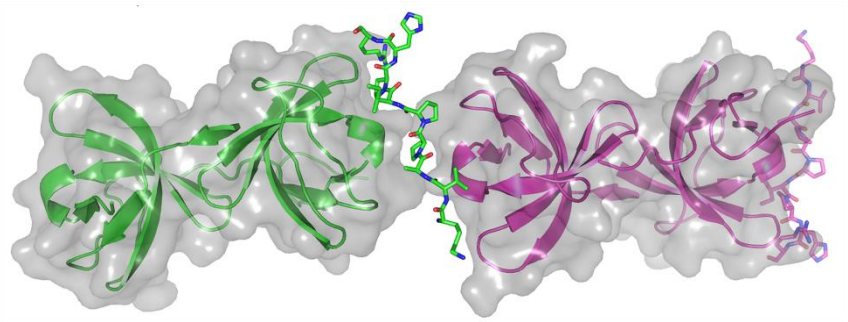
Structural and Functional Characterisation of Unconventional myosin IB from *Entamoeba histolytica*



Thesis submitted to Jawaharlal Nehru University
for the award of degree of

DOCTOR OF PHILOSOPHY

2017

By

Gunjan Gautam

School of Life Sciences
Jawaharlal Nehru University
New Delhi – 110067
INDIA

SCHOOL OF LIFE SCIENCES

Jawaharlal Nehru University

New Delhi – 110067

INDIA

CERTIFICATE

This is to certify that the research work embodied in this thesis entitled as "**Structural and Functional Characterisation of Unconventional myosin IB from *Entamoeba histolytica***" submitted for the award of degree of **Doctor of Philosophy**, has been carried out by **Ms. GUNJAN GAUTAM** under the guidance and supervision of Prof. Samudrala Gourinath, Professor, Structural Biology Laboratory at the School of Life Sciences, Jawaharlal Nehru University, New Delhi, India.

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

Gunjan Gautam
(Ph.D. Candidate)

Prof. Samudrala Gourinath
(supervisor)

for **Prof. S. K. Goswami**
(Dean)

SCHOOL OF LIFE SCIENCES

Jawaharlal Nehru University

New Delhi – 110067

INDIA

CERTIFICATE OF ORIGINALITY

The research work embodied in this thesis entitled as "**Structural and Functional Characterisation of Unconventional myosin IB from *Entamoeba histolytica***" has been carried out by me at the School of Life Sciences, Jawaharlal Nehru University, New Delhi, India. The thesis has been subjected to plagiarism check by **'Tunitin'** software. The work submitted for the consideration of award of Ph.D. is original.

Gunjan Gautam

SCHOOL OF LIFE SCIENCES

Jawaharlal Nehru University

New Delhi – 110067

INDIA

Supervisor's Certificate for Exclusion of Self-Published work

The content of the Chapter 1 entitled "Structural Studies of SH3 domain of EhMyosin IB" have been published in

- **Gunjan Gautam**, Syed Arif Abdul Rehman, Preeti Pandey, and Samudrala Gourinath. Crystal structure of the PEG-bound SH3 domain of myosin IB from *Entamoeba histolytica* reveals its mode of ligand recognition. Acta Cryst. (2017). D73.

The published work has been included in the thesis and has not been submitted for any degree to any University/ Institute.

Gunjan Gautam
(Ph.D. Candidate)

Prof. Samudrala Gourinath
(supervisor)

ACKNOWLEDGEMENTS

"Patience, persistence and perspiration make an unbeatable combination for success"- Napoleon Hill.

Thesis is a long journey with ups and downs and I have not travelled through this all by myself. The support, love and encouragement of a numerous people have kept me going. It's now time to express my heartfelt gratitude to all of them.

At this moment of accomplishment, first and the foremost, I would like to express my sincere gratitude towards my supervisor and mentor Prof. Samudrala Gourinath. I am deeply indebted to him for believing in me in giving me an opportunity to work with him. I have never met a person as much patient as he is. His ways to deal with situations and his logical attitude towards problems had kept me going. His critical analysis to refine the outcome will be the best thing I have learned with him. Despite his busy schedules, he has always fulfilled his responsibilities as a guide and encouraged me when things didn't work out. His fun filled yet extensive lab meetings provided me an opportunity to discuss the problems at common platform with the whole group. The after meeting lunch treat by him provided the whole group a wonderful opportunity to interact with each other. Time spent with his family was wonderful. His wife Dr. Neelima Alam has always welcomed us with so much warmth at her in-house parties. She is a lady with elegance both in thoughts and manners. She has always motivated us towards life. The connection I could make with Mansha, his daughter can't be explained in words. Her innocence and affection has always moved me. Sir, your constant support, guidance, encouragement and invaluable criticism has helped me to go ahead in difficult times and deal with every problem during my PhD. Thank you, for being there always and supporting me in times of distress.

I am thankful to the current and previous Deans of School of Life Sciences for providing excellent work environment and facilities at SLS. I am also thankful to all the faculty members at SLS for extending their support.

I sincerely thank Mr. Mishra, Mr. Amar Chand and other CIF staff for the technical help. Special thanks to Tripti, for her extraordinary help during live-cell experiments. I would also like to thank the administrative staff of SLS especially Late Talwar sir, Meenu ma'am and Sunita ma'am for their help regarding any official matter through these five years.

I acknowledge CSIR for the financial assistance!

I want to thank the laboratory staff members Ved Prakash and Krishna for providing the necessary technical support and always being around to extend their help.

I feel immense pleasure to thank my seniors Sudhir kumar, Isha Raj, Arif Sir, Faisal sir, Aruna, Nitesh kumar, Sanjeev kumar, Rohit sir, Preeti ma'am, Mohit sir, Shadab Sir, Vijayan sir, and Tamanna ma'am who have always encouraged and supported me in the lab. I would specially like to thank Nitesh sir for creating a friendly environment in the lab and helpful nature during my starting days in the lab; and Arif sir for teaching me with basic scientific skills of structural biology. Special thanks to Dr. Tamanna Anwar, Dr. Aruna Murmu and Preeti ma'am, for their scientific and friendly advices during our refreshing tea sessions.

The support of good colleagues is necessary for a fun filled and motivating working environment. I have been lucky to have my batchmate and friend, Sudhaker with me in lab during all these years, who has always supported me and helped me whenever I needed. His friendly and fun-filled nature always kept the environment light and helped a lot in times of distress. Decoding his telgu mixed hindi was much of a task, but had been a part of most of the joyous moments for the entire lab. His scientific inputs has always helped me many a times, during my Ph.D. years. My loving juniors, witty Suneeta; irritating, jolly and helpful Dhakaram; hardworking Poonam; caring and funny Priya; lovely Pragya; PJ king Debrat, and adorable Preeti. I am greatly indebted to you all for making an excellent team to work with. Special thanks to Preeti, for helping me during the draft of thesis. I will cherish the fun times with my lovely juniors,

Someone once told me that, for a ph.D. need a good supervisor and one very good friend. It turned out to be really true. This journey would have not been easy without the company my close friends; Aruna, Shagufta, Preeti and Silky. Special thanks to you guys!! I will always cherish the moments spent with them. They have always extended an ear to all my frustrations. Their advice both on the professional and personal front were helpful. My tea sessions with Aruna used to be the most refreshing part of the day. From discussing every aspect of our life, we shared a good camaraderie. I will definitely miss her company. She has been a senior, a friend, and a sister to me. Shagy's weekend trips at my hostel room and Saturday night cooking and chatting sessions with Silky have been an awesome part of my these Ph.D. years in JNU. I was lucky to have so good friends who never felt bad when I didn't responded to their messages.

I would like to thank all my PhD batchmates specially Kailash, Suchi, Pareeta, Sabir, Pam, Ranjana, Vipin, Poonam, Pratyusha, and Prabhu, at SLS for their help in one way or another during my PhD tenure. I am also thankful to my M.Sc. Batchmates and friends Anita, and Janesh for all the help. Special

thanks to my dear friend Anita. Her jolly nature, life controversies and 'let go' attitude had always amazing expedition for us. She has always shared the ups and downs of professional as well as personal front with me. Thank you for being there!

I am extremely lucky to have wonderful roommates cum friends; Story-teller Marnyier, Nimisha and the tea queen Neha. They have always looked out for me and been there whenever I needed them. Nimisha and me shared the largest part of my Ph.D. years, her wonderful nature full of warmth and care always touched me. She had been a great friend. She and Kailash were always there to help me. We shared a good companionship. Special mention to my latest roomie and the 'Tea Queen', Neha. Though we spent little time with each other but each moment will be memorable to me. She had supported me a lot during my thesis writing with our tea sessions and adjusting nature.

The main credit of my thesis goes to my family. Family is everything. I am extremely grateful to my parents. They always motivated me to strive for the best with honesty which has helped me in becoming a better person in life. I am at a loss of words to express my thanks to my Mom and Dad. Their constant support, encouragement, blessings and unending love is the reason for my success. My mom has always put up with me, even in my worse moods and uplifted my spirits with her positive attitude. My Dad has always been my role model. His hard working, responsible, logical attitude towards life and emerging as a winner in worse situations in life had always inspired me and kept hope alive in me during my Ph.D. Special thanks to my loving siblings Montu, Little and Deepu Didi, for their constant support. They have always cheered me up and been my strength throughout. The situational jokes and leg pulling with little has always cheered up my mood. Love you guys! I owe my thesis to you all.

Above all, I am extremely thankful to the almighty God. He provided me with strength and courage at every step of my life. My belief that he's there for me never stopped me from making efforts.

I owe this thesis to all these wonderful people associated with me.

Gunjan Gautam

July 2017.

Table of Contents

Table of Figures

List of Tables

List of Abbreviations

α	Alpha
β	Beta
µ	micro
Å	Angstrom
σ	Sigma
Ac	*Acanthamooeba castallani*
ATP	Adenosine Triphosphate
ADF	Actin Depolymerization Factor
APS	Ammonium Per Sulfate
bp	base pair
BSA	Bovine Serum Albumin
Ca	Calcium
CaBPs	Calcium Binding Proteins
$CaCl_2$	Calcium Chloride
CBB	Coomassie Brilliant Blue
CCP4	Collaborative Computational Project 4
$CHCl_3$	Chloroform
Cl	Chloride
COOT	Crystallographic Object-Oriented Toolkit
Dd	*Dictyostelium discoideum*
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylene Diamine Tetra Acetate
G-Buffer	G-Actin Buffer
F-Buffer	F-Actin Buffer
GalNAc	N-Acetylglucosamine
GPC	Gel-Permeation Chromatography
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEK	HEPES, EDTA, KCl
hr	Hour
Ig	Immunoglobulin
kb	Kilobase

KCl	Potassium Chloride
kDa	Kilo Dalton
LB	Luria Broth
M	Marker
mA	milli Ampere
MALDI	Matrix Assisted Laser Desorption Ionization
mg	milligram
$MgCl_2$	Magnesium Chloride
Min	minute
ml	millilitre
mM	milliMolar
MPD	2-methyl-2,4-pentandiol
MS	Mass Spectrometry
My	Myosin
NaN_3	Sodium Azide
nm	Nanometer
°C	degree centigrade
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene Glycol
PFA	Paraformaldehyde
PIC	Protease inhibitor cocktail
PMSF	Phenyl Methyl Sulphonyl Fluoride
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid
rpm	revolutions per minute
SAD	Single-wavelength Anomalous Dispersion
Sc	*Saccharomyces cerevisiae*
SDS	Sodium Dodecyl Sulphate
sec, s	seconds
SH3	Serine homology 3
TAE	Tris-Acetate-EDTA
TEMED	N,N,N',N', Tetramethylethylenediamine
T_m	Annealing temperature

Tris	Tris (hyroxymethyl) amino ethane
U	Unit
v/v	volume/volume
w/v	weight/volume
WHO	World Health Organization
β-ME	β -Mercaptoethanol

1. Review of Literature

1.1 Discovery and classification of Unconventional myosins.

Unconventional myosins came into the picture with the discovery of Myosin I from *Acanthamoeba* (Pollard and Korn, 1973) which was initially believed to be a proteolytically degraded product of the then well studied conventional myosin II (Korn and Hammer, 1990). When two different groups independently isolated *Acanthamoeba* myosin II (Maruta & Korn 1977, Pollard, Stafford & Porter 1978), it was proved that the previously isolated molecule belonged to a new class of myosin which unlike dimeric Myosin II, exist as a monomer i.e. has a single heavy chain with a single motor head domain and hence, were called Myosin 'I' and also lacked a coiled-coil tail. The difference in the actin binding kinetics and structure of Myosin I when compared to conventional Myosin II, categorized these myosins to be unconventional.

According to a recent report based on phylogenetic analysis and genomic survey, Myosin superfamily comprises 31 classes [1] and 15 classes belong to the unconventional myosins. In this classification, Myosin I forms the largest class of unconventional myosin which has been further divided into 4 subclasses [2]. On the basis of amino acid sequence, myosin I can be divided into two main groups: Amoeboid Myosin I from lower eukaryotes including rat myr3 and human myosin IC homologs, and Myosin I from higher eukaryotes.

1.2 Evolution of myosin I and divergence in eukaryotes:

Evolution of myosin protein started with eukaryotic evolution. Myosin gene family is widely distributed across all the taxonomic genera. No classical myosin genes have been found in prokaryotes. Among eukaryotes, myosin genes prevail from

yeast to higher eukaryotes. *S.cerevisae* are the simplest eukaryotes to possess myosin genes in the genome. A few genera like red algae, diplomonad protists lack genes for myosin head domain. A recent hypothesis states three myosin subfamilies to be the most ancient ones which include (i) MSD subfamily containing myosin II, (ii) MYTH4/FERM domain containing myosin families and (iii) myosin I.

Myosin I subfamily possess membrane binding, TH1 domain. It is found in chromalveolates, excavates, opisthokonts and Amoebozoa but not in Plantae, inferring that it was cenancestral but was lost by Plantae. Among the ancestral families, myosin I and MYTH4/FERM both the families have an SH3 domain in the tail region only in unikonts (animals, fungi, Choanozoa and Amoebozoa) but not in bikonts (plants, chromists, and all other protozoa), while MSD subfamily has an N-terminal SH3 domain. [3]. According to a recent hypothesis, the first cenancestral myosin gene must have an SH3 like domain as a result of insertion which might have been incorporated into the three ancestral subfamilies because of gene duplication, deletions, and divergence in the course of eukaryotic evolution.[3].

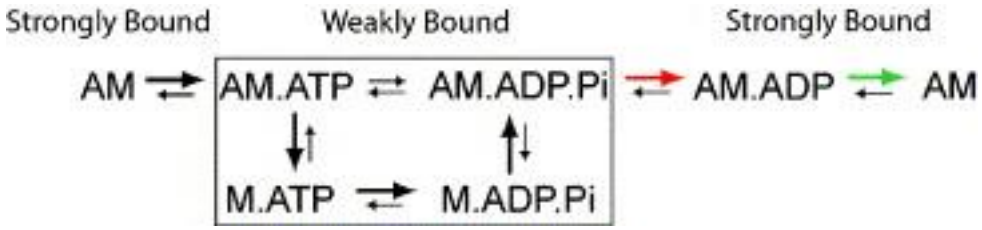
The primary life processes like cytokinesis, pseudopodial movement, vesicle movement and phagocytosis are vital for a eukaryotic cell and partitioned between these three ancestral myosin families. But still, the basis of partitioning during the course of evolution needs to be explored. Also, the presence of SH3 domain in two of three ancestral myosin increases the importance of this domain for the vital processes of the early eukaryotic cell.

1.3 Domain architecture of Myosin I

Myosin I are monomeric entities with a single heavy chain and lack filamentous coiled-coil tail as present in Myosin II. A single myosin I polypeptide chain can be divided into three distinct regions: Motor head domain, neck domain, and Tail domain.

1.3.1 Motor Domain:

Myosin I has a single motor domain also known as the 'catalytic head domain' which is about 80kDa in size. Myosin head has an adenosine nucleotide binding site and an ATP-dependent actin binding site. The hydrolysis of ATP and release of the product are crucial for the catalytic activity of myosin motor which determines and regulates its movement on filamentous actin. In absence of nucleotide, motor domain is strongly bound to the actin filament, known as the 'Rigor state' of myosin head domain. When ATP binds at nucleotide binding site, the affinity of the motor domain for filamentous actin weakens and the motor domain dissociates from actin filament. Bound ATP is then hydrolyzed into ADP +Pi which leads to a swing in the lever arm located in the head domain. As soon as the product of ATP hydrolysis is released, motor domain again tightly binds the filamentous actin. Thus, the ATPase cycle regulates the motility of myosin on the actin filament [4]. Motor domain of different myosins varies in 'duty ratio' i.e. the span of time in an ATPase cycle for which myosin is strongly bound to actin, and rate of ATP hydrolysis [5].



Myosin ATPase cycle [5]

Kinetic studies of Myosin I motors of *Acanthamoeba* Myosin IA and Myosin IB, brushed border myosin, Myo 1a, and human myosin IC (Myo Ie) stated that myosin I are low duty myosins i.e. spend maximum time in nucleotide bound state weakly bound to actin filament.[6-8]. Compared to muscle myosin, myo 1a spend maximum time weakly bound to F-actin, still it shows a slow rate of ATP hydrolysis

and ADP release, and ATP induced dissociation from actin filament as compared to muscle myosin [7]. Myosin I lacks the first 80 amino acids present in myosin II but shares many conserved residues present in motor domain of myosin II and other myosins. [9].

1.3.2 Neck Domain:

Motor domain is followed by an alpha-helical neck domain also known as 'Light chain Binding Domain (LCBD)'. This region contains one or more stretches of 'IQ' motif which has a consensus sequence 'IQXXXRGXXXR' in which Isoleucine-glutamine residues are characteristically positioned at the start of the sequence [10,11]. The number of IQ motif present differs among different myosin I isoforms. *Acanthamoeba* myosin IA has three IQ motifs while myosin IB and myosin IC, each have one IQ motif [12]. Calmodulins binds at the IQ motif of myosin I in yeast and vertebrates and they are not dependent on calcium for the binding response. The affinity of calmodulin for IQ motifs as well as the sequence of IQ motifs varies among myosins. The number of IQ motifs determines the step size of a myosin motor [12].

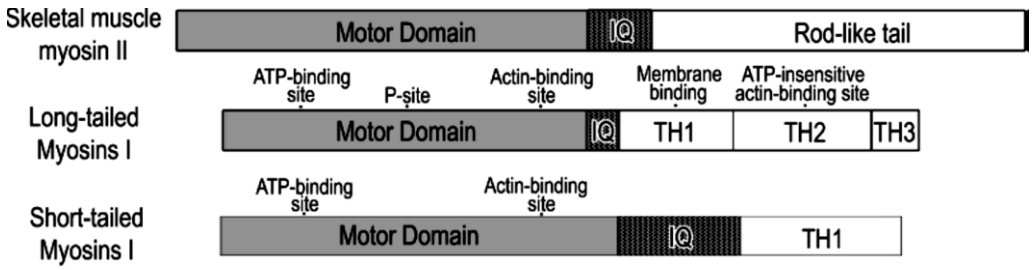
Among Amoebazoa, light chain of *Acanthamoeba* myosin IC resembles calmodulin and binds to the single IQ motif present. This could be one of the light chains for AcMyosin IA, which has 3 IQ motifs. Three class I myosins present in *Dictyostelium*, associate with unique light chains while DdMyoA binds calmodulin [12]. The light chain of DdMyoB, is much smaller and resembles one lobe of calmodulin. [13]. Calcium binding protein 5, which is structurally similar to essential light chain of myosin than to calmodulin has been postulated to be the light chain of myosin IB of *Entamoeba histolytica* [14].

IQ motifs had also been found to serve as a binding site for proteins other than calmodulins. Vertebrate MyoIc interacts with the receptors present at the tip of hair cells through its IQ motif. [15]

1.3.3 Tail Domain:

Myosin I have a diverse architecture of tail according to their in-vivo function. Based on the structure of tail domain, myosin I has been classified into two types: Long tailed Myosin I and Short tailed myosin I.

The tail of myosin I is divided into three distinct regions known as 'tail homology' (TH) regions: TH1, TH2 and TH3. Myosin I which has all the three tail homology domains are known as 'classical' or 'amoeboid' or long-tailed' myosin I while myosin which possess only TH1 in their tail domain are known as 'short-tailed' myosin I.



Schematic representation of primary structure of myosin I molecules.[12]

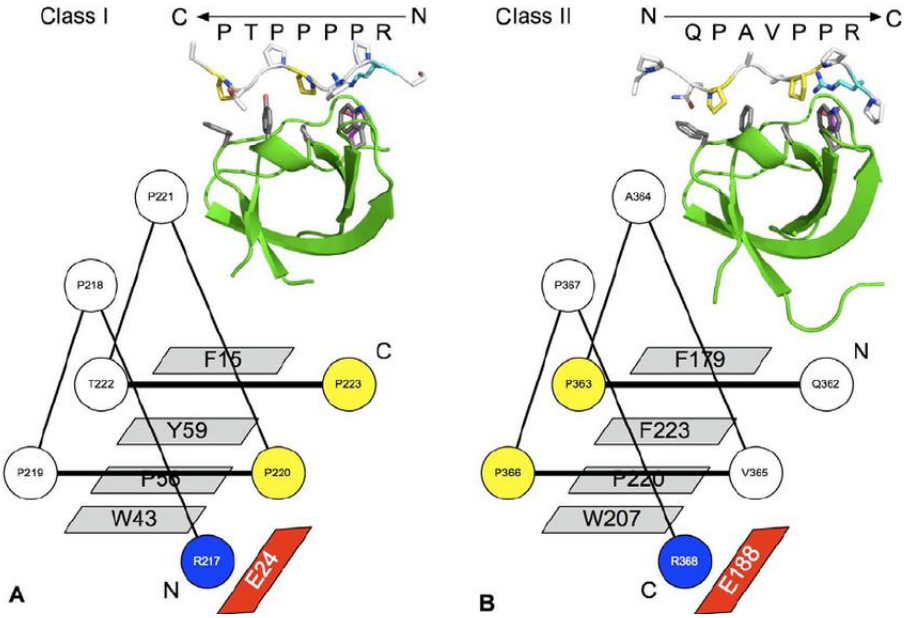
TH1 region is highly basic in nature and involved in interaction of myosin I with acidic phospholipids and membranes. It does not have the same affinity for all phospholipids [16]. TH1 region in *Acanthamoeba* Myosin IC also has cryptic actin binding site [17]. Myosin I in non-vertebrates as well as vertebrates has always been seen to be associated with the membranes.

TH2 region is known as ATP-independent actin binding site. This region is rich in glycine, proline and alanine or glutamic acid, hence is also called GPA or GPQ rich region. TH2 is believed to be involved in cross-linking of actin filaments within the cell.

TH3 domain in myosin I is basically an SH3 domain (serine homology domain 3). Sh3 domain are present in a number of signalling proteins and have a variety of effectors.

1.4 SH3 domain

Sh3 domains are protein-protein interaction modules found in a variety of signalling proteins. It is a small domain of about 60 amino acid in length. Structurally, SH3 domain is composed of five antiparallel β-sheets arranged in a β-barrel kind of a structure connected by two short and a longer loop and a 3_{10} helix. Ligands recognised by SH3 domain contains a stretch of PXXP motif forming a poly-proline helix. A pair of 'XP' residues fits in each of the two consensus pockets on the surface of SH3 domain. Sh3 domain ligands have been further divided into two types: Class I and class II with '(R/K)xxPxxP ' and 'xPxxPx(R/K)' motifs respectively which binds onto the SH3 surface in two opposite directions. Lysine and arginine residues increases the specificity of interaction and occupies 'specificity pocket' in continuation with the 'xP' pockets on the SH3 surface.



Representation of SH3 domain ligand recognition in a 'plus' (C←N) (left) or 'minus' (N→C orientation (right) [18]

Apart from these conventional ligands, many atypical ligands have been reported to be recognised by the SH3 domain [18]. Though all the SH3 domains recognise polyproline motifs still they are very specific for their respective ligands.

1.5 An overview of Sh3 domains of amoeboid myosin I:

SH3 domain in long-tailed amoeboid myosins directly or indirectly regulate Arp2/3 both in lower and higher organisms, in turn plays a significant role in actin nucleation and organisation. In vivo functions of Sh3 domain of different amoeboid myosins are described one by one in this section.

1.5.1 Class I myosin in *Acanthamooeba castallani* (Ac):

Acanthamoeba has three class I myosins isoforms: Myosin IA, IB and IC. All the three are long-tailed myosin. Localization pattern of active form i.e. phosphorylated form of these myosin I isoforms demonstrated the function of each isoform. Myosin IA is involved in phagocytosis; Myosin IB is associated with both phagocytosis as well as pinocytosis; and myosin IC found to be associated with contractile vacuoles [17,19].

SH3 domain of AcMyosin IA and IC interacts with Acan125, a homologue of CARMIL protein through polyproline motif present in it [12,20-22]. SH3 domain of AcMyosin IB had also been crystallized with a polyproline motif of Acan125. (PDBid 2DRK, 2DRM). Acan125 also colocalizes on cellular organelles with AcMyosin I [22]. The C-terminal of Acan125 has leucine rich repeat regions similar to WASP protein which binds to Arp2/3 complexes and activates them for nucleation of filamentous actin [23].

1.5.2 Class I Myosin in *Dictyostelium* discoideum (Dd):

Dictyostelium has seven myosin I isoforms, MyoA, MyoE and MyoF are short-tailed isoforms while MyoB, MyoC and MyoD are the classical amoeboid myosin with all the three tail homology domains. MyoK, the seventh isoform is however, unique myosin isoform which lack tail region and has an additional actin binding loop in its motor domain. Overexpression of MyoB had been observed to overcome the effect of MyoA deletion and maintain normal growth, cortical actin distribution and endocytosis. Functional SH3 domain and phosphorylation site are important for MyoB to produce such effect. [24]

The SH3 domain of MyoB and MyoC, binds a 116kDa protein, initially known as P116 but now is named as CARMIL (homologous to Acan125 of *Acanthamoeba*). It colocalises with myosin I and Arp2/3 at the leading edges of moving *Dd*.[25]. It leads to the formation of a multimeric complex containing capping protein, Arp2/3 complex and myosin I which leads to the nucleation and branching of actin filament in locomoting cells. Also, an increase or decrease in the expression of mammalian CARMIL protein modulates the lamellipodial protrusions [26]. CARMIL shares homology with WASP proteins.

1.5.3 Myosin I of Saccharomyces cerevisiae (Sc):

Sc has two highly similar myosin I proteins, Myo3 and Myo5. Both belong to long-tailed myosin I class. Expression of full length Myo5 has been found to rescue the myo3/myo5 null double mutants and restores normal wild type phenotype. While Sh3 deleted Myo5 fails to overcome the effect of myo3-/5- double mutation. Myo5 SH3 domain control its function by controlling localization. [27].

Sh3 domains of Myo3 and Myo5 interact with Bee1p/Las17p, homologous to WASP-like adaptor protein and verpolin (Vrp1p), a homologue of human WIP adaptor protein (WASP interacting protein). These adaptor proteins link various molecules to actin assembly. Interaction with Vrp1 is crucial for localization of

Myo5 to cortical actin [27]. Bee1p interacts with Arp2/3 which nucleates actin filament and hence, Myo3 and Myo5 binding to Bee1p indirectly regulates actin dynamics within the cells at sites of polarised growth and organisation of F-actin [28]. Moreover, Myo5 and Myo3 can also directly interact with Arp2/3 complex. TEDS site phosphorylation is crucial to stimulate this Arp2/3 mediated actin organisation [12,28].

1.5.4 *Aspergillus nidulans* (An) Class I myosin:

Aspergillus has a single myosin I, MyoA which is an essential gene and the mutants show defect in protein secretion and hyphal growth [29]. MyoA co-localises with the cortical actin patches present all over the cell. In contrast to *Dictyostelium* and *Saccharomyces*, deletion of Sh3 domain and mutation in TEDS phosphorylation site gives wildtype phenotype and does not affect MyoA function. It was hypothesized that instead of interacting with a proline rich ligand though SH3 domain like Bee1p, MyoA might interact with a SH3 containing ligand to regulate the actin assembly. The function of MyoA SH3 domain could be to mask its adjacent polyproline region as in AcMyosin IA [21,30,31].

1.5.5 Amoeboid Class I myosin in Vertebrates:

Vertebrates have eight isoforms of myosinI, of which two MyoIE and MyoIF are amoeboid type myosin I with a Sh3 domain in their C-terminal tail. MyoIF is expressed in hematopoietic cells, predominantly while MyoIE is highly expressed in all the mammalian cells. It is found associated with sites of cell–cell contact and in macrophages, it is localized to phagocytic cups [32]. Its homologue in rat is called myr3 and in humans, it was previously known as Myosin IC. MyoIE SH3 domain when overexpressed alone recruits WIRE, WIP and their interacting partner N-WASP proteins to activate Arp2/3 complex and initiate actin polymerization during clathrin-mediated endocytosis [33]. MyoIF depleted neutrophils reduces actin polymerisation in cortical actin [34]. MyoIF exists as a fusion protein with Mixed Lineage Leukemia (MLL) gene in acute infant myeloid

leukemia patients. It has been hypothesized that the SH3 domain of MyoIF interacts with the proline rich region in MLL and induces its oligomerisation [35].

1.6 Role of Myosin I in actin assembly and endocytosis.

There are three main models which have been proposed for the assembly of actin filaments during diverse cellular events. In the cytosol, a pool of G-actin bound to ATP is present bound to profilin in a readily polymerizable state. This profilin-actin complex recognises polyproline sequence, and is recruited at the site of actin polymerisation via several activator proteins like VASP, WASp etc. [36]. The basic proposed models of actin assembly are as follows:

a. De-novo actin nucleation takes place by means of Arp2/3 complex. Nucleation occurs mainly by branching of the existing actin filaments, and the force to push the cellular membranes is generated by polymerization of ATP-G-actin itself. The newly generated actin filament is rapidly capped by capping proteins followed by stabilization by several actin binding proteins. ADP/cofilin binding to ADP bound F-actin initiates the actin depolymerisation. Profilin reconstitutes the pool of ATP-G-actin by acting as a nucleotide exchange factor [37].
b. The alternative mechanism proposes that extracellular signals are relayed through effector proteins like Rho family GTPases to the VASP, SCAR or WASp family proteins, leading to localized recruitment and activation of the Arp2/3 complex. Recruitment and phosphorylation of filamin protein takes place by PAK kinase which together lead to orthogonally cross-linked actin network [38].
c. Another model to generate unbranched actin filament, involves filamin protein which act under the influence of Rho GTPases. Filamin is able to nucleate actin filament and remains associated with the barbed end of growing F-actin [39].

d. The dendritic nucleation model defines the formation of branched actin filaments mediated by Arp2/3 complex which are activated by Nucleation promoting factors as well as pre-existing filaments. Growing filaments are capped by capping proteins and cofilin starts filament depolymerisation by binding to ADP-F-actin. Profilin maintains the pool of ATP-G-actin [36,40].

On the basis of immunolocalization studies in amoeba and higher eukaryotes, class I myosin have been found to be associated with dynamic region of cell cortex such as sites of particle ingestion, pseudopods of migrating cells, endocytic structures, cell-cell junctions and microvilli [12,41]. In soil amoeba, Dd myoB and myoK have been reported to localize in phagocytic cups and membrane ruffles. Overexpression of myoB leads to decrease in micropinocytosis [42-44]. Cells lacking myok and myoB have reduced phagocytosis. Moreover, myoB null mutant have defect in membrane recycling from endosomes [42,45,46]. Myosin I mutant and overexpressed cell in *Acanthamoeba* and *Entamoeba* have also shown defects in endocytic processes and cell motility [47,48].

Similar results have been seen in Sc Myo3 and Myo5 double mutants, which has shown impaired endocytosis, growth defects and accumulation of intracellular vesicles. In *Aspergillus*, focal secretion and generation of cell polarity requires a functional MyoA [49]. In higher eukaryotes, Myo Ie and Myo If are homologues of amoeboid myosin I have been found to be involved in endocytosis [41]. Many short tailed myosins in lower as well as higher organisms associate with endosomes and lysosomes and drives the vesicular transport [50].

Cell movement, phagocytosis and macropinocytosis all are highly dependent on actin cytoskeleton for extension of membrane protrusions. Sh3 domain located in the tail region of myosin I plays a prominent role in bridging the process of actin assembly and endocytosis. In yeast, Sc the two myosins recuits Arp2/3 complex at the site of actin nucleation by direct interaction with acidic region present in the c-terminal tail [51-54]. Similar acidic region are present in class I myosin of other fungus like *Candida albicans, S. pombe and Aspergillus*; and Arp2/3 recruiting proteins like Las17p and WASp [36]. Sh3 domain of Myo3 and

Myo5 interacts with Vrp1 and Las17p, essential for recruitment of myosin I to cortical actin nucleation sites and arp2/3 recruitment as well [52]. WH2 domain present in Vrp1 and Las17p are able to bind G-actin during actin polymerisation independent of profilin. In amoeba, SH3 domain of myosin I interacts with adapter protein like CARMIL and Acan125 which recruits and activate Arp2/3. Two different ways have been proposed for class I myosins to concentrate G-actin for polymerisation. They can either directly bind to profilin actin (DdMyoK) or via WH2 domain of adaptor proteins (WASp, CARMIL and Vrp1) [55].

Vertebrate myosin Ie , has been proposed to be involved in clathrin-mediated endocytosis and recuits the actin-assembly factors N-WASP, WIP (WASP-interacting protein) and WASP-binding protein, WIRE as well as proteins, such as synaptojanin-1 and dynamin involved in other clathrin-mediated endocytic [33,56].

Myosin I are plays a significant role in bridging endocytic process both directly by actin interaction with motor domain and G-actin transport, and indirectly, by recruiting factors associated with nucleation, polymerization and stabilization of actin filaments.

1.7 *Entamoeba histolytica*: A highly motile phagocytic protozoan.

Entamoeba histolytica is a highly motile, unicellular parasitic protozoan that causes amoebic dysentery (Amoebiasis) disease in humans. It is a common health hazard in developing countries. Though 90% of the cases are asymptomatic, still it is the third leading cause of death among parasitic diseases resulting in 100,000 deaths annually [57].

Scientific Classification:

Kingdom: Eukaryota

Phylum: Amoebazoa

Class: Archamoeba

Order: Amoebida

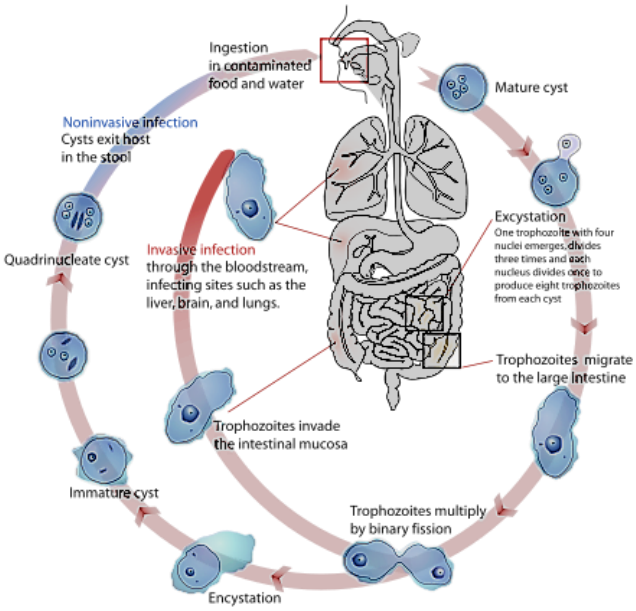
Genus: *Entamoeba*

Species: *histolytica*

Binomial Nomenclature: *Entamoeba histolytica*

1.7.1 Lifecycle and Pathogenesis:

Entamoeba has 2 stage life cycle: infective cyst stage and motile trophozoites stage. Infection begins in humans by ingestion of cyst via contaminated food and water. Each cyst divides and give rise to eight trophozoites. These trophozoites migrate to large intestine and may invade intestinal mucosa causing intestinal amoebiasis. Infection becomes invasive and more chronic if the trophozites reach liver, brain and blood through blood stream leading to formation of life-threatening abscess [58]. However, the sequence of events which determine encystation and excystation and transform the commensal pathogen into an invasive virulent form is unknown.



