ($\Delta \Delta G$) that reflects the difference in affinity [ΔG_{tot} (total free-energy change) of binding] in the presence and absence of these factors

(e.g. $\Delta \Delta G_{\text{sel}} = \Delta G_{\text{tot, presence of Mg}} - \Delta G_{\text{tot, absence of Mg}}$). So only one term that is independent of any other factor is ΔG_{intr} otherwise for the calculation of contribution towards calcium binding affinity by other terms in the equation (1) one has to take into account other factors.

4.2.2 Determinants of intrinsic Ca²⁺ affinity

The EF-hand presents an intriguing mystery, as the Ca²⁺ dissociation constants found in this protein family range from 10⁻⁹ to 10⁻⁴ M. If one considers thermodynamics and the following well-known equations, an understanding can be made:

$$\Delta G = -RT \cdot \ln K_a \quad (2)$$

Where R is the gas constant and T is the temperature.

$$\Delta G = \Delta H - T\Delta S \quad (3)$$

Where ΔH is the change in the enthalpy of the system, T is the temperature in degrees kelvin and ΔS is the change in entropy. The intrinsic Ca²⁺ affinity is determined by the difference in the Gibbs' free energy (ΔG) between the unbound and bound states. Clearly, the larger and more negative the ΔG , the higher the affinity. At a constant temperature,

phenomena that increase ΔS or decrease ΔH between the unbound and bound states should lead to higher affinity.

One factor mentioned in equation (3) is the entropy and this can be analysed further in order to understand the complexities of the interaction between the calcium and EF hand loop.

4.2.3 Entropic contributions to affinity

It is a well documented phenomenon that on binding of the calcium to the EF hand loop there is a release of frozen water molecules from its coordination spheres to the bulk solvent. An event that is fairly constant per bound calcium and this lead to increase in solvent entropy (Linse et al., 1995). This increase in solvent entropy is large enough to favour the increase in ΔG . As a result of the exchange of water molecules with the bulk solvent there is an increase in calcium affinity. So greater the exchange of water molecules, greater is the calcium affinity.

The above mentioned fact can be highlighted by the identity of ninth and twelfth loop positions. The coordinating side chain of a glutamic acid or glutamine residue in the ninth position is large enough to form linkage with the bound calcium ion directly. If the coordinating side chain is small, it has to take the help of water molecules to bridge the gap (Rake et al., 1996). The consequence of this is apparent when the EF-hands of parvalbumin & subfamily member oncomodulin are taken into account. In case of EF2 of oncomodulin there is an aspartic acid residue in the ninth position and a water molecule in its co-ordination sphere instead of glutamic acid. The resulting loss in increased solvent entropy due to the more incomplete loop chelation is thought to contribute to the lower affinity of this site compared with that of EF3 of oncomodulin and those found in parvalbumin (Lee et al., 2004).

In cases where an aspartic acid residue is found instead of a glutamic acid residue in position 12 of the EF-loop the above mentioned explanation can be applied in the same manner. This can be further corroborated by the crystal structure of CIB (calcium and integrin binding protein) where in the aspartic residue has to rely upon the water

molecule to complete the coordination sphere (Gentry et al., 2005). Consequently there is a fivefold decrease in affinity of EF3 (Asp12) compared to EF4 (Glu12) of CIB (calcium and integrin binding protein) (Amniuk et al., 2004).

The quantum of the favorable entropy term in eqn (3) can be decreased if unfavorable factors have a role to play in the entropy of Ca^{2+} binding to EF-hands. There are enough evidences to show that to bring a change in conformation of EF hand loop more energy is spent in comparison to the others. Both crystallographic B-factors and NMR relaxation give evidences that a glycine triplet in the *N*-terminal part of the EF1 loop of sTnC is more disordered, and consequently more stable, in terms of conformational entropic free energy, than EF2 (Li, E., et al., 1995; Strynadka et al., 1997). The expenditure of energy in bringing about an order in this flexible loop is reflected in its tenfold lower affinity compared with EF2 (Gagne et al., 1997). If there is an availability of 'preformed' EF-hands, such as the ψ - hand of calbindin D9K, the entropy cost to bring about an order in the loop is reduced and consequently there is an increase in the calcium affinity of this site (Ke et al., 1991). There is every possibility that EF-hands put to use such conformational entropy costs due to loop flexibility in the apo form to modulate their Ca^{2+} affinity (Agne et al., 1998).

A second unfavourable entropic factor is the Ca^{2+} -induced exposure of hydrophobic surfaces, as experienced by the Ca^{2+} sensors, an energy expensive process (Foguel et al., 1996). The amount of energy consumed to expose the hydrophobic surface is considerable and is evident from the reduction in the calcium affinity. This consumption of energy act as an opposite to the favourable release of water molecules (Nelson et al., 1998). Closed compact structure of calcium sensors does not permit calcium to bind to EF hand. For the calcium binding to happen it has to come in the open conformation. If the calcium could be accommodated in the closed structure, there would be stronger binding of the calcium to the EF hand loop.

These are some of the reasons which may provide a greater platform for the better understanding of the variation in calcium binding affinity of EF hands.

CHAPTER 5

An Overview of OASS

5.1 General description of OASS

The cysteine biosynthetic pathway is a platform for the assimilation of inorganic sulfur into organic compounds in bacteria, plants, and parasitic protists, including *E. histolytica*, *T. vaginalis*, and *Trypanosoma cruzi*, which is the main causal agent of dreaded disease of Chagas' disease (American trypanosomiasis) Nozaki et al. 1998; 1999; 2005; 2000; 2001; Westrop et al., 2006). The pathway has been thoroughly analysed in bacteria and plants (Hell et al., 1994; Ogasawara et al. 1994; Saito et al., 1992). In plants, the whole pathway involves more than one compartments of the cell (Saito et al., 1994). In *E. histolytica*, where typical eukaryotic cell organelles like mitochondria and chloroplasts are absent, the pathway exists exclusively in the cytosol. In case of *T. vaginalis* and *T. cruzi* the two enzymes in the pathway does not have organelle targeting sequences so the whole pathway is accomplished in the cytosol. In bacteria and plants, which has the ability to reduce sulfate into sulfide via sulfite, first brings the sulfate from outside the cell by sulfate transporters (Saito et al., 1992; 1994). After sulfate activation and reduction, incorporated sulfate (+6) is reduced to sulfide (-2), which receives an alanyl moiety from a donor molecule. Serine O-acetyltransferase (SAT) (EC 2.3.1.30) is instrumental in the production of the alanyl donor O-acetyl- serine from serine and acetyl-CoA (Nozaki et al., 1999) (fig. 5.3.1). CS [O-acetyl- L-serine (thiol)-lyase] (EC 4.2.99.8) then comes into picture and catalyzes the production of L-cysteine by the transportation of the alanyl moiety from the O-acetylserine to sulfide (Nozaki et al. 1998, 2000). But in case of *E. histolytica* and *T. vaginalis* which do not have sufficient avenues to utilize the sulfate via sulfate reduction pathway, take the help of sulfide derived from the iron-sulfur proteins present in ingested bacteria which is residing in either intestine or vagina. In contrast to these parasites, animals do not have the sulfur-incorporation pathway and thus dependent on the methionine from any external source for the sulfur requirement.

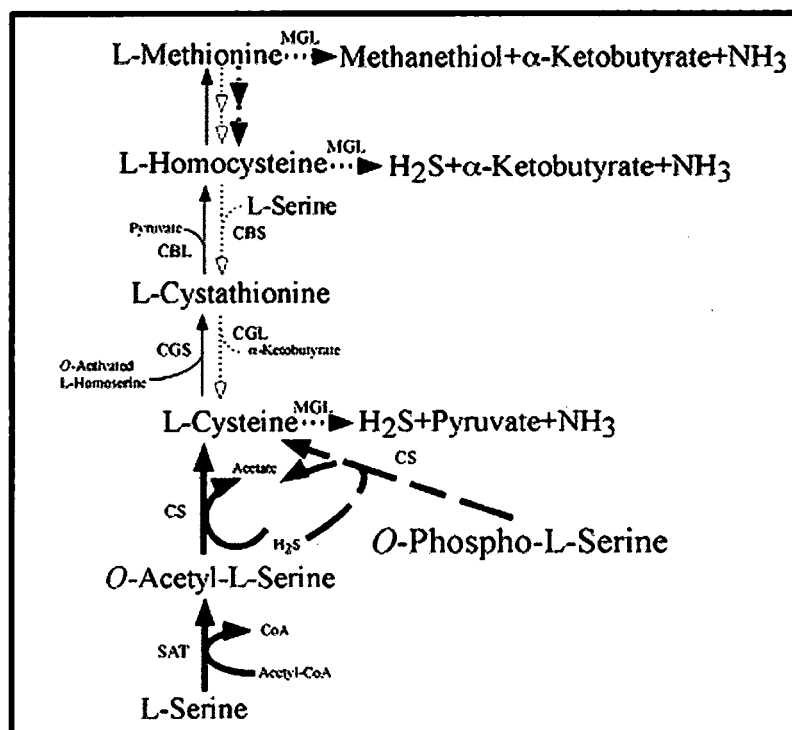


Fig. 5.3.1 General scheme of transsulfuration, cysteine biosynthesis, and sulfur amino acid degradation. The schematic diagram shows all pathways present in bacteria. Open arrows with thin dotted lines depict pathways present in mammals; arrows with thick dotted lines represent pathways present in both *E. histolytica* and *T. vaginalis*; and arrows with thick unbroken lines or thick broken lines represent pathways present only in *E. histolytica* or *T. vaginalis*, respectively. CBS, cystathionine -synthase; CBL, cystathionine -lyase. Adapted from (Vahab et al., 2007)

5.1.1 *Entamoeba histolytica* OASS

Entamoeba histolytica, the causative agent of human amebiasis, is an enteric protozoan parasite, and causes amebic colitis and extraintestinal abscesses (e.g. hepatic, pulmonary, and cerebral) in approximately fifty million inhabitants of endemic areas (The World Health Report, 1995). *E. histolytica* trophozoites which are mainly found in the colonic lumen, which has very less oxygen that is essentially an anaerobic environment, has an access to many bacterial and host derived food. However, the level of oxygen stress they face during tissue invasion, metastasis, and extraintestinal propagation is huge. Despite being an anaerobe *E. histolytica* trophozoites have been proven beyond doubt that it has

some inclination for oxygen and can withstand oxygen tension (Weinbach et al., 1974). Thus, *E. histolytica* must possess mechanisms to detoxify the reactive oxygen species produced by the mammalian host and within their own cells. However, *E. histolytica* does not have well developed antioxidant defense mechanisms present in aerobic or other aerotolerant cells, such as catalase, peroxidase, reduced glutathione, and the glutathione-recycling enzymes glutathione peroxidase and glutathione reductase (Mehlotra et al., 1996). Instead, they have different way of dealing with this kind of situation and have alternative mechanisms for detoxification similar to those known to exist in certain prokaryotes (Bruchhaus et al., 1995).

The biological significance of cysteine synthesis in “amitochondriate” protists is still not yet fully understood. It was shown that cysteine is a major intracellular thiol in these organisms (Brown et al., 1993; Fahey et al., 1984; Gillin et al., 1980b; Gillin et al., 1984). L-Cysteine could not be replaced by any other thiols or reducing agents in the growth medium (Gillin F. D., and L. S. Diamond 1981d), which proves that cysteine is indispensable for the existence of *E. histolytica* and *G. intestinalis* (Brown et al. 1993, Gillin et al. 1981c). Moreover, cysteine also takes part in the antioxidative defense of *E. histolytica* (Nozaki et al., 1999). Cysteine has significant contribution to the oxygen defense mechanisms in this glutathione-deficient organism. It has also been experimentally proven that amoebae need a high amount of extracellular cysteine for attachment to matrix, elongation, motility, and growth in vitro (Gillin et al., 1980b, 1981d).

In nature, the cysteine biosynthetic pathway is the platform for the assimilation of inorganic sulfur. They produce cysteine by transporting alanyl moiety of O-acetylserine, to sulfide, which is synthesized by reduction of sulfate via sulfite. This reaction is catalyzed by cysteine synthase (CS) (O-acetyl-L- serine (thiol)-lyase). This pyridoxal phosphate-dependent enzyme exists in a variety of bacteria (Byrne et al., 1988; Ogasawara et al., 1994) and plants (Hell et al., 1994; Noji et al., 1994; Romer et al., 1992; Saito et al., 1992).

From the recent study it has been shown that *E. histolytica* possesses three allelic SAT isotypes and three allelic isotypes of CS (Nozaki et al., 1998; 1999; 2000; 2001). There are several unique features of the amoebic SAT and CS which does not match to any other organism. It has been shown that SAT1 (previously named SAT) is under strict control in the cysteine biosynthetic pathway in *E. histolytica* (Nozaki et al., 1999). SAT1 which has an allosteric site is under the negative feedback control of cysteine. Negative regulation by L-cysteine is unique to the amoebic SAT1 enzyme. The other unique finding about the amoebic SAT1 is that there is a lack of protein - protein interaction with CS. In both bacteria and plants, CS and SAT does interact and form a heteromeric complex with a very high molecular mass of several hundred kilodaltons (Droux et al., 1998; 1992; Dan et al., 2000). However, both CS1 (and CS2) and SAT1 form a homodimer, but there is no evidence that they interact with each other under physiological conditions. There has been a lot of biochemical and genetic method to prove that there is no CS - SAT interaction. The methods employed to prove are (i) separation by conventional chromatography during purification from the crude cell lysate, (ii) and to coimmunoaffinity purification of the proteins, and (iii) the yeast two-hybrid system (Nozaki et al., 1999).

Among the three CS isotypes (CS1 to 3), two CS proteins (CS1 and CS2) are identical except for two conservative amino acid changes (Nozaki et al., 1998). CS3 is only among three which is quite dissimilar from the other two isotypes, with approximately 83% amino acid identity to CS1 and CS2 (Nozaki et al. 1998; 1999; 2000; 2001). While both CS1 and CS2 are present in the cytoplasm of *E. histolytica*, similar to prokaryotic CysK and CysM, the intracellular distribution of CS3 is not clear. One experiment showed that CS1, CS2, and CS3 rescued the growth defect of a CysK- deficient *E. coli* strain, thereby proving that all three isoforms of CS can even act as functional as CysK in a heterologous organism (Nozaki et al., 1998). The non virulent *E. dispar* also has two CS isotypes, CS1 and CS2 (82 to 83% mutual identity) (Nozaki et al., 2000). *E. dispar* CS1 and CS2 matches with its counterpart *E. histolytica* CS1/2 and 3, respectively. After pooling these informations together it can be understood that *Entamoeba* possesses at least two classes of CS isotypes, each with a distinct pI. But one thing is common to the CS isoforms that

is all of these CS isotypes lack signal sequences or organelle-targeting sequences, suggesting that they have only cytosolic presence. The presence of multiple cytosolic CS isotypes even in the non virulent *E. dispar* species gives ample indication that this enzyme is not having direct linkage with pathogenicity of the amoeba but that it plays an important housekeeping role in *Entamoeba*.

Over expression of CS (2- to 3-fold) but not of SAT (13-fold) by introducing multicopy plasmids in the organism lead to 2 times increase in the intracellular thiol content and the hydrogen peroxide resistance (Nozaki et al., 1999). These data indicate that the intracellular concentration of CS1 (or CS2) but not SAT1 mainly affects the thiol content. One of the major unsolved questions related to the pathway is why these protists possess apparently redundant systems, while they have lost many other metabolic pathways by reductive evolution.

5.1.2 *T. vaginalis* OASS

Recently, sulfur-assimilatory cysteine biosynthesis has been elucidated in *T. vaginalis* (Westrop et al., 2006). There is no SAT in *T. vaginalis* where as it has six copies of CS. Enzymological characterization of CS indicates that *T. vaginalis* CS has unique aspect that it can utilize both O-phosphoserine as well as O-acetylserine as an alanyl donor. Since *T. vaginalis* lacks SAT and therefore it does not have capacity to produce O-acetylserine. But *T. vaginalis* also has three copies each of 3-phosphoglycerate dehydrogenase and phosphoserine aminotransferase and these enzymes produce the substrate for CS that is O-phosphoserine from 3-phosphoglycerate taken from glycolysis.

5.1.3 Bacterial OASS

Cysteine is an essential amino acid that performs important aspect in the catalytic activity and structure of many proteins. The joining of two sulphurs that is disulfide bonds between cysteine residues is the most important aspect in the activation of bacterial transcriptional regulators such as OxyR (Zheng et al., 1998) and the molecular chaperone Hsp33 (Jakob et al., 1999). Disulfide bonds are not only required for proper folding but

also for stability of some proteins, particularly those found in extracytoplasmic compartments (Wedemeyer et al., 2000).

In bacteria, cysteine is synthesized from serine by incorporation of sulfide or thiosulfate. Sulfide is obtained by employing two methods that is either inorganic sulfate is transported from the outside and then reduced or from organic sulfonate compounds such as taurine (Van der Ploeg et al. 1996; 1998). The last step in cysteine biosynthesis is catalyzed by either O-acetylserine (thiol)-lyase A or O-acetylserine (thiol)-lyase B, encoded by the genes *cysK* and *cysM*, respectively (Fimmel et al., 1977; Hulanicka et al., 1974). The CysK and CysM proteins from *Escherichia coli* are 43% identical. The difference between CysK and CysM is that CysK produces cysteine from O-acetylserine and sulfide, while the CysM protein synthesizes cysteine from thiosulfate instead of sulfide. The reaction between O-acetylserine and thiosulfate produces S-sulfocysteine, which is converted into cysteine by an as yet uncharacterized mechanism (Nakamura et al., 1984). It is suggested that the O-acetylserine (thiol)-lyase B isozyme is preferentially used during specific environment of less oxygen that is during anaerobic growth conditions (Kredich et al., 1996). In *E. coli*, cysteine performs various functions and one of them is that it can be used to donate the sulfur moiety for methionine biosynthesis in a set of reactions known as the trans-sulfuration pathway. This pathway can be reversed in *Bacillus subtilis*, which can therefore use methionine as its sole source of sulfur (Grundy et al., 2002). The genes taking part in cysteine biosynthesis and sulfur incorporation in *E. coli* and *Salmonella enterica serovar Typhimurium* have been well analyzed (Kredich et al., 1996). In recent times, cysteine biosynthesis has been looked into in the gram-positive bacteria *B. subtilis* (Van der Ploeg et al., 2001) and *Lactococcus lactis*, and also in the archaeon genus *Methanosarcina* (Borup et al., 2000; Kitabatake et al., 2000). In contrast, cysteine biosynthesis and sulfur assimilation in the gram-positive *Staphylococcus aureus* (human pathogen capable of causing a variety of infections, ranging from minor skin and wound infections to life-threatening diseases (Lowy et al., 1998) have not been well studied.

5.1.4 Plant OASS

In plants making of cysteine from reduced sulphide in chloroplasts, mitochondria, and the cytoplasm is the end product in environmental sulfur assimilation. Cysteine is the metabolic sulfide donor for all cellular components containing reduced sulfur. In bacteria and plants, cysteine is not only necessary for the protein structure but also the starting material for most of the sulfur-containing metabolites such as methionine, glutathione, phytochelatins, iron-sulfur clusters, vitamin cofactors, and multiple secondary metabolites. Two enzymes are instrumental in the production of cysteine in plants and bacteria [fig. 5.3.2 (a)]. Serine acetyltransferase (SAT4 ; EC 2.3.1.30) transporting acetate from acetyl-CoA to serine, thereby producing O-acetylserine. O-Acetylserine sulfhydrylase (OASS or O-acetylserine(thiol)lyase; EC 4.2.99.8) uses pyridoxal 5 prime phosphate (PLP) as a cofactor to synthesize cysteine from O-acetylserine and sulfide. Association of these two enzymes into an assembly which is called the cysteine synthase complex coordinates sulfate assimilation and modulates cysteine synthesis at the cellular level (Kredich et al., 1969; Ruffet et al., 1994; Droux et al., 1998). The whole complex has one SAT hexamer and two OASS dimers (Kredich et al., 1969), and when the two enzymes are present in association SAT activity is enhanced and OASS activity decreases (Droux et al., 1998).

Environmental stresses changes the expression profile and enzymatic activity of OASS. Although OASS is constitutively expressed, there are other factors like sulfur, nitrogen, and carbon starvation conditions and abiotic stresses like salt and heavy metal exposure that modulates (increase) expression profile of OASS in *Arabidopsis thaliana* (AtOASS) (Takahashi et al., 1996; Dominguez-Solis et al., 2001). It has been shown that over expression of OASS in transgenic plants has a lot of positive impact (improves) on heavy metal tolerance (Dominguez-Solis et al., 2001), increases cysteine biosynthesis in response to sulfur stress (Saito et al., 1994), and is a kind of insurance against oxidative stress (Youssefian et al., 2001; Noji et al., 2001). In the plant cells, there are three isoforms, CS-A, CS-B, and CS-C, present in cytoplasm, chloroplasts, and mitochondria, respectively. This was exhibited by molecular cloning of the corresponding genes (Saito et al., 1992; 1993; 1994) as well as physical studies by organelle separation (Fankhauser

et al., 1976). Plant CSs were also shown to catalyze the production of beta-substituted alanines besides cysteine, and are considered to be involved, at least in part, in detoxifying internal toxins like cyanide, pyrazole, and 3,4-dihydroxypyridine (Ikegami et al., 1994). The OASS from different plant organelles share approximately 40% amino acid sequence identity with the bacterial enzymes (Bonner et al., 2005).

