Cloning, Over-expression and Biochemical Characterization of Putative Protease PFI1625c from *Plasmodium falciparum* 3D7

A Thesis
Submitted in Partial
Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

by

Ms. Kimjolly Lhouvum

Department of Biotechnology
Indian Institute of Technology Guwahati
Guwahati 781039, Assam, India

July 2014

Dedicated to my Parents,

Siblings and

All who have taken pain for

my Thesis

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOTECHNOLOGY

STATEMENT

I hereby declare that the matter embodied in this thesis entitled "**Cloning, Over-expression and Biochemical Characterization of Putative Protease PFI1625c from *Plasmodium falciparum* 3D7"** is the result of investigations carried out by me in the Department of Biotechnology, Indian Institute of Technology Guwahati, India, under the supervision of **Dr. Vishal Trivedi.**

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work of other investigators are referred.

July, 2014

Kimjolly Lhouvum

Roll No: 09610622

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DEPARTMENT OF BIOTECHNOLOGY

CERTIFICATE

It is certified that the work described in this thesis entitled "**Cloning, Over-expression and Biochemical Characterization of Putative Protease PFI1625c from *Plasmodium falciparum* 3D7" by Ms. Kimjolly Lhouvum (**Roll No: 09610622), submitted to Indian Institute of Technology Guwahati, India for the award of degree of Doctor of Philosophy, is an authentic record of results obtained from the research work carried out under my supervision at the Department of Biotechnology, Indian Institute of Technology Guwahati, India, and this work has not been submitted elsewhere for a degree.

Dr. Vishal Trivedi

(Supervisor)

Acknowledgements

It is exceedingly a great pleasure to express my gratitude for all kinds of affections showered upon me throughout my research. I am really honored by the opportunity to be a part of this department where since I joined, I was enriched with excellent basic skills and knowledge by highly ambitious scientists and my colleagues.

First of all, I would like to take the opportunity to express my heartiest gratitude to my supervisor Dr. Vishal Trivedi, Associate Professor, Department of Biotechnology, who, despite his hectic schedule, devoted his precious time to inculcate my mind in the field of my work, equipped me with the best way he could technically and logically and rendering me his constant tireless guidance throughout with patience, understanding and kindness. I feel fortunate throughout my research for my guide being accessible and available to discuss about my work at any given circumstances.

I extend my gratitude to all my Doctoral Committee members namely, Dr. Nitin Choudhary, Dr. Vikash Kumar Dubey and Dr. Debasis Manna for their valuable contributions at each step of my work to make my work accomplished.

I am also thankful to the successive Heads of Department of Biotechnology, Indian Institute of Technology, Prof. Arun Goyal and Prof. V.V Dasu for providing me the departmental facilities to carry out my research work.

Worthy to mention, I owe my gratitude to the non-teaching staffs of Department of Biotechnology, who assisted me technically and availing me all the starting materials for my work as a foundation.

I am indeed grateful to all faculties who poured me a fountain of invaluable knowledge during course work namely, Dr. Vikash Kumar Dubey, Dr. Lingraj Sahoo, Dr. Pakshirajan, Dr. Rakhi Chaturvedi, Dr. Bithiah Grace Jaganathan, Dr. Anil K. Limaye and my supervisor Dr. Vishal Trivedi.

My special thanks go to Dr. Nitin Choudhary for not only providing me with facilities to carry out a part of my work but also imparting me with knowledge and giving me valuable suggestions and solving my doubts. I extend my special thanks to Dr. Vibin Ramakrishnan for kindly allowing me to use their facilities when I needed the most and I deeply appreciate the help and cooperation I received from his students especially to Prakash.

I am immensely indebted to all my Malaria Research Group members, past and present namely, Dr. Rohitash Deshmukh, Rajesh, Balaji, Suman, Arnish, Anish, Rakesh, Prasad, Sushant, Sourav, Ananya and Ankita for being a part of the family and contributing in countless number of ways throughout my research.

There have been a number of generous contributions from different angles and for this, I would like to express my gratitude to Dr. P. Saravanan, Dr. Abhay Narayan Singh, Babina Chakma and Rituparna for rendering help experimentally and logically. I also extend my gratitude to my neighboring labs and its members (VKD lab and AML lab) for their generosity and co-operation whenever I need an instant help.

Mental happiness is an essence to hard work and ultimately friends are main source of mental happiness and de-stressing agents. Their presences have made my stay in IIT Guwahati more incredible and momentous. Mention may be made of Amrita, Yagom, Darilang, James, Moa, Sangtea and Longshi.

A time has come when I can take the privilege to express my indebtedness I owe to my parents and family members without whose moral, emotional and spiritual support, nothing would have been possible. I extend my gratefulness to all my friends, families and IITG prayer group who lifted me up when I was down through their constant humble prayers.

Last but not the least, I give all the glory to Almighty God Whose faithfulness, Blessings and Answers to our prayers have led to all my achievements and turning all impossibilities possible as the Bible says "For God, Nothing is Impossible and What God has started for you, He will finish".

Kimjolly Lhouvum

July, 2014

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

UNITS

Å	Angstrom	**kCal**	Kilo calorie
bp	Base pairs	**mAU**	Milli absorbance units
Da	Dalton	**mg**	Milligram
µg	Microgram	**mg/ml**	Milligram per Milliliter
µl	Microlitre	**M**	Molar
µg/ml	Microgram per Milliliter	**min**	Minute
µm	Micrometer	**mM**	Millimolar
µM	Micromolar	**nm**	Nanometer
°C	Degree Celsius	**nM**	Nanomolar
gm	Gram	**rpm**	Revolution per Minute
gm/L	Gram/Liter	**w/v**	Weight/Volume
hrs	Hours	**v/v**	Volume/Volume
kDa	Kilo Dalton	**V**	Volt

ABBREVIATIONS

ACE	Atomic contact energy
ACT	Artemisinin combination therapies
AQ	Amodiaquine
BF	Bright field
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
CBEV	curling-buckling-eversion-vesiculation
CQ	Chloroquine
CTAB	Cetyl Trimethyl Ammonium Bromide
DAB	Diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
DHFR	Dihydrofolatereductase
DHPS	Dihydropteroate synthase
DMSO	Dimethylsufoxide
DOPE	Discrete optimized protein energy
DPAP	Calpain-like dipeptidyl peptidase
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent Assay
FITC	fluorescein isothiocyanate
FPLC	Fast protein liquid chromatography
FV	Food vacuole
HAP	Histo aspartic protease
Hb	Hemoglobin
Hb	Hemoglobin
HCl	Hydrochloric acid
HRP	Horseradish peroxidase
IC50	50% inhibitory concentration
IDA	Iminodiacetic acid
IDE	Insulin degrading enzyme
IPTG	Isopropyl β-D-1-thiogalacto pyranoside
iRBC	Infected red blood cells
LB	Luria Bertani
TEMED	Tetramethylethylenediamine
Tris	Tris(hydroxymethyl)amino methane)
LMW	Low molecular weight
M18AAP	M18 Aspartylaminopeptidase
MetAP1	Methionine aminopeptidase
MMP	Matrix metalloprotease
MPP	Mitochondrial processing peptidase
MQ	Mefloquine
MSP	Merozoite surface proteins
NCBI	National center for biotechnology information
NJ	Neighbor joining
NR	Non redundant
OD	Optical density
OPD	O-Phenylenediamine
OPP	Organelle processing peptidase
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCD	Programmed cell death
PCI	Phenol:Chloroform:Isoamyl alcohol
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethelene glycol
PfA-M1	M1 Alanyl aminopeptidase
PfAPP	Aminoacylprolineamino peptidase
PfLAP	M17 Leucylaminopeptidase
PfROM	*P. falciparum* rhomboid
PMs	Plasmepsins
PM-V	Plasmepsin-V
PVDF	Polyvinylidene fluoride
RBC	Red blood cells
RMSD	Root mean square deviation
RT-PCR	Real Time Polymerase Chain Reaction
SDS	Sodium dodecyl sulfate
SERA	Serine repeat antigen
SPP	Stromal processing peptidase
TCA	Trichloroacetic acid

Chapter-1: Introduction and Review of Literature

Introduction and Review of Literature

1.1 Introduction

Despite the consistent steps taken by World Health Organization, *Plasmodium falciparum* infection still remains a global threat. There were an estimated 207 million cases of malaria in 2012 as reported in December 2013 and nearly 1 million deaths per year in which 90% of all malaria deaths occur in sub-Saharan Africa [1-3]. This includes only 10% of cases estimated to occur globally due to lack of surveillance system especially in poor countries with high malaria cases [1, 2]. Malaria is caused by Plasmodium species viz, *P.falciparum, P.vivax, P.ovale, P.malariae* and the fifth most recently reported, *P.knowlesi* which is extremely rare [4]. *P.falciparum* infection is found to be most lethal and accounts for nearly 1 million deaths each year [1]. Common symptoms of malaria include fever, chills, sweat, headache, nausea, vomiting, dry cough, muscle or back pain. *P.falciparum* infection leads to severe malaria and affect the nervous, respiratory, renal, liver, spleen and hematopoietic system [5, 6]. The malaria parasite requires two hosts to complete its life cycle. The sexual life cycle is completed in invertebrate host while the asexual cycle in vertebrate hosts. Human malaria begins with asymptomatic liver stage followed by symptomatic RBC stage. The liver stage is the first gate of malaria infection and it is marked by generation of large number of merozoites from invading sporozoites. The RBC phase is an intervening stage and is responsible for the disease development. During the RBC stage, the parasite undergoes three conspicuous developmental morphologies; ring, trophozoite and schizont. Malaria has become a threat to global health mainly due to increase in resistance of parasites to the existing drugs including newly discovered artemisinin. The increase in resistance leads to put pressure for exploration of novel drug targets from parasite. There are various pathways which are crucial for parasite survival. Various proteases have been explored and they play significant roles in controlling these pathways. The hemoglobin digestion is carried out by aspartic, cysteine and metalloproteases. Merozoite invasion and egress also require proteolytic activity to accomplish these events, mainly controlled by serine or metalloproteases. There are a number of putative proteases present in the genome of *P.falciparum* out of which about 85% remain uncharacterized including several members from metalloproteases family.

1.2 Life cycle of malaria parasite.

Malaria is a protozoan disease where the parasite completes its life cycle in two hosts; vertebrate (man) and invertebrate host (female *Anopheles* mosquito). The infection starts when the malaria infected mosquitos bite human and inject sporozoites into the blood stream. These sporozoites then migrate to the liver, where they pass through kupffer cells, invade hepatocytes and develop into liver merozoites. In *P.falciparum*, the number of merozoites released by single sporozoite at the end of cycle can reach upto 30,000 [7]. The merozoites are released from liver into the blood stream and infect erythrocytes. Within RBCs merozoites go through ring, trophozoite, schizont stage and the daughter merozoites are released from RBC through egress. The released merozoites infect another erythrocyte to continue the asexual life cycle (Figure 1.1). All clinical symptoms and pathology arise mostly due to the asexual blood stages. *P.falciparum* infection is lethal because it crosses blood brain barrier and can cause cerebral malaria [8, 9]. Some of the parasites in intra-erythrocytic stages reproduces sexually and develop into male or female gametocytes. These are taken up by the mosquito during feeding which develop into gametes in the insect gut and fuse to form zygotes. The zygote develops to form an invasive ookinete that traverses the midgut. The ookinete transforms into an oocyst from

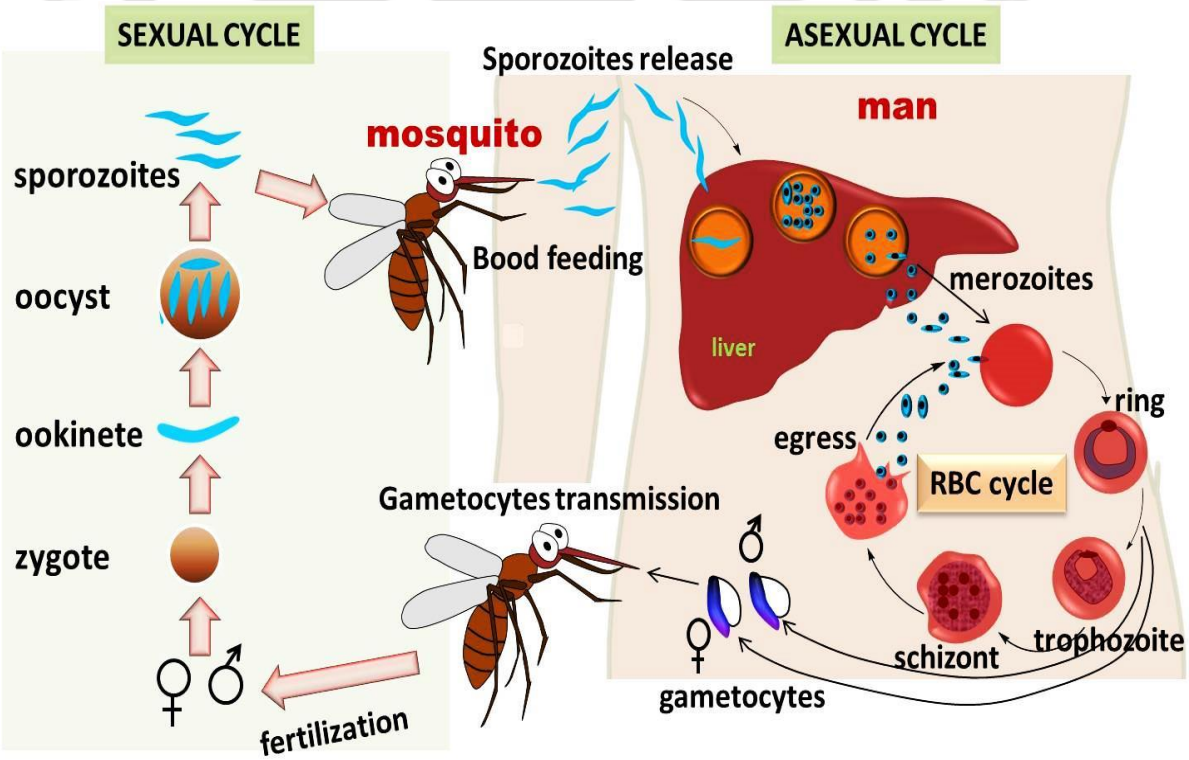


Figure 1.1.Schematic diagram of malaria parasite life cycle in mosquito (sexual phase) and human (asexual phase) host.

which sporozoites are released. The sporozoites then migrate to the salivary glands for injection into a human host during the next blood meal [10].

1.2.1 Significance of different molecular events during parasite life cycle

The RBC stage of malaria parasite starts soon after the release of merozoites from liver cells. In the RBC, the parasite undergoes three conspicuous developmental stages namely the ring, trophozoite and schizont. There are three critical events taking place in the RBC which may be briefly described below;

1) Invasion: This is the first essential step of parasite to start its life cycle inside the RBC. It is initiated by an invasion machine "merozoite" as a function for continuing the replicative cycle of asexual stage [11]. It involves five steps as illustrated in Figure 1.2A. These events are: (a) Recognition and attachment of RBC surface receptors by merozoites surface proteins (MSPs), (b) Reorientation of merozoite to attach with the apical portion to the RBC surface receptor, (c) Formation of tight junctions involving high affinity receptor-ligand interactions, (d) Movement of tight junction to posterior pole without actually penetrating the RBC membrane which is powered by the parasite's actin-myosin motor. Here shedding of merozoite surface coat and proteolytic removal of adhesive protein at the junctions take place and (e) Resealing of the RBC membrane creating an envelope for the invading parasite which is known as parasitophorous vacuole. The process of invasion requires involvement of various apical membrane proteins and other enzymatic activity [10, 12-14].

2) Growth phase involving hemoglobin digestion: Degradation of hemoglobin (Hb) inside the RBC is essential for the development of parasite and to maintain homeostasis [15, 16]. Soon after the merzoites invade erythrocytes, it undergoes a series of morphological changes, the first morphology is called the ring or young trophozoite with the newly formed hemoglobin-laden food vacuole appears like a ring in giemsa-stained blood smears. The next developmental morphology is called the trophozoite in which stage the cell enlarges as hemoglobin degradation takes place actively followed by formation of hemozoin or dark brown pigment. The last developmental morphology is called the schizont which marks the end of the trophozoite stage and the beginning of the nuclear division followed by cell division [17]. The intra-erythrocytic stages of *P.falciparum* rapidly consume almost all the cell hemoglobin and utilize it as nutrition source for its growth and development. The mechanism of hemoglobin digestion may be described briefly under the following points;

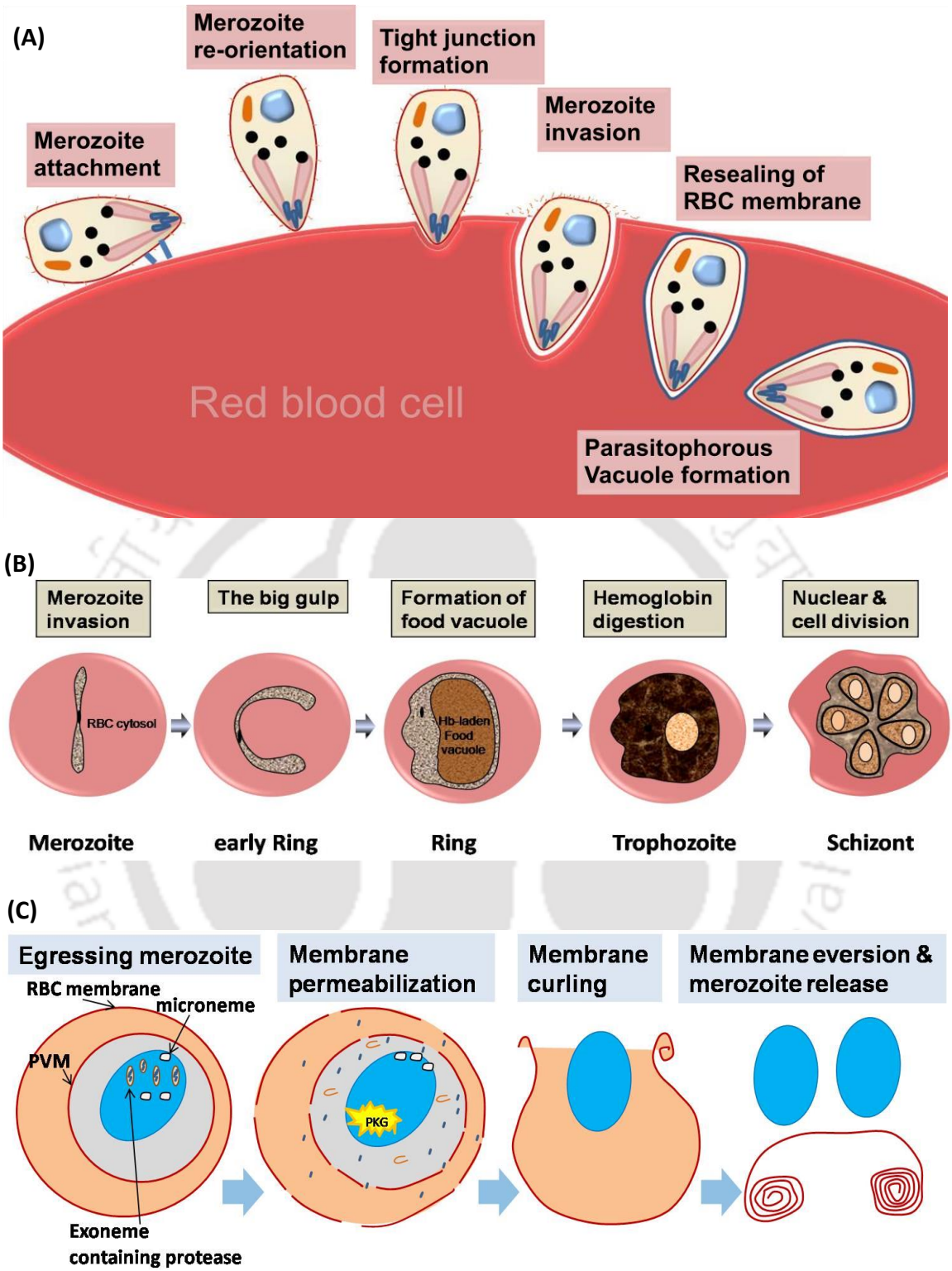


Figure 1.2. Molecular events taking place during RBC stages of parasite life cycle. (A) The steps involved in the invasion of merozoite to RBC. **(B)** Schematic illustration of Hb digestion and developmental stages of asexual parasites **(C)** Different hypothesis to explain the mechanism of merozoite egress from infected RBC.

i) Hemoglobin uptake: The uptake of hemoglobin is the first step which involves four distinct pathways. The first pathway is called the big gulp where the RBC cytoplasm is engulfed in the process of ring stage transformation giving rise to large vacuole filled

with hemoglobin (Figure 1.2B). The second pathway is the use of cytostome or parasite mouth which is laden with hemoglobin. The cytostome then buds off by endocytosis and form a part of the lysosomal compartment or the food vacuole. The third and fourth pathways involve cytostomal tube and phagotrophy similar to the first and second event respectively except that the latter two are actin-independent [18].

ii) Hemoglobin digestion: After the uptake of hemoglobin is completed inside the food vacuole, hemoglobin digestion takes place in a semi-ordered fashion which involves cascades of proteolytic activity [19]. Degradation process is initiated by aspartic proteases by cleaving a conserved hinge region in the alpha chain of native hemoglobin. The denatured fragment is then cleaved by cysteine proteases and aspartic proteases. Further downstream in the pathway involves metalloprotease which acts on 15-20 long oligopeptides followed by dipeptidylaminopeptidase which acts on 5-10 long peptides. The final dipeptide is transported into the cytosol and cleaved by metalloproteases into individual amino acids [20]. The toxic free heme derived after hemoglobin digestion is converted into non-toxic bio-crystal hemozoin [21].

3) Egress: Parasite egress is a fundamental step in the life cycle of *P.falciparum*. The parasite replicates by a process called schizogeny (schizont) during which time, multiple copies of genetic material is formed along with other cellular ingredients resulting in independent infective merozoites. These merozoites need to exit from the RBC to initiate new invasion by rupturing the RBC membrane in a process called egress. There are different models explaining the mechanism of egress however, each of them is inconclusive and un-explanatory. The underlying proteases which may be involved in egress are needed to be explored further. In one of the studies, egress or merozoite release involves change in RBC membrane conformation and primary rupture of the parasitophorous vacuole membrane followed by rupture of the RBC plasma membrane [22, 23]. Permeabilization of the erythrocyte membrane involves calcium-dependent perforin-like protein [24]. A number of parasite proteases and organelles involved in egress are released under the action of cGMP-dependent protein kinase [25]. Another hypothesis of egress involves swelling of the infected or pregnant RBC due to osmotic stress followed by rupture of the RBC membrane releasing the merozoites. During hypotonic shock, pregnant RBC undergoes a series of transformation known as curling-buckling-eversion-vesiculation (CBEV) which is how the membrane ruptures and releases the merozoites [26]. The different models of egress are combined for complete pictorial representation (Figure 1.2C).

1.2.2 Drug resistance in malaria

Malaria has become a major health issue mainly due to development of drug resistance in parasites. In India, resistance of *P.falciparum* to chloroquine, the cheapest and the most used drug was first reported in the year 1973 from Diphu of Karbi-Anglong district in Assam state [27]. Drug resistance in *P. falciparum* is not confined to chloroquine alone, but also to the other currently used antimalarials including artemisinin [28-30]. As a result of rapid resistance in *P.faciparum*, cases of *P.falciparum* infection increases despite the decrease in total malaria cases [31]. In general, resistance appears to occur through spontaneous mutations (single or multiple) in a given gene that confers reduced sensitivity to a given drug or class of drugs. Resistance also develops more quickly when large populations of parasites are exposed to drug pressure. This removes sensitive parasites, while resistant parasites survive [32]. List of existing drugs with their mode of action and mechanism of resistance is given in Table 1.1 and briefly discussed below:-

(i) Chloroquine acts by getting accumulated in the food vacuole where it inhibits heme polymerization. Resistant strains are able to efflux the drug by an active pump mechanism and release the drug at least 40 times faster than sensitive strains, thereby rendering the drug ineffective. It is associated with mutation in the *pfmdr-1* and *pfcrt* genes [33-35].

(ii) Antifolate combination drugs act through sequential and synergistic blockade of 2 key enzymes dihydrofolatereductase (DHFR) and dihydropteroate synthase (DHPS) involved in folate synthesis. Resistance develops with increasing number of mutations in the gene encoding these enzymes [36, 37].

(iii) Atovaquone acts through inhibition of electron transport at the cytochrome *bc1* complex. Resistance is due to a single mutation at position 268 (Y268S) of cytochrome b [38].

(iv) Mefloquine resistance is associated with increase in copy number and polymorphism of *pfmdr-1* gene [32, 39, 44].

(v) Quinine resistance mechanism is still unclear but is implicated to be conferred by *pfmdr-1* and *pfcrt* mutations associated with chloroquine. Sodium/hydrogen exchanger (*pfnhe-1*) gene and mutations in multidrug resistance protein gene may contribute to quinine resistance [45].

(vi) Artemisinin acts by alkylating heme and other essential parasitic proteins and its activity against plasmodium requires hemoglobin uptake and digestion. The mechanism of artemisinin resistance is still under study yet a molecular marker known as K-13 propeller is shown to be responsible for artemisinin resistance. Artemisinin combination

therapies (ACT's) are now recommended to avoid resistance mechanisms [46-49].

Table 1.1. List of antimalarial drugs in use*

Drug name	Mechanism of action	Mechanism of drug resistance
Mefloquine + artesunate	Synergistic effect of individual action and reduce mefloquine side effects	Increase production of target gene *pfmdr-1*
Sulfadoxine/ pyrimethamine + artesunate	Synergistic blockade of folic acid synthesis	Not known
Lumefantrine + artemether	Similar to halofantrine	Not known
Chloroquine (CQ)	Inhibits hemoglobin digestion and heme polymerization	Efflux of drug due to mutation in *pfcrt* gene
Antifolate	Blocks enzymes involved in folate synthesis	Decreased sensitivity due to mutations in *dhfr* and *dhps* genes
Amodiaquine (AQ)	Inhibits heme polymerization	Decreased drug sensitivity due to mutation in *pfcrt* SVMNT allele
Sulfadoxine/ pyrimethamine	Inhibit folate synthesis	Decreased susceptibility due to mutation in *dhfr* and *dhps* genes respectively
Mefloquine (MQ)	May inhibit ingestion of hemoglobin	Increase production of target gene *pfmdr1*
Halofantrine	Unknown mechanism may act as schizonticides	Still unclear
Quinine	Accumulate in the food vacuole and act as schizonticidal	Still unclear may be similar to chloroquine
Tetracycline/ doxycycline	Blocks protein synthesis	Not known
Clindamycin	Act as blood schizonticides	Not known
Atovaquone /proguanil	Atovaquone inhibits of electron transport at the cytochrome bc1complex. Proguanil inhibits folate synthesis	Single mutation in Y268S of *Cyt* band double mutation of *dhfr* gene
Artemisinin compounds (Artesunate, artemisinin, artemether)	produce carbon-centered radicals that alkylate heme and other critical parasite proteins	Mutation in K-13 propeller
Primaquine	Effective against gametocytes	Not yet reported
Bulaquine	Effective against gametocytes	Not yet reported
Ozonide OZ439	Similar to artemisinin compounds	Not yet reported

*** Data for antimalarial drugs is compiled from WHO 2001, different literatures mentioned in the section 1.2.2 and [37, 39-43]**

1.2.3 Potential drug targets in *P.falciparum*

In the light of malaria, chloroquine was known to be the most effective and first line of defense against malaria. However, due to development of resistance in *P.falciparum*, chloroquine is surpassed and replaced by artemisinin. Artemisinin is a plant-derived antimalarial compound which was known to provide chemically safe with long therapeutic life. Yet, the efficacy of artemisinin also been challenged with the report of

resistance against it. The problems of malaria became more intense when most of the anti-malarial drugs including artemisinin conferred resistance [30]. Parasite resistance towards drugs is difficult to control because virulence genes expression occur through chromatin modification in parasite [50]. Thus the limitation of antimalarial chemotherapy underlines the constant need for identifying new drugs. Hence, identifying new drug targets in malaria parasite biology have become a core emphasis in malaria research. Some of the notable pathways of potential drug targets include 1) invasion of fresh RBC by merozoites, 2) Assimilation and biogenesis and 3) egress of merozoites (Figure 1.4). Target pathways of interest in the parasite biology mainly involve the proteome viz, proteases, kinases and secretome [51-55]. These proteins are shown to have relevance in the life cycle of malaria parasite and each has their own essence and cruciality. Protease raises interest as they are involved in hemoglobin digestion, the most conspicuous and detectable event in the RBC stages of malaria due to its hemozoin or pigment formation [56, 57].