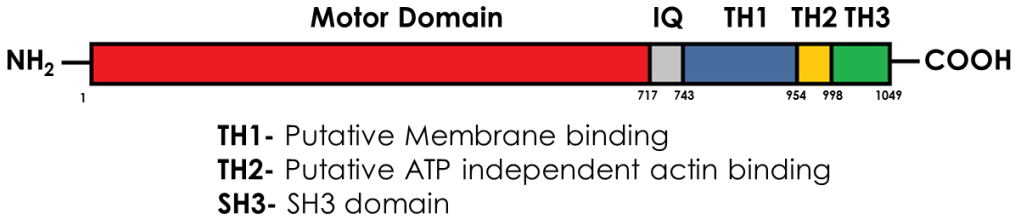
Life cycle of *E.histolytica*

(Adapted from internet from *Entamoeba* histolytica life cycle-en.svg)

During invasive amoebiasis, *Entamoeba* first adheres to the cell surface via several cell surface receptors and accessory effectors like Gal/GalNAc-specific lectins, SREHP, KERP, TMK, *Eh*LPPG,etc. Adherence is followed by cytolysis of the target cells by involvement of amoebapores and cysteine proteases. The final step is phagocytosis of target cells (intestinal epithelial cells, immune and non-immune cells. Phagocytic deficient mutants of *E.histolytica* are avirulent implying the importance of phagocytosis in *Entamoeba* pathogenesis. Phagocytosis in associated with intensive cytoskeletal remodelling which involve actin filaments, several actin binding proteins and myosin [59].

1.8 Unconventional myosin I in *E. histolytica*.

Entamoeba genome encodes for only two myosin isoforms: Myosin II and unconventional myosin IB. Cytoskeletal studies on have revealed that very low levels of conventional myosin II is present in the pseudopods and phagocytic cup [60]. The gene encoding myosin I in *E. histolytica* was found to be homologous to myosin IB of other organisms and was amplified as 3150 bp fragment, 1049 amino acids in length with a molecular weight of 130kDa [61].



*Eh*Myosin IB is a long tailed myosin I with a motor head domain, neck containing a single IQ motif and a tail with membrane binding domain followed by a short ATP independent actin binding region and a c-terminal SH3 domain. It shows around 70% similarity with myosin IB from *Dictyostelium* and *Acanthamoeba*. Few features of motor domain like presence of glutamic acid at consensus phosphorylation site as in metazoans and absence of DALAK sequences speculates that *Eh*Myosin I has evolved in a similar way as metazoan myosins [61]. Calcium

binding protein-5 (CaBP5) which has much similarity to the essential light chain than to calmodulin binds to the IQ motif present in the neck region [14]. The tail of *Eh*myosin IB has GPA rich domain much shorter than in other amoeba like *Dictyostelium* and *Acanthamoeba*. The Sh3 domain of other myosins have been involved in recruiting proteins involved in actin polymerisation during endocytosis but nothing is known regarding the function of *Eh*Myosin IB Sh3 domain.

In resting *E.histolytica* cells, myosin IB is localised in the cytoplasm, beneath the plasma membrane, and enriched at the cortical region. Myosin IB relocalizes to the pseudopods where actin filaments exist in gelled state in polarized amoeba cells undergoing phagocytosis [48]. Overexpression of myosin IB delays phagocytic cup formation which had been co-related with an increase in viscosity of actin cytoskeleton. The presence of two actin binding site leading to crosslinking of cytoplasmic actin filament has been associated with the increase in cytoplasmic density [62].

In contrast to closely related amoeba, like *Dictyostelium* and *Acanthamoeba* which have seven and three myosin I isoforms which carry out different as well as overlapping function during cell growth, motility and vesicular trafficking processes; *Entamoeba* has only unconventional myosin IB. Involvement of Myosin I during endocytosis and actin assembly in other organisms, encouraged us to study the role of this protein and its putative downstream and upstream effectors that takes part in the signalling cascade via Myosin IB during phagocytosis of RBC and intestinal cells by this highly motile parasitic pathogen.

1.9 References:

[1] Sebé-Pedrós, A., Grau-Bové, X., Richards, T.A. and Ruiz-Trillo, I. (2014). Evolution and Classification of Myosins, a Paneukaryotic Whole-Genome Approach. Genome Biology and Evolution 6, 290-305.

[2] Hodge, T., Jamie, M. and Cope, T.V. (2000). A myosin family tree. Journal of Cell Science 113, 3353-3354.

[3] Richards, T.A. and Cavalier-Smith, T. (2005). Myosin domain evolution and the primary divergence of eukaryotes. Nature 436, 1113-1118.

[4] Hartman, M.A. and Spudich, J.A. (2012). The myosin superfamily at a glance. J Cell Sci 125, 1627-1632.

[5] O'Connell, C.B., Tyska, M.J. and Mooseker, M.S. (2007). Myosin at work: Motor adaptations for a variety of cellular functions. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1773, 615-630.

[6] Ostap, E.M. and Pollard, T.D. (1996). Biochemical kinetic characterization of the Acanthamoeba myosin-I ATPase. The Journal of Cell Biology 132, 1053-1060.

[7] Jontes, J.D., Milligan, R.A., Pollard, T.D. and Ostap, E.M. (1997). Kinetic characterization of brush border myosin-I ATPase. Proceedings of the National Academy of Sciences 94, 14332-14337.

[8] Mezgueldi, M.E., Tang, N., Rosenfeld, S.S. and Ostap, E.M. (2002). The Kinetic Mechanism of Myo1e (Human Myosin-IC). Journal of Biological Chemistry 277, 21514-21521.

[9] Pollard, T.D., Doberstein, S.K. and Zot, H.G. (1991). Myosin-I. Annu Rev Physiol 53, 653-81.

[10] Cheney, M.S.M.a.R.E. (1995). Unconventional Myosins. Annual Review of Cell and Developmental Biology 11, 633-675.

[11] Mooseker, M.S. and Cheney, R.E. (1995). Unconventional myosins. Annu Rev Cell Dev Biol 11, 633-75.

[12] Coluccio, L.M. (1997). Myosin I. American Journal of Physiology - Cell Physiology 273, C347-C359.

[13] Côté, G.P., Albanesi, J.P., Ueno, T., Hammer, J.A. and Korn, E.D. (1985). Purification from Dictyostelium discoideum of a low-molecular-weight myosin that resembles myosin I from Acanthamoeba castellanii. Journal of Biological Chemistry 260, 4543-6.

[14] Kumar, S. et al. (2014). Crystal Structure of Calcium Binding Protein-5 from Entamoeba histolytica and Its Involvement in Initiation of Phagocytosis of Human Erythrocytes. Plos Pathogens 10

[15] Cyr, J.L., Dumont, R.A. and Gillespie, P.G. (2002). Myosin-1c Interacts with Hair-Cell Receptors through Its Calmodulin-Binding IQ Domains. The Journal of Neuroscience 22, 2487-2495.

[16] Hokanson, D.E., Laakso, J.M., Lin, T., Sept, D. and Ostap, E.M. (2006). Myo1c Binds Phosphoinositides through a Putative Pleckstrin Homology Domain. Molecular Biology of the Cell 17, 4856-4865.

[17] Doberstein, S. and Pollard, T. (1992). Localization and specificity of the phospholipid and actin binding sites on the tail of Acanthamoeba myosin IC. The Journal of Cell Biology 117, 1241-1249.

[18] Kaneko, T., Li, L. and Li, S. (2008). The SH3 domain--a family of versatile peptide- and protein-recognition module. Front Biosci 13, 4938-4952.

[19] Baines, I.C., Corigliano-Murphy, A. and Korn, E.D. (1995). Quantification and localization of phosphorylated myosin I isoforms in Acanthamoeba castellanii. The Journal of Cell Biology 130, 591-603.

[20] Xu, P., Mitchelhill, K.I., Kobe, B., Kemp, B.E. and Zot, H.G. (1997). The myosin-I-binding protein Acan125 binds the SH3 domain and belongs to the superfamily of leucine-rich repeat proteins. Proceedings of the National Academy of Sciences 94, 3685-3690.

[21] Osherov, N. and May, G.S. (2000). In vivo function of class I myosins. Cell Motility and the Cytoskeleton 47, 163-173.

[22] Xu, P., Zot, A.S. and Zot, H.G. (1995). Identification of Acan125 as a myosin-I-binding protein present with myosin-I on cellular organelles of Acanthamoeba. J Biol Chem 270, 25316-9.

[23] Svitkina, T.M. and Borisy, G.G. (1999). Progress in protrusion: the tell-tale scar. Trends in Biochemical Sciences 24, 432-436.

[24] Novak, K.D. and Titus, M.A. (1998). The Myosin I SH3 Domain and TEDS Rule Phosphorylation Site are Required for In Vivo Function. Molecular Biology of the Cell 9, 75-88.

[25] Jung, G., Remmert, K., Wu, X., Volosky, J.M. and Hammer, J.A. (2001). The Dictyostelium Carmil Protein Links Capping Protein and the Arp2/3 Complex to Type I Myosins through Their Sh3 Domains. The Journal of Cell Biology 153, 1479-1498.

[26] Yang, C., Pring, M., Wear, M.A., Huang, M., Cooper, J.A., Svitkina, T.M. and Zigmond, S.H. (2005). Mammalian CARMIL inhibits actin filament capping by capping protein. Dev Cell 9, 209-21.

[27] Anderson, B.L., Boldogh, I., Evangelista, M., Boone, C., Greene, L.A. and Pon, L.A. (1998). The Src homology domain 3 (SH3) of a yeast type I myosin, Myo5p, binds to verprolin and is required for targeting to sites of actin polarization. The Journal of cell biology 141, 1357-1370.

[28] Lechler, T., Shevchenko, A. and Li, R. (2000). Direct involvement of yeast type I myosins in Cdc42-dependent actin polymerization. J Cell Biol 148, 363-73.

[29] McGoldrick, C.A., Gruver, C. and May, G.S. (1995). myoA of Aspergillus nidulans encodes an essential myosin I required for secretion and polarized growth. J Cell Biol 128, 577-87.

[30] Osherov, N., Yamashita, R.A., Chung, Y.-S. and May, G.S. (1998). Structural Requirements for in Vivo Myosin I Function in Aspergillus nidulans. Journal of Biological Chemistry 273, 27017-27025.

[31] Yamashita, R.A., Osherov, N. and May, G.S. (2000). Localization of wild type and mutant class I myosin proteins in Aspergillus nidulans using GFP-fusion proteins. Cell Motility and the Cytoskeleton 45, 163-172.

[32] Swanson, J.A., Johnson, M.T., Beningo, K., Post, P., Mooseker, M. and Araki, N. (1999). A contractile activity that closes phagosomes in macrophages. J Cell Sci 112 (Pt 3), 307-16.

[33] Cheng, J., Grassart, A. and Drubin, D.G. (2012). Myosin 1E coordinates actin assembly and cargo trafficking during clathrin-mediated endocytosis. Mol Biol Cell 23, 2891-904.

[34] Kim, S.V. et al. (2006). Modulation of Cell Adhesion and Motility in the Immune System by Myo1f. Science 314, 136-139.

[35] Duhoux, F.P., Ameye, G., Libouton, J.M., Bahloula, K., Iossifidis, S., Chantrain, C.F., Demoulin, J.B. and Poirel, H.A. (2011). The t(11;19)(q23;p13) fusing MLL with MYO1F is recurrent in infant acute myeloid leukemias. Leuk Res 35, e171-2.

[36] Soldati, T. (2003). Unconventional myosins, actin dynamics and endocytosis: a menage a trois? Traffic 4, 358-366.

[37] Blanchoin, L., Pollard, T.D. and Mullins, R.D. (2000). Interactions of ADF/cofilin, Arp2/3 complex, capping protein and profilin in remodeling of branched actin filament networks. Current Biology 10, 1273-1282.

[38] Pollard, T.D., Blanchoin, L. and Mullins, R.D. (2000). Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. Annual review of biophysics and biomolecular structure 29, 545-576.

[39] Pruyne, D., Evangelista, M., Yang, C., Bi, E., Zigmond, S., Bretscher, A. and Boone, C. (2002). Role of formins in actin assembly: nucleation and barbed-end association. Science 297, 612-615.

[40] Nicholson-Dykstra, S., Higgs, H.N. and Harris, E.S. (2005). Actin Dynamics: Growth from Dendritic Branches. Current Biology 15, R346-R357.

[41] Stoffler, H., Honnert, U., Bauer, C.A., Hofer, D., Schwarz, H., Muller, R., Drenckhahn, D. and Bahler, M. (1998). Targeting of the myosin-I myr 3 to intercellular adherens type junctions induced by dominant active Cdc42 in HeLa cells. Journal of Cell Science 111, 2779-2788.

[42] Schwarz, E.C., Neuhaus, E.M., Kistler, C., Henkel, A.W. and Soldati, T. (2000). Dictyostelium myosin IK is involved in the maintenance of cortical tension and affects motility and phagocytosis. J Cell Sci 113, 621-633.

[43] Fukui, Y., Lynch, T.J., Brzeska, H. and Korn, E.D. (1989). Myosin I is located at the leading edges of locomoting Dictyostelium amoebae. Nature 341, 328-331.

[44] Novak, K.D., Peterson, M.D., Reedy, M.C. and Titus, M.A. (1995). Dictyostelium myosin I double mutants exhibit conditional defects in pinocytosis. The Journal of Cell Biology 131, 1205-1221.

[45] Jung, G. and Hammer, J.A. (1990). Generation and characterization of Dictyostelium cells deficient in a myosin I heavy chain isoform. The Journal of Cell Biology 110, 1955-1964.

[46] Jung, G., Wu, X. and Hammer, J.A. (1996). Dictyostelium mutants lacking multiple classic myosin I isoforms reveal combinations of shared and distinct functions. The Journal of Cell Biology 133, 305-323.

[47] Baines, I.C., Brzeska, H. and Korn, E.D. (1992). Differential localization of Acanthamoeba myosin I isoforms. The Journal of Cell Biology 119, 1193-1203.

[48] Voigt, H., Olivo, J.C., Sansonetti, P. and Guillen, N. (1999). Myosin IB from Entamoeba histolytica is involved in phagocytosis of human erythrocytes. Journal of Cell Science 112, 1191-1201.

[49] McGoldrick, C.A., Gruver, C. and May, G.S. (1995). myoA of Aspergillus nidulans encodes an essential myosin I required for secretion and polarized growth. The Journal of Cell Biology 128, 577-587.

[50] Raposo, G., Cordonnier, M.-N., Tenza, D., Menichi, B., Dürrbach, A., Louvard, D. and Coudrier, E. (1999). Association of myosin I alpha with endosomes and lysosomes in mammalian cells. Molecular biology of the cell 10, 1477-1494.

[51] Evangelista, M. et al. (2000). A role for myosin-I in actin assembly through interactions with Vrp1p, Bee1p, and the Arp2/3 complex. The Journal of cell biology 148, 353-362.

[52] Lee, W.-L., Bezanilla, M. and Pollard, T.D. (2000). Fission yeast myosin-I, Myo1p, stimulates actin assembly by Arp2/3 complex and shares functions with WASp. The Journal of cell biology 151, 789-800.

[53] Lechler, T., Shevchenko, A., Shevchenko, A. and Li, R. (2000). Direct involvement of yeast type I myosins in Cdc42-dependent actin polymerization. The Journal of cell biology 148, 363-374.

[54] Geli, M.I., Lombardi, R., Schmelzl, B. and Riezman, H. (2000). An intact SH3 domain is required for myosin I-induced actin polymerization. The EMBO Journal 19, 4281-4291.

[55] Paunola, E., Mattila, P.K. and Lappalainen, P. (2002). WH2 domain: a small, versatile adapter for actin monomers. FEBS letters 513, 92-97.

[56] Krendel, M., Osterweil, E.K. and Mooseker, M.S. (2007). Myosin 1E interacts with synaptojanin-1 and dynamin and is involved in endocytosis. FEBS Letters 581, 644-650.

[57] Stanley, S.L., Jr. (2003). Amoebiasis. The Lancet 361, 1025-1034.

[58] Ravdin, J. and Petri, W. (1990). Entamoeba histolytica (amebiasis). Principles and practice of infectious diseases., 2036-2049.

[59] Nozaki, T. and Bhattacharya, A. (2014) Amebiasis: Biology and Pathogenesis of Entamoeba, Springer

[60] Arhets, P., Gounon, P., Sansonetti, P. and Guillen, N. (1995). Myosin II is involved in capping and uroid formation in the human pathogen Entamoeba histolytica. Infection and Immunity 63, 4358-4367.

[61] Vargas, M., Voigt, H., Sansonetti, P. and Guillen, N. (1997). Molecular characterization of myosin IB from the lower eukaryote Entamoeba histolytica, a human parasite. Mol Biochem Parasitol 86, 61-73.

[62] Marion, S., Wilhelm, C., Voigt, H., Bacri, J.-C. and Guillén, N. (2004). Overexpression of myosin IB in living Entamoeba histolytica enhances cytoplasm viscosity and reduces phagocytosis. Journal of Cell Science 117, 3271-3279.

2. CHAPTER I

Structural Studies of SH3 domain of *Eh*Myosin IB

2.1 Introduction:

Myosin I are single-headed actin based molecular motors which forms the largest class of unconventional myosins [1]. Unlike conventional two-headed myosins, they show relatively slow kinetics and do not form filaments [1]; they instead are involved in organelle movement, vesicle trafficking, phagocytosis, pinocytosis, membrane ruffling, or formation of lamellipodia [2]. Myosin I was the first unconventional myosin to be discovered, and it was first discovered in *Acanthamoeba*. *Acanthamoeba* has 3 myosin I isoforms denoted as myosin IA, IB, and IC, and these isoforms show differential localization patterns and are associated with different membrane and actin-based cellular functions [3]. *Dictyostelium* has six myosin I isoforms; out of which single null mutants of only MyoB and MyoC show reduced phagocytosis. [4]. Interestingly, only these isoforms, MyoB and MyoC have a c-terminal SH3 domain while SH3 domain although present in myoD but is followed by short membrane binding region [4]. In contrast to these closely related amoebae, *Entamoeba* has only a single unconventional myosin I isoform, myosin IB. It will be interesting to study, how myosin IB is solely carrying out a different variety of functions in *Entamoeba*.

Myosin IB of *Entamoeba* localizes in phagocytic cups from initiation until cup closure during erythrophagocytosis while its light chain CaBP5 leaves after the closure of the phagocytic cup [5]. Overexpression of myosin IB has been shown to lead to a decrease in phagocytosis by *Entamoeba* cells due to defect in the phagosome initiation process [6] while downregulation of its light chain CaBP5 expression, led to a reduction in phagocytosis [5]. Also, over-expression of actin binding protein, coactosin led to 80% reduction in the number of phagocytic cups leading to a decrease in phagocytosis similar myosin IB [7] while other actin binding protein, ArpC1 showed significant reduction in phagocytosis, on suppressed expression [8]. Phagocytosis is a crucial event for E. histolytica pathogenesis during invasive amoebiasis and involves intensive cytoskeletal

remodeling. Highly infective amoebae have been associated with high motility and phagocytic behavior.

Myosin IB from E. histolytica is a long-tailed myosin I with a C-terminal SH3 domain [9]. SH3 domains are important for proper functioning of such long-tailed myosin Is. The SH3 domain of Myo5p in S. cerevisiae is required for localization of actin patches at the site of polarized cell surface growth and performs this localization by interacting with Vrp1p [10]. An SH3 deletion mutant of *Dictyostelium* myoB failed to complement the myoB null phenotype [11]. The SH3 domain of *Dictyostelium* myoB interacts with the CARMIL protein, which in turn binds to the Arp2/3 complex and capping proteins leading to the formation of a large complex associated with organization and regulation of actin cytoskeleton in the cortical region and phagocytic cup. [12].

The SH3 domain of myosin IB from E. histolytica (*Eh*MySH3) has a high sequence homology with the SH3 domains of other myosins and tyrosine kinases (**Error! Reference source not found.**). The SH3 domains recognize ligands with the consensus PXXP motif and each 'XP' pair binds into two highly conserved pockets [13,14]. Such polyproline ligands have been further divided into two classes — class I ligands [(R/K)xxPxxP] and class II ligands [xPxxPx(R/K)] — on the basis of whether a positively charged residue, which can be arginine or lysine and is located at the N-terminus or C-terminus of the motif. Positively charged residue occupies a third negatively charged pocket, known as specificity pocket and decides the orientation (+ or -) of peptide binding. Class I and II ligands occupy the same three pockets but in opposite orientation which changes the directionality of amide backbone and interchanges the positioning of critical prolines in the 'XP' pockets [15].

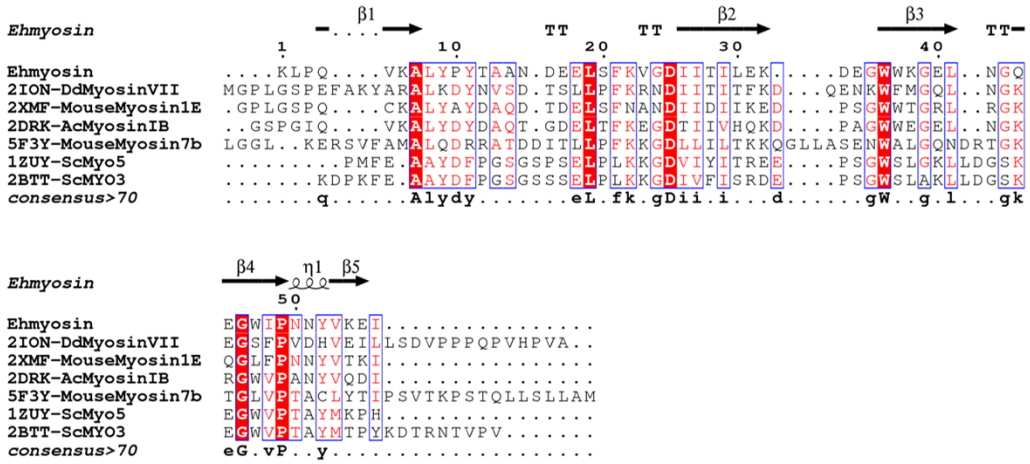


Figure 2.1-I: Structural alignment of SH3 domains.

Structural alignment of SH3 domains from various myosins and from Src using ESPript [16], showing mostly conserved residues outlined in rectangular blocks and totally conserved residues with a red background. (*Eh-Entamoeba histolytica*; Dd-*Dictyostelium discoideum*; Ac-*Acanthamoeba castellanii*; Sc- *Saccharomyces cerevisiae*).

This chapter mainly deals with the determination of the high resolution crystal structure of the myosin IB SH3 domain from *E.histolytica* at 1.7 A^{o} in two different conditions, including one where the SH3 domain was observed to be bound to PEG molecules. Comparison of the structure of the PEG-bound SH3 domain with peptide-bound SH3 structures indicated that PEG molecules were bound at the protein-protein interaction surface. On the basis of PEG-bound *Eh*MySH3 structure we have studied the

2.2 Materials and Methods:

2.2.1 Cloning of C-terminal SH3 domain of unconventional myosin IB from *E.histolytica* (*Eh*MySH3):

EhMySH3 construct DNA sequence (nucleotides position 3007-3161) obtained from NCBI:

ttgccacaagttttgccacaagttaaagcactctatccatatactgctgctaatgacgaagagttatcattcaaagta
ggtgatatcattactattctcgaaaaagatgaaggttggtggaaaggagaattgaatggacaagaaggctggatt
cctaacaactatgtcaaagaaatttaa

EhMySH3 Protein Sequence (amino acid position 994-1049) obtained from NCBI:

LPQVKALYPYTAANDEELSFKVGDIITILEKDEGWWKGELNGQEGWIPNNYVKEI

2.2.1.1 Primer Designing:

The *Entamoeba* histolytica (HM-1: IMSS) unconventional myosin IB gene sequence was taken from NCBI database and gene specific primers were designed for *Eh*MySH3 domain construct on the basis of minimum Tm difference between forward primer and reverse primer using freely available oligo software for Tm calculation. Following are the primers sequences:

FwMy1SH3 :	5'- CTAGCTAGCATGTTGCCACAAGTTAAAGCACTCTATC- 3'	Nhe I
RvMy1TD:	5'-CCGCTCGAGAATTTCTTTGACATAGTTGTTAGGAATC- 3'	Xho I

2.2.1.2 PCR amplification:

Myosin IB SH3 domain (*Eh*MySH3) construct (amino acid positions 994-1049) was PCR amplified using FwMy1SH3 and RvMy1TD primers from the genomic DNA of *Entamoeba histolytica* strain HM1: IMSS. The amplified fragment was cloned in

pET21c vector (Novagen) between NheI and XhoI restriction sites with a C-terminal hexa-histidine tag. Reaction mix for PCR amplification was prepared as mentioned in appendix (A.4) and run in a thermocycler for 30 cycles with cycling parameters as mentioned in appendix (A.4) with 53^{o}C as annealing temperature. The amplified PCR product and the vector pET21c were double digested with Nhe1/Xho1 restriction enzymes and gel purified. These digested and purified products were ligated using T4DNA ligase and kept at 16°C for 16hrs. The ligated mixture was then transformed to E.Coli DH5α cells, plated on ampicillin containing LB agar plates and kept at 37°C for overnight. The colonies were screened for positive clones by colony PCR. Plasmid was isolated from positive colonies. The clone was further confirmed by double digestion of isolated plasmid. (Detailed steps are given in appendix A.4).

2.2.2 Overexpression and purification of *Eh*MySH3:

2.2.2.1 Overexpression of *Eh*MySH3 protein in *E.coli*:

The cloned *Eh*MySH3 gene was expressed in E coli BL21 (DE3) cells, which were grown in LB media supplemented with 100µg/ml of ampicillin at 37ºC inoculated with 1% inoculums of overnight grown primary culture. The culture was induced by 0.2mM IPTG concentration at OD600 = 0.6 and grown at 37ºC for another 5 hours. The bacterial cells were harvested by centrifugation at 6000 rpm for 10 minutes at 4ºC. The harvested cells were resuspended in lysis buffer (50mM Tris pH 8.0, 150mM NaCl.). The resuspened cells were subjected to five cycles of freeze and thaw. The thawed cells were subjected to sonication at 4°C for 5 x 30sec at 20 amplitude with an interval of 30 seconds. The cell lysate was centrifuged at 13,000 rpm for 30 min. The pellet was discarded and supernatant was then loaded to Ni-NTA column pre-equilibrated with lysis buffer.

2.2.2.2 Purification of *Eh*MySH3 by Nickel-NTA chromatography:

The clear supernatant obtained was loaded onto Ni-NTA column (Sigma) prepacked with Ni-sepharose beads (GE healthcare). The *Eh*MySH3 protein was eluted with elution buffer (50 mM Tris, 150 mM NaCl, 100 mM imidazole) after three wash cycles with wash

buffers (50 mM Tris pH 7.4; 50 mM Tris pH 7.4, 300 mM NaCl; and 50 mM Tris pH 7.4, 150 mM NaCl, 20 mM imidazole). The eluted fractions were checked for purity on 15% SDS PAGE prepared as mentioned in appendix (A.3).

2.2.2.3 Purification of *Eh*MySH3 by Gel Permeation chromatography:

The eluted fractions from Ni+2- NTA column chromatography were pooled, concentrated using 3kDa cut-off centricon (Amicon, Millipore) and injected into Superdex G75 16/60 column (GE Healthcare) pre-equilibrated in buffer containing 20 mM Tris pH7.4, 150mM NaCl.. The elution profile of MY1SH3 was monitored using UV280 detector where the protein eluted at 70-80ml. The peak fractions were pooled together and checked on 15% SDS-PAGE.

2.2.3 Crystallization of *Eh*MySH3:

Purified *Eh*MySH3 protein was concentrated to 30 mg/ml and was subjected to crystallization using hanging drop vapour diffusion method using mosquito (TTP Labtech) with several commercial screens from Hampton Research and Molecular Dimensions. After a week, initial hit was obtained in crystal screen HT (Hampton Research) condition 2.0M Ammonium sulphate, 5%(v/v) isopropanol at 16˚C which has morphology of multiple crystals fused together in a rock-like structure and could not be broken up into single crystal. Better crystals were obtained using Additive screen (Hampton Research) and 2.0 M ammonium sulphate, 5% v/v isopropanol as reservoir buffer. Crystallization drop was set with protein and buffer in 1:1 ratio in a 2ul drop containing 0.5 µl of various additives, was allowed to equilibrate with the reservoir solutions (which did not contain the additives) in 24-well plates (Corning) at 16˚C and 4˚C. Relatively large crystals were obtained at 4˚C with 2.2 M ammonium sulphate, 5% v/v isopropanol as the reservoir solution, together with a 2 µl drop having a protein-to-buffer ratio of 1:1 and to which 0.5 µl of ethanol was added (**Error! Reference source not found.**). While these crystals were relatively large, they were still found to be fused together.

Another crystallization condition was found at 4˚C in crystal screen II (Hampton Research) condition 0.2M Ammonium sulphate, 30% PEG 8000. These crystals

were plate like single crystals but were obtained after one month. Crystals in both the conditions with and without PEG were temperature sensitive. Hence, all of the work, from observing the plates to freezing the crystals in cryo-loops, was done in the cold room.

2.2.4 Data collection, Structure determination and Refinement:

The X-ray diffraction data for crystals from both conditions were collected at the BM14 synchrotron beam line (ESRF, Grenoble, France). Crystals were flash frozen using mother liquor as cryoprotectant in both cases. The data sets were indexed and scaled using HKL2000 (Otwinowski and Minor, 1997). Each structure was determined by performing molecular replacement using the PHASER MR Program (McCoy et al., 2007) of the CCP4 program suite (Winn et al., 2011). Here, the second SH3 domain of ITSN1 from Homo sapiens, PDB code 4IIM (Structural Genomics Consortium, unpublished work) was used as the template for the *Eh*MySH3 native crystal data, and the *Eh*MySH3 structure was used as the template for the PEG-bound *Eh*MySH3 crystal data. After a few cycles of manual building, each model was refined using REFMAC5 (Murshudov et al., 2011) of the CCP4 suite followed by PHENIX refine (Afonine et al., 2012). Data statistics for both crystals are given in ***Table 2.2-1***

Table 2.2-1: Data collection and refinement statistics:

Crystallographic Data	***Eh*MySH3**	**PEG-*Eh*MySH3**
PDB code	5XGG	5XG9
X-ray Source	ERSF BEAMLINE BM14	ERSF BEAMLINE BM14
Space group	$P2_1$	C2
Wavelength (Å)	0.91	0.97
Unit Cell Parameters		
a,b,c (Å)	52.2 77.5 61.8	106.5, 79.6, 88.5
α, β, γ (°)	90, 108.2, 90	90, 122.6 , 90
Resolution range (Å)	39.55 - 1.73 (1.73-1.70)[a]	31.6 - 1.78 (1.84-1.78)
R_{sym} or R_{merge} (%)	0.11 (0.43)	0.06 (0.47)
CC1/2	0.997	0.85
Completeness (%)	99.9 (99.4)	98.9 (100)
Redundancy	6.5 (6.2)	3.8(3.8)
Mosaicity	0.4	0.7
Average I/σ(I)	25.1 (4.2)	27.14 (2.3)
No. of molecules in asymmetric unit	6	8
Refinement Statistics		
Resolution range (A°)	30.55 - 1.70	31.6 - 1.78
Total no. of observations	324450 (44626)	109431 (10854)
No. of unique observations	52037 (7477)	58346 (5567)
R_{work}	0.19 (0.18)	0.18 (0.19)
R_{free}	0.21 (0.21)	0.23 (0.24)
Mean B factor ($A^{\circ 2}$)	17.92	22.1
No. of atoms	Protein 3215/ Water 650	Protein 3752/ ligand 155
RMSD deviations		
Bonds (A°)	0.007	0.019
Bond angles (°)	1.07	1.86
Dihedral angles (°)	14.28	14.92
Cross validation error	0.180	0.215
Ramachandran Statistics		
Most Favoured Region (%)	93.4	97.7
Additionally Allowed Region (%)	4.6	2.3
Generously Allowed Region (%)	2.0	0.0
Disallowed region (%)	0	0.0

[a]Values in parenthesis are of highest resolution shell.

2.3 Results:

2.3.1 Cloning of *Eh*MySH3:

*Eh*MySH3 domain was successfully cloned in pET21c vector between NheI and XhoI restriction site. Cloning was confirmed by colony PCR and restriction digestion of recombinant plasmid with the before mentioned enzymes.

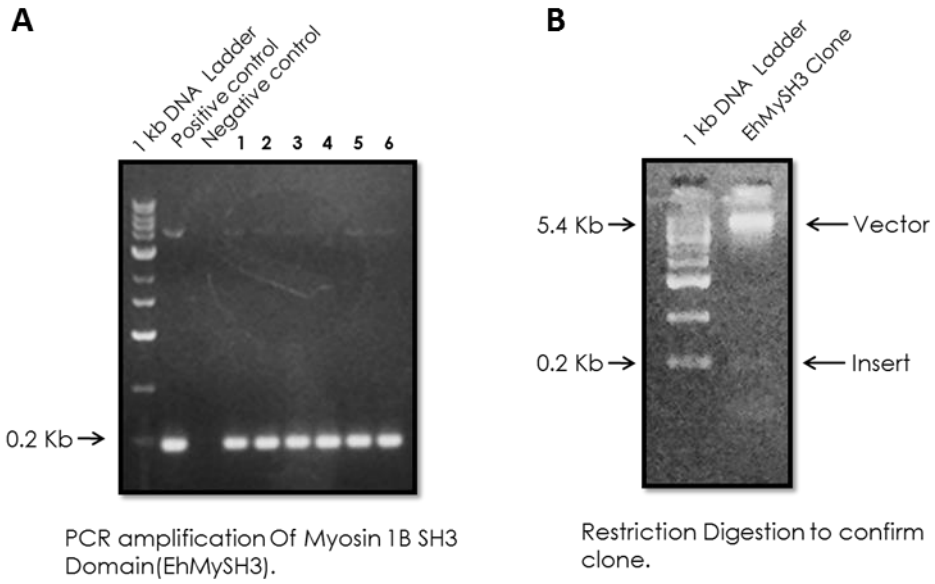


Figure 2.3-I: Cloning of EhMySH3.

(A) Agarose gel showing amplified *Eh*MySH3 domain corresponding to 0.2kb band after PCR. (B) Confirmation of *Eh*MySH3 clone by double digestion with NheI and XhoI. Presence of 0.2kb confirms *Eh*MySH3 gene and 5.4kb band corresponds to pET21c vector.

2.3.2 Overexpression, purification and crystallization of *Eh*MySH3:

Recombinant *Eh*MySh3 was overexpressed in BL21 (DE3) *E.coli* cells and purified using Nickel-NTA chromatography followed by second round of purification by gel

permeation chromatography. Elutions after each chromatography were run on 15% SDS-PAGE to check the purity of purified protein.