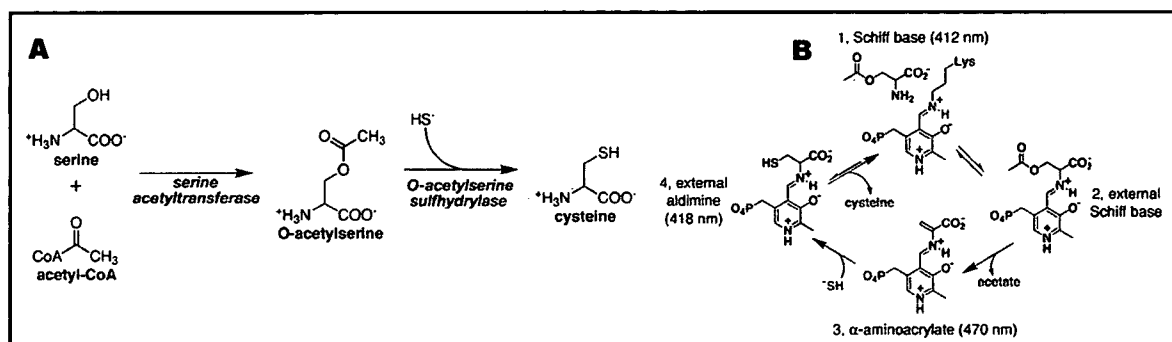


Fig. 5.3.2 Cysteine biosynthesis. Adapted from (Bonner et al., 2005)

The chemical reaction mechanism of OASS is well studied (Tai et al., 2001b). The enzyme active site contains PLP linked to a lysine as an internal Schiff base [fig. 5.3.2 (b), step 1)]. Binding of O-acetylserine displaces the lysine [fig. 5.3.2 (b), step 2)], initiating the first half-reaction yielding α -aminoacrylate intermediate linked to PLP [fig. 5.3.2 (b), step 3)]. The second half-reaction involves sulfide addition to the intermediate, thereby generating an external aldimine with the amino acid [fig. 5.3.2 (b), step 4)]. The active site lysine reacts with this intermediate, releasing cysteine and regenerating the Schiff base (Cook et al., 1992; Schnackerz et al., 1995).

5.2 Regulation of OASS

Multienzyme complexes are key themes in a variety of cellular processes, including translation, transcription, gene expression and signal transduction. Organizing macromolecules into complexes offers cells with a way of associating molecular networks and of controlling metabolism by bringing together key enzymes or channeling metabolites between enzyme active sites (Srere, 1987; Winkel, 2004). Kredich et al.,

(1969) is given credit for discovering one of the earliest examples of a macromolecular complex in primary metabolism by isolating the Cys synthase complex but elucidating molecular basis for assembly of this multienzyme complex is only starting up. In plants and bacteria, Cys biosynthesis happens in two stages (Wirtz and Droux, 2005). Ser acetyltransferase (SAT) produces O-acetylserine by transporting acetate from acetyl-CoA to Ser. Next, O-acetylserine sulfhydrylase (OASS) uses pyridoxal phosphate (PLP) as a cofactor to form Cys from O-acetylserine and sulfide.

In plants, strict control of the pathway by two mechanisms keeps intra cellular Cys levels at the required level (Saito et al., 1994). First, feedback inhibition of SAT by Cys can regulate synthesis of the amino acid (Noji et al., 1998). The second level of regulation is done by the association of SAT and OASS to make the Cys synthase complex. The function of the complex is not metabolic channeling, since O-acetylserine freely diffuses out of the complex (Kredich et al., 1969; Cook and Wedding, 1977; Droux et al., 1998). Instead, interaction of SAT and OASS regulates sulfate assimilation and controls Cys synthesis at the cellular level in plants (Hell and Hillebrand, 2001) (fig. 5.3.3). When there is enough amount of sulphur present in cells, the two enzymes associate and form the complex in which SAT activity increases and opposite to this, OASS activity decreases (Saito et al., 1995; Droux et al., 1998). This leads to the formation of O-acetylserine. Under sulphur deficient condition, O-acetylserine keeps piling up because free OASS is unable to produce Cys due to an absence of sulfide. Elevated O-acetylserine levels dissociate the complex, which down regulates SAT. Very high level of O-acetylserine concentration induces genes encoding sulfate transporters, ATP sulfurylase, OASS, and SAT (Hopkins et al., 2005). This results in more of sulfur uptake and reduction. As sulfur levels keep rising up, free OASS catalyzes Cys formation, which reduces O-acetylserine levels. This permits association of SAT and OASS, activation of SAT, and resumption of Cys biosynthesis. The exact analysis in terms of molecular mechanism of how OASS and SAT form the Cys synthase complex are unclear. One SAT hexamer (each Mr of; 180 kDa) and two OASS dimers (each Mr of; 70 kDa) comprise the complex in plants and bacteria (Kredich et al., 1969; Zhu et al., 1998).

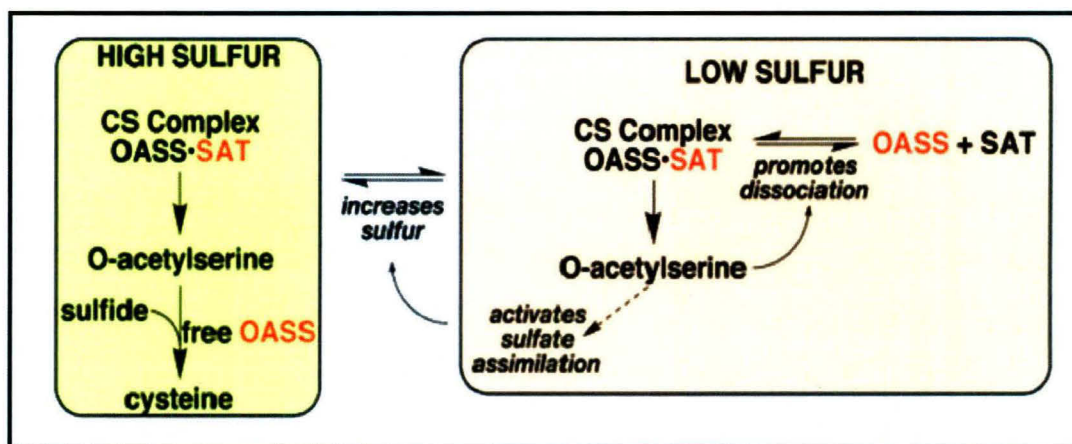


Fig. 5.3.3 Regulation of Cys Synthesis by Formation of the Cys Synthase Complex.
Adapted from (Julie et al., 2006)

Demonstration of the interaction between OASS and SAT in plants and bacteria has employed multiple approaches, including size exclusion chromatography, yeast two hybrid analysis, surface plasmon resonance, and fluorescence spectroscopy (Mino et al., 1999; 2000; Wirtz et al., 2001; Berkowitz et al., 2002; Bonner et al., 2005; Campanini et al., 2005).

On further analysing the protein–protein interaction regions in the Cys synthase complex it is found that the C terminus of SAT plays an very important part in association with OASS in plants and bacteria (Mino et al., 1999; 2000; Wirtz et al., 2001). From analysis of the structure of *Haemophilus influenzae* OASS (Hi-OASS) in complex with a peptide (a part of the C terminus of SAT from the same organism), and protein–protein interaction studies of *Arabidopsis* OASS (At-OASS) and At-SAT indicated that the OASS active site is a central theme in the SAT interaction site (Bonner et al., 2005; Huang et al., 2005). Changes in the C termini sequence of SAT and differences in the OASS active site may give evidence in support of the specificity for formation of the Cys synthase complex in different organism like plants and bacteria. These results provide new insights into the molecular mechanism underlying formation of the plant Cys synthase complex.

CHAPTER 6

Expression, Purification, Crystallization & Structural Analysis of EhOASS and in Complex with Cysteine

6.1 Introduction

The cysteine biosynthetic pathway contributes significantly for incorporation of inorganic sulfur into organic compounds. In bacteria and plants, L-cysteine is the precursor of most sulfur containing metabolites including methionine and glutathione. Sulfide reacts with O-Acetylserine, which is produced from serine and acetyl-CoA by serine acetyl transferase (SAT). This final reaction that forms L-Cysteine by transfer of the alanyl group of O-Acetylserine to sulfide is catalyzed by Cysteine Synthase (CS: O-acetyl-L-serine sulfhydrylase (OASS), EC 4.2.99.8).

In plants and bacteria, the intra cellular level of Cys is strictly regulated by two mechanisms. The first mode of control involves feedback inhibition of SAT by Cys, which regulates the SAT activity and thus production of Cys (Noji et al., 1998). Second mechanism involves association of OASS and SAT to form Cysteine synthase complex and this complex formation is favored under excess sulfur concentrations. In the complex form SAT activity increases while the OASS activity is curtailed (Saito et al., 1995; Droux et al., 1998).

SAT exists as hexamer in solution while OASS as a dimer. One SAT hexamer and two OASS dimers interact with each other to form the Cysteine synthase complex (Droux et al., 1998). Various biochemical studies and structural studies have revealed that the C-terminal end of SAT interacts with the active site of OASS (Mino et al., 1999; 2000; Huang et al., 2005; Francois et al., 2006; Schnell et al., 2007).

Entamoeba histolytica, the causative agent of human amebiasis, is an enteric protozoan parasite, causes amebic colitis and extra intestinal abscesses (e.g. hepatic, pulmonary, and cerebral) (WHO report, 1995) in approximately fifty million inhabitants of endemic

areas. Cysteine was shown to be the major thiol in *E. histolytica* (Nozaki et al., 1998) and therefore, was assumed to play an important role in oxygen defense mechanisms in this glutathione-deficient organism. Similar to other bacteria and plants, the major route of cysteine biosynthesis in this parasite is the condensation of *O*-Acetylserine with sulfide by the “*de novo* cysteine biosynthetic pathway” involving key enzymes Serine acetyl transferase (EhSAT) and Cysteine Synthase (EhCS) or *O*-Acetylserine Sulfhydrylase (EhOASS). Unlike in other bacteria and plants, EhSAT and EhOASS do not interact with each other and does not form Cysteine Synthase complex (Nozaki et al., 1999).

Various biochemical and structural studies on OASS from various sources have indicated that it undergoes structural transition along the reaction pathway (Schnackerz et al., 1995; McClure & Cook, 1994; Benci et al., 1999; Burkhard et al., 2000; Bonner et al., 2005). The smaller N-terminal domain was seen close to the large C-terminal domain in PLP cross linking Lys mutated and Met complexed OASS structures, better described as external aldimine bound structures (Burkhard et al., 2000; Bonner et al., 2005) compared to native structures (Burkhard et al., 1998; Bonner et al., 2005; Schnell et al., 2007) indicating that the enzyme is in a substrate-bound state while the reaction is taking place at the active site.

Here the crystal structure of native EhOASS at 1.86Å resolution and in complex with its product, Cysteine at 2.4Å resolution is discussed. Both biochemical and structural studies with cysteine indicate, the EhOASS shows binding affinity with its product cysteine. The Cysteine bound structure also shows the movement of N-terminal domain and this structure may represent the state of the EhOASS just before the product release. The extended N-terminal region is involved in the extensive dimeric interface in a domain-swapping manner, favoring formation of a strong dimer. And C-terminal helix positioning in EhOASS may influence its critical interactions with other proteins.

6.2 Materials and Methods

6.2.1 Over expression and purification

pET28a-EhCS was transformed into *Escherichia coli* strain BLR (DE3) cells, which were grown in LB media supplemented with 50 µg/ml kanamycin at 310 K. When the OD₆₀₀ reached 0.5, over expression of EhCS was induced by the introduction of 0.5 mM isopropyl β-d-thiogalactopyranoside (IPTG) into the culture and incubation at 303 K for an additional 3 h. The cells were harvested by centrifugation at 6000g for 5 min at 277 K, suspended in buffer A (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 100 mM PMSF) and lysed with 1 mg/ml lysozyme and 0.1 % Triton-X. After sonication on ice a clear supernatant was obtained by centrifugation at 15000 rev/min for 15 min at 277 K. The first chromatographic step involved the selection of His-tagged protein on a Ni-NTA column (Sigma-Aldrich) pre-equilibrated with buffer A (lysis buffer) and 5 mM imidazole. The column was washed with five column volumes of buffer B (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 30 mM imidazole) at 277 K. The bound protein was eluted as 1 ml fractions with buffer B containing 300 mM imidazole. The fractions with optimum absorbance at 280 nm were checked for homogeneity on SDS-PAGE (Laemmli, 1970), pooled and subjected to 40 % ammonium sulfate precipitation. The precipitant was dissolved in 0.5 ml buffer C (100 mM Tris-HCl pH 8.0, 50 mM NaCl, 10 mM β-mercaptoethanol). The final purification step involved gel filtration on a HiLoad Superdex 75G 16/60 column (GE Amersham Biosciences). The column was pre-equilibrated with buffer A (50 mM Tris, 150 mM NaCl, 10 mM β-mercaptoethanol pH 8.0). The purity of the protein was assessed on SDS-PAGE and the purified protein was concentrated using YM3K Centricon tubes (Amicon) to a final concentration of 15 mg/ml as estimated by the BCA assay (Smith et al., 1985). For the complete details of Competent cell preparation, Agarose gel electrophoresis, Transformation process, LB media, Culture growth condition, Solution preparation and SDS PAGE refer to Chapter 2 material method section.

6.2.2 Crystallization

Purified EhCS was concentrated to a final concentration of 15 mg/ml in 50 mM Tris-HCl pH 8.0 buffer containing 150 mM NaCl for crystallization trials. Crystallization trials were carried out at 289 and 277 K using the hanging-drop method by mixing equal volumes (3 μ l) of protein and reservoir solutions. The drops were equilibrated against 0.5 ml of the same precipitant solution. Several crystallization conditions were tested with PEG, ammonium sulfate and sodium malonate. After obtaining hit and optimization of these conditions with ammonium sulfate resulted in the production of good -quality single crystals (ranging in size from 0.3 x 0.2 x 0.2 to 0.5 x 0.2 x 0.2 mm) using 2.5 M ammonium sulfate as precipitant in 100 mM Tris pH 7.0. Elongated hexagonal shaped crystals of EhCS were obtained in 20 d at 289 K.

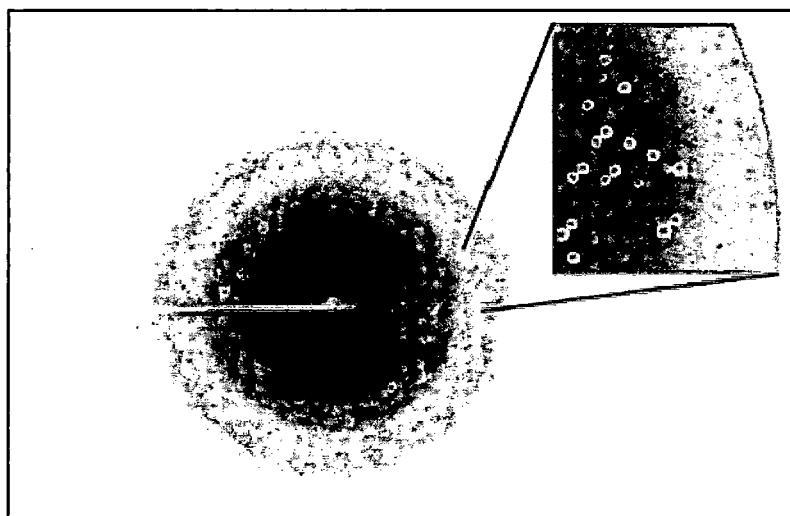
6.2.3 Co-Crystallization of EhOASS with Cysteine

The recombinant EhOASS was expressed and purified as described. The protein was concentrated to 15 mg/ml and 1M Cysteine stock solution was added to this protein to make final concentration of cystiene as 5 mM in 50 mM Tris buffer (pH 8.0) containing 150 mM NaCl for crystallization trails. Crystallization trials were carried out at 289K using the hanging drop method by mixing equal volumes (3 μ l) of protein and reservoir solution. The crystals of EhCS-Cys complex were obtained using 2.3 M ammonium sulfate as precipitant, 5 mM Cysteine in 100 mM Tris pH 7.2, which is almost similar to native EhCS crystallization condition.

6.2.4 Collection and processing of diffraction data (EhOASS)

A single crystal of EhOASS was separated from the drop using a cryoloop (Hampton Research). The crystal was then flash-cooled to 100 K in a nitrogen-gas stream. Prior to flash-freezing, the crystal was soaked sequentially for 15 s in reservoir solution supplemented with 5, 10, 15 and 20 % glycerol as a cryoprotectant. Diffraction data were collected on a MAR imaging plate using a Rigaku MicroMax-007 X-ray generator with X-ray optics to focus the beam (International Center of Genetic Engineering and

Biotechnology, New Delhi). The data were indexed, integrated and scaled using HKL-2000 (Otwinowski & Minor, 1997).



Diffraction pattern of EhCS crystals to 1.86 Å resolution. The data were collected using a Rigaku MicroMax-007 generator and a MAR imaging plate. The imaging plate was adjusted to a distance of 150 mm and the crystals were exposed for 90 s per frame. Diffraction spots were observed to the edges of the image plate, as shown in the enlargement.

6.2.5 Data collection and Processing Cys-EhOASS

The crystals of Cys-EhOASS complex were washed with reservoir solution to remove precipitate and were equilibrated in reservoir solution supplemented with sequential increase of 5, 10, 15 and 20 % glycerol. These crystals were picked up with cryoloop and then flash cooled to 100K in a nitrogen-gas stream. Diffraction data were collected on a MAR imaging plate using Rigaku generator (National Institute of Immunology, New Delhi). The data were indexed, integrated and scaled using HKL-2000 (Otwinowski & Minor, 1997).

6.2.6 Structure determination and refinement of EhOASS

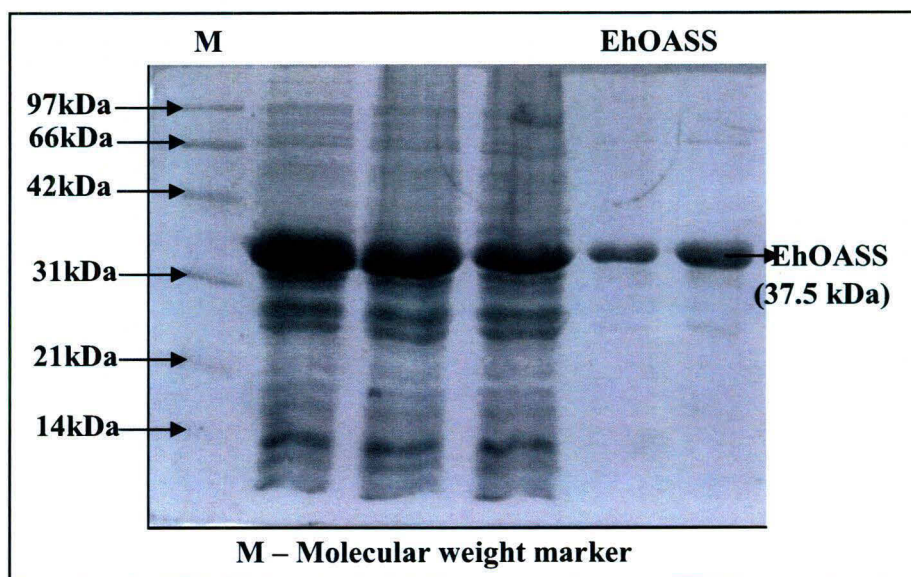
The EhOASS structure was determined by molecular replacement method using *Salmonella typhimurium* OASS as search model (1OAS). The protein was crystallized in P41 space group with two molecules in an asymmetric unit (Table 6.4.1). Initially the structure was refined as rigid body refinement with two rigid bodies as movable domain (residues from 64 to 165) and non-movable domain (residues from 1 to 63 and 166 to 336). Later the model was submitted to ARP/WARP for auto-building (Perrakis et al., 1999), which successfully built up 90 % of side chains in one molecule and 55 % of side chains in the other one with good quality electron density and 24 % R-factor. The remaining parts of the structure were built manually with COOT (Emsley and Cowton, 2004) and refined by iterative model building using COOT graphics package combined with REFMAC (CCP4; Murshudov, et al., 1997). The final model was refined well with good electron density (fig. 6.3.2) and crystallographic R factor as well as Free_R factor values (Table 6.4.1). The PLP molecules, water molecules and sulfate molecules were added manually, where justified by $F_o - F_c$ electron density at $>3\sigma$ contour level. The water molecules were initially picked up by COOT graphics package, later they were manually checked for electron density, justified by their hydrogen bond interactions with the protein. The final model of dimer consists of 5117, 30, 366 protein, PLP and water atoms respectively (Table 6.4.1).

6.2.7 Structure determination and refinement of EhOASS in complex with Cysteine

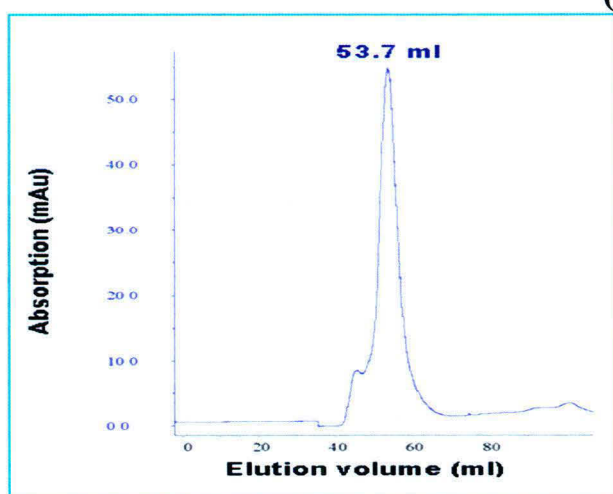
The Native EhCS structure was used as model for solving the structure of EhOASS-cysteine complex by the molecular replacement, and as expected the solution was with very good correlation coefficient (70.5 %) and R-factor (35 %). Initially the model was refined as rigid body refinement using CNS with three rigid bodies and the structure was further improved by iterative model building either by COOT graphics package (Emsley et al., 2004) or by O graphics package (Jones et al., 1991) combined with conjugate-gradient minimization with bulk solvent correction in CNS (Brunger et al., 1998). Initially in one monomer of the EhOASS dimer, the movable domain was not well

defined and extra difference Fourier map was observed from the active site to the end of movable domain. After consecutive cycles of refinement and model building, this region was completely retraceable into the new electron density keeping the Cysteine at the active site.