1.3 Proteases of *P.falciparum*

Proteolysis is a metabolic event required by all living organisms from human to virus for growth and control of physiological processes and also for synthesis of proteins from amino acids derived after proteolysis of ingested food. This event is carried out by hydrolyzing a peptide bond through nucleophilic attack. There are four major classes of proteases namely, aspartic, cysteine, serine and metallo-protease along with recent inclusion of threonine protease [58]. In parasitic organisms, proteases constitute major virulence factors in development of various diseases such as malaria, leishmaniasis, african sleeping sickness etc [59]. In malaria, proteases are one of the most essential factors for parasite growth and virulence. They carry out obligatory functions throughout the life cycle of parasites. During the RBC stages, malarial proteases take major parts in the crucial steps such as invasion, hemoglobin digestion and egress (Figure 1.3). In *Plasmodium falciparum*, the comparative genomic analysis shows 92 predicted proteases out of which about 14 are characterized [60]. These 92 putative proteases belong to 26 families of five clans (as shown in table below) of which 11% aspartic, 36% cysteine, 22% metallo, 17% serine, and 14% threonine. Combining the complementary results from microarray, RT-PCR, and proteomics analysis, 88 out of 92 putative proteases were transcribed and 67 were translated at some stage in the life cycle. The remaining four may be expressed at extra-erythrocytic stages or may be pseudogenes.

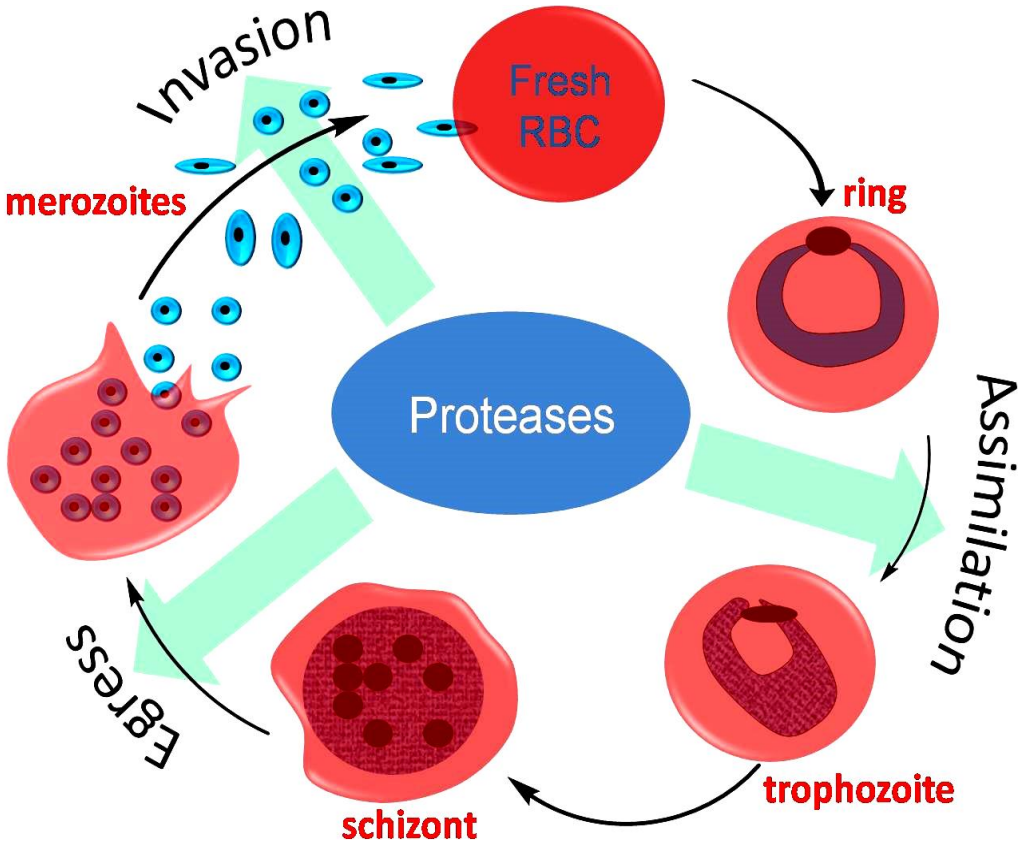


Figure 1.3. Schematic representation of different role of proteases in the RBC stages of *P.falciparum* life cycle.

1.3.1 Aspartic proteases:

Aspartic proteases are a group of proteases that have two highly conserved aspartic acid residues at the active site. They work optimally in acidic conditions such as stomach and food vacuoles thus limiting their location of activity and therefore constitute a lesser population in protease family [61]. They play important role in digestion of food and increase of blood pressure [62].

Classification of aspartic proteases: In *P.falciparum*, aspartic proteases constitute about 11% of the total proteases. Aspartic protease family A1 is identified in plasmodium genome and they are known as plasmepsins. There are 10 plasmepsins (I to X) identified in the genome and presented in Table 1.2. Plasmepsin III is also known as histo-aspartic protease (HAP) due to its difference in active site residues. Out of the 10 plasmepsins, plasmepsins I-V and IX-X are expressed while plasmepsins VI, VII and VIII are not expressed during erythrocytic stages [60, 63].

Table 1.2. Members of the aspartic proteases in *P.falciparum* genome*			
Protease family	**Name of the enzyme**	**PlasmoDB ID**	**Charaterization status**
A1	Plasmepsin I	PF14_0076	Yes
	Plasmepsin II	PF14_0077	Yes
	Histoaspartic protease (plasmepsin III)	PF14_0078	Yes
	Plasmepsin IV	PF14_0075	Yes
	Plasmepsin V	PF13_0133	No
	Plasmepsin VI	PFC0495w	No
	Plasmepsin VII	PF10_0329	No
	Plasmepsin VIII	PF14_0623	No
	Plasmepsin IX	PF14_0281	No
	Plasmepsin X	PF08_0108	No
*** Data is obtained from plasmodium database (PlasmoDB.org) and literature searches.**			

Mechanism of action: Plasmepsins possess two aspartic acid residues at their active sites except for plasmepsin III (also known as histo aspartic protease) [20]. Generally, one aspartate residue acts as a general acid, whereas the other is in the ionized form and acts as a general base to extract a proton from the water molecule. The activated water molecule is the nucleophile that attacks the scissile bond. The resultant nucleophilic attack of water molecule is the hydrolysis of the substrate into two peptide fragments (Figure 1.4). In HAP, at its pH optimum near 6, the active-site histidine may function in the protonated state and act as a general acid, if its pKa remains near 6–7, the pKa of histidine amino acid [20].

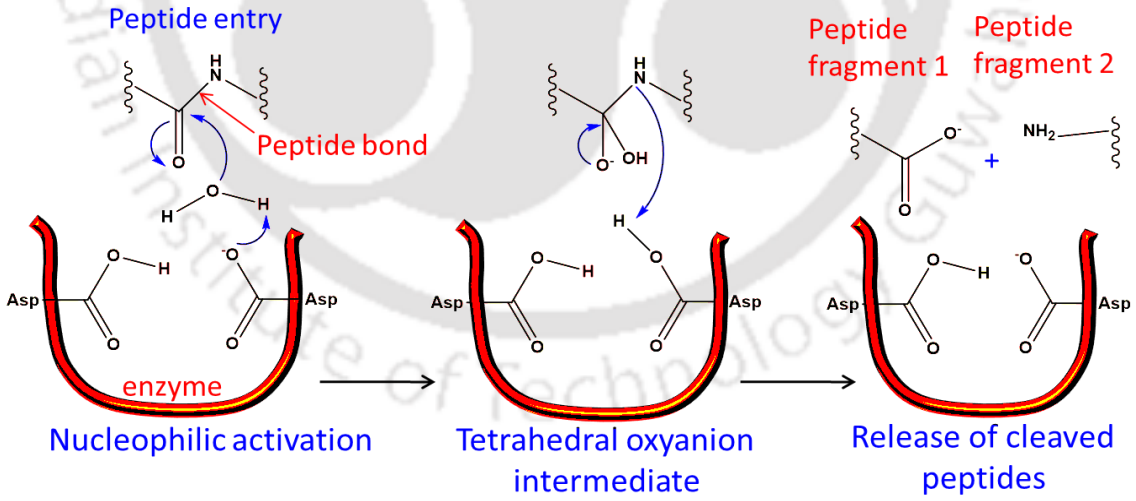


Figure 1.4. Catalytic mechanism of aspartic proteases. It involves two aspartic acid residues, one in ionized form which acts as general base and the unionized form acts as general acid.

Physiological roles of aspartic proteases: Aspartic proteases are known to be the key initiator for hemoglobin digestion. Most of the aspartic proteases play role in hemoglobin digestion which is discussed below;

(i) Hemoglobin digestion: In *P. falciparum*, several proteases have been implicated in the process of hemoglobin degradation. Hemoglobin is endocytosed from the erythrocyte cytosol and trafficked to an acidic food vacuole (FV), where it is degraded by parasite proteases proposed to act in a semi-ordered fashion [19, 64]. Aspartic proteases plasmepsins (PMs) initiate the degradative process by cleaving the native hemoglobin molecule in a highly conserved hinge region between Phe33 and Leu34 [65]. Subsequently cysteine protease and metalloprotease act downstream to degrade hemoglobin into smaller peptides (Figure 1.5). Out of 10 PMs present in *P.falciparum* genome, PM I, II, IV and HAP localizes to the food vacuole (FV), and participate in hemoglobin degradation [20]. Plasmepsins are also a part of protein complex involved in hemozoin formation and therefore it is plausible that plasmepsins play a part in hemozoin formation [56].

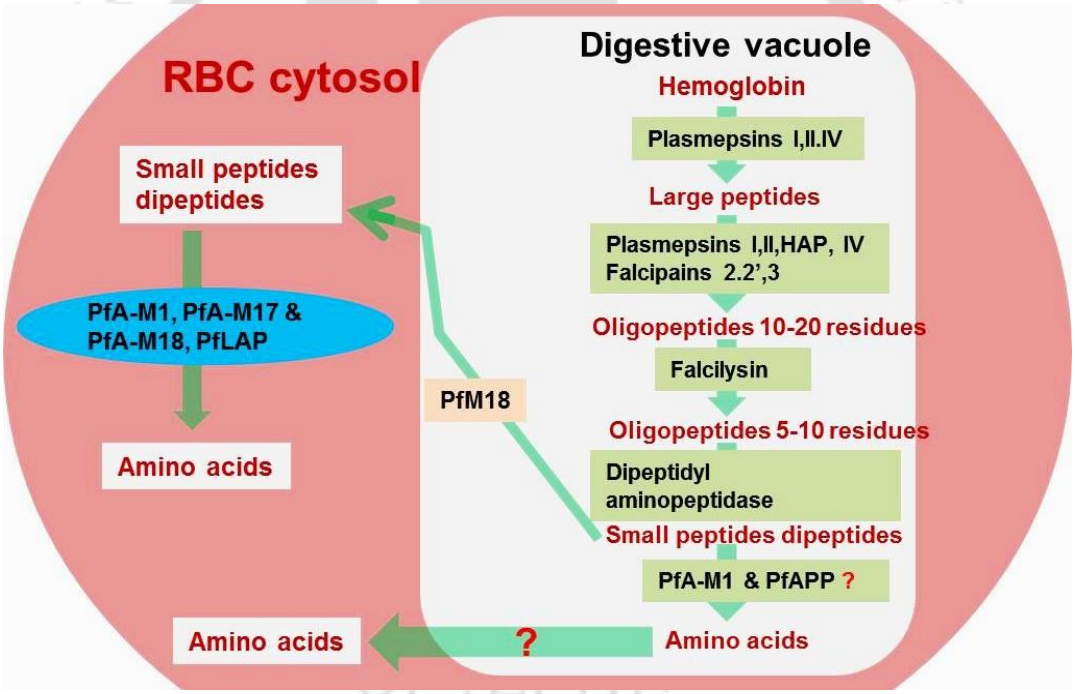


Figure 1.5. Role of proteases in hemoglobin digestion inside the digestive food vacuole as well as cytosol. Digestion is initiated by proteases of aspartic class (plasmepsins) and subsequently by cysteine proteases (falcipains and DPAP) and metalloproteases (falcilysin, PfA-M1,M17,M18, PfLAP and PfAPP). Some of the terminal processing pathway is still unclear.

(ii) Protein trafficking: Recently it is reported that plasmepsin V (PM-V) directs malaria effectors proteins to the host cell by cleaving them at PEXEL motifs and produce mature proteins which are exported. Immuno-fluorescence microscopy detects that PM-V localizes to the endoplasmic reticulum [66-68].

Therapeutic potentials of aspartic proteases: Plasmepsins have always been a target for antimalarial drug development as they are mainly involved in the hemoglobin digestion. The structure of plasmepsins III and IV differ considerably from human counterparts hence they are better candidates for antimalarial drug development [69]. A number of potent inhibitors are identified and tested against plasmepsins as given in Table 1.4 [70-77]. As anticipated, these inhibitors block the pathway and lead to parasite death. Recent gene knockout studied have shown that plasmepsins I-IV are not essential for the completion of the *Plasmodium* life cycle. However aspartic protease inhibitors used on these knockouts effectively kill the cells, suggesting that the additional proteases are vital for *Plasmodium* life cycle that may belongs to plasmepsins or new aspartic protease [78]. Hence exploration of new inhibitors against plasmepsins is a constant challenge in the anti-malarial drug development [79].

1.3.2 Cysteine proteases

Cysteine proteases constitute the largest protease class in *P.falciparum* genome which is 36% of the total *P.falciparum* proteases [60]. They are characterized by the presence of a catalytic cysteine, which mediates protein hydrolysis via nucleophilic attack on the carbonyl carbon of susceptible peptide bond. In human, they carry out multifunctional roles which include apoptosis, MHC class II immune responses, pro-hormone processing, and extracellular matrix remodeling important for bone development [80].

Classification of cysteine proteases: Cysteine proteases are classified as different families in *P.falciparum*. The families of cysteine proteases identified in the plasmodium genome consist of C1, C2, C12, C13, C14, C19, C48 and C56 (Table 1.3). The C1 family comprises four falcipains, two dipeptide-peptidase I (DPP I), one cathepsin and nine serine repeat antigens (SERAs). C2 family comprises only one protease calpain. C12 family consists of ubiquitin carboxyl-terminal hydrolase, family 1 (UCH1) proteases and C12 family has GPI8p transaminidase. C14 family is constituted by metacaspase while C19 family consists of UCH2 proteases. C48 family has two proteases namely sumo1 and Ubiquitin-like protease-2 (UIp2) peptidase and C56 has Dj-1 peptidase [60, 81, 82].

Mechanism of action: Cysteine proteases are sub-divided into clans sharing no sequence or structural identity [83, 84]. Clan CA, which is a C1 family (papain family) and C2 family (calpain) in *P.falciparum* utilise catalytic Cys, His, and Asn residues [85]. These are the best characterised cysteine proteases comprising of four falcipains, nine SERA, three dipeptidyl peptidase and a calpain homologue. A second clan CD, which is a C13

and C14 family in *P.falciparum* utilises a catalytic His-Cys dyad for activity [83] where transamidase and metacaspase are present in plasmodium. Clan CE, which is characterised by catalytic residues in the order; His, Glu (or Asp), Cys, is also represented in the *P. falciparum* genome. In general His-Cys diad catalytic system, the histidine activates the cysteine residue which becomes nucleophile and attack the incoming protein substrate resulting in the cleavage of the substrate and releasing of cleaved substrate. The cleaved substrate is released by nucleophile activation by water molecule which is represented in Figure 1.6.

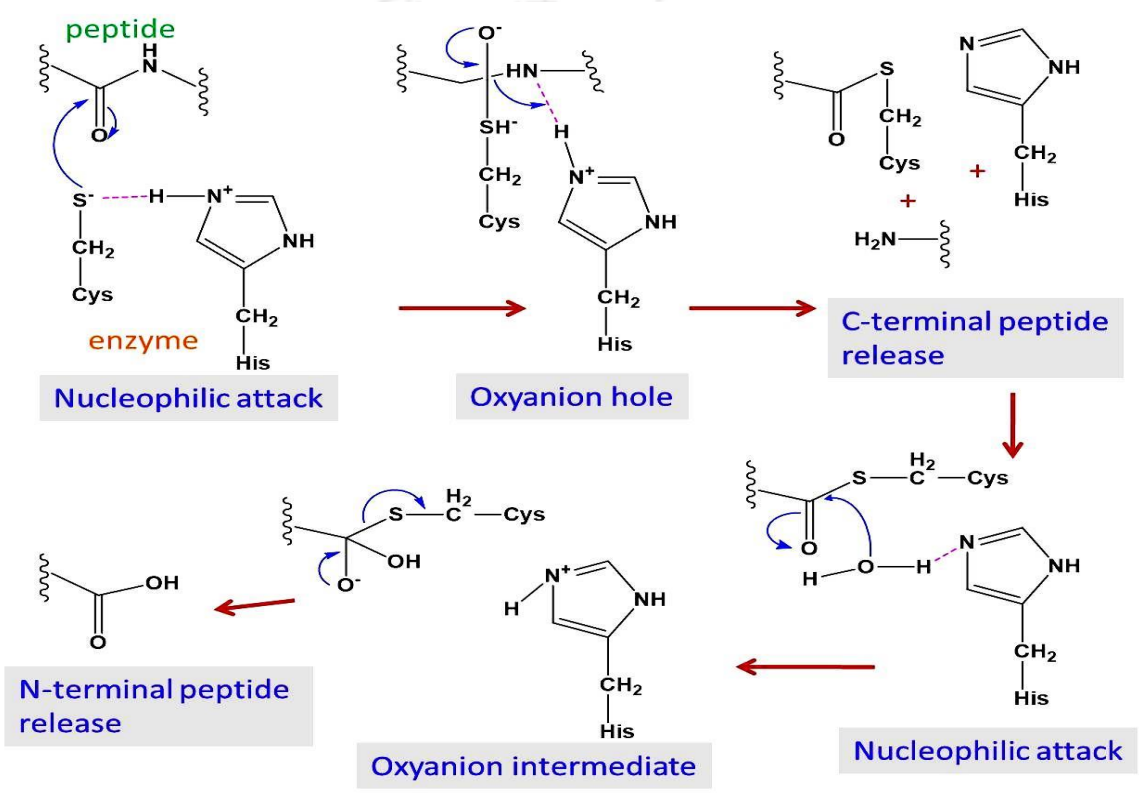


Figure 1.6. Generalized catalytic mechanism of His-Cys diad of cysteine protease. The catalytic cysteine is activated by histidine which enables the initiation of nucleophilic attack and release of peptides and aided subsequently by water molecule.

Physiological role of cysteine proteases: Cysteine proteases following its largest population in proteases of plasmodium, they also perform a wide spectrum of roles in the parasite biology**.** Studies on protease inhibitors have suggested that malaria cysteine proteases have diverse functions and are important for the hydrolysis of hemoglobin, erythrocyte rupture, and erythrocyte invasion [86] which are discussed below;

(i) Hemoglobin digestion: The best characterized Plasmodium cysteine proteases for hemoglobin digestion are from the falcipain family [87]. In *P. falciparum*, four falcipains (falcipain-1, falcipain-2 and -2′, and falcipain-3) have been described at intra-erythrocytic

Table 1.3. Members of the cysteine proteases in *P.falciparum* genome*

Cysteine protease family	Protease name	PlasmoDB ID	Characterization status
C1	Falcipain 1	PF14_0553	Yes
	Falcipain 2	PF11_0165	Yes
	Falcipain 3	PF11_0162	Yes
	Papain	PF11_0161	No
	DPP I	PFL2290w	No
	DPP I	PFD0230c	No
	Cathepsin C	PF11_0174	No
	SERA 1	PFB0360c	No
	SERA 2	PFB0325c	No
	SERA 3	PFB0330c	No
	SERA 4	PFB0335c	No
	SERA 5	PFB0340c	Yes
	SERA 6	PFB0345c	No
	SERA 7	PFB0350c	No
	SERA 8	PFB0355c	No
	Papain (SERA 9)	PFI0135c	No
	DPAP1	Not assigned	Yes
	DPAP3	Not assigned	Yes
C2	Calpain	MAL13P1.310	No
C12	UCH1	PF14_0577	No
	UCH1	PF11_0177	No
C13	GPI8p transamidase	PF11_0298	No
C14	Metacaspase	PF13_0289	No
C19	UCH2	PFA0220w	No
	UCH2	PFD0165w	No
	UCH2	PFD0680c	No
	UCH2	PFE1355c	No
	UCH2	PFE0835w0	No
	UCH2	MAL7P1.147	No
	UCH2	PFI0225w	No
	UCH2	PF13_0096	No
	UCH2	PF14_0145	No
C48	Sumo 1 protease	PFL1635w	No
	Ulp2 peptidase	MAL8P1.157	No
C56	DJ-1 peptidase	MAL6P1.153	No

*** Data is obtained from plasmodium database (PlasmoDB.org) and from literature searches.**

stage of parasite life cycle. Falcipain-2, 2′and 3 are involved in hemoglobin digestion while falcipain-1 is involved in invasion of merozoites (Figure 1.6). Gene disruption of falcipain-2 confirms its critical role in hemoglobin digestion. Falcipain-2 and falcipain-3 have been localized to the food vacuole by cell fractionation. These proteases are synthesized as integral membrane proteins, and subsequently hydrolyzed to release soluble active proteases [88, 89]. Falcipain 2 is also involved in hemozoin formation as evident from its presence in hemozoin formation protein complex [56]. The calpain-like

dipeptidyl peptidase I (DPAP1) also hydrolyses the protein complex [56]. The calpain-like dipeptidyl peptidase I (DPAP1) also hydrolyses the hemoglobin oligopeptides into 2-10 amino acid residues which are exported to the cytoplasm where they are cleaved into single amino acid units by the cytosolic neutral aminopeptidases [64, 90]. Inhibitor studies have implicated that the food vacuole PMs and falcipain cysteine proteases act co-operatively in Hb hydrolysis [91].

(ii) Invasion: Among the four falcipains that have been described at intra-erythrocytic stage of parasite life cycle, only falcipain-1 is involved in invasion by merozoites [89].

(iii) Erythrocyte rupture: The serine-repeat antigens are one of the cysteine proteases which have cysteine protease motifs, but in few members, the active site cys is replaced by a ser. One of these proteins, SERA-5, was recently shown to have serine protease activity and play a role in erythrocyte rupture (Figure 1.8). It localizes to the parasitophorous vacuole and is implicated to have intra-erythrocytic functions. DPAP3 is another protease involved in parasitic egress by maturing PfSUB2 [92-95]. Falcipain-2 is also found to be involved in merozoite egress which is revealed by treatment of falcipain-2 with si-RNA resulting in blocking of erythrocyte membrane rupture [96].

Therapeutic potentials of cysteine proteases: Rosenthal and team have demonstrated that cysteine proteases have potent in-vitro and in-vivo antimalarial effects and these have been supported by inhibiting the proteases with specific inhibitors. Few of selected cysteine protease inhibitors are given in Table 1.4. Specific inhibitors such as peptidylfluoromethyl ketone [97], vinyl sulfone [98], and aldehyde inhibited falcipain activities and disrupt the development of cultured parasites at nano-molar (nM) concentration [99]. The inhibition of parasite development was accompanied by inhibition of specific step in hemoglobin hydrolysis, marked by the appearance of swollen, dark-staining food vacuoles. Non-peptidyl inhibitors of falcipains are more potent than peptidyl inhibitors and extensive studies are underway for antimalarial drug development [100]. Flavonoids, peptidomimetics, suramic analogues, symplostatin-4 and different synthesized compounds are also shown to have potent inhibitory effect on falcipain-2 and 3 [76, 101-107]. Specific inhibitors of DPAPs and SERAs also effectively reduced parasite growth [93, 108]. There is a recent report on programmed cell death (PCD) pathway in *P.falciparum* which has general features of mammalian apoptosis that is mediated by clan CA cysteine proteases probably the calpains or cathepsins [109] which is an unexplored potential pathway in the life of a parasite.

Table 1.4 Role of *P.falciparum* proteases during asexual stages and their inhibitors*

Type of protease	Name of protease	Physiological role	Inhibitor	IC_{50}
Aspartic	Plasmepsin I	Hemoglobin digestion	Asp20-21 Peptidomimetic SC-50083 Azanorbornenes	84 µM 2 µM 6 nM
	Plasmepsin II	Hemoglobin digestion	Pepstasin A, Halofantrine KNI-727 Flavonoid myricetin Benzimidazole derivatives	16 µM 2-5 µM 6-10 µM 7.2 µM 2-48 µM
	Plasmepsin IV	Hemoglobin digestion	GBIII32	19.9 nM
	Histo-aspartic	Hemoglobin digestion	Pepstasin A KNI-10006	1 µM 0.6 µM
	Plasmepsin V	Trafficking effector proteins to host cell	Pepstasin A HIV protease inhibitors	1 µM 50 µM
Cysteine	Falcipain-1	RBC invasion by merozoites	E-64 Leupeptin E-64d Peptidyl epoxide YA29	3.4 µM 6.8 µM 1.5 µM 11.2 µM
	Falcipain-2′2	Hemoglobin digestion	Fluoromethyl ketone Vinyl sulfone MLA myricetin morin peptidomimetic compounds small compounds suramin analogues	3.38 nM 2-13 nM 1.3 nM 1.5 µM 1.7 µM 6.33 µM 7 µM 36 nM
	Falcipain-3	Hemoglobin digestion	MLA compound II Compound III	2.4nM EC 4 µM EC 1.5 µM
	Falcipains	Hemoglobin digestion	MG132 Symplostatin 4	47.6 nM 8.5-140 nM
	Calpain like-DPAP1	Oligopeptide digestion	Compound ALLN Ala-4(I)Phe-DMK ML4118S	72 nM 2.8 nM 19 nM
	DPAP3	Parasitic egress by maturing PfSUB1	Vinyl sulfone SAK1 ML4118S	15 µM 100 nM
	SERA-5	Erythrocyte rupture	Cyclic peptide SPB1	1.7 µM
Serine	Subtilisin-like-PfSUB1	Mediates RBC rupture and invasion	Chloroisocoumarin (JCP104), MRT12113 alpha-ketoamides Hydrazone 2	18 µM 0.3 µM 1 µM 20 µM
	PfSUB2	Release of MSPs during invasion	Propeptide PfSUB2PD	300nM
	PfSUB3	Indirectly in motility, virulence & immune evasion	PMSF	NA
	gp76	Involved in invasion by degrading band 3 and glycophorin A		Not available
	Rhomboid-1	Cleaves transmembrane region of proteins involved in invasion	TCLK,PMSF,DCI	Not available
	Rhomboid-4	Cleaves transmembrane region of proteins involved in invasion	TCLK,PMSF,DCI	Not available

Table 1.4 continued…

Metallo	Falcilysin	Oligopeptide cleavage	1,10 phenanthroline	25-29 µM
	PfM18AAP	Release of Hb derived peptides and interaction with spectrins and membrane proteins	Phosphonic dipeptide Phosphonic analogs	(Ki)4.4 µM (Ki) 11 nM
	PfA-M1	Aminopeptide cleavage and reinvasion	Bestatin, Compound 4 Phosphonic analogues Phenyl bestatin 1 Phosphonopeptides	8-14 µM 24-62 µM (Ki) 11 µM 3.2 µM 14 µM
	PfLAP	Aminopeptide cleavage	Methyl ester bestasin, Nitrobestasin Bestatin Aminoalkylphosphonate Phosphinate dipeptide analog	8 µM 20.5 µM 25µM 14 µM 12-40 µM
	MetAP1b	Protein biogenesis	XC11	112nM
	MetAP2	Protein biogenesis	Fumagillin (Fumarranol) TNP-470	0.20 µM NA
Threonine	HslV/Clp Qorthologue	threonine protease-like activity	Lactacystin and Bortezomib	NA
	Proteasome	proteasome activity	Epoxomicin MG132 peptidosulfonyl fluoride Lactocystin&MLN-273	MIC 17nM 47.6 nM IC20 nM NA

*** Data is compiled through literature search. GBII32=Amide-containing Diphenyl urea. MLA=morpholinocarbonyl-leucinehomophenylalanine aldehyde**

1.3.3 Serine proteases

Serine proteases are a group of proteases characterized by the presence of serine within the catalytic triad along with histidine and aspartic acid [110]. They are optimally active in non-acidic environment and are thus not found in acidic compartments like the food vacuole in parasites. In human, serine proteases carry out certain physiological roles especially in digestion, blood clotting and complement system [111].

Classification of serine proteases: Serine proteases constitute 17% of total proteases identified in *P.falciparum* genome. The representatives from 8 families are identified consisting of a total of 16 proteases (Table 1.5). This includes families of chymotrypsin (S1 family), subtilisin (S8 family), signal peptidase (S16, S26A and S26B families), peptidase clp (S14 family), prolyloligopeptidase (S9 family) and rhomboid proteases (S54 family) [60].

Mechanism of action: Generally, serine proteases activity involves three catalytic residues (Asp, His and Ser) forming a catalytic triad. The oxygen atom (OH) of serine acts as a neocleophile which is activated by aspartic acid and histidine of the triad. This allows attack on the carbonyl atom of the incoming substrates. Then one half of the cleaved incoming substrate is released while the remaining is still attached to the catalytic

serine. The next nucleophilic attack is activated by water molecule and by expense of the water molecule; the other half of the substrate is released as shown in Figure 1.7.