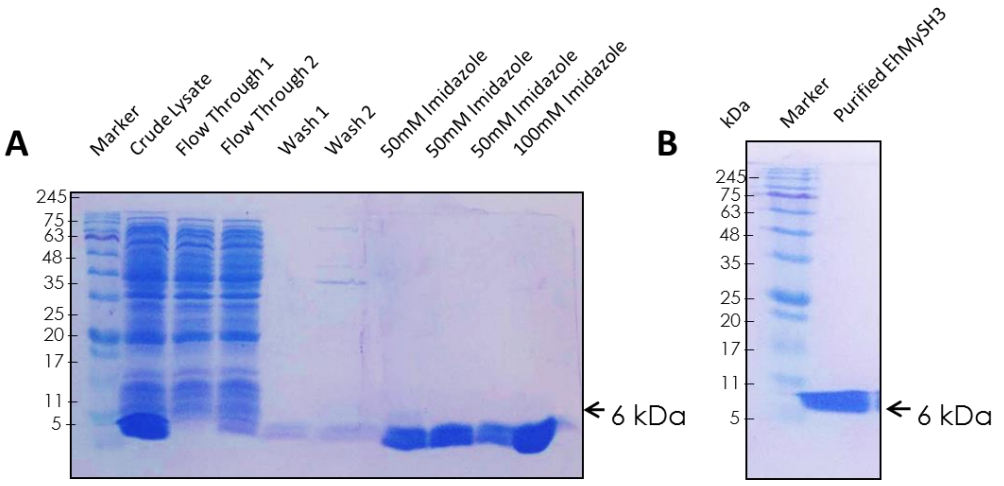


Figure 2.3-II: Overexpression and purification of EhMySh3.

(A) 15% SDS PAGE showing overexpressed *Eh*MySH3 recombinant protein eluted using Ni-NTA chromatography. (B) 15% SDS PAGE showing purified *Eh*MySH3 after GPC.

The gel permeation chromatography (GPC) profile of *Eh*MySH3 showed that it formed a trimer in solution (**Error! Reference source not found.**). *Eh*MySH3 has a molecular weight of 6 kDa, but gave a major peak corresponding to 18 kDa on a Superdex G75 16/60 column (GE Healthcare) that was pre-calibrated using a low-molecular-mass gel filtration calibration kit (Sigma). Protein eluted during GPC was collected in different fractions, and the peak fraction was run on 15% SDS-PAGE and stained with Coomassie blue (**Error! Reference source not found.**). Purified *Eh*MySH3 was found to be >95% pure.

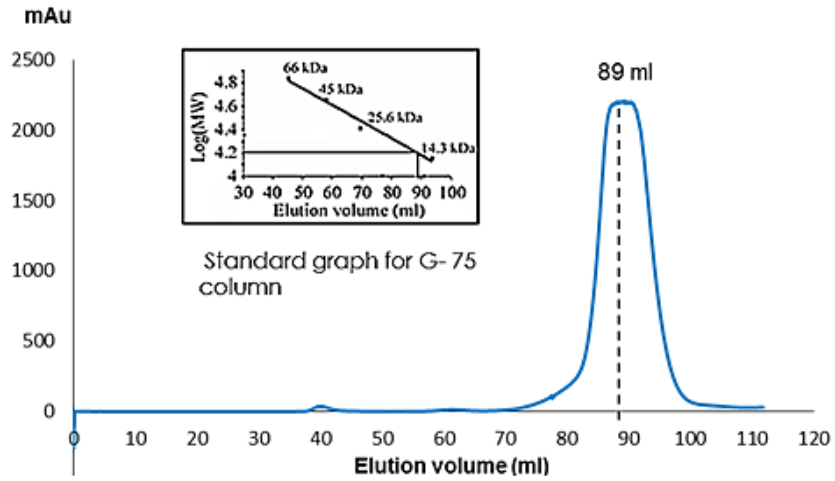


Figure 2.3-III: Elution profile of EhMySH3 showing major peak corresponding to 18 kDa during GPC using a G-75 Sephadex column.

*Eh*MySH3 was crystallized in two conditions. The SH3 domain that crystallised in the 0.2 M ammonium sulphate, 30% PEG 8000 condition was observed to be in complex with PEG, and is denoted here as PEG-*Eh*MySH3. The SH3 domain crystallized from the other condition did not bind to any (non-water) molecule, and is simply denoted as *Eh*MySH3.

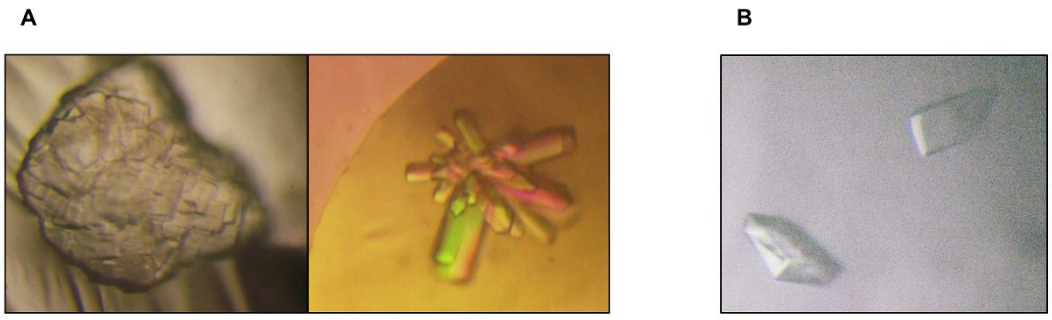


Figure 2.3-IV: Crystallization of EhMySh3 and PEG-EhMySH3.

(A) Left images shows initial *Eh*MySH3 crystals grown at 16°C from 2.0 M ammonium sulphate with 5% (v/v) 2-propanol (Crystal Screen II, Hampton Research). The right image shows improved crystals grown at 4ºC from 2.2 M ammonium sulphate with 5% v/v 2-propanol and 20% ethanol, which was added as an additive only in the drop; diffraction data was collected from these improved crystals. (B) PEG-*Eh*MySH3 crystals, obtained after one month at

4˚C from the crystal screen II (Hampton Research) condition consisting of 0.2 M ammonium sulphate and 30% PEG 8000.

2.3.3 Overall structure of *Eh*Myosin IB SH3 domain.

The crystal structures of both *Eh*MySH3 and PEG-*Eh*MySH3 were determined at 1.7 Å resolution (***Table 2.2-1***). Both the *Eh*MySH3 structure and PEG- *Eh*MySH3 structure showed the typical SH3 domain fold with five β-strands arranged in two sets of antiparallel β-sheets, a long loop commonly known as the RT loop, two small β-hairpins, viz. n-Src loop and distal loop, followed by a short 310 helix between strands β4 and β5. One of the two sets of β-sheets was observed to include β1, β2 and β5, with the other containing β2, β3 and β4. The two sets of antiparallel β-sheets run perpendicular to each other, together forming a β-barrel resembling a Greek-key motif (**Error! Reference source not found.** A).

β2 was observed to be shared by both sheets and to be the longest of the β-strands, spanning from residue D25 to K32, with its N-terminal residues highly conserved. Preceding β2 in the crystal structure is the RT loop, which is in general the longest hairpin of the SH3 fold, and was observed in the current crystal structure to extend from residue A7 to G24, and to be stabilized by several intra-loop hydrogen bonds between the peptide backbone atoms. Residues at the N- and C-termini of the RT loop are quite conserved (**Error! Reference source not found.**), with a highly conserved E18 residue in the middle of the loop. The orientation of the RT loop residues is important for ligand interaction. Another hairpin observed in the crystal structure was the short n-Src loop, spanning residues D33 to W36 and positioned near the ligand-binding site. Glutamic acid and aspartic acid residues were observed in the n-Src loop, near the hydrophobic pockets, and generally create a region of negative charge at this site, and may thus influence the ligand specificity of the *Eh*MySH3 domain.

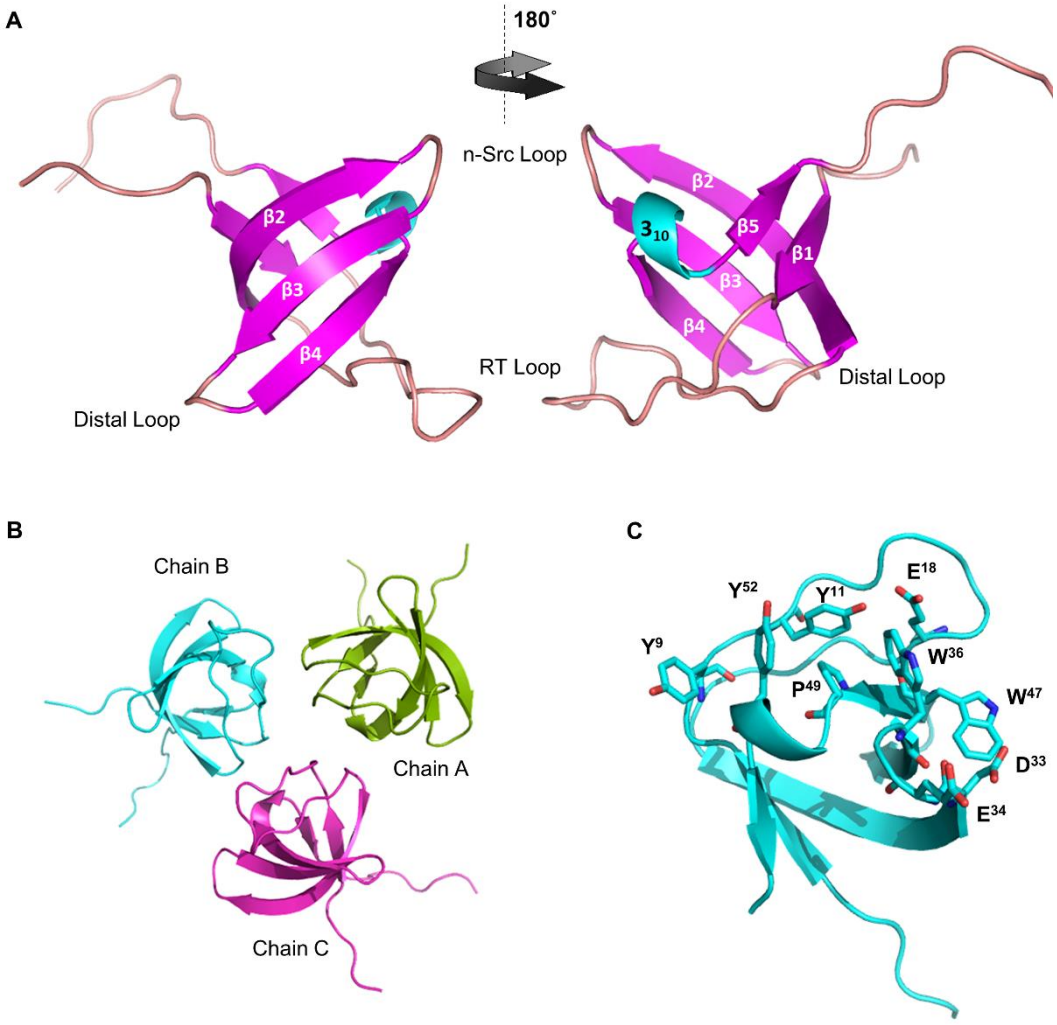


Figure 2.3-V: Crystal structure of EhMySH3.

(A) Structure of *Eh*MySH3 showing a β-barrel kind of a structure with a characteristic long RT loop shown from both the front and rear. (B) Three molecules of *Eh*MySH3 representing one trimer observed in the asymmetric unit. (C) *Eh*MySH3 structure showing residues crucial for ligand interactions.

2.3.4 Arrangement of molecules in the asymmetric unit.

We have got *Eh*MySH3 crystals in two different crystallization conditions as described in material and methods. Crystals obtained from condition where ammonium sulphate was used as a precipitant crystallized in the P21 space group with six molecules per asymmetric unit (denoted as A to F). Superimposing each of chains B to F, on chain A revealed the structures of all of the chains to be highly similar, except for minor deviations in the n-Src loop and RT-loop region. Moreover, on expanding the number of symmetry mates using PyMOL [17], we identified two unique trimers: one trimeric unit consisting of chains A, B and C, another unit including chains E, F and D interacting with symmetry mates of E and F (**Error! Reference source not found.** B). This crystallographic observation and the size of the *Eh*MySH3 protein eluted from the gel filtration column as described above both revealed the tendency of this protein domain to form trimers.

Crystals obtained from the second condition, where PEG 8000 was used as a precipitant, diffracted in C2 space group with eight molecules of *Eh*MySH3, along with short segments of PEG bound in the hydrophobic ligand-binding pockets of six chains A, B, C, D, G and H of these eight *Eh*MySH3 molecules. Two *Eh*MySH3 molecules and bound PEG were observed in a heterodimeric state. SH3 chains E and F showed no bound PEG. Chains B-C and Chains D-H formed dimers, with each chain bound to segments of PEG (**Error! Reference source not found.** A). We had also observed dimers of chain A with F and G with E with PEG bound only to one of these chains, with the other chain in an unbound state (**Error! Reference source not found.** C). Chains B and D were observed to share a longer segment of PEG fitting in the P4 specificity pockets of the two SH3 chains (**Error! Reference source not found.** B).

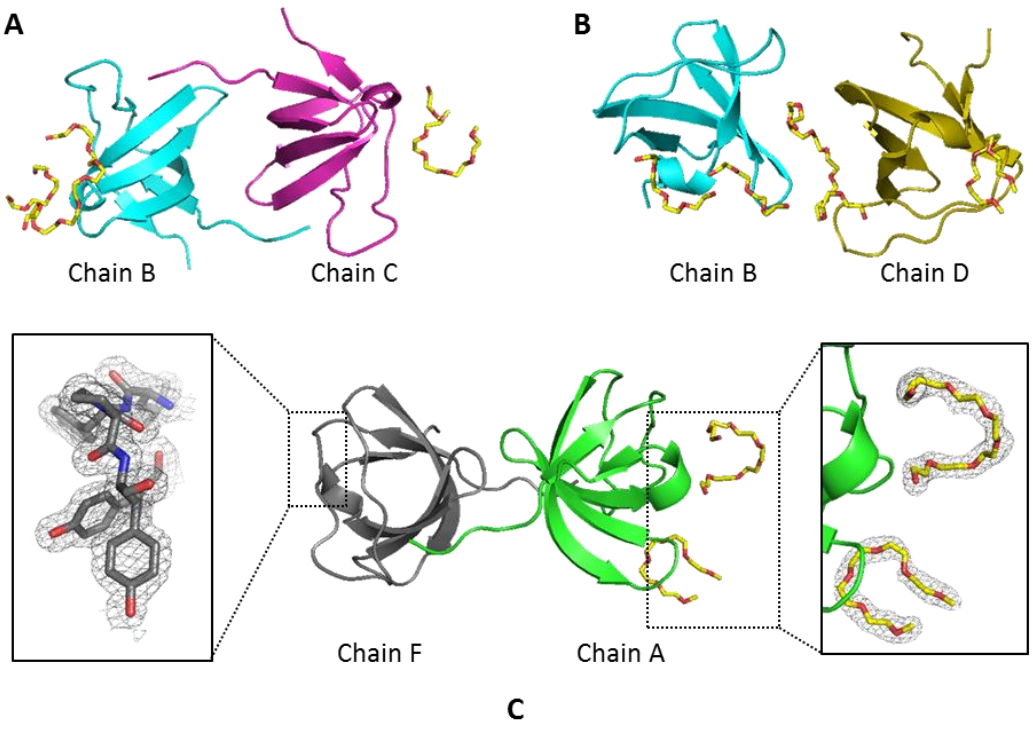


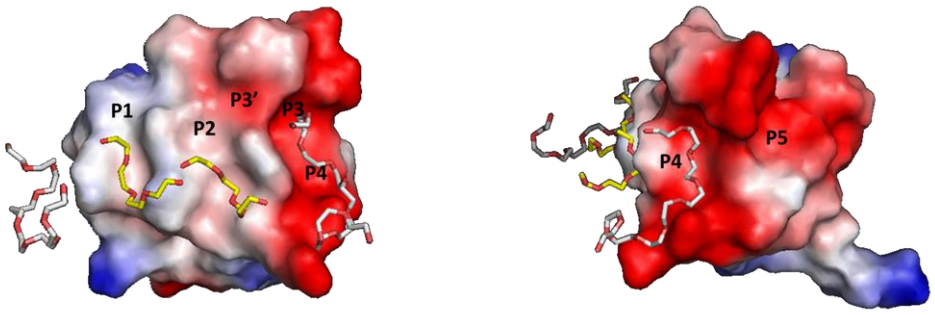
Figure 2.3-VI: Crystal structure of PEG-EhMySH3.

(A, B) Representatives of heterotrimers and heterotetramers made by different molecules of *Eh*MySH3 and bound PEG residues present in the asymmetric unit of PEG-*Eh*MySH3. Chains B and C formed dimers, with both chains in the PEG-bound state. Chains B and D were observed to share a longer piece of PEG, with it bound in the P4 pocket of both chains. (C) Chain A is bound to PEG, while chain F is in the unbound state. The left inset shows an expanded view of a stick diagram of a portion of the RT loop and 310 helix, and the right inset shows an expanded view of bound segments of PEG and their corresponding 1σ 2Fo – Fc electron density at a resolution of 1.7Aº.

2.3.5 Ligand-binding interface of *Eh*MySH3.

SH3 domains recognise ligands with a consensus xPxxP motif [14]. In the PEG-bound *Eh*MySH3 structure, PEG specifically occupied hydrophobic pockets meant for polyproline ligands as seen in previously reported SH3 structures [18,19]. On the basis of bound PEG molecules and previously reported structures, we identified four hydrophobic pockets, denoted as P1, P2, P3 and P4, in the *Eh*MySH3

domain structure (**Error! Reference source not found.**). One 'xP' pair of residues has been previously reported to fit in canonical pocket P1, which is flanked by residues Y9 and Y52 in *Eh*MySH3. The other 'xP' pair of residues has been shown to occupy P2, made up of residues P49 and W36 in *Eh*MySH3. These four mentioned residues flanking P1 and P2 are highly conserved in the SH3 domain family (**Error! Reference source not found.**). The P3 and P4 are additional ligand-binding pockets and have been shown to be responsible for the high affinity and high specificity of SH3 domains towards their respective polyproline ligands (Saksela and Permi, 2012). P3 is well known as the 'specificity pocket', and in our *Eh*MySH3 structures was observed to be enclosed by residues E18, W47 and W36, with the side chain of E18 generally imparting a negative charge to P3. Nearby P3 in a few SH3 structures lies another pocket P3'; this pocket in our *Eh*MySH3 structures was observed to include residue Y11 and E18, which would impart a negative charge. Either P3 or P3' is expected to be occupied by a consensus arginine or lysine residue present in class I and class II polyproline ligands. The *Eh*MySH3 crystal structures showed an atypical P4 pocket, much flatter and wider than that initially reported in mouse β-Pix SH3 domains, and consisting of residues W36, W47, D33 and E34 with a region of negative charge towards the n-Src loop due to side chain carboxylate of D33 (**Error! Reference source not found.** C). As mentioned above, PEG was observed to be bound in six chains out of total eight present in the asymmetric unit of the PEG-*Eh*MySH3 structure, with different segments of PEG occupying the different pockets. Chains A and B showed all four pockets occupied by PEG, hence we set out to describe the four pockets on the basis of these two chains (**Error! Reference source not found.**).



Chain B (Front view) **Chain B (Side View)**

Figure 2.3-VII: Surface charge diagram of chain B of PEG-EhMySH3 showing binding pockets.

Surface charge diagram of chain B of PEG-*Eh*MySH3 molecule with all of its pockets occupied by PEG (PEG segments are drawn as stick figures). We have identified two additional pockets, P3' and P5, on the basis of a similarity with the rat β-Pix SH3 domain. The figure was prepared with PyMOL (Delano, 2002) using the ABPS2.1 plugin. Negatively charged (red, PyMOL scale of -3), positively charged (blue, PyMOL scale of +3), and neutral (grey) surfaces are shown.

2.3.6 Difference between unbound and PEG-bound *Eh*MySH3.

We superimposed three chains of PEG-*Eh*MySH3: chain B, which bound PEG, and chains E and F, to which no PEG was seen to be bound. Both the alignment of chain E on chain B and that of chain F on chain B yielded an RMS deviation of only 0.15 Å for all Cα atoms. . The only major difference was seen in the n-Src loop region, with backbone carbonyls of residues D33 and E34 of chain B flipped by 180˚ relative to those of chains E and F, so as to move the residues away from pocket P4 to accommodate bound PEG. This movement coincided with lengthening of the already longest β-strand, β2, by one more residue in chain B (**Error! Reference source not found.** C). This structural deviation emphasizes the concept that the

binding of PEG to *Eh*MySH3 in the crystal mimicked what may be expected to occur when a physiological ligand binds to *Eh*MySH3.

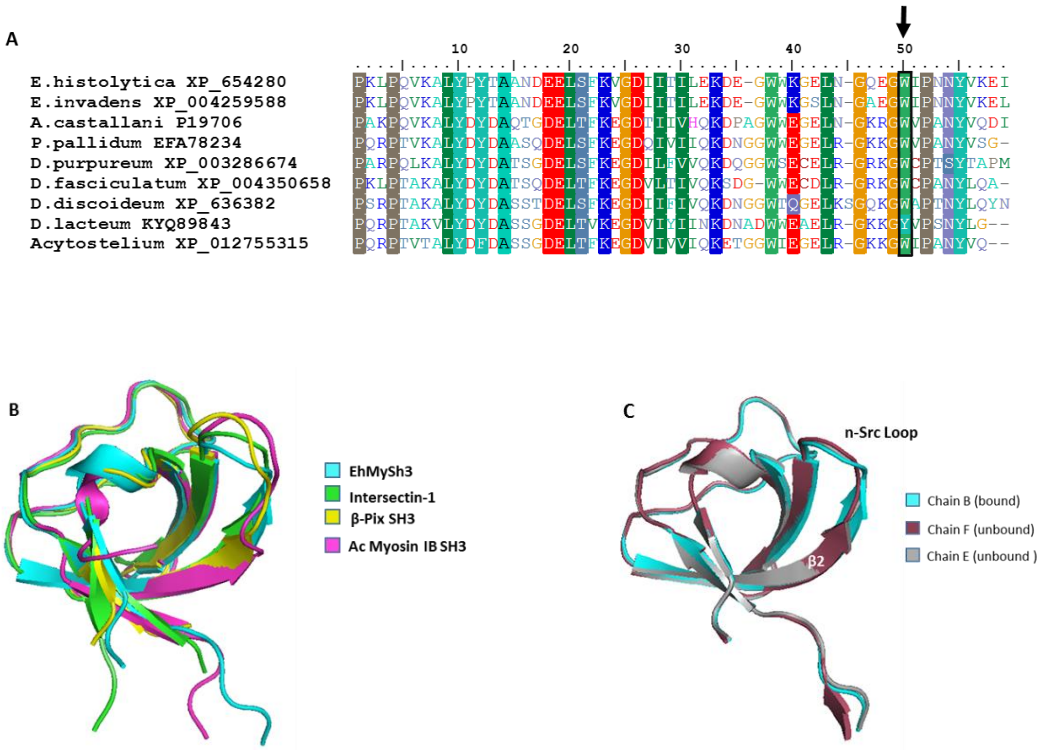


Figure 2.3-VIII: Comparison of EhMySH3 with other myosins.

(A) A multiple sequence alignment of SH3 domains of myosin I from various organisms of amoebozoa genera, showing a conserved tryptophan residue corresponding to W47 of E. histolytica, marked with an arrow. (B) Superimposition of PEG-*Eh*MySH3 with the peptide-bound SH3 domain structures of intersectin-I (4IIM), *Acanthamoeba* myosin IB SH3 domain (2DRK), and rat β-Pix SH3 domain (2p4r). (C) Superimposition of three chains of PEG-*Eh*MySH3: chain B, which is bound to PEG in all four pockets, and chains E and F, to which no PEG was seen bound. This comparison revealed a longer β2 strand in chain B, a slight change in the position of the n-Src loop, and a coinciding wider P4 pocket. In panels B and C, only the protein portions of the structures are shown, and they were drawn using PyMOL (Delano, 2002).

2.3.7 Comparison with SH3 domains of other myosins

A PDB search for the available SH3 domains belonging to different myosins allowed us to compare them to our *Eh*MySH3 structure. A multiple sequence alignment of all myosin SH3 structures and Src SH3 using ESpript (Robert and Gouet, 2014) showed them to be highly similar in the RT loop, β2, β3, β4, distal loop and 310 helix region. A marked difference was seen only in the n-Src loop region and β1. This difference in n-Src loop region lead to differences in the depth of the P4 pocket as well as surface charge distribution (**Error! Reference source not found.**). Also, the W47 residue, which appeared to make the P4 pocket wide and flat, was found to be conserved in amoebozoa and yeast myosin I SH3 (**Error! Reference source not found.** A). In higher eukaryotes, however, the corresponding position was replaced in most cases by leucine, e.g., in mouse myosin IE SH3 (PDB 2XMF) and mouse myosin 7b SH3 (PDB 5F3Y), or methionine. In the *Dictyostelium* myosin VII SH3 domain (PDB 2ION), a serine residue was found in place of this tryptophan at the P4 pocket. Thus, a tryptophan residue at this position appears to be particular to class I myosin SH3 domains of amoebozoa, and not present in all the myosin classes. The preference of tryptophan in the P4 pocket of amoebozoa myosin I SH3 domains was also confirmed when compared across all non-redundant protein database (**Error! Reference source not found.** A).

*Eh*MySH3 shares a high sequence identity, of around 58%, with the *Acanthamoeba* myosin IB SH3 domain, but the structure of the *Acanthamoeba* myosin IB SH3 domain (PDB 2DRM) did not yield a molecular replacement solution for the *Eh*MySH3 diffraction data. These structures superimposed with an RMS deviation of 1.2 Å, with the major difference between them being only in β2 and the n-Src loop region (**Error! Reference source not found.** B). We solved the structure of *Eh*MySH3 by using the second SH3 domain of intersectin-I (PDB 4IIM) as the search model; these two SH3 domains showed an RMS deviation of 0.6 Å and sequence identity of 50%.

2.3.8 Comparison with peptide-bound SH3 structures

We sought to predict what type of peptide ligand *Eh*MySH3 would preferentially bind, and for this goal, we first performed a DALI search [20] based on the *Eh*MySH3 structure. The structures of β-pix complexed with the atrophin-interacting protein 4 (AIP4) peptide (PDB code 2P4R), the PAK2-β-pix SH3 domain (PDB code 2DF6) and the second SH3 domain of ITSN1 with a synthetic peptide (PDB code 4IIO) were found to be the most similar to the structure of *Eh*MySH3, with RMSDs of 0.6 Å, 0.7 Å and 0.9 Å, respectively, and all having the P4 specificity pocket. Moreover, the ligand-binding surface of *Eh*MySH3 showed a shape and surface charge distribution quite similar to those of the β-pix SH3 domain, and less similar to those of the SH3 domain of human intersectin-2 (**Error! Reference source not found.**).

The structure of the AIP4 peptide bound to β-pix (PDB i.d. 2p4r) and other previous investigations revealed AIP4 to bind to the β-pix SH3 domain in a class I orientation, and with a high affinity of about 7 μM [21]. This crystal structure also revealed the β-pix SH3 domain to have additional pockets in the specificity zone, apart from canonical pockets P1, P2 and P3 [21]. It showed a specificity pocket P4 made hydrophobic by W43 and W54 from the β4 strand and containing a specific interaction formed by two hydrogen bonds connecting W54 and E45 of β-pix with the backbone carbonyl atom of P212 and the carboxylate group of R211 of AIP4 [21]. We observed in *Eh*MySH3 a similar region of negative charge at the P3 pocket, due to residue E18, but a wider and much flatter P4 pocket surrounded by highly acidic charges like in β-pix, though *Eh*MySH3 showed a shallow P4 pocket towards the n-Src loop. The P4 pockets of several SH3 domains such as Abl kinase SH3, the insulin receptor tyrosine kinase SH3 (IRTKS) domain, and β-pix SH3 have been observed to be occupied by proline residues of high-affinity ligands [19]. Another pocket marked as P5 in **Error! Reference source not found.**, occupied by P212 of the AIP4 peptide ligand in the β-pix SH3 domain structure, is also present

in the *Eh*MySH3 crystal structure. The AIP4 peptide is an example of a unique polyproline ligand that has an additional N-terminal PPII helix that interacts with additional specificity pockets present in β-pix.

The β-pix SH3 domain has also been reported to make high-affinity interactions with PAK2 (PDB 2DF6), a peptide with an atypical non-canonical PxxxPR motif (Hoelz et al., 2006). A salt bridge between an arginine of the PAK peptide and a conserved aspartate residue of the RT loop was indicated to be responsible for the relatively high-affinity binding of PAK2 to the β-pix SH3 domain. The β-pix pockets and residues involved in accommodating the PxxxPR portion of the ligand are the canonical residues of SH3 involved in ligand binding, with the same residues found in *Eh*MySH3 as well. The difference between *Eh*MySH3 and β-pix appears to lie in the involvement of the P4 pocket. The portion of the P4 pocket near the n-Src loop was measured to be shallower and narrower in the *Eh*MySH3 crystal structure than in the β-pix SH3 structure. Also, in the 2DF6 structure, consensus R186 was observed to occupy the P3 pocket and to interact with both residues D23 and E24 of the RT loop. At the position corresponding to D23 of 2DF6, *Eh*MySH3 instead has a glutamate, E17, and hence an arginine residue from a ligand would not be able to penetrate as deep into this pocket of *Eh*MySH3 as it was observed to do so in the 2DF6 structure. Although the arrangement of PEG in the P4 pocket supports the prediction that EHMySH3 prefers to bind 2DF6-type peptide ligands, the affinity would be expected to also depend on other non-consensus residues surrounding ligand-binding surface. Examples of other SH3 domains have been reported where atypical peptides interact with all four pockets on the SH3 surface, with such examples including C-terminal Src kinase/ tyrosine phosphatase PEP (Csk/PEP) and Gads T cell adaptor C-terminal SH3 domain with the SLP-76 peptide [19].

2.3.9 Peptide docking studies:

To achieve a more detailed understanding of the peptide preferences of *Eh*MySH3, we docked both AIP4 and PAK2 peptides onto the *Eh*MySH3 domain structure using HADDOCK webserver 2.2 [22]. The electrostatic and van der Waals interaction energies for docked AIP4-*Eh*MySH3 were calculated to be -242.5 kcal/mol and -56.9 kcal/mol, respectively, whereas for PAK2-*Eh*MySH3 structure, they were calculated to be -151.6 kcal/mol and -38.6 kcal/mol, respectively. Moreover, docked AIP4 occupies the consensus hydrophobic pockets P1 and P2 with arginine residue making a salt bridge with the residue in P3' pocket and the additional N-terminal PPII helix of AIP4 occupying P4 and P5. In contrast, the docked Pak2 peptide did not occupy the consensus pockets for polyproline. These experimental and computational structural studies together suggest that E. histolytica proteins with an AIP4 type of peptide sequence around the PXXP motif may bind the SH3 domain of *Entamoeba* myosin IB. Also, additional pockets in the specificity zone, i.e., P4 and P5, are expected to be key determinants for high-affinity interactions of *Eh*MySH3 with its ligand.

Table 2.3-1: Docking statistics of AIP4 and PAK2 peptides onto the EhMySH3 domain structure using HADDOCK webserver 2.2.

	AIP4 - *Eh*MySH3	**PAK2 - *Eh*MySH3**
HADDOCK score	-111.1 +/- 2.9	-86.7 +/- 1.6
Cluster size	10	400
RMSD from the overall lowest-energy structure	0.5 +/- 0.3	0.9 +/- 0.5
Van der Waals energy	-56.9 +/- 3.5	-38.6 +/- 1.9
Electrostatic energy	-242.5 +/- 38.4	-151.6 +/- 7.1
Desolvation energy	-7.3 +/- 4.8	-19.1 +/- 2.4
Restraints violation energy	17.1 +/- 15.55	13.7 +/- 14.99
Buried Surface Area	1312.8 +/- 35.6	1035.5 +/- 10.0

Z-Score	-1.1	0.0

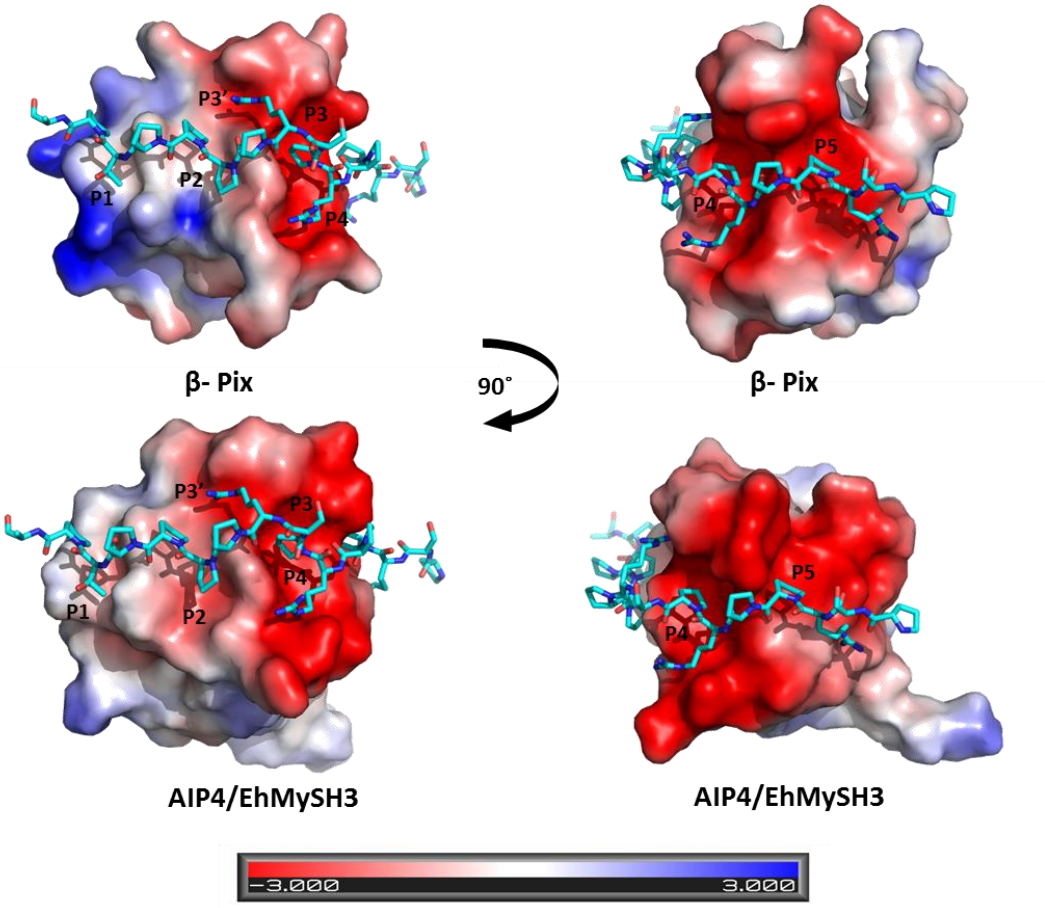


Figure 2.3-IX: Structural comparison of EhMySH3 with β-Pix SH3 domain.

Comparison of AIP4 peptide docked on *Eh*MySH3 using HADDOCK 2.2 webserver and actual crystal structure of rat AIP4 bound β-Pix Sh3 domain (2P4R) on the basis of similarity in surface charge and ligand binding pockets. The images on the left were rotated by 90º clockwise (along an imaginary axis on the plane of the page) to obtain the images on the right. The figure was prepared with PyMOL [17] using the ABPS2.1 plugin [23]. Negatively charged (red, PyMOL scale of -3), positively charged (blue, PyMOL scale of +3), and neutral (grey) surfaces are shown.