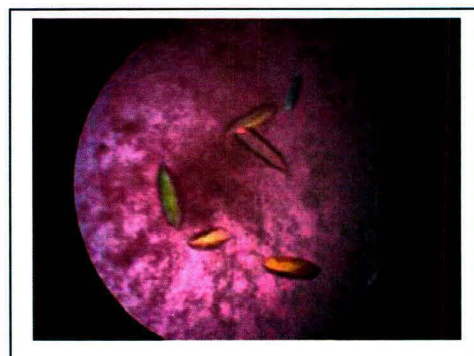
6.3 Results and Discussion



(a)



(b)



(c)

Fig. 6.3.1 (a) 12% SDS PAGE showing purified EhOASS. (b) GPC profile of EhOASS. (c) EhOASS crystals.

Table 6.4.1
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

Data Set	EhCS	EhCS-Cys
Crystallographic data		
X-ray source generator	Rigaku MicroMAX-007	Rigaku rotating anode
Wave length (Å)	1.54	1.54
Space group	P4 ₁	P4 ₁
Unit-cell parameters (Å)		
<i>a</i>	80.31	80.36
<i>b</i>	80.31	80.36
<i>c</i>	112.21	111.761
Resolution range (Å)	50-1.86	50-2.4
R _{sym} (%)	4.3	7.9(47.8)
Completeness(%)	99.8	99.7(99.4)
Total No. of observations	392808	109003
No. of unique observations	59571	27772
Redundancy	6.6	3.9
Average <i>I</i> / σ (<i>I</i>)	16.2 (5.55)	9.2(2.16)
Crystal mosaicity(°)	0.4	0.78
Refinement		
Resolution	30-1.86 (1.908-1.86)	50-2.4 (2.55-2.4)
R factor (%)	18.0 (21.5)	19.4 (29.6)
Free R factor (%)	21.15 (25.4)	23.5 (37.2)
Mean B factors	25.19	42.3
Number of atoms		
Protein/PLP/water/other	5117/30/366/10	4996/30/263/37
RMS deviations		
Bonds (Å)	0.015	0.007
Bond angles (°)	1.8	1.6
Dihedral angles (°)	23.9	23.2
Improper Angles (°)	1.9	0.86
Cross validated error	0.21	0.34

Values in parentheses are for last resolution shell.

Free_R factor was calculated with a subset of 5% randomly selected reflections.

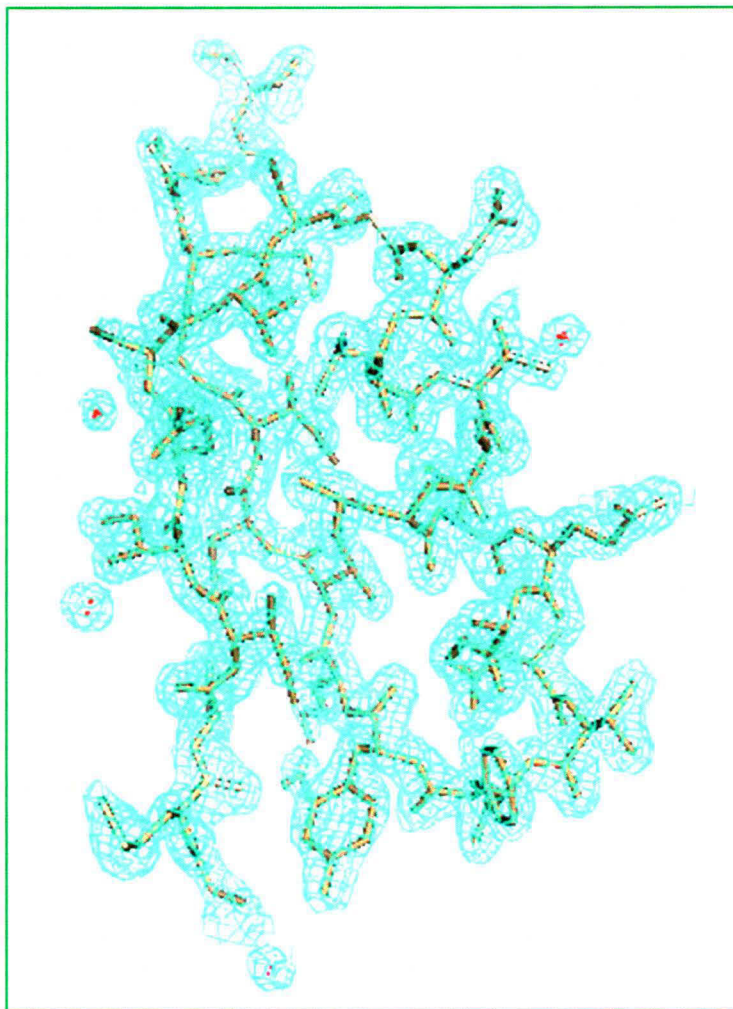


Fig. 6.3.2 Representative region the electron density map ($2F_o-F_c$) with a final model superimposed on it. The map was countered at 1.5σ .

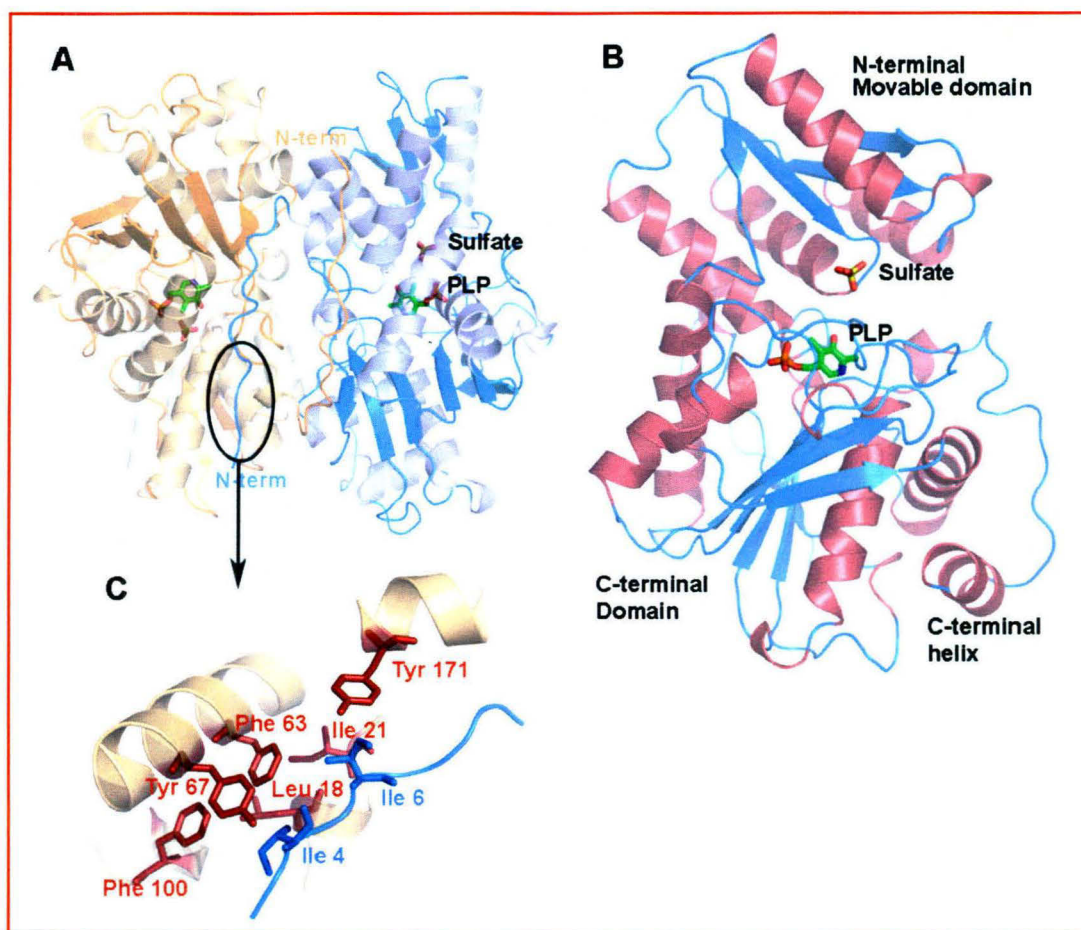


Fig. 6.3.3 Structure of *E. histolytica* OASS.

(A) Dimeric form of the native EhOASS seen in the asymmetric unit displayed as a ribbon diagram generated by PyMOL (DeLano, 2002), two molecules are shown in different shades with sulfate and PLP at the active site. Swapping of the N-terminal region can also be seen.

(B) Magnified view of the monomeric form of EhOASS, showing domain arrangement and active site cleft.

(C) The magnified view of the N-terminal region of one molecule (cyan) making several interactions with the other molecule.

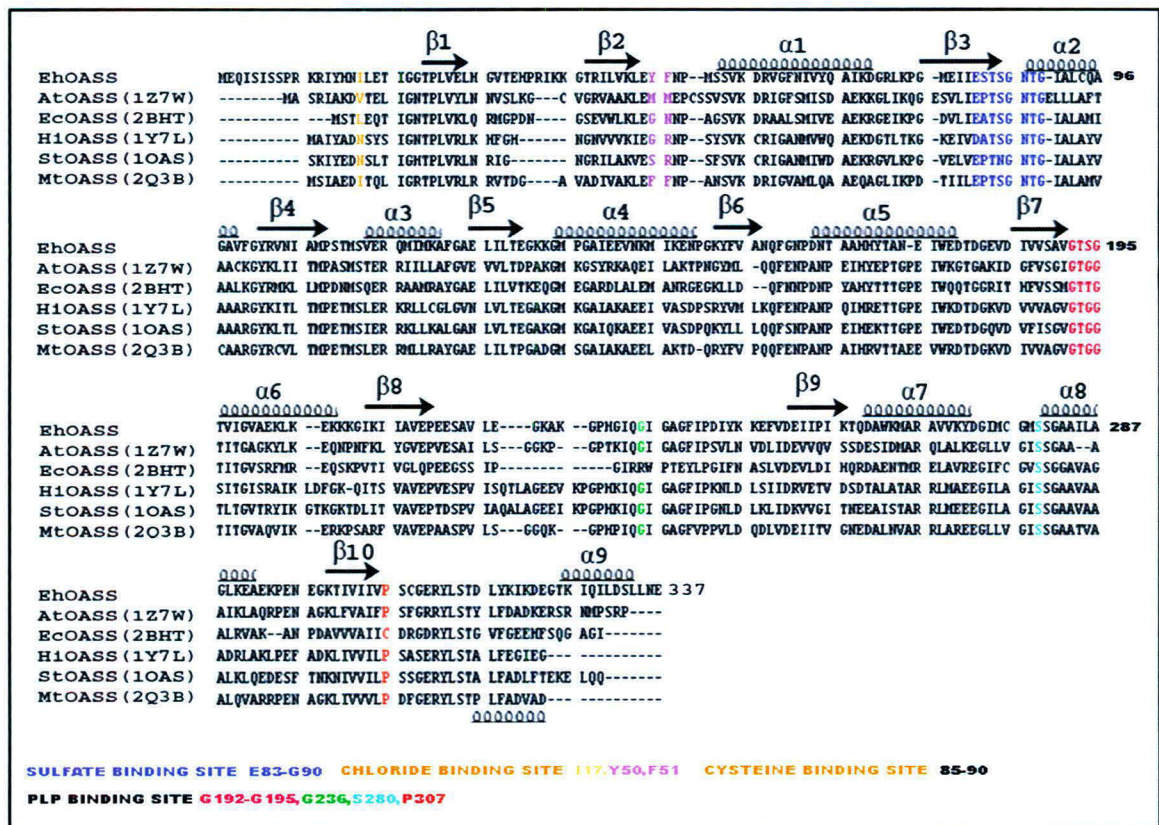


Fig. 6.3.4 Sequence homology of various OASS sequences. The secondary structures are indicated along the sequence according to the EhOASS. The helix $\alpha 9$ position in EhOASS and StOASS is different. The $\alpha 9$ positioned below represents the StOASS structure.

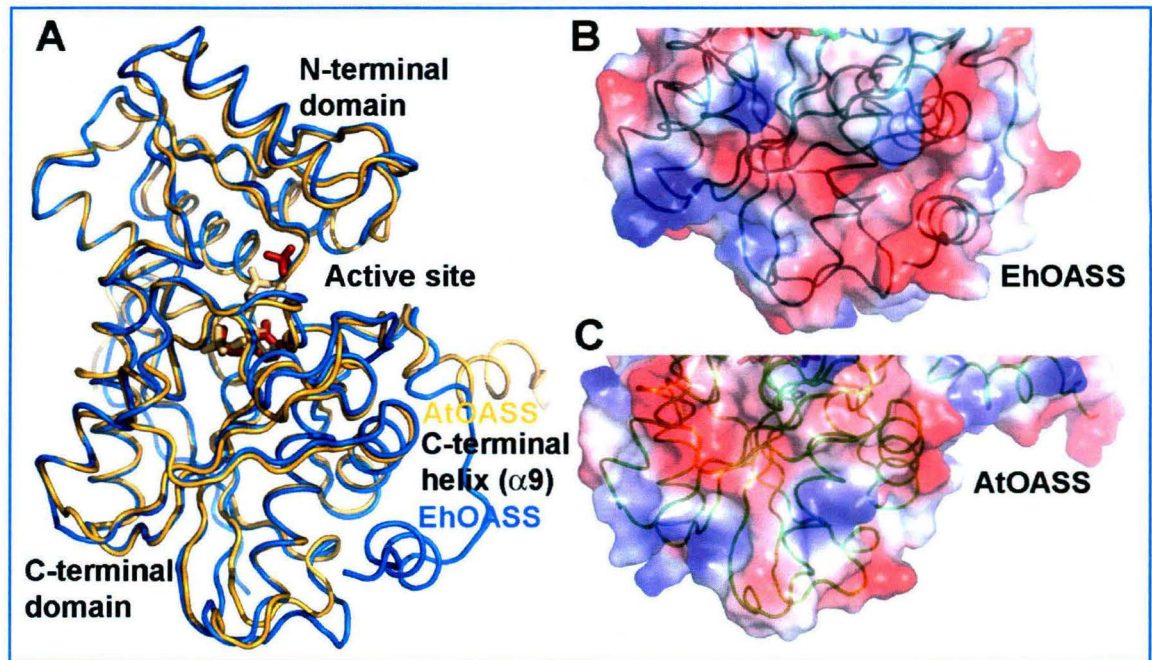


Fig. 6.3.5 Comparison of EhOASS and AtOASS and difference in the c-terminal helix position

(A) Schematic overlay of Native EhOASS (cyan) and AtOASS (gold), shows difference in orientation of C-terminal helix.

(B) The space filled charged surface of the C-terminal domain of EhOASS and (C) AtOASS. Due to the different orientation of C-terminal helix, the cleft in the AtOASS C-terminal domain is clearly seen and this cleft is filled by its own C-terminal helix in EhOASS.

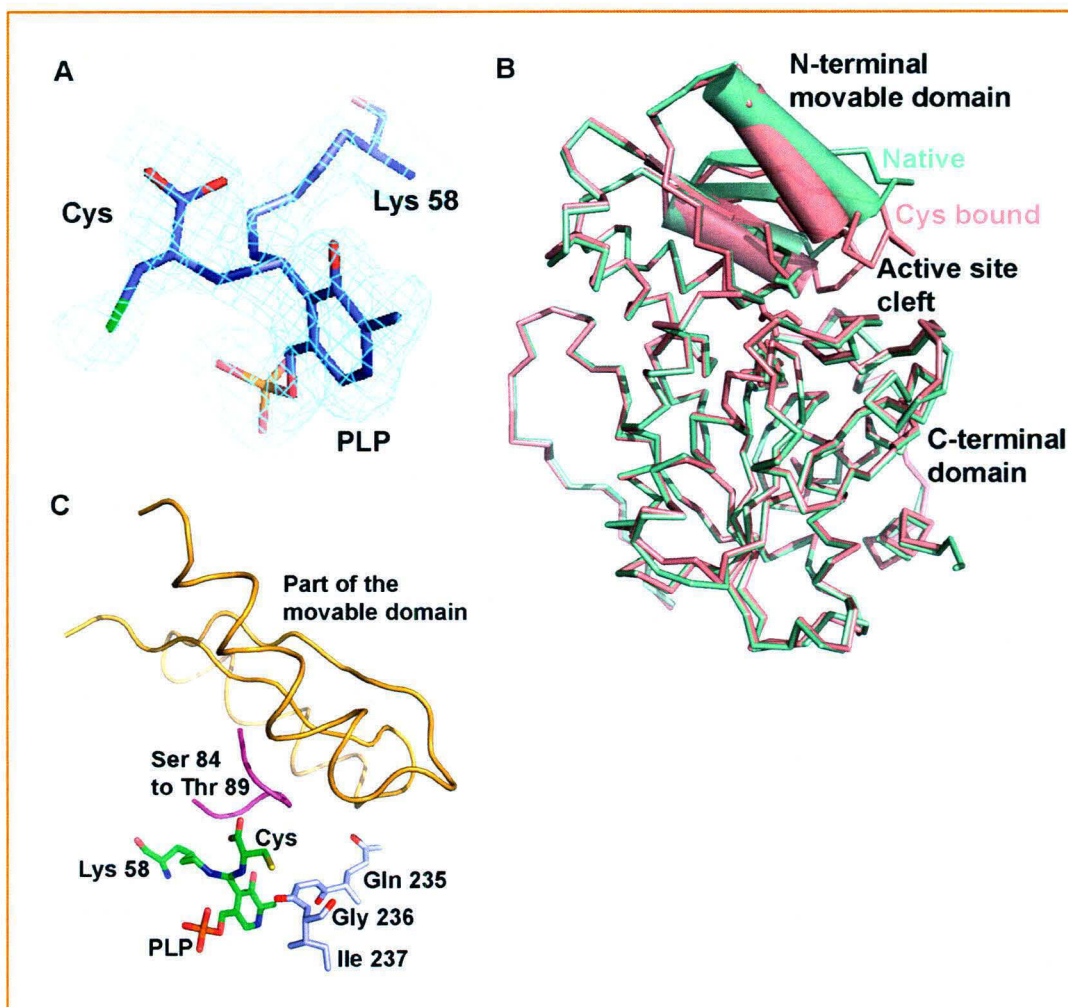


Fig. 6.3.6 Cysteine bound EhOASS

(A) The Cys, PLP and Lys at the active site of Cys-EhOASS complex are shown with the electron density (2Fo-Fc) map. The amino group of Cys is close ($\sim 2\text{\AA}$) to the reaction center of PLP and Lys.

(B) Comparison of native EhOASS with Cys bound EhOASS. The secondary structures of the movable domain, which had larger difference showed as cylinders and ribbons.

(C) Cys is shown at the active site with neighboring residues and part of the movable domain (102 to 150), which moves close to the active site.

6.5.1 Overall structure

Cysteine synthase from *E. histolytica* was expressed in BLR (DE3) cells, resulting in approximately 100% soluble protein. The Recombinant EhOASS was having a C-terminal six histidine tag. And it was purified by affinity and size-exclusion chromatography. The sample was 99 % pure as estimated by SDS-PAGE [(fig. 6.3.1) (a)]. The purified EhOASS was yellow in color indicating bound PLP (pyridoxal 5'-phosphate) to the protein and crystals grown were also yellowish. The retention time of the protein on the size-exclusion column indicated that the protein is a dimer of ~75kDa [fig. 6.3.1 (b)]. Many drops containing ammonium sulfate as a precipitant at 289 K developed phase separation in 1 or 2 d. Optimization of these conditions with ammonium sulfate resulted in the production of good -quality single crystals (ranging in size from 0.3 x 0.2 x 0.2 to 0.5 x 0.2 x 0.2 mm) using 2.5 M ammonium sulfate as precipitant in 100 mM Tris pH 7.0. Elongated hexagonal shaped crystals of EhCS were obtained in 20 d at 289 K [(fig. 6.3.1 (c)]. The crystals belong to the tetragonal space group P41, with unit cell parameters $a = 80.3$, $b = 80.3$, $c = 112.2 \text{ \AA}$. A complete X-ray diffraction data set of about seven fold redundancy was collected to 1.86 \AA resolution. The volume of the asymmetric unit allows the presence of a dimer, giving a Matthews volume (V_M) of $2.5 \text{ \AA}^3 \text{ Da}^{-1}$ and a solvent content of 50.7 % (Matthews, 1968). This is consistent with the solution obtained by molecular replacement and allowed us to assign the space group as P41. The final model was well refined to R and Free_R of 18 % and 21.1% respectively (Table 6.4.1).

The EhOASS structure reveals that the poly peptide folds into two sub units [(fig. 6.3.3 (b))] with each sub unit consisting of a α/β domain. These two domains are described as a smaller N-terminal movable domain and a larger C-terminal non-movable domain (Burkhard et al., 2000). The smaller N-terminal movable domain (residues 57 to 164) consists of four β -sheets surrounded by four α -helices. The larger C-terminal domain (1 to 56 and 165 to 336 residues) folds into six β sheets surrounded by four major α -helices. The reaction center is located in between these two domains and the cofactor PLP cross linked to Lys 58 is located between these domains.

Each asymmetric unit consists of two molecules of EhOASS forming a dimer [(fig. 6.3.3 (a)]. These two molecules are related to each other by 2-fold non crystallographic symmetric axis. The extended N-terminal region and the first two β -strands belonging to the C-terminal domain make extensive interactions at the interface of two molecules in the dimer. The dimer interface buries 2952 Å² per monomer or about 19 % of the surface area of the monomer. The dimeric interface consists of mixture of hydrophobic and hydrogen bond interactions.