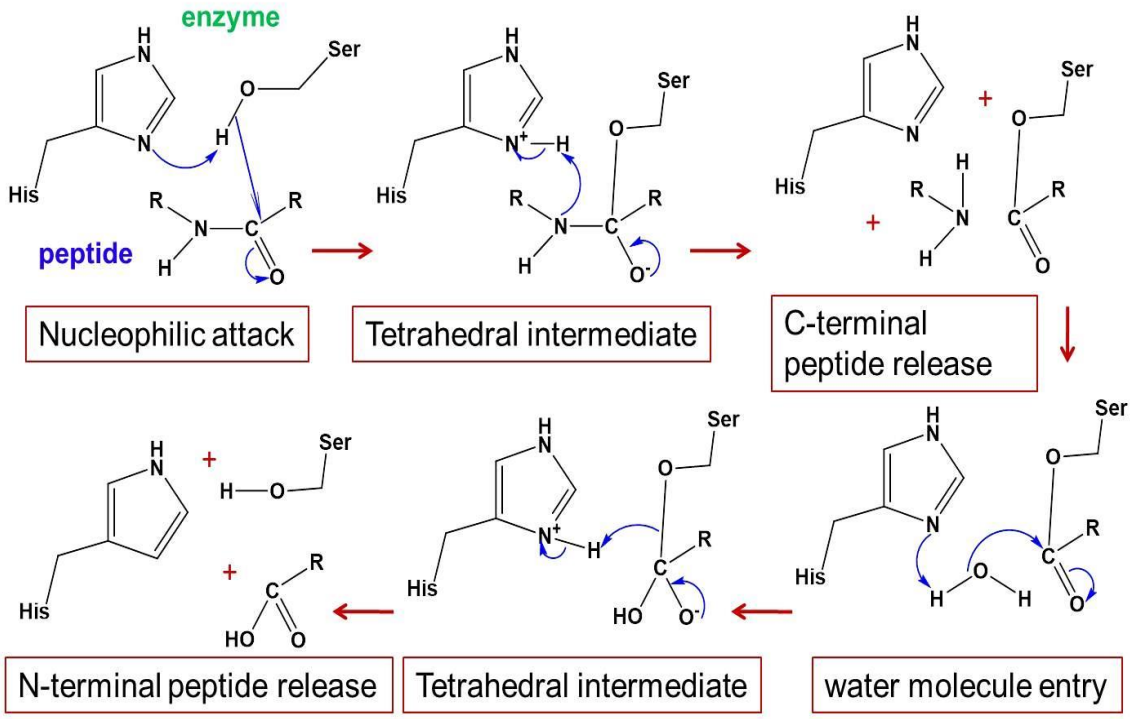


Figure 1.7. Generalized mechanism of action of serine proteases. The catalytic serine is activated by histidine which enables the initiation of nucleophilic attack and release of peptides and aided subsequently by water molecule.

Physiological roles of serine proteases: Clinical manifestations of malaria are caused by the asexual (RBC stage) life cycle when schizont ruptures and releases daughter merozoites to invade fresh RBCs. Studies with broad spectrum protease inhibitors have strongly implicated parasite protease activity in the process. Serine proteases take major part in this process of rupture and invasion which are discussed below;

(i) Invasion: The genome of *P.falciparum* encodes three subtilisin like serine proteases; PfSUB1, PfSUB2 and PfSUB3. Among them PfSUB1 has multiple roles during RBC stages [112]. PfSUB1 mediates RBC rupture by maturing the protein of SERA family (Figure 1.8). Inhibition studies have also implicated the involvement of PfSUB1 in invasion of new RBCs [113-115]. PfSUB1 accumulates in exonemes in daughter merozoites which is discharged into the PV, where it rapidly mediates activation of SERA-4, 5, 6, and other PV-resident SERA family members by proteolytic cleavage. It also cleaves a number of merozoite surface proteins MSP1/6/7 prior to invasion [115-117]. Recently, it was identified that it is essential for complete development of the liver stage schizont and production of merozoites [118].

Table 1.5. Members of the serine proteases in *P.falciparum* genome*

Serine protease family	Protease name	PlasmoDB ID	Characterization status
S1	DegP protease	MAL8P1.126	No
	Neurotypsin-like	PF14_0067	No
S8	Subtilase 1	PFE0370c	Yes
	Subtilase 2	PF11_0381	Yes
	Subtilase-like	PFE0355c	No
S9	ACPH	PFC0950c	No
S14	Clp	PFC0310c	No
	Clp	PF14_0348	No
	ClpB	PF08_0063	No
	ClpB	PF14_0063	No
	ClpC	PF11_0175	No
S16	Lon	PF14_0147	No
S26A	SP1	PF13_0118	No
S26B	Signalase	MAL13P1.167	No
S54	Rhomboid 4	PFE0340c	Yes
	Rhomboid 3	MAL8P1.16	No
	Rhomboid 1	PF11_0150	Yes
	Rhomboid 6	PF13_0241	No
	Rhomboid 7	PF13_0312	No
	Rhomboid 8	PF14_0110	No
	Rhomboid 9	PFE0755c	No
	Rhomboid 10	MAL6P1.241	No

*** Data is compiled from plasmodium database (PlasmoDB.org) and through literature searches.**

PfSUB2 was recently identified as the 'sheddase' responsible for the release of merozoite surface proteins during erythrocyte invasion. It is localized in the microneme of merozoite. Upon merozoite release it is secreted onto the parasite surface and translocated to its posterior pole in an actin-dependent manner, a trafficking pattern predicted of the sheddase. It has potent shedding activity for PTRAMP, MSP1 and AM1 [119-122]. The protease gp76 has been suggested to be involved in invasion by degrading band 3 and glycophorin A [123]**.** Another member of serine proteases i.e., rhomboid proteases are membrane proteins responsible for processing of adhesins involved in attachment, invasion, motility, immune evasion etc [124, 125]. Recently a new merozoite organelle known as the mononeme was identified that contains *P. falciparum* rhomboid-1 (PfROM1), a protease that cleaves the transmembrane regions of proteins involved in invasion [126]. Rhomboid 1 and 4 are shown to cleave diverse transmembrane adhesions such as TRAP, CTRP, MTRAP, PFF0800c, EBA-175, BAEBL, JESEBL, MAEBL,AMA1, Rh1, Rh2a, Rh2b, and Rh4 [127]. PfROM4 was localized to the sporozoite surface, could be the possibly the main factor for the function. Thus they are involved in all invasive stages of parasite life cycle [127]. Rhomboid proteases are also involved in parasitophorous vacuole formation which was tested in *P.yoelii [128]*.

(ii) ***Egress:*** Serine protease PfSUB1 is a versatile protease playing various roles in hepatic stage and merozoite invasion of fresh RBCs. Recently it was reported to also have a key role in egress of merozoites in calcium dependent manner which was confirmed by mutagenesis and inhibition study [114, 118, 129-131]. Besides invasion and egress, PfSUB3 was found to have specificity towards profilin which is a multifunctional protein, suggesting PfSUB3 to be involved in the processes of motility, virulence and immune evasion [132, 133]

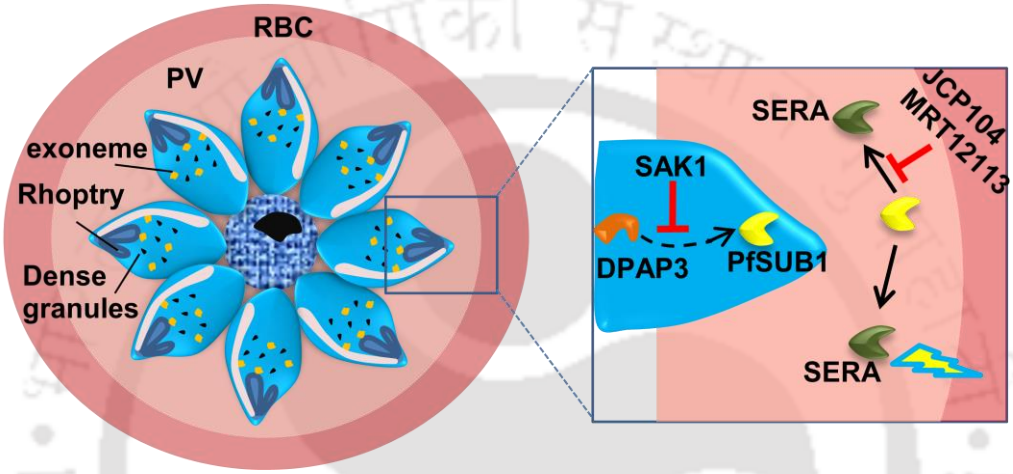


Figure 1.8. Role of serine protease (PfSUB1) during egress of merozoites. PfSUB1 indirectly mediates RBC rupture by maturing cysteine proteases (SERA family). Blocking of either PfSUB1 or SERA using specific inhibitors results in blocking of egress.

Therapeutic potentials of serine proteases: Serine proteases in *P.falciparum* are mostly uncharacterized and not many inhibitors have been synthesized as compared to cysteine and aspartic proteases. However, potent inhibitors have been identified for some serine proteases of essential roles (Table 1.4). PfSUB1 inhibitors result in blocking egress and invasive capabilities of merozoites [113]. Selective PfSUB2 inhibitor also blocks the shedding of MSP1 and AM1 effectively thus hindering invasion. Recent findings on rhomboid inhibitors against PfROM1 and 4 have also shown significant effects on invasion and sporozoite motility which ultimately inhibit sporozoite infection to host cells [127]. PfROM1 is also shown to play important role in formation of parasitophorous vacuole [134]. Owing to these potent effects of inhibitors on parasitic growth and development, serine proteases are also in the limelight to be used as drug targets.

1.3.4 Threonine proteases:

Threonine proteases are new class of proteases unidentified until the discovery of autocatalytic property in 20S proteasome [135]. Proteasomes are the best known

threonine proteases which utilize a particular N-terminal threonine residue and are found in all eukaryotes and archaea bacteria predecessor ClpQ/hslV is present. They represent an important check point for protein quality control by degrading the used or misfolded proteins into amino acids for synthesis of new ones [136]. Proteins that are damaged, misfolded or have fulfilled their physiological role are tagged by a multi-ubiquitin chain which in turn is recognized and bound by subunits of the 19S regulatory particle of the 26S proteasome. The regulatory particle deubiquitinates the target protein unfolds it and threads the linearized polypeptide chain into the 20S core particle of the proteasome where it is degraded (Figure 1.9). These proteasome and its prokaryotic predecessor constitute the T1 family of threonine proteases as T1A and T1B respectively.

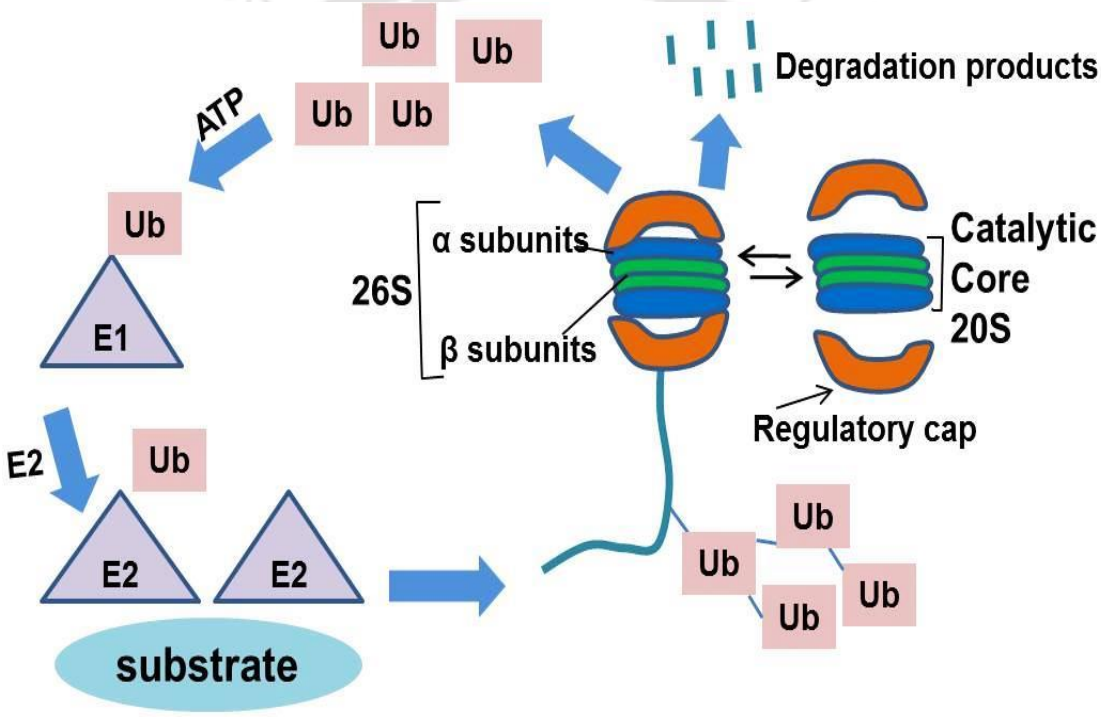


Figure 1.9. Schematic representation of proteasome degradation pathway. The 20S and the regulatory cap form the catalytic 26S system. The unused unbiquitinated protein is recognized by regulatory cap and degraded by the 20S core releasing the ubiquitins.

Features of threonine proteases and their roles in *P.falciparum*: Threonine proteases constitute 14% of the total proteases. Exceptionally, *P.falciparum* genome possesses both eukaryotic proteasome and prokaryotic predecessor (PfhslV in *P.falciparum*) which are encoded by 14 genes and one gene respectively [137]. The proteasome is composed of 14α subunits and 14β subunits whereas the PfhslV is composed of two subunits (Table 1.6). The former is expressed throughout the life cycle while the latter is expressed only at schizont and merozoites. Although activity of threonine proteases is detected in

P.falciparum, the individual component of threonine proteases in the proteosome complex is still uncharacterized except for PfhslV [60, 138, 139]. The prokaryotic homologue Pfhs1V protease is found to be involved in growth and development of mitochondria in *P.falciparum* which is evident from the mutagenesis experiment. Mutation in the protease blocked normal development and function of mitochondria. This partially characterized PfhslV/PfCIpQ protease is ATP dependent threonine protease and is localized in the cytosol [140]. Recently a ubiquitinin proteasome complex of *P.falciparum* was reported to be involved in melatonin signaling in co-ordination with atypical kinase PfPK7 [141].

Table 1.6. Members of the threonine proteases in *P.falciparum* genome*

Threonine protease family	Protease name	PlasmoDB ID	Characterization status
T1	Proteasome α1	PF14_0716	No
	Proteasome α2	MAL6P1.88	No
	Proteasome α3	PFC0745c	No
	Proteasome α4	PF13_0282	No
	Proteasome α5	PF07_0112	No
	Proteasome α6	MAL8P1.128	No
	Proteasome α7	MAL13P1.270	No
	Proteasome β1	PFE0915c	No
	Proteasome β2	MAL8P1.142	No
	Proteasome β3	FA0400c	No
	Proteasome β4	PF14_0676	No
	Proteasome β5	PF10_0111	No
	Proteasome β6	PFI1545c	No
	Proteasome β7	PF13_0156	No
	PfhslV	Not assigned	Yes

*** Data is compiled from plasmodium database (PlasmoDB.org) and through literature searches.**

Therapeutic potentials of threonine proteases: Threonine proteases are new class of protease identified in plasmodium genome and most of which remains uncharacterized. However, evidences of threonine proteases with multiple functions are identified through inhibitors study showing its potentiality as a drug target [139, 142, 143]. The only characterized PfhslV/PfCIpQ protease is a novel drug target candidate that has no homolog in the human host. Silencing of PfHslV reduced the growth of parasites [144]. Its expression in schizont and merozoite implicates its role in RBC stages. Its roles in hepatic and sexual stages have also been implicated by its response to proteasome inhibitors during these stages [145]. Treatment of parasites with specific threonine protease inhibitor such as epoxomicin, MG132 and lactacystin blocked catalytically active proteasome subunits and delayed parasitic development. This led to the accumulation of ubiquitinated proteins and finally to parasite death suggesting a crucial role played by proteasome [138, 144, 146]. A new class of proteasome inhibitors

peptidosulfonylflorides also showed effective growth inhibition against multi-drug resistant and chloroquine sensitive *P.falciparum* parasites [147]. All these reasons interest researchers to further study the insights of threonine proteases for development of inhibitors against these proteases for use as antimalarial drugs.

1.3.5 Metalloproteases

Metalloproteases represent a class of protease (22%) is the second largest protease class after cysteine proteases. Metalloprotease hydrolyzes peptide bond by the action of water molecule which is activated by forming a complex with bivalent metal ions. Most metalloproteases are characterized by presence of catalytic zinc ion although some enzymes may exploit manganese, cobalt, nickel and copper ions. In some metalloproteases, two metal ions act co-catalytically. The metal ion is complexed by three conserved amino acid residues; His, Asp, Glu or Lys [148]. According to MEROPS database, metalloproteases are grouped in 14 different clans which altogether form 98 families.

Classification of metalloproteases: In plasmodium genome, the metalloproteases are not yet assigned any clan and they are classified under different families [60]. The metalloprotease families so far identified comprises: M1, M3, M16, M17, M18, M22, M24A, M24B, M41 and M76 (Table 1.7). The M1 family consists of alanylaminopeptidase (AMPN) which has been characterized. M3 family comprises of dipeptidyl-carboxypeptidase (Dcp) and neurolysin both of which are still uncharacterized. M16 family consists of mitochondrial processing peptidases MPPa, MPPb, M16 peptidase, insulysin, falcilysin and pitrilysin out of which falcilysin has been characterized. M17 family is represented by leucylaminopeptidase (AMPL) which has been recently characterized. Aspartylaminopeptidase (DNPE) represents the M18 family which is recently characterized. M22 family consists of one uncharacterized enzyme glycoprotease (GCP). M24A family consists of four methionine aminopeptidase (AMPM) 1a, 1b, 1c and 2 out of which AMPM2 is partially studied. M24B family has one enzyme aminopeptidase P (AMPP) which in partially characterized. M41 family comprises of four FtsH homologues out of which FtsH peptidase with gene ID PFL1925w is characterized. Besides the metalloproteases identified by Yimin Wu and team, there is another putative metalloprotease family M76 obtained from plasmodium database. MEROPS database also includes a number of putative peptidase and non-peptidase homologues which are not assigned any identity and do not have any locus in the

plasmodium genome [149]. There were only two metalloproteases (AMPN and AMPM) characterized as reported by Yimin and team nevertheless with rapid progress in research, another four metalloproteases (PfLAP, PfAPP, PfM18AAP and FtSH) have been characterized recently and are under study (Table 1.7).

Mechanism of action: Most of the putative metalloproteases in plasmodium utilize zinc ion for its activity while some are not assigned the type of metal they bind. Interestingly, only one metalloprotease is shown to bind manganese ion which is the M17 leucylaminopeptidase (PfLAP) as listed in Table 1.7.

The M16 family of plasmodium represented by insulin degrading enzyme (pitrilysin) is one of the zinc ion binding metalloproteases. The M16 family is characterized by an inverted active site motif of HXXEH, which is the inversion of the well-known zincin metalloprotease thermolysin. Thus they are categorized as inverzincins under zinc metalloproteases [150]. In the zinc binding motif, the two histidine and a glutamic acid residue located 75–84 residues C-terminal to this motif are involved in coordinating the catalytic zinc ion [151]. Inverzincins and thermolysin follow similar mechanism of action for activity where two histidine and one glutamic acid in the motif HXXEH are involved in coordinating with zinc ion (Figure 1.10). Another glutamic acid initiates the action by deprotonating water molecule that completes the coordination sphere of the zinc. Other residues that generally support the catalytic complex include tyrosine, threonine and arginine [152].

Physiological roles of metalloproteases: Metalloproteases mostly have signal peptides targeting into the mitochondria as obtained from plasmodium database. This shows that they are involved in processing of mitochondrial proteins. There are few members that have mitochondrial signal peptides but also form a cellular component in food vacuole and cytosol as shown in Table 1.7. Some enzymes are also shown to have both mitochondria and apicoplast signal peptides which suggest their involvement in dual organelles proteins processing. There are certain putative crucial roles of metalloproteases as predicted by plasmodium database which are listed in Table 1.7. Based on the characteristic features, metalloproteases are found to perform several functions from hemoglobin digestion to host proteins interaction to reinvasion. The various physiological roles of metalloprotease are discussed below;

(i) Hemoglobin digestion: The M16 family is divided into sub family A and B. Sub family A comprises the oligoendopeptidases involved in processing biologically important peptides. Sub family B includes the organelle processing peptidases such as

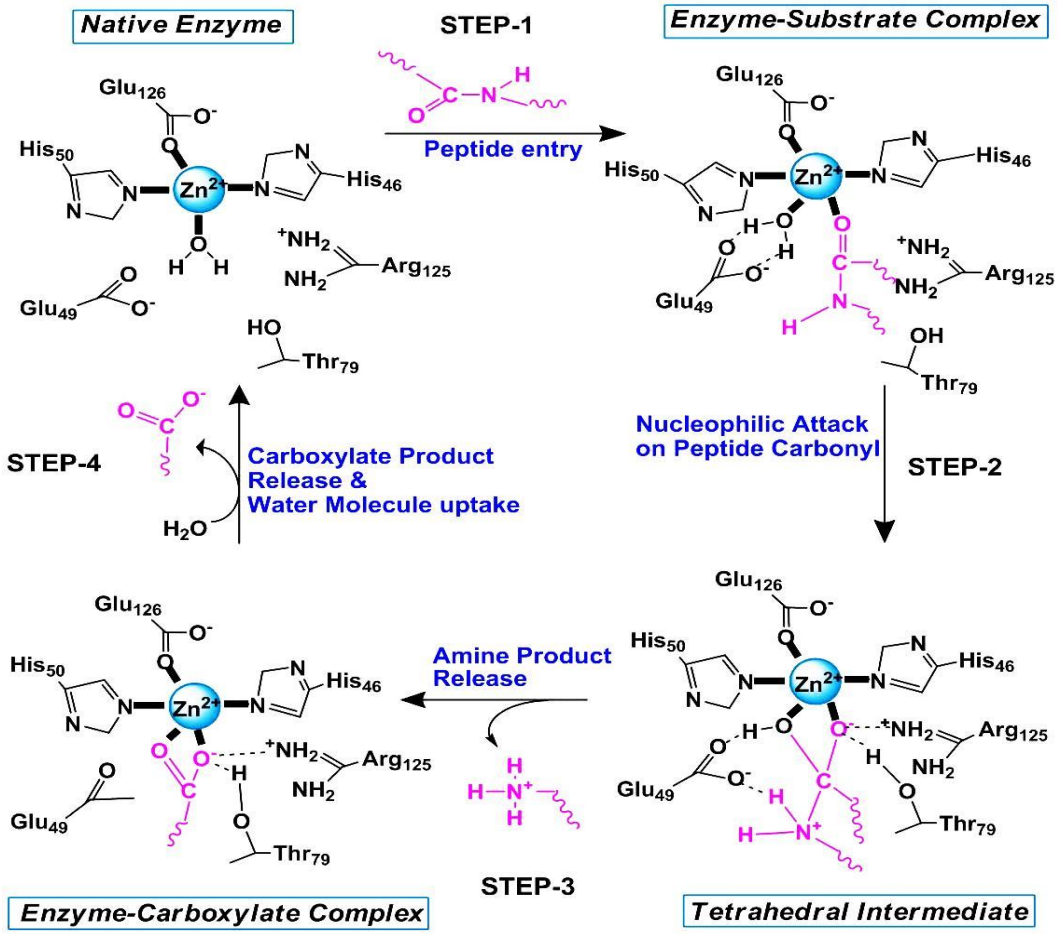


Figure 1.10. Mechanism of action of inverzincin family of metalloproteases. The zinc binding motif HXXEH and water molecule co-ordinate with zinc ion. A catalytic residue glutamic acid nearby the motif is positioned by water molecule and initiates a nucleophilic attack on peptide carbonyl which releases the cleaved peptides at the expense of water molecule.

mitochondrial processing peptidase (MPP) and the chloroplast stromal processing peptidase; both enzymes are involved in removing signal peptides. A zinc metalloprotease falcilysin is the only characterized protease of M16 family belonging to subfamily A. Falcilysin prefers to act on smaller peptides up to 20 amino acids in length (Figure 1.5). The globin peptides that arise from degradation of hemoglobin by plasmepsins and falcipains serve as substrates for this enzyme [154]. A metalloprotease PfM18AAP is suggested to play a role in the release of host hemoglobin-derived peptides into the cytosol (Figure 1.5) [155]. It is expressed in the cytosol and exported outside the parasite between the PVM and the parasite membrane [156]. Aminopeptidases such as M1 alanylaminopeptidase (PfA-M1), aminopeptidase P (PfAPP) and M17 leucylamino peptidase (PfLAP) are also involved in terminal hemoglobin digestion inside the food vacuole while PfAPP and PfLAP are implicated to perform digestion outside the food vacuole as well (Figure 1.5). Inhibition studies show that PfA-M1 mediated cell death is

due to disturbance in hemoglobin digestion whereas PfLAP mediated cell death is outside food vacuole suggesting another role. PfA-M1 and PfAPP are localized in the food vacuole of the parasite while PfLAP in the cytosol [157-166].

Table 1.7. Members of the metalloproteases in *P.falciparum* genome along with their putative pathway and activity*

Family	PlasmoDB gene ID and name	Signal peptide and cellular component	Putative Pathway	Predicted activity	Character-istic status
M1	MAL13P1.56	Api Mt Cellular component of Nuc & FV	1,2,3, 4,5	M1-family alanylaminopeptidase	Yes
M3	MAL13P1.184	Mt	6,7,8,9	Mitochondrial intermediate processing peptidase. M3A unassigned.	No
M16	MAL5P1.231, PFE1155c	Mt	2,4,5,7,8,10	MPP alpha subunit putative	No
	PFI1625c	Mt	5,6,7,8,10	OPP/MPP beta subunit	No
	PF11_0189	NA	5,8	insulinase, putative	No
	PF11_0226	NA	8,11	subfamily M16A peptidase	No
	PF13_0322	Api, FV	2,3,4,5,8	Falcilysin, Hb digestion	Yes
	PF14_0382	Api	5,8	CPP/SPP	No
M17	PF14_0439	Api	2,3,4,8,12	M17 leucylamino-Peptidase (PfLAP)	Yes
M18	PFI1570c	Vacuole	2,3,5,8	M18AAP aspartylaminopeptidase	Yes
M22	PF10_0299	NA	5,8	glycoprotease, putative	No
M24A	MAL5P1.272, PFE1360c,	NA	3,8,13,14	Methionylaminopeptidase (MetAP1a)	No
	MAL8P1.140	Api	3,8,9,14	methionylaminopeptidase 1 (MetAP1c)	No
	PF10_0150	NA	2,3,8,13,14,	Methionylaminopeptidase 1b (MetAP1b)	No
	PF14_0327	NA	3,8,13,14	methionylaminopeptidase 2 (MetAP2)	Yes
M24B	PF14_0517	Api & Mt Cellular component of Nuc & FV	2,3,4,8	Aminoacylprolineaminopeptid-ase (PfAPP)	Yes
M41	PF11_0203	Mt	5,8,15	m-AAA protease	No
	MAL12P1.383, PFL1925w	Mt	4,8	FtsH, putative; Tg AFG3 ATPase family protein.	Yes
	PF14_0616	Mt	8,15,16	ATP-dependent protease la, putative, family M41 unassigned peptidases	No
M76	PF14_0396	Mt	11,17	Unknown	No

*** Data is compiled from plasmodium database (PlasmoDB.org) and through literature searches. Api=apicoplast. Mt=mitochondria. Nuc=nucleus. FV=food vacuole. OPP=organelle processing peptidase. CPP=chloroplast processing peptidase. SPP=stromal processing peptidase. Putative pathways are annotated in numeric forms for convenience;**

1) ER-to Golgi, translocation and quality control. 2) S-Glutathionylated proteins. 3) Hb digestion & ferriproto-porphyrin IX polymerization. 4) Total palmitome of *P.falciparum*. 5) Zn ion binding. 6) Topogenesis of the Rieske-FeS protein. 7) Import of proteins into the Mt. 8) Peptidases and proteases. 9) Metal binding NA. 10) Proteins of detergent-resistant membranes. 11) Metal ion binding. 12) Mn ion binding. 13) Co-translational cleavage of N-terminal Met residues & N-terminal acetylation. 14) Protein biosynthesis. 15) Chaperone network and protein quality control of Mt. 16) Pathways of peptide export from the Mt. 17) M76, putative unassigned peptidases.

(ii) Invasion: A metalloprotease PfM18AAP is suggested to play a role in the release of host hemoglobin-derived peptides and is also characterized to interact with RBC spectrins and other membrane proteins, suggesting its role in invasion step in RBC stages [155]. It is evident by its presence in the cytosol and exported outside the parasite between the PVM and the parasite membrane. PfA-M1 have been described to be involved during reinvasion [158]. Recent studies also proposed the localization of PfA-M1 outside the food vacuole where it is trafficked via the parasitophorous vacuole and marginally delivered to the food vacuole [167].

(iii) Peptide degradation: The sub family B of M16 class includes uncharacterized organelle processing peptidease. Falcilysin is one enzyme with multifunctional capabilities. Immunofluorescence microscopy reveals two locations of falcilysin inside and outside the food vacuole which is confirmed to be the apicoplast, suggesting a role in trans-peptide degradation [168].

(iii) Protein biogenesis: Metalloproteases are believed to be involved in protein biogenesis as shown in Table 1.8 however, exact pathway is not characterized. Methionine metalloprotease (MetAP1b and MetAP2) inhibitors showed decrease in proliferation of parasites in-vitro and in-vivo suggesting its roles in protein biogenesis as given in Table 1.7. [169-171]. Little is known about organelle biogenesis in parasite and there is a protease analogue of AAA+/FtsH in *P.falciparum* (PfFtsH) and it is believed to be involved in organelle biogenesis [171].