2.3.10 Can *Eh*MySH3 form ligand-mediated SH3 dimers?

Chain B of PEG-*Eh*MySH3 with PEG bound in pockets P4 and P2 forms a dimeric arrangement with the symmetry-related chain B' and its bound PEG. The structure of bound PEG segments together resembles a continuous polyproline peptide-like conformation. We may thus picture our crystal as containing a heterotrimer composed of two symmtery-related SH3 domains and PEG in between, analogous to the heterotrimeric arrangements (with peptide ligands instead of PEG) in the β-pix SH3/cbl-B (PDB 2AK5) and Cin85/cbl-B (PDB 2BZ8) structures (**Error! Reference source not found.**) [24]. Mediating SH3 dimerization is a characteristic of pseudo-symmetrical ligands that simultaneously bind two SH3 domain molecules, one in the class I orientation and other in the class II orientation [18]. Rat β-pix SH3 was shown to form a heterotrimer with the cbl-B peptide in solution as well [24]. The *Eh*MySH3-PEG heterotrimer is different from the superSH3 domain formed in p47phox since it does not involve the 'GWW' motif characteristics of superSH3 [25]. This kind of an arrangement of *Eh*MySH3 provides it with the possibility of engaging in peptide-mediated dimer formation, and hence forming multimeric protein complexes. Myosin X has been shown to form antiparallel dimers through a dimerization motif present in its tail region, in contrast to two-headed myosins, which form parallel dimers resulting from the coiled-coil regions of their tails. This arrangement of myosin X has been shown to enable it to walk efficiently on both single actin fibres as well as bundled actin [26]. If two myosin I molecules bind to a target that has a pseudo-symmetrical polyproline motif, they can form dimers and carry the target on an actin filament similar to myosin X.

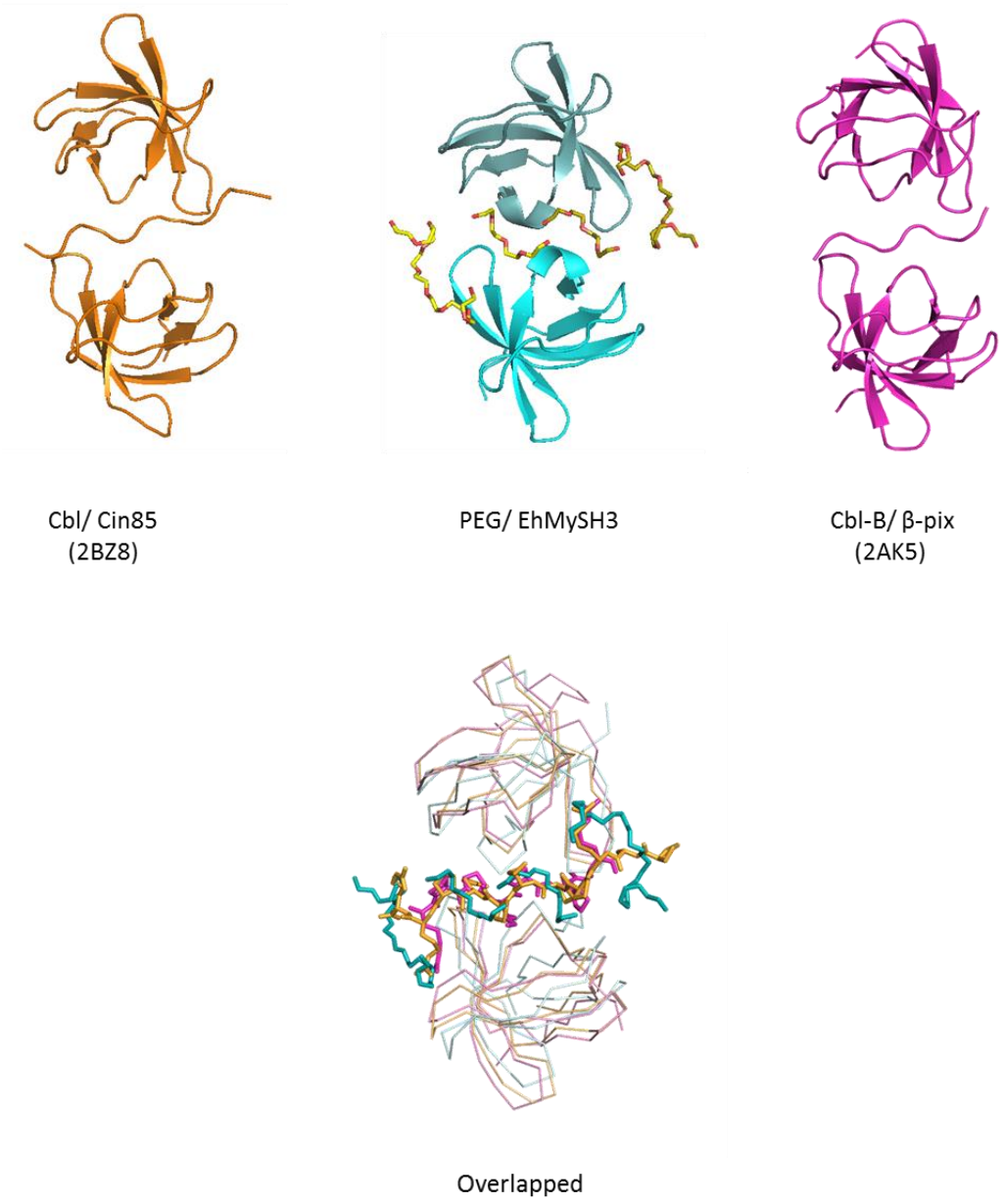


Figure 2.3-X: Ligand-mediated dimerization.

Cin85 and β-pix each form a heterotrimer when bound to the cbl-B peptide. A similar kind of structure was observed for chain B of *Eh*MySH3 with its symmetry mates and four segments of PEG bound in between. Two of these four segments of PEG are bound primarily to one *Eh*MySH3 chain, and the other two to the symmetry mate, leading to a pseudo-symmetrical arrangement

as found for the cbl-B peptide. Using PyMOL, all three structures superimposed well, supporting the possibility of a ligand-mediated dimerization of the *Eh*MySH3 domain occurring physiologically.

2.4 Conclusion:

In this chapter, we determined the crystal structure of the C-terminal SH3 domain of unconventional myosin IB from the human pathogen *Entamoeba histolytica* (*Eh*MySH3) with and without bound PEG. In the PEG-bound *Eh*MySH3 structure, PEG appeared to have mimicked physiological ligand by occupying the conserved ligand-binding pockets on the *Eh*MySH3 surface. These bound PEG molecules highlighted the importance of the *Eh*MySH3 domain P4 pocket, located between residue W47 and the well-conserved residue W36. The conservation of tryptophan residue corresponding to W47 in amoebozoa myosin I also suggested the P4 pocket to be crucial for the specificity and affinity of *Eh*MySH3 for its ligands. Comparing various myosins revealed *Eh*MySH3 to be similar to the SH3 domain of myosin IB from *Acanthamoeba*, except in the β2 sheet and n-Src region. However, according to the charge distribution, on the ligand-binding surface as well as adjacent C-terminal region in '+' peptide-binding orientation, *Eh*MySH3 also shares structural similarity with the rat β-pix SH3 domain bound to the AIP4 peptide. These analyses led us to recognise two other regions of negative surface charge on *Eh*MySh3 surface, i.e., P3' and P5, which could be essential for its interaction with AIP4-like ligands. The arrangement of chain B of PEG-*Eh*MySH3 with its symmetry mate revealed that a ligand may induce the dimerization of *Eh*MySH3 and hence lead to the formation of a heterotrimeric protein-ligand-protein complex. Myosin IB in *E. histolytica* has been found to be actively involved with the actin cytoskeleton in phagocytic cup formation but the role played by its C-terminal SH3 domain is still under study. This SH3 may be responsible for the formation of a multimeric-protein complex with various proteins along with actin which are involved in phagocytic process. *Eh*Myosin IB is different from other amoeboid myosins and lies in between metazoan and amoeboid myosins [9]. Hence, the results of the current study about the SH3 domain of myosin I has a

larger implication and are expected to be highly useful for understanding and predicting possible binding partners of unconventional myosin IB of *E. histolytica*.

2.5 References:

[1] Pollard, T.D., Doberstein, S.K. and Zot, H.G. (1991). Myosin-I. Annu Rev Physiol 53, 653-81.

[2] Mooseker, M.S. and Cheney, R.E. (1995). Unconventional myosins. Annu Rev Cell Dev Biol 11, 633-75.

[3] Baines, I.C., Brzeska, H. and Korn, E.D. (1992). Differential localization of *Acanthamoeba* myosin I isoforms. J Cell Biol 119, 1193-203.

[4] Ostap, E.M. and Pollard, T.D. (1996). Overlapping functions of myosin-I isoforms? J Cell Biol 133, 221-4.

[5] Kumar, S. et al. (2014). Crystal Structure of Calcium Binding Protein-5 from *Entamoeba* histolytica and Its Involvement in Initiation of Phagocytosis of Human Erythrocytes. Plos Pathogens 10

[6] Voigt, H., Olivo, J.C., Sansonetti, P. and Guillen, N. (1999). Myosin IB from *Entamoeba* histolytica is involved in phagocytosis of human erythrocytes. Journal of Cell Science 112, 1191-1201.

[7] Kumar, N., Somlata, Mazumder, M., Dutta, P., Maiti, S. and Gourinath, S. (2014). *Eh*Coactosin Stabilizes Actin Filaments in the Protist Parasite *Entamoeba* histolytica. Plos Pathogens 10

[8] Babuta, M., Mansuri, M.S., Bhattacharya, S. and Bhattacharya, A. (2015). The *Entamoeba* histolytica, Arp2/3 Complex Is Recruited to Phagocytic Cups through an Atypical Kinase *Eh*AK1. Plos Pathogens 11

[9] Vargas, M., Voigt, H., Sansonetti, P. and Guillen, N. (1997). Molecular characterization of myosin IB from the lower eukaryote *Entamoeba* histolytica, a human parasite. Mol Biochem Parasitol 86, 61-73.

[10] Anderson, B.L., Boldogh, I., Evangelista, M., Boone, C., Greene, L.A. and Pon, L.A. (1998). The Src Homology Domain 3 (SH3) of a Yeast Type I Myosin, Myo5p, Binds

to Verprolin and Is Required for Targeting to Sites of Actin Polarization. The Journal of Cell Biology 141, 1357-1370.

[11] Novak, K.D. and Titus, M.A. (1998). The myosin I SH3 domain and TEDS rule phosphorylation site are required for in vivo function. Mol Biol Cell 9, 75-88.

[12] Jung, G., Remmert, K., Wu, X., Volosky, J.M. and Hammer, J.A. (2001). The *Dictyostelium* Carmil Protein Links Capping Protein and the Arp2/3 Complex to Type I Myosins through Their Sh3 Domains. The Journal of Cell Biology 153, 1479-1498.

[13] Ren, R., Mayer, B., Cicchetti, P. and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. Science 259, 1157-1161.

[14] Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W. and Schrelber, S.L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 76, 933-945.

[15] Feng, S., Chen, J., Yu, H., Simon, J. and Schreiber, S. (1994). Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. Science 266, 1241-1247.

[16] Robert, X. and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Research 42, W320-W324.

[17] Delano, W.L. (2002). The PyMOL Molecular Graphics System.

[18] Kaneko, T., Li, L. and Li, S. (2008). The SH3 domain--a family of versatile peptide- and protein-recognition module. Front Biosci 13, 4938-4952.

[19] Saksela, K. and Permi, P. (2012). SH3 domain ligand binding: What's the consensus and where's the specificity? FEBS Lett 586, 2609-14.

[20] Holm, L. and Rosenström, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Research 38, W545-W549.

[21] Janz, J.M., Sakmar, T.P. and Min, K.C. (2007). A novel interaction between atrophin-interacting protein 4 and beta-p21-activated kinase-interactive exchange factor is mediated by an SH3 domain. J Biol Chem 282, 28893-903.

[22] van Zundert, G.C.P. et al. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. Journal of Molecular Biology 428, 720-725.

[23] Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. Proceedings of the National Academy of Sciences 98, 10037-10041.

[24] Jozic, D. et al. (2005). Cbl promotes clustering of endocytic adaptor proteins. Nat Struct Mol Biol 12, 972-9.

[25] Groemping, Y., Lapouge, K., Smerdon, S.J. and Rittinger, K. (2003). Molecular basis of phosphorylation-induced activation of the NADPH oxidase. Cell 113, 343-55.

[26] Lu, Q., Ye, F., Wei, Z., Wen, Z. and Zhang, M. (2012). Antiparallel coiled-coil-mediated dimerization of myosin X. Proc Natl Acad Sci U S A 109, 17388-93.

3. CHAPTER II

Functional characterisation of SH3 domain of EhMyosin IB

3.1 Introduction

Phagocytosis is a kind of endocytosis which involves ingestion of large particles (>0.5μM) by the cells and is initiated by attachment of particles to the cell surface receptors (Fc receptors, FcRs or receptors for complement, CRs) leading to activation of different signaling pathways that facilitates internalization of the particles [1]. The phagocytic process has been studied in much detail in professional phagocytes like macrophages, dendritic cells, and neutrophils. These carry out phagocytosis for a defense mechanism, hence they internalize and degrade the particle and eventually do antigen presentation. FcR and CR pathways though morphologically different but both lead to the activation of Ras family GTPases. Rho GTPases play a central role in the membrane and cytoskeletal rearrangements during the phagocytic process [2,3].

Lower eukaryotes like *Dictyostelium* and *Entamoeba* uses the phagocytic process for feeding. During Fc receptor-mediated endocytosis, several GEFs (GTPase exchange factors) came into play which activate cdc42, Rac1 and RhoA GTPases that promote actin remodeling leading to protrusion of the cup as well as phagosome closure. Phagocytosis is quite divergent in *E. histolytica* cells when compared to professional phagocytes. *Entamoeba* lacks WASP/SCAR adaptor proteins which activate arp2/3 complex to actin nucleation sites after Cdc42 mediated recruitment, during FcR pathway, and has numerous actin binding proteins to modify F-actin [4]. *E. histolytica* has only one unconventional myosin, Myosin IB and lacks other myosins like myosin X, *Dd*Myosin VII important for membrane extension during phagocytosis; and myosin Ic, Ixb, and V which may be involved in force generation during pseudopod extension [5].

Membrane phospholipids play a major role in membrane trafficking, cytoskeletal rearrangements and cell surface receptor signaling [6]. PI3P (phosphatidylinositol-3-phosphate) is the product of class II and III PI 3 kinases on membrane PIs. Formation of PI3P on phagosome surface results in the recruitment of several PI3P-binding proteins including EEA1 and Hrs [7]. Moreover, inhibition of PI3P

blocks phagosome maturation. PI3P is involved in endocytosis and phagocytosis in *E.histolytica* cells as well [4,8]. FYVE domains are zinc finger domains which specifically identifies PI(3)P on the membrane. FYVE domain containing Rho GEF, *Eh*FP4 was found to be involved in phagocytosis. *Entamoeba* genome has 12 such FYVE family Rho GEFs, out of which *Eh*FP5, 6, 7 and 10 are highly expressed [4].

E. histolytica lacks effectors of PI3P including EEA1 and Hrs [4,9]. *Eh*FP4 has several Rac/ Rho GTPases as its downstream effectors during the phagocytic process [4]. In FcR mediated phagocytosis, several GTPases act upon effectors and adaptor protein like WASP and other WASP related proteins which leads to actin polymerization at the site of cup formation. WASP-related proteins Vrp1, Bee1p and Las17p interacts with SH3 domain of Myo3 and Myo5 of *S.cerevisae* and helps in recruitment of these myosin Is, and Arp2/3 at cortical actin nucleation sites. In *Dictyostelium* and *Acanthamoeba*, SH3 domain of myosin I interact with adapter proteins like CARMIL and Acan125 respectively, which recruits and activate Arp2/3 [10]. Formation of multiprotein complexes via these interactions is essential for rearrangement of the actin cytoskeleton during endocytosis. However, *Entamoeba* lacks WASP/SCAR protein though it has several actin binding proteins which recruit Arp2/3 for actin polymerization. The molecules involved in the phagocytic process in *Entamoeba* are quite diverse as compared to the classical phagocytic cells suggesting it to be significantly divergent organism [4].

During phagocytosis in *E. histolytica*, Myosin IB has been found to be present from initiation till the pinching of the phagocytic vesicle [11]. Sh3 domains of some class I myosins have been seen to recruit adaptor proteins which regulate cortical actin dynamics during the endocytic process. Till now, there are no studies regarding the function and significance of C-terminal SH3 domain of *Eh*Myosin IB.

In the present chapter, we have identified an FYVE domain containing Rho-GEF, *Eh*FP10 to be one of the interacting protein of *Eh*Myosin IB Sh3 domain. We have validated the finding through a cocrystal structure of *Eh*MYSH3 bound to a proline rich peptide from *Eh*FP10. The interaction has been seen in-vivo as well in GFP-tagged *Eh*FP10 overexpressed cells activated for erythrophagocytosis. *Eh*FP10 and

*Eh*Myosin IB colocalizes in the phagocytic cup. *Eh*Myosin IB Sh3 domain interacts with APC domain which is also known to bind actin filament but the interaction has no effect on its actin binding activity. This is also the first report showing the association of microtubules with the phagocytic cup in *E. histolytica* in GFP-*Eh*FP10 overexpressing cell lines. The absence of PI3P effector proteins like EEA, Hrs in *Entamoeba* and absence of homologous proteins of *Eh*FP10, indicates the possibility for this interaction to be unique to *E. histolytica*. However, this is in accordance with the reports implicating the association of myosin I Sh3 domains with proteins involved in actin remodeling and cytoskeletal rearrangements in other organisms.

3.2 Material and methods:

E.histolytica's genomic DNA was obtained from Prof. A. Bhattacharya lab SLS, JNU. All In-Vivo studies were performed and culture facility were provided by our collaborator Prof. Alok Bhattacharya S.L.S, J.N.U. Mr. Sabir Ali helped me in growth and maintenance of parasite *E. histolytica* strain HM-IMSS and all in-vivo studies.

3.2.1 Cloning of various constructs:

The primers for all the constructs were designed as described in chapter1 (2.2.1). The primer sequence of all the gene with their respective vector backbone and restriction site details are given in the table below:

S.No.	Construct Name	Amino acid position (start – end) [accession no.]	Primer Sequence	Plasmid Vector	Restriction enzyme site
1.	*Eh*My1TD	1045 – 1049 [EAL48894]	**FwMy1TD:** 5'-CGCCATATGTCTTCTGTTTCTTCTTATGCTGCTC-3' **RvMy1TD :** 5'-CCGCTCGAGAATTTCTTTGACATAGTTGTTAGGAATC- 3'	pET21c	NdeI - Xho
2.	GSTMySH3	994 – 1049 [EAL48894]	**FwGSTSH3:** 5'-CCGGGGATCCCAAGTTAAAGCACTCTATCCATATACTG-3' **RvGSTSH3:** 5'-CCCCTCGAGTTAAATTTCTTTGACATAGTTGTTAGGAA-3'	pGEX-6P2	BamHI - XhoI
3.	*Eh*GEFD	113 – 438 [EAL46050]	**FwGefDHPH:** 5'-CATGCCATGGCTAAACCTGAAACTAAAGCATTATC-3'	pET28(b)	NcoI - XhoI

			RvGEFDHPH: 5'- CCGCTCGAGTAAAGTTTTTAA TTTTGACTTTTCC-3'		
4.	GEFAPC	631- 876 [EAL46050]	**FwAPCGEF:** 5'- CATGCCATGGCTGTCCCTCAA CAACAACCAGAA-3' **RvGEFfull:** 5'- CCGCTCGAGTTTTGTAACCCC TCTTTTTGG 3'	pET28(b)	NcoI - XhoI
5.	NGFPGEF	1 – 876 [EAL46050]	**FwGFPNterGEF:** 5'- CCGCTCGAGATGACTGAAAA GCTAAGTGATTGG-3' **RvGFPNterGEF:** 5'- CCGGGGATCCTTATTTTGTAA CCCCTCTTTTTGG-3'	p*Eh*- NEOGFP	XhoI - BamHI

All the genes were amplified by PCR using respective forward and reverse primers from the genomic DNA of *Entamoeba histolytica* strain HM1: IMSS. Reaction mix for PCR amplification was prepared as mentioned in the appendix (A.4) and run in a thermocycler for 30 cycles with cycling parameters as mentioned in the appendix (A.4). The PCR products and the vector were double digested with the pre-mentioned restriction enzymes and gel purified. These digested and purified products were then ligated using T4DNA ligase and kept at 16°C for 16hrs. The ligated mixture was then transformed into *E.Coli* DH5α, plated on ampicillin containing LB agar plates and kept at 37°C for overnight. The colonies were screened for positive clones by colony PCR. The plasmid was isolated from positive colonies. The clone was further confirmed by double digestion of isolated plasmids and gene sequencing. (Detailed steps are given in appendix A.4).

3.2.2 Protein overexpression and purification:

*Eh*MySH3 was purified as mentioned in Chapter 1 (2.2.2). *Eh*My1TD, GSTMySH3, and *Eh*GEFD recombinant plasmid were transformed in BL21 (DE3) and GEFAPC in BL21 (C41) *E.coli* strain. The secondary culture was grown at 37˚C using 1% primary culture grown overnight from the single BL21 colony and was induced with 0.2mM IPTG for 4 hours. Pellets were resuspended in their respective lysis buffer and sonicated following 4-5 freeze-thaw cycles. After sonication, cell lysate was centrifuged at 13,000 rpm for 30 minutes. Soup obtained was loaded on to Ni-NTA column (Sigma) with Ni-sepharose (GE healthcare).

*Eh*My1TD, GSTMySH3, and *Eh*GEFD proteins were eluted with their respective elution buffers following a wash cycle (30ml) with wash buffers. GEFAPC protein co-expressed with a chaperone. To remove the chaperone, after the protein was bound to Ni-sepharose, it was incubated for ½hr to 1 hr in incubation buffer (2mM Mg-ATP, 50mM Tris pH-7.5, 300mM NaCl, 10mM $MgCl_2$, 5% glycerol, 10mM imidazole and 5mM β-ME) and then was washed with small amount of heat-denatured *E.coli* protein in the wash buffer. After 3-4 cycles of incubation followed by washes, GEFAPC protein was eluted in the elution buffer.

S.No.	Protein	Lysis Buffer	Wash Buffer	Elution Buffer	GPC Buffer
1.	GSTSH3	50mM Tris pH-7.4 150mM NaCl 5% glycerol 3mM β-ME	50mM Tris pH7.4 300mM NaCl 3mM β-ME	50mM Tris pH-7.4 150mM NaCl 20mM reduced glutathione 5% glycerol 3mM β-ME	50mM Tris pH-7.4 150mM NaCl 5% glycerol 3mM β-ME
2.	*Eh*My1TD	50mM Tris pH-7.4 150mM NaCl	**Wash buffer1:** 50mM Tris pH-7.4 **Wash buffer2:** 50mM Tris pH7.4 300mM NaCl	50mM Tris pH-7.4 150mM NaCl 100mM Imidazole	50mM Tris pH-7.4 150mM NaCl

			Wash buffer3: 50mM Tris pH7.4 150mM NaCl 20mM imidazole		
3.	GEFAPC	50mM Tris pH-7.4 150mM NaCl 10mM $MgCl_2$ 5% glycerol 10mM imidazole 5mM β-ME 0.1% Triton-X 100	50mM Tris pH-7.4 300mM NaCl 10mM $MgCl_2$ 5% glycerol 10mM imidazole 5mM β-ME	50mM Tris pH-7.4 150mM NaCl 100mM Imidazole 5% glycerol 5mM β-ME	50mM Tris pH-7.4 150mM NaCl 5% glycerol 5mM β-ME
4.	*Eh*GEFD	50mM Tris pH-7.4 200mM NaCl 5% glycerol 3mM β-ME	50mM Tris pH-7.4 200mM NaCl 5% glycerol 5mM imidazole 3mM β-ME	50mM Tris pH-7.4 200mM NaCl 100mM Imidazole 5% glycerol 3mM β-ME	30mM Tris pH-7.4 200mM NaCl 5% glycerol 3mM β-ME

Purified proteins were concentrated using centricon (Amicon, Millipore) and further purified by Gel Permeation Chromatography using Superdex G75 16/60 column and G200 10/300 columns (GE Healthcare) pre-equilibrated in the buffer. The peak fractions were pooled together and checked on SDS-PAGE. GEFAPC showed a regular degradation pattern soon after chaperone removal, hence concentrated protein was frozen immediately after purification and stored at -80˚C.

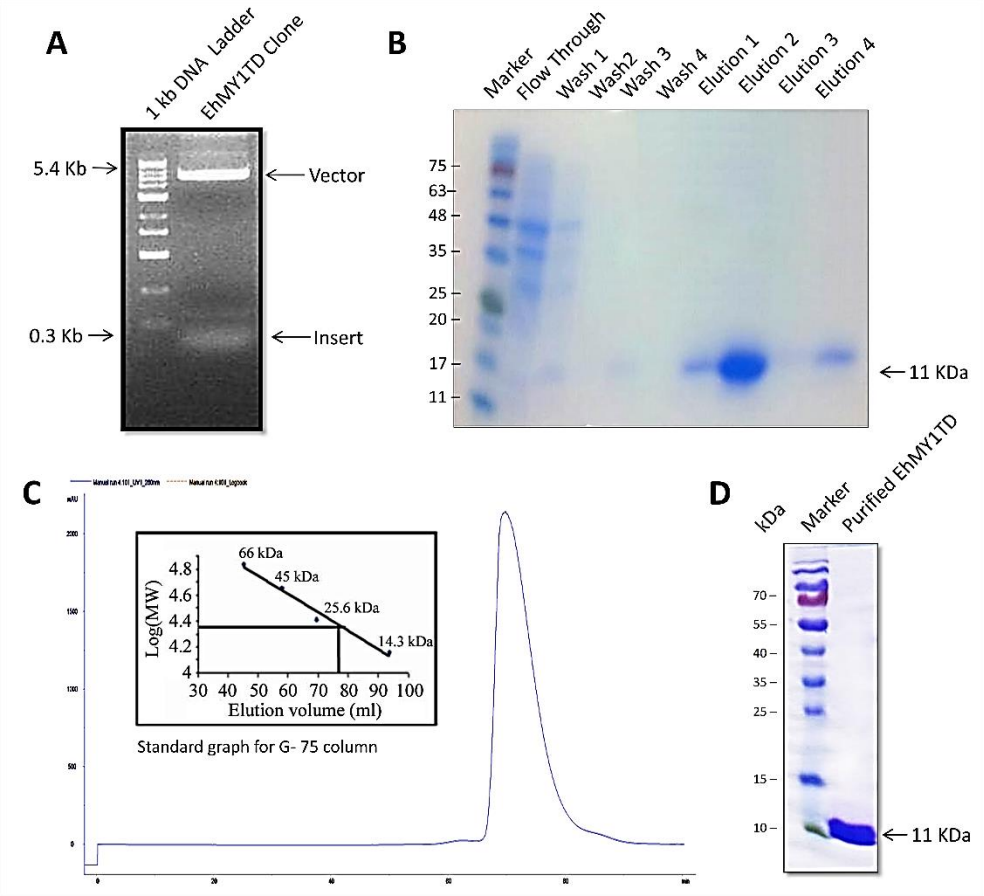


Figure 3.2-I: Cloning, Overexpression, and Purification of EhMy1TD.

(A) Confirmation of *Eh*My1TD clone by double digestion with NdeI and XhoI. The presence of 0.3kb confirms *Eh*My1TD gene and 5.4kb band corresponds to pET21c vector. (B) Purification of overexpressed *Eh*My1TD by Ni2+-NTA chromatography. (C) Purification of *Eh*My1TD by gel permeation chromatography (GPC) with a peak at 78ml. (D) SDS-PAGE showing about 90% purified *Eh*My1TD protein with a band at 11kDa.

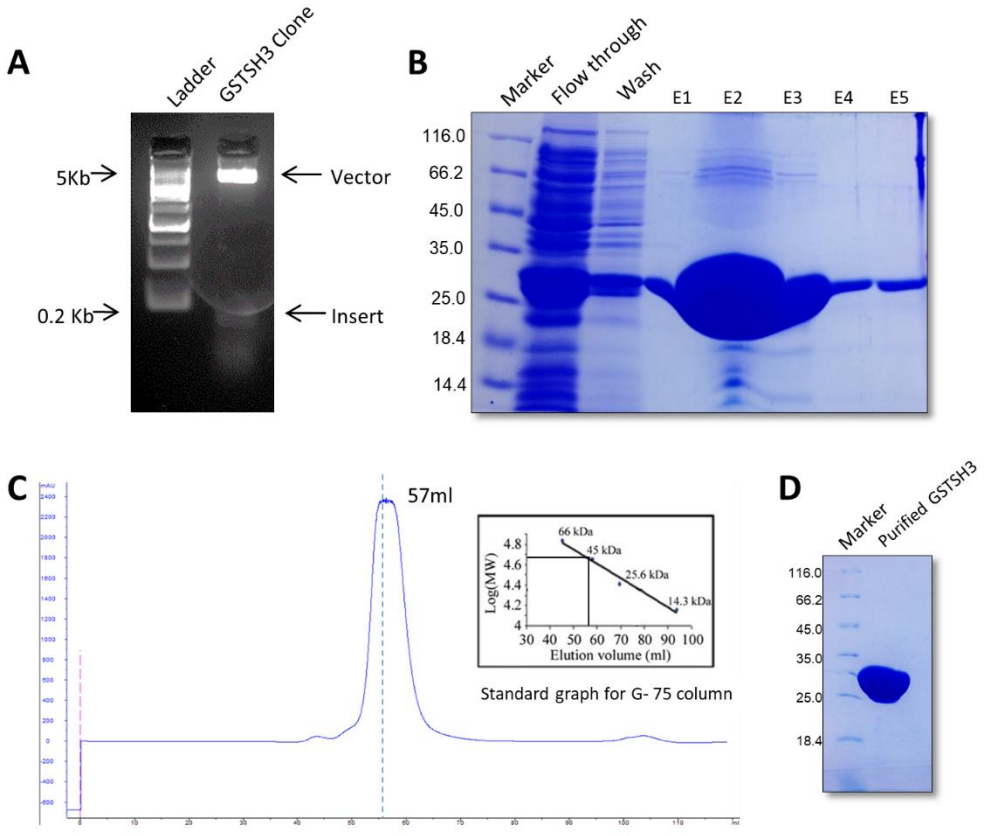


Figure 3.2-II: Cloning, Overexpression, and Purification of GSTSH3.

(A) Confirmation of GSTSH3 clone by double digestion with BamHI and XhoI. The presence of 0.2kb confirms *Eh*MySH3 gene and 5.4kb band corresponds to pGEX-6P2 vector. (B) Purification of overexpressed GSTSH3 by column chromatography using glutathione resin (GE Healthcare). (C) Purification of GSTSH3 by gel permeation chromatography (GPC) with a peak at 57ml. (D) SDS-PAGE showing about 95% purified GSTSH3 protein with a band at 28kDa.

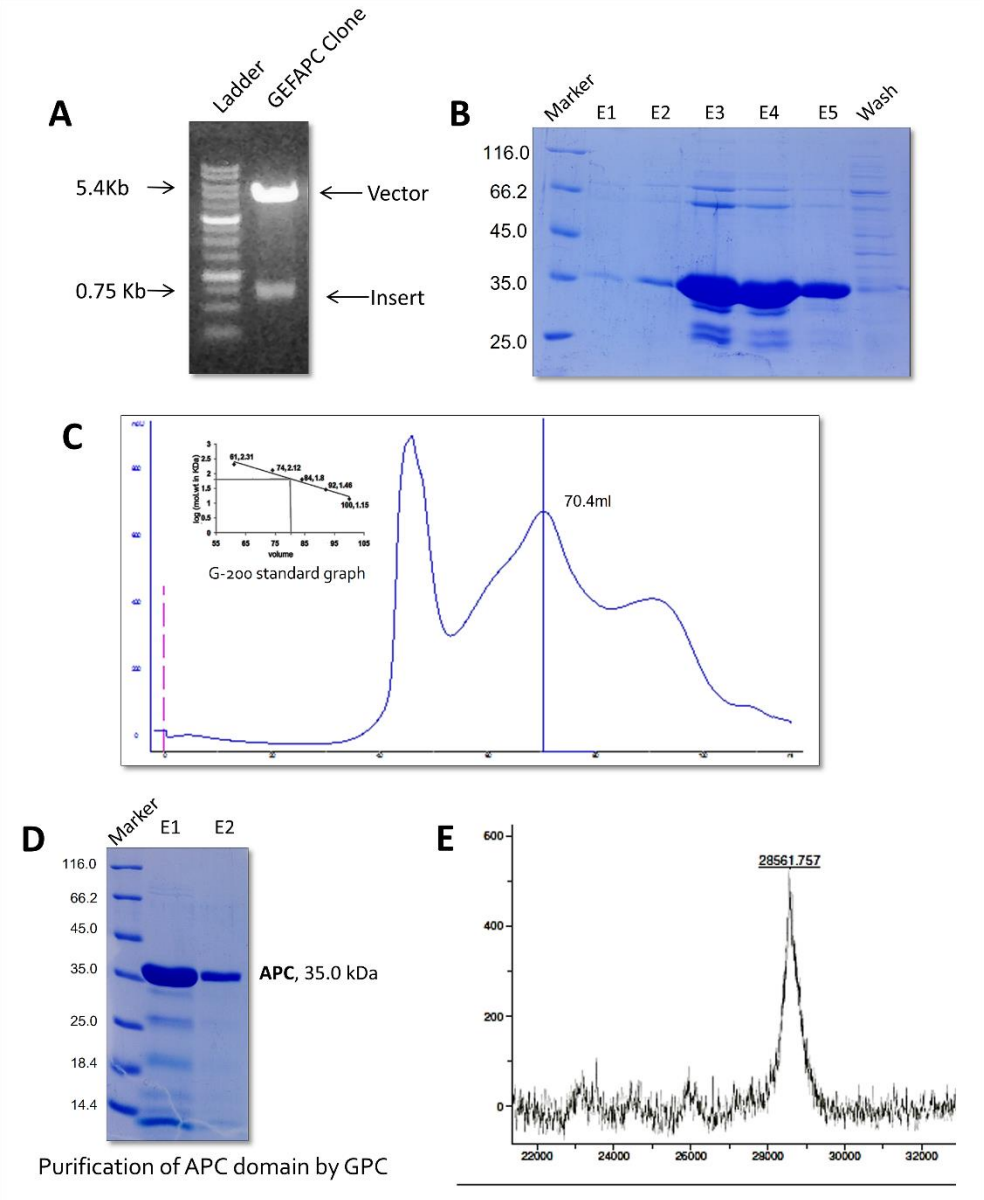


Figure 3.2-III: Cloning, Overexpression, and Purification of GEFAPC:

(A) Confirmation of GEFAPC clone by double digestion with NcoI and XhoI. The presence of 0.75kb confirms GEFAPC gene and 5.4kb band corresponds to pET28b vector. (B) Purification of overexpressed GEFAPC by Ni2+-NTA chromatography. (C) Purification of GEFAPC by gel permeation chromatography (GPC) with a major peak at 70.4ml. (D) SDS-PAGE showing about 90% purified GEFAPC protein with a band at 35kDa with lower degradation bands. (E) The molecular weight of GEFAPC is 28kDa, which has been confirmed by intact mass analysis of purified GEFAPC, but it runs at 35 kDa in 12% SDS-PAGE.