6.5.2 Comparison with other structures

The overall structure of EhOASS is similar to *Salmonella typhimurium* OASS (StOASS) (Buckard et al., 1998; 1999; 2000), *Arabidopsis thaliana* OASS (AtOASS) (Bonner et al., 2005; Francois et al., 2007), *E.coli* OASS (EcOASS) (Claus et al., 2005), *Haemophilus influenzae* OASS (HiOASS) (Haung et al., 2005) and *Mycobacterium tuberculosis* OASS (MtOASS) (Schnell et al., 2007). EhOASS shows 50 %, 47 %, 46 %, 43 % and 39 % sequence identity to MtOASS, StOASS, AtOASS, HiOASS and EcOASS respectively. EhOASS resembles StOASS (pdbcode - 1OAS) with an r.m.s deviation of 1.03 Å for 299 C α atoms, AtOASS (pdbcode - 1Z7W) with 0.9 Å for 302 C α atoms, EcOASS (pdbcode - 2BHT) with 1.17 Å for 242 C α atoms, HiOASS (pdbcode - 1Y7Y) with 0.96 Å for 284 C α atoms and MtOASS (pdbcode - 2Q3B) with 0.88 Å for 297 C α atoms. Even though EhOASS exhibits more sequence identity (47 %) with StOASS, it is structurally more similar to AtOASS, MtOASS and least similar to EcOASS. The structures of AtOASS and StOASS have been determined in three different states and we further compared EhOASS with the structures of all available states of AtOASS and StOASS. The StOASS structure is determined in native form, i.e open conformation (Buckard et al., 1998; 1OAS), external aldemine bound, closed conformation (Bukhard et al., 1999; 1D6S) and inhibited form or chloride / sulfate bound form (Bukhard et al., 2000; 1FCJ). The EhOASS structure is more similar to the native, open conformation with rmsd of 1.03 Å for 299 C α atoms and least similar to aldemine bound closed conformation with rmsd of 1.28 Å for 279 C α atoms. Interestingly, even though the EhOASS is bound to sulfate in the crystal structure, it is more similar to native structure of StOASS rather than the

sulfate and chloride bound form of StOASS. The structure of OASS from *Arabidopsis thaliana* (AtOASS) is also determined in three different states, in complex with sulfate (Bonner et al., 2005; 1Z7W), bound with external aldemine (lys mutated protein in complex with Met) (Bonner et al., 2005; 1Z7Y) and complexed with C-terminal peptide of SAT (Francois et al., 2006; 2ISQ). Unlike different states of StOASS structure, the various forms of AtOASS are more similar to each other. In comparison with these structures EhOASS is more similar to the sulfate bound form of AtOASS. Overall EhOASS is more similar to native or open state of both StOASS and AtOASS. Even though the overall structure is similar the major differences were observed in chloride binding site, N-terminal region and the C-terminal region.

Table 6.4.2 : Surface area buried at the interface of the dimer

	Area at the interface	% of surface area buried by each monomer
EhOASS (present structure)	5905 Å ²	19.1
StOASS (1OAS)	4785 Å ²	17.3
AtOASS (1Z7W)	5683 Å ²	19.6
EcOASS (2BHT)	2960 Å ²	11.1
MtOASS (2Q3B)	4260 Å ²	17.1
HiOASS (1Y7L)	3821 Å ²	13.8

6.5.3 Dimer interface and Extended N-terminal region

The extended N-terminal region of EhOASS is part of the dimeric interface and contributes in making it as the largest interface buried in the dimeric form of EhOASS compared to other OASS structures (Table 6.4.2). AtOASS has the second largest interface area contributed from extended C-terminal helix. The N-terminal part of the EhOASS is extended by about 10 extra residues compared to other OASS structures and these residues participate in interesting domain swapping dimeric interactions. The

dimeric interface consists of mixture of hydrophobic and hydrogen bonds dominated by some crucial hydrophobic interactions. The extended structure of N-terminal region is well traced, as it makes large number of interactions with the N-terminal domain of its dimeric counterpart. Ile4 of one monomer enters into the hydrophobic pocket surrounded by Leu 18, Phe 63, Tyr 67 and Phe100 of another monomer [(fig. 6.3.3 (c)]. Similarly Ile6 of one molecule interacts with Ile 21, Phe 63 and Tyr 171 of another molecule. The N-terminal region also makes various hydrophilic interactions like carbonyl groups of Ile 6 and Ser 5 makes hydrogen bonds with Arg12 of another molecule. There are many structured water molecules between two molecules making lot of inter molecular interactions. These extended dimeric interface interactions should contribute significantly in favoring the formation of stable EhOASS dimer.

6.5.4 C-terminal region

The C-terminal helix ($\alpha 9$) of EhOASS is quite differently/distinctly oriented compared to AtOASS structure, while in other OASS structures it is not well traced. The Sequence homology is also least in this region and the helix position in the sequence is not conserved (fig. 6.3.4). In AtOASS the C-terminal helix starts from 320 and ends at 327 (according to EhOASS numbers), where as in EhOASS the C-terminal helix begins at 326 and ends at 334. This helix in EhOASS ($\alpha 9$) is positioned in a groove formed between $\alpha 7$, $\alpha 8$ and the loop Gly 31 to Lys 40, which is part of the C-terminal domain. Whereas in AtOASS $\alpha 9$ is not making any intra molecular interactions rather it makes several inter molecular interactions with the N-terminal domain of the neighboring molecule, which is involved in dimeric interactions. In EhOASS, the interactions between $\alpha 9$ and rest of the groove are dominated by hydrophobic interactions and further stabilized by few hydrogen bonds and salt bridges. The hydrophobic residues 330 Ile, 331 Lue, 334 Leu and 335 Leu from C-terminal helix interact with hydrophobic patch consisting of Val 32, Met 265, Ala 268, Tyr 272, Leu 286 and Leu 289 (fig. 6.3.5) in the groove of the C-terminal domain. These interactions are further stabilized by hydrogen bonds and salt bridge between Ser 333 - Lys 264, Asp 332 – His 35 and Lys 327 – Glu 34.

The hydrophobic patch in the groove of the C-terminal domain formed by helix $\alpha 7$ residues and hydrophilic residues contributed by loop (31-40) in EhOASS, which are not conserved in other OASS sequences. These differences in the C-terminal groove may be responsible for loss of interactions with C-terminal helix and may contribute for altered orientation of the C-terminal helix in other OASS structures. This groove in MtOASS is bound to MPD (2-Methyl 2, 4 pentanediol), which further confirms that the groove can participate in protein-protein or protein-ligand interactions. The exposed groove in other OASS structures may contribute to the protein-protein interactions, like interactions with SAT and these interactions may be absent in the case of EhOASS.

6.5.5 Active site

The PLP (Pyridoxal Phosphate) binding site is highly conserved in all OASS structures and it is positioned in between N-terminal and C-terminal domains. PLP is covalently linked to the side chain of Lys 58 in C-terminal domain of EhOASS and the aromatic ring of PLP is positioned in between three loops from C-terminal domain, interacting with Gly 236, Ser 280 and Pro 307. The N1 atom of the pyridine ring forms a hydrogen bond with the -OH group of Ser 280. Ser 280 and surrounding residues of the loop are well conserved in all OASS structures, while Gly 236 and Pro 307 are conserved in most of the OASS structures except in EcoASS, where the active site is modified to accommodate the larger substrates (Claus et al., 2005). The phosphate group of PLP forms several hydrogen bonds with glycine rich loop ($^{191}\text{GTSGT}^{195}$) in between $\beta 7$ and $\alpha 6$, and three structured water molecules. In this loop 1st, 2nd and 4th residues are well conserved and other two residues are replaced with homologous residues. The 3' hydroxyl group of PLP ring makes a hydrogen bond with the Asn 88 from a well conserved loop between $\beta 3$ and $\alpha 2$ of N-terminal domain. This loop from Glu 83 to Gly 90 is also involved in sulfate binding, where sulfate makes two hydrogen bonds with backbone nitrogen of Ser86 and Gly 87. Sulfate also participates in hydrogen bond with the side chain of Gln159 from N-terminal domain. Overall OASS structures are well conserved in the active site with similar PLP and sulfate binding sites except few minor

differences in the PLP binding loops.

6.5.6 SAT binding site

The biochemical and structural studies of OASS and SAT has revealed that the C-terminal end of SAT interacts with active site of OASS and forms a large Cysteine synthase complex in various organisms (Huang et al., 2005; Francois et al., 2006; Schnell et al., 2007). These enzymes are strictly regulated by feedback inhibition of SAT by Cysteine and SAT exhibits inhibitory association with OASS (Noji et al., 1998) by closing the active site. Interestingly in the case of *Entamoeba histolytica*, Nozaki et al (1999) showed by various biochemical studies that SAT does not interact with OASS and they do not form a Cysteine synthase complex.

Various structures like HiOASS, StOASS and MtOASS with small C-terminal peptide of respective SAT's have revealed that the C-terminal Ile binds at the active site of the OASS just besides the PLP (Huang et al., 2005; Francois et al., 2006; Schnell et al., 2007). At the C-terminal end of SATs only the hydrophobic amino acid Ile is most conserved. The C-terminal peptide of SAT interacts mainly at the active site of OASS, specifically with the loop between $\beta 8$ and $\beta 9$ and with an α -helix $\alpha 4$ at the active site (Francois et al., 2006). Sequence comparison of various OASS reveals that the active site region is over all conserved except the loop between $\beta 8$, $\beta 9$ which exhibits high sequence variability (fig. 6.3.4) but it is structurally conserved. The sequence variations in the loop between $\beta 8$, $\beta 9$ in OASS and in the C-terminal region of SAT may be responsible for specific binding and recognition between the respective partners in order to form Cysteine synthase complex.

In the case of *Entamoeba histolytica*, the C-terminal end residue Ile in SAT and the active site structure of OASS are well conserved, also various ligand binding affinity is comparable to other OASSs. The structure of EhOASS-Cysteine complex (fig. 6.3.6) shows that Cysteine binds at the active site of EhOASS, similar to external aldimine (Met) bound structures. The Ile at C-terminal end of EhSAT can also bind in similar way

almost at the same location as Cys, to form the EhOASS-EhSAT (Cysteine synthase) complex. Interestingly Nozaki et al., (1999) showed that these two enzymes do not interact with each other in this organism. There are two possible structural explanations for this; 1) the loop between $\beta 8$ and $\beta 9$ of EhOASS and the C-terminal region of the SAT may not complement each other. 2) The groove formed between $\alpha 7$, $\alpha 8$ and N-terminal loop (31 to 40) in various OASS may participate in making further interactions with SAT but in EhOASS this groove is occupied by its own C-terminal helix (fig. 6.3.5). Due to the lack of availability of this groove, the OASS-SAT interactions may be very weak in *Entamoeba histolytica* disabling them in forming the Cysteine synthase complex.

6.5.7 Chloride binding site

Unlike StOASS both EhOASS and AtOASS do not have a chloride binding site at the dimeric interface. In StOASS, the binding of chloride strengthens dimer formation and this is also evident from its chloride binding studies (Tai et al., 2001a). The chloride binding residues are not conserved in these structures (fig. 6.3.4). In StOASS Ser, Arg and Asn, which participate in chloride binding, are replaced by large or hydrophobic residues like Tyr 50, Phe 51 and Ile17 in EhOASS and equivalent residues in AtOASS are changed to Met, Met and Val. The replacement of hydrophilic residues in StOASS by hydrophobic residues in EhOASS and AtOASS structures, incapacitate them in chloride binding. Hence the dimer formation is unaffected by chloride in EhOASS and AtOASS. In these two structures the dimeric interface is larger than the other structures (Table 6.4.2), thus this extended dimeric interface may be compensating for the loss of chloride binding site in forming strong dimers.

6.5.8 EhOASS in complex with Cysteine

The EhOASS-Cysteine complex is also crystallized as dimer in $P4_1$ space group, with one dimer or two EhOASS molecules in an asymmetric unit. Only one EhOASS molecule was seen bound to Cysteine and other molecule was observed with Sulfate in the active site, bound in the similar manner as observed in the native structure. Cysteine is well placed with over 75 % occupancy with good electron density [fig. 6.3.6 (a)] causing

structural changes in the movable N-terminal domain. The structure of cysteine bound molecule is similar to the native structure with rmsd of 0.66 Å for 317 C α atoms, except the crucial differences starting from active site region to one end of N-terminal movable domain. Part of the movable domain (64 to 165) is twisted about 15 degrees, which causes closure of the active site [fig. 6.3.6 (b)]. The closure of the active site is located at the middle of β -sheets in the movable domain, where the twisting of β -sheets is centered. The distance between C α of 231 Pro from C-terminal domain and C α of 133 Lys movable N-terminal domain changed from 15 Å to 7 Å after the cysteine bound at the active site. Overall structure of cysteine bound EhOASS structure is very similar to the Lys mutated; Met complexed or external aldemine bound StOASS structure (Buckard et al., 1999).

In the Cys bound structure at the active site, the amino group of cysteine is in close proximity with C4A of PLP and amino group of Lys 58 [(fig. 6.3.6 (a)]. But PLP still seems to be bound to Lys 58 in the EhOASS-Cys complex, as indicated by inter atomic distance between C4A of PLP and amino group of Lys 58 (1.76 Å), while it is at about 2 Å distance with amino group of Cys. The amino group of Cys also interacts with backbone carbonyl group of Gly 236 (2.8 Å) at the active site. The carboxy group of Cys makes several hydrogen bonds with side chains of Thr 85 (2.53 Å), Ser 86 (3.4 Å) and is also located in such a way that four backbone amide groups of the turn formed by residues from 85 to 89 (TSGNT) are very close to it (3.3 - 3.8 Å) at the active site [(fig. 6.3.6 (c)]. These residues are well conserved among the cysteine synthase family. The close proximity between amino group of Cys with C4A of PLP makes the conditions potentially favorable for forming a covalent bond between the two or it can be a snap shot depicting the beginning of dissociation between the product (Cys) and cofactor (PLP).

The residues Thr 85 and Ser 86 move close to the Cys in the EhOSS-Cys complex as compared to the native structure in order to form hydrogen bonds with carboxy group of Cys. The twist in the β 3 of the movable domain accommodates this movement and to complement this movement β -sheet 4 and 5 also undergo twisting, which consequently causes the dislocation of helix 4. Overall this twist in the β -sheets as well as the shift in

α -helix causes the closure of active site in EhOASS-Cys complex [fig. 6.3.3 (b)]. By moving the N-terminal domain close to C-terminal domain, it closes the active site and protects from bulk solvent. The closure of the active site also helps in holding the reactants close to each other to form the product. Once the product is formed either the affinity towards the final product is reduced or the affinity for the reactants should be higher to replace the final product. Thus this structure may represent the snap shot of the EhOASS just before releasing the final product Cysteine.

Protein Data Bank accession codes

Atomic coordinates and structure factors for the EhOASS and its complex with Cysteine have been deposited with RCSB Protein Data Bank with PDB IDs: 2PQM and 3BM5.

6.5.9 Conclusion

Entamoeba histolytica uses Cysteine as a major reducing agent, which plays a vital role in the survival mechanism of this organism, including its antioxidative defense, matrix adhesion, elongation, motility and growth in this glutathione deficient organism. In this organism the OASS and SAT interaction is missing and thus the OASS inhibition by SAT is absent. There could also be possibility that some or the other kind of regulation be existing in this organism which could regulate cysteine metabolism there by helping *Entamoeba* in surviving the oxygen stress of the host.

CHAPTER 7

Spectral and Kinetic Studies of EhOASS

7.1 Introduction

7.1.1 Cysteine Biosynthesis

The biosynthesis of L-cysteine in enteric bacteria, such as *Salmonella typhimurium* and *Escherichia coli*, and in plants proceeds via a two-step pathway. The amino acid precursor of L-cysteine is L-serine, which is first acetylated at its β -hydroxyl group by acetyl-CoA to give O-acetyl-L-serine (OAS), a reaction catalyzed by the enzyme serine acetyltransferase (SAT; E.C. 2.3.1.30) (Kredich et al., 1966). The final step in cysteine synthesis, replacement of the acetate side chain, is catalyzed by OASS (E.C. 4.2.99.8) with contribution from inorganic sulfide as the thiol donor (Kredich et al., 1966).

7.1.2 Active site changes in OASS during cysteine biosynthesis

The enzyme active site contains PLP linked to a lysine as an internal Schiff base. Binding of O-acetylserine displaces the lysine and initiates the first half-reaction, yielding alpha-aminoacrylate intermediate linked to PLP. This leads to shifting of the absorption maximum to 470 nm (Cook & Wedding, 1976). The second half-reaction involves sulfide addition to the intermediate, thereby generating an external aldimine with the amino acid. The external aldimine formation also brings about a change in absorption maximum wherein λ_{\max} shifts to 418 nm (Cook & Wedding, 1976). The active site lysine reacts with this intermediate, releasing cysteine and regenerating the Schiff base (Cook et al., 1992; Schnackerz et al., 1995). In one of the studies it has been proven that binding of methionine, a substrate analog, with *Salmonella* OASS brings about an active site loop rearrangement. But binding of methionine with *Salmonella* OASS as an external aldimine does not complete the reaction cycle (Burkhard et al., 1998; Burkhard et al., 1999). So various kinds of rearrangement of the active site loop of OASS take place upon binding with various ligands. These rearrangements result in the shifting of the OASS spectrum peak to different wavelengths depending upon the type of ligand interaction

with the OASS. In the light of the above mentioned finding here also it has been investigated what sort of EhOASS active site loop rearrangement takes place on binding with different ligands.

7.1.3 Kinetic mechanism of OASS

The OASS-A has previously been shown to have a ping pong kinetic mechanism that requires the elimination of acetate from OAS in the first half-reaction to generate alpha - aminoacrylate in Schiff base with the active site PLP (Cook & Wedding, 1976). The Michael addition of sulfide to the alpha - aminoacrylate intermediate then occurs in the second half reaction to produce the final product L-cysteine. Two dead end complexes, E:sulfide and F:OAS, were observed consistent with the ping pong nature of the mechanism (where F is alpha - aminoacrylate in Schiff base with the active site PLP). In this study an effort has been made to understand what kinetic mechanism EhOASS employs to produce cysteine from its reactants.

7.1.4 OASS inhibition

In case of *Salmonella* OASS a number of studies has been carried out with respect to sulphate as an inhibitor. In these studies by varying one of the two substrate that is either OAS or TNB and keeping ammonium sulfate fixed various plots have been obtained. In the light of these experiments it was shown that sulfate was potent inhibitor of EhOASS with complex kinetics (Tai et al., 2001a). The kinetics with each of the variable was highly complex. In the present study kinetics of sulfate with respect to EhOASS is being deciphered.

7.2 Materials and Methods

7.2.1 Sources of materials

Chemicals - O-Acetyl-L-serine, DTNB and DTT were obtained from Sigma. The DTNB was dissolved in ethanol. The reduced form of DTNB that is TNB was prepared fresh prior to use by reduction with DTT in slight molar excess to DTNB. The TNB was then used without further treatment. Neither reduced nor oxidized DTT has any effect on the OASS reaction. All other reagents and chemicals were obtained from commercially available sources and were of the highest quality available.

7.2.2 Over expression and purification of EhOASS

For details of the over expression and purification of EhOASS refer to the chapter 6 materials and methods section.

7.2.3 Ligand Binding and Spectrum Studies

Binding of O-acetylserine, cysteine, and methionine was measured by monitoring the change in absorbance of the PLP using a Amersham biosciences Ultraspec 2100 UV-visible spectrophotometer. Titration of AtOASS (200 μ g) with O-acetylserine (0–220 μ M) was performed in 0.1 M Mes (pH 6.0). Titration of EhOASS (200 μ g) with cysteine (0–20 mM) or methionine (0–20 mM) was performed in 0.1 M Hepes (pH 7.5). All UV-visible spectra were obtained using Amersham biosciences Ultraspec 2100 UV-visible spectrophotometer and recording the absorbance at wavelengths from 240 to 600 nm using a light path of 1 cm. The K_d value for each ligand was calculated by fitting the data to a reversible two-state model of binding, $\Delta A = (\Delta A_{\max} [L]) / (K_d + [L])$, using Graph Pad prism software, where ΔA is the change in absorbance at a given wavelength in the presence of ligand at concentration $[L]$. The data were also plotted as a linear transform of the same equation ($1/\Delta A$ versus $1/[L]$).

7.2.4 Kinetics and Ammonium Sulphate inhibition studies

All spectra were collected at 25°C in reaction cuvettes with 1 cm in path length and 400 micro liters in volume. Absorption spectra were measured utilizing a Amersham biosciences Ultraspec 2100 spectrophotometer. Steady-state kinetic parameters were determined by initial velocity experiments, in which product formation was linear over the time monitored (1–10 min) using the above assay with varied concentrations of O-acetylserine (0–8 mM).

Measurements were made in 100 mM Hepes, pH 7. All assay buffers used in these studies were titrated with NaOH. The protein concentration of the purified enzyme was determined from BCA assay (Smith et al., 1985).

A convenient method for the assay of EhOASS activity was developed on the basis of the use of the chromophoric TNB as a substrate for the sulfhydrylase. The system consists of Amersham Biosciences ultraspec 2100 with a 412 nm wavelength filter. Reaction mixtures were 400 microliter in volume in a fluorescent cuvette with the following final reaction component concentrations: 100 mM Hepes, pH 7; 100 micomolar TNB; and variable OAS concentration. The reaction was initiated by the addition of Eh OASS to the sample including water, buffer (Hepes, pH 7), sample, and TNB. Spectra were recorded prior to each of the experiments using the appropriate buffer solution as a blank and the blank solution was having the entire component except the enzyme itself. The 10 scans in each experiment are collected at specific intervals. The interval is of one minute.

The disappearance of TNB was followed spectrophotometrically at 412 nm (Cook et al, 1991). The disappearance of TNB at 412 nm resulting from formation of S-(3-carboxy-4-nitrophenyl)cysteine, was monitored continuously as the reaction proceeded. Data were collected using Amersham Biosciences ultraspec 2100. A typical assay in a final volume of 400 microliter contained the following: 100 mM Hepes pH - 7, OAS - 2 mM; TNB - 0.1 mM. The reaction was initiated with the addition of OASS. Initial rates were calculated using a ϵ_{412} of $13\ 600\ \text{M}^{-1}\ \text{cm}^{-1}$ for TNB (Ellman, 1959).

7.2.5 Initial Velocity Studies

Initial velocity patterns were obtained by varying one reactant over a range of concentrations less than K_m and at fixed concentrations of the second reactant. Dead-end inhibition patterns were obtained by measuring the initial velocity at varying concentrations of one reactant with the second maintained at fixed level, at several different fixed concentrations of inhibitor including zero.