Therapeutic potentials of metalloproteases: The roles of metalloproteases in parasite are extensive and crucial but they are unexplored so far. Inhibitors of characterized metalloproteases already result in parasite deaths or decrease in proliferation rate with potent IC_{50} values as shown in Table 1.4 [172-178]. Most of the metalloprotease with potential functions remains uncharacterized. Even the most well studied hemoglobin degradation pathway is not clearly concluded at the terminal part which might involve metalloproteases (Figure 1.5). Protein processing or maturation of proteins and proteases is least studied process in parasite. The M16 family is one of the promising groups with putative protein processing for mitochondria and apicoplast but most of which remained unexplored. Furthermore, the molecular mechanism involved in the severity of malaria is not fully understood.

1.3 Significance and aim of the work

RBC stages have long been a research interest as all clinical symptoms arise during the asexual blood cycle. Chemical drugs therapy has been the only remedy used lately. However, failure to treatment due to development of resistance by parasites to all the existing drugs becomes a challenge to search for new drugs. Emergence of resistance to the most effective drugs i.e., artemisinin, has also been recently reported [30]. The limitations of antimalarial chemotherapy underscore the need for new drugs.

Malarial proteases have long been considered potential targets for chemotherapy due to their crucial roles in RBC stages of the parasitic life cycle from invasion through egress. Further, comparative genome analysis [60] revealed large number of putative potential proteases, most of which remained uncharacterized so far. Metalloproteases are one class of proteases in *P.falciparum* involving in all aspects of parasite life cycle such as hemoglobin digestion, interaction with RBC membrane proteins, protein biogenesis etc. Recent findings showed blocking of invasion by metalloprotease inhibitor 1,10-phenanthroline suggesting its role in invasion [179].

Among the metalloprotease class, matrix metalloproteases are so far remain uncharacterized in *P.falciparum.* They are believed to be one of the key factors inducing pathogenesis of severe malaria. They are involved in the modification of microvascular endothelium in human host which is induced by parasite hemozoin [180].

Besides information obtained from literature, plasmodium database has listed various putative functional roles of metalloproteases which are crucial to parasite survival. The curated metabolic pathways include protein processing, protein modification and trafficking, protein biogenesis, protein quality control, protease and peptidase functions.

PFI1625c is a putative organelle processing peptidase present in *P.falciparum* genome. Protein expression is observed throughout the erythrocytic stages predicting a role in erythrocytic stages (www.plasmoDB.org). Literature studies and *in silico* comparison implicates that it belongs to M16 sub family B of metalloprotease [60]. Plasmodium database has listed a number of putative pathways for PFI1625c which include proteins of detergent-resistant membranes, topogenesis of the Rieske-FeS protein, import of proteins into the mitochondrion, nuclear genes with mitochondrial signal sequences, peptidases and proteases and Zn ion binding. Thus its role in processing the organelle proteins for degradation of transit peptides or cleavage of other proteins in the RBC stages as well as its role in egress and reinvasion (as evident from metalloprotease inhibition study) would help study the effect on parasite survival. The crucial roles played

by PFI1625c as a matrix metalloprotease in parasite survival has raised our research interest on PFI1625c as a potential target and developing sustainable or non-toxic drugs. So the proposed potential pathway of PFI1625c during the erythrocyte stages is represented as a schematic diagram in Figure 1.11.

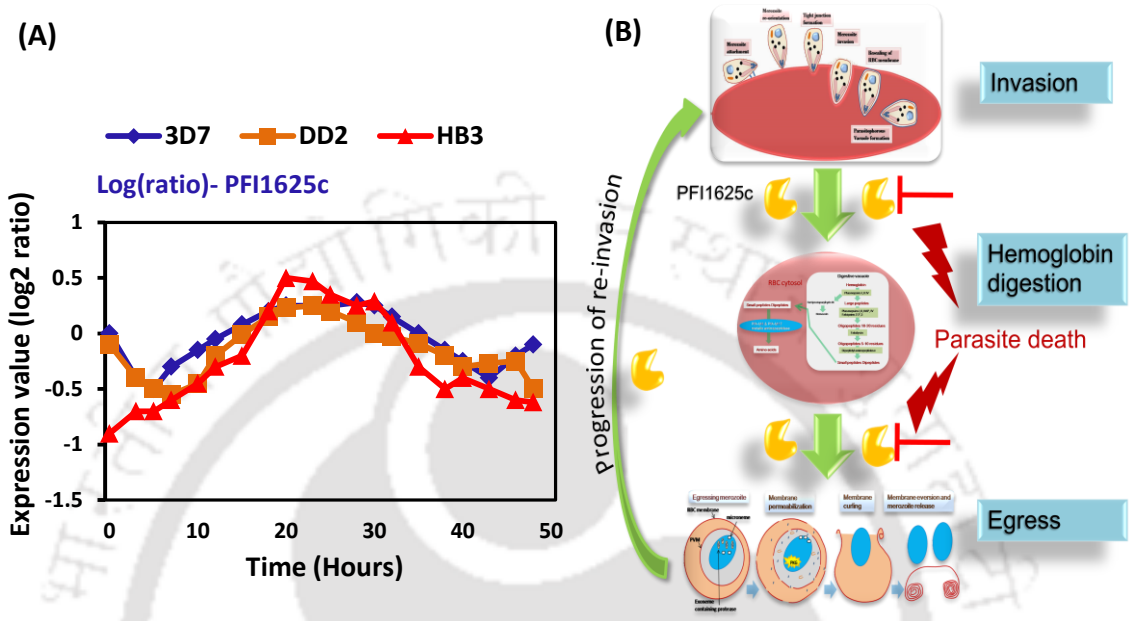


Figure 1.11. Proposed role(s) of PFI1625c within parasite. (A) Expression pattern of PFI1625c from three *P.falciparum* strains by whole genome micro array analysis (date is obtained from plasmodb.org and replotted. **(B)** Based on preliminary experiments, the proposed pathways for PFI1625c.

Based on the preliminary study of PFI1625c, the work presented in this thesis is therefore divided into the following objectives;

1. **Cloning, over-expression, purification and preliminary localization study of PFI1625c from *P.falciparum* 3D7**
2. **Structural characterization of putative PFI1625c from *P.falciparum* 3D7**
3. **Biochemical characterization of PFI1625c from *P.falciparum* 3D7**
4. **Designing and screening potent inhibitor against PFI1625c and understanding the mode of inhibition**

References: Please refer to bibliography section from page number 108-119.

Chapter-2: Cloning, Over-expression, Purification and Preliminary Localization study of PFI1625c from *P.falciparum* 3D7

Cloning, Over-expression, Purification and Preliminary Localization study of PFI1625c from *P.falciparum* 3D7

2.1 Introduction

Proteases are a class of enzymes which take part in all major functions in the life cycle of malaria parasite. With the ability of the parasite to confer resistance towards existing drugs, there is a constant need to develop new drugs. Proteases are potential targets for anti-malarial drug development as they perform critical events such as invasion, digestion and egress. PFI1625c is a putative protein found in the genome of *P.falciparum*. Preliminary studies showed that it may be similar to organelle processing peptidase, but detail biochemical and functional role of this protease is not explored in *P.falciparum*. Characterization of PFI1625c will help us to understand the role(s) of this protein in the life cycle of parasite and allow us to exploit it as target for anti-malarial drug development. In this chapter, I have discussed cloning, over-expression in bacterial expression system and affinity purification from the over-expressing cells. In addition, we tried to discuss the preliminary immuno-localization study of PFI1625c and its functional relevance in *P.falciparum* 3D7.

2.2 Experimental procedures

2.2.1 Isolation of genomic DNA: Malaria parasite *P.falciparum* 3D7 was cultured using candle-jar technique of Trager-Jansen [181]. The infected RBCs (iRBC) were centrifuged at 800g for 5mins at room temperature. The RBC pellet was re-suspended in 3 volumes of ice-cold phosphate buffer saline (PBS) containing 0.05% saponin and left on ice for 10mins. The parasite was pelleted at 4000g for 5mins and the supernatant was discarded. The pellet was washed twice with PBS and after the last wash, re-suspended in 1/100 volume of original PBS. Proteinase K (final concentration of 10µg/ml) and SDS (final concentration of 0.5%) was added and the mixture was carefully mixed by inverting the tube slowly. The mixture was incubated for 2-12hrs at 56-65°C and 1 volume of Phenol:Chloroform:Isoamyl alcohol (PCI) was added to the mixture. The tubes were inverted for approximately 10mins until a complete mixture of the organic and aqueous phase occurred. The mixture was centrifuged at 15000g for 10mins. The upper aqueous phase was transferred into a new tube and PCI addition step was repeated. To the aqueous phase, 1 volume of chloroform was added and both phases were mixed carefully for 5mins and centrifuged at 15000g for 5mins. The aqueous phase was transferred to a new

tube and 1/10 volume of 3M Na acetate pH 4.5 and 2.5 volume of 95% ethanol were added. The tube was inverted several times to mix the solution completely and frozen at -20°C for at least 1hr. The mixture was then centrifuged at 15000g for 10mins and supernatant was discarded. Ice cold ethanol (500-1000µl of 70%) was added and vortex quickly. The mixture was centrifuged at 15000g for 5mins and supernatant was discarded and the pellet was air dried. The dried pellet was dissolved in TE buffer. The quality of the DNA was checked in 0.5% agarose gel.

2.2.2 Amplification of PFI1625c gene from genomic DNA: PFI1625c gene was amplified from *P.falciparum 3D7* genomic using site specific forward (5'-GAA**GGATCC**ATGTGGAAAAGAAAAGTGGTAAACGT-3) and reverse primers (5'-AAT**CTCGAG**AGTCTTTTGCCTGAGGTCGATGTAC-3). The PCR condition was set for 35 cycles with initial denaturation at 94°C for 3mins followed by 10 cycles of (94°C for 1min, 50°C for 1min, 65°C for 4min) and 25 cycles of (94°C for 1min, 55°C for 1min, 72°C for 3min) and a final extension at 72°C for 10mins. The PCR product was analyzed on 0.8% agarose gel. The reaction components and condition are described in Appendix-I.

2.2.3 Cloning of PFI1625c in pTZ57R/T cloning vector: The amplified product of PFI1625c (1.4 kb) was eluted from the gel using gel extraction kit (Himedia). The detail protocol of gel extraction is given in Appendix-I. The eluted PFI1625c gene was cloned in pTZ5R/T vector by following manufacturer's instructions [182]. The ligation mixture was transformed in *E.coli* DH5α cells and grown overnight at 37°C. Plasmids were isolated from transformed colonies by lithium chloride method and checked for amplification of PFI1625c by PCR using site specific primers. The PCR product was checked on 0.8% agarose gel. The positive colonies were selected for double restriction digestion analysis using BamHI and XhoI restriction enzymes by incubating for 3hrs at 37°C. The digestion product was checked in 0.8% agarose gel.

2.2.4 Cloning of PFI1625c in *E.coli* expression vector: The digested product of PFI1625c from pTZ57R/T vector was eluted using gel extraction kit (Himedia). The pET22b expression vector was pre-digested with BamHI and XhoI by incubating for 3hrs at 37°C. The PFI1625c gene was ligated into the pET22b vector by T4 DNA ligase (Fermentas) by incubating overnight at 4°C. The plasmids were transformed in *E.coli* BL21 (DE3) cells and grown overnight at 37°C. pET22b plasmids were isolated from the cells and checked for amplification using colony PCR. A double restriction digestion was

carried out using BamHI and XhoI from the colonies which show amplification in PCR. PFI1625c was also cloned in pET23a expression vector using the same strategy as described for pET22b. The components and conditions of ligation, transformation and restriction digestion are described in Appendix-I.

2.2.5 Optimization of PFI1625c over-expression in *E.coli* expression system:

The pET22b plasmid containing PFI1625c gene was transformed in different *E.coli* strains (BL21 (DE3), Codon Plus, C41(DE3) and Rosetta to check suitable host for over-expression in respective selective media. Transformed BL21 (DE3), and C41(DE3) required only ampicillin (50µg/ml) as selective antibiotic while Codon Plus and Rosetta required ampicillin (50µg/ml) and chloramphenicol (35µg/ml) as selective antibiotic. The transformed colonies were inoculated in 5ml LB containing selective antibiotics and incubated at 37°C till OD reached 0.5-0.6.The cells were harvested and cell lysate was checked in SDS PAGE and expression of PFI1625c was detected in western blotting using anti-his antibodies and chemiluminescent peroxidase substrate. Optimization of IPTG concentrations (0.1-1mM), induction time (3-11hrs) and induction temperature (25-37°C) were also optimized for transformed BL21(DE3) cells.

2.2.6 Western blotting for detection of PFI1625c: Transformed bacterial cells were centrifuged at 12000 rpm for 1min. The pellet was dissolved in loading buffer and heated at 95°C for 5 to 10mins. Samples were first run in 10% SDS PAGE (in duplicate), one of the gels was stained with coomassie stain and the other gel was used for probing PFI1625c with anti-His antibodies in western blotting [183]. The protein samples in the gel were blotted on to nitrocellulose membrane at 70V for about 5hrs or 120V for 3hrs in cooled condition. Membrane was stained with ponceau S to check the transfer of protein bands. The membrane was blocked with 5% skim milk overnight at 4°C or 90mins at 37°C. The blot was probed with anti-His antibody (sigma) at 1:3000 dilutions and incubated overnight at 4°C. The membrane was washed 4-5 times with wash buffer containing 0.05% Tween-20. Secondary antibody (HRP-conjugated anti-mouse antibody) was added at 1:2000 dilutions and incubated over-night at 4°C. The membrane was washed 4-5 times with wash buffer containing 0.05% Tween-20. The blot was developed either by using colorimetric substrate diaminobenzidine (DAB) or chemiluminescent peroxidase substrate. For colorimetric substrate, DAB solution (Appendix-II) was poured on the membrane in dark for 10mins or until the appearance of brown color bands. The

chemiluminescent substrate was poured on the membrane in dark and observed with Molecular Imaging System (Gel Logic 1500).

2.2.7 Optimizing solubility of PFI1625c: *E.coli* strains BL21(DE3), Codon Plus and Rosetta transformants were grown and induced at 18°C for 24hrs using 0.8mM IPTG to express PFI1625c. The cultures were transferred to 2ml eppendorf and pelleted by centrifuging at 12000rpm for 1min and re-suspended in 0.2ml of lysis buffer (Appendix-II). The culture suspension was sonicated for 2mins with an interval of 10sec pulse off with amplitude of 30%. The bacterial lysate was centrifuged at 13000rpm for 20mins and supernatant was collected. The whole cell, supernatant and pellet were loaded in SDS PAGE to check the solubility of PFI1625c. Solubility was optimized for BL21(DE3) using different pH (4-9) of lysis buffers. Western blotting was performed to confirm the presence of PFI1625c.

2.2.8 Denaturation and Refolding of PFI1625c: PFI1626c was expressed well in BL21(DE3) cells but it was expressed present in inclusion bodies. Solubility was minimal even in optimal condition so, to overcome the solubility problem, we tried whether denaturating and refolding might give us native active protein. For this, BL21(DE3) transformants were grown in 1L Luria Bertani (LB) media to obtain sufficient protein in inclusion bodies. The bacterial culture was centrifuged at 6000 rpm for 10mins. The cell pellet was re-suspended and sonicated in lysis buffer for 10mins with an interval of 10secs pulse off with amplitude of 32%. The sonicated lysate was centrifuged at 4°C for 30mins. The inclusion bodies containing the PFI1625c were solubilized in either 8M urea or Tris buffer pH 12 containing 2M urea at 40°C for about 30mins with constant agitation. The solution was centrifuged down at 12,500 rpm for 30mins and the supernatant containing the solubilized PFI1625c (~15ml) was dialyzed against (2L) 10mM Tris buffer pH 8.5 at 4°C in cold room. The buffer was changed 4-5 times or until complete removal of urea was achieved. The protease activity was monitored to follow the refolding process. The refolded protein was present in native condition as checked by gel filtration chromatography.

2.2.9 Purification of PFI1625c: PFI1625c was purified by Ni-IDA affinity chromatography. Firstly nickel-IDA affinity column was prepared by packing 1-2 ml of iminodiacetic acid (IDA) sepharose matrix (GE Healthcare) in a column. The matrix was washed with distilled water to remove ethanol used for storage. About 10 ml of nickel sulfate (100mM) was passed to the matrix to charge the IDA-sepharose matrix. The

column was equilibrated with 5 column volume of binding buffer, Tris buffer pH 8.8 containing 250mM NaCl. The refolded protein mixture (filtrated) was loaded onto the column. The column was washed with binding buffer containing 20mM of imidazole to remove non-specific proteins bound to column. The bound PFI1625c was eluted using step gradient of imidazole (100mM, 200mM and 250mM) in the elution buffer (Appendix-II). The purity of eluted fractions was checked in 10% SDS PAGE.

2.2.10 Determination of native molecular weight: The oligomeric status or native molecular weight of the purified protein is determined by size exclusion chromatography using gel filtration column (AKTA Purifier FPLC). The gel filtration column was pre-equilibrated with degassed Tris buffer pH 8.5 containing 100mM NaCl. Standard proteins such as bovine serum albumin, ovalbumin, conalbumin, hemoglobin and carbonic anhydrase were obtained from LMW gel filtration calibration kit (GE Healthcare). About 2ml of each standard protein (1mg/ml) was injected into the sample loop and was run at a flow rate of 0.2ml/min during loading. Afterwards, flow rate was increased to 0.5ml/min and chromatography was performed at room temperature. The running conditions were kept identical for the standard proteins as well as PFI1625c. Blue Dextran was also run to determine the void volume. The elution volume for each protein was recorded (Table 2.1). To determine the native molecular weight of PFI1625c, calibration curve was drawn between log molecular weight (Log M_r) and distribution co-efficient (K_{av}). The K_{av} of each protein was calculated using the formula given in equation (1).

$$K_{av} = \frac{(Ve - Vo)}{(Vt - Vo)} \quad \text{..... (1)}$$

Where, Ve = elution volume of protein sample, Vo = void volume of gel filtration column and Vt = total volume of gel filtration column.

The K_{av} of PFI1625c was calculated using equation (1) as given below:

$$K_{av} \text{ of PFI1625c} = \frac{(244-110)}{(330-110)} = 0.705263$$

The sub unit molecular weight of PFI1625c was calculated from 10% SDS PAGE. Purified PFI1625c fractions were run under denaturing condition in 10% SDS PAGE. Pre-stained protein marker (Fisher Scientific) of known molecular weight (33-135kDa) was also run along with PFI1625c fractions. The migration distance for each protein band

present in the ladder and PFI1625c was measured to obtain relative mobility (R_f) values. The R_f value was calculated using the formula given in equation (2).

$$Rf = \frac{\text{migration distance of protein}}{\text{migration distance of dye front}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

A graph was plotted between R_f values and log molecular weight (Log M_r) of marker proteins and the sub unit molecular weight of PFI1625c was calculated from the equation obtained from the graph. The oligomeric status of PFI1625c was determined by comparing the native molecular weight (obtained from gel filtration) and sub unit molecular weight obtained from 10% SDS PAGE.

Table 2.1 Native molecular weight and elution volume of standard proteins

Standard protein	Molecular weight (kDa)	Elution volume (ml)
Conalbumin	75	223
BSA	66	260
Hemoglobin	65	235
Ovalbumin	43	255
Carbonic anhydrase	29	270
Blue Dextran	2000	110
PFI1625c	52*	244

* value obtained from calibration curve

2.2.11 Antibody generation: For antibody generation, large amount of pure PFI1625c protein was required. To achieve this, concentrated partially purified PFI1625c was loaded in 10% SDS PAGE in a large well formed by fusing 5 wells. A small strip of PAGE was excised from the unstained gel and was stained to detect the position of PFI1625c. PFI1625c region was excised from the unstained gel and placed in dialysis bag filled with 5-7 ml of SDS PAGE running buffer. The dialysis bag was placed in horizontal electrophoresis tank filled with SDS PAGE running buffer (Appendix-II). The protein was electro eluted from the gel by performing electrophoresis at 80-100 volts overnight under cold condition to avoid heating. The gel piece was removed from the bag and the eluted protein was dialyzed to remove SDS and then lyophilized. The protein powder was then saved for antibody development. The eluted protein was checked in 10% SDS PAGE and western blot using anti-His antibody. Injection of antigen into the female rabbit (New Zealand strain) as animal model for antibody production was carried out at NIPER, Guwahati, in collaboration with Dr. M. Lahkar. Purified PFI1625c protein was dissolved in sterile water at a concentration of ~400µg/0.5ml. The first antigen injection was carried out by preparing an emulsion of PFI1625c protein with Freund's

complete adjuvant using a syringe (Appendix-II). The hair at the trunk region of rabbit was removed to have better access to the skin for injection. The exposed skin was swabbed with 70% alcohol to avoid infection at the site of injection. The emulsion was injected sub-cutaneously at different spots in the exposed skin of the rabbit. On the 32nd day post injection, booster injection was given by preparing the emulsion of PFI1625c in Freund's incomplete adjuvant (Appendix-II). On the 10^{th} day of booster injection, blood was collected from the ear vein in a sterile tube. To obtain a clear serum, the collected blood was kept at 4°C overnight in order to settle the RBC and avoid lysis. After overnight incubation at 4°C, serum was obtained as a clear supernatant. Using the clear serum, ELISA was performed to determine the antibody titer. Next day, blood was drawn from the animal through cardiac puncture. Serum was collected as clear supernatant by settling the RBC at 4°C overnight. The serum was dispensed in 1.5ml eppendorf and was stored at -80°C.

2.2.12 Estimation of anti-PFI1625c antibody titer: The titer value of generated anti-PFI1625c antibody was estimated by Enzyme-linked Immunosorbent Assay (ELISA) as described [185]. The serum collected from the rabbit was diluted to 1:100, 1:500, 1:1000, 1:5000, 1:10000, 1:20000, 1:40000 and 1:80000 using PBS (Appendix-II). PFI1625c protein was dissolved in bicarbonate buffer pH 9.2 at a concentration of 25µg/ml. Protein solution (50µl) prepared in coating buffer (bicarbonate buffer pH 9.2), was coated in ELISA plate wells and incubated for 8-10hrs at 4°C. The unbound protein was removed by washing gently with wash buffer (PBS containing1% tween-20) and the wells were blocked with 5% skim milk for 2hrs at 37°C. The excess skim milk was washed off with wash buffer. The diluted serum (50µl) serum was added to the wells and incubated for about 2hrs at 37°C. The unbound serum was washed off with wash buffer 3-4 times. Secondary antibody (HRP-conjugated anti-rabbit antibody) was added to the wells and incubated for about 2hrs at 37°C. The unbound antibody was removed by washing 3-4 times with wash buffer. HRP substrate O-Phenylenediamine (OPD) was prepared as described in Appendix-II. About 100µl each of the substrate solution was then added to the wells in dark condition and enzymatic reaction was allowed for 10mins to develop color. The optical density of the color was measured at 460nm in spectrophotometer (Spectramax M2, Molecular Devices). Dot blot was performed to further confirm the presence of anti-PFI1625c antibody in the serum using pure PFI1625c

samples. Western blotting was performed to check the specificity of the produced PFI1625c-antibody.

2.2.13 Dot blot for analyzing generated anti-PFI1625c antibody: For confirmation of anti-PFI1625c antibody in serum of rabbit, PFI1625c samples from different sources were used for blotting. Pure PFI1625c was used as +ve control and BSA was used as –ve control. Samples for detection are blotted on nitrocellulose membrane in a small volume at a time using microtip in order to get intense dot. Each blotting was followed by intermittent drying using hair dryer (Panasonic). After blotting was completed, it was air dried completely at 37°C for 30mins. The dry membrane was blocked with 5% skim milk in PBS for overnight at 4°C. The membrane was washed with wash buffer containing PBS and 0.05% tween-20 to remove excess of skim milk. Generated primary anti-PFI1625c antibody at 1:1000 dilutions was added onto the membrane and incubated at 37°C for 2hrs. The primary antibody was washed off 4-5 times with wash buffer with proper shaking for 5-10mins each. Secondary antibody (HRP-conjugated anti-rabbit antibody) was added onto the membrane and wrapped in polyethene and incubated at 37°C for 2hrs. The membrane was washed 4-5 times with wash buffer. DAB (1mg/ml) was poured immediately onto the membrane in dark condition and allowed it for about 10mins or until brown coloration developed.

2.2.14 Stage specific immuno-localization of PFI1625c: Synchronous *P.falciparum* culture was harvested at about 3-5% parasitemia and smeared on a glass cover slip followed by fixing with methanol at room temperature. The cover slips were placed in 6-wells plate and hydrated with PBS at pH 7.4. The cells were permeabilized on cover slips using 0.1% triton x-100 prepared in PBS for 30mins at room temperature. Permeabilized cells on cover slips were washed 4-5 times with PBS for complete removal of triton x-100. The cells were blocked with 5% bovine serum albumin (BSA) in PBS for 30mins at room temperature. Smears were subsequently incubated over night at 4°C with anti-PFI1625c antibody at 1:50 dilutions (in 2% BSA prepared in PBS). The smears were washed thoroughly with PBS at room temperature to remove unbound or excess antibodies. After washing, anti-rabbit IgG labeled with fluorescein isothiocyanate (FITC) was added on to the cover slips at a dilution of 1:500 and incubated for 5-6hrs at 4°C. The cover slips were washed thoroughly with PBS at room temperature. The cover slips were mounted on a glass slide with mounting media containing DAPI (4',6-diamidino-2-phenylindole). DAPI fluorescence was observed under fluorescence microscope (Nikon

Eclipse 80i) with excitation wavelength and emission wavelength 350nm and 470nm respectively. FITC fluorescence was observed under excitation and emission wavelength 495nm and 519nm respectively.

2.3 Results

2.3.1 PFI1625c gene was cloned in *E. coli* expression vector

PFI1625c gene amplification from genomic DNA of *P.falciparum* shows an amplified product of ~ 1.4 kb (Figure 2.1A). The amplified gene product was sequenced and it matches well with the nucleotide sequence of PFI1625c given in plasmodium genome database (http://www.plasmodb.org/plasmo/showRecord.do?name=GeneRecordClasses.GeneRecordClass&project_id=&primary_key=PFI1625c). Cloning of the amplified DNA fragment in pTZ57R/T cloning vector showed a product of approximately 1.4kb DNA band in PCR as well as double restriction digestion from colonies 16, 17 and 22 (Figure 2.1B and C). Subsequent sub-cloning in pET22b expression vector also showed amplification of approximately 1.4 kb DNA from colonies 3, 18, 32 and 34 (Figure 2.1D). Double digestion of pET23a and pET22b recombinant plasmid gives1.4 kb digestion product as observed (Figure 2.1E and F). Undigested pET23a and pET22b did not show any band. The Results in Figure 2.1 confirms that PFI1625c gene was cloned in *E.coli* expression vectors (pET23a and pET22b).

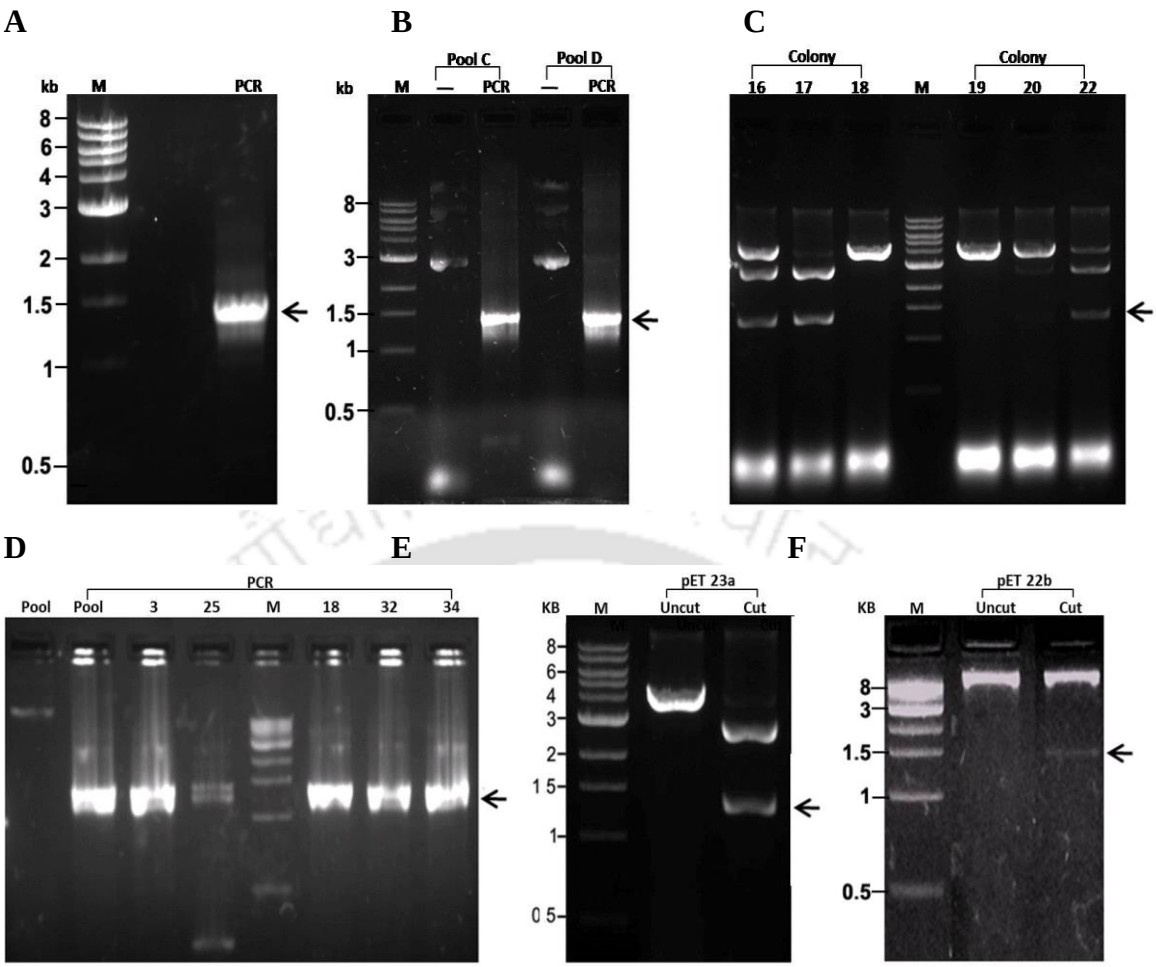


Figure 2.1. Cloning of PFI1625c into the *E.coli* expression vector. (A) PCR amplification of PFI1625c from *P.falciparum* genomic DNA. **(B)** Confirmation of PFI1625c cloning in pTZ57R/T vector. Different colonies were picked from the transformed LB-amplicillin plate and grown overnight. A 5ml bacterial culture was pooled together (indicated as PoolC and poolD for convenience. **(C)** Double digestion of pTZ57R/T vectors from different colonies. **(D)** PCR product of PFI1625c from pET22b vector. Pool is a combination of different colonies assigned as c3-c34. **(E)** Double digestion product of PFI1625c from pET23a vector showing uncut pET23a and cut pET23a. **(F)** Double digestion product of PFI1625c from pET22b vector showing uncut pET22b and cut pET22b.