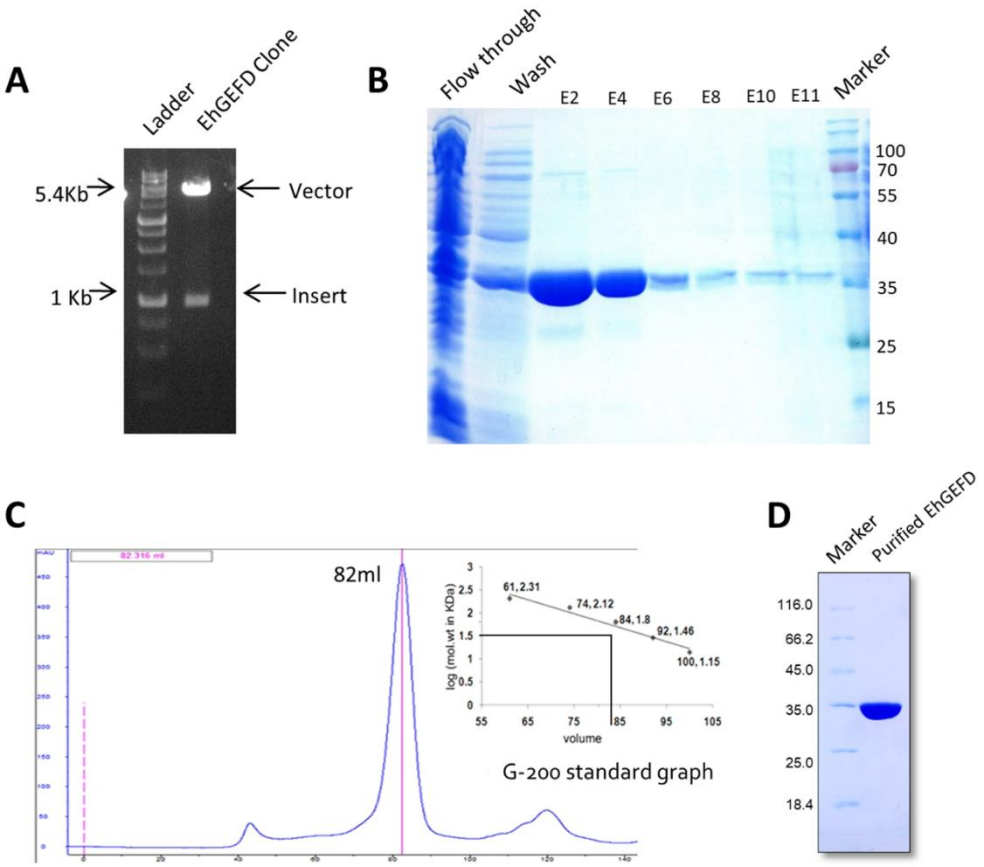


Figure 3.2-IV: Cloning, Overexpression, and Purification of EhGEFD.

(A) Confirmation of *Eh*GEFD clone by double digestion with NcoI and XhoI. The presence of 1kb confirms *Eh*GEFD gene and 5.4kb band corresponds to pET28b vector. (B) Purification of overexpressed *Eh*GEFD by Ni2+-NTA chromatography. (C) Purification of *Eh*GEFD by gel permeation chromatography (GPC) with a major peak at 82ml. (D) SDS-PAGE showing about 90% purified *Eh*GEFD protein with a band at 35kDa.

3.2.3 Crystallization and data collection:

For co-crystallization with all 5 synthesized peptides, 30mg/ml of *Eh*MySH3 protein was mixed with peptides in different molar ratios and incubated overnight at 4°C. Crystallization trials were done with several commercial screens from Hampton Research and Molecular Dimensions. After 4 days, small needle-like crystals appeared at 4˚C for *Eh*MySH3-P2 complex which was mixed in a molar ratio of 1:2, in crystal screen II (Hampton Research) conditions 0.2M Ammonium sulphate, 30% PEG 8000 and 0.2M Ammonium sulphate, 30% PEG 4000. These

crystals were used as seeds and bigger crystals were obtained by macroseeding in 0.2M ammonium Sulphate and 30% PEG 8000 condition. Diffractable crystals were flash frozen using mother liquor as a cryoprotectant in cold-room itself since the crystals were very temperature sensitive.

Figure 3.2-V: Crystals of EhMySH3-P2 peptide complex:

(A)Initial hits of were obtained in 30% PEG 8000, 0.2M Ammonium Sulphate condition of Crystal Screen II (Hampton Research) at 4°C. (B) Diffractable were crystals obtained after macroseeding with a few crystals from initial hits in the same condition.

3.2.4 Structure Determination:

The X-ray data for the peptide bound crystals of *Eh*MySH3 was collected at the BM14 synchrotron beamline (ESRF, Grenoble, France). The data sets were indexed and scaled using HKL2000 [12]. The structure was determined by molecular replacement by PHASER MR Program of CCP4 program suite [13,14] using PDB ID 5XGG (Crystal Structure C-terminal SH3 domain of Myosin IB from *Entamoeba* histolytica) as a template for *Eh*MySH3-P2 peptide crystals. After few cycles of manual building in COOT [15], refinement was done using REFMAC5 [16] in CCP4 suite followed by PHENIX refine[17] resulting in a model with R-factor 0.19 (R_{free} 0.24). Data statistics are given in Table 3.2-1.

Table 3.2-1: Data collection and refinement statistics for the EhMySH3-P2 complex.

Values given in brackets are for higher resolution shell.

Crystallographic Data	***Eh*MySH3-P2 complex**
X-ray Source	ERSF BEAMLINE BM14
Space group	$P2_12_12_1$
Wavelength (Å)	0.95
Unit Cell Parameters	
a,b,c (Å)	29.00, 60.0, 95.3
α, β, γ (°)	90, 90 , 90
Resolution range (Å)	50.0-1.98 (2.01-1.98)
R_{sym} or R_{merge} (%)	0.10 (0.66)
CC1/2	0.9
Completeness (%)	99.7 (98.4)
Redundancy	14 (12.3)
Mosaicity	0.3
Average I/σ(I)	27.7 (3.3)
No. of molecules in asymmetric unit	2
Refinement Statistics	
Resolution range (A°)	37.32-1.98 (2.05-1.98)
Total no. of observations	109431 (10854)
No. of unique observations	12139 (1153)
R_{work}	0.19 (0.23)
R_{free}	0.23 (0.25)
Mean B factor ($A^{\circ 2}$)	32.04
No. of atoms	Protein 1099/ Water 51/ ligand 15
RMSD deviations	
Bonds (A°)	0.02
Bond angles (°)	1.97
Dihedral angles (°)	20.3
Cross validation error	0.14
Ramachandran Statistics	
Most Favoured Region (%)	91.7
Allowed Region (%)	7.3
Generously Allowed Region (%)	0.6
Disallowed region (%)	0

3.2.5 Binding studies with SPR (Surface Plasmon Resonance):

For binding studies of *Eh*MySH3 with selected peptides, Autolab SPR was used at Advanced Instrumentation Research Facility, Jawaharlal N*eh*ru University, New Delhi, India. Here, the 11-mercaptoundecanoic acid (11MUA) monolayer on the surface of the gold chip (Autolab) was activated by ***N***-hydroxysuccinimide (NHS; 0.05 M)/***N***-ethyl-***N***-(diethyl aminopropyl) and carbodiimide (EDC; 0.2 M). 37µM of *Eh*Mysh3 protein (ligand) was immobilized in channel 1 on the activated surface of the gold chip in 10 mM sodium acetate buffer (pH 5.0) up to 700RU. During experimentation, channel 1 with immobilized *Eh*Mysh3 was used to pass the analytes (peptides) while channel 2 was used as blank (the signals of the analyte with a ligand-free surface). After ligand immobilization, the surface was blocked with 100mM ethanolamine at pH 8.5, followed by regeneration using 50 mM NaOH. Different concentrations of each peptide were injected at the rate of 20 µl/min across the sensor chip surface and the association/dissociation kinetics was monitored for 200s/ 100s respectively. All the dilutions were done in running buffer, 10mMHEPES, 150mMNaCl, 3 mM EDTA, 0.05% NP-40 surfactant [pH 7.4].

For interaction studies with purified GEFAPC domain, 1.175 ng/mm^3 of *Eh*MySH3 was immobilized on the activated chip surface and different concentrations of GEFAPC domain was passed over it. Association kinetics was studied for 400s while the dissociation was seen for 150s. 1M NaCl was used for regeneration of the chip surface in case of protein-protein interactions. Spr studies with F-actin was carried out by immobilizing 0.67ng/mm^3 of filamentous actin on chip surface using buffer 50mM sodium acetate pH 4.1, 0.05mM $CaCl_2$, 0.05mM ATP and 0.125mM DTT. (0.5mg/ml F-actin solution was obtained by dissolving Actin powder (Sigma) in 50mM sodium acetate pH 4.0, 12.5mM KCl, 0.5 mM $MgCl_2$ and 0.25mM ATP).

All of the data were recorded at 25°C. For kinetic evaluation, differential sensorgrams were analyzed using Autolab SPR Kinetic Evaluation software.

3.2.6 GST Pull down Assay:

Purified GSTSH3 and GST alone were incubated for 1hr at 4˚C with Glutathione sepharose (GE Healthcare Lifesciences), pre-equilibrated in equilibration buffer (50mM Tris pH-7.4, 150mM NaCl). Then, it was washed thrice with the equilibration buffer and incubated with the same amount of purified GEFAPC for 2 hrs at 4 ˚C. Following 3-4 washes with wash buffer (50mM Tris pH-7.4, 300mM NaCl), proteins were eluted with 20mM reduced glutathione, 50mM Tris pH-7.4, and 150mM NaCl. The fractions were loaded on to 12% SDS-PAGE and analyzed.

3.2.7 Actin co-sedimentation Assay:

23µM of purified G-actin was dissolved in G-buffer (20mM Tris-Cl, 0.5mM DTT, 0.2mM ATP, 0.1mM $CaCl_2$and 0.1mM NaN_3) and was allowed to polymerise for 1½ hour at RT by adding I_{mix} solution (50mM KCl, 2mM $MgCl_2$, and 0.5mM ATP) to it. After polymerization, different concentrations of target protein were added in a total volume of 200µl and incubated for 30 minutes at RT. The mixture was centrifuged at 40,000 rpm for 2 hours in Beckman ultracentrifuge. Pellet and supernatant fractions were collected and analyzed on 12% SDS-PAGE followed by Coomassie blue staining.

3.2.8 F-actin bundling Assay:

Polymerized F-actin was prepared as described previously. After polymerization, different concentrations of target protein were added in a total volume of 50µl and incubated for 30 minutes at RT. The mixture was centrifuged at 12,000 rpm for 10 minutes. Pellet and supernatant fractions were collected and analyzed on 12% SDS-PAGE followed by Coomassie blue staining.

3.2.9 Antibody generation:

The purified antigenic proteins (*Eh*My1TD for m*Eh*Myo1B and *Eh*GEFD for mGEFD) were dialyzed against PBS. Six mice were immunized subcutaneously with 100 µg of protein per mice per injection with an interval of 2 weeks between each injection. The first dose of the protein was emulsified with complete Freund's adjuvant while the following doses were emulsified with incomplete Freund's adjuvant. Following the inoculation series, the serum was stored in aliquots at -80°C. rGEFAPC polyclonal antibody was custom generated commercially (Link biotech). To check the titer of the polyclonal antibody, western analysis of the *E. histolytica* lysate was performed.

3.2.10 *E. histolytica* culture maintenance:

All the culture work was done in collaboration with Prof. Alok Bhattacharya using his in-house culturing facility. Culture maintenance was done with the help of Mr. Sabir Ali. *E. histolytica* strain HM-1-IMSS trophozoites were grown axenically in TYI-S-33 medium [18]. The cells were maintained and grown in TYI-33 medium complemented with 15% adult bovine serum, 1X Diamond's vitamin mix and antibiotic (125 µl of 250 U/ml Benzyl Penicillin and 0.25 mg/ml Streptomycin per 90 ml of medium).

3.2.11 Transfection of *E. histolytica* trophozoites by electroporation:

Transfection was performed by electroporation as described previously [19]. Briefly, trophozoites in log phase were harvested and washed with PBS followed by incomplete cytomix buffer [10 mM K_2HPO_4/KH_2PO_4 (pH 7.6), 120 mM KCl, 0.15 mM $CaCl_2$, 25 mM HEPES (pH 7.4), 2 mM EGTA, 5 mM $MgCl_2$]. The washed cells were then resuspended in 0.8 ml of complete cytomix buffer (incomplete cytomix containing 4 mM adenosine triphosphate, 10 mM glutathione) containing 200 µg

of plasmid DNA and subjected to two consecutive pulses of 3000 V/cm (1.2 kV) at 25 µF (Bio-Rad, electroporator). The transfectants were initially allowed to grow without any selection. Drug selection was initiated after 2 days of transfection in the presence of 10 g/ml G-418 for constructs with luciferase reporter gene or 10 g/ml of hygromycin B was used for tetracycline inducible constructs.

3.2.12 Preparation of *E. histolytica* cell lysate:

One million trophozoites growing in log phase were harvested at 280 xg for 7 min at 4°C. The pellet was then washed with chilled PBS pH 7.2 and pelleted as before. For probing *Eh*FP10, washed cells were fixed with 3.7% pre-warmed paraformald*eh*yde for 60 min at room temperature. After the cells were fixed, they were again harvested at 280 xg for 7 min. In both the cases, the pellet was heated at 80°C for 1 min. It was then resuspended in 2x SDS dye without β-ME. The lysate was heated at 100°C for 7 minutes and centrifuged at 13000 rpm for 15min to pellet down the debris. The supernatant was aliquoted and stored and quantification was done by BCA.

3.2.13 Protein Estimation by BCA assay:

The amount of protein was estimated by the bicinchoninic acid (BCA) assay using BSA as the standard. The working solution was prepared by mixing BCA (Sigma) and 4% copper sulphate in a ratio of 50:1. Equal volumes of the sample and the working solution were mixed in a microtiter plate and incubated at 37°C till a purple color develops in the lowest concentration of BSA. The absorbance was taken at 560 nm using a microtiter plate reader (Bio-Rad, USA).

3.2.14 Western blotting:

SDS-Polyacrylamide gel to be transferred was incubated in transfer buffer for 5 mins. The treated gel was placed on two sheets of Whatman 3 mm paper cut to the size of the gel, saturated with transfer buffer. A sheet of PVDF membrane pre-

activated by soaking in methanol followed by transfer buffer was placed on the gel taking care that no air bubble(s) were trapped in between the membrane and the gel. Three sheets of Whatman 3 mm paper were placed enclosing the membrane and gel. The transfer was done in wet transfer apparatus (Genei) at a costant voltage of 50mV for 500s.The membrane was then stained with Ponceau S and was blocked with 5 % skimmed milk powder in PBS-T (PBS containing 0.05 % Tween 20) for 1 ½ hour at RT. Primary antibody followed by secondary antibody incubation was done in 3 % milk powder in PBS-T with shaking at RT for 2 hrs and 30min, respectively. The blots were washed thoroughly with PBS-T after every incubation with antibody. The secondary antibody used was horse radish peroxidase conjugated IgG. Band detection was done using ECL kit (Millipore).The antigens were detected with polyclonal antibodies m*Eh*My1TD for Myosin IB and mGEFD for *Eh*FP10 raised mice at a dilution of 1:2500 and 1:2000 respectively, followed by secondary anti-mice immunoglobulins conjugated to HRPO of dilution 1:10,000 (Sigma). ECL reagents were used for visualization (Millipore).

3.2.15 Immunofluorescence staining:

E. histolytica trophozoites and transfectants were resuspended in incomplete TY1-33 medium and transferred onto acetone-cleaned coverslips placed in a petri-dish. The cells were allowed to adhere for 5 min at 37°C. The culture medium was discarded and the cells were fixed with 3.7% pre-warmed paraformald*eh*yde for 30 min. After fixation, the cells were permeabilized with 0.1% Triton X-100/PBS for 5 min, washed with PBS and then quenched for 30 min in PBS containing 50 mM NH_4Cl. The coverslips were blocked with 1% BSA/PBS for 2h, followed by incubation with primary antibody at 37°C for 1.5h. The coverslips were washed 3 times with 1% BSA/PBS before incubation with secondary antibody for 45 min at 37°C. Antibody dilutions used were: m*Eh*Myo1B at 1:100, anti-GFP (Sigma) at 1:100, rGEFAPC at 1: 150, anti-rabbit/mice Alexa 488, Alexa 556 and Pacific blue-410 (Molecular Probes) at 1:250, TRITC-Phalloidin at 1:250. Cells were further washed with 1% BSA/PBS twice and then PBS once and mounted on a glass slide using DABCO (1, 4-diazbicyclo (2,2,2) octane (55) 2.5% in 80% Glycerol). The edges

of the coverslip were sealed with nail-paint to avoid drying. Confocal images were visualized using an Olympus Fluoview FV1000 laser scanning microscope.

3.2.16 Fluorescent labeling of RBCs:

RBCs were stained with CFDA (Carboxyfluorescein succinimidyl ester, Thermo Fisher Scientific Cat no. C1157). RBCs were collected by pricking a finger with a needle and collecting the blood in PBS. The cells were washed with PBS twice followed by incubation in PBS containing 10 µM CFDA for 10 min at 37°C with intermittent tapping. The reaction was stopped by washing the RBCs with PBS and then kept in ice.

3.2.17 Time-lapse imaging:

Amoebic cells expressing n-terminal GFP tagged *Eh*FP10, NGFPGEF were plated onto a 35 mm glass bottom dish and allowed to adhere to the dish. Labeled RBC or TRICT-Dextran (Sigma) containing media was added to it and time-lapse imaging was done using a Spinning Disk confocal microscope (Nikon A1R, Optics- Plan Apo VC606 oil DIC N2, Camera- Nikon A1, NA-1.4, RI-1.515). The temperature was maintained at 37°C with the help of a chamber provided along with the microscope. The images were captured at 500ms interval. The raw images were processed using NIS element 3.20analysis software.

3.3 Results:

3.3.1 Identification of potential binding partner for the *Eh*MySH3 domain.

For our study, we utilized a previously published proteomics data [20] which was obtained after Myosin IB was overexpressed *in E. histolytica* cells activated for erythrophagocytosis. Since we were focusing on the binding partners for the SH3 domain of *Eh*Myosin IB, hence, we tried to screen all the proteins in the cytoplasmic fraction for the presence of PXXP as well as class I and II motifs using SH3 HUNTER web server [21]. Out of all the predicted proteins, those which had signature PXXP motifs, similar to the ones in yeast MYO3 and MYO5 (similarity was predicted by the software itself) were chosen for further studies. An FYVE family Rho-GEF, *Eh*FP10 (Accession no. EAL46050.1) was seen to harbor a maximum number of predicted motifs i.e. 13, some of them were continuous and overlapping each other. *Eh*FP10 is the longest among 12 and has a PH, FYVE and a C-terminal APC domain (GEF-APC). *Eh*FP10 is also among highly expressed *Eh*FP genes [4]. A Rho family GTPase, Rac G (Accession no. EAL46413) were also predicted to have PxxP type motifs. Previously no unconventional myosin I homologues have been reported to be involved in regulation of Rho signaling, hence we proceeded further with *Eh*FP10 and Rac G in our current study.

Sh3 domain recognizes PXXP motif in proteins which are further classified into class I [(K/R)xxPxxP] and class II [xPxxPx(K/R)]. In our previous analysis of PEG bound SH3 structure (chapter 1), we had made predictions about the ligand preferences of *Eh*MySh3. Hence, to confirm the type of ligand interactions preferred by *Eh*MySH3 domain, we selected peptide sequences from *Eh*FP10 belonging to class I, class II and an overlapping sequence belonging to both the classes, on the basis of prediction score as given by SH3 HUNTER (Table 3.3-1) and two sequences from Rac G was also chosen but both were PXXP type. All the

synthesized peptide sequences were as they were present in their respective proteins, no extra amino acids were added.

Table 3.3-1: List of peptides selected on the basis of SH3 Hunter software (http://cbm.bio.uniroma2.it/SH3-Hunter/).

Accession No.	**Protein**	**Peptides**
EAL46050.1	RhoGEF-PH-FYVE	656-661 **PPIPHR** 794-800 **RPQPPTP** 515-520 **PKLPPK** 661-667 **RSLPPKP** 572-577 **PTVPPK**
EAL52122.1	p21-activated kinase	151-156 **PSVPAP**
EAL49974.1	Rho guanine nucleotide exchange factor, putative	447-452 **PELPQY** 413-419 **VAPPSLP** 444-450 **REPPELP** 416-421 **PSLPNI**
EAL51743.1	protein kinase2	354-359 **PWIPPV**
EAL46413.1	Rho family GTPase	108-113 **PNVPII**
EAL44666.2	gamma-adaptin	670-675 **PINPTP**
EAL51552.1	phosphatidylinositol 3-kinase, putative	536-541 **PLNPRV**
EAL50557.1	myotubularin	189-194 **PDLPST**
EAL46989.1	protein kinase with WD repeats	329-334 **PNVPKE**

3.3.2 *Eh*MySH3 showed a higher binding response for pseudo-symmetrical peptide, P2 from *Eh*FP10.

*Eh*MySH3-peptide interaction was studied using Autolab SPR by passing different peptides on immobilized *Eh*MySH3. In SPR studies, the P2 peptide, which was a pseudo-symmetrical ligand containing an overlapping class I and class II motif from

*Eh*FP10 gave the highest response relative to others peptides as was predicted in chapter 1, while P4 peptide (class I) gave the least response (Figure 3.3-I). The binding affinity of the peptides with the *Eh*MySH3 domain could not be calculated since the molecular weight of all the peptides was around 1000 Da, which lies in the minimum detectable range for the instrument. Hence, on the basis of relative response, we proceeded for co-crystallization of the *Eh*MySH3 with the high response peptides.

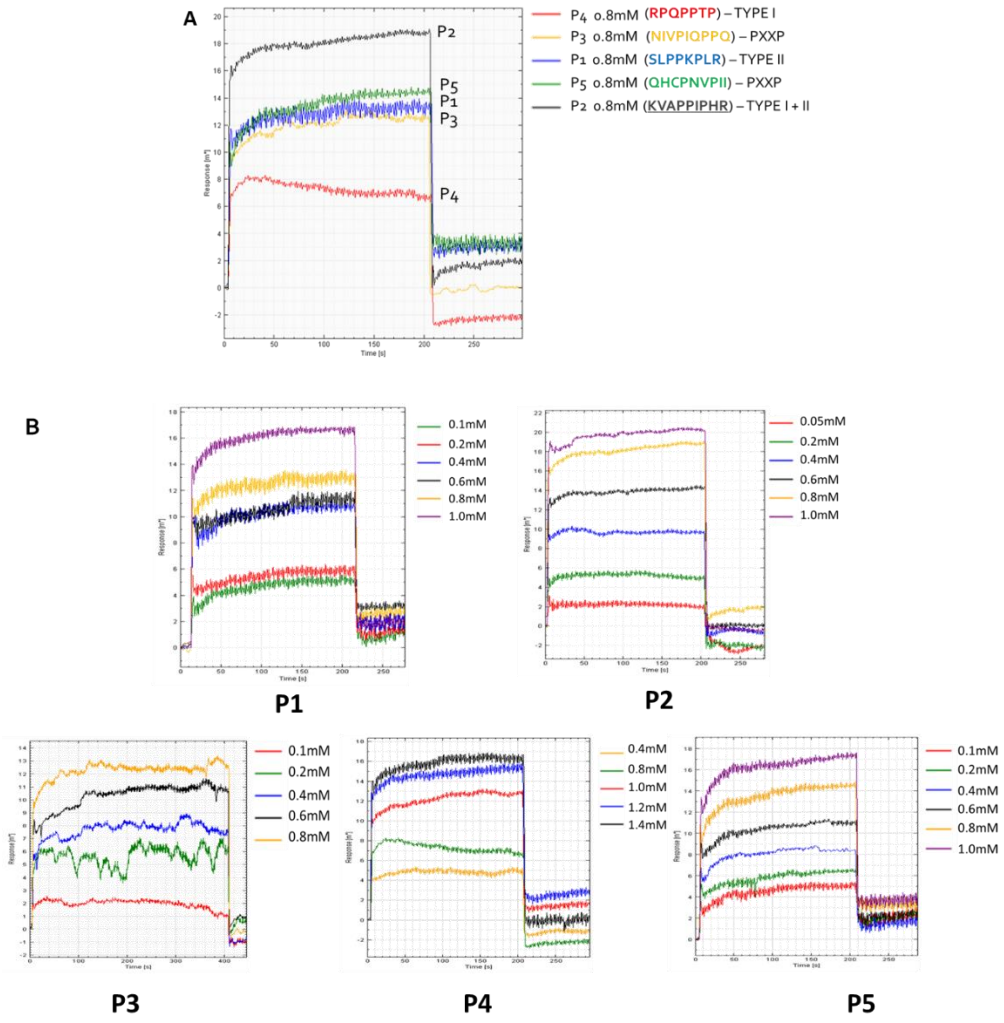


Figure 3.3-I: Binding Studies with SPR.

(A) SPR sensorgram depicting binding response when 0.8mM of each peptide was passed over immobilized *Eh*MySH3 protein in Autolab SPR. (B) SPR kinetic analysis with each of the selected peptide over immobilized *Eh*MySH3 protein as a ligand.

3.3.3 *Eh*MySH3 prefers class II mode for interaction with ligands/ P2 peptide.

We could co-crystallize P2 with *Eh*MySH3. In *Eh*MYSH3-P2 complex structure, asymmetric unit is composed of two molecules of *Eh*MySH3 and one peptide, P2 molecule near chain B (**Error! Reference source not found.** A). On expanding to symmetry molecules, P2 was found to be shared by Chain A and Chain B of *Eh*MySH3 (**Error! Reference source not found.** C). A Close analysis of electrostatic interactions made by P2 with both the chains revealed that P2 makes more interactions with chain B than chain A. P2 binds with *Eh*MySH3 chain B in class II fashion, i.e. [**xP**x**xP**x(**K/R**)]. A^{35} and P^{36} forms the first XP motif which fits in the first consensus hydrophobic pocket made by Y^{9} and Y^{52} residues of *Eh*MySH3 and I^{38} and P^{39} forms the second XP motif which occupies the second hydrophobic pocket made by P^{49} and W^{36} residues of *Eh*MySH3. The C-terminal R^{41} residue of P2 interacts with E^{34} residue in the n-src loop of *Eh*MySH3 forming the specificity pocket of *Eh*MySH3. In other known ligand bound SH3 structures, terminal Arg or Lys residues interact with a residue in the RT loop which forms the specificity pocket. In our case, though the density of the terminal R^{41} residue is not complete, but after refinement, we get a better density near n-src loop though there is some green density near E^{18} in the RT loop, which could be one of the probable orientation for R^{41} residue (**Error! Reference source not found.** D). Also, when we superimposed unbound native Sh3 structure with *Eh*MySH3-P2 complex, the only difference was seen in β2 and n-src loop, where β2 becomes one residue longer making n-src bent towards the R^{41} of the peptide, P2. Overall, we could conclude that *Eh*FP10 peptide, P2 binds with *Eh*MySH3 and leads to clustering of the two *Eh*MySH3 domains due to its pseudo symmetrical sequence, but prefer class II mode of peptide binding orientation. (**Error! Reference source not found.**)

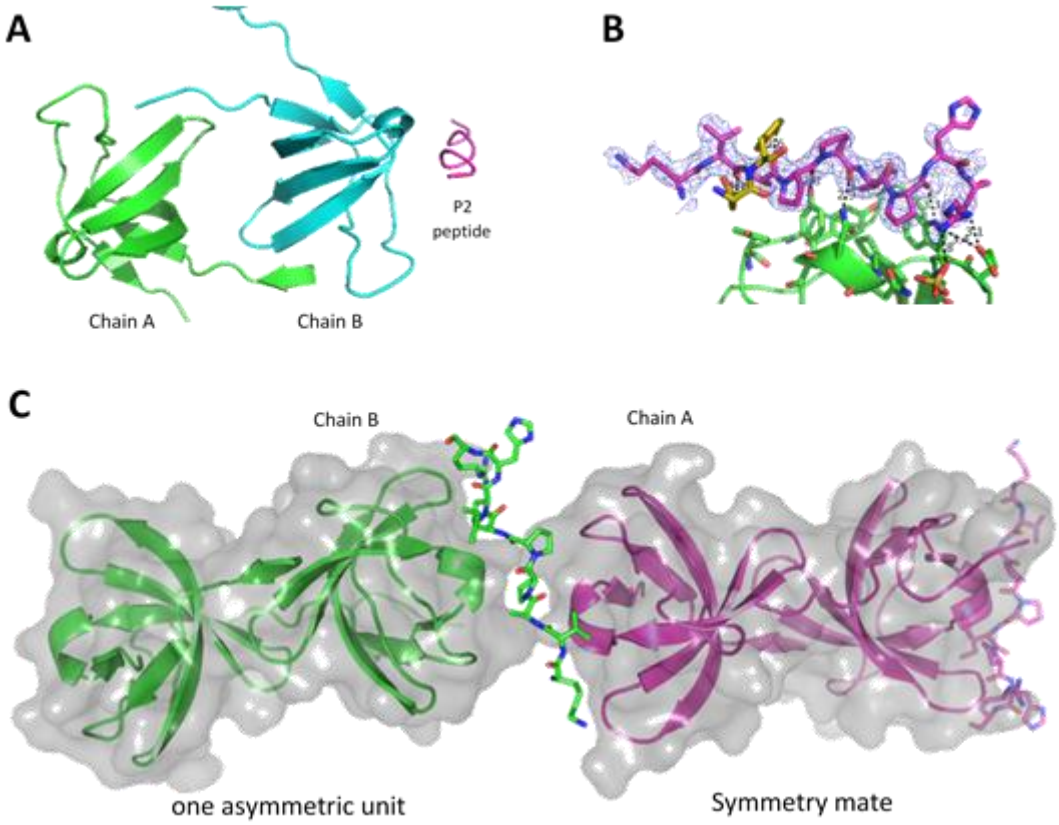


Figure 3.3-II: Crystal structure of EhFP10-P2 complex.

(A) An asymmetric unit of the complex is composed of two protein molecules and one molecule of P2 near chain B. (B) 2Fo-Fc electron density map of the peptide, P2 co-crystallized with *Eh*MySH3 at a 1.5σ cut-off. (C) The arrangement of the *Eh*MySH3-P2 complex with its symmetry mate showing P2 peptide being shared by two molecules of *Eh*MySH3 polypeptide.

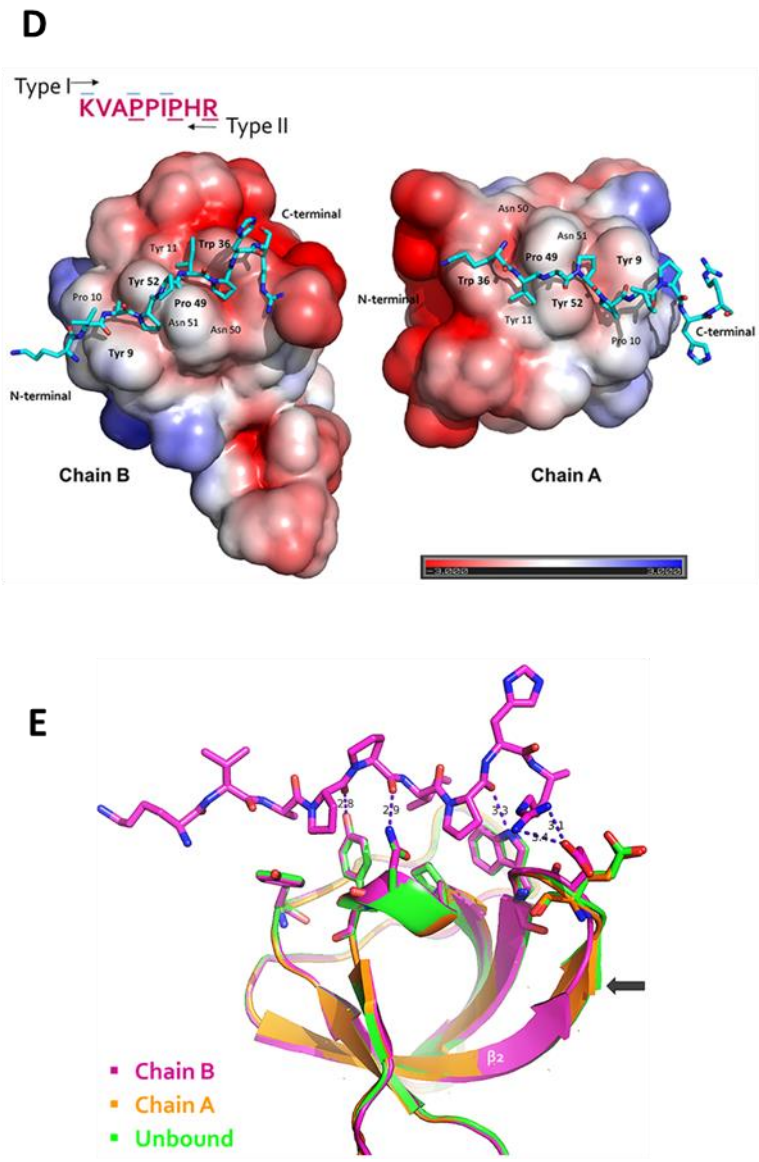


Figure 3.3-II: Crystal structure of EhFP10-P2 complex (cont.).

(D) Surface-charge representation of bound peptide with both the chains with c-terminal ARG residue located in the specificity depicting class II orientation of binding with Chain B. (D) Superimposed Chain A, B and native unbound *Eh*MySH3 shown the only visible difference in β2-strand on peptide binding.

3.3.4 *Eh*Myosin IB interacts with c-terminal APC domain of an FYVE family Rho guanine nucleotide exchange factor (*Eh*FP10).

*Eh*FP10 peptide, P2 is located between 653-661 residues in the c-terminal APC basic domain of FYVE family Rho guanine nucleotide exchange factor (*Eh*FP10) from *E.histolytica*. An analysis of domain organization of all 12 FYVE family GEFs, present in E.histolytica, *Eh*FP10 was found to be the only GEF which has a c-terminal APC domain. Also, our BLAST search could not find any such GEF with APC basic domain in other closely related amoeba like *Acanthamoeba*, *Dictyostelium* as well as in humans. Though it is present in all *Entamoeba* species.

To further strengthen *Eh*MySH3-GEFAPC interaction pull-down with recombinant *Eh*GEFAPC protein was performed. In GST pull down assay, purified protein got eluted along with GST tagged *Eh*MySH3 protein. In **Error! Reference source not found.** A, APC protein could be seen along with GSTSH3 protein in Lane 2. No APC protein was present with GST alone. The affinity of *Eh*MySH3-GEFAPC interaction was found to be 200nM by SPR studies where different concentrations of GEFAPC protein was passed over immobilized *Eh*MySH3 (Figure 3.2-III).

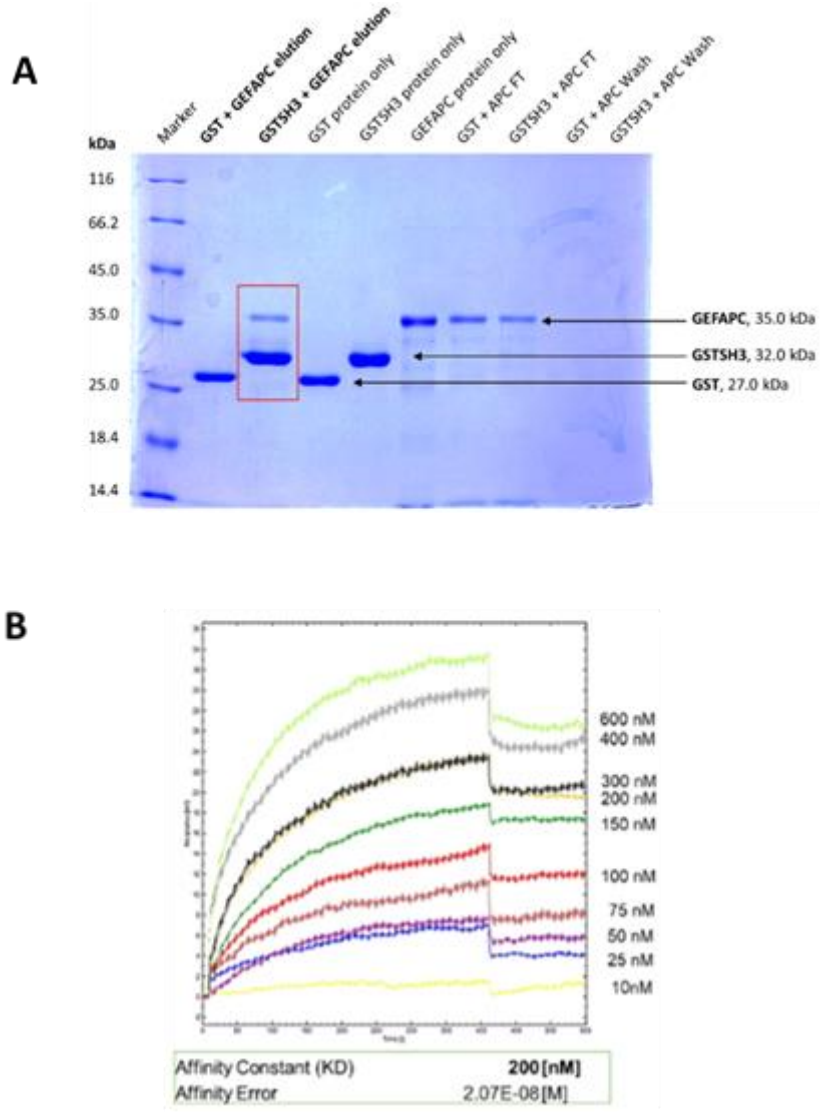


Figure 3.3-III : EhMySH3 interacts with APC domain of EhFP10.