O-Acetylserine sulfahydrylase activity was monitored using TNB as the nucleophilic substrate. The intrinsic absorbance of TNB and its disappearance was monitored continually at 412 nm (ϵ_{412} , $13600 \text{ M}^{-1}\text{cm}^{-1}$) using a recording spectrophotometer. All assays were performed in 100 mM Hepes, pH 7, at 25°C. Initial rates were obtained at varying concentrations of OAS with the other reactant fixed and at different fixed concentrations of sulfate (including zero).

7.3 Results and Discussion

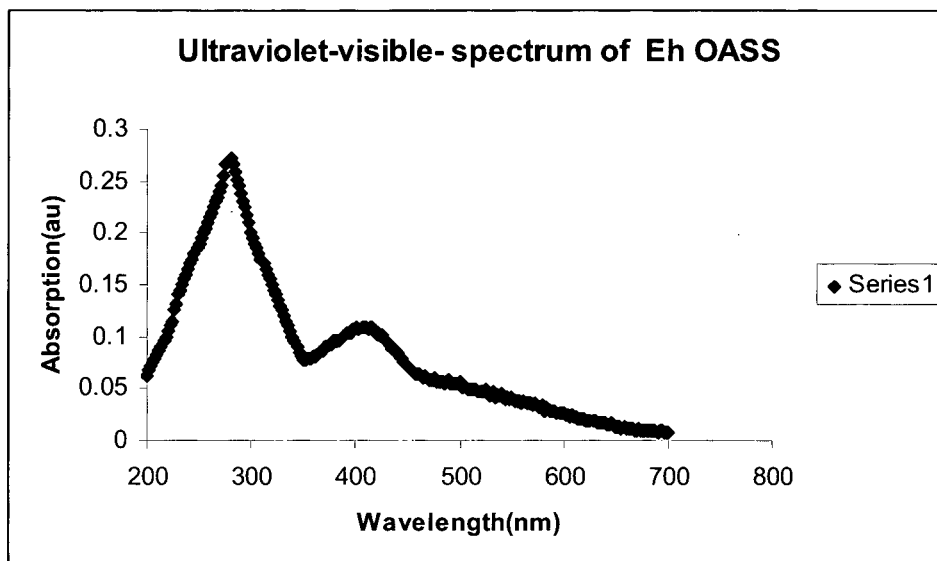


Fig. 7.3.1 Ultraviolet-visible spectrum of EhOASS at pH 7,100 mM Hepes.

Absorption spectra for EhOASS

The absorbance spectrum for OASS gives two peaks first at 280 nm and second at 412 nm (fig. 7.3.1). The absorption at 280 nm is because of the presence of aromatic amino acids in the sequence of EhOASS and absorption of 412 nm is because of the internal Schiff's base formation between the active Lysine and active PLP of EhOASS.

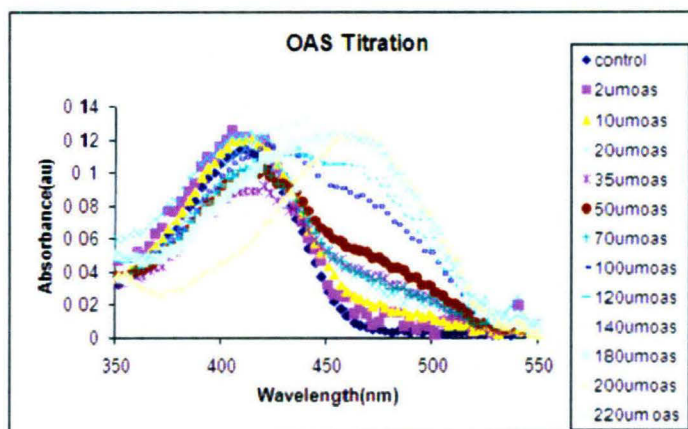


Fig. 7.3.2 (A) Spectrum of EhOASS with various concentrations of OAS.

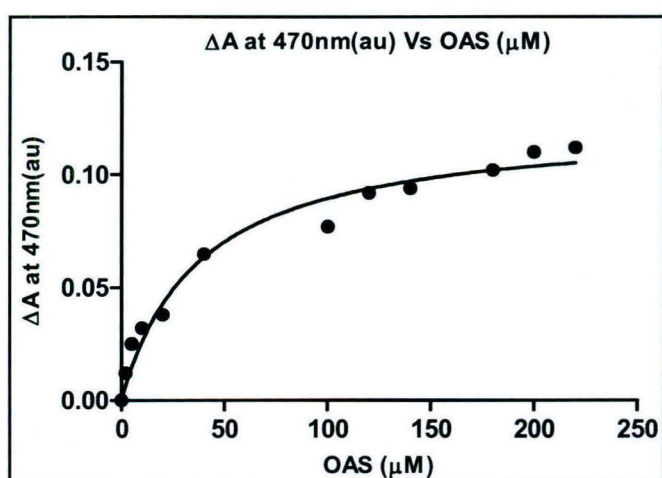


Fig. 6.3.2 (B) The plot of EhOASS between changes in the Abs. at 470 nm versus OAS conc.

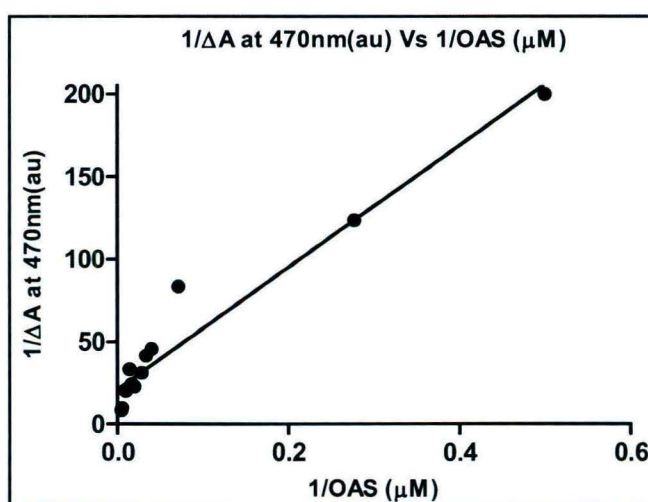


Fig.6.3.2(C)EhOASS Double reciprocal plot between changes in the abs. at 470nm Vs OAS conc.

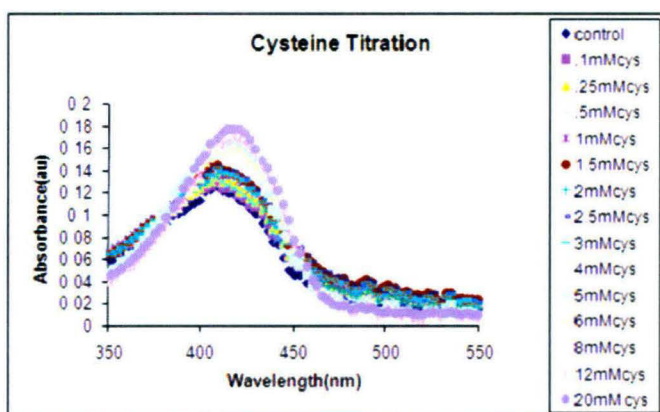


Fig. 6.3.2 (D) Spectrum of EhOASS with various concentrations of Cys.

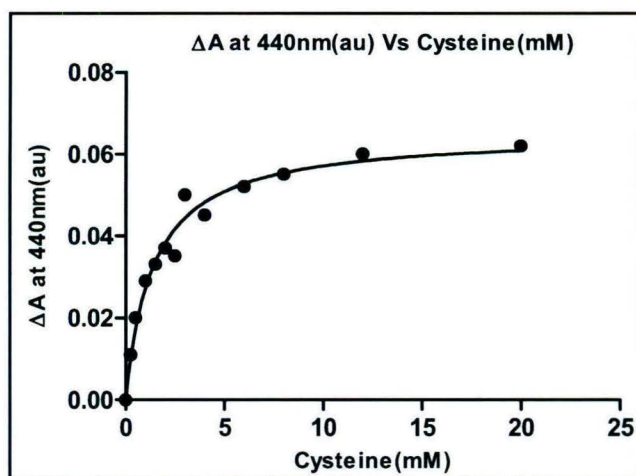


Fig. 6.3.2 (E) The plot of EhOASS between changes in abs. at 440 nm versus Cys Conc.

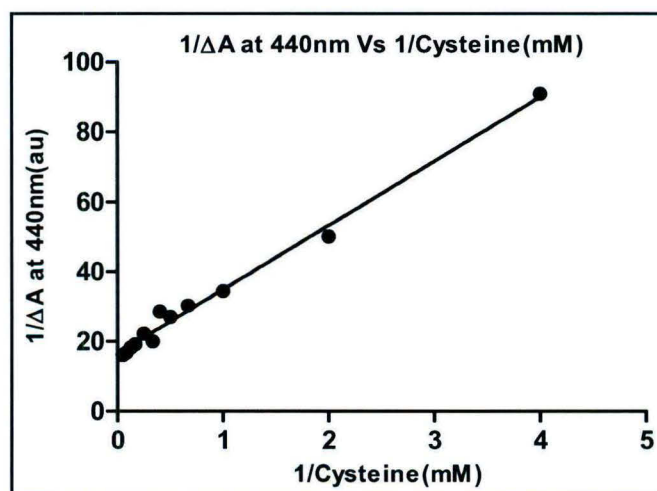


Fig. 6.3.2 (F) Double reciprocal plot of EhOASS between changes in abs. at 440nm Vs Cys conc.

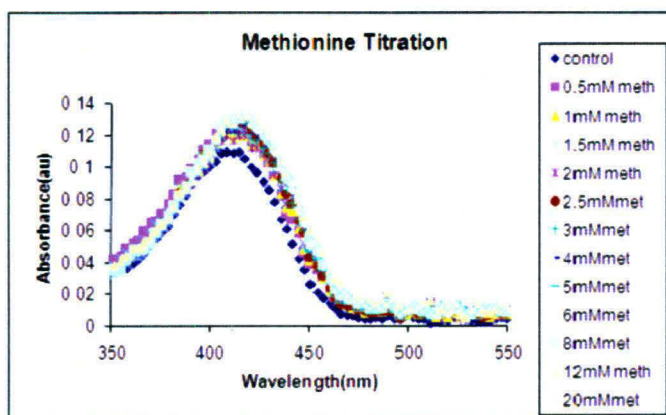


Fig. 6.3.2 (G) Spectrum of EhOASS with various concentrations of Met.

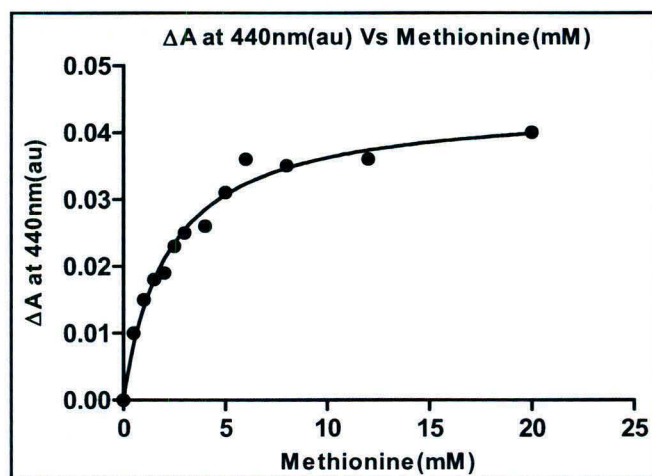


Fig. 6.3.2 (H) The plot of EhOASS between changes in abs. at 440nm versus Met. conc.

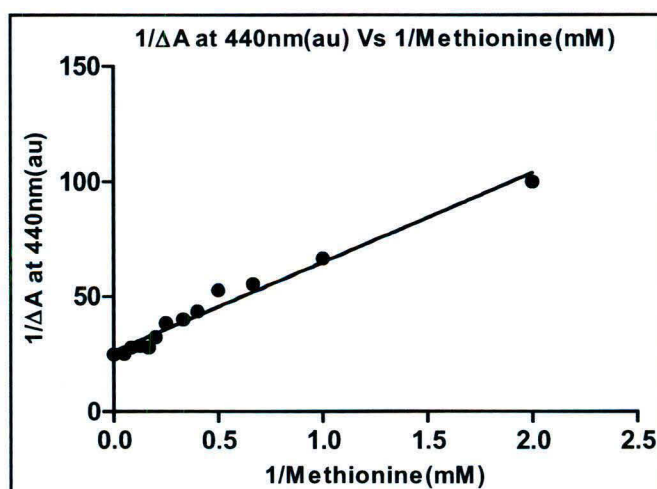


Fig. 6.3.2 (I) Double reciprocal of EhOASS plot between changes in abs. at 440nm Vs Met conc.

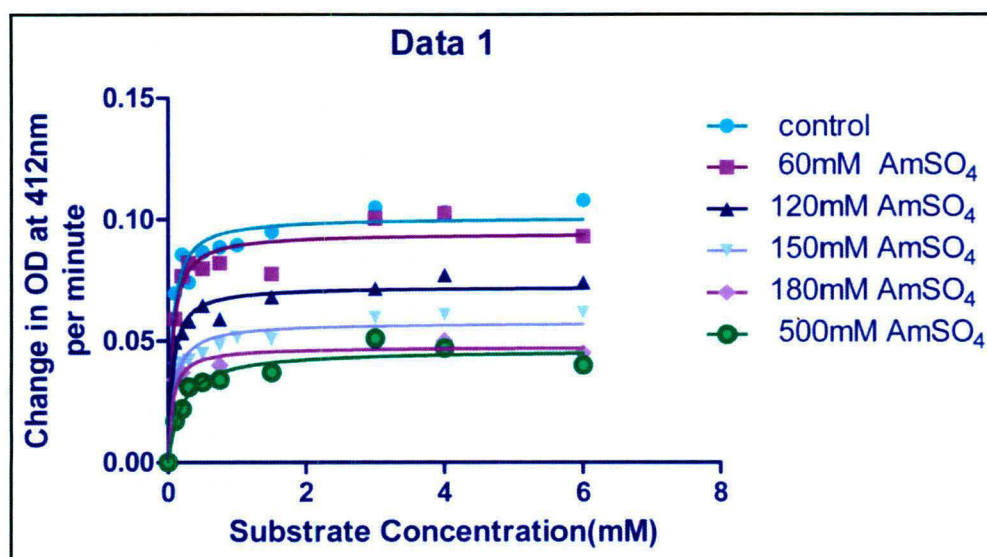


Fig. 7.3.3 The plot of EhOASS between changes in OD. at 412 nm Vs substrate Concentration (mM).

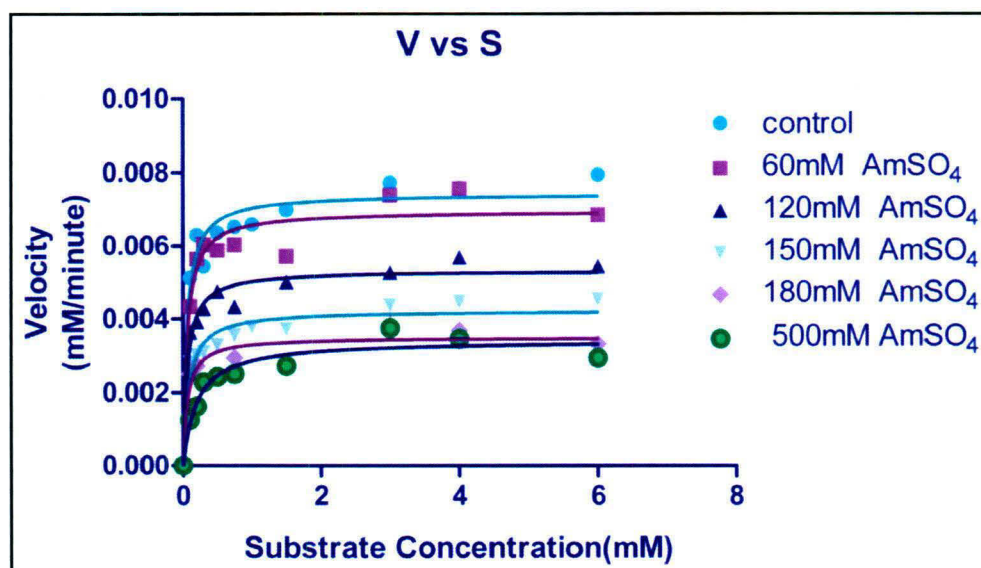


Fig. 7.3.4 Inhibition of EhOASS by sulfate at varying levels of OAS. The concentration of TNB was fixed at 0.1mM, and all rates were measured at pH 7.0, 100mM Hepes and 25°C. The fixed sulfate concentration in mM were used. The plot between V vs S at 412 nm.

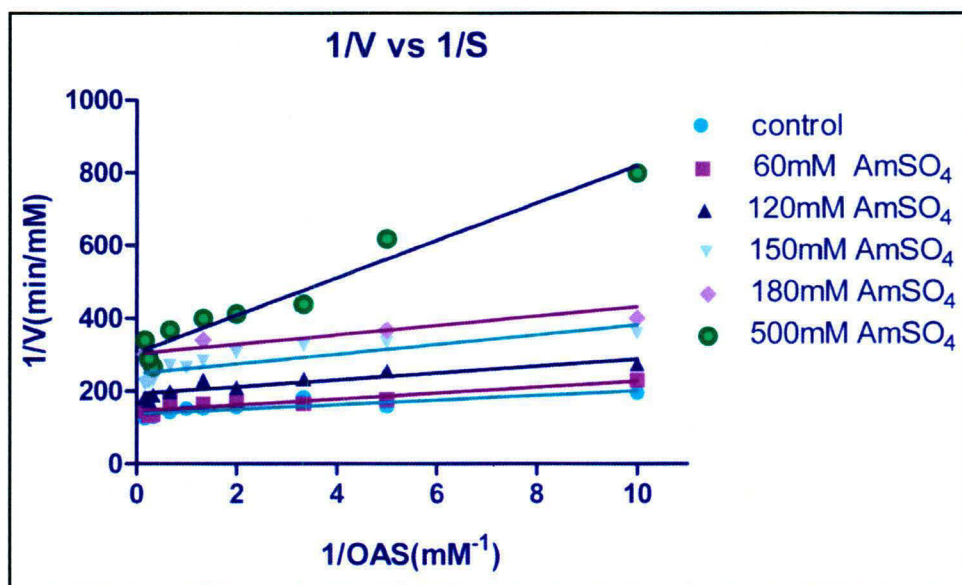
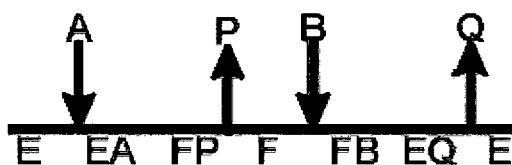


Fig. 7.3.5 Dead end inhibition of EhOASS by sulfate at varying levels of OAS. The concentration of TNB was fixed at 0.1 mM, and all rates were measured at pH – 7.0, 100mM Hepes at 25 °C. The fixed sulfate concentration in mM was used. Curves are theoretical based on fit using the equation 1 while points are experimental. The double reciprocal plot between 1/V vs 1/S at 412 nm.

7.4.1 Kinetics mechanism of EhOASS

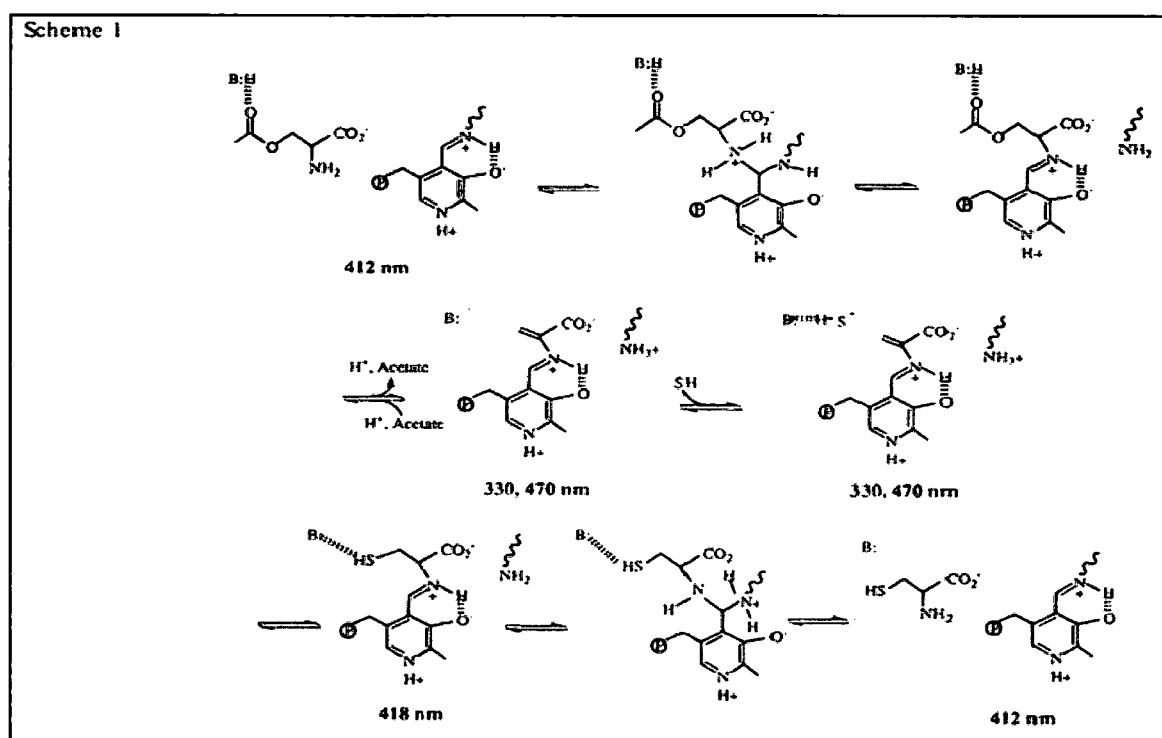
As has been reported with *Salmonella* OASS, EhOASS also somehow follows ping-pong kinetics mechanism (Chia-Hui Tai et al., 1993). The enzyme has a classical ping-pong kinetic mechanism. But till 8mM of OAS concentration no substrate inhibition is observed. The distinguishing feature of enzymes following ping-pong (double displacement) mechanism, is that at least one product is released from the enzyme before all of the substrates have bound. In the ping – pong kinetics the first substrate gets attached with the enzyme (E) through a bond. This leads to the formation of the modified form of the enzyme (E) that is F form. Here a part of the substrate already attached is released before the attachment of the second substrate. After the release of the part of the substrate, the other substrate binds and subsequently the final product is released. After the release of the product, the modification that was done as a result of the attachment of the first substrate is also removed and unmodified form of the enzyme that is E is also released. The whole scheme is summarized diagrammatically below.