2.3.2 PFI1625c over expressing well in BL21 (DE3) and Codon Plus

The theoretical sub unit molecular weight of PFI1625c as obtained from plasmodium database is 55.7kDa whereas recombinant PFI1625c is ~57kDa. The over expression of PFI1625c in pET22b expression vector using different *E.coli* strains shows that significant expression of PFI1625c gene is observed in induced Codon Plus and BL21 (DE3) strains and no expression observed in un-induced cells as observed in SDS PAGE and western blot (Figure 2.2A). Low expression is also observed in Rosetta strain. Since the expression level was similar and yield of bacterial pellet is less in codon plus, we have preferred BL21 (DE3) for subsequent study. Western Blot result showed maximum

expression of BL21 (DE3) at 0.8mM IPTG concentration (Figure 2.2B). Induction for 11h showed highest expression than the lower induction time points (Figure 2.2C) and lower temperature such as 25°C was more preferred than higher temperatures (Figure 2.2D). Hence, for over-expression of PFI1625c, 0.8mM IPTG, 11hrs induction at 25°C are the optimal conditions to obtain best yield and production in bacterial expression system.

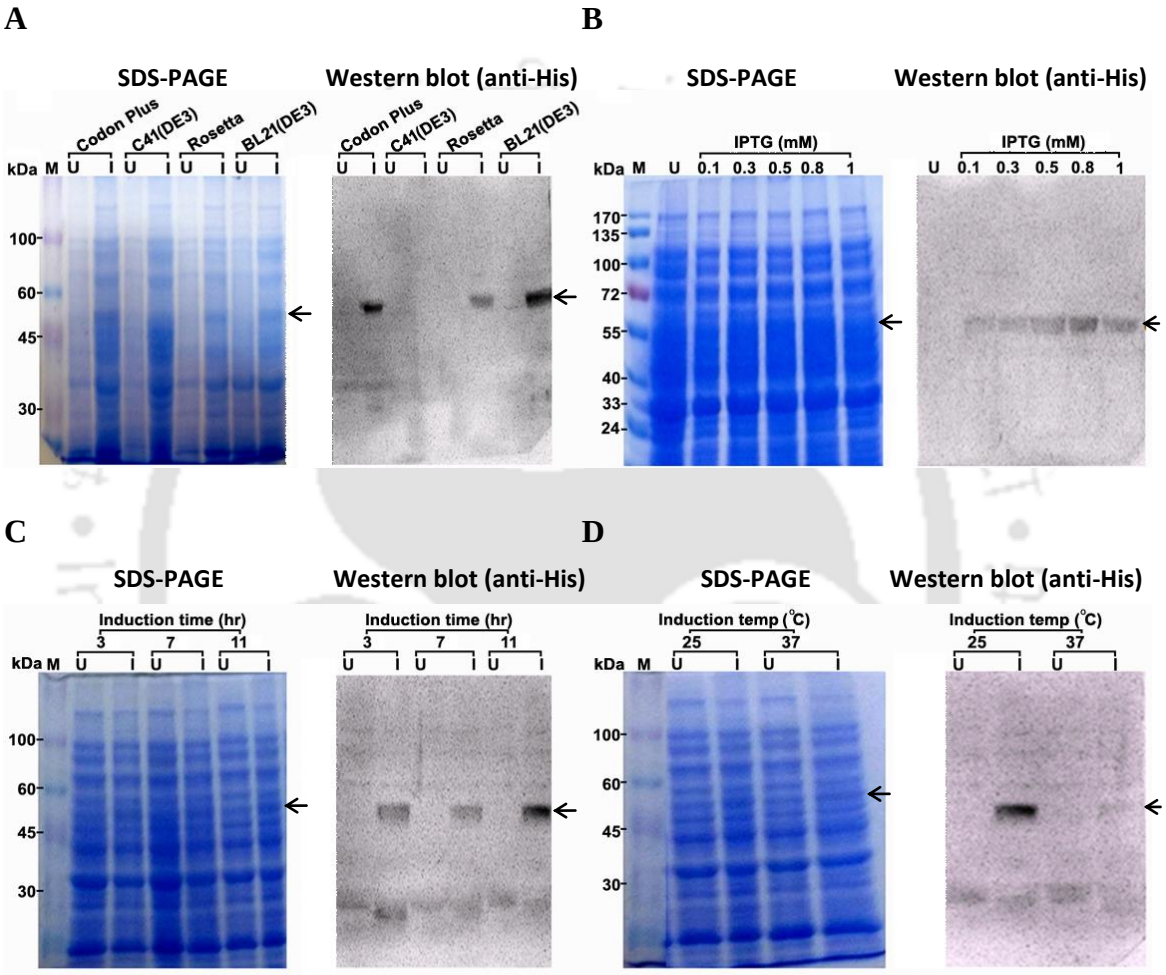


Figure 2.2. Optimization of different parameters for over-expression of PFI1625c in *E.coli* expressing system. (A) Expression analysis in different *E.coli* strains. **(B)** IPTG curve of PFI1625c expression using BL21(DE3) strain. **(C)** Induction Time curve with BL21(DE3) expressing PFI1625c cells. **(D)** Induction Temperature curve with BL21(DE3) expressing PFI1625c cells. Both SDS-PAGE and Western blot profiles are shown. Arrow indicates the band of PFI1625c. U=uninduced; I=induced.

2.3.3 PFI1625c is not present in the soluble fraction

Solubility test of PFI1625c in BL21(DE3), Codon Plus and Rosetta showed that PFI1625c was expressed as inclusion bodies as shown in SDS PAGE which was further confirmed by western blotting. Cell lysate also showed presence of PFI1625c but the protein band was absent in supernatant (Figure 2.3 A and B). Optimization of lysis buffer

pH did not improve solubility which was confirmed by western blot (Figure 2.3 C-E).

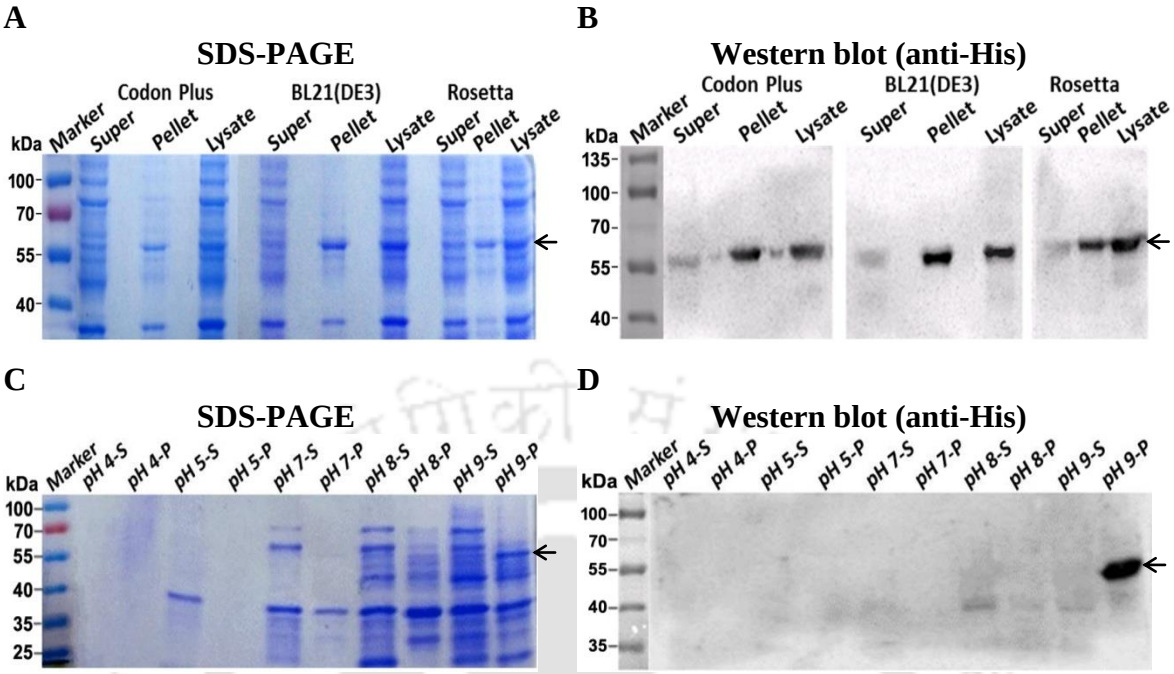


Figure 2.3. Testing of different conditions for obtaining soluble PFI1625c. Arrow indicates PFI1625c band. (A) SDS PAGE for solubility test of PFI1625c in different *E.coli* strains. **(B)** Western blot of (A). **(C)** Solubility optimization using lysis buffer of different pH. S=supernatant after sonication, P=pellet after sonication. **(D)** Western blot of (B).

2.3.4 PFI1625c is a monomeric protein

Dialyzed refolded PFI1625c protein bound to Ni-IDA affinity column was eluted using increasing concentration of imidazole in the elution buffer. The eluted fractions were analyzed in 10% SDS PAGE. PFI1625c started eluting at 100mM imidazole along faint non-specific band. Pure prominent protein band was observed in the fractions containing 200mM imidazole which is the size of PFI1625c protein of ~57kDa (Figure 2.4). This shows that PFI1625c protein was purified to homogenity. The yield of PFI1625c protein after purification was ~1mg/L. The native molecular weight determination using gel filtration chromatography showed that PFI1625c eluted at 244ml, region corresponding to the elution volume and molecular weight between hemoglobin and ovalbumin (Table 2.1). Elution profile of PFI1625c is shown in Figure 2.5A. The molecular weight obtained from 10% SDS-PAGE and gel filtration was 55.5kDa and 52kDa respectively. The molecular weight of sub unit PFI1625c was calculated from the SDS PAGE and it matches that of monomeric status of gel elution (Figure 2.4B and Figure 2.5B). This shows that PFI1625c is a monomer which is supported by the fact that monomeric molecular weight of recombinant PFI1625c is 57kDa and it eluted between hemoglobin

(65kDa) and ovalbumin (43kDa). This confirms that PFI1625c is a monomeric protein. The elution of PFI1625c at the expected size also affirms that it is properly refolded into native conformation. The R^2 value obtained from 10% SDS PAGE was 0.99 where the R^2 value obtained from gel filtration was 0.98.

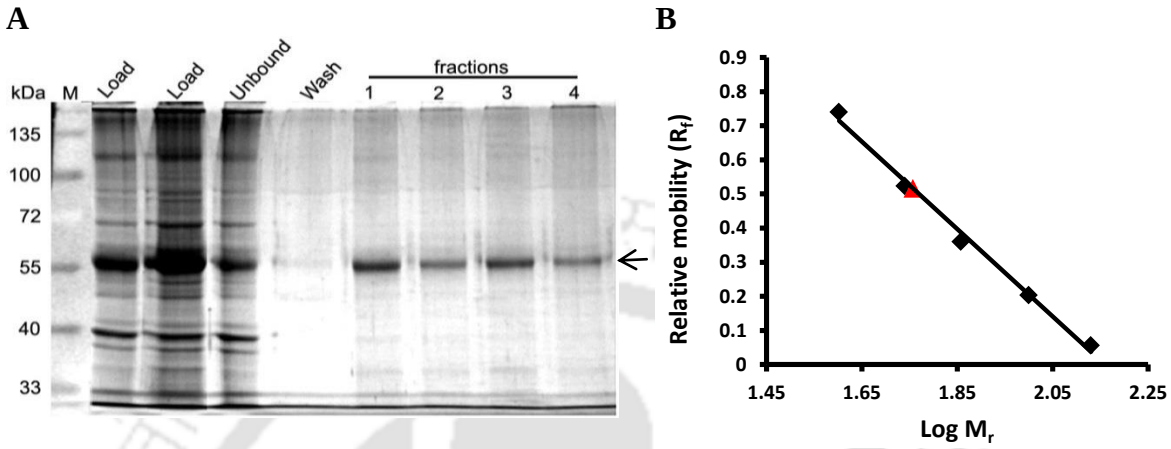


Figure 2.4. Purification and determination of sub-unit molecular weight of PFI1625c. (A) Elution profile of PFI1625c**.** Arrow indicates the purified fraction of PFI1625c which is about 57kDa in size. M=protein marker. Fractions1-4 are 200mM imidazole fractions. **(B)** Molecular weight determination of PFI1625c from graph by plotting R_f values of protein markers and their corresponding molecular weight in log scale. PFI1625c is highlighted in red color triangle.

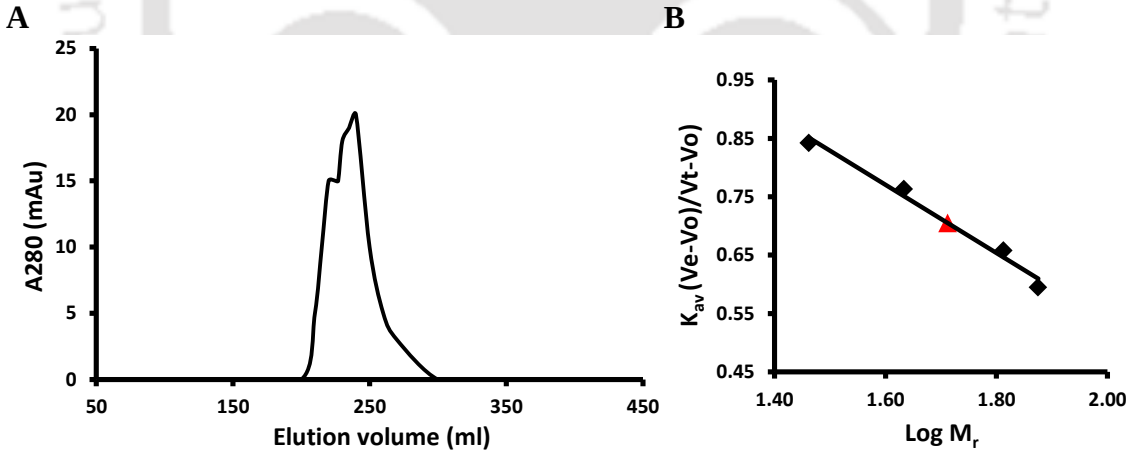


Figure 2.5. Determination of native molecular weight and oligomeric status of PFI1625c. (A) Gel filtration profile of PFI1625c. **(B)** Molecular weight determination of PFI1625c from graph by plotting R_f values of protein markers and their corresponding molecular weight in log scale. The position of PFI1625c is highlighted in red color triangle.

2.2.5 Anti-PFI1625c antibodies were produced with high titer

Electro elution of PFI1625c for antibody production showed single protein band and chemiluminescent signal of 57kDa in 10% SDS PAGE and western blot respectively indicates that PFI1625c protein is eluted successfully (Figure 2.6B). In ELISA the antibody titer was observed at 1:20000 dilutions as shown (Figure 2.6E and F). This

indicates a good titer and presence of high amount of anti-PFI1625c antibody in the serum. Dot blot was performed to detect the presence of PFI1625c antibody using a color substrate which turns brownish upon enzymatic activity. Dot blot shows a brown coloration at the spotted region while negative control BSA did not show any color (Figure 2.6C). Western blotting is done to test the specificity of the produced antibody. Uninduced bacterial lysate was used as control for western blotting. Western blot shows an intense brown color at the region of PFI1625c band in induced sample while uninduced sample shows little or no coloration. The presence of non-specific bands in western blot may be due to excess primary antibody used (Figure 2.6D).

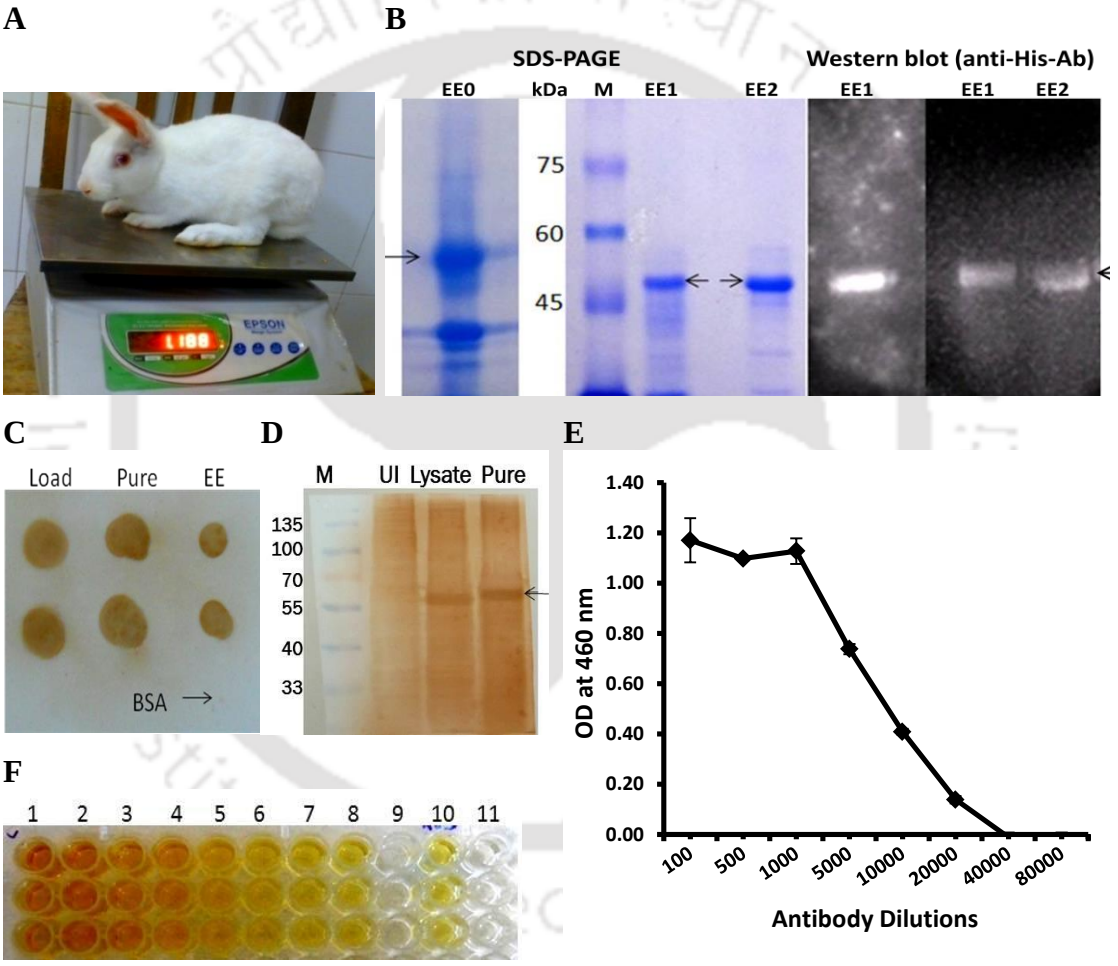


Figure 2.6 Production of anti-PFI1625c antibody in rabbit. **(A)** Model organism (female rabbit, NZ strain)) used in the production of antibody. **(B)** SDS-PAGE and western blot of prepared antigen (PFI1625c). Arrows indicate PFI1625c band. M=protein marker, EE0=crude PFI1625c, EE1 & EE2= electro-eluted pure PFI1625c. **(C & D)** Confirmation of anti-PFI1625c antibody by dot blot and western blot. **(C)** Load=crude PFI1625c, pure=column purified PFI1625c, EE=electro-eluted PFI1625c. BSA did not show any color. **(D)** UI=uninduced transformed BL21(DE3) cells, lysate=induced bacterial lysate. **(E)** ELISA of serum for estimation of antibody titer. **(F)** Color development of ELISA. Number 1-11 denotes antibody dilution of 100, 500, 1000, 5000, 10000, 20000, 40000, 80000, only buffer, only 1˚ antibody and only OPD substrate respectively.

2.2.5 PFI1625c is expressed at all stages and it is a non-nuclear protein

Localization studies help us to initiate our understanding the possible roles of a protein. Proteases that are localized inside food vacuole are involved mainly in hemoglobin digestion whereas those that are localized in the cytosol or parasite membrane are mainly involved in protein processing and egress [78, 133]. It is therefore, important to study the localization of PFI1625c inside or outside the parasite. Using the generated anti-PFI1625c antibody, stage specific immunolocalization study of PFI1625c was performed for ring, trophozoite and schizont stage. Merozoites were also observed in late schizont stage. Bright field (BF) was used to identify the parasitized RBCs and determination of life-cycle stage. FITC fluorescence localized the presence of PFI1625c and DAPI fluorescence tracked the nucleus of the parasite. When observed under fluorescence microscope, fluorescence signals of FITC and DAPI were observed in all the stages where parasite is located. FITC florescence showed that PFI1625c is expressed in all stages of the parasite (Figure 2.7A-D). Individual or segregated merozoites could be seen in late schizonts and they also showed high expression (Figure 2.7D). It was observed that PFI1625c and nuclear fluorescence signals were not co-localizing with each other clearly indicated that PFI1625c may be non-nuclear protein. There was no FITC fluorescence in the food vacuole which is marked by the presence of unstained hemozoin crystal in trophozoite stage and large food vacuole in late schizont stage (Figure 2.7B and D). Wherever there was dark hemozoin pigment, FITC signal was not visible. This shows that PFI1625c is probably not localized in food vacuole. When the localization of PFI1625c was closely observed, FITC signal appeared as a small vesicles which is more obvious in trophozoite stage (Figure 2.7B F1 and F2) and schizont stage (Figure 2.7C F1-F3) which become concentrated in the peripheral region towards the membrane. PFI1625c signal was observed throughout the entire cytosol and was not confined in sub cellular organelles. It is clear that PFI1625c is a non-nuclear protein (may be cytosolic) which might be exported to different location through secretory vesicles.

2.4 Discussion

E.coli is a system of choice for over- expression of heterologous protein for its inexpensive culture conditions and easy to grow. Although serves a good choice for protein expression, over-expression of eukaryotic recombinant protein results in improperly folded proteins or inclusion bodies which becomes a major challenge. Incompatibility of codon usage between the native gene sequence of heterologous protein

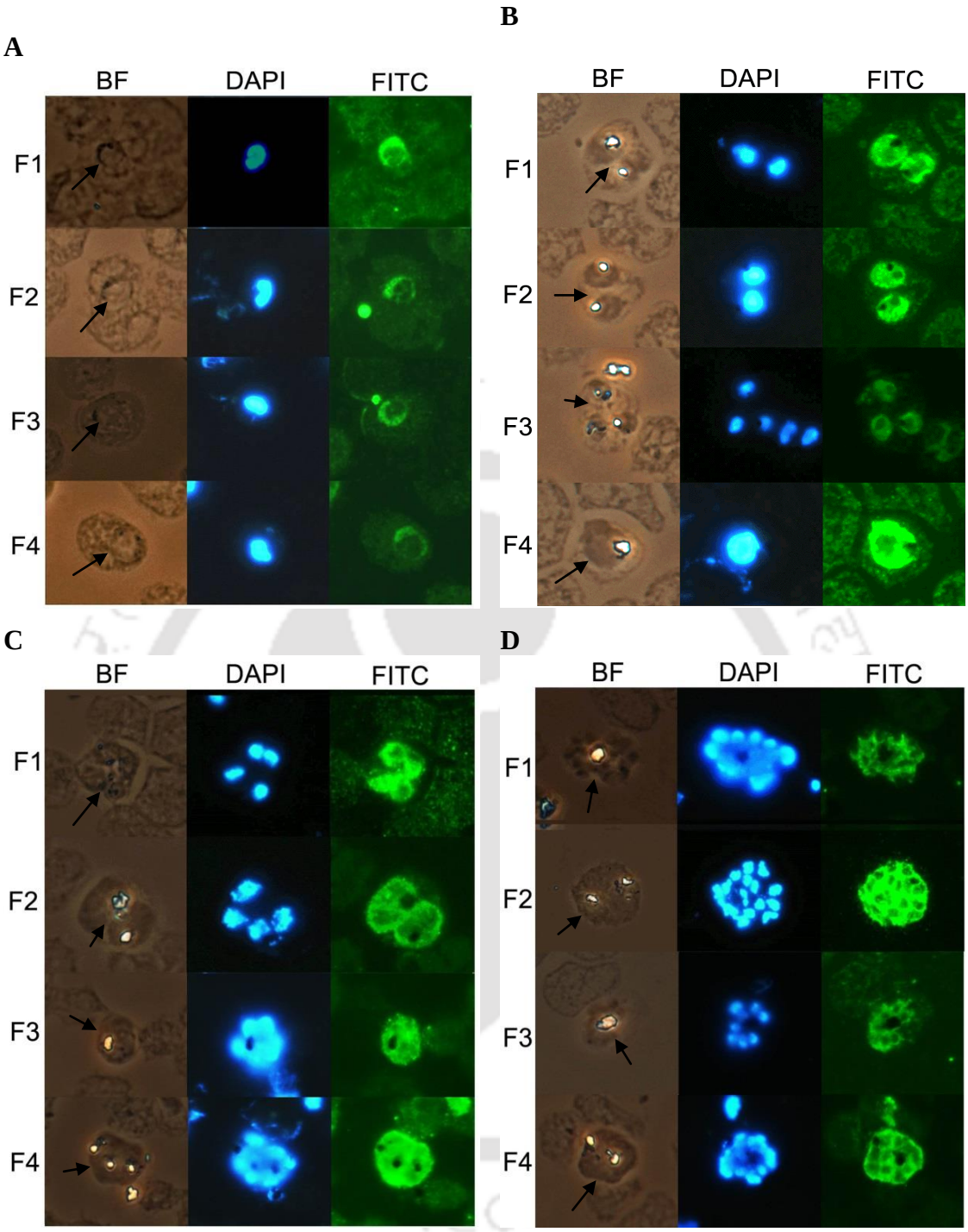


Figure 2.7 Immuno-localization of PFI1625c at different stages. (A) Ring stage where parasite has a single nucleus and no hemozoin is formed; BF=bright field where parasites are visible and highlighted in arrows, DAPI=4',6-diamidino-2-phenylindole that stains the nucleus of the parasites, FITC=Fluorescein isothiocyanate that stains PFI1625c. F1-F4=different fields annotated as 1 to 4. **(B)** Trophozoite stage where hemozoin starts to form and division of nucleus starts taking place in late trophozoite. **(C)** Schizont stage where cells grow and nuclear division continues to take place and parasites occupy the entire RBC. **(D)** Late schizont stage and formation of cell wall for each nucleus takes place to form individual merozoite.

and that of host is the cause of low expression and improper folding. The high A/T versus C/G content and frequent repeats of lysine and arginine repeats in *P.falciparum* genome are thought to be the main reason for early termination in the mRNA translation process [186]. PFI1625c was not present in soluble fraction.

Since pET22b has a pelB leader sequence, PFI1625c protein was believed to might have exported in the periplasmic layer of the bacterial membrane as it was not detected in the soluble fraction. So attempt was made to extract the periplasmic fraction. However, there was no detection of PFI1625c in the periplasm which suggested its presence as inclusion bodies. The presence of PFI1625c as inclusion bodies was confirmed by western blot (Figure 2.3). Standardization of western blot was performed in order to obtain specific signal. Optimum condition was obtained when antibodies were diluted to 1:3000 (primary antibody) and 1:2000 (secondary antibody). PVDF produced lesser background noise compared to nitrocellulose membrane.