(A) 12 % SDS PAGE showing results after GST-pull down assay. GEFAPC was seen only in GSTSH3 fraction, not in GST alone. (B) SPR sensorgram depicting results after different concentrations of GEFAPC was passed over immobilized EhMySH3 protein. KD was calculated to be 200nM.

3.3.5 *Eh*FP10 localizes in pseudopods during phagocytosis as well as pinocytosis in *E. histolytica* cells.

Unconventional Myosin IB (*Eh*MyosinIB) has been found to be localized in the cytoplasm of *E.histolytica* and is involved in the phagocytic process where it localizes into the tip of phagocytic cup till the pinching of vesicles [11]. Since *Eh*FP10 has been confirmed to be one of the interacting partners of *Eh*MyosinIB, so we proceeded to study the cytoskeletal distribution of *Eh*FP10 within motile *E. histolytica* trophozoites. For the localization studies, N-terminal GFP tagged full-length *Eh*FP10 (GFP-*Eh*FP10) constructs were used. Expression of the GFP-*Eh*FP10 was confirmed by using in house raised *Eh*FP10 polyclonal antibody as well as a commercial available anti-GFP polyclonal antibody (Sigma) (**Error! Reference source not found.**).

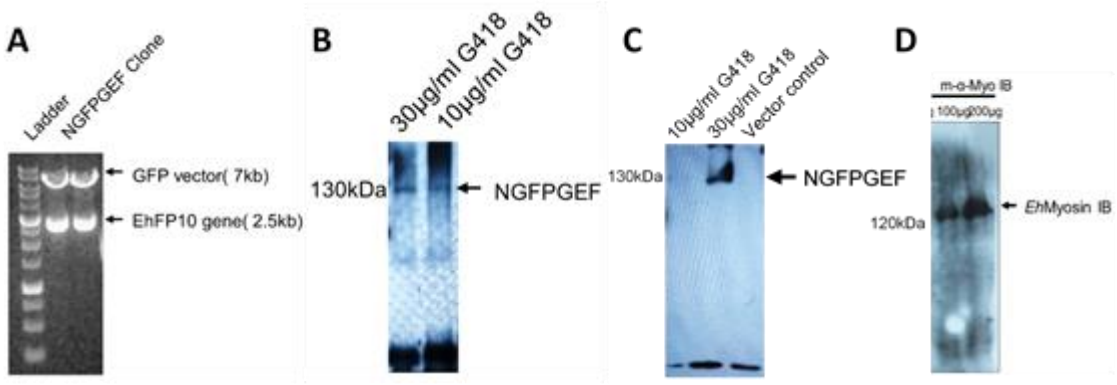


Figure 3.3-IV: Construct and antibody generation for in-vivo studies.

(A) Confirmation of NGFPGEF clone by double digestion with XhoI and BamHI. The presence of 2.5kb confirms *Eh*FP10 gene and 7kb band corresponds to p*Eh*-NEOGFP vector. (B, C) Expression of GFP tagged *Eh*FP10 protein within *E. histolytica* trophozoites induced at different concentration of G418 drug showing a band at 130kDa in western analysis probed by the mGEFD antibody(B) and r-anti-GFP antibody (C). (D) Expression myosin IB within wild type *E. histolytica* trophozoites by m*Eh*My1TD antibody showing a band at 120kDa.

Immunofluorescence studies revealed GFP-*Eh*FP10 to be a cytoplasmic protein as it was present all over the cytoplasm in steady state. Also, in some of the cells, it was found to be enriched in membrane invaginations and small cup like projections (**Error! Reference source not found.**)

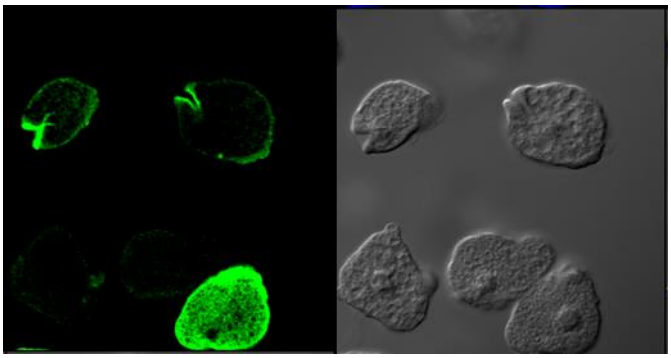


Figure 3.3-V: Immunofluoresence reveals EhFP10 to be a cytoplasmic protein enriched near membrane invaginations.

During live cell imaging with normal steady state motile *E.histolytica* cells, we could visualize GFP fluorescence to enrich near the membrane around the site of endosome formation, intensifies and move along the membrane till it reaches the site of cup formation. GFP-*Eh*FP10 was found to till the scission of the cup and internalized vesicles.

To confirm the endocytic process to be a pinocytosis, we labeled the media with Tritc-dextran dye. The presence of red fluorescence within internalized vesicles engulfed by GFP-*Eh*FP10 enriched pseudopods confirmed that *Eh*FP10 is involved in fluid phase macropinocytosis [22] in *E.histolytica*. *Eh*FP10 was present from initiation till scission of pinocytic vesicle as well as internalized macropinosomes (Figure 3.3-VII). A single macropinocytic event was completed in 4-5 seconds.

When GFP-*Eh*FP10 overexpressed cells were activated for erythrophagocytosis, *Eh*FP10 was seen to be enriched at the site of pseudopod formation near the cell membrane as soon as RBC comes in contact with *E.histolytica* cells, later it enriches into the pseudopods during the cup progression and was present till phagosome scission took place similar to pinocytosis (Figure 3.3-VIII). However, *Eh*FP10 was not found to be associated with internalized phagosome membrane as seen in the

case of pinocytosis. An erythrophagocytic process took around 7-8 seconds to complete.

To sum up, *Eh*FP10 was found enriched at the membrane around the site of pinocytosis and phagocytosis. The enriched in the pseudopods during cup progression and was present till the scission of the vesicle. It was found to be associated with the membrane of only with internalized macropinosomes but not phagosomes. A single pinocytic event was faster as compared to the phagocytic event.

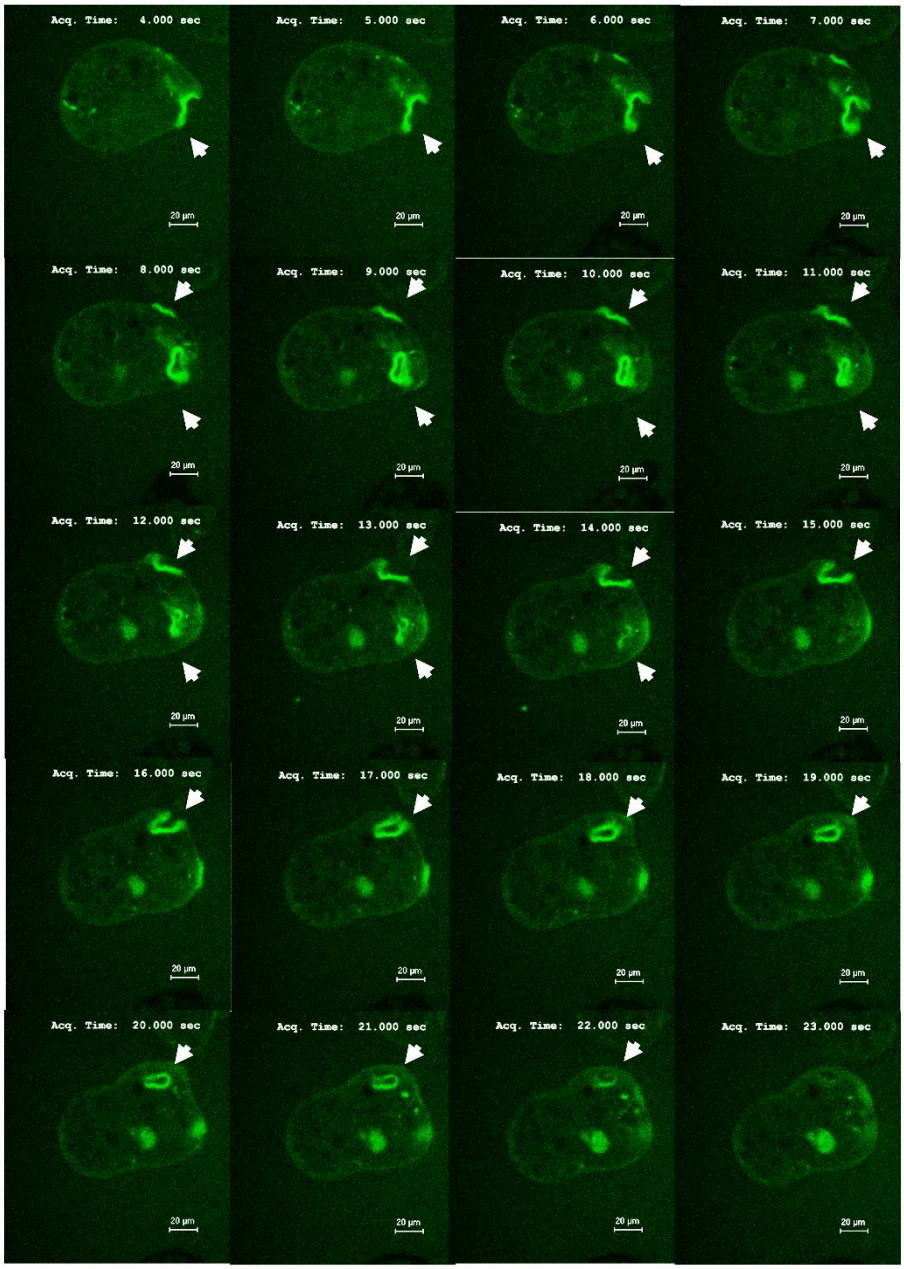


Figure 3.3-VI: Localization of EhFP10 within E.histolytica cells:

A montage representing time series of images of GFP-*Eh*ARPC1 expressing trophozoites. A number of pseudopods leading to cup formation, formed in different directions were marked by white arrowheads. Scale bar represents 20µm.

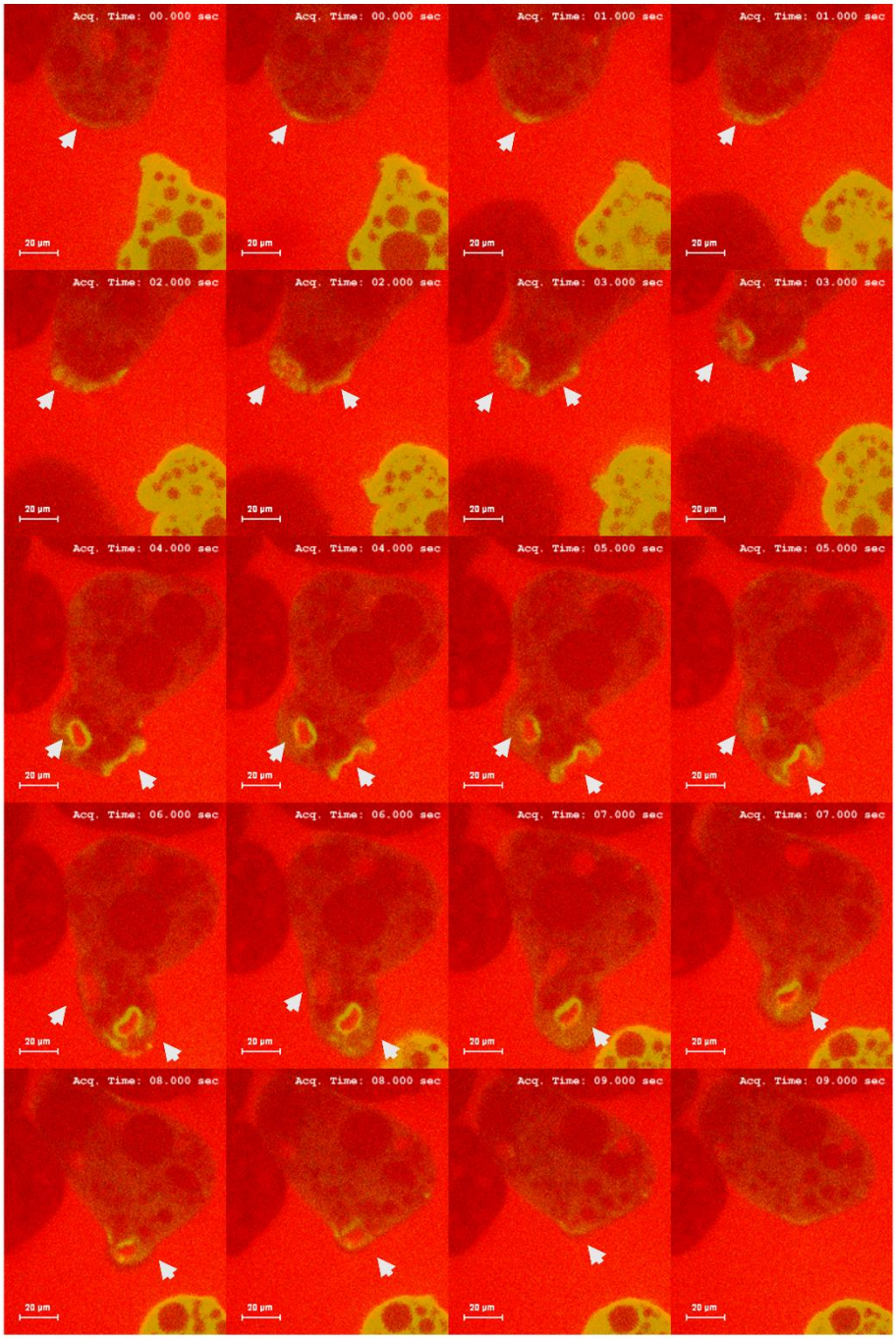


Figure 3.3-VII: Involvement of EhFP10 in pinocytosis.

A montage representing time series of images of GFP-*Eh*ARPC1 expressing trophozoites undergoing pinocytosis. Enrichment of *Eh*FP10 in progressing pseudopods and engulfment of labeled media (red colour) within the vesicle proves the endocytic process to be a pinocytic event. *Eh*FP10 leaves the pinosome membrane after internalization of the pinocytic vesicle. A number of pseudopods leading to pinocytic cup formation were marked by white arrowheads. Scale bar represents 20µm.

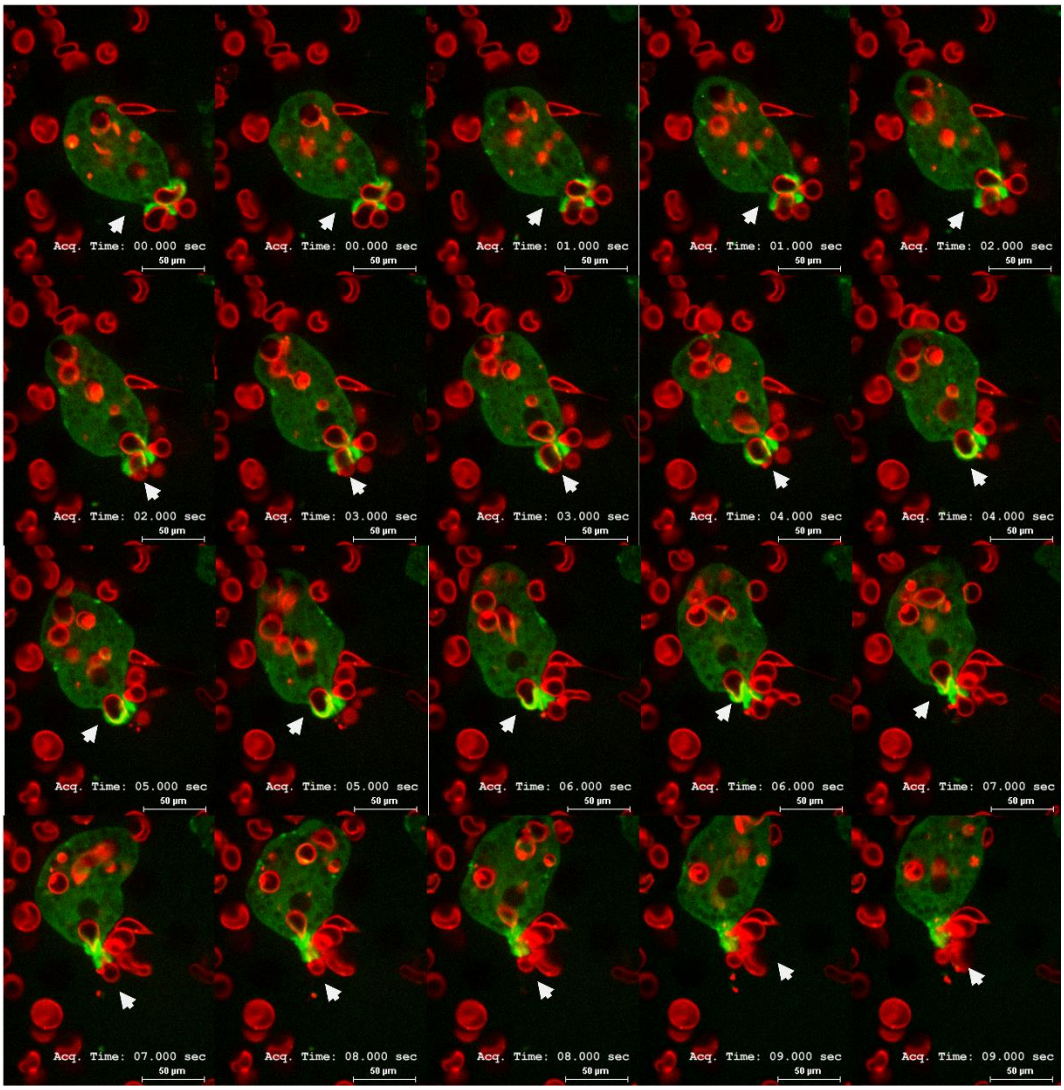


Figure 3.3-VIII: Involvement of EhFP10 in the phagocytic process in E. histolytica.

A montage representing time series of images of GFP-*Eh*ARPC1 expressing trophozoites undergoing phagocytosis. Enrichment of *Eh*FP10 in progressing pseudopods and in cups encircling RBC (labeled red in colour) leading to engulfment RBCs confirms the involvement of *Eh*FP10 in erythrophagocytosis. *Eh*FP10 is not present in the membrane of internalized phagosomes. A number of pseudopods leading to phagocytic cup formation were marked by white arrowheads. Scale bar represents 50µm.

3.3.6 C-terminal APC domain of *Eh*FP10 binds with actin and bundles actin filaments like classical APC basic domain and colocalizes with microtubules and actin filaments.

APC basic domain belongs to the adenomatous polyposis coli protein family and is known to be microtubule binding domain [23]. APC basic domain has been found to be localized with the actin cytoskeleton, involved in actin binding and preferably bundling of the actin filament. It could also crosslink actin filament and microtubules and mediates cytoskeleton crosstalk [24]. As reported before, cytoplasmic microtubules have not been observed in trophozoites of *E. histolytica* cells. (Isaura M et. al., Archives of Medical Research, Volume 37, Issue 2, February 2006). But, in GFP-*Eh*FP10 overexpressed cell lines, microtubules could be observed associated with phagocytic cup and colocalizes with GEF and actin filament, in immunofluorescence studies (Figure 3.3-IX, Figure 3.3-X). Cytoskeletal dynamics in *E. histolytica* is largely dependent on actin cytoskeleton hence, focussed on GEFAPC-actin interactions in our study.

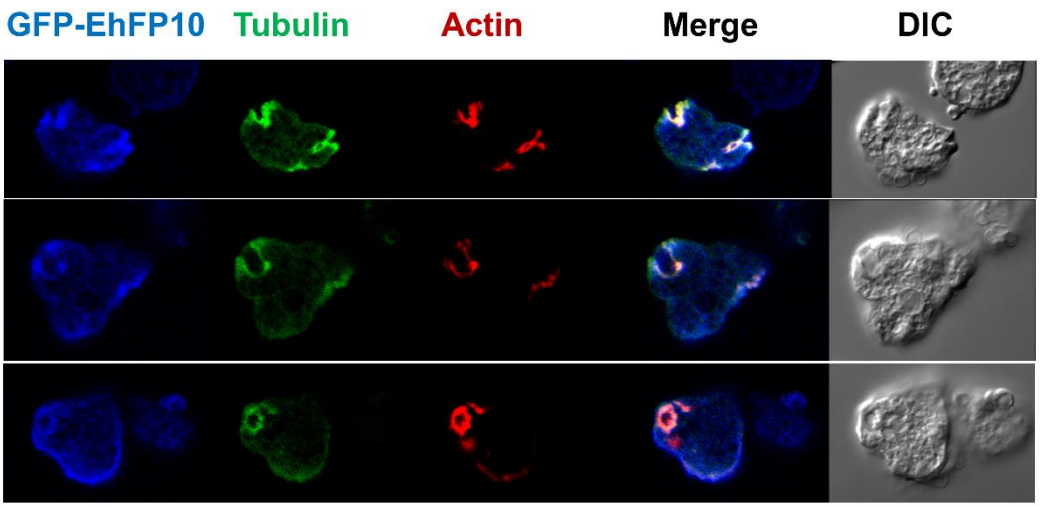


Figure 3.3-IX: Co-localization of EhFP10 with Microtubules in phagocytic cups.

In immunofluorescence studies, *Eh*FP10 was found to colocalize with microtubules in the phagocytic cups in *E.histolytica* trophozoites at all the stages of phagocytosis from initiation to internalized vesicles. Actin was used as a marker for the phagocytic cup.

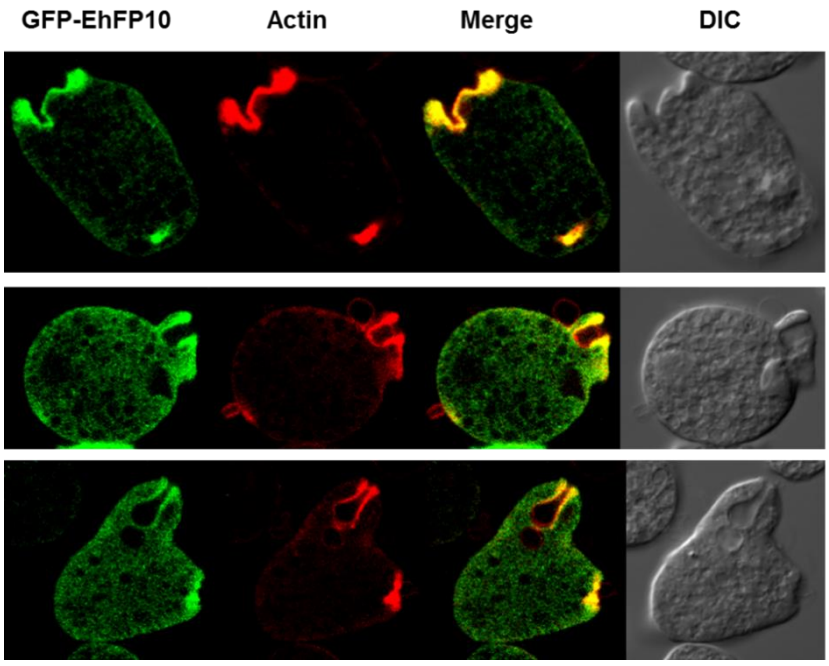


Figure 3.3-X: Co-localization of EhFP10 with actin NGFPGEF overexpressed cell lines.

*Eh*FP10 was found to co-localise with actin in phagocytic cups of *E.histolytica* trophozoites activated for erythrophagocytosis.

Similar to classical APC basic domain, purified recombinant *Eh*FP10-APC domain was also observed to bind and bundle purified rabbit muscle F-actin in centrifugation based actin binding (Figure 3.3-XII A) and bundling assays (Figure 3.3-XII C, D). At high-speed centrifugation, F-actin goes into the pellet fraction while G-actin stays in the supernatant, which forms the basis of the co-sedimentation assay. At low-speed centrifugation, only bundled thicker actin filament settles and forms the pellet while thin actin-filaments along with G-actin stay in the supernatant. During both the assays, the percentage of F-actin was found to be increased in the pellet fraction with an increase in the concentration of GEFAPC protein. Bound fraction of GEFAPC along with F-actin could also be seen in the pellet fraction while protein alone stayed in the supernatant. Imaging through TEM microscopy also depicted the formation of thicker actin bundle with increasing concentration of GEFAPC protein in actin bundling experiment. The Affinity of *Eh*FP10 with actin was calculated with SPR to be around 87nM, almost more than twice of *Eh*FP10-*Eh*MySH3 (Figure 3.3-XII B).

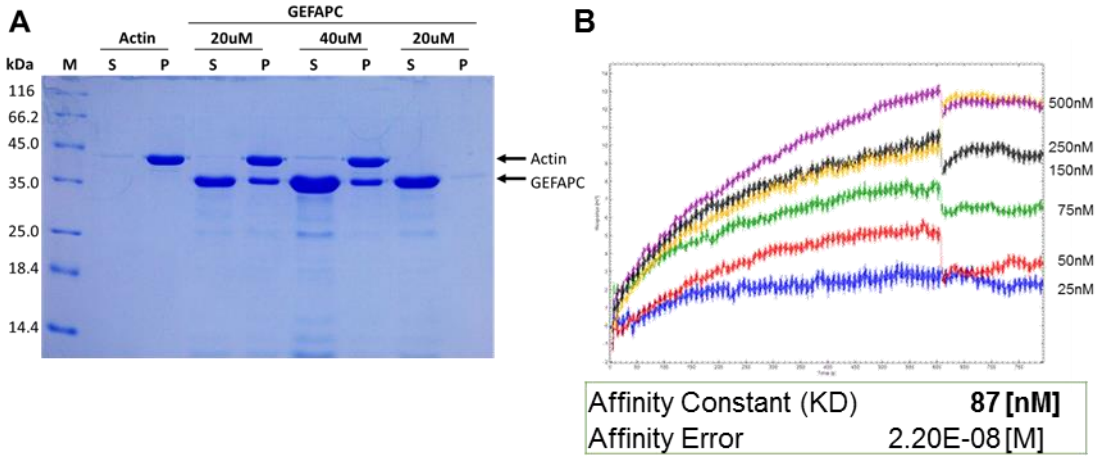


Figure 3.3-XI: APC domain of EhFP10 binds and bundles actin filament.

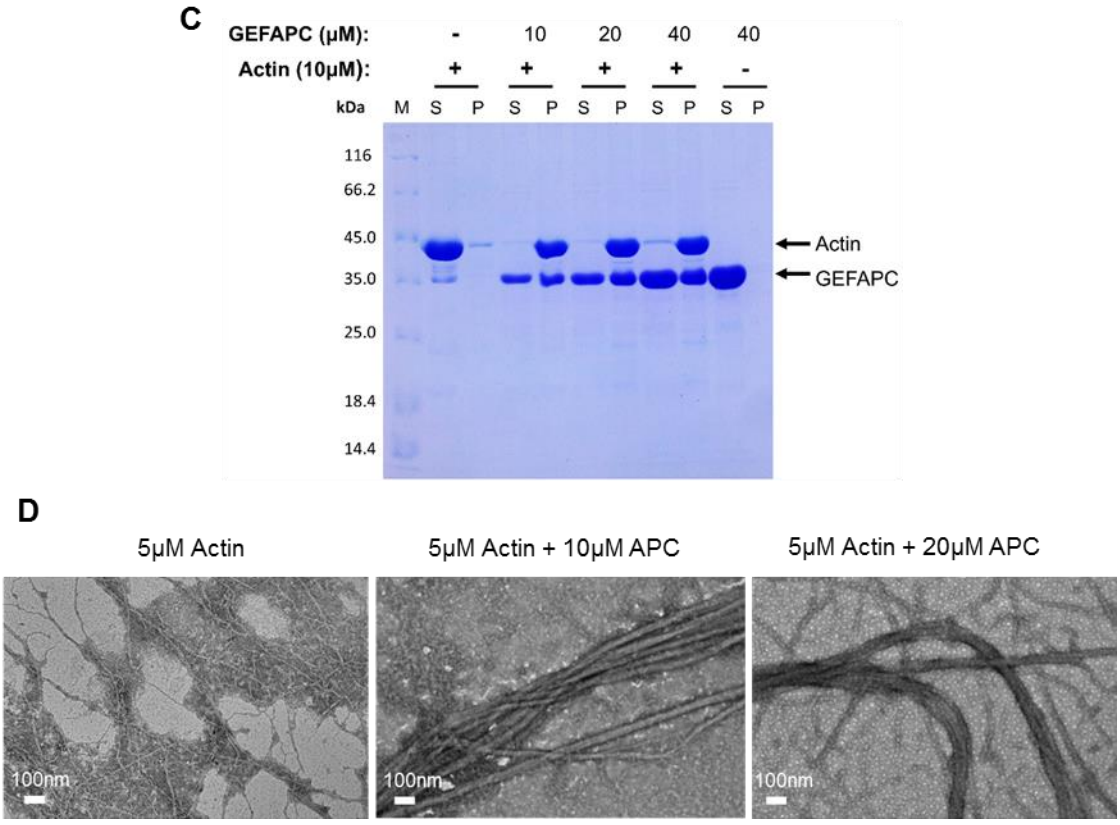


Figure 3.3-XII: APC domain of EhFP10 binds and bundles actin filament.

(A) Purified APC protein was seen to co-sediment with filamentous rabbit muscle actin and was found in pellet fraction during actin co-sedimentation assay. (B) SPR sensorgram plotted after different concentrations of GEFAPC protein was passed over immobilized F-actin having a KD value of 87nM after kinetic evaluation. (C) APC was found to bundle F-actin as only F-actin stays in supernatant while heavier bundled actin goes into pellet along with APC at low-speed centrifugation. As the concentration of actin increases more amount of bundled actin was seen in the pellet. (D) TEM images showing changes in the thickness of F-actin as APC concentration increases. Scale bar represents 100nm.

3.3.7 *Eh*Myosin IB co-localizes with *Eh*FP10 as well as F-actin during phagocytosis as well as pinocytosis.

Till now, we had proved that the SH3 domain of *Eh*Myosin IB interacts with the APC domain of *Eh*FP10. In immunofluorescence studies with GFP-*Eh*FP10 overexpressed cells, myosin IB was found to be colocalized with into the phagocytic as well as the pinocytic cup (Figure 3.3-XIII). *Eh*Myosin IB was found to co-localize with *Eh*FP10 as well as actin from initiation of the cup till scission of the vesicle. The *Eh*FP10 and *Eh*Myosin IB colocalizes at the proximal end of progressing pseudopods while actin localizes more in the base of the cup. As the cup was near to closing all the three colocalizes very well. Later on, *Eh*FP10 and *Eh*MySH3 move more towards the tip of the closing pseudopods (fig.). In previous studies, myosin IB has been reported to help in the pinching of the vesicle. On the basis of localization studies, it can be said that *Eh*FP10 also plays a role in the process. Neither *Eh*FP10 nor *Eh*Myosin IB was seen to be present within the internalized phagocytic vesicle but were present only at the tip (Figure 3.3-XIII A).

During pinocytosis, *Eh*FP10 and Myosin IB colocalizes very well from initiation, pseudopod extension, cup closure as well as within the internalized pinosomes encircling the vesicle (Figure 3.3-XIII B). The presence of *Eh*FP10 and *Eh*Myosin IB encircling the internalized vesicle only during pinocytosis and not in phagocytosis was quite interesting.

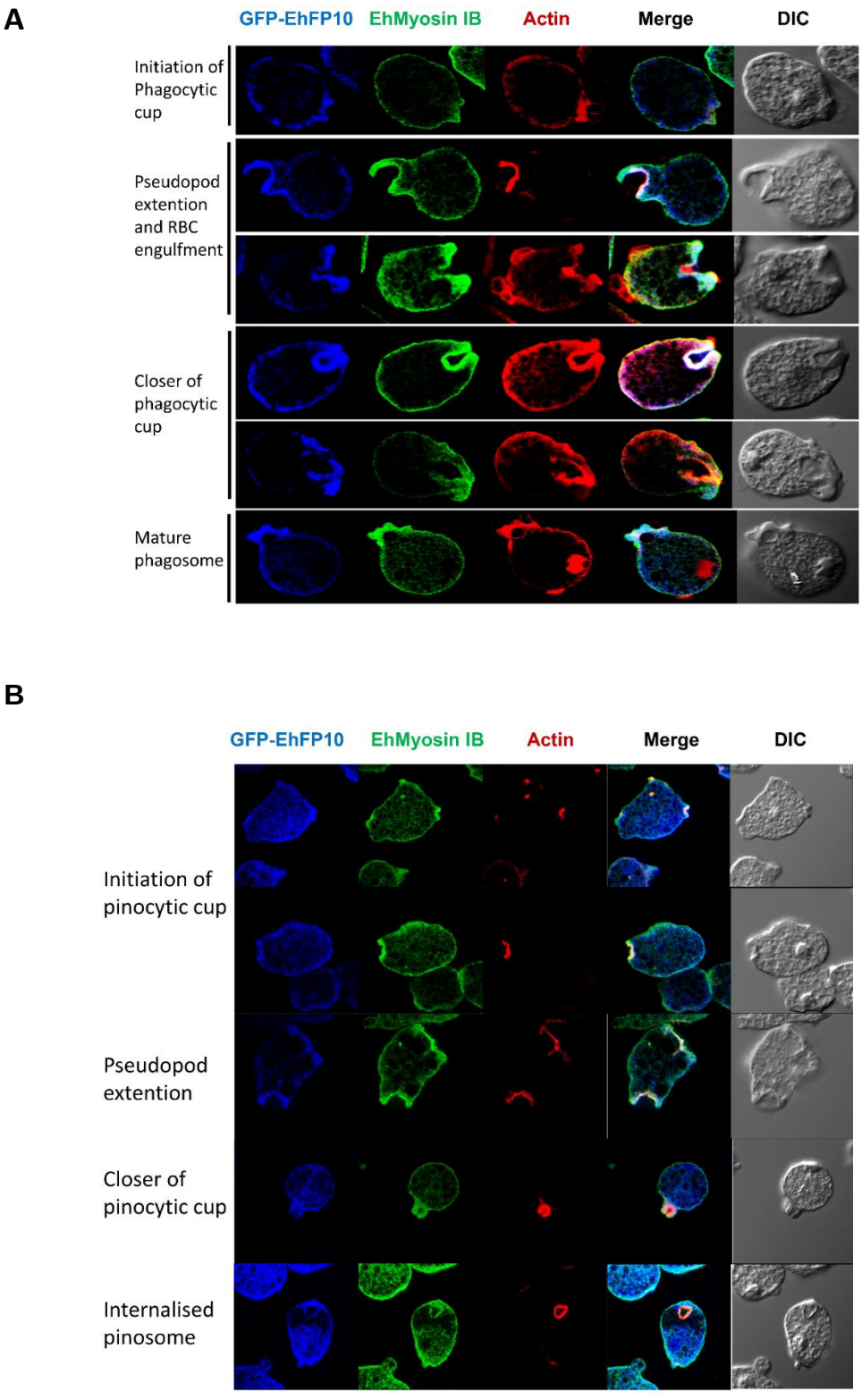


Figure 3.3-XIII: Colocalization of Myosin IB with EhFP10 during phagocytosis and pinocytosis.