In case of EhOASS evidence of the modified form of the enzyme that is F is given by the absorption of EhOASS alpha - aminoacrylate at 470 nm and 330 nm [fig 7.3.2 (a)] (Chia-Hui Tai et al., 1993). And also at the end of the cycle that is after the release of the product, peak absorption shifts to the absorption maximum of the native enzyme form that is at 412 nm which is due to PLP linked to active lysine of the enzyme EhOASS. The detailed explanation of the kinetics mechanism of EhOASS is as follows -

The kinetic mechanism for EhOASS (Scheme 1) takes a path in which OAS gets attached with its α -amine unprotonated to facilitate a nucleophilic attack on C 4' of the protonated Schiff base, and with the acetyl carbonyl hydrogen-bonded to a protonated enzyme group, which aids in the β -elimination of acetate (Cook et al., 1976). Then comes the picture of active lysine that was in Schiff base linkage with the active site PLP, removes hydrogen from the α -carbon which lead to the β -elimination reaction, and as a result a

proton along with the acetate product is given out. All these reactions give rise to intermediate product known as alpha - aminoacrylate ($\lambda_{\text{max}} = 470 \text{ nm}$) [fig 7.3.2 (a)]. Sulfide likely binds as HS^- and this lead to the nucleophilic attack on the alpha - aminoacrylate intermediate. Thereafter there is protonation of the alpha - carbon by the enzyme lysine. The HS^- attached with the enzyme group might be instrumental in the β elimination of acetate.



Scheme 1 Kinetic Mechanism of OASS. Adapted from (Ellika et al., 1996)

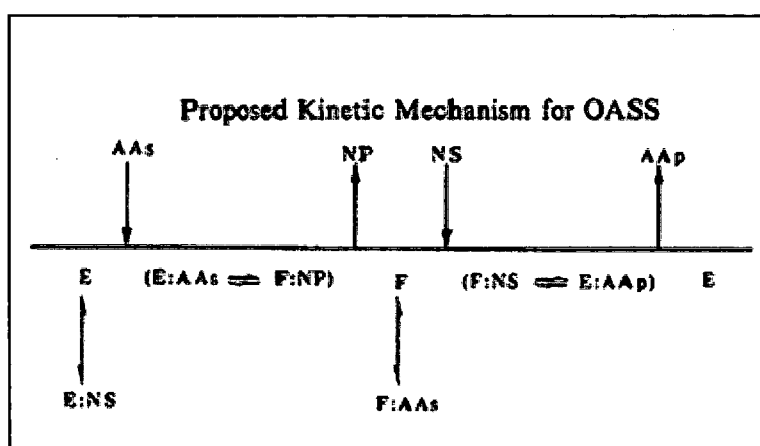
Ist half reaction has external Schiff base formation as intermediate which in the next step is converted into alpha - aminoacrylate with the β - elimination of acetate group [fig 7.3.2 (a)]. In the second half reaction where there is addition of sulphide group to alpha - aminoacrylate there is no intermediates formation.

In most of OASS cases like *Salmonella* OASS (Chia-Hui Tai et al., 1993) substrate inhibition was observed. But in case of EhOASS the substrate inhibition part was not shown by either of the substrate pair. The reason attributed to this phenomenon could be

low concentration of OAS and TNB used in the kinetics experiments. In case of OAS till 8 mM of concentration was used and till this concentration no substrate inhibition was observed and in case of TNB only 100 micromolar was used. And no inhibition because of TNB or OAS was found during the kinetics experiment.

7.4.1.1 Initial Velocity Patterns.

The initial velocity patterns obtained with all reactant for EhOASS is consistent with a ping pong mechanism. The mechanism generally adheres to that shown in Scheme 11, where AAs, NP, NS, and AAP are amino acid substrate, nucleophilic product, nucleophilic substrate, and amino acid product, respectively. E form of the enzyme is the native enzyme EhOASS. F is EhOASS bound with OAS. Double arrow indicates reversible reaction. Nucleophilic product is acetate. Amino acid product is cysteine. Nucleophilic substrate is sulphide. So the whole reaction proceeds in this way - E combines with the OAS and gives rise to F and acetate. F combines with sulphide to give rise to cysteine.



Scheme II Adapted from (Chia-Hui Tai et al., 1993)

7.4.2 Spectrum studies of EhOASS with different ligands

(a) EhOASS and OAS

The ultraviolet-visible spectrum of EhOASS exhibits an absorption maximum at 412 nm due to the formation of a protonated Schiff base between an active site lysine and the active site PLP (Cook & Wedding, 1976; Cook et al., 1992). Rapid-Scanning measurements were carried out to obtain information on the identity and rates of appearance and decomposition of transients during the pre-steady state time course of the OASS reaction.

During the reaction of OASS with OAS, the reactant in the first half reaction of the ping-pong mechanism gives the spectra shown in fig 7.3.2 (a). In this, one spectrum for EhOASS is prior to addition of OAS. Upon addition of OAS, the 412 nm absorption band of free enzyme resulting from the internal Schiff base (Cook et al., 1992) is shifted slightly to higher wavelength as seen in the value of the λ_{\max} for spectrum in comparison to the spectrum which is without OAS, and the lack of a single isosbestic point for other spectra [fig 7.3.2 (a)]. The rapid shift in the wavelength which signals formation of the external Schiff base of OAS (Schnackerz et al., 1995) is followed by a decrease in the absorbance at 412 nm – 420 nm (disappearance of the absorbance at 412 nm) with the concomitant appearance of new bands at 330 and 470 nm (new absorption maxima) [fig 7.3.2 (a)] indicative of the formation of a protonated Schiff base between PLP and alpha - aminoacrylate upon the β - elimination of acetate from OAS (Cook & Wedding, 1976; Schnackerz et al., 1979; 1995; Cook et al., 1992).

The 330 nm band also likely reflects alpha - aminoacrylate intermediate. It is suggested that the 330 nm band is due to an electronic transition of the alpha - aminoacrylate and likely represents a second tautomeric form of the intermediate, or a form in which the intramolecular hydrogen bond between the imine nitrogen and the 3-hydroxyl of PLP is for some reason not allowed (Cook et al., 1992).

There is an indication of pH dependence of dissociation constants for decomposition of the external Schiff bases as evident from the *Salmonella* OASS but in case of Eh OASS it has not been investigated.

Titration of OAS with EhOASS and change in the absorbance at 470nm gives a hyperbolic trend and yields OAS binding constant of $17.61 \pm 4.1 \mu\text{M}$ [fig 7.3.2 (b) (c)].

(b) EhOASS and Cysteine

Cysteine forms external Aldimine structure with EhOASS. The addition of L-cysteine to EhOASS lead to shifting of λ_{max} or peak absorption of EhOASS to 418 nm [fig 7.3.2 (d)] from 412 nm which is the normal absorption of EhOASS when it is without a ligand or unliganded form of EhOASS (Schnackerz et al., 1979; 1995). Cysteine formation in the EhOASS product cycle takes place at the end of the cycle and here there is shift in the absorption spectrum of OASS to 418 nm due to the formation of external aldimine structure [fig 6.3.2 (d)]. So shift in the spectrum of EhOASS with cysteine from 412 nm to 418 nm is an indication of the last part of the EhOASS product cycle. It is one step before the release of the product from EhOASS product cycle. There is also change (increase) in absorption spectrum for EhOASS at 470 nm with the increasing concentration of cysteine.

Titration of Cys with enzyme EhOASS yielded binding constants $1.129 \pm 0.25 \text{ mM}$ [fig 7.3.2 (e) (f)].

(c) EhOASS and Methionine

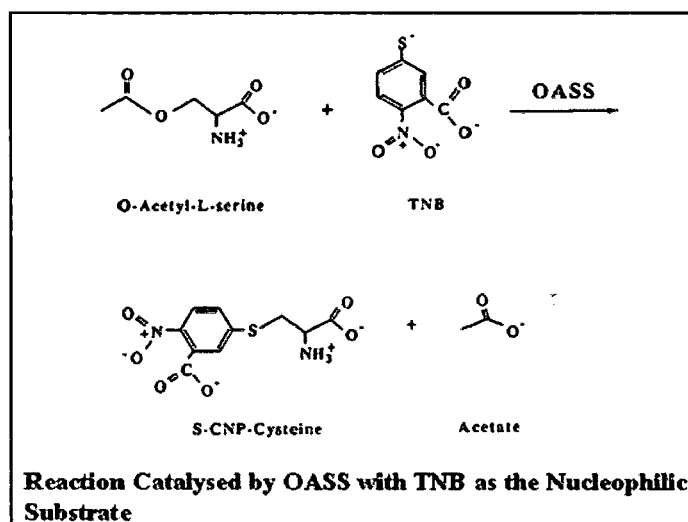
It is same as that of cysteine spectrum. Methionine forms external Aldimine structure with EhOASS. The λ_{max} or peak absorption of EhOASS shifts to 418 nm from 412 nm [fig 7.3.2 (g)] which is the normal absorption of EhOASS when it is without a ligand or unliganded form of EhOASS (Bonner et al., 2005). Methionine is known to be a substrate analog for EhOASS so this would well justify the shift in absorption spectrum to 418 nm (λ_{max}) [fig 7.3.2 (g)] which is the λ_{max} of EhOASS during the product formation. There is also change (increase) in absorption spectrum for EhOASS at 470 nm with the increasing concentration of methionine. The spectrum for the Methionine with EhOASS is somewhat similar or identical to the spectrum of cysteine with EhOASS.

Titration of Met with enzyme EhOASS yielded binding constants $1.491 \pm 0.35 \text{ mM}$ [fig 7.3.2 (h) (i)].

7.4.3 Kinetics of EhOASS with Ammonium Sulphate as an inhibitor

(a) Mercaptonitrobenzoate as an Alternative Substrate

Chia- Hui Tai (1993) developed an assay using TNB as an alternative substrate for sulfide. *Salmonella* OASS transforms OAS to alpha - aminoacrylate in Schiff base with PLP and λ_{\max} for this is at 320 and 470 nm (Cook et al. Wedding, 1976; Cook et al., 1991). Addition of TNB to the alpha - aminoacrylate leads to the shifting of absorbance from 470 nm to 340 nm, (Chai-Hui Tai et al., 1993). A difference spectrum was observed on an Amersham Biosciences ultraspec 2100 UV-Visible spectrophotometer. This revealed the diminishing of TNB at 412 nm and the simultaneous appearance of another compound absorbing at 340 nm, due to the formation of the another species, S-(3-carboxy-4-nitro- phenyl)-L-cysteine, Scheme III. Similar spectra were also observed for the alpha - aminoacrylate intermediate and S-CNP-cysteine formation catalyzed by EhOASS. The velocity Vs substrate concentration is linear for given concentration of reactants for EhOASS as observed with *Salmonella* OASS (Chia-Hui Tai et al., 1993). The saturation curves for OAS at a fixed concentration of the TNB are hyperbolic as observed with *Salmonella* OASS (Chia-Hui Tai et al., 1993). Here TNB acts as a nucleophilic substrate.



Scheme III Adapted from (Chia-Hui Tai et al., 1993)

In all cases, initial velocities obtained varying one reactant at different fixed levels of the second gave a series of parallel lines. The OASS has previously been shown to have a ping pong kinetic mechanism that requires the β elimination of acetate from OAS in the first half-reaction to generate α -aminoacrylate in Schiff base with the active site PLP (Cook & Wedding, 1976). The Michael addition of sulfide to the α -aminoacrylate intermediate then occurs in the second half-reaction to produce the final product L-cysteine.

(b) Ammonium sulphate as an inhibitor of EhOASS

Ammonium Sulfate inhibition

In this assay OAS was varied at different fixed levels of ammonium sulfate and the data obtained from this assay indicated that there is a non competitive inhibition by ammonium sulfate with respect to OAS (at fixed level of TNB) [fig. (7.3.3), (7.3.4)]. Competitive inhibition is predicted between a substrate and its analog when they bind to the same enzyme form (same site), while a non competitive pattern is predicted for the analog and the other substrate when they bind to different enzyme forms (different site). Non competitive inhibition by ammonium sulfate indicates that ammonium sulfate does not compete with OAS with respect to the binding to the EhOASS (internal Schiff base form) and it may happen that after binding of ammonium sulfate to EhOASS, it brings about structural change in the EhOASS which either make the enzyme unfit for binding to OAS or it combines with OAS but does not give rise to product. Slope inhibition data indicate that there must be some allosteric site for the binding of ammonium sulfate to the EhOASS (fig. 7.3.5). In nut shell sulfate acts as a dead end analogue of the nucleophilic substrate.

Data Processing - Data were fitted using the FORTRAN programs developed by Cleland (Cleland et al., 1979).

$$v = VA / [K_a(1 + I / K_{is}) A(1 + I / K_{ii})] \text{ -----(1)}$$

v and V are initial and maximum velocities, respectively, A represents the variable reactant, and I represents the inhibitor sulfate. The constant K_a represents the K_m for A , while K_{is} and K_{ii} represent slope and intercept inhibition constants, respectively.

The fit to linear noncompetitive gives:

$$K = 0.05 \pm 0.01$$

$$V = 7.5 \pm 0.3 \quad (\text{velocities moved by three decimal points})$$

$$K_{is} = 86 \pm 40$$

$$K_{ii} = 250 \pm 40$$

7.4.4 Kinetics of Cysteine with EhOASS

Cysteine is a product catalyzed by EhOASS using OAS and sulphide as substrates. In case of *Salmonella*, cysteine has been shown to be an inhibitor of OASS (Chia-Hui Tai et al., 1993). In case of *Salmonella* it acts as a potent inhibitor of OASS and this falls in the category of product inhibition. And in case of *Salmonella* OASS it acts as a negative modulator. These results are quiet in contradiction with what was observed with the Eh OASS. In case of EhOASS no inhibition was shown by EhOASS on the addition of cysteine to EhOASS reaction catalyzing the formation of the product using the TNB and OAS as substrate pairs. Starting from 2 mM of cysteine concentration to 100 mM of cysteine concentration was used but no inhibition was observed on the EhOASS reaction. In a nut shell it proves that cysteine does not exercise a negative feedback inhibition of EhOASS thereby allowing it to perform without the negative feedback regulation by the cysteine.

7.4.5 Conclusion

These binding constants with different ligands and also dead inhibition studies with ammonium sulfate of EhOASS are comparable to other OASS enzymes indicating that there is not much change in active site region and active site dynamics. Structure also shows that active site and substrate binding area is over all well conserved.

REFERENCES

- Agne, S. M., Tsuda, S., Spyropoulos, L., Kay, L. E. and Sykes, B. D. (1998). Backbone and methyl dynamics of the regulatory domain of troponin C: anisotropic rotational diffusion and contribution of conformational entropy to calcium affinity. *J. Mol. Biol.* 278, 667–686.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* 15:403–410.
- Amniuk, A. P., Nguyen, L. T., Hoang, T. T. and Vogel, H. J. (2004). Metal ion binding properties and conformational states of calcium- and integrin-binding protein. *Biochemistry*, 43, 2558–2568.
- András F, Andrej S (2003). ModLoop: automated modeling of loops in protein structures. *Bioinformatics* 19:2500–2501. doi:10.1093/bioinformatics/btg362.
- Ankri, S., Stolarsky, T., Bracha, R., Padilla-Vaca, F., and Mirelman, D. (1999). Antisense inhibition of expression of cysteine proteinases affects *Entamoeba histolytica*-induced formation of liver abscess in hamsters. *Infect Immun* 67, 421–422.
- 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.
- 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.
- Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992). Backbone dynamics of calmodulin studied by ¹⁵N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* 31, 5269–5278.
- Benci, S., Vaccari, S., Mozzarelli, A., and Cook, P. F. (1999). Time-resolved fluorescence of O-acetylserine sulphydrylase. *Biochim Biophys Acta.* 1429, 317–330.
- Berkowitz, O., Wirtz, M., Wolf, A., Kuhlmann, J., and Hell, R. (2002). Use of biomolecular interaction analysis to elucidate the regulatory mechanism of the cysteine synthase complex from *Arabidopsis thaliana*. *J. Biol. Chem.* 277, 30629–30634.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000). The Protein Data Bank. *Nucleic Acids Res.* 28:235–242. doi:10.1093/nar/28.1.235.
- Bernal, R. M., Tovar, R., Santos, J. I., and Munoz, M. L. (1998). Possible role of calmodulin in excystation of *Giardia lamblia*. *Parasitol Res* 84, 687–693.
- 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.
- 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.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., *et al.* (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309, 416–422.
- Bhattacharya, A., Arya, R., Clark, C. G., and Ackers, J. P. (2000). Absence of lipophosphoglycan-like glycoconjugates in *Entamoeba dispar*. *Parasitology* 120 (Pt 1), 31–35.
- Bhattacharya, A., Padhan, N., Jain, R., and Bhattacharya, S. (2006). Calcium-binding proteins of *Entamoeba histolytica*. *Arch Med Res* 37, 221–225.

- Bhattacharya, A., Prasad, R., and Sacks, D. L. (1992). Identification and partial characterization of a lipophosphoglycan from a pathogenic strain of *Entamoeba histolytica*. *Mol Biochem Parasitol* 56, 161-168.
- Bhattacharya, S., Bunick, C. G., and Chazin, W. J. (2004). Target selectivity in EF-hand calcium binding proteins. *Biochim Biophys Acta* 1742, 69-79.
- Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7, 1513-1523.
- Blaustein, M. P., and Lederer, W. J. (1999). Sodium/calcium exchange: its physiological implications. *Physiol Rev* 79, 763-854.
- Bonner, E. R., Cahoon, R. E., Knapke, S. M., and Jez, J. M. (2005). Molecular basis of cysteine biosynthesis in plants: structural and functional analysis of O-acetylserine sulfhydrylase from *Arabidopsis thaliana*. *J. Biol. Chem.* 280, 38803-38813.
- Borup, B., and J. G. Ferry. (2000). O-Acetylserine sulfhydrylase from *Methanosarcina thermophila*. *J. Bacteriol.* 182:45-50.
- Brown, D. M., J. A. Upcroft, and P. Upcroft. 1993. Cysteine is the major low-molecular weight thiol in *Giardia duodenalis*. *Mol. Biochem. Parasitol.* 61:155-158.
- Bruchhaus I, Tannich E. (1995). Identification of an *Entamoeba histolytica* gene encoding a protein homologous to prokaryotic disulphide oxidoreductases. *Mol Biochem Parasitol*, 70:187-91.
- Brumpt. E. (1925). Etude sommaire de l'*Entamoeba dispar* n. sp. Amibe a kystes quadinnuclees, parasite de l'homme. *Bull. De l'Acad. Med (Paris)*. 94: 942.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D*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.
- Burchard, G. D., and Bilke, R. (1992). Adherence of pathogenic and non-pathogenic *Entamoeba histolytica* strains to neutrophils. *Parasitol Res* 78, 146-153.
- Burgoyne, R. D., and Weiss, J. L. (2001). The neuronal calcium sensor family of Ca²⁺-binding proteins. *Biochem J* 353, 1-12.
- Burkhard, P., Jagannatha Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F., and Jansonius, J. N. (1998). Three-dimensional structure of O-acetylserine sulfhydrylase from *Salmonella typhimurium*. *J.Mol.Biol.* 283, 121-133.
- Burkhard, P., Tai, C. H., Ristroph, C. M., Cook, P. F., and Jansonius, J. N. (1999). Ligand binding induces a large conformational change in O-acetylserine sulfhydrylase from *Salmonella typhimurium*. *J. Mol. Biol.* 291, 941-953.
- Burkhard, P., Tai, C. H., Jansonius, J. N., and Cook, P. F. (2000). Identification of an allosteric anion-binding site on O-acetylserine sulfhydrylase: structure of the enzyme with chloride bound. *J. Mol. Biol.* 303, 279-286.