Conditions of growth such as temperature, time, pH of media, inducer (IPTG) concentration etc have profound effect on the formation of inclusion bodies. Soluble protein was obtained in several proteins by changing these conditions [186-188]. Optimization of growth parameters was performed to obtain soluble PFI1625c. But these parameters only helps in increasing amount of expressed protein but did not have any improvement in achieving soluble protein. Growth of expressing cells under osmotic stress and exposure to heat stress increases the level of soluble protein. This stress increases the concentration of osmolytes and chaperones which assist in protein refolding and prevent aggregation. Chemical chaperones such as sorbitol, arginine and PEG have been shown to increase solubility of recombinant proteins in *E.coli* [189, 190]. Different concentrations of sorbitol, arginine and PEG were used in growth media (LB) for PFI1625c expressing *E.coli* BL21(DE3) cells but did not improve the solubility of PFI1625c. Heat shock was also given to expressing cells (prior to induction) at 45°C for 30mins but did not give better result.

As attempts to express PFI1625c as soluble protein have failed, expressing BL21(DE3) cells were sonicated by using different pH in the lysis buffer (pH 4-9) to check if there is any improvement in acidic or alkaline environment but not too high to denature the protein native structure. In all the pH range, there was no soluble PFI1625c except for its presence in the sonicated pellet (Figure 2.3D). In the western blot of solubility test for different strains (Figure 2.3A and B) negligible signal was observed in the sonicated supernatant of all the strains. That was due to incomplete centrifugation of the sonicated

sample as subsequent experiments never showed any signal in the supernatant. During optimization of lysis buffer pH, the acidic pH was neutralized before loading in SDS PAGE.

A number of recombinant *P.falciparum* proteases have been studied by denaturing and refolding the inclusion bodies into native structure such as plasmepsin I, II, IV, falcipain, M18 aminipeptidase etc [20, 155, 184]. We also tried to denature PFI1625c inclusion bodies using different denaturants such as 6M guanidine HCl, 2% SDS, 1% β-mercaptoethanol, 8M urea, Tris pH 12.5 or combination of these denaturants. 6M Guanidine HCl was not able to solubilize as compared to 8M urea. Using 8M urea with 2% SDS and 1% β-mercaptoethanol solubilized almost 80% of the inclusion bodies but it hampered the native condition of PFI1625c. Among the denaturants used, 8M urea and Tris pH 12.5 produced best results as the protein was obtained in properly folded condition as evident from the gel filtration fraction. But pH 12.5 containing 2M urea was found to solubilize larger amount of protein and required less time for dialysis so it was routinely used for denaturation. Since our target was to obtain more expressed protein, lengthening culture duration up to 24hrs and low (up to 18°C) were used.

Aggregation of recombinant proteins during refolding is another limiting step. Folding of protein requires natural environment aided by molecular chaperones, pH, temperature, co-factors, salts etc [191]. Reduced/oxidized glutathione, a redox reagent for disulfide bond formation by cysteine residues did not have any effect on PFI1625c solubility and refolding mechanism. This is however, relevant for PFI1625c which does not have disulfide bridges. Different chaotropic salts or additives such as SDS, triton X-100, tween 20, PEG, DTT, metal ions etc in the lysis buffer and denaturants which are reported to facilitate protein folding [192, 193] did not improve refolding of PFI1625c. Aggregation level was similar with or without additives so we processed PFI1625c without additives and metal ions.

Codon optimization of eukaryotic gene has been a great success in solving biased codon usage in *E.coli* expression system. Expression of recombinant proteins in soluble and native conformation has been achieved including malaria vaccine candidate EBA-175 [194]. To enhance expression and solubility of PFI1625c, the gene sequence was submitted to GeneArt® (Life Technologies) for codon optimization. The pMK-RQ plasmid containing synthetic PFI1625c was sub cloned in pET23a expression vector and transformed in *E.coli* BL21(DE3). Expression was highly enhanced however, protein was

not expressed in soluble fraction so it was denatured and refolded but it was functionally not active. Therefore, native protein with low yield was subsequently used.

Purification of PFI1625c was performed by eluting with increasing imidazole concentration from 50mM-250mM. The yield of purified protein after denaturation and refolding was about 1mg/L. Purification of recombinant proteins usually results in low yield as observed in other recombinant parasite proteins such as falcipain which also give the same yield [184].

Localization of PFI1625c showed that PFI1625c was expressed throughout the stages. By comparing with the nuclear stain, PFI1625c was found to be localized outside the nucleus. Food vacuole appeared as dark spot in FITC fluorescence which shows that PFI1625c is not present in food vacuole. Thus it may be a cytosolic protein involved in processing of cytosolic proteins or it may be localized in other cytoplasmic organelles. Several cytosolic proteases are involved in hemoglobin digestion such as metalloproteases PfLAP and PfDAP [157]. Localization of PFI1625c is comparable to a known cytosolic metalloprotease M17 aminopeptidase (PfLAP) which supports our data suggesting the presence of PFI1625c in the cytosol [165]. Similar type of PFI1625c localization is observed in well-known metalloprotease falcilysin (localized in peripheral membrane), SERAs (localized in the parasitophorous vacuole) and in PfSUB2 which is transported through vesicles [121, 151, 195]. A vesicular transport is a well-known mechanism for transport of PfSUB2 [121]. Appearance of PFI1625c in small vesicles during late trophozoite and schizont stage suggests that PFI1625c might be transported through secretory vesicles (exonemes or micronemes). Similar vesicular transport is seen in serine protease PfSUB2 which is mainly involved in membrane rupture and merozoite release [121]. The location of PFI1625c became more peripheral towards the parasite membrane which becomes obvious in the schizont stage. Similarity with falcilysin location as well as SERAs location in the late stage shows that it might be tanslocated during the late stage towards the parasite membrane and into the parasitophorous vacuole [151, 195]. It shows the possible involvement of PFI1625c during egress and invasion step. PFI1625c might also be secreted outside the RBC but it is not possible to detect by localization studies because during smear preparation only RBCs are fixed and other liquid components including culture supernatant are washed off. This detection can be achieved by dot blot of culture supernatant.

References: Please refer to bibliography section from page number 108-119.

2.5 Appendix-I: Materials and Molecular biology Protocols

Materials: Site-specific primers were purchased from Metabion International AG, vectors were purchased from MBI Fermentas and Novagen. Restriction enzymes were from New England Biolabs. LB media, gel extraction silica kit, OPD etc were purchased from Himedia. Agarose IPTG, Nitrocellulose membrane, PVDF membranes, chemiluminescent peroxidase substrate, DAB, antibodies, imidazole, Freund's complete and incomplete adjuvants were purchased from Sigma, St. Louis, MO. Ni-IDA sepharose, column and low molecular weight (LMW) calibration kit for gel filtration were purchased from GE Healthcare. FITC and DAPI were purchased from BD Biosciences. Other chemicals and reagents were of analytical and molecular biology grade.

2.5.1 Competent cells preparation: The competent cells of *E.coli* strains were prepared using PEG-DMSO method as described in [196]. *E.coli* bacterial strains DH5α, BL21(DE3), Codon Plus, C41(DE3) and Rosetta were grown in LB media at 37°C. When OD of the culture reached 0.5, cells were harvested. The cells were re-suspended gently on ice with one-tenth volume of Transformation storage solution (TSS buffer). The re-suspended cells were incubated on ice for 15mins and dispensed as 100µl aliquot in 1.5ml eppendorf tubes.

Recipe for preparation of TSS buffer	
Components	**Amount**
LB broth	85% v/v
Polyethelene glycol	10% w/v
Dimethylsulfoxide	5% v/v
$MgCl_2$	50mM

2.5.2 Isolation of plasmids: Plasmid isolation was performed using LiCl method as described [197]. The bacterial culture was taken and pelleted down at 10000 rpm for 30secs. Then 100µl of TELT solution was added to the pellet and re-suspended by vortexing. Then 100µl of 1:1 PCI was added and vortexed for 5secs and let it stand for ≤ 15mins. The sample was centrifuged at 15000x g for 1min. The upper aqueous layer which contains DNA was transferred carefully in clean eppendorf tube. Then two volume of chilled 100% ethanol was added for washing the plasmid DNA. The DNA was then precipitated and centrifuged at 13000rpm for 5mins. The DNA pellet was air dried to removes traces of ethanol. The pellet was dissolved in 30µl sterile water and heated at 60°C for 10mins and given a short spin. The plasmid was ready to be used.

TELT solution

Components	Stock	Working soln (10ml)
Tris base	1M pH 8	0.5ml
EDTA	0.1M pH 8	6.25ml
Triton X100		0.4ml
LiCl2	10M	2.5ml

2.5.3 Transformation of competent cells: The protocol of transformation was followed as described in [198]. Standard plasmid of known concentration was taken and added in the competent cells aliquot to count transformation efficiency. The mixture was kept in ice for 30mins. Then heat shock was given at 42ºC for 45secs and immediately kept in ice for 2mins. Known volume of plain LB was then added to the cell mixture and incubated at 37ºC for about 1hr. Then 200µl of the cell mixture was taken and plated on LB agar ampicillin plate (100µg/ml) and incubated overnight at 37ºC. The number of colonies grown on the plate was counted and transformation efficiency was determined per µg plasmid.

2.5.4 Extraction of DNA from agarose gel: Extraction of DNA was performed by following Gel extraction kit (Himedia). The DNA band in agarose gel was excised by using scalpel blade under UV light. The gel slice was weighed and accordingly 3 volume of chaotropic salt solution was added per gel slice. The solution was incubated at 55°C for 5-10mins with intermittent mixing every 2-3mins until the agarose gel dissolved completely. Glass powder suspension (GPS) beads of about 10-20µl was added to the solution, mixed well and incubated for 10mins at RT with occasional mixing. The mixture was centrifuged at 1000rpm for 3mins and the pellet was then rinsed with 500µl of sterile water solution by re-suspending it by pipetting. The mixture was centrifuged at 10000rpm for 1min and supernatant was discarded. The washing was repeated two times. The supernatant was removed completely after last wash and the pellet was air dried for 5mins. The beads were re-suspended in 15-50µl of TE buffer or sterile water and incubated at 55°C for 10mins with intermittent mixing. The mixture was centrifuged at 10000rpm for 1min and the supernatant containing DNA was carefully transferred into a new tube.

Composition of Sterile water solution	9:1:10 ratio of water, concentrated wash solution (SWS) and 100% ethanol

2.5.5 PCR amplification recipe

Components	Final concentration	For 25µl reaction volume	For 50µl reaction volume
DNA	50ng	1.5µl	3µl
Forward primer	1µM	0.5µl	1µl
Reverse primer	1µM	0.5µl	1µl
2x DreamTaq Green PCR master mix	1x	12.5 µl	25 µl
Water	-	10µl	20µl
Total reaction volume	-	25µl	50µl

2.5.6 Ligation and restriction digestion recipes

Components for ligation	Final concentration
DNA	75ng
Vector	165ng
Ligation buffer	1x
T4 DNA Ligase	5U
Total volume with water	30µl

Components for Restriction Digestion	Final concentration
Vector DNA	300ng
NEB buffer 3	1x
BSA	5µg
XhoI & BamHI	10U each
Total volume with water	50µl

2.6 Appendix-II: Media and Buffers solution

i. Luria Bertani media and agar (LB broth and LB agar): LB broth was used for growing bacterial culture. It is composed of peptone (1%), yeast extract (0.5%), and sodium chloride (0.5%). Components were dissolved in distilled water and autoclaved for 20mins at 121°C and 15lb pressure. For LB agar solid media, agar was added in media to a final concentration of 1.5%.

iv. Diaminobenzidine (DAB) substrate: 1x PBS containing 1mg/ml of DAB and 0.05% H_2O_2

v. Chemiluminescent peroxidase substrate: Add equal volume (1:1) of chemiluminescent substrate and solution-2 supplied in the kit. Dissolve the solution well in dark and keep for 30mins prior to use.

vi. Western blotting transfer buffer: Dissolve 5.63gm of glycine and 1.21gm of tris base in 400ml of distilled water. Add 20% of methanol (100ml) to the solution and mix well.

vii. O-Phenylenediamine (OPD) substrate: Dissolve1mg/ml of OPD using citrate buffer pH 5.6 containing 0.03% hydrogen peroxide.

ii. Lysis and binding buffer: 100mM of tris buffer pH 8.8 containing 250mM NaCl.

iii. Western blot wash buffer: 1x PBS containing 0.05 % tween 20.

viii. Antigen preparation for injection into animal model: Using a syringe, mix 0.5ml of sterile PFI1625c antigen (400µg/ml) and 0.5ml of Freund's complete adjuvant until formation of micelles. Similarly prepare for Freund's incomplete adjuvant.

ix. Phosphate buffer saline (PBS) and Tris acetate EDTA (TAE) buffer

PBS buffer composition	
Components	**For 10x (gm)**
NaCl	80gm
KCl	2gm
Na_2HPO_4	14.4gm
KH_2PO_4	2.4gm
Adjust pH& volume	7.4 and 1L

TAE buffer composition	
Components	**For 50x (Molar)**
Tris base	2M
Glacial acetic acid	1M
EDTA	100ml of 0.5M
Adjust pH	8.4
Final volume	1L

x. Agarose gel: Electrophoretic analysis of DNA was carried out using agarose gel. Agarose gel was prepared by dissolving appropriate amount of agarose in 1x TAE buffer. Dissolution was achieved by heating the solution in microwave oven for several minutes (2 minutes onwards) with intermittent shaking. For plasmid DNA detection, 0.8-1% agarose was used whereas for genomic DNA detection, 0.5% agarose was used.

xi. SDS-PAGE: Electrophoretic analysis of protein was carried out using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE). The components of SDS-PAGE from preparation of the gel to de-staining of the gel are described from Table A.1 through Table A.6.

Table A.1 Ingredients and preparation of stock reagents for SDS-PAGE

Stock reagent	Preparation
Acrylamide solution (30%)	1gm of N'N'-methylene-bis acrylamide was dissolved in 50ml ultrapure deionized water collected at 18 MΩcm (Millipore, Milli-Q water purification system) in amber colored bottle. On complete dissolving, 29gm acrylamide was added to it and stirred on a magnetic stirrer till a clear solution was formed. The final volume was adjusted to 100ml. The solution was filtered (Whatman No. 1) and stored at 4°C in dark.
TrisHCl (1.5 M, pH 8.8)	54.45gm Tris base was dissolved in 150ml deionized water. The pH of solution was adjusted to 8.8 using HCl and volume made to 300 ml. It was stored at 4°C.
TrisHCl (1M, pH 6.8)	Tris base 6gm was dissolved in 60ml deionized water. The pH of solution was adjusted to 6.8 using HCl and volume made to 100 ml. It was stored at 4°C.
SDS (10%, w/v)	10gm sodium dodecyl sulfate (SDS) was dissolved in 60 ml deionized water. The volume made to 100 ml.
APS (10%,w/v)	100mg ammonium per sulfate (APS) was dissolved in 1ml water.

Table A.2 Recipe for preparation of separating gel of SDS-PAGE (10ml)				
Components (ml)	**8%**	**10%**	**12%**	**15%**
Deionized water	4.6	4	3.3	2.3
30% acrylamide solution	2.7	3.3	4	5
1.5M Tris pH 8.8	2.5	2.5	2.5	2.5
10% w/v SDS solution	0.1	0.1	0.1	0.1
10% w/v APS solution	0.1	0.1	0.1	0.1
TEMED	0.006	0.004	0.004	0.004

Table A.3 Recipe for preparation of stacking gel of SDS-PAGE (5%)		
Components (ml)	**5ml**	**10ml**
Deionized water	3.4	6.8
30% acrylamide solution	0.83	2.0
1M Tris pH 6.8	0.63	2.5
10% w/v SDS solution	0.05	0.1
10% w/v APS solution	0.05	0.1
TEMED	0.005	0.010

Table A.4 Recipe for preparation of 10x loading dye for SDS-PAGE		
Components	**Proportion**	**Requirement (5ml)**
Glycerol	50%	2.5ml
Bromophenol blue	0.2%	10mg
β-mercaptoethanol	10%	0.5ml
SDS	1%	0.5ml of 10% stock
Tris	250mM	1.25ml of 1M stock

Table A.5 Recipe for the preparation of 5x running buffer and loading buffer	
Solution	**Preparation**
5x Running Buffer	15gm Tris base, 5gm SDS and 72gm glycine were dissolved in 800ml of deionized water. The pH was adjusted to 8.3 and volume was adjusted to 1000ml. The solution was filtered (Whatman, Filter No. 1) and stored at 4°C. The buffer (5x) was diluted to 1x and pre-warmed at 37°C before use.
5x Loading Buffer	10ml 0.5M Tris (pH 6.8), 1.6ml SDS 10%, 10 ml glycerol, 0.4ml β-mercaptoethanol and 0.4ml 0.5%(w/v) bromophenol blue were dissolved in 3ml deionized water and pH was adjusted to 6.8. The final concentration of buffer was 1x by mixing 1 volume of 5x sample loading buffer to 4 volumes of sample (protein) before loading in gel.

Table A.6 Recipe for the preparation of staining and de-staining solution (100ml)		
Components	**De-staining solution**	**Staining solution**
Methanol	40ml	40ml
Acetic acid	10ml	10ml
Distilled water	50ml	50ml
Coomassie Brilliant Blue R250	-	0.25gm

Chapter-3: Structural Characterization of Putative PFI1625c from *P.falciparum* 3D7

Structural Characterization of Putative PFI1625c from *P.falciparum* 3D7

3.1 Introduction

The structure of a molecule reveals a number of biochemical and functional properties in the light of drug discovery which increase efficiency and speed up the processes. Different methods like structure based drug design utilize the knowledge from 3D protein structures to identify and select therapeutically relevant target, characterize binding pockets and properties, prepare target specific compound libraries, identify hits by docking and study nature of interactions and optimize lead compounds. Homology modeling is a methodology to construct protein structure based on amino acid sequence of target and template protein with the assumptions that proteins with similar sequence have similar structures. The process of homology modeling consists of (i) identification of known template 3D structure, (ii) sequence alignment of target and template proteins, (iii) model building of the target and (iv) validation of models and refinement. PFI1625c (PF3D7_0933600) is a putative uncharacterized protein present in the genome of *P.falciparum* 3D7. *In Silico* characterization of PFI1625c was performed to decipher the structural and functional properties which are discussed in the following sections.

3.2 Experimental procedures

3.2.1 Retrieval of PFI1625c gene: PFI1625c gene is located on chromosome 9 from 1,331,349 to 1,332,803 in *P.falciparum* 3D7 genome. The sequence of PFI1625c was retrieved from plasmodium database (http://plasmodb.org/plasmo/showRecord.do?name=GeneRecordClasses.GeneRecordClass&source_id=PF3D7_0933600&project_id=PlasmoDB). It is 1455 base pairs in length and it codes for 484 amino acids of 55.7kDa. The isoelectric point (pI) of PFI1625c as obtained from plasmodium database is 6.78.

3.2.2 Sequence analysis and alignment: The NCBI PSI-BLAST algorithm (http://blast.ncbi.nlm.nih.gov/Blast) was used to identify homologous structure of PFI1625c by searching the protein sequences in the Protein Data Bank (PDB). Amino Acid Sequence of bc1 core protein of chicken, bovine, yeast mitochondrial processing peptidase and bacillus peptidase was retrieved from NCBI database and multiple sequence alignment was performed using Clustal W 2.0.11 [199].

3.2.3 Molecular modeling of PFI1625c: The amino acid sequence of PFI1625c was retrieved from the Plasmodium Database (http://www.plasmodb.org). PFI1625c 3D

structure modeling was performed in the following sequential steps: template selection, sequence-template alignment, model building, refinement and validation. The suitable template was identified by searching PFI1625c into the protein data bank of NCBI using PSI-blast (http://blast.ncbi.nlm.nih.gov/Blast). The top hits were further analyzed for least number of gaps and highest sequence identity. Crystal structure of mitochondrial processing peptidase from yeast (PDB code 1HR6, β-subunit) was found to be suitable [200] with 38% identity over 435 residues. The 3D model of PFI1625c was generated using MODELLER 9v9 [201]. Auto model module of modeler 9v9 was used to generate 100 initial models with a ga431 score of 1 and these models were ranked based on their DOPE scores and molpdf.

3.2.4 Structural validation of PFI1625c model: Top ten models of PFI1625c having the lowest DOPE scores and molpdf were selected for validation. The stereo-chemical quality of each model was confirmed by PROCHECK and Ramachandran plot [202]. The statistics of non-bonded interactions between different atom types were analyzed by ERRAT program which gives a measure of the structural error at each residue in the protein [203]. The compatibility of the atomic model (3D) with its own amino acid sequence was determined by Verify 3D [204]. The model with the least number of residues in the disallowed region was selected and energy minimized using Steepest Descent (SD) algorithm with GROMOS96 43a1 force fields in GROMACS 4.0.7 package (http://www.gromacs.org/). This process was repeated until most of the residues are not below 95% in ERRAT plot. The final model was checked by Verify-3D, PROCHECK and Ramachandran plot.

3.2.5: Hydropathy plot of PFI1625c: To check the hydrophobicity of PFI1625c domains that involve in membrane spanning, potential antigenic sites or exposed regions, the amino acid sequence of PFI1625c was imported into the online based hydropathy plot. The method used was the Kyte-Doolittle scale with a window size of 19 [205]. Amino acid residues are assigned a scale or value and according to this scale, the regions which fall below 0 are hydrophilic and those falls above 0 are hydrophobic whereas those above 1.8 are transmembrane protein.

3.2.6: Phylogenetic distance with human metalloprotease: To validate PFI1625c as a targetable protein for drugs, the phylogenetic distance of PFI1625c with similar human metalloprotease was studied and for this, PFI1625c was blasted into NCBI non-redundant protein database (nr). The top hits proteases and PFI1625c was aligned

using multiple sequence alignment tool Clustal W 2.0.11. The aligned sequences were imported into phylogenetic analysis package Phylogeny.fr which uses the algorithm of BioNJ which is a newer version of neighbor joining [206]. Regions of sequences that were difficult to align were removed by Gblocks present in Phylogeny.fr package. To determine the strength of the groupings, bootstrap values for nodes were calculated by analyzing 500 bootstrap replicate data sets. Structural comparison and active site conformation was also studied between PFI1625c and closest homologue of human metalloprotease insulysin.

3.3 Results

3.3.1 PFI1625c may belong to metalloprotease class

The BLASTP result shows that all the top hits proteins belong to metalloprotease class. They are yeast mitochondrial processing peptidase, bacillus peptidase, bovine and chicken bc1 core proteins. Multiple sequence alignment with top hits implicates that PFI1625c belongs to the metalloprotease family. The characteristic conserved sequences pattern present in PFI1625c matches with other known metalloprotease (Figure 3.1). A characteristic zinc binding motif (HXXEH) and active site residues specific for a metalloprotease class are also present in PFI1625c and bacillus peptidase. This motif is a characteristic feature of inverzincins family. Sequence analysis showed that conserved zinc co-ordinate E-160 (yeast metalloprotease) which is present in different motif to co-ordinate the zinc ion, is also conserved in PFI1625c and other similar metalloproteases. Important substrate binding residues such as E-160, D-164 and F-77 are also conserved in PFI1625c which is also conserved entirely or partly in other metalloproteases. Two substrate binding scaffold from different domains are present in yeast metalloprotease (HLNAY & SYMSFSTSY) out of which one scaffold HLNAY is present near the zinc binding motif and this scaffold is conserved in PFI1625c and other metalloproteases. The other scaffold is conserved but not entirely conserved in all metalloproteases including PFI1625c. Conserved sequence comparison is given in Table 3.1. To further validate PFI1625c as metalloprotease, the amino acid sequence was analyzed by ProtIdent web server [207] which is developed by fusing the functional domain and sequential evolution information. The first layer identifies the query protein as protease or non-protease while the second layer identifies the protease class. ProtIdent server indicates that PFI1625c is a protease belonging to the metalloprotease family.

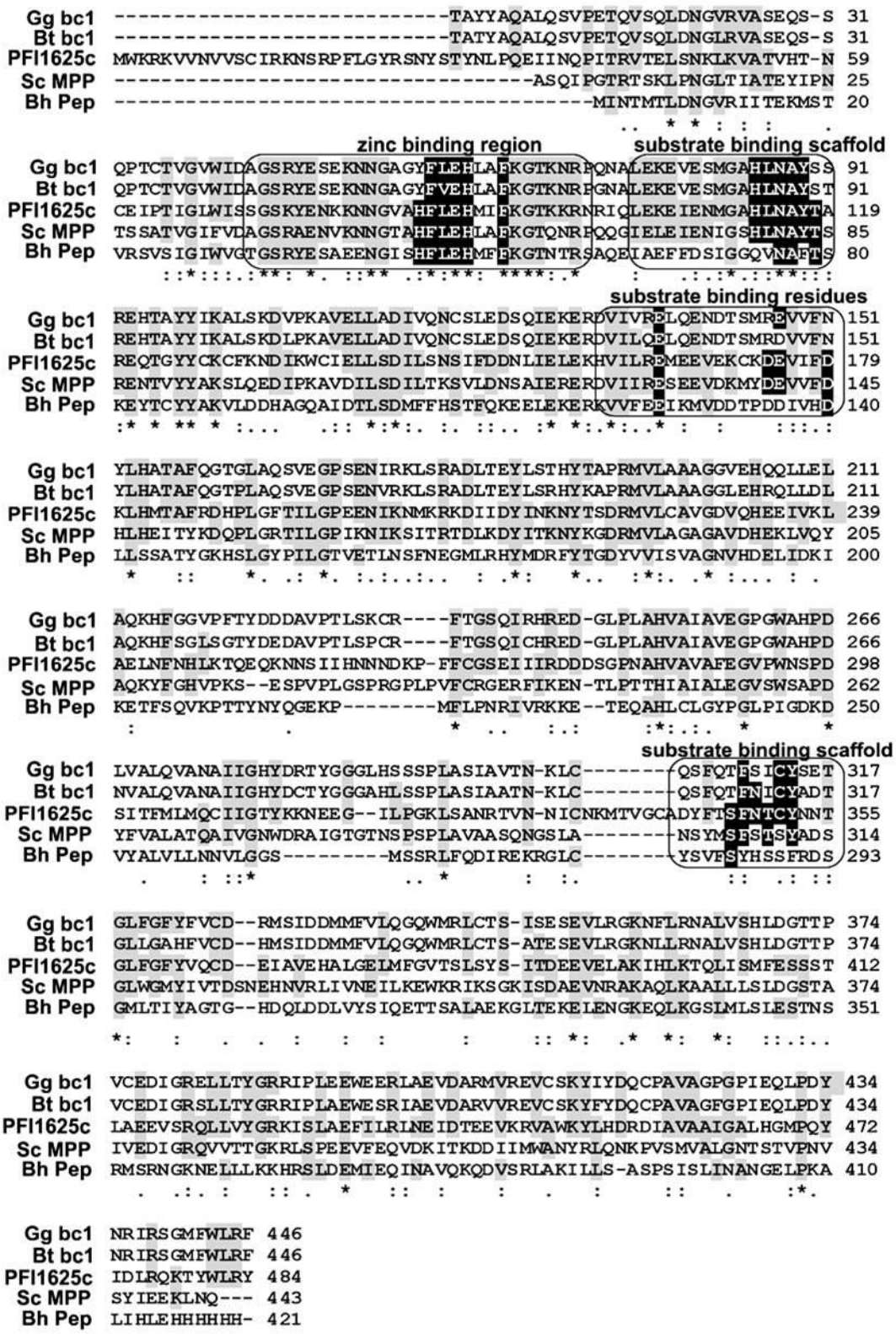


Figure 3.1 Characterization of different structural elements in PFI1625c. Multiple sequence alignments of PFI1625c from *P.falciparum* with bc1 core protein of *G. gallus* (Gg bc1), *B. taurus* (Bt bc1), metalloprotease of *S. cerevisiae* (Sc MPP) and protease from *B. halodurans* (Bh Pep). Similar and identical residues are indicated by standard annotation. The zinc binding region and substrate binding scaffold are enclosed within the rectangular box to highlight the presence of these regions in PFI1625c.

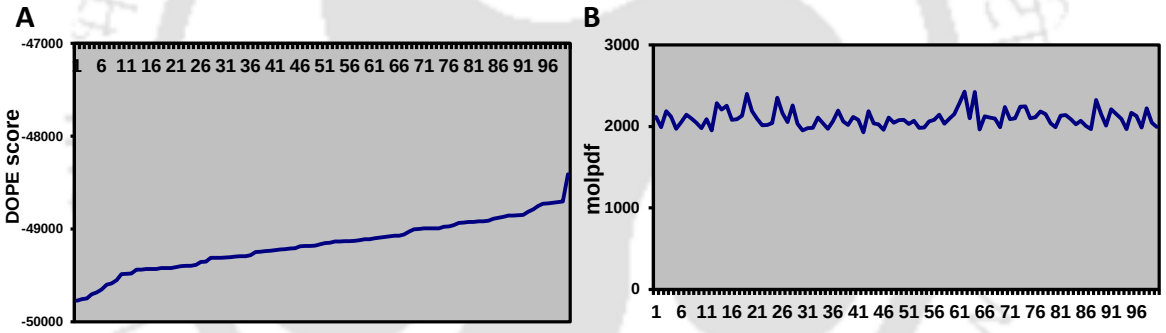
Table 3.1 Comparison of conserved residues of top hits metalloproteases

Protein name	motif	Residue	Binding residues	Binding scaffold	Substrate binding region
ScMPP	HXXEH	Glu-150	E-160, D-164 & F-77	HLNAY/SYMSFSTSY	Rich in
PFI1625C	HXXEH	Present	Present	HLNAY/YFTSFNTCY	negatively
Bt bc1	YXXEH	Present	F-77 present	HLNAY/SFQTFNICY	charged
Gg bc1	YXXEH	Present	E-160 and F-77 present	HLNAY/SFQTFSICY	amino acid
Bh pep	HXXEH	Present	D-164 and F-77 present	QVNAF/SVFSYHFFS	residues

Conserved sequence comparison of PFI1625c with bc1 core protein of *G. gallus* (Gg bc1), *B. taurus* (Bt bc1), metalloprotease of *S. cerevisiae* (Sc MPP) and protease from *B. halodurans* (Bh Pep).