(A) Myosin IB was found to colocalize with *Eh*FP10 at different stages of phagocytosis. (B) Myosin IB also colocalized with *Eh*FP10 during pinocytosis from initiation to cup closure as well as internalized pinosomes.

3.3.8 *Eh*MySH3 binds to *Eh*FP10 and inhibits its F-actin bundling activity

The SH3 domain of myosin IB from *E. histolytica*, *Eh*MySH3 didn't show any binding to rabbit muscle filamentous actin during centrifugation based actin co-sedimentation assay. *Eh*MySH3 could be seen in supernatant while filamentous actin was present in the pellet fraction (**Error! Reference source not found.** A). *Eh*FP10 had been found to bind actin filaments as well as bundle the thinner filaments into thick bundles. When actin bundling assay was performed in the presence and absence of *Eh*MySH3 at same GEFAPC concentration, an increase in the percentage of actin present in the supernatant fraction was observed in the presence of *Eh*MySH3. This states that *Eh*Myosin IB interaction with the FYVE-family GEF, *Eh*FP10 inhibits its bundling activity. Though at high GEFAPC concentration i.e. 40µM GEFAPC was able to overcome the inhibitory effect of *Eh*MySH3 and bundle majority of the actin filaments (**Error! Reference source not found.** C).

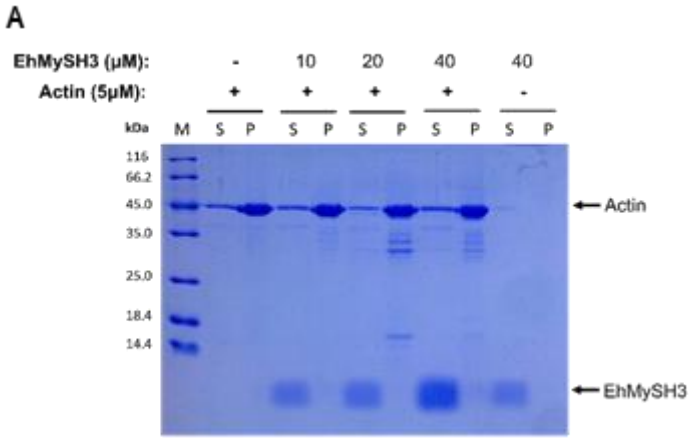


Figure 3.3-XIV: Sh3 domain of Ehmyosin IB interacts with APC domain of EhFP10 and inhibits its actin bundling activity.

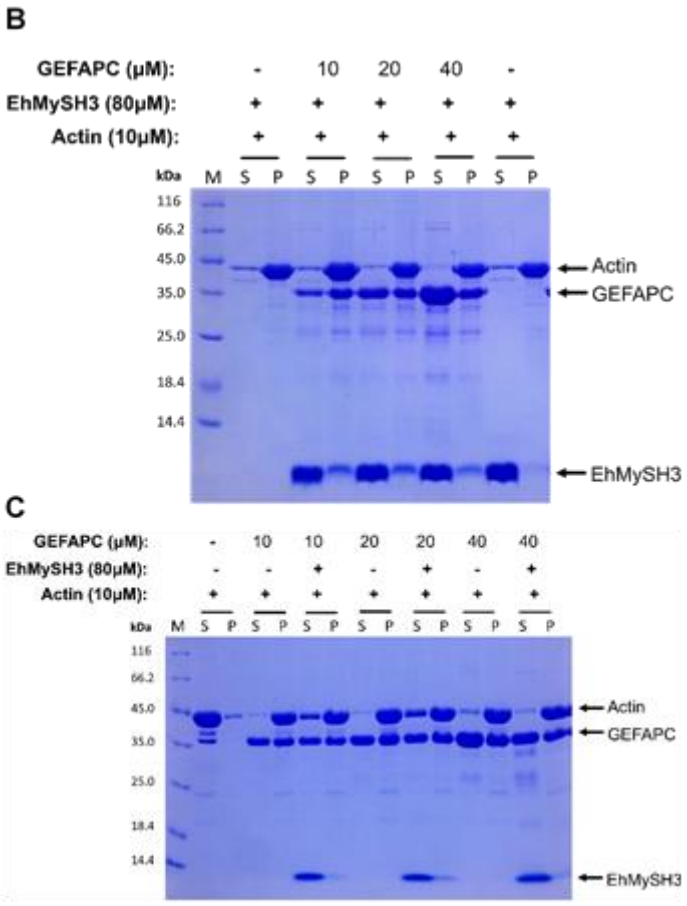


Figure 3.3-XV: Sh3 domain of Ehmyosin IB interacts with APC domain of EhFP10 and inhibits its actin bundling activity.

(A) The SH3 domain of EhMyosin IB did not bind to actin filament during actin co-sedimentation assay and can be seen only in the supernatant fraction. (B) A part of the EhMySH3 domain can also be seen in the pellet fraction in the presence of GEFAPC domain which binds with actin filament and bound fraction was seen in the pellet fraction during actin co-sedimentation assay. (C) During actin bundling assay, in the presence of EhMySH3 domain, a fraction of actin filament remains in the supernatant i.e. cannot form bundles. As the ratio of GEFAPC: EhMySH3 increased, GEFAPC could overcome the inhibitory effect of EhMySH3.

3.4 Conclusion:

On the basis of structural characterization of the c-terminal SH3 domain of myosin IB i.e. *Eh*MySH3 (Chapter 1), a few proteins were shortlisted which could be a probable binding partner of *Eh*MySH3. On the basis of SPR studies, a peptide from

*Eh*FP10, an FYVE domain containing GEF was found to show the best binding response. To validate the interaction, peptide P2 was successfully co-crystallized with the *Eh*MySH3 domain in $P2_12_12_1$ space group. Structure studies of *Eh*MySH3-P2 peptide complex revealed that P2 leads to clustering of the *Eh*MySH3 domain, and is shared by two molecules of *Eh*MySh3. Further analysis led us to predict that *Eh*MySH3 shows a preference for class II polyproline ligands and preferably binds to P2 in a class II manner.

Since *Eh*FP10 has not been characterized in *E. histolytica*, localization studies of *Eh*FP10 were done via time-lapse imaging of GFP-tagged *Eh*FP10 expressing *E. histolytica* trophozoites. *Eh*FP10 was found to suddenly appear near the membrane at the site of pinocytosis and phagocytosis in normal resting trophozoites and in cells induced for erythrophagocytosis respectively. *Eh*FP10 was present from initiation till the closure of the cup in both phagocytosis and pinocytosis. *Eh*FP10 was associated with the membrane encircling the newly internalized pinosomes and disappears after a while, but not in newly formed phagosomes. This interesting observation states that there are differences in the regulation of both the endocytic processes.

*Eh*Myosin IB was found to colocalize with *Eh*FP10 at all stages during pinocytosis and phagocytosis. P2 peptide resides in the c-terminal APC domain of *Eh*FP10. Thus, the interaction was further validated by in vitro pull down assay with APC domain of *Eh*FP10. APC domain of *Eh*FP10 was seen to bind with filamentous actin and form thicker actin bundles as seen for classical APC basic domains. Interaction of *Eh*MySh3 with APC domain of *Eh*FP10 was seen to inhibit the bundling activity by *Eh*FP10 at high *Eh*MySH3: *Eh*FP10 ratios while as the concentration of APC increases, it was seen to overcome the inhibition and bundle the actin filaments efficiently. This observation gives the functional relevance of *Eh*MySH3 interaction with *Eh*FP10 protein during phagocytosis and pinocytosis in *E. histolytica*.

The known effectors of the SH3 domain of unconventional myosin I have been found to regulate the actin dynamics directly or indirectly via arp2/3 complex. Several GTPases act upon these effectors and adaptor protein like WASP and other WASP related proteins like Vrp1, Bee1p, and Las17p which leads to actin

polymerization at the site of cup formation. As *E. histolytica* lacks WASP/SCAR protein, this highly expressing, unique, modular FYVE family Rho GTPase Exchange Factor, *Eh*FP10 with a c-terminal tail that can regulate actin dynamics is of great advantage to the pathogen. Our findings reveal a new mode of regulation of endocytic processes in highly motile phagocytosing gastric pathogen like *E. histolytica.*

3.5 References:

[1] Somlata and Bhattacharya, A. (2015) Phagocytosis in *Entamoeba* histolytica. In Amebiasis: Biology and Pathogenesis of *Entamoeba* (Nozaki, T. and Bhattacharya, A., ed.^eds), pp. 189-206. Springer Japan, Tokyo.

[2] Chimini, G. and Chavrier, P. (2000). Function of Rho family proteins in actin dynamics during phagocytosis and engulfment. Nature cell biology 2, E191-E196.

[3] Niedergang, F. and Chavrier, P. (2004). Signaling and membrane dynamics during phagocytosis: many roads lead to the phagos (R) ome. Current opinion in cell biology 16, 422-428.

[4] Nakada-Tsukui, K., Okada, H., Mitra, B.N. and Nozaki, T. (2009). Phosphatidylinositol-phosphates mediate cytoskeletal reorganization during phagocytosis via a unique modular protein consisting of RhoGEF/DH and FYVE domains in the parasitic protozoon *Entamoeba* histolytica. Cellular Microbiology 11, 1471-1491.

[5] Labruyère, E. and Guillén, N. (2006). Host tissue invasion by *Entamoeba* histolytica is powered by motility and phagocytosis. Archives of medical research 37, 252-257.

[6] Lindmo, K. and Stenmark, H. (2006). Regulation of membrane traffic by phosphoinositide 3-kinases. Journal of cell science 119, 605-614.

[7] Stenmark, H., Aasland, R. and Driscoll, P.C. (2002). The phosphatidylinositol 3-phosphate-binding FYVE finger. Febs Letters 513, 77-84.

[8] Powell, R.R., Welter, B.H., Hwu, R., Bowersox, B., Attaway, C. and Temesvari, L.A. (2006). *Entamoeba* histolytica: FYVE-finger domains, phosphatidylinositol 3-

phosphate biosensors, associate with phagosomes but not fluid filled endosomes. Exp Parasitol 112, 221-31.

[9] Clark, C.G. et al. (2007). Structure and content of the *Entamoeba* histolytica genome. Adv Parasitol 65, 51-190.

[10] Osherov, N. and May, G.S. (2000). In vivo function of class I myosins. Cell Motility and the Cytoskeleton 47, 163-173.

[11] Voigt, H., Olivo, J.C., Sansonetti, P. and Guillen, N. (1999). Myosin IB from *Entamoeba* histolytica is involved in phagocytosis of human erythrocytes. Journal of Cell Science 112, 1191-1201.

[12] Otwinowski, Z. and Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276, 307-326.

[13] Winn, M.D. et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallographica Section D: Biological Crystallography 67, 235-242.

[14] McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr 40, 658-674.

[15] Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66, 486-501.

[16] Murshudov, G.N. et al. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallographica Section D: Biological Crystallography 67, 355-367.

[17] Afonine, P.V. et al. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallographica Section D 68, 352-367.

[18] Diamond, L.S., Harlow, D.R. and Cunnick, C.C. (1978). A new medium for the axenic cultivation of *Entamoeba* histolytica and other *Entamoeba*. Transactions of the Royal Society of Tropical Medicine and Hygiene 72, 431-432.

[19] Sahoo, N., Chakrabarty, P., Yadava, N., Bhattacharya, S. and Bhattacharya, A. (2000). Calcium binding protein of *Entamoeba* histolytica. Archives of medical research 31, S57-S59.

[20] Marion, S., Laurent, C. and Guillén, N. (2005). Signalization and cytoskeleton activity through myosin IB during the early steps of phagocytosis in *Entamoeba* histolytica: a proteomic approach. Cellular microbiology 7, 1504-1518.

[21] Ferraro, E., Peluso, D., Via, A., Ausiello, G. and Helmer-Citterich, M. (2007). SH3-Hunter: discovery of SH3 domain interaction sites in proteins. Nucleic acids research 35, W451-W454.

[22] Meza, I. and Clarke, M. (2004). Dynamics of endocytic traffic of *Entamoeba* histolytica revealed by confocal microscopy and flow cytometry. Cell Motility and the Cytoskeleton 59, 215-226.

[23] Fearnhead, N.S., Britton, M.P. and Bodmer, W.F. (2001). The abc of apc. Human molecular genetics 10, 721-733.

[24] Moseley, J.B., Bartolini, F., Okada, K., Wen, Y., Gundersen, G.G. and Goode, B.L. (2007). Regulated Binding of Adenomatous Polyposis Coli Protein to Actin. Journal of Biological Chemistry 282, 12661-12668.

4. CHAPTER III

Expression of *Eh*Myosin IB Motor domain in Insect cells

4.1 Introduction:

Motor domain is highly conserved in myosin superfamily and forms the basis of myosin classification [1-3]. It is believed that motor domains have co-evolved with the tail domain [4]. The mechanochemical properties of the motor domain of a particular myosin is decided by its cellular function, which is largely governed by the tail domain [4,5]. Detailed mechano-chemical studies regarding the motor function of myosins have largely been done on amoeboid myosins from *Acanthamoeba* and *Dictyostelium*. Motor domain of *E. histolytica* myosin I was found to be 54% identical and 71% similar to myosin IB of *Acanthamoeba* and 54% identical and 70% similar to myosin IB from *Dictyostelium* [6].

Vertebrate unconventional myosins I, Myo1a, Myo1b and Myo1e; and amoeboid myosin from *Acanthamoeba* spend the maximum duration of their ATPase cycle as weakly bound to actin filament [7-9]. Hence, Myosin I are shown to have low duty ratio. This suggests that it will be difficult for myosin I to singly carry cargoes over long distances and hence must cluster together for efficient transport of cargo to distant locations [10,11]. On the basis of single molecule studies, the motor domain of those myosins which are meant to carry heavy loads or cross-link actin filament for maintaining cortical tension during the transport, have been described to show biphasic ATP-induced dissociation from actin, slow ADP release rate and, two step power stroke along with ATP-induced conformational change [9,11,12]. However, *Dictyostelium* MyoE also shows biphasic ATP-induced actin dissociation but is not a slow motor.

The actin activated Mg^{2+}- ATPase activity of myosin I motors in *Acanthamoeba* and *Dictyostelium* are regulated by phosphorylation of a serine or threonine residue located at TEDS site, by PAK (p21-activated kinases). However, vertebrate Myo IE gets phosphorylated at glutamate residue located in place of serine. But, *Dictyostelium* myosins, MyoA and MyoC are not regulated by PAK mediated phosphorylation. These observations suggest that though Motor domains are highly conserved part of a myosin motor but still have structural differences to carry out their destined functions.

Motor domain of unconventional myosin IB is the most conserved but still, sequence comparison has revealed certain crucial differences from other amoeboid myosins. *Eh*Myosin IB has a glutamate, E^{348} residue located at the phosphorylation site similar to vertebrate myosin I, in place of serine or threonine residues seen in amoeboid myosin [6]. About 16 residues upstream of E^{348}, the NAIAK motif is found instead of consensus DALAK motif, typical of all unconventional myosin I studied so far. The ATP-binding site GESGAGKT and ATP dependent actin-binding site are located between amino acid residues 122-129, and 588–612 respectively in *Eh*Myosin IB [6]. These differences suggest that *Eh*Myosin IB is quite divergent from other amoeboid myosins. Hence, we tried to express motor domain constructs to study the mechanochemical properties and kinetics of the motor domain of *Eh*MyosinIB.

4.2 Materials and Methods:

4.2.1 Cloning of different constructs of *Eh*Myosin IB:

The complete sequence of each construct is given in appendix (A.5). The primers were designed as described in chapter1 (2.2.1).

S.No.	Construct Name	Amino acid position (start – end)	Primer Sequence
1.	BacMyFull	1 - 1049	**BACMIQ_FP:** 5'- CGCGGATCCAGTGTAGGATTCAAGAGAAATACTG-3' **RvBacMyFull :** 5'-CCGCTCGAGTTAAATTTCTTTGACATAGTTGTTAG-3'
2.	BacMyMIQ	1 - 766	**BACMIQ_FP:** 5'- CGCGGATCCAGTGTAGGATTCAAGAGAAATACTG-3' **RvMYMIQBAC:** 5'-CCGCTCGAGTTATAATGACCTCTTTCTATTTCTATAACC-3'
3.	BacMyMD	1 - 744	**BACMIQ_FP:** 5'- CGCGGATCCAGTGTAGGATTCAAGAGAAATACTG-3' **RvMy1MD:** 5'- CCG CTC GAG GTC CAT TTT TTC AAG CAT TTC- 3'

BacMyMD (2.1kb), BacMYMIQ (2.2kb), and BacMYFULL (3.1kb) genes were amplified by PCR using respective forward and reverse primers from the genomic DNA of *Entamoeba histolytica* strain HM1: IMSS. Reaction mix for PCR amplification was prepared as

mentioned in the appendix (A.4) and run in a thermocycler for 30 cycles with cycling parameters as mentioned in the appendix (A.4). The PCR products and the vector pFastBac-HTB were double digested with BamH1 and Xho1 restriction enzymes and gel purified. These digested and purified products were then ligated using T4DNA ligase and kept at 16°C for 16hrs. The ligated mixture was then transformed to *E.Coli* DH5α, plated on ampicillin containing LB agar plates and kept at 37°C for overnight. The colonies were screened for positive clones by colony PCR. The plasmid was isolated from positive colonies. The clone was further confirmed by double digestion of isolated plasmids and gene sequencing. (Detailed steps are given in appendix A.4).

4.2.2 Generation of Recombinant Bacmid by transformation into DH10Bac *E.coli* cells:

After the gene of interest was successfully cloned in pFastBac-HTB vector, recombinant bacmid was generated by site specific transposition of the expression cassette present in pFastBac clones into bacmid present in DH10Bac *E.coli* cells. DH10Bac competent cells were thawed on ice and 10ng of vector constructs were added to them. After 30 min of incubation on ice, a heat shock for 45 s at 42 °C was performed. The cells were then kept on ice for 5 mins followed by addition of 900 µl of LB medium and the tubes were incubated at 37 °C at 220 rpm for 4 h. The grown cells were diluted 10 times and 100 times with LB media, 100 µl of each dilution were spread on LB agar plates containing 50 µg/ml kanamycin, 7 µg/ml gentamicin, 10 µg/ml tetracycline, 100 µg/ml X-Gal, and 40 µg/ml IPTG to select for DH10Bac transformants. The plates were covered with aluminium foil and incubated for 48 hours at 37 °C. Consequently, few white colonies were picked for growth in 10 mL of LB supplemented with 50 µg/mL kanamycin, 7 µg/mL gentamicin and 10 µg/mL tetracycline for further bacmid amplification. Bacmid was isolated using the protocol as prescribed in Bac-to-Bac Baculovirus expression manual from Invitrogen.

Recombinant bacmid containing desired gene constructs were screened by PCR amplification of the gene of interest using M13 forward and reverse primers and isolated Bacmid DNA as templates. The PCR cycling parameters with Taq DNA polymerase are given below:

Initial Denaturation	95^{o}C	3 min	
Denaturation	95^{o}C	45 sec	
Annealing	55^{o}C	45 sec	30 cycles
Extension	72^{o}C	5 min	
Final Extension	72^{o}C	7 min	
Hold	4^{o}C	∞	

4.2.3 Insect cell culture and maintenance:

*Sf*9 cell line (derived from pupal ovarian tissue of fall army worm *Spodoptera frugiperda*) were obtained from cell line repository, NCCS (Pune, India). Cells were obtained as an adherent culture in media supplemented with 5% Fetal Bovine Serum. Cells were adapted to grow in serum free media EX-CELL® 420 (Sigma) as a suspension culture and the culture was maintained in the mid-log phase by regular passages and kept at 27°C, as per Invitrogen BEVS manual [13].

For normal maintenance cells were dislodged via tapping and rinsing and an aliquot was taken for viability determination and cell counting. 10µl cell suspension was mixed with the same amount of trypan blue solution (0.08 % (w/v) in PBS) and a hemocytometer was loaded with cells and counted. Cells were seeded at a minimum density of 5 X 10^5 cells/ml in a T75 flask or T250. Standard densities before passaging the cells were 1.2 x10^7 to 2.4 x10^7 cells, corresponding to 85-95 % confluency.

Cell preservation was done in mid log phase of cell growth (3-4 x10^6 cells /ml) and a viability >95 %. Cells were harvested and centrifuged at 600x g for 8 min and resuspended to a final concentration of 1x10^7 cells /ml in 7.5 % (v/v) DMSO, 46.25 % conditioned and 46.25 % fresh medium. Aliquots were immediately frozen with

sequential freezing at 4°C, then -20°C followed by storage at -80 °C and finally in liquid nitrogen.

4.2.4 Transfection and recombinant Baculoviral stock generation:

Log phase *Sf*9 cells with a viability of >95 % were seeded into a 6-well plate (9 x10^6 cells/well in 1 ml) in the antibiotic free medium. For transfection 1 µg bacmid DNA (1-2 µl) was prepared in 98 µl medium in one tube and in a second tube 2 µl, Escort IV transfection reagent (Sigma) was mixed with 98 µl medium. These two premixes were combined and incubated for 15-30 min at RT. As a control a MOCK transfection, without the addition of any bacmid DNA, was performed in parallel. This mixture was added dropwise to the cells and incubated for 5-6 h at 27 °C before the medium was exchanged for further cultivation. P1 Virus stocks were harvested in the late infection stages 48-72 h post transfection/infection, harvested supernatant was cleared from cells debris via centrifugation for 5 min at 500 x g; virus stocks were stored protected from light at 4 °C. P2 was generated to higher viral titre of the stocks by adding 300 µl virus stock to a well containing 1 x10^6 cells/ml. Preliminary expression experiments were performed until P3, using 200 µl to 300 µl of P2 recombinant virus stock and the infected cells were harvested for expression analysis by western blotting using anti-His antibody.

4.2.5 *Sf9* cell lysate preparation for western blotting:

72 hours post infection by P3 generation viruses, cells of a 6-well plate were harvested via rinsing in 1 mL PBS, centrifuged at 500x g at 4 °C and the supernatant was discarded. The cells pellets were stored at - 20 °C and resuspended in RIPA buffer (Appendix A.3). Cells were disrupted via mechanical shearing through a needle and syringes, cracked cells were further incubated for 30 mins on ice. Cell debris was separated via centrifugation (20,000 x g, 15 min at 4 °C). The protein amount in the supernatant was assayed via Bradford [14] and 20 µg total protein content was applied to a 12 % SDS-PAGE for further western blot analysis.

4.2.6 Viral Plaque Assay:

2 ml of Mid-log phase *Sf*9 cell suspension (5 X 10^5 cells/ml) was plated into each well of two 6-well plates. Cells were allowed to adhere for 1 hour at room temperature. Eight serial dilutions of baculoviral stock were prepared (10^{-1} ,10^{-2} , 10^{-3} ,10^{-4} ,10^{-5} ,10^{-6} ,10^{-7} ,10^{-8}) and stocks with concentration 10^{-4} to 10^{-8} were used in the assay. After the cells were adhered to the bottom and looked 50% confluent, 1ml of the medium from each well was replaced by appropriate virus dilution. Again, the cells were incubated for 1 hr at room temperature. Only medium (without virus) was used as a negative control. Media from each well was replaced by 2ml of plaquing medium (appendix A.3) quickly before the cells start to desiccate. Agarose overlay was allowed to solidify at room temperature, later the plates were shifted to 27°C incubator for 7-10 days until plaques were visible and countable. The Agarose overlay was stained for 3 hr at 37°C by adding 0.1 volume of MTT solution (5 mg/ml in PBS) (Sigma, M2128). The monolayer appeared blue-black and plaques were visible as clear areas.

$$Titre\ (pfu/ml) = number\ of\ plaques\ X\ dilution\ factor\ X\ \frac{1}{(ml\ of\ inoculum\ per\ well)}$$

4.2.7 Overexpression of BacMyMIQ construct in *Sf*9 cells:

For expression analysis in a 6 well plate, cells were infected when they were in the mid-logarithmic phase of growth at a density of 1 x 10^6 to 2 x 10^6 cells/ml. Cells were infected with P3 viral stock of BacMyMIQ at different MOIs. MOI is defined as the multiplicity of infection is the ratio of infectious virus particles added per cell. The cells were scrapped after 72 hrs and the lysate was prepared as mentioned in 4.2.5 and expression was analyzed by western blotting using anti-his antibody against the hexa-histidine tag present in the recombinant protein.

Bulk expression was checked in 500ml culture harboring a density of 2 x 10^6 cells/ml, infected with P3 stock at a suitable MOI. Cells were pelleted after 72 hrs in a Beckman centrifuge at 4000 rpm and stored at -80°C.

4.2.8 Purification of BacMyMIQ by Ni-NTA chromatography:

Pellet was resuspended in 30ml lysis buffer (50mM Tris pH 7.4, 300mM KCl, 4mM $MgCl_2$, 2mM ATP, 10mM β-Me, 4mM imidazole pH 8.0, 0.2mM EGTA, 1% Triton-X, protease inhibitor cocktail). Lysis was done by 3-4 freeze thaw cycle followed by sonication and then, ultracentrifugation at 138,000×*g* for 35 min. The clear supernatant obtained was loaded onto Ni-NTA column (Sigma) prepacked with Ni-sepharose beads (GE healthcare) pre-equilibrated with equilibration buffer (50mM Tris pH 7.4, 600mM KCl, 4mM $MgCl_2$, 0.5mM ATP, 7mM β-Me, 20mM imidazole pH 8.0, 0.1mM EGTA, 0.2% Triton-X). The MyMIQ protein was eluted with elution buffer (50 mM Tris 7.4, 300 mM KCl, 100 mM imidazole, 4mM $MgCl_2$, 0.1mM EGTA) after washing with wash buffer (50 mM Tris 7.4, 300 mM KCl, 40 mM imidazole, 4mM $MgCl_2$, 0.1mM EGTA) [15]. The eluted fractions were checked for purity on 10% SDS PAGE prepared as mentioned in appendix (A.3).

4.3 Results:

4.3.1 Cloning of *Eh*Myosin IB constructs:

Three *Eh*MyosinIB constructs; motor domain BacMyMD (2.1kb), motor domain with IQ motif BacMyMIQ (2.2kb) and full length BacMyFull (3.1kb) constructs were successfully cloned between BamH1 and Xho1 sites in pFastBac-HTB vector for expression in *Sf*9 insect cells. Clones were confirmed by double digestion with respective restriction enzymes.

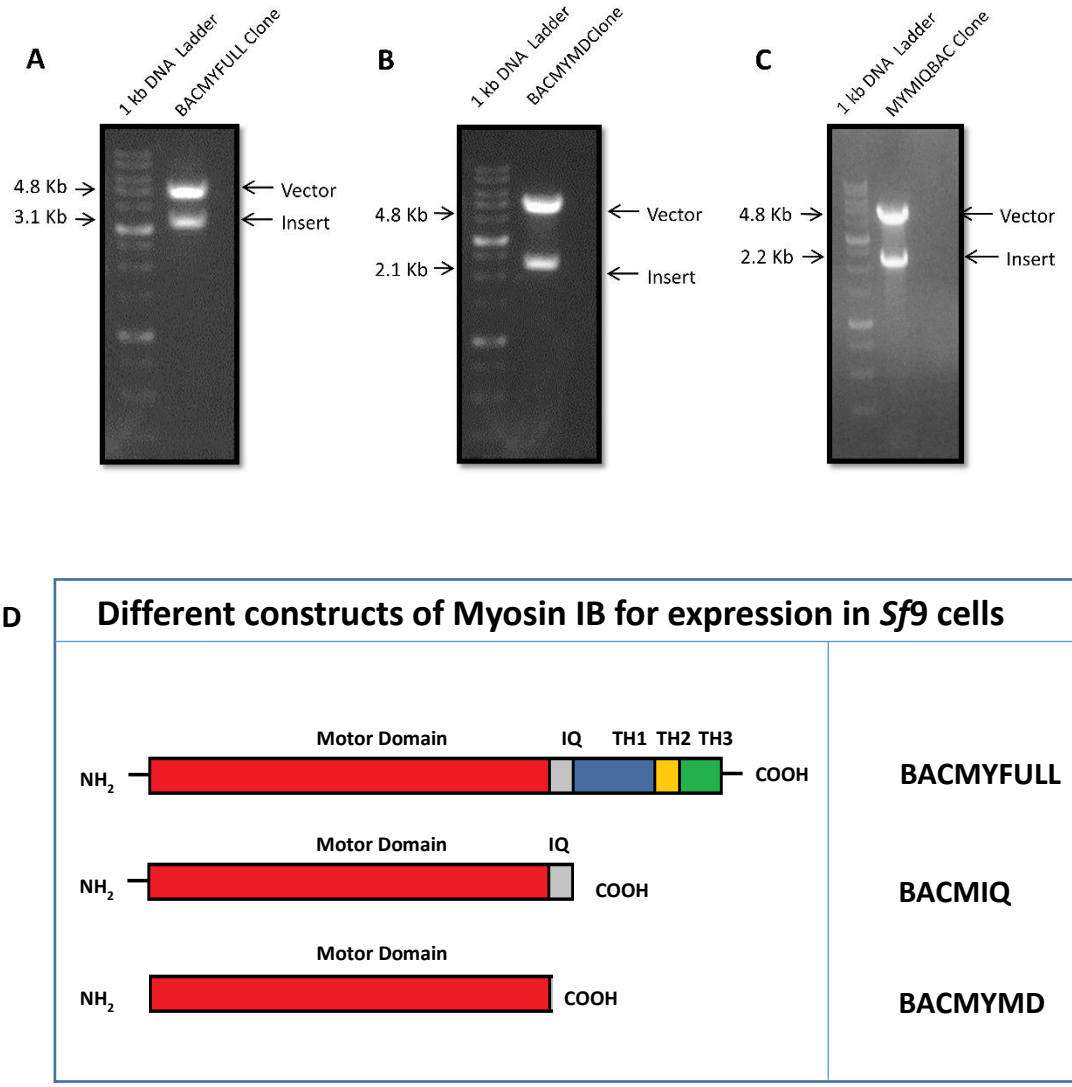


Figure 4.3-I: Cloning of myosin IB constructs for expression in insect cells.

Agarose gel showing restriction digestion with BamH1 and Xho1 of clones, (A) BacMyFull corresponding to 3.1 kb and vector pFasatBac-HTB at 4.8 kb, (B) BacMyMD

corresponding to 2.1 kb and vector pFasatBac-HTB at 4.8 kb, and (C) BacMyMIQ corresponding to 2.2 kb and vector pFasatBac-HTB at 4.8 kb. (D) Schematic representation of the three motor domain constructs of *Eh*Myosin IB cloned for expression in insect cells (*Sf*9).

4.3.2 Generation of recombinant Bacmid DNA:

The expression cassette of BacMyMD, BacMyMIQ and BacMyFull was successfully inserted into bacmid DNA through transposition. Insertion of the mini-Tn7 of pFastBacHT into the mini-attTn7 attachment site on the bacmid disrupt the expression of the LacZα peptide, so colonies containing the recombinant bacmid were white. These white colonies were screened by PCR with M13 promoter primers which can be seen as a total amplicon of size (2430 kb + gene of interest) on agarose gel. All the constructs were successfully inserted into bacmid DNA which were then transfected in *Sf*9 cells. All the clones were also confirmed by sequencing.

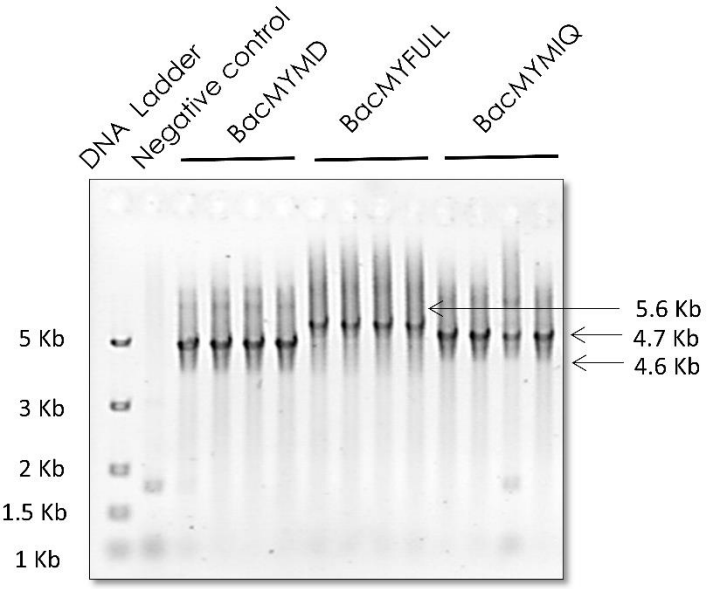


Figure 4.3-II: Bacmid confirmation for different Myosin IB constructs.

Agarose gel showing gene amplification of recombinant bacmid constructs; BacMyMD, BacMyFull, and BacMyMIQ after colony PCR using M13 promoter primers. All the constructs showed amplification at required sizes: BacMyMD at 4.6kb, BacMyFull at 5.6kb and BacMyMIQ at 4.7kb.

4.3.3 Generation of high titre recombinant baculoviruses:

Sf9 cell were transfected with recombinant bacmid DNA. The generated baculoviruses were collected and labelled as P1 generation viral stock. Subsequent infection of mid-log phase cells with one generation viral stock resulted in next generation higher titre stock. Viruses were amplified for three generations resulting in high titre, P3 stock. Titre was measured by plaque assay.P3 stock of BacMyFull and BacMyMIQ resulted in a titre of 1 x 10^{-8} pfu/ml while the titre of BacMyMD P3 stock was calculated to be 1.8 X 10^{7} pfu/ml.

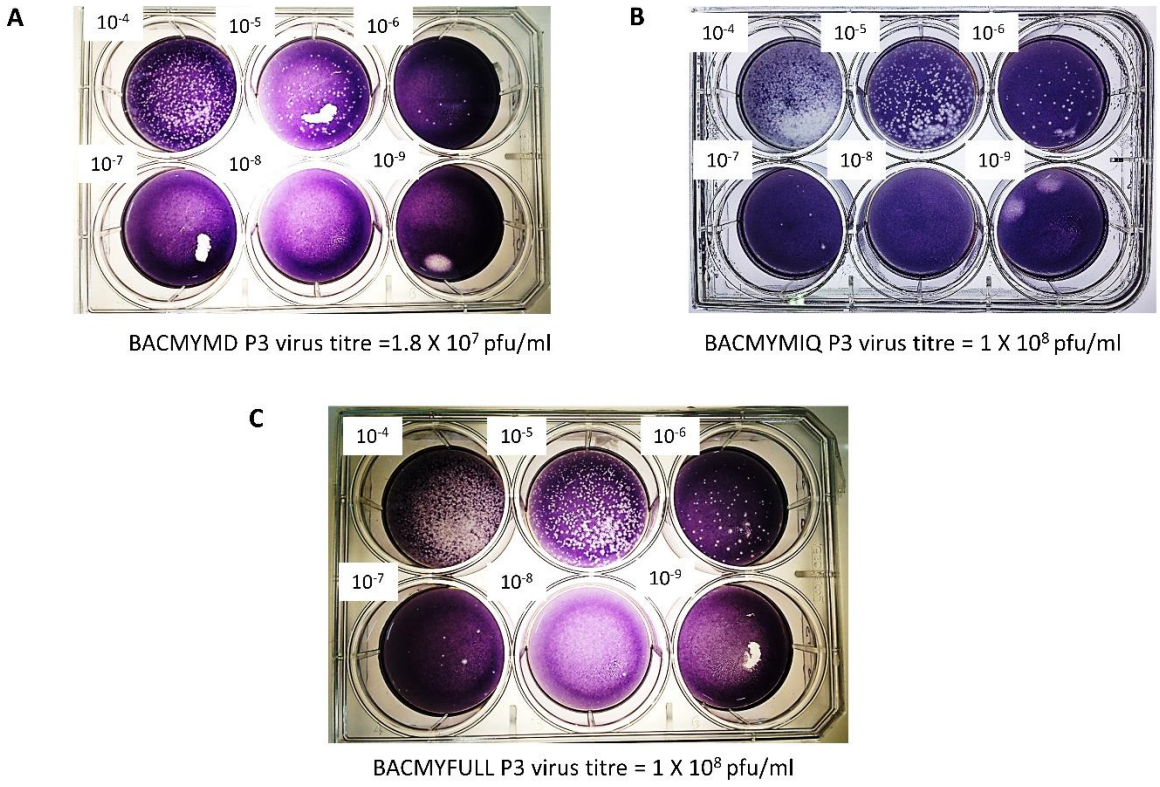


Figure 4.3-III: Baculoviral titre measurement by Viral Plaque Assay.