- Byrne, C. R., Monroe, R. S., Ward, K. A., and Kredich, N. M. (1988). DNA sequences of the *cysK* regions of *Salmonella typhimurium* and *Escherichia coli* and linkage of the *cysK* regions to *ptsH*. *J. Bacteriol.* 170, 3150-3157.
- Campanini, B., Speroni, F., Salsi, E., Cook, P.F., Roderick, S.L., Huang, B., Bettati, S., and Mozzarelli, A. (2005). Interaction of serine acetyltransferase with O- acetylserine sulfhydrylase active site: Evidence from fluorescence spectroscopy. *Protein Sci.* 14, 2115– 2124.
- 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.
- 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.
- Cheng, X. J., Hughes, M. A., Huston, C. D., Loftus, B., Gilchrist, C. A., Lockhart, L. A., Ghosh, S., Miller-Sims, V., Mann, B. J., Petri, W. A., Jr., and Tachibana, H. (2001). Intermediate subunit of the Gal/GalNAc lectin of *Entamoeba histolytica* is a member of a gene family containing multiple CXXC sequence motifs. *Infect Immun* 69, 5892-5898.
- Chia-Hui Tai, Srinivasa R. Nalabolu, Tony M. Jacobson, David E. Minter and Paul F. Cook P L (1993). Kinetic mechanisms of the A and B isozymes of O-acetylserine sulfhydrylase from *Salmonella typhimurium* LT-2 using the natural and alternate reactants. *Biochemistry*, 32, 6433-6442.
- Chou, J. J., Li, S., Klee, C. B., and Bax, A. (2001). Solution structure of Ca⁽²⁺⁾-calmodulin reveals flexible hand-like properties of its domains. *Nat Struct Biol* 8, 990-997.
- Clapham, D. E. (1995). Calcium signaling. *Cell* 80, 259-268.
- Clark, C. G. (2000). Cryptic genetic variation in parasitic protozoa. *J Med Microbiol* 49, 489-491.
- Claus, M. T., Zocher, G. E., Maier, T. H., and Schulz, G. F. (2005). Structure of the O-acetylserine sulfhydrylase isoenzyme CysM from *Escherichia coli*. *Biochemistry*. 44, 8620-8626.
- Cleland, W. W. (1979) Statistical analysis of enzyme kinetic data. *Methods Enzymol.* 63, 103-138.
- Cook, P. F., Hara, S., Nalabolu, S., and Schnackerz, K. D. (1992). pH dependence of the absorbance and ³¹P NMR spectra of O-acetylserine sulfhydrylase in the absence and presence of O-acetyl-L-serine. *Biochemistry* 31, 2298–2303.
- Cook, P.F., and Wedding, R.T. (1977). Initial kinetic characterization of the multienzyme complex, cysteine synthetase. *Arch. Biochem. Biophys.* 178, 293–302.
- Cook, P. F., & Wedding R. T. (1976). A reaction mechanism from steady state kinetic studies for O-acetylserine sulfhydrylase from *Salmonella typhimurium* LT-2. *J. Biol. Chem.* 251, 2023-2029.
- Cook, P. F., Nalabolu, S. R., & Tai, C. H. (1991). In *Enzymes Dependent on Pyridoxal Phosphate & Other Carbonyl Compounds as Cofactors* (Fukui, T., Kagamiyama, H., Soda, K., & Wada, H., Eds.) pp 321-323, Pergamon Press, Tokyo.
- Daum, S., Tai, C. H., and Cook, P. F. (2003). Characterization of the S272A,D site-directed mutations of O-acetylserine sulfhydrylase: involvement of the pyridine ring in the alpha,beta-elimination reaction. *Biochemistry* 42, 106–113.

- DeLano WL (2002). The PyMOL user's manual. San Carlos, CA, USA: DeLano Scientific.
- 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.
- Docampo, R., and Moreno, S. N. (2001). The acidocalcisome. *Mol Biochem Parasitol* 114, 151-159.
- Dominguez-Solis, J. R., Gutierrez-Alcala, G., Romero, J. C., and Gotor, C. (2001). The cytosolic O-acetylserine(thiol)lyase gene is regulated by heavy metals and can function in cadmium tolerance. *J. Biol. Chem.* 276, 9297-9302.
- Droux, M., Ruffet, M.L., Douce, R., and Job, D. (1998). Interactions between serine acetyltransferase and O-acetylserine (thiol) lyase in higher plants structural and kinetic properties of the free and bound enzymes. *Eur. J. Biochem.* 255, 235-245.
- Dvorak, J. A., Kobayashi, S., Alling, D. W., and Hallahan, C. W. (1995). Elucidation of the DNA synthetic cycle of *Entamoeba* spp. using flow cytometry and mathematical modeling. *J Eukaryot Microbiol* 42, 610-616.
- Eichinger, L., Pachebat, J. A., Glockner, G., Rajandream, M. A., Sugang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., *et al.* (2005). The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 435, 43-57.
- Eilika U. Woehl, Chia-Hui Tai, Michael F. Dunn, and Paul F. Cook (1996). Formation of the α -Aminoacrylate Intermediate Limits the Overall Reaction Catalyzed by O-Acetylserine Sulfhydrylase. *Biochemistry*, 35, 4776-4783
- Ellman, G. L. (1959) Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* 82, 70.
- Emsley, P., and Cowton, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*60. 2126-2132.
- Fahey RC, Newton GL, Arrick B, Overdank Bogart T, Aley SB. (1984). *Entamoeba histolytica*: a eukaryote without glutathione metabolism. *Science*, 224:70-2.
- Fallon, J. L., and Quioco, F. A. (2003). A closed compact structure of native $\text{Ca}^{(2+)}$ -calmodulin. *Structure* 11, 1303-1307.
- Fankhauser H, Brunold C, Erismann KH. (1976). Subcellular localization of O-acetylserine sulfhydrylase in spinach leaves. *Experientia*, 32:1494-7.
- Favaron, M., and Bernardi, P. (1985). Tissue-specific modulation of the mitochondrial calcium uniporter by magnesium ions. *FEBS Lett* 183, 260-264.
- Fimmel, A. L., and R. E. Loughlin. (1977). Isolation and characterization of cysK mutants of *Escherichia coli* K12. *J. Gen. Microbiol.* 103:37-43.
- Foguel, D., Suarez, M. C., Barbosa, C., Rodrigues, Jr, J. J., Sorenson, M. M., Smillie, L. B. and Silva, J. L. (1996). Mimicry of the calcium-induced conformational state of troponin C by low temperature under pressure. *Proc. Natl. Acad. Sci. U.S.A.* 93, 10642-10646.
- Francois, J. A., Kumaran, S., and Jez, J. M. (2006). Structural basis for interaction of O-acetylserine sulfhydrylase and serine acetyltransferase in the Arabidopsis cysteine synthase complex. *Plant Cell.* 18, 3647-3655.

- Gagne, S. M., Li, M. X. and Sykes, B. D. (1997) Mechanism of direct coupling Between binding and induced structural change in regulatory calcium binding proteins. *Biochemistry* 36, 4386–4392
- Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., *et al.* (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498-511.
- Gathiram, V., and Jackson, T. F. (1987). A longitudinal study of asymptomatic carriers of pathogenic zymodemes of *Entamoeba histolytica*. *S Afr Med J* 72, 669-672.
- 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 Ca²⁺ signalling to membrane dynamics. *Nat Rev Mol Cell Biol* 6, 449-461.
- Gifford, J. L., Walsh, M. P., and Vogel, H. J. (2007). Structures and metal-ion-binding properties of the Ca²⁺-binding helix-loop-helix EF-hand motifs. *Biochem J* 405, 199-221.
- Gilchrist, C. A., Leo, M., Line, C. G., Mann, B. J., and Petri, W. A., Jr. (2003). Calcium modulates promoter occupancy by the *Entamoeba histolytica* Ca²⁺-binding transcription factor URE3-BP. *J Biol Chem* 278, 4646-4653.
- Gilligan, D. M., and Satir, B. H. (1983). Stimulation and inhibition of secretion in *Paramecium*: role of divalent cations. *J Cell Biol* 97, 224-234.
- Gillin, F. D., and L. S. Diamond. (1980a). Attachment and short-term maintenance of motility and viability of *Entamoeba histolytica* in a defined medium. *J. Protozool.* 27:220–225.
- Gillin, F. D., D. S. Reiner, R. B. Levy, and P. A. Henkart. (1984). Thiol groups on the surface of anaerobic parasitic protozoa. *Mol. Biochem. Parasitol.* 13:1–12.
- Gillin, F. D., and L. S. Diamond. (1981c). *Entamoeba histolytica* and *Giardia lamblia*: growth responses to reducing agents. *Exp. Parasitol.* 51:382–391.
- Gillin FD, Diamond LS. (1981d). *Entamoeba histolytica* and *Giardia lamblia*: effects of cysteine and oxygen tension on trophozoite attachment to glass and survival in culture media. *Exp Parasitol*, 52:9–17.
- Gillin FD, Diamond LS. (1980b). Attachment of *Entamoeba histolytica* to glass in a defined maintenance medium: specific requirement for cysteine and ascorbic acid. *J Protozool*, 27:474–8.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166, 557-580.
- Hell, R., R. Jost, O. Berkowitz, and M. Wirtz. (2002). Molecular and biochemical analysis of the enzymes of cysteine biosynthesis in the plant *Arabidopsis thaliana*. *Amino Acids* 22:245–257.
- Hell, R., and Hillebrand, H. (2001). Plant concepts for mineral acquisition and allocation. *Curr. Opin. Biotechnol.* 12, 161–168.
- Hell, R., Bork, C., Bogdanova, N., Frolov, I., and Hauschild, R.(1994). Isolation and characterization of two cDNAs encoding for compartment specific isoforms of O-acetylserine (thiol) lyase from *Arabidopsis thaliana*. *FEBS Lett.* 351, 57–62.

- Hoeflich, K. P., and Ikura, M. (2002). Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell* 108, 739-742.
- Hooft RWW, Vriend G, Sander C, Abola EE (1996). Errors in protein structures. *Nature* 381:272–272. doi:10.1038/381272a0.
- Hopkins, L., Parmar, S., Blaszczyk, A., Hesse, H., Hoefgen, R., and Hawkesford, M.J. (2005). O-acetylserine and the regulation of expression of genes encoding components for sulfate uptake and assimilation in potato. *Plant Physiol.* 138, 433–440.
- Huang, B., Vetting, M. W., and Roderick, S. L. (2005). The active site of O-acetylserine sulfhydrylase is the anchor point for holoenzyme complex formation with serine acetyltransferase. *J. Bacteriol.* 187, 3201-3205.
- Hulanicka, M. D., N. M. Kredich, and D. M. Treiman. (1974). The structural gene for O-acetylserine sulfhydrylase A in *Salmonella typhimurium*. Identity with the *trzA* locus. *J. Biol. Chem.* 249:867–872.
- Humphrey W, Dalke A, Schulten K (1996). VMD—visual molecular dynamics. *J Mol Graph* 14:33–38. doi:10.1016/0263-7855(96)00018-5.
- Hung, C. C., Ji, D. D., Sun, H. Y., Lee, Y. T., Hsu, S. Y., Chang, S. Y., Wu, C. H., Chan, Y. H., Hsiao, C. F., Liu, W. C., and Colebunders, R. (2008). Increased Risk for *Entamoeba histolytica* Infection and Invasive Amebiasis in HIV Seropositive Men Who Have Sex with Men in Taiwan. *PLoS Negl Trop Dis* 2, e175.
- Huston, C.D., Haque, R. and Petri, W.A., Jr. (1999) Molecular-based diagnosis of *Entamoeba histolytica* infection.
- Ikegami F, Murakoshi I. (1994). Enzymic synthesis of non protein beta-substituted alanines and some higher homologues in plants. *Phytochemistry* ,35:1089–104.
- Jakob, U., W. Muse, M. Eser, and J. C. Bardwell. 1999. Chaperone activity with a redox switch. *Cell* 96:341–352.
- Jessica L. Gifford, Michael P. Walsh and Hans J. Vogel (2007) Structures and metal-ion-binding properties of the Ca²⁺-binding helix–loop–helix EF-hand motifs. *Biochem. J.* 405 (199–221).
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47. 110-119.
- Julie A. Francois, Sangaralingam Kumaran, and Joseph M. Jeza, 2006. Structural Basis for Interaction of O-Acetylserine Sulfhydrylase and Serine Acetyltransferase in the Arabidopsis Cysteine Synthase Complex. *The Plant Cell Preview*.
- Katz, M., Despommier, D.D. and Gwadz, R. (1989). *Entamoeba histolytica*. In *Parasitic Diseases* (2nd edn), Springer-Verlag, pp. 136-143.
- Kawasaki, H., Nakayama, S., and Kretsinger, R. H. (1998). Classification and evolution of EF-hand proteins. *Biometals* 11, 277-295.
- Ke, M., Forsen, S. and Chazin, W. J. (1991). Molecular basis for co-operativity in Ca²⁺ binding to calbindin D9k. 1H nuclear magnetic resonance studies of (Cd²⁺)1-bovine calbindin D9k. *J. Mol. Biol.* 220, 173–189.

- Kessin, R. H., and Franke, J. (1986). Secreted adenylate cyclase of *Bordetella pertussis*: calmodulin requirements and partial purification of two forms. *J Bacteriol* 166, 290-296.
- Keystone, J. S., Keystone, D. L., and Proctor, E. M. (1980). Intestinal parasitic infections in homosexual men: prevalence, symptoms and factors in transmission. *Can Med Assoc J* 123, 512-514.
- Klee, C. B., Crouch, T. H., and Krinks, M. H. (1979). Calcineurin: a calcium- and calmodulin-binding protein of the nervous system. *Proc Natl Acad Sci U S A* 76, 6270-6273.
- Kleinsmith, Hardin & Bertoni, *The World of the Cell*, 7th edition Becker
- Kopriva, S. (2006). Regulation of sulfate assimilation in *Arabidopsis* and beyond. *Ann. Bot. (Lond.)* 97, 479-495.
- Kredich, N. M., Becker, M. A., and Tomkins, G. M. (1969). Purification and characterization of cysteine synthetase, a bifunctional protein complex, from *Salmonella typhimurium*. *J. Biol. Chem.* 244, 2428-2439
- Kredich, N.M. (1996). Biosynthesis of cysteine, p. 514-527. In F. C. Neidhardt, R. Curtis III, J. L. Ingraham, M. Riley, M. Schaechter, and E. E. Umbarger (ed.), *Escherichia coli and Salmonella: cellular and molecular biology*, vol. 1. American Society for Microbiology, Washington, D.C.
- Kredich NM, Tomkins GM (1966). The enzymic synthesis of L-cysteine in *Escherichia coli* and *Salmonella typhimurium*. *J Biol Chem.* 241(21):4955-65.
- 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.
- Krishna, H., Jain, R., Kashav, T., Wadhwa, D., Alam, N., and Gourinath, S. (2007). Crystallization and preliminary crystallographic analysis of cysteine synthase from *Entamoeba histolytica*. *Acta. Crystallogr. F.* 63, 512-515.
- 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, 990-998.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Laskowski RA, MacArthur MW, Moss DS, Thornton J (1993). M: PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283-291. doi:10.1107/S0021889892009944.
- Lee, Y. H., Tanner, J. J., Larson, J. D. and Henzl, M. T. (2004). Crystal structure of a high affinity variant of rat α -parvalbumin. *Biochemistry* 43, 10008-10017.
- Leippe, M. (1997). Amoebapores. *Parasitol Today* 13, 178-183.
- Leustek, T., and Saito, K. (1999). Sulfate transport and assimilation. *Plant Physiol.* 120, 637-643
- Leustek, T., Martin, M.N., Bick, J.A., and Davies, J. P. (2000). Pathways and regulation of sulfur metabolism revealed through molecular and genetic studies. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51, 141-165.

- Li, M. X., Gagne, S. M., Tsuda, S., Kay, C. M., Smillie, L. B. and Sykes, B. D. (1995). Calcium binding to the regulatory N-domain of skeletal muscle troponin C occurs in a stepwise manner. *Biochemistry* 34, 8330–8340.
- Li, E., Yang, W. G., Zhang, T., and Stanley, S. L., Jr. (1995). Interaction of laminin with *Entamoeba histolytica* cysteine proteinases and its effect on amebic pathogenesis. *Infect Immun* 63, 4150–4153.
- Lin, Y. M., Liu, Y. P., and Cheung, W. Y. (1974). Cyclic 3':5'-nucleotide phosphodiesterase. Purification, characterization, and active form of the protein activator from bovine brain. *J Biol Chem* 249, 4943–4954.
- Linse, S. and Forsen, S. (1995). Determinants that govern high-affinity calcium binding. *Adv. Second Messenger Phosphoprotein Res.* 30, 89–151.
- Loftus, B., Anderson, I., Davies, R., Alsmark, U. C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R. P., Mann, B. J., *et al.* (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433, 865–868.
- Losch, F. A. (1875). Massive development of amebas in the large intestine. *Am J Trop Med Hyg.* 24:383–392.
- Lovett, J. L., and Sibley, L. D. (2003). Intracellular calcium stores in *Toxoplasma gondii* govern invasion of host cells. *J Cell Sci* 116, 3009–3016.
- Lowy, F. D. (1998). Staphylococcus aureus infections. *N. Engl. J. Med.* 339: 520–532.
- Lu HG, Zhong L, Chang KP, Docampo R. Intracellular Ca²⁺ pool content and signaling and expression of a calcium pump are linked to virulence in *Leishmania mexicana* amazonensis amastigotes. *J Biol Chem* 1997;272:9464–9473.
- Mann, B. J. (2002). Structure and function of the *Entamoeba histolytica* Gal/GalNAc lectin. *Int Rev Cytol* 216, 59–80.
- Mao, C., Kim, S. H., Almenoff, J. S., Rudner, X. L., Kearney, D. M., and Kindman, L. A. (1996). Molecular cloning and characterization of SCaMPER, a sphingolipid Ca²⁺ release-mediating protein from endoplasmic reticulum. *Proc Natl Acad Sci U S A* 93, 1993–1996.
- Marinets, A., Zhang, T., Guillen, N., Gounon, P., Bohle, B., Vollmann, U., Scheiner, O., Wiedermann, G., Stanley, S. L., and Duchene, 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, 1557–1565.
- 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.
- 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.
- Maurer, P., Hohenester, E., and Engel, J. (1996). Extracellular calcium-binding proteins. *Curr Opin Cell Biol* 8, 609–617.
- McKenzie, H. A., and White, F. H., Jr. (1991). Lysozyme and alpha-lactalbumin: structure, function, and interrelationships. *Adv Protein Chem* 41, 173–315.

- McClure, G. D. Jr. and Cook, P. F. (1994). Product binding to the alpha-carboxyl subsite results in a conformational change at the active site of O-acetylserine sulfhydrylase-A: evidence from fluorescence spectroscopy. *Biochemistry*. 33, 1674-1683.
- Medvedev, S.S. (2005). Calcium signaling system in plants. *Russian Journal of Plant Physiology*, 52(2): p. 249 - 270.
- Mehlotra R. (1996). Antioxidant defense mechanisms in parasitic protozoa. *Crit Rev Microbiol*,22:295–314.
- Meza, I. (2000). Extracellular matrix-induced signaling in *Entamoeba histolytica*: its role in invasiveness. *Parasitol Today* 16, 23-28.
- Mino, K., Yamanoue, T., Sakiyama, T., Eisaki, N., Matsuyama, A. and Nakanishi, K. (1999). Purification and characterization of serine acetyltransferase from *Escherichia coli* partially truncated at the C-terminal region. *Biosci Biotechnol Biochem*. 63, 168-179.
- Mino, K., Yamanoue, T., Sakiyama, T., Eisaki, N., Matsuyama, A and Nakanishi, K. (2000). Effects of bienzyme complex formation of cysteine synthetase from *Escherichia coli* on some properties and kinetics. *Biosci Biotechnol Biochem*. 64, 1628-1640.
- Murshudov, G. N., Vagin, A. A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallogr. D*53. 240-255.
- 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.
- Nakamura, T., H. Iwahashi, and Y. Eguchi. (1984). Enzymatic proof for the identity of the S-sulfocysteine synthase and cysteine synthase B of *Salmonella typhimurium*. *J. Bacteriol*. 158:1122–1127.
- Nalefski, E. A., and Falke, J. J. (1996). The C2 domain calcium-binding motif: structural and functional diversity. *Protein Sci* 5, 2375-2390.
- Nelson, M. R. and Chazin, W. J. (1998) An interaction-based analysis of calcium induced conformational changes in Ca²⁺ sensor proteins. *Protein Sci*. 7, 270–282
- 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.
- Nickel, R., Ott, C., Dandekar, T., and Leippe, M. (1999). Pore-forming peptides of *Entamoeba dispar*. Similarity and divergence to amoebapores in structure, expression and activity. *Eur J Biochem* 265, 1002-1007.
- Noji, M., Murakoshi, I., and Saito K (1994). Molecular cloning of a cysteine synthase cDNA from *Citrullus vulgaris* (watermelon) by genetic complementation in an *Escherichia coli* Cys-auxotroph. *Mol Gen Genet*, 244, 57–66.
- Noji, M., Inoue, K., Kimura, N., Gouda, A., and Saito, K. (1998). Isoform-dependent differences in feedback regulation and subcellular localization of serine acetyltransferase involved in cysteine biosynthesis from *Arabidopsis thaliana*. *J. Biol. Chem*. 273, 32739-32745.
- Noji, M., and K. Saito. (2002). Molecular and biochemical analysis of serine acetyltransferase and cysteine synthase towards sulfur metabolic engineering in plants. *Amino Acids* 22:231–243.

- Noji, M., Saito, M., Nakamura, M., Aono, M., Saji, H., and Saito, K. (2001). Cysteine Synthase Overexpression in Tobacco Confers Tolerance to Sulfur-Containing Environmental Pollutants *Plant Physiol.* 126, 973–980.
- Nozaki, T., Asai, T., Kobayashi, S., Ikegami, F., Noji, M., Saito, K., and Takeuchi, T. (1998). Molecular cloning and characterization of the genes encoding two isoforms of cysteine synthase in the enteric protozoan parasite *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* 97, 33–44.
- Nozaki, T., Asai, T., Lidya B. S., Seiki, K., Miki, N., & Takeuchi, T. (1999). Characterization of the gene encoding serine acetyltransferase, a regulated enzyme of cysteine biosynthesis from the protist parasites *Entamoeba histolytica* and *Entamoeba dispar*. Regulation and possible function of the cysteine biosynthetic pathway in *Entamoeba*. *J. Biol. Chem.* 274, 32445–32452.
- Nozaki, T., V. Ali, and M. Tokoro. (2005). Sulfur-containing amino acid metabolism in parasitic protozoa. *Adv. Parasitol.* 60C:1–99.
- Nozaki, T., Y. Shigeta, Y. Saito-Nakano, M. Imada, and W. D. Kruger. (2001). Characterization of transsulfuration and cysteine biosynthetic pathways in the protozoan hemoflagellate, *Trypanosoma cruzi*. Isolation and molecular characterization of cystathionine beta-synthase and serine acetyltransferase from *Trypanosoma*. *J. Biol. Chem.* 276:6516–6523.
- Nozaki, T., M. Tokoro, M. Imada, Y. Saito, Y. Abe, Y. Shigeta, and T. Takeuchi. (2000). Cloning and biochemical characterization of genes encoding two isozymes of cysteine synthase from *Entamoeba dispar*. *Mol. Biochem. Parasitol.* 107:129–133
- Ogasawara, N., Nakai, S., Yoshikawa, H. (1994). Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. *DNA Res.* 1, 1–14.
- 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^{2+} -calmodulin-dependent kinase kinase. *Nat Struct Biol* 6, 819-824.
- Otwinowski, Z., AND Minor, W. (1997) Processing of X-ray diffraction and data collection in oscillation mode. *Methods Enzymol.* 276, 307–326.
- Pan, C. Q., and Lazarus, R. A. (1999). Ca^{2+} -dependent activity of human DNase I and its hyperactive variants. *Protein Sci* 8, 1780-1788.
- Pattni, K., and Banting, G. (2004). Ins(1,4,5)P3 metabolism and the family of IP3-3Kinases. *Cell Signal* 16, 643-654.
- Perrakis, A., Morris, R.M. and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* 6, 458-463. Public Health Agency of Canada (PHAC). (2001). Infectious substances: *Entamoeba histolytica*. Office of Laboratory Security. Material Safety Data Sheet.
- Pozzan, T., Rizzuto, R., Volpe, P., and Meldolesi, J. (1994). Molecular and cellular physiology of intracellular calcium stores. *Physiol Rev* 74, 595-636.
- 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.
- Prusch, R. D., and Hannafin, J. A. (1979). Calcium distribution in *Amoeba proteus*. *J Gen Physiol* 74, 511-521.