3.3.2 PFI1625c model verification

The sequence homology between PFI1625c and 1HR6 was 38% identity and 59% similarity. The DOPE score is best in the first ranked models whereas molpdf is almost constant in all models (Figure 3.2A and B). All models give a GA341 score of 1 which shows a good quality model. Based on the DOPE score and molpdf best 10 models were selected (Figure 3.2C).



C

Filename	molpdf	DOPE score	GA341 score
PFI1625c.B99990009.pdb	2117.91113	-49772.4375	1
PFI1625c.B99990059.pdb	1991.3208	-49757.51563	1
PFI1625c.B99990048.pdb	2187.45605	-49750.05469	1
PFI1625c.B99990040.pdb	2119.91968	-49702.57813	1
PFI1625c.B99990010.pdb	1968.86743	-49682.69141	1
PFI1625c.B99990081.pdb	2053.15942	-49651.76563	1
PFI1625c.B99990038.pdb	2145.3374	-49602.33594	1
PFI1625c.B99990006.pdb	2095.44751	-49586.73828	1
PFI1625c.B99990015.pdb	2044.11133	-49551.95703	1
PFI1625c.B99990097.pdb	1978.07947	-49487.35938	1
PFI1625c.B99990078.pdb	2087.16895	-49483.52344	1
PFI1625c.B99990064.pdb	1951.82019	-49478.53906	1
PFI1625c.B99990001.pdb	2283.32568	-49439.48047	1
PFI1625c.B99990021.pdb	2204.83667	-49437.35156	1
PFI1625c.B99990007.pdb	2253.32202	-49433.04688	1
PFI1625c.B99990069.pdb	2080.30298	-49432.62891	1
PFI1625c.B99990018.pdb	2086.84473	-49431.93359	1
PFI1625c.B99990035.pdb	2131.07886	-49422.68359	1
PFI1625c.B99990071.pdb	2399.00195	-49422.51953	1
PFI1625c.B99990056.pdb	2186.53174	-49419.94922	1

D

Figure 3.2 Criteria for selection of 3-D molecular models for validation. (A & B) DOPE score and molpdf of total 100 generated models where DOPE score is best at top rank models whereas molpdf is almost constant throughout the models. **(C)** Ranking of generated models based on their DOPE score and molpdf. Here 20 top models are highlighted where top 10 models (yellow) are taken for validation. **(D)** Structural alignment of generated PFI1625c model (green) with template model 1HR6 (magenta) showing a good RMSD value of 1.52Å.

The PROCHECK result of top PFI1625c model generated from 1HR6 template shows good quality structure. The Ramachandran plot drawn through PROCHECK program validated the model with 89.7% of the total amino acids in the allowed conformation. 8.8% falls in the additional allowed region. 1.5% falls in the generously allowed region while no amino acid falls in the disallowed region (Figure 3.3A). This stipulates that protein backbone dihedral angles phi (φ) and psi (ψ) occupied reasonably accurate positions in the selected 3-D model. The ERRAT plot also shows the overall quality factor of 90.071. This shows that most of the residues fall below 95% cut off value and thus is a good structure (Figure 3.3B). Compatibility of the atomic model (3D) with its own amino acid sequence as verified by Verify_3D program shows compatibility of 80.56% (Figure 3.3C). The PFI1625c structure superimposes onto 1HR6 B protein with a root-mean-square-deviation (RMSD) of 1.52Å which shows a highly successful model (Figure 3.2D). The above results confirm that the 3D model constructed for PFI1625c by comparative modeling is reliable for detail structural analysis. Overall no short contact was observed in the final model.

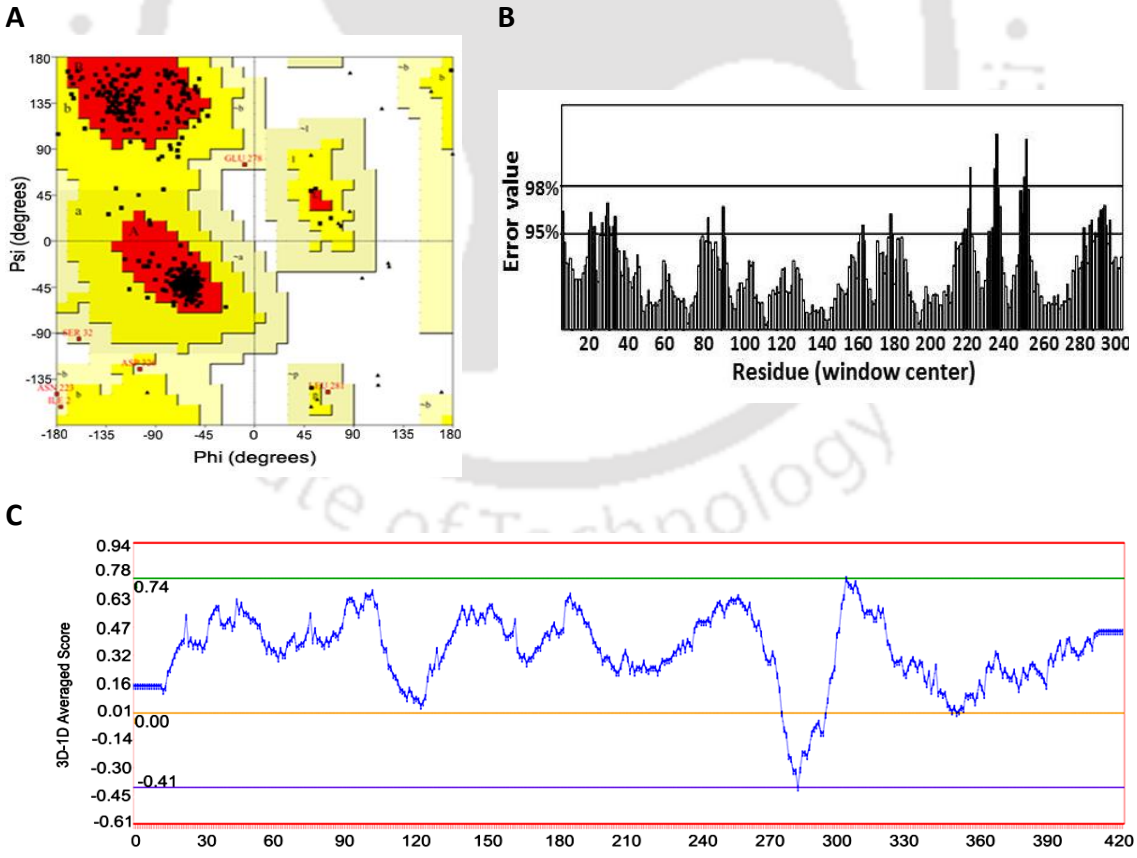


Figure 3.3 Validation of PFI1625c molecular model. (A) Ramachandran plot of validated model showing all residues in the allowed region. **(B)** ERRAT plot showing the overall quality factor of 90.071. **(C)** Verify_3D graph showing the compatibility of atomic model with its own amino acid sequence with a value of 80.56%.

3.3.3 The overall structure of PFI1625c

The complete structure of PFI1625c can be divided into domains I and II. There is a sequence difference between two domains but the α-helix and β-sheet folds in each domain to give similar folding topology of a non-super-imposable image (Figure 3.4A). Electrostatic representation showed an organized cavity with charge distribution which is negatively charged at the exterior and neutral or hydrophobic at the interior (Figure 3.4B). A loop of 25 residues connects α-helix 11 of a domain I and β-sheet 7 of domain II. The α-helices and β-sheets are equally distributed in both the two domains as represented in the topology diagram.

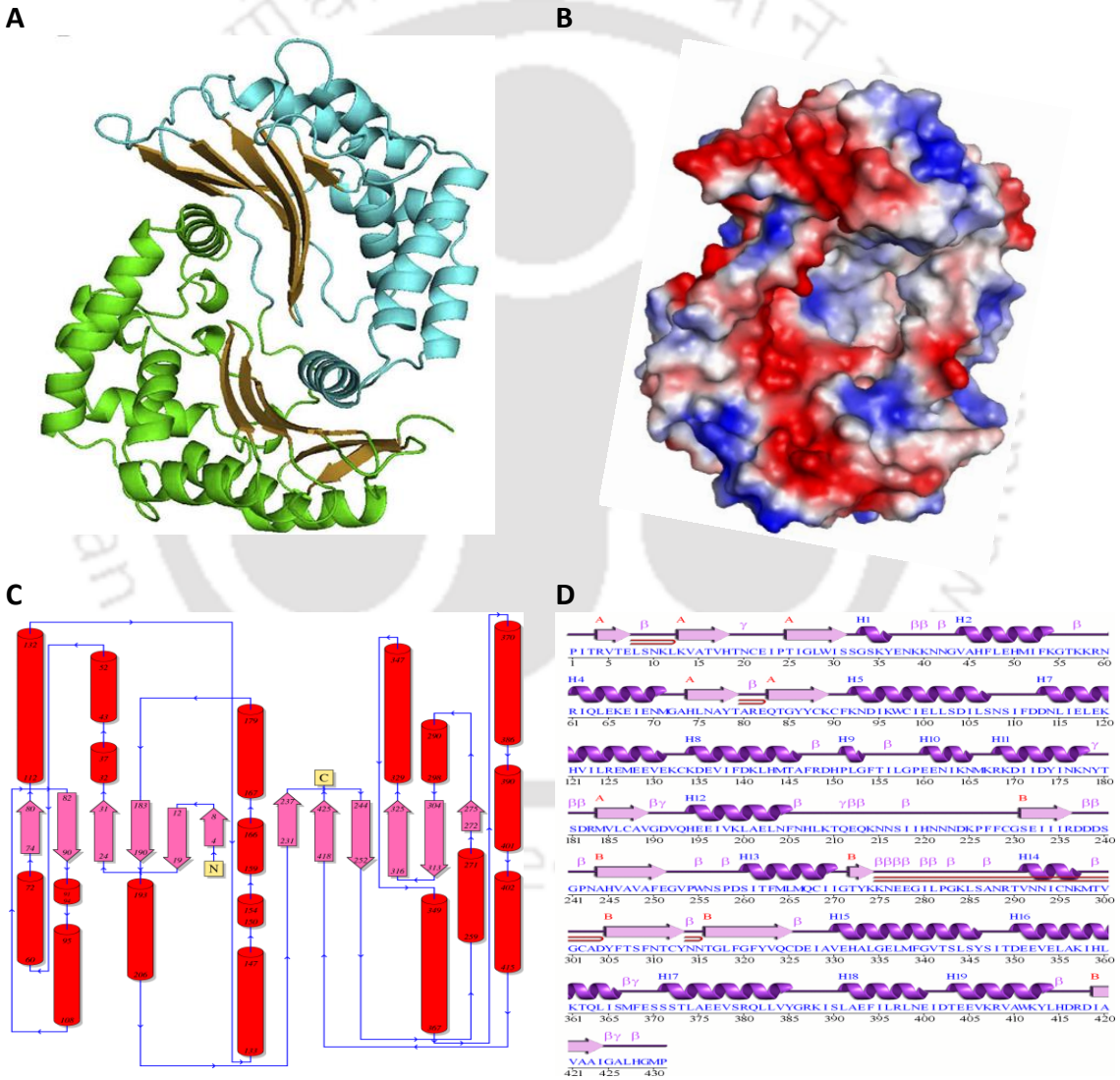


Figure 3.4 Overall structure of PFI1625c. (A) 3-D structure of PFI1625c consists of two domains. Domain I is colored green and domain 2 is colored cyan. The β-sheets present in respective domains are colored in brown. **(B)** Electrostatic representation showing charge distribution and organized cavity. **(C)** Topology diagram of PFI1625c to depict the organization of structural elements. The α-helices are colored red while β-sheets are colored pink and structural elements are connected by blue colored lines. **(D)** Secondary structure showing the backbone dihedral angles and organization.

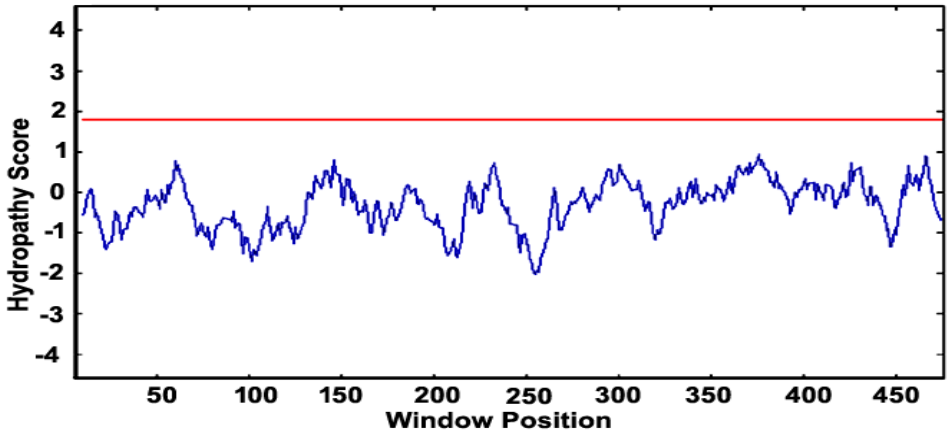


Figure 3.5 Kyte-Doolittle Hydropathy plot of PFI1625c with window size of 19. Regions above 0 are non-polar whereas regions below 0 are polar. The red line indicates a critical value of 1.8. Any segment that reaches above 1.8 suggests transmembrane region. This indicates that PFI1625c is not a transmembrane protein.

The β-Sheets of each domain are projected into the interior core of the PFI1625c and consist of two hairpin structures and Ψ-loop motif. The Ψ-loop structures from each domain are composed of two anti-parallel β-strands. The α-helices are present toward the exterior portion of the PFI1625c to shield the interior β- sheets (Figure 3.4C). The secondary structure showed the exact pattern of backbone dihedral angles where altogether there are 2 β-sheets, 4 β-hairpins, 2 Ψ-loop, 2 β-bulges, 12 strands, 19 α-helices, 24 helix-helix interacs, 35 β -turns and 9 γ-turns (Figure 3.4D). These arrangements give rise to a complete protease structure with an organized active site cavity and intact catalytic site present in other metalloproteases. Hydropathy plot showed that most of the residues fall below 0 which means PFI1625c is more hydrophilic in nature. No segment falls above a score of 1.8 that shows that there is no membrane-spanning and antigenic domain of PFI1625c which suggested that PFI1625c is not a transmembrane protein (Figure 3.5).

3.3.4 Zinc binding site

A characteristic feature of metalloprotease is the presence of the metal binding motif [208]. Multiple Sequence alignment indicates that a conserved zinc binding motif (HXXEH) with H-46, E-49, H-50 residues are present in PFI1625c where zinc co-ordinates E-126 (Figure 3.6A). The E-126, another zinc binding residue from other helix comes closer with the H-46 and H-50 to form the perfect geometry that fits the zinc metal. The E-49 is not co-coordinating the zinc ion but is predicted to initiate the catalysis by activating water molecule for nucleophilic attack of incoming substrate as observed in the template metalloprotease. The structure of the protein implies that α-helix6 and 2

from a domain I come closer to each other with a unique folding pattern to align the zinc binding residues (H-46, E-49, H-50) to form a catalytic triad with E-126 (Figure 3.6A and B).

3.3.5 Substrate binding site

Substrate binding region of PFI1625c is constituted by E-136, D-140 and F-53 (Figure 3.6B). The active site is present inside a well-organized large cavity in the center of the structure. The active site cavity closer to domain I is negatively charged whereas rest cavity is of mixed environment (Figure 3.6C). The domain I region of the active site is negatively charged to facilitate binding of positively charged residues of the substrate while the remaining portion of the cavity consists of mixed charges with a hydrophobic

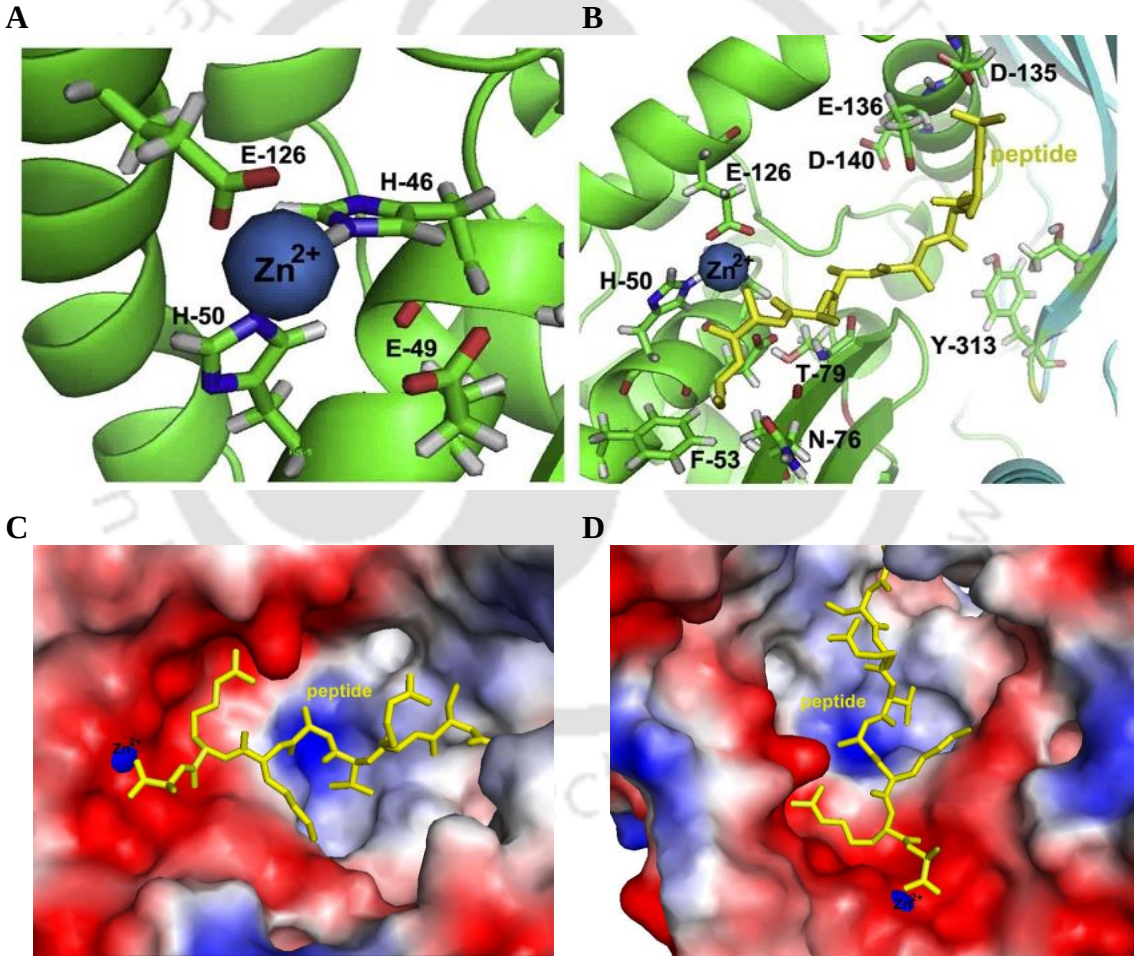


Figure 3.6 PFI1625c is a metalloprotease with a well-defined active site. (A) Zinc binding motif of PFI1625c present in the domain I. The zinc ion is shown in blue and the residues interacting with Zn^{2+} to stabilize it within the catalytic site are shown with corresponding label. **(B)** Active site of PFI1625c with bound peptide substrate. The template peptide (LSRVAKRA) was taken from yeast mitochondrial processing peptidase (PDB ID 1HR9) and modeled into the active site. The residues interacting with peptide substrate are labeled and shown. PFI1625c surface representation in **(C)** horizontal and **(D)** vertical view with bound template peptide in yellow. Red represents −ve charge, blue represents +ve charge.

environment toward the interior to provide docking sites for hydrophobic residues of a peptide substrate. The cavity extends toward the domain II which may assist in binding the large substrates to allow access of the cleavage site by the catalytic residues present in the domain I (Figure 3.6D). Substrate binding scaffolds comprise of 2 β-sheets, contributed from both domains. The β-sheet from domain I has conserved HLNAY residues while the β-sheet from domain I has SFNTCY residues.

3.3.6 Interaction of PFI1625c with template peptide

PFI1625c superimpose well with the template yeast mitochondrial processing peptidase (1HR6). Interaction of PFI1625c was studied with malate dehydrogenase signal peptide (LSRVAKRA) which is complexed with yeast mitochondrial processing peptidase mutant (PDB code 1HR9). The super imposed template peptide binds PFI1625c similar to the co-crystal structure of mitochondrial processing peptidase bound to signal peptide (PDB Code: 1HR9). The LSRVA occupy the more hydrophobic region of the active site pocket while KRA occupy the more negatively charged region (Figure 3.7A). The interaction between PFI1625c and peptide was studied from Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC) and distance of interaction was measured using Swiss-PdbViewer (http://www.expasy.org/spdbv/). LigPlot+ v.1.4.5 [209] was also used to generate PFI1625c-peptide interaction. The first amino acid in the peptide i.e., L1 (non-polar) lies in the hydrophobic region and is bound by non-polar residue M-264 and neutral residues Q-267 and Q-363 at a distance of hydrophobic and polar bond (Figure 3.7B). Second residue S2 is close to neutral residues N310 and Q267 of PFI1625c. Third residue R3 is not close to any residue the nearest being F309 of PFI1625c. Residues in the middle of peptide such as V4 and A5 interact with A80 and R81 of PFI1625c with a Van der Waals force. A strong ionic bond is formed by K6 of peptide with E376 of PFI1625c and R7 of peptide with E136 of PFI1625c. K6 of peptide interacts with Y78 of PFI1625c at a polar bond distance. The last residue A8 of peptide forms polar contact with two of the zinc binding residues H46 and H50. The catalytic site of PFI1625c shows all the type of interactions with template peptide such as ionic bond, hydrogen bond, hydrophobic bond and Van der Waals force. The type of atom involved in the interaction and the distance as calculated from Swiss-PdbViewer are given in Table 3.2.

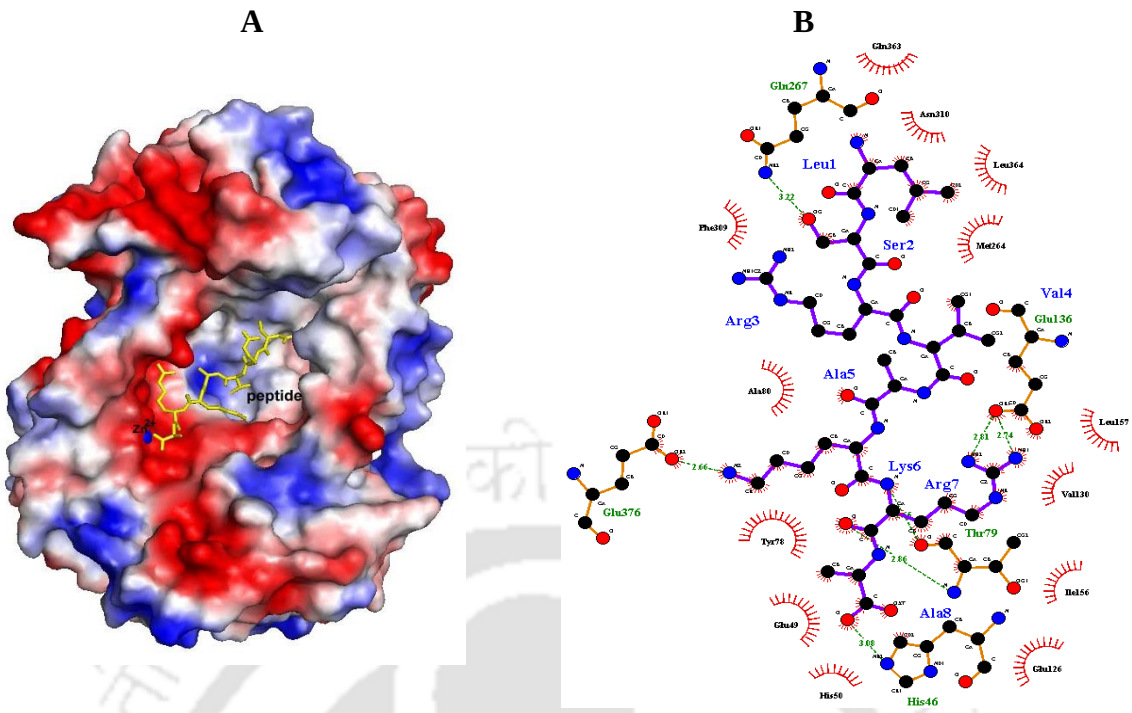


Figure 3.7 Interaction of template peptide (LSRVAKRA) with PFI1625c. (A) Surface representation of PFI1625c with template peptide. Red represents −ve charge, blue represents +ve charge. **(B)** Interaction of template peptide (purple backbone) with active site residues of PFI1625c (red) as generated by LigPlot+ v.1.4.5.

Table 3.2: Interaction of the peptide within the active site of PFI1625c*

Template peptide	Protein Residue	Atom of protein	Distance (Å)
L1	M264	CE	2.52
L1	Q363	NE2	1.37
L1	M264	CE	2.63
S2	N310	O	3.15
S2	N310	H	2.49
S2	Q267	NE2	2.44
R3	F309	CB	6.05
V4	A80	CA	3.82
A5	R81	NE	3.74
K6	OE2	NZ	2.66
K6	Y78	OE2	2.66
R7	E136	OE1	2.74
A8	H46	NE2	2.16
A8	H50	NE2	2.71

*The interaction distance is obtained by manually measuring two closely located residues in Swiss-PdbViewer and Pymol. The polar contacts are obtained from pymol.

3.3.7 PFI1625c is evolutionarily distant from human metalloprotease

Phylogenetic analysis of non-redundant database by phylogeny.fr neighbor-joining shows an unrooted phylogram joining two taxa by a single node. The phylogram is represented in radial form so that the branch length corresponding to genetic change can be easily estimated. There are three distinct clusters obtained from the phylogenetic tree and can be

classified into three clades as Clade-1, Clade-2 and Clade-3. PFI1625c forms a separate cluster in Clade-1 with putative proteases of other apicomplexan species as they are evolutionarily close to each other and high sequence identity. The nearest homologues to PFI1625c were found to be metalloprotease from cryptosporium, theileria species and *Babesia bovis*. The genetic change as a function of mutations or substitutions per 100 nucleotides occurring during evolutionary change is shown in the tree with a reference line segment of 0.3. The branch length in phylogram directly corresponds to the number of mutation. Human metalloprotease is found to be highly diversified into distinct group (Clade-3) and is evolutionary and taxonomically most distant from PFI1625c. Considering the branch length, there is about 5times genetic mutations and substitutions occurred between PFI1625c and human metalloprotease during evolutionary change (Figure 3.8). The bootstrap values are all above 60% which shows the reliability of the tree. This evolutionary difference support PFI1625c to be a candidate for drug target.