Viral Plaque Assay to determine the titre of generated baculovirus. The titre of recombinant baculoviruses of motor domain construct, BacMyMD was 1.8 x10^{7}, BacMyMIQ was 1 x10^{8}, and BacMyFull was 1 x10^{8}.

4.3.4 Expression optimization of Motor domain construct, BacMyMIQ.

Small scale expression of all three constructs was done by infecting the mid-log phase *Sf*9 cells with P1 stocks of each construct. Cell were checked for expression after 72 hrs. All the three constructs showed bands at required molecular weights. Expression conditions were optimized for BacMyMIQ.

BacMyMIQ was expressed maximally when the cells were infected at an MOI of 5 and were allowed to grow for 72 hrs. Expression increased with increasing MOI. In the western blot band was detected at around 85 kDa.

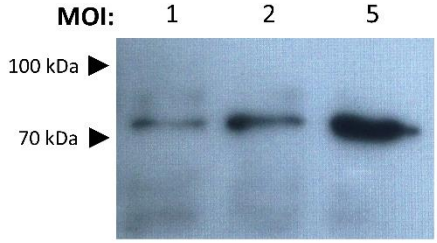


Figure 4.3-IV: Expression analysis of overexpressed BacMyMIQ by western blotting.

Expression analysis of BacMyMIQ (motor domain + IQ motif) construct by infection of insect cells at different MOI with P1 viruses. Expression was found to be increased with increasing MOI. Maximal expression was seen at an MOI of 5.

4.3.5 Purification of Motor domain of *Eh*Myosin IB:

Presence of N-terminal histidine tag facilitated the purification of BacMyMIQ protein using Ni-NTA chromatography as mentioned in materials and methods 4.2.8. Very small amount of protein was eluted at 100mM imidazole concentration with few contaminating bands of impurities. Presence of protein was confirmed by MALDI-TOF analysis of band around 88kDa.

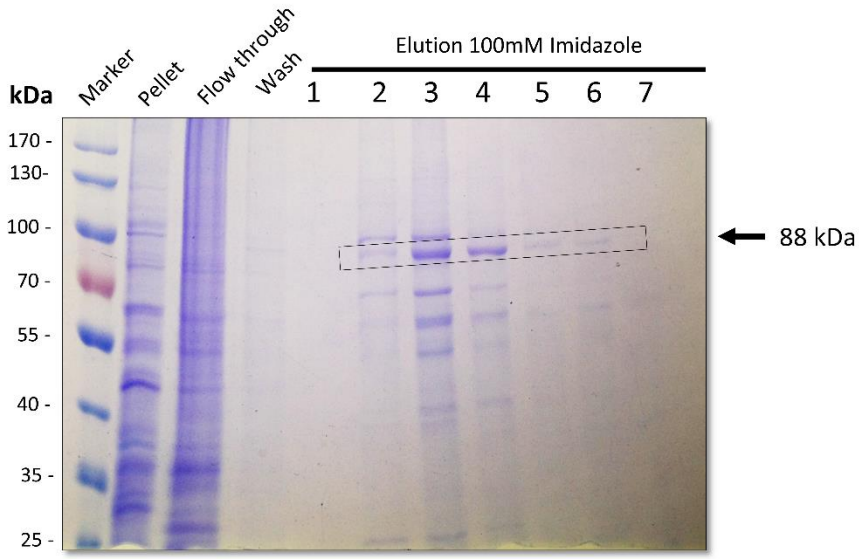


Figure 4.3-V: 8% SDS-PAGE showing overexpressed purified Motor domain +IQ motif of EhMyosin IB.

4.4 Conclusion and Future directions:

Recombinant baculoviruses of motor domain constructs and full length *Eh*Myosin IB were successfully generated. Expression of one of the construct with motor domain along with IQ motif, BacMyMIQ was checked in small scale culture and purified with half litre culture. Expression of protein was confirmed by MALDI-TOF analysis.

Future directions include purification of BacMyMIQ from bulk culture for kinetic studies of motor domain and structure studies with motor domain.

4.5 References:

[1] Cheney, M.S.M.a.R.E. (1995). Unconventional Myosins. Annual Review of Cell and Developmental Biology 11, 633-675.

[2] Cheney, R.E., Riley, M.A. and Mooseker, M.S. (1993). Phylogenetic analysis of the myosin superfamily. Cytoskeleton 24, 215-223.

[3] Goodson, H.V. and Spudich, J.A. (1993). Molecular evolution of the myosin family: relationships derived from comparisons of amino acid sequences. Proceedings of the National Academy of Sciences 90, 659-663.

[4] Korn, E.D. (2000). Coevolution of head, neck, and tail domains of myosin heavy chains. Proceedings of the National Academy of Sciences 97, 12559-12564.

[5] O'Connell, C.B., Tyska, M.J. and Mooseker, M.S. (2007). Myosin at work: motor adaptations for a variety of cellular functions. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1773, 615-630.

[6] Vargas, M., Voigt, H., Sansonetti, P. and Guillen, N. (1997). Molecular characterization of myosin IB from the lower eukaryote *Entamoeba* histolytica, a human parasite. Mol Biochem Parasitol 86, 61-73.

[7] Jontes, J.D., Milligan, R.A., Pollard, T.D. and Ostap, E.M. (1997). Kinetic characterization of brush border myosin-I ATPase. Proceedings of the National Academy of Sciences 94, 14332-14337.

[8] Lewis, J.H., Lin, T., Hokanson, D.E. and Ostap, E.M. (2006). Temperature dependence of nucleotide association and kinetic characterization of myo1b. Biochemistry 45, 11589-11597.

[9] El Mezgueldi, M., Tang, N., Rosenfeld, S.S. and Ostap, E.M. (2002). The kinetic mechanism of Myo1e (human myosin-IC). Journal of biological chemistry 277, 21514-21521.

[10] Ostap, E.M. and Pollard, T.D. (1996). Overlapping functions of myosin-I isoforms? J Cell Biol 133, 221-4.

[11] Coluccio, L.M. (1997). Myosin I. American Journal of Physiology - Cell Physiology 273, C347-C359.

[12] Geeves, M.A., Perreault-Micale, C. and Coluccio, L.M. (2000). Kinetic analyses of a truncated mammalian myosin I suggest a novel isomerization event preceding nucleotide binding. Journal of Biological Chemistry 275, 21624-21630.

[13] Manual, I. Guide to Baculovirus Expression Vector Systems (BEVS) and Insect Cell Culture Techniques.

[14] Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry 72, 248-254.

[15] Heissler, S.M. and Manstein, D.J. (2012). Functional characterization of the human myosin-7a motor domain. Cellular and Molecular Life Sciences 69, 299-311.

5. Summary

Unconventional Myosin I forms the largest class of myosin superfamily and has been associated with several membrane associated processes like vesicular transport, membrane ruffling, formation of filpodia, etc. Phagocytosis is one of the crucial event associated with the pathogenesis of *Entamoeba histolytica,* a gastric pathogen that causes Amoebiasis disease in humans. E. histolytica has only two myosins: Myosin II and Myosin IB in its genome. Unconventional myosin IB has been found to be involved in phagocytosis and is the only unconventional myosin present in *Entamoeba* unlike other closely related organisms like *Dictyostelium* and *Acanthamoeba*, which have several isoforms of myosin I which are associated with different functions. On the basis of sequence similarity, EhMyosinIB motor domain as well as the tail domain has been found to be different from long tailed myosin I of other amoeba.

Phagocytic event in *E.histolytica* largely depends on a regulated interplay between several signalling molecules and cytoskeletal proteins. Myosin IB in *E. histolytica* has been found to be actively involved with the actin cytoskeleton in phagocytic cup formation but the role played by the c-terminal SH3 domain of EhMyosin IB during phagocytosis has not been fully understood and also no known interacting partners of the same has been reported till now. In our attempt to study the role of C-terminal SH3 domain of *EhMyosin* IB, we proceeded with the study of structural details of EhMyosin IB SH3 domain named as EhMySH3 in the study.

Chapter I deals with the determination of the crystal structure of *Eh*MySH3 in two forms: Native form (EhMySH3) and PEG bound form (PEG-EhMySH3). In the PEG-bound *Eh*MySH3 structure, PEG appeared to have mimicked physiological ligand by occupying the conserved ligand-binding pockets on the *Eh*MySH3 surface. These bound PEG molecules highlighted the importance of the *Eh*MySH3 domain P4 pocket, located between residue W47 and the well-conserved residue W36. The conservation of tryptophan residue corresponding to W47 in amoebozoa myosin I, also suggested the P4 pocket to be crucial for the specificity and affinity

of *Eh*MySH3 for its ligands. Structural comparisons made with the SH3 domains of various myosins revealed *Eh*MySH3 to be similar to the SH3 domain of myosin IB from *Acanthamoeba*, except in the β2 sheet and n-Src region. However, according to the charge distribution, on the ligand-binding surface as well as adjacent C-terminal region in '+' peptide-binding orientation, *Eh*MySH3 shared structural similarity with the rat β-pix SH3 domain bound to the AIP4 peptide. These analyses led us to recognise two other regions of negative surface charge on *Eh*MySh3 surface, i.e., P3' and P5, which could be essential for its interaction with ligands. The arrangement of chain B of PEG-*Eh*MySH3 with its symmetry molecule revealed that a pseudo-symmetric ligand may induce the dimerization of *Eh*MySH3 and hence lead to the formation of a heterotrimeric protein-ligand-protein complex. This led us to conclude that SH3 of EhMyosinIB may be responsible for the formation of a multimeric-protein complex with various proteins along with actin which are involved in phagocytic process.

On the basis of structural characterization of the c-terminal SH3 domain of myosin IB i.e. *Eh*MySH3 and the kind of ligand it will prefer (Chapter 1), we next proceeded with identification of probable ligands of EhMyosin IB Sh3 domain in Chapter II. On the basis of a previously published proteomic data, a few proteins were shortlisted which could be a probable binding partner of *Eh*MySH3. On the basis of SPR studies, a peptide from *Eh*FP10, an FYVE domain containing GEF was found to show the best binding response. Peptide P2 was also successfully co-crystallized with the *Eh*MySH3 domain in $P2_12_12_1$ space group. Structural studies of *Eh*MySH3-P2 peptide complex revealed that P2 leads to clustering of the *Eh*MySH3 domain, and is shared by two molecules of *Eh*MySh3 as has been concluded in our previous results. Further analysis led us to predict that *Eh*MySH3 shows a preference for class II polyproline ligands and preferably binds to P2 in a class II manner.

Since *Eh*FP10 has not been characterized in *E. histolytica*, localization studies of *Eh*FP10 were done via time-lapse imaging of GFP-tagged *Eh*FP10 expressing *E. histolytica* trophozoites. *Eh*FP10 was found to suddenly appear near the membrane at the site of pinocytosis in normal resting trophozoites and in phagocytosis, in cells induced for erythrophagocytosis respectively. *Eh*FP10 was

present from initiation till the closure of the cup in both phagocytosis and pinocytosis. *Eh*FP10 was associated with the membrane encircling the newly internalized pinosomes and disappears after a while, but not in newly formed phagosomes. This interesting observation states that there are differences in the regulation of both the endocytic processes.

*Eh*Myosin IB was found to colocalize with *Eh*FP10 at all stages during pinocytosis and phagocytosis. P2 peptide resided in the c-terminal APC domain of *Eh*FP10. Hence, the interaction was further validated by in vitro pull down assay of APC domain of *Eh*FP10 with bound EhMySh3. APC domain of *Eh*FP10 was seen to bind with filamentous actin and form thicker actin bundles as known for classical APC basic domains. Interaction of *Eh*MySh3 with APC domain of *Eh*FP10 was seen to inhibit the bundling activity by *Eh*FP10 at high *Eh*MySH3: *Eh*FP10 ratios while as the concentration of APC increased, EhFP10 could overcome the inhibition and bundle the actin filaments efficiently. This observation gave the functional relevance of *Eh*MySH3 interaction with *Eh*FP10 protein during phagocytosis and pinocytosis in *E. histolytica*.

The known effectors of the SH3 domain of unconventional myosin I have been found to regulate the actin dynamics directly or indirectly via arp2/3 complex. Several GTPases act upon these effectors and adaptor protein like WASP and other WASP related proteins like Vrp1, Bee1p, and Las17p which leads to actin polymerization at the site of cup formation. As *E. histolytica* lacks WASP/SCAR protein, this highly expressing, unique, modular FYVE family Rho GTPase Exchange Factor, *Eh*FP10 with a c-terminal tail that can regulate actin dynamics is of great advantage to the pathogen. Our findings reveal a new mode of regulation of endocytic processes in highly motile phagocytosing gastric pathogen like *E. histolytica.*

To get the complete overview of how EhMySH3 and EhFP10 regulate eachother during the endocytic processes in *Entamoeba* and kinetics of the EhMyosin IB motor domain during movement on actin filament, we tried to overexpress full length EhMyosinIB and different motor domain constructs in insect cells (Chapter

3). We could successfully express the myosin IB motor domain + IQ motif construct and purify the protein from small scale culture. Further work remains to be done.

*Eh*Myosin IB is different from other amoeboid myosins and lies in between metazoan and amoeboid myosins. Hence, the results of the current study about the SH3 domain of myosin I has a larger implication and are expected to be highly useful for understanding and predicting possible binding partners of unconventional myosin IB of *E. histolytica*.

6. Appendix

A.1 Sources of materials:

Escherichia coli strains DH5α and BL21 (DE3) were obtained from Bethesda Research Labs (B.R.L, U.S.A.), and used for recombinant protein over-expression. *Sf9* cells were obtained from NCCS cell line repository, (Pune, India). E. histolytica strain HM-1-IMSS were obtained from our collaborator Prof.Alok Bhattacharya's (Jawaharlal Nehru University, New Delhi) lab. All PCR amplifications were performed in any of these thermocyclers: Techne Touchgene gradient, Eppendorf thermal cycler or Infinigen thermal cycler. Expression vectors pET-28b, pGEX-6p2 was obtained from Novagen and pFastBac from Invitrogen. Molecular biological reagents, media components and other chemicals were purchased from Fermentas (now Thermo Fisher Scientific), New England Biolabs (NEB, UK), Sigma-Aldrich (USA), Real Biotech Corporation (Taiwan), Merck (Germany), HiMedia (India), Difco, Qualigens (India) and Sisco Research Laboratories (India). PCR purification kits, Gel extraction kits and plasmid purification kits were purchased from Qiagen and MDI. Ni-NTA resin was purchased from GE healthcare. Chromatographic columns were obtained from Sigma-Aldrich (U.S.A.), Centricons were obtained from Millipore. Crystallization reagents were obtained from Hampton (U.S.A.), Molecular Dimensions (U.K) and Sigma-Aldrich (U.S.A.).

A.2 Organisms and growth conditions:

E. coli DH5α and BL21 (DE3) cells from an agar stab or frozen glycerol stock were first streaked on an LB plate (containing the appropriate antibiotic wherever necessary) and allowed to grow overnight at 37°C. Liquid cultures in LB medium were initiated from a single colony and were grown with constant shaking at 220 rpm at 37°C. The cells 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.

Trophozoites of E. histolytica were grown axenically in TYI-S-33 medium as described by Diamond et al., 1978. The cells were maintained and grown in TYI-33 medium complemented with 15% adult bovine serum, 1X Diamond's vitamin mix and antibiotic (125 µl of 250 U/ml Benzyl Penicillin and 0.25 mg/ml Streptomycin per 90 ml of medium).

A.3 Buffers and Solutions:

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

1) Luria Broth (100ml):

Yeast Extract	0.5 g
NaCl	0.5 g
Tryptone	1g
Water	100ml

Solution was mixed well and autoclaved. For Solid media, 2% agar was added.

2) TYI-33 medium composition

Potassium phosphate, dibasic	1.0 g
Potassium phosphate, monobasic	0.6 g
Biosate peptone	30.0 g
Dextrose	10.0 g
Sodium chloride	2.0 g
L-Cysteine hydrochloride	1.0 g

Ascorbic acid	1.0 g
Ferric ammonium citrate	22.8 mg

To these components 700 ml of double distilled water was added and pH was adjusted to 6.8 using 5 N NaOH. The volume was made up to 900 ml (10 units) and filtered using Whatman Filter paper #1. This was then aliquoted and autoclaved. Incomplete medium was stored at -20^{o}C.

Medium was completed by adding 15% heat inactivated adult bovine serum, 1X Diamond's vitamin mix and 125 µl of antibiotic (250 U/ml Benzyl Penicillin and 0.25 mg/ml Streptomycin). This completed media (TYI-S-33) was used to grow *E. histolytica* (Diamond et al.1978).

3) **Plaquing Media:** Combine 30 ml of Sf-900 Medium (1.3X) and 10ml of the melted 4% Agarose Gel in the empty 100 ml bottle and mix gently.

4) Antibiotics Stocks:

Name of the Antibiotic	**Stock Concentration**
Kanamycin	100mg/ml
Ampicillin	100mg/ml
G418	10 mg/ml
Hygromycin	10 mg/ml

The solutions were sterile filtered and stored at -20°C.

5) SOLUTION FOR AGAROSE GEL ELECTROPHORESIS:

1) 50X TAE:

Tris-base	242 g
Glacial Acetic acid	57.1 ml
0.5 M EDTA (pH 8.0)	100 ml

The final volume was made up to 1 litre.

2) ***Ethidium bromide:*** 10mg/ml; light sensitive and prepared in autoclaved water, wrapped in aluminum foil and kept in dark at RT.

3) 6X DNA loading buffer

Bromophenol blue	0.25% (W/V)
Xylene cyanol FF	0.25%
Glycerol	30% (V/V)

Preparation of IPTG (1M Stock concentration): Required amount of IPTG was weighed and dissolved in milliQ water, which was sterilized by filtration using filter of 0.22 µm pore size.

4) SDS PAGE COMPONENTS:

a) ***Acrylamide solution (30%):*** 29.2 gm acrylamide and 0.8 gm bis-acrylamide were dissolved in about 60 ml water and kept at RT for some time until acrylamide dissolved, then volume was made up to 100 ml and filtered by Whatman no 3 filter paper. Reagent was prepared and stored in dark colored bottle at 4^0C.

b) Separating and stacking gel:

A. Components	**% Separating gel (5ml)**					**5% stacking gel (2ml)**
	6	**8**	**10**	**12**	**15**	
H_2O	2.6	2.3	1.9	1.6	1.1	1.4
30% acrylamide	1.0	1.3	1.7	2.0	2.5	0.33
1.5 M Tris (pH 8.8)	1.3	1.3	1.3	1.3	1.3	-----
1 M Tris-Cl pH (6.8)	-----					0.25
10% SDS	0.05					0.02
10% APS*	0.05					0.02

TEMED	0.04	0.002

**APS (Ammonium per sulphate): (10%), freshly prepared (100mg/ml in water; light sensitive)*

c) 5X-SDS sample dye:

Tris-Cl (pH 6.8)	225 mM
SDS	5% W/V
Bromophenol blue	0.05% W/V
Glycerol	50% V/V
DTT or β-Mercaptoethanol	250 mM

The final volume was made up to 10 ml.

d) 10X SDS running buffer (for 1 L) pH 8.3

Glycine	144.2 g
SDS	10 g
Tris	30.3 g

e) Staining solution (Coomassie brilliant blue R-250, for 100 ml)

Coomassie brilliant blue R-250	0.25 gm
Glacial acetic acid	10 ml
Methanol	50 ml
Water	40 ml

f) De-staining solution (for 100 ml)

Methanol	50 ml
Glacial acetic acid	10 ml
Water	40 l

5) PLASMID ISOLATION SOLUTIONS

Solution I (GTE), 100 ml (Re-suspension solution); Stored at 4°C

50 mM Glucose	1.711g
10 mM Tris pH-8	2.5 ml
pH-10 mM EDTA pH 8	2 ml

Solution II, 10 ml (Freshly made)

10 N NaOH	0.2 ml (0.2M NaOH)
10% SDS	1ml (1% SDS)

Solution III, 100ml (Neutralization solution); Stored at 4°C

5M Potassium acetate	60 ml
Glacial acetic acid	11.5 ml
Water	28.5 ml

6) G-Buffer

Tris(pH 7.5)	20mM
DTT	0.5mM
ATP	0.2mM
$CaCl_2$	0.1mM
NaN_3	0.1mM

7) F-Buffer

Tris(pH 7.5)	20mM

DTT	0.5mM
ATP	0.2mM
$MgCl_2$	2mM
KCl	50mM

8) RIPA Buffer:

Component	Concentration
Tris HCl, pH 7.5	50 mM
NaCl	150 mM
NP-40 (4-Nonylphenyl-polyethylene glycol)	1 % (v/v)
DOC (sodium deoxycholate)	0.5 % (w/v)
SDS	0.1 % (w/v)
ß-Glycerolphosphate	2 mM
Na3VO4	1 mM
PMSF (Phenylmethanesulfonyl fluoride)	0.4 mM
EGTA	1 mM
NaF	1 mM
Protease-Inhibitor (Roche, Germany)	1 tablet 50 mL-1

9) 1X Transfer Buffer (1L):

Glycine	14.4g
10% SDS	5ml
Tris	3.03 g
Methanol	200ml

A.4: Cloning protocols:

1) PCR cocktail:

PCR Reaction mixture	***Vol. (µL)***
DNA (genomic) (100ng/µL)	*1.0*
Buffer (5X HF Buffer)	*10.0*
$MgCl_2$ (50mM)	*1.0*
dNTPs (10mM)	*1.0*
Fp (10µM)	*2.5*
Rp (10µM)	*2.5*
Phusion enzyme(2.0 units/µL)	*0.5*
Double distilled water	*31.5*
Total volume	*50*

2) PCR cycling protocol:

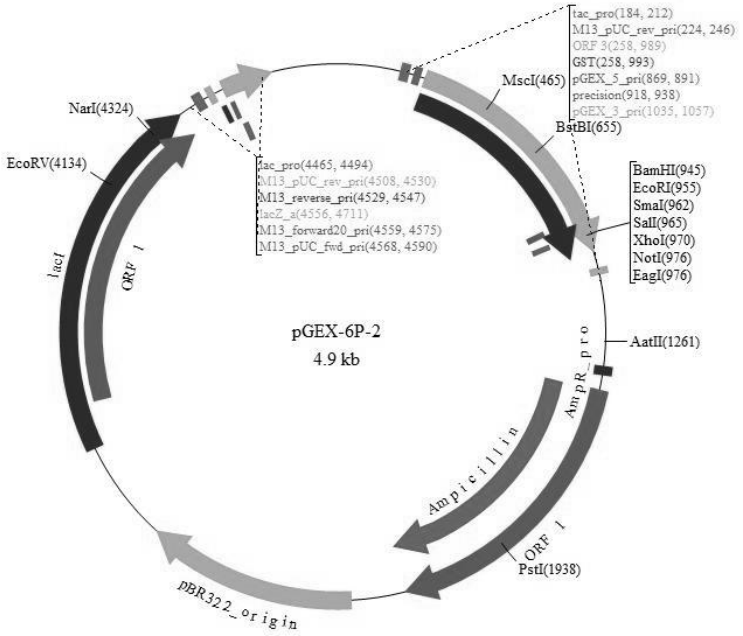
Initial Denaturation	98^{o}C	30 sec	
Denaturation	98^{o}C	15 sec	30 cycles
Annealing	55-65^{o}C	30 sec	
Extension	72^{o}C	30 sec	
Final Extension	72^{o}C	10 min	
Hold	4^{o}C	∞	

*Annealing temperature was optimized for each construct.

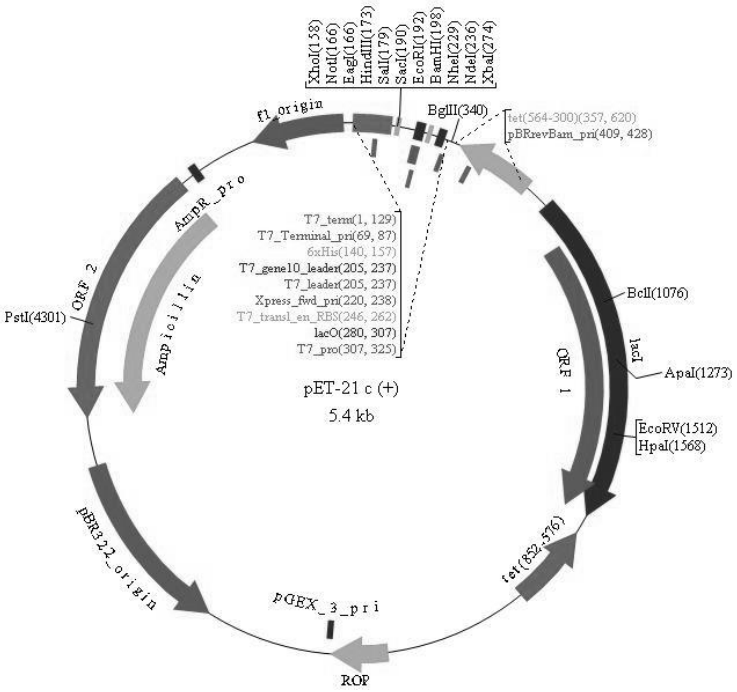
3) Restriction Digestion:

Components	Volume
DNA	5 (1µg/ µl)
10X Buffer	2µl
Restriction enzyme1	1µl
Restriction enzyme 2	1µl
MQ	11µl
total	20

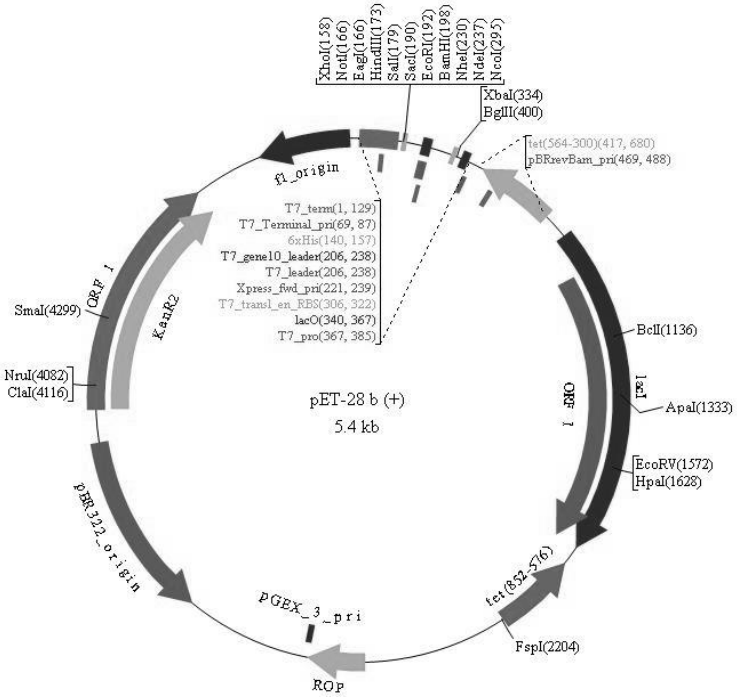
A.5: Plasmid Vectors used:



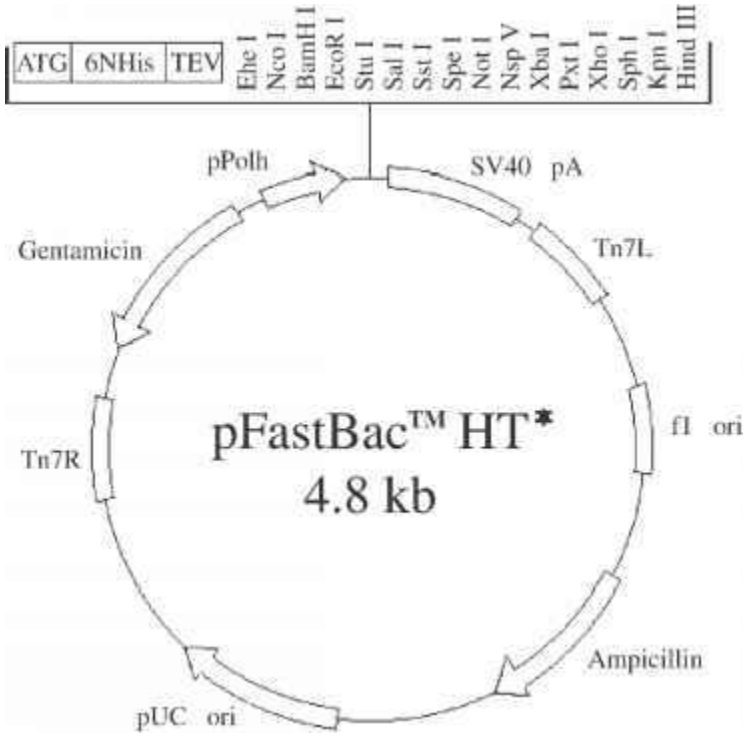
Source: http://www.genomex.com/vector_maps/pGEX-6P-2_map.pdf



Source: http://www.genomex.com/vector_maps/pET21_map.pdf



Source: http://www.genomex.com/vector_maps/pET28_map.pdf



http://www.flandershealth.us/therapeutic-proteins/images/9405_24_21.jpg

A.6: Protein sequences used under study:

Unconventional myosin IB *[Entamoeba histolytica* HM-1:IMSS]

>XP_654280.1 unconventional myosin IB [*Entamoeba* histolytica HM-1:IMSS]

MSVGFKRNTEKERENEIAAFKSKKGFVQKGVDDLVMLPKIDENEIVDNLKKRYNYDAIYTNIGPVLIVIN
PFKDLGLTTDEYVHLYKGKFRQELPPHIFALSEETYRNMKNEKCNQCCIISGESGAGKTVAAKQIMQYIA
AVSGGNEKVVYVKNVILDSNPLLEAFGNAKTLRNNNSSRFGKYFEINFDDHADPVGGTITNYLLEKSRVC
TQQTGERNYHIFYQLLAGAPENYYNDFYLTTPDYFVYTNQSNCMVVDGIDDKKDYADCVKAMNTIGISAD
EQYWIFQLVAAILHLGNVYFTENSSGYAEITDKSALETAAYLLEVDPGQLEHVMLVRKLTTGVGARAEVF
DSPLTVEQAEATKNAIAKEIYDRIFTMLVQKVNIALQKHGIEHQCVIGILDIFGFEIFPVNGFEQFCINY
VNEKLQQYFIELTLKTEQEEYIAEGIKWTPIKYFNNKVVCDLIEGKNPPGMFSLLDDICATMHAQTEGAD
FKFLEKCCSVHSTNQHFIPSGETFTVQHYAGPVSYYCEGFCEKNKDTLFLECVECLQTSNNQLLYDLFPW
SAEQEAAKTGQKKRPTTAGFKIKTSANELMAALSQCQPHYIRCIKPNETKKPQDWDADRVKHQVKYLGLL
ENVKVRRAGFAYRAPFARFLQTYKKLSPKTWGTWGEWTGDAKEGCRLILEGLTNIEPGQWQFGTTKLFIR
LPESIFAIEEMLEKMDFDKAVEIQKAWKGYRNRKRALSQRAEAASILKGKKERRRESVDYTNNKWTADYI
NYENEIPLQEAMEGHMDELVNFDDWVMRLNKKGKIENRALLLTDKAIYIAQRQIKKKVLKYIVTFRGDLS

QISEVVMSTMQDNFMVIRVRGAYDSVLLCPHKTEFLMMLLENYRALMNGQSFPINFSDSISFRGKDEKKD
KLLVFKKDDGPKTLPVEEQKSFSKHVLVRTGLPKETDTTPGSYNSSVSSYAAPTASSSNNYSAPPMKSAP
PAPQVAAPPPPKPKLPQVKALYPYTAANDEELSFKVGDIITILEKDEGWWKGELNGQEGWIPNNYVKEI

Rho guanine nucleotide exchange factor, putative *[Entamoeba histolytica* HM-1:IMSS]

>EAL46050.1 Rho guanine nucleotide exchange factor, putative [*Entamoeba* histolytica HM-1:IMSS]

MTEKLSDWEVLDENFNKTNTDIPPQLPPKPERLTIAGHHTQKAEIKSPSENTGKIISNDKEKDSVPSQQQ
RSSLQIGDSVELSENQNLIDNSISKNQPSKDVFETTNEPIKPKPETKALSKEDYRERIVNEMFDTEKSYV
NSMEICIKGYYEPLIQSGHSVAPADKVNAVFLHFQSVLSINKELLKNMTELKEKGELSTRLGEAFSQFIP
MMNVYKLFLGNSDTSLQFLVELEKSSKFNDILDLLRSHLPGDNQLDLRSYLIMPVQRLPRYKLLLTDLIK
HTDDDFVDKPKLIDALDKISKLATLVNEVIKERSRNQKLLELVDKIEGLSHELVTPYREYIKSGSLLKIC
RKDNRERYFYLFNDLLVYGIGDENKIKVSQEFQLESLKINLNQNVPNSFQILAKKSFSVIAKDETDRDEW
MKAINDAIQLEKSKLKTLKRDKKAEDEYIAPVWVQDTQNCQVCGAKFTTLFRKHHCRKCGLCICSNCSKQ
TIIINGKKERVCDSCANNNKKLGESTLTEQNSSESKDTLLSTSTEQDQQLTTSSSNDDNKKIDVQPILQE
ESNNEQSTSVDVENMNNNEQPLELPKLPPKRKSGFVKESVQQKQQPDILQLKENIVDSPSNEKVVDSQED
VPQQQPEELIPPTVPPKRKTQKTLTQEEGVKQENQPLKRNDSKTLQQDKPSSIETQQKEQPSSNKQEPIE
LCSQPTGKLQNSVNGITPTPSKKVAPPIPHRSLPPKPLRSSVITSPLTNTSSQPLPKQNPFVPPQQKTPT
KPSPQQLGCKQTQSQPTSYRQTTLNKTTTQQNSARLAASLNQSSPRTKSVQPEVVKSTPIKEDISQYRSV
RDSPFLKQQRAQKESELSQVSPRRPQPPTPKRGVTK

7. List of Publication:

Published

- **Gunjan Gautam**, Syed Arif Abdul Rehman, Preeti Pandey, and Samudrala Gourinath. Crystal structure of the PEG-bound SH3 domain of myosin IB from Entamoeba histolytica reveals its mode of ligand recognition. Acta Cryst. (2017). D73

- Priya Rani, **Gunjan Gautam**, Kongara Hanumantha Rao, Swagata Ghosh, Samudrala Gourinath, Suman Kumar Dhar, Asis Datta. Cloning, expression, purification and crystallization of a novel glcnac metabolic protein, gig2 (duf1479) from Pathogenic fungus *Candida albicans*. JOURNAL OF PROTEINS AND PROTEOMICS, 8(2), 2017, pp. 127-1327

- Pandey P, Verma V, **Gautam G**, Kumari N, Dhar SK, Gourinath S. Targeting the β-clamp in Helicobacter pylori with FDA-approved drugs reveals micromolar inhibition by diflunisal. FEBS *Letters* (2017). doi: 10.1002/1873-3468.12734.

Manuscript under preparation

- **Gunjan Gautam**, Sabir Ali, Alok Bhattacharya, Samudrala Gourinath. A novel interaction between EhMyosin IB and RhoGEF during phagocytosis in E. histolytica.