- Que, X., and Reed, S. L. (2000). Cysteine proteinases and the pathogenesis of amebiasis. *Clin Microbiol Rev* 13, 196-206.
- R, D. B. (2004). The neuronal calcium-sensor proteins. *Biochim Biophys Acta* 1742, 59-68.
- Rake, S. K. and Falke, J. J. (1996). Kinetic tuning of the EF-hand calcium binding motif: the gateway residue independently adjusts (i) barrier height and (ii) equilibrium. *Biochemistry* 35, 1753-1760.
- Ramu C, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31(13):3497- 3500. doi:10.1093/nar/gkg546.
- Rasmussen H, J.P., Lake W, Goodman Db. (1976). Calcium ion as second messenger. *Clinical Endocrinology (Oxf)*, 5 Suppl.: p. 11S-27S.
- Ravdin, J. I., and Guerrant, R. L. (1981). Role of adherence in cytopathogenic mechanisms of *Entamoeba histolytica*. Study with mammalian tissue culture cells and human erythrocytes. *J Clin Invest* 68, 1305-1313.
- 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., Murphy, C. F., Guerrant, R. L., and Long-Krug, S. A. (1985). Effect of antagonists of calcium and phospholipase A on the cytopathogenicity of *Entamoeba histolytica*. *J Infect Dis* 152, 542-549.
- Rege, V. S., Kredich, N. M., Tai, C. H., Karsten, W. E., Schnackerz, K. D., and Cook, P. F. (1996). A change in the internal aldimine lysine (K42) in O-acetylserine sulfhydrylase to alanine indicates its importance in transamination and as a general base catalyst. *Biochemistry* 35, 13485-13493.
- Romer, S., d'Harlingue, A., Camara, B., Schantz, R., Kuntz (1992). Cysteine synthase from *Capsicum annuum* chromoplasts. Characterization and cDNA cloning of an up-regulated enzyme during fruit development. *J. Biol. Chem.* 267, 7966-70.
- Ruffet, M. L., Droux, M., and Dounce, R. (1994). Purification and Kinetic Properties of Serine Acetyltransferase Free of O-Acetylserine(thiol)lyase from Spinach Chloroplasts. *Plant Physiol.* 104, 597-604.
- 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.
- Saito, K., Miura, N., Yamazaki, M., Hirano, H., Murakoshi, I., (1992). Molecular cloning and bacterial expression of cDNA encoding a plant cysteine synthase. *Proc. Natl. Acad. Sci. USA*, 89, 8078-82.
- Saito, K., Yokoyama, H., Noji, M. & Murakoshi, I. (1995). Molecular cloning and characterization of a plant serine acetyltransferase playing a regulatory role in cysteine biosynthesis from watermelon. *J Biol Chem.*270:16321-16321.
- Saito, K. (2004). Sulfur Assimilatory Metabolism. The Long and Smelling Road. *Plant Physiol.* 136, 2443-2450

- Saito, K., Kurosawa, M., Tatsuguchi, K., Takagi, Y., and Murakoshi, I. (1994). Modulation of cysteine biosynthesis in chloroplasts of transgenic tobacco overexpressing cysteine synthase (O-acetylserine(thiol)lyase). *Plant Physiol.* 106, 887–895.
- Saito K, Tatsuguchi K, Murakoshi I, Hirano H. (1993). cDNA cloning and expression of cysteine synthase B localized in chloroplasts of *Spinacia oleracea*. *FEBS Lett*, 324:247–52.
- Sali A, Blundell T (1993). L: Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815. doi:10.1006/jmbi.1993.1626.
- Schaudinn, F. (1903). The development of some rhizopoda (preliminary report). *Arbeiten aus dem Kaiserlichen Gesundheitsamte.* 19:547.
- Scheibel, L. W. (1992). Role of calcium/calmodulin-mediated processes in protozoa. *Int Rev Cytol* 134, 165-242.
- Schnackerz, K. D., Tai, C. H., Simmons, J. W. 3rd, Jacobson, T. M., Rao, G. S., and Cook, P. F. (1995). Identification and spectral characterization of the external aldimine of the O-acetylserine sulphydrylase reaction. *Biochemistry.* 34, 12152-12160.
- Schnackerz, K. D., Ehrlich, J. H., Giessman, W., & Reed, T. A. (1979) Mechanism of action of D-serine dehydratase. Identification of a transient intermediate. *Biochemistry* 18, 3557-3563.
- Schnell, R., Oehlmann, W., Singh M., and Schneider, G. (2007). Structural Insights into Catalysis and Inhibition of O-Acetylserine Sulphydrylase from *Mycobacterium tuberculosis*: crystal structures of the enzyme {alpha}-aminoacrylate intermediate and an enzyme-inhibitor complex. *J. Biol. Chem.* 282, 23473-23481.
- Schroder, B., Schlumbohm, C., Kaune, R., and Breves, G. (1996). Role of calbindin-D9k in buffering cytosolic free Ca²⁺ ions in pig duodenal enterocytes. *J Physiol* 492 (Pt 3), 715-722.
- Seaton, B. A., and Dedman, J. R. (1998). Annexins. *Biometals* 11, 399-404.
- 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.
- Shandar Ahmad, M. Michael Gromiha, Hamed Fawareh and Akinori Sarai *BMC Bioinformatics* (2004) 5:51
- Sippl M J (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins* 17:355–362. doi:10.1002/prot.340170404.
- Srere, P.A. (1987). Complexes of sequential metabolic enzymes. *Annu. Rev. Biochem.* 56, 89–124.
- Stanley, S. L., Jr., Zhang, T., Rubin, D., and Li, E. (1995). Role of the *Entamoeba histolytica* cysteine proteinase in amebic liver abscess formation in severe combined immunodeficient mice. *Infect Immun* 63, 1587-1590.
- 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., Cherney, M., Sielecki, A. R., Li, M. X., Smillie, L. B. and James, M. N. (1997). Structural details of a calcium-induced molecular switch: X-ray Crystallographic analysis of the

- calcium-saturated N-terminal domain of troponin C at 1.75Å resolution. *J. Mol. Biol.* 273, 238–255.
- Sutton, R. B., and Sprang, S. R. (1998). Structure of the protein kinase C beta phospholipid-binding C2 domain complexed with Ca²⁺. *Structure* 6, 1395-1405.
- Sutton, R. B., Davletov, B. A., Berghuis, A. M., Sudhof, T. C., and Sprang, S. R. (1995). Structure of the first C2 domain of synaptotagmin I: a novel Ca²⁺/phospholipid-binding fold. *Cell* 80, 929-938.
- Tai, C. H., Burkhard, P., Gani, D., Jenn, T., Johnson, C., and Cook, P. F. (2001a). Characterization of the allosteric anion – binding site of O-acetylserine sulfhydrylase. *Biochemistry*. 40, 7446-7452.
- Tai, C. H., and Cook, P. F. (2001b). Pyridoxal 5'-phosphate-dependent alpha,beta- elimination reactions: mechanism of O-acetylserine sulfhydrylase. *Acc. Chem. Res.* 34, 49–59
- Takahashi, H., and Saito, K. (1996). Subcellular Localization of Spinach Cysteine Synthase Isoforms and Regulation of Their Gene Expression by Nitrogen and Sulfur. *Plant Physiol.* 112, 273–280
- 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 U S A* 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⁽²⁺⁾ affinities and target recognition. *J Biol Chem* 275, 10514-10518.
- Tardieux, I., Webster, P., Ravesloot, J., Boron, W., Lunn, J. A., Heuser, J. E., and Andrews, N. W. (1992). Lysosome recruitment and fusion are early events required for *trypanosome* invasion of mammalian cells. *Cell* 71, 1117-1130.
- The World Health Report (1995): Bridging the Gaps; Report of the Director-General. Geneva: World Health Organization, 16, 377–385.
- Vagin A, Teplyakov A MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* 1997;30:1022– 1025.
- Vahab Ali and Tomoyoshi Nozaki (2007). Current Therapeutics, Their Problems , and Sulfur-Containing-Amino-Acid Metabolism as a Novel Target against Infections by “Amitochondriate” Protozoan Parasites *CLINICAL MICROBIOLOGY REVIEWS*. p. 164–187.
- Van der Ploeg, J. R., M. A. Weiss, E. Saller, H. Nashimoto, N. Saito, M. A. Kertesz, and T. Leisinger. (1996). Identification of sulfate starvation-regulated genes in *Escherichia coli*: a gene cluster involved in the utilization of taurine as a sulfur source. *J. Bacteriol.* 178:5438–5446.
- Van der Ploeg, J. R., M. Barone, and T. Leisinger. (2001). Functional analysis of the *Bacillus subtilis* *cysK* and *cysJI* genes. *FEMS Microbiol. Lett.* 201:29–35.
- Van der Ploeg, J. R., N. J. Cummings, T. Leisinger, and I. F. Connerton. (1998). *Bacillus subtilis* genes for the utilization of sulfur from aliphatic sulfonates. *Microbiology* 144:2555–2561.
- Vieira MC, Moreno SN. (2000). Mobilization of intracellular calcium upon attachment of *Toxoplasma gondii* tachyzoites to human fibroblasts is required for invasion. *Mol Biochem Parasitol.* 106:157–162.

- Vogel, H. J. (1994). The Merck Frosst Award Lecture (1994). Calmodulin: a versatile calcium mediator protein. *Biochem Cell Biol* 72, 357-376.
- Voorheis, H. P., Bowles, D. J., and Smith, G. A. (1982). Characteristics of the release of the surface coat protein from bloodstream forms of *Trypanosoma brucei*. *J Biol Chem* 257, 2300-2304.
- Vriend G (1990). WHAT IF: a molecular modeling and drug design program. *J Mol Graph* 8:52-56. doi:10.1016/0263-7855 (90)80070-V.
- WHO. (1998). "The world health report 1998. Life in the 21st century: A vision for all." World Health Organization, Geneva, Switzerland.
- Wedemeyer, W. J., E. Welker, M. Narayan, and H. A. Scheraga. (2000). Disulfide bonds and protein folding. *Biochemistry* 39:4207-1426.
- 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* 56, 1485-1491.
- Weinbach EC, Diamond L. (1974). *Entamoeba histolytica* I aerobic metabolism. *Exp Parasitol* ,35:232-43.
- Weis, W. I., Taylor, M. E., and Drickamer, K. (1998). The C-type lectin superfamily in the immune system. *Immunol Rev* 163, 19-34.
- Westrop, G. D., G. Goodall, J. C. Mottram, and G. H. Coombs. (2006). Cysteine biosynthesis in *Trichomonas vaginalis* involves cysteine synthase utilizing O-phosphoserine. *J. Biol. Chem.* 281:25062-25075.
- Wiederstein M, Sippl MJ (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 10:1093. doi:10.1093/nar/10.3.1093
- Wilihoeft, U., Campos-Gongora, E., Touzni, S., Bruchhaus, I., and Tannich, E. (2001). Introns of *Entamoeba histolytica* and *Entamoeba dispar*. *Protist* 152, 149-156.
- Winkel, B.S.J. (2004). Metabolic channeling in plants. *Annu. Rev. Plant Biol.* 55, 85-107.
- Wirtz, M., Berkowitz, O., Droux, M., and Hell, R. (2001). The cysteine synthase complex in plants: Mitochondrial serine acetyltransferase from *Arabidopsis thaliana* carries a bifunctional domain for catalysis and protein-protein interaction. *Eur. J. Biochem.* 268, 686-693.
- Wirtz, M., and Droux, M. (2005). Synthesis of the sulfur amino acids: Cysteine and methionine. *Photosynth. Res.* 86, 345-346.
- WWW.biochem.arizona.edu/classes/bioc462/462
- 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.
- Yakubu MA, Majumder S, Kierszenbaum F. (1994). Changes in *Trypanosoma cruzi* infectivity by treatments that affect calcium ion levels. *Mol Biochem Parasitol* ,1:119-125.
- Youssefian, S., Nakamura, M., Orudjev, E., and Kondo, N. (2001). Increased Cysteine Biosynthesis Capacity of Transgenic Tobacco Overexpressing an O-Acetylserine(thiol) Lyase Modifies Plant Responses to Oxidative Stress *Plant Physiol.* 126, 1001-1011.

Zhang, M., and Yuan, T. (1998). Molecular mechanisms of calmodulin's functional versatility. *Biochem Cell Biol* 76, 313-323.

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.

Zheng, M., F. Aslund, and G. Storz. (1998). Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* 279:1718–1721.

Zhu, X., Yamaguchi, T., and Masada, M. (1998). Complexes of serine acetyltransferase and isozymes of cysteine synthase in spinach leaves. *Biosci. Biotechnol. Biochem.* 62, 947–952.

Appendices

List of Figures

- Figure 1.6.1** Life cycle of *Entamoeba histolytica* and the clinical manifestations of infection in humans
- Figure 1.6.2** Mechanisms responsible for calcium homeostasis in the cell
- Figure 1.6.3** Different calcium binding motifs
- Figure 1.6.4** EF hand Ca²⁺ binding motif
- Figure 1.6.5** Three dimensional structure of an EF-hand motif from parvalbumin
- Figure 1.6.6** Schematic representation of calmodulin
- Figure 2.3.1** 14 % SDS PAGE showing purified Eh CaBP2 In lane 1
- Figure 2.3.2** Crystallization of EhCaBP2, Sr-EhCaBP2, IQ- EhCaBP2
- Figure 3.3.1** Multiple sequence alignment of nine EhCaBPs with Calmodulin and Skeletal Essential Light Chain
- Figure 3.3.2** Cartoon representations of models of seven EhCaBPs
- Figure 3.3.3** Cartoon representations of models of seven EhCaBPs with labelled EF hands
- Figure 3.3.4** Ramachandran plots for the models of seven EhCaBPs proteins
- Figure 3.3.5** ProsaWeb analysis of theoretical 3D models of seven EhCaBPs
- Figure 3.3.6** Electrostatic molecular surface representation of the models of seven EhCaBPs
- Figure 3.3.7** Graphical representation of solvent accessibility of seven EhCaBPs
- Figure 3.3.8** Composition of amino acids in the seven EhCaBPs and Calmodulin and Groupwise Composition of amino acids in the seven EhCaBPs and Cam
- Figure 3.3.9** Superimposed structures of theoretical 3D models of the five EhCaBP proteins with (C-alpha atoms only) on their respective templates

- Figure 3.3.10** An alternative models for CaBP7 and CaBP9 showing structural part, Ramachandran plot, Z score and energy plots
- Figure 3.3.11** Phylogenetic relationship among nine EhCaBPs and 27 hCaBPs proteins of *Entamoeba histolytica*
- Figure 5.3.1** General scheme of transsulfuration, cysteine biosynthesis, and sulfur Amino acid degradation
- Figure 5.3.2** Cysteine biosynthesis.
- Figure 5.3.3** Regulation of Cys Synthesis by Formation of the Cys Synthase Complex
- Figure 6.3.1** 12% SDS PAGE showing purified EhOASS, GPC profile of EhOASS, EhOASS crystals
- Figure 6.3.2** Representative region the electron density map (2Fo-Fc) with a final model superimposed on it
- Figure 6.3.3** Structure of *E. histolytica* OASS
- Figure 6.3.4** Sequence homology of various OASS sequences
- Figure 6.3.5** Comparison of EhOASS and AtOASS and difference in the C-terminal Helix Position
- Figure 6.3.6** Cysteine bound EhOASS
- Figure 7.3.1** Ultraviolet-visible spectrum of EhOASS
- Figure 7.3.2** Spectral Studies of EhOASS with OAS, Methionine and Cysteine, Changes in Absorbance of EhOASS with OAS, Methionine and Cysteine at 440 nm, Double reciprocal plot between changes in abs. of EhOASS at 440 nm OAS, Methionine and Cysteine
- Figure 7.3.3** The plot between changes in OD. at 412 nm Vs substrate conc. (mM)
- Figure 7.3.4** Inhibition of OASS by sulfate at varying levels of OAS
- Figure 7.3.5** Dead end inhibition of OASS by sulfate at varying levels of OAS

List of Tables

Table 1.7	List of Ca ²⁺ binding proteins
Table 2.3.3	Crystallization of EhCaBP2 and its complexes
Table 2.3.4	Data Collection Statistics Of EhCaBP2 and SrEhCaBP2
Table 2.3.5	Data Collection Statistics Of IQEhCaBP2
Table 3.5.1	Members of the EhCaBP family and the sequences used for modeling
Table 3.5.2	Ramachandran plot statistics of EhCaBP proteins studied
Table 3.5.3	Charge of the proteins studied
Table 3.5.4	Comparison of models with their respective templates
Table 3.5.5	Ramachandran plot statistics of alternative model of EhCaBP (7 & 9)
Table 3.5.6	Sequence identity of EhCaBPs with EhCaBP1
Table 4.3	Calcium Binding Constants for Calcium Binding Proteins
Table 6.4.1	Data-collection and refinement statistics of EhCS and EhCS- Cys
Table 6.4.2	Surface area buried at the interface of the dimer

Sequences

CaBP2 sequence

MAEALFKQLDANGDGSVSYEEVKAFVSSKRPIKNEQLQLIFKAIDIDGNGEIDLAEFT
KFAAAVKEQDLSDEKVLKILYKLMADGDGKLTKEEVTTFFKFGYEVVDQIMKADA
NGDGYITLLEFLAFNL

CaBP3 sequence

MSEQKKVLTAAEQQYKEAFQLFDKDNDNKLTAEEELGTVMRALGANPTKQKISEIVKDY
DKDNSGKFDQETFLTIMLEYGQEV DSTEDIKKA FEIFDKEKNGYISASELKHVLTTLGE
KLTEQEVDDLKEIGVEEGLIN VDDFVKLITSK

CaBP4 sequence

MSSLTEEQKKKFKATFKHYDKDKGQITFEELGQILRGMGRNTTDVEIYQMQQIHGSDK
IDEATYLALLAQKLQEPDSVEEIQKAFDTFFGPGKTTITSDEFKAMMEFGERVSEDEA
DELIKDTGLNKEGCIDVKAFIDHVFESK

CaBP5 sequence

MQKHNE DLKESFLFDGDGDGYLTLNEFESLVRVLGVVMETSIAI ASTYNSNSKVRGMSY
ELFTSCFSQLKTKSFNKDEIKTAINVLDKDKGFIPAIELRRILSTIGDNMEQKEITDL
FTFMGIDEQGVVKVDDFINQLMTVFK

CaBP6 sequence

MSMEIEAPNANTQKIRDCFNFYDRDYDGKIDVKQLGTLIRSLGCAPTEDEVNSYIKEFA
IEGETFQIEQFELIMEREQSKPDTREIKLRKA FEVFDQDKDGKIKASDLAHLNLT TVGDK
MTKEEVEKVFSILGITMESDIDLATFLKLVAL

CaBP7 sequence

MNNTQIEFIFNLLDKDKSGYLS PDEFCEGLIEFYHITDDKKESYREIFNDIFLLADGKG
LFNSKDNKITLREFKNICSLLP TETKKT DVIIGTVAFRLVDSNNSGKVNKKELTAFLKK
TGVHLKKGEIDALMDSIDEDGDGKISFQEFMLLYDLN

CaBP9 sequence

MPEDKPQGKTIEEFFAEIDLDKDGSVNVDEYFNGVKMWRNDITDDDKPSQALLFHLADL
NDDGEIDIRQFARLIAILNRGFGKDIKSVFTAVFRLLDIKDQGGKIGAPELERLLKMG
QIESDDIEGFMQVDQDLQDGFISLKEFLAHFVKESKDGQ

CaBP10 sequence

MSRRRDEEVDQELVGELKDAFDMFDSSKKGYLKDDVKKLFKTNGIRVTDEDLDAAFKE
ADADGDKKIECMEFINMMTGKMTASTEQKLTEAFKVFDPPEEKVIDSKELTEALLNIG
ERCTTSEVGELKTVAENQEGQIRYELFIQAVFAKK

Cysteine Synthase sequence

MEQISISSPRKRIYHNILETIGGTPLVELHGVTEHPRIKKGTRILVKLEYFNPMSVVD
RVGFNIVYQAIKDGRLKPGMEIESTSGNTGIALCQAGAVFGYRVNIAMPSTMSVERQMI
MKAFGAELILTEGKKGMPGAIEEVNMIKENPGKYFVANQFGNPDNTAAHHTANEIWE
DTDGEVDIVVSAVGTSGTVIGVAEKLKEKKKGIKI IAVEPEESAVLEGKAKGPHGIQGI
GAGFIPDIYKKEFVDEI IPIKTQDAWKMARAVVKYDGIMCGMSSGAAI LAGLKEAEKPE
NEGKTIVIIVPSCGERYLSTDLYKIKDEGTKIQILDSLLNE