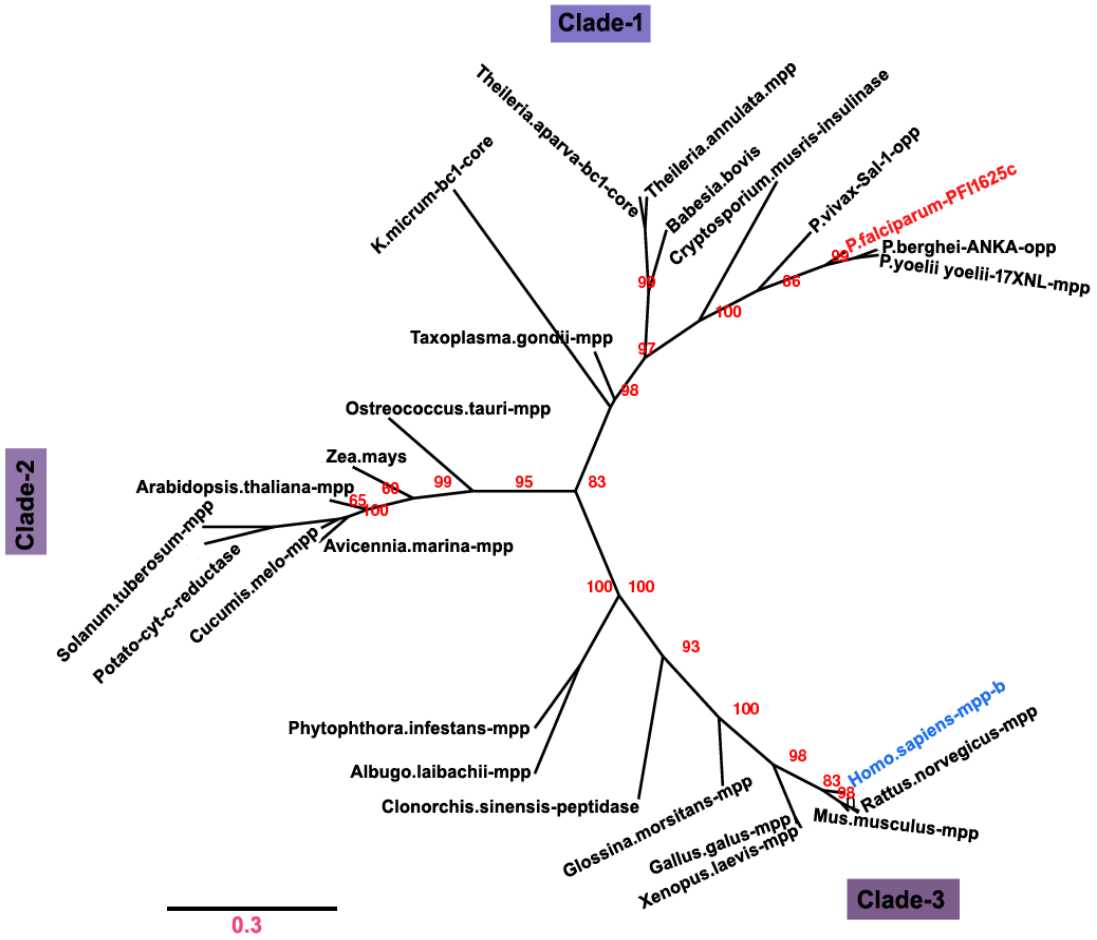


Figure 3.8 Evolutionary relationship of PFI1625c with proteases present in other organism. Three distinct clusters are formed and are assigned as Clade-1 to 3. Human metalloprotease (highlighted as blue) is evolutionarily diverse from PFI1625c (highlighted as red) and form a distinct cluster in Clade-3. Bootstrap confidence values (%) are given in red. A reference line segment with value (0.3) in pink denotes the mutations or substitution occurring during evolutionary change.

3.3.8 PFI1625c is a suitable target for drug development

PFI1625c and human IDE differ considerably in their structural organization. Multiple sequence alignment of PFI1625c with top hits showed that the zinc binding motif (HXXEH) is conserved in human IDE. The other zinc co-ordinate E-126 is also present in human IDE. However, significant differences are observed in other conserved regions such as substrate binding motifs and substrate binding residues. The characteristic substrate binding motifs HLNAYT and SFNTXY are not conserved in human IDE where NAFT and SLN are present in human IDE. The F-53 which is present near the zinc binding motif to accommodate incoming hydrophobic substrate residue is not conserved in human IDE. Another conserved residue E136 is also absent in human IDE. Structural comparison of PFI1625c with human IDE showed a vast difference in the active site

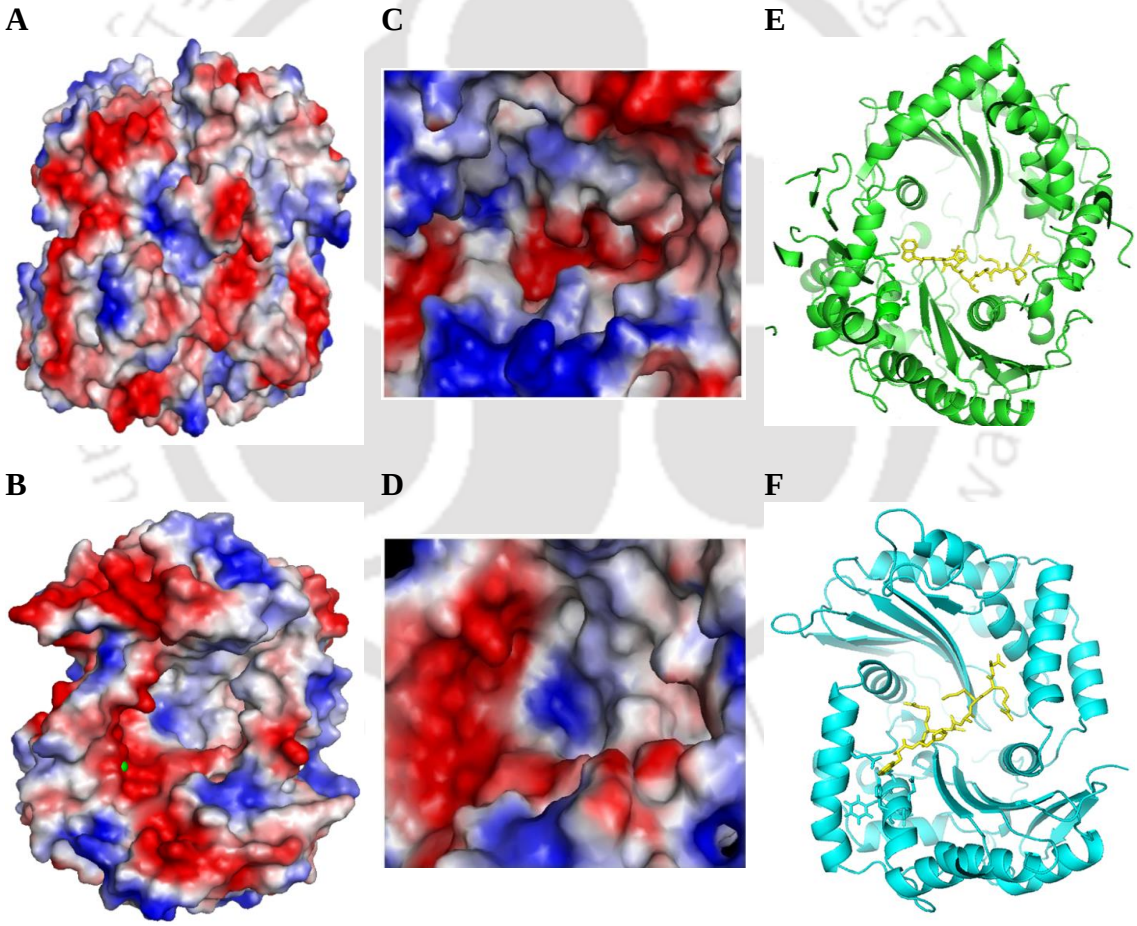


Figure 3.9 The electrostatic representation of the active site and interaction of bound peptide within human insulin degrading enzyme (IDE) and modeled structure of PFI1625c. (A & B) Overall surface representation of **(A)** IDE and **(B)** PFI1625c. Blue is +ve charge and red is –ve charge. Active site cavity is visible in PFI1625c marked by Zinc ion (green dot). **(C & D)** Surface Structure of the peptide binding pocket of IDE and PFI1625c. **(E & F)** Interaction of bound peptide with key residues present within the binding pocket of IDE (green) and PFI1625c (cyan). Peptide (yellow) is present in an extended conformation.

cavity in which the active site of human IDE is completely buried (Figure 3.9A) whereas that of PFI1625c is exposed and the zinc binding motif coordinating the zinc ion can be visualized (Figure 3.9B). The residues covering the active site cavity of human IDE were removed for convenient illustration. The active site of human IDE is not organized unlike PFI1625c which is a large organized cavity as shown in the surface representation (Figure 3.9C and D). The active site of human IDE has a region of +vely charged exterior and mixed residues in the interior whereas PFI1625c is more –vely charged in the exterior and has more hydrophobic interior. Their differences in interaction with incoming peptide are illustrated in (Figure 3.9E and F). All these differences in evolutionary, sequential and structural organization support PFI1625c as a suitable target.

3.4 Discussion

PFI1625c is evolutionarily distant from human metalloprotease. Phylogenetic tree shows that they form a separate and farthest cluster with about 5 times of 0.3 genetic mutations occurring between the two. PFI1625c and human metalloprotease also differs considerably in their active site conformation and charge distribution. This strongly validates PFI1625c to be a targetable protein. Besides its difference with human protease, it also forms separate branches with *P.vivax*. Thus, specific drug against *P.falciparum* can be achieved. PFI1625c has high similarity with other known metalloproteases with a conserved metal binding motif HXXEH [208]. This is a characteristic feature of inverzincins which is an inverted motif present in thermolysin. A typical Zn^{2+}dependent metalloprotease follows a mechanism involving different distinct reaction intermediate [210]. Structural determinants play a crucial role in substrate recognition and specificity [211]. Superimposed model of PFI1625c and yeast mitochondrial processing peptidase (PDB Code 1HR9) was used to obtain the 1HR9 peptide (LSRVAKRA). The root mean square deviation between generated model and crystallographic complex (1HR9) is 1.52Å. This gives us an idea to initiate our study how signal peptide interacts with PFI1625c [200]. Although LigPlot is a robust tool for generating protein-ligand interaction, it does not show all the type of interactions and distance. It showed distance only for ionic and polar bonds. Therefore the use of manual measurement was necessary. However, the distances obtained from LigPlot and Swiss-PdbViewer was more or less the same which shows the reliability of SwisspdbViewer. Based on the stereospecific positions of different residues within the PFI1625c active site, a schematic cleavage mechanism is given in Figure 3.10.

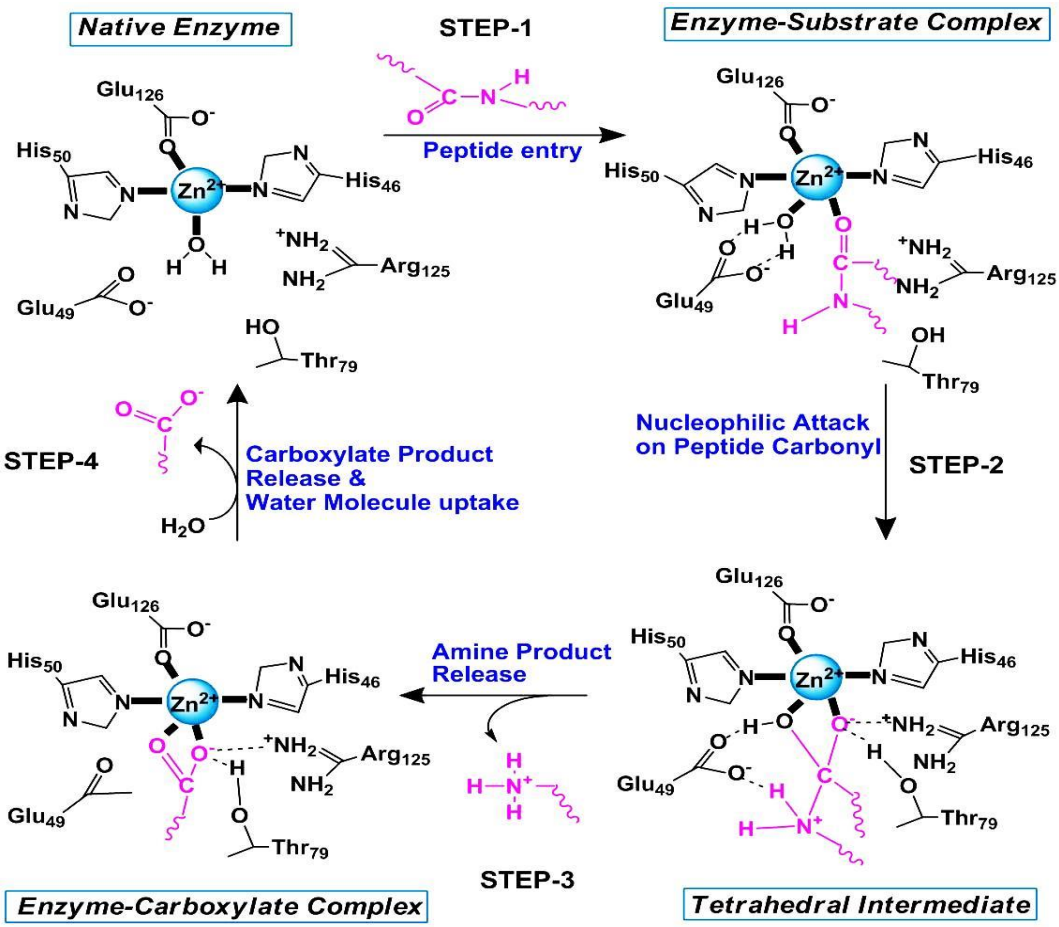


Figure 3.10 Proposed catalytic mechanism of PFI1625c mediated peptide cleavage. Based on the stereospecific positions of different residues within the PFI1625c active site, a schematic cleavage mechanism is proposed.

Comparing PFI1625c with metalloprotease shows that E-136 and D-140 present near the zinc binding site to form the S2 and S3 sites. The negatively charged sites form a salt-bridge with an arginine present in an incoming peptide to facilitate strong binding (Figure 3.7B). The F-53 lies at the S1ŝite and is believed to interact with the aromatic residues found in the substrate. The residue E-49 which lies within hydrogen bonding distance to a water molecule coordinated to zinc is predicted to polarize this water molecule, thereby aligning it for nucleophilic attack on the bond of the substrate (Figure 3.10). The active site of PFI1625cor other metalloproteases differs from thermolysin in the absence of analogous residues Y-157 and H-231 which are proposed to stabilize the oxyanion of a tetrahedral intermediate. Solvent molecules or similar charged amino acids T-79 and R-125 could be involved in hydrogen bonding to stabilize oxyanion of a tetrahedral intermediate in PFI1625c (Figure 3.10).

References: Please refer to bibliography section from page number 108-119.

Chapter-4: Biochemical Characterization of Putative PFI1625c from *P.falciparum* 3D7

Biochemical Characterization of Putative PFI1625c from *P.falciparum* 3D7

4.1 Introduction

Metalloproteases constitute second largest protease class (22%) in *P.falciparum* genome. Metalloproteases are involved in different biological activity imparting to parasitic growth as in the invasion step during RBC stages of malaria [179], organelle processing activity and hemoglobin digestion [159]. They also process different proteins to reach their target organelles like apicoplast [168]. Besides the well documented metalloproteases, there are a number of uncharacterized proteases present in the *P.falciparum* genome. Metalloprotease specific inhibitor studies showed that complete inhibition is not achieved, indicating role of additional uncharacterized metalloproteases might be required for invasion [179], protein biogenesis [169, 170] and gametogenesis [212]. The degradation of hemoglobin and export of degraded amino acids into cytosol are not fully understood. Uncharacterized metalloproteases might be involved in later part of hemoglobin digestion as they are mostly involved in degradation of oligopeptides and dipeptides [165]. PFI1625c is a putative uncharacterized protein identified in *P.falciparum* genome. Bioinformatics analysis and *in silico* molecular modeling studies implicate that it may belong to metalloprotease family. However, it needs to be validated by additional in-depth biochemical experiments. In this chapter, we have discussed biochemical characterization of PFI625c and its relevance in drug discovery & diagnosis of *P.falciparum*. The chapter is described in the following sections:

4.2 Experimental procedures

4.2.1 Protease assay: Protease assay was performed for PFI1625c to study the biochemical properties. Assay buffer was prepared using Tris base and the pH was adjusted to 8.5. Pure PFI165c and substrates were prepared in assay buffer. 100µl of PFI1625c containing 20µg was pipetted out into eppendorf tube in triplicate. A working solution of substrate (at a final concentration of 1%) was added to the enzyme solution. The reaction volume was adjusted to 0.5ml using assay buffer (Tris pH 8.5). A blank solution containing only substrate and assay buffer was also prepared with a final volume of 0.5ml. The reaction mixture containing enzyme and substrate as well as blank solution was incubated for about 2hrs at 37˚C. The reaction was stopped by precipitating with 10% Trichloroacetic acid (TCA) by mixing gently and incubated for 5mins at room temperature. The precipitation was pelleted by centrifugation at 13,000rpm for 10mins at

room temperature. The supernatant containing the cleaved peptide was collected and optical density (OD) of the supernatant was measured at 280 nm using UV-Visible spectrophotometer (Cary 100 Bio). The protease activity was calculated from the absorbance (after subtracting the blank value) and Michaelis-Menten curve was drawn.

4.2.9 Gelatin zymography: Gelatin zymography was performed to check the activity of PFI1625c towards matrix protein gelatin, a standard assay for matrix metaloproteases. Matrix metalloprotease-2 (MMP2) and BSA were used as positive and negative control respectively. Firstly, gelatin stock (1%) was prepared in distilled water by incubating at 37°C with intermittent mixing. After the gelatin was completely dissolved, 1ml of the stock solution was incorporated in 8% resolving gel of SDS PAGE to achieve a final gelatin concentration of 0.1%. The solution was mixed well to completely dissolve the gelatin. The solution was carefully poured into the glass plate of SDS PAGE apparatus (Biorad). After resolving gel polymerized, stacking gel (not containing gelatin) was poured above resolving gel. Samples were prepared in native PAGE loading buffer without heating. PFI1625c along with control samples and protein marker were loaded in the gel. The gel was run at low voltage ~ 50V until bromophenol dye reached near the end. The gel was washed with wash buffer containing 2.5% Triton X-100 in water (3 x 30mins each) with constant shaking for complete removal of SDS. Activation buffer (Table 4.1) was then poured on to the gel for activation of PFI1625c. The gel was equilibrated with activation buffer by constant shaking at room temperature for about 30mins. The gel in activation buffer was kept at 37°C for about 38h. The activation buffer was removed and the gel was stained with coomassie brilliant blue followed by destaining until clear zone was observed. PFI1625c activity in gelatin zymography was performed in 8% SDS PAGE and western blot along with molecular weight marker, negative control BSA and positive control matrix metalloprotease-2 (MMP2).

Table 4.1 Preparation of reagents for gelatin zymography

1. Recipe for 8% resolving gel (10ml)	
Components	**Ml**
Deionized water	3.6
1% gelatin	1.0
30% acrylamide solution	2.7
1.5M Tris pH 8.8	2.5
10% w/v SDS solution	0.1
10% w/v APS solution	0.1
TEMED	0.006

2. Activation buffer	
Components	**Requirement**
Tris buffer pH 7.5	10mM
Triton X-100	1.25%
$ZnCl_2$	1µM
$CaCl_2$	5mM
3. Wash buffer composition	
Triton X-100	2.5% in water

4.2.10 Detection of PFI1625c in malaria culture supernatant using generated anti-PFI1625c antibody: Dot blot was performed to detect the presence of PFI1625c in different samples. For detection of PFI1625c in malaria culture supernatant at different stages, the ring, trophozoite and schizont stages from synchronous *P.falciparum* culture was harvested at about 3-5 % parasitemia and the culture supernatant of different stages was taken for dot blot. For detection in culture supernatant at different parasitemia, asynchronous parasite culture was harvested at different parasitemia (2.5%. 5% and 10%). Pure PFI1625c was used as +ve control and for comparative analysis whereas BSA was used as –ve control. Samples for detection are blotted on nitrocellulose membrane in a small volume at a time using micro tip in order to get intense dot. Each blotting was followed by intermittent drying using hair dryer. After blotting was completed, it was air dried completely at 37°C for 30mins. The dry membrane was blocked in 5% skim milk in PBS for overnight at 4°C. The membrane was washed with wash buffer PBS + 0.05% tween-20 to remove excess of skim milk. Generated 1° antibody at 1:1000 was added onto the membrane and incubated at 37°C for 2hrs. The 1° antibody was then washed off 4-5 times with wash buffer with proper shaking for 5-10mins each. 2° antibody (HRP-conjugated) was then added onto the membrane wrapped in polyethene and incubated 37°C for 2hrs. the membrane was washed 4-5 times as before. Substrate DAB (Diaminobenzedene) at 1mg/ml was prepared in PBS containing 0.03% H2O2. The substrate solution was poured immediately onto the membrane in dark condition and allowed it for about 10mins or until brown coloration developed. For detection of PFI1625c at different stages, chemiluminescent peroxidase substrate was also used for development of blot and was observed under Molecular Imaging System (Gel Logic 1500). Densities of chemiluminescent signals at different stages were analyzed using image analyzing tool using NIH-Image J.

4.3 Results

4.3.1 PFI1625c is protease in nature

Our previous *in silico* studies have shown the similarity of PFI1625c with known proteins such as cytochrome bc-1 core protein of mitochondria from bovine, metalloprotease from yeast and peptidase from bacteria [153]. Proteolytic activity of enzyme was performed to check the enzymatic and protease nature of PFI1625c. Enzymatic nature was checked by using different amount of PFI1625c ranging from 3-40µg with fixed amount of substrate**.** Common protease substrate bovine substrate albumin (BSA) was prepared in 1x PBS.1%

of BSA was used as substrate in 0.5ml reaction volume. The reaction mixture containing substrate and enzyme was incubated at 37˚C for 2hrs and reaction was stopped by precipitating with 10% TCA. Michaelis-Menten curve was drawn from OD values obtained as a result of substrate cleavage. Interestingly, PFI1625c cleaves BSA in a dose dependent manner (Figure 4.1). It confirms enzymatic nature of PFI1625c and cleavage of BSA confirms the claim that PFI1625c is a protease.

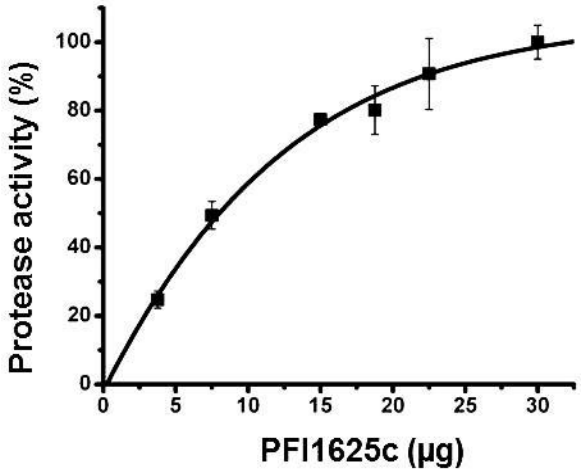


Figure 4.1 Testing the enzymatic nature of PFI1625c as protease. PFI1625c exhibits enzyme dose response curve with standard protease substrate BSA.

4.3.2 PFI1625c prefers hemoglobin and matrix protein gelatin as substrate

PFI1625c is an uncharacterized protease present in *P.falciparum*. Localization of PFI1625c was performed at different stages inside the RBC and it was expressed throughout the stages from ring to schizont. Preliminary Localization indicates that it is a non-nuclear protease. It may be a cytosolic protease or other organelles and matrix protease. Hemoglobin is the most abundant protein found in the food vacuole. It is also found in the cytosol in the form of di-peptides. Proteases are believed to be involved in terminal degradation of hemoglobin inside the food vacuole and in the cytosol [155, 157]. To decipher the substrate preference of PFI1625c, different protein substrates were analyzed including biological substrate hemoglobin and matrix protein gelatin. Other substrates used in the assay were BSA, ovalbumin, skim milk and casein. Each substrate was used at a final concentration of 1% (prepared in assay buffer Tris pH 8.5). To each substrate solution, 20µg of PFI1625c was added. The reaction mixture was incubated for about 2hrs at 37˚C. The reaction was stopped by precipitating with 10% TCA. The change in activity for each substrate was calculated from the OD values by considering PFI1625c activity as 100%. With a fixed substrate concentration (1%), PFI1625c was

found to cleave BSA, ovalbumin, hemoglobin and gelatin. Highest activity was observed, when gelatin was used as substrate. No Activity was found for skim milk and casein (Figure 4.2A). To further test the specificity, Michalis-Menten curve was drawn by taking different substrate concentrations (0.25-10%) of selected substrates (BSA, ovalbumin, hemoglobin and gelatin) with the fixed enzyme concentration (20µg). It was observed that, PFI1625c exhibits a dose response curve when gelatin and hemoglobin used as substrate (Figure 4.2B and C). To determine the Km value for hemoglobin, Lineweaver Burk plot was drawn and it was found to be 48.4µM which is in agreement to the value (~46µM) obtained from Michaelis-Menten curve (Figure 4.2D). BSA and ovalbumin did not show any significant curve as observed with hemoglobin and gelatin (data not shown). It indicates the preference of PFI1625c towards hemoglobin and gelatin as substrate.

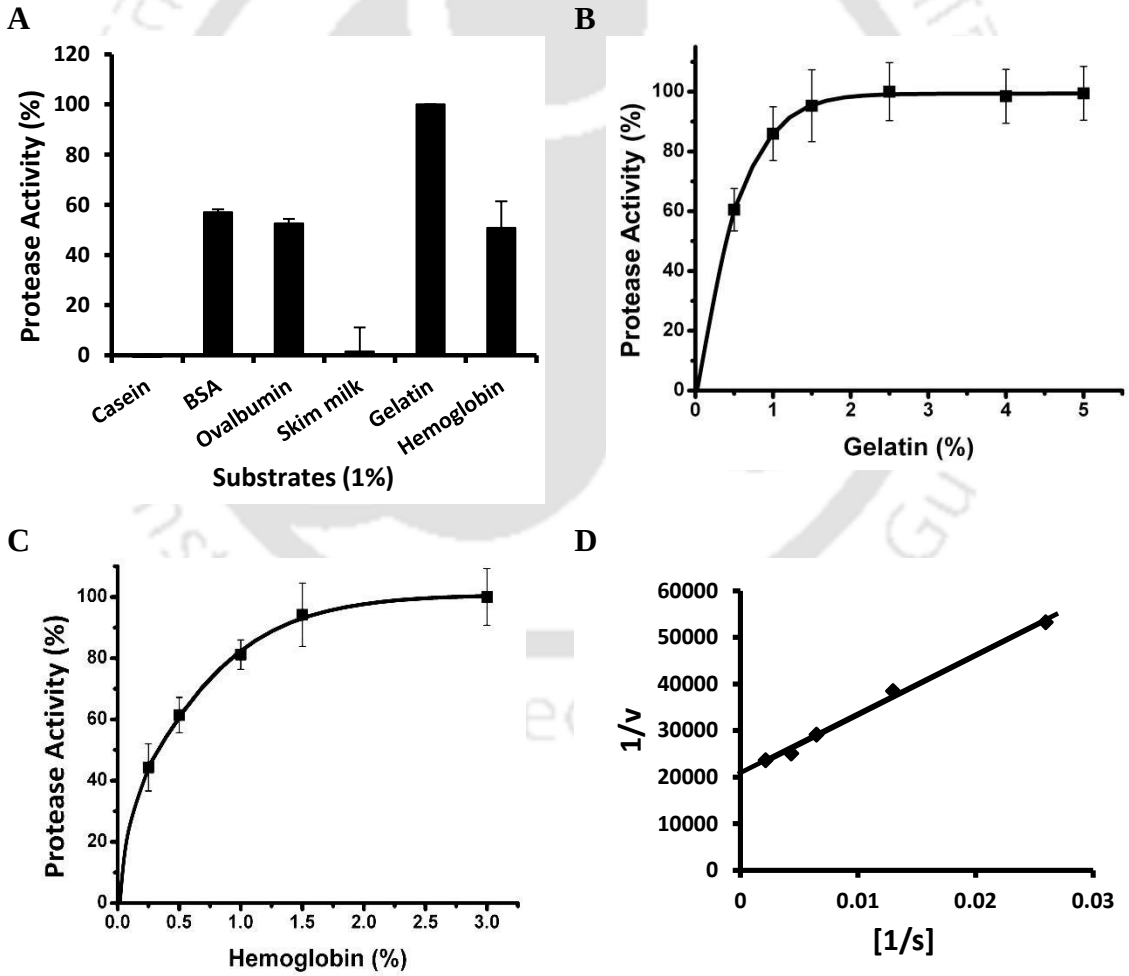


Figure 4.2 Substrate specificity of PFI1625c. **(A)** PFI1625c activity towards different protein substrates (casein, ovalbumin, skim milk, gelatin and hemoglobin). **(B & C)** PFI1625c activity using different concentrations of hemoglobin and gelatin. **(D)** Lineweaver Burk plot of hemoglobin at different substrate concentrations.

4.3.3 Kinetic parameters of PFI1625c catalyzed gelatin degradation

Matrix metalloproteases are slow acting and require substrates in micromolar range for its maximal activity [213]. To determine the Michaelis-Menten constant Km value of PFI1625c, it was incubated with different concentrations of gelatin (0.25-2.5%) for about 2hrs at 37˚C.The reaction was stopped by precipitating with 10% TCA. Michaelis-Menten curve was analyzed from OD values obtained as a result of substrate cleavage. The Km value obtained from Michaelis-Menten curve was 25µM (Figure 4.3(i) A). Lineweaver Burk plot was drawn to get more accurate Km value and it was found to be 30µM (Figure 4.3(i) B) which is comparable with other metalloproteases of malaria parasite. The substrate preference of PFI1625c towards matrix protein gelatin was assayed for gelatin zymography, a standard and reliable assay to show gelatin degradation by matrix metaloproteases. Matrix metalloprotease-2 (MMP-2) was used as positive control whereas BSA was used as negative control. Gelatin zymography showed a clearance zone at the region where PFI1625c protein (mol. wt. 57kDa) was present as shown in Figure 4.3(ii). MMP-2 (mol wt. 63) also showed clear zone just a little above PFI1625c but no clearance zone was observed in BSA. The clear zone formed by PFI1625c activity is significantly good for recombinant proteases. The PFI1625c band was confirmed by SDS PAGE and western blot using anti-His antibody specific for PFI1625c. SDS PAGE showed a PFI1625c protein band at the region where clearance zone was observed in gelatin zymography. BSA protein band is clearly visible in SDS PAGE as shown in Figure 4.3(ii). Western blot showed an intense signal at PFI1625c protein band as it is clear from the molecular weight marker as shown in Figure 4.3(ii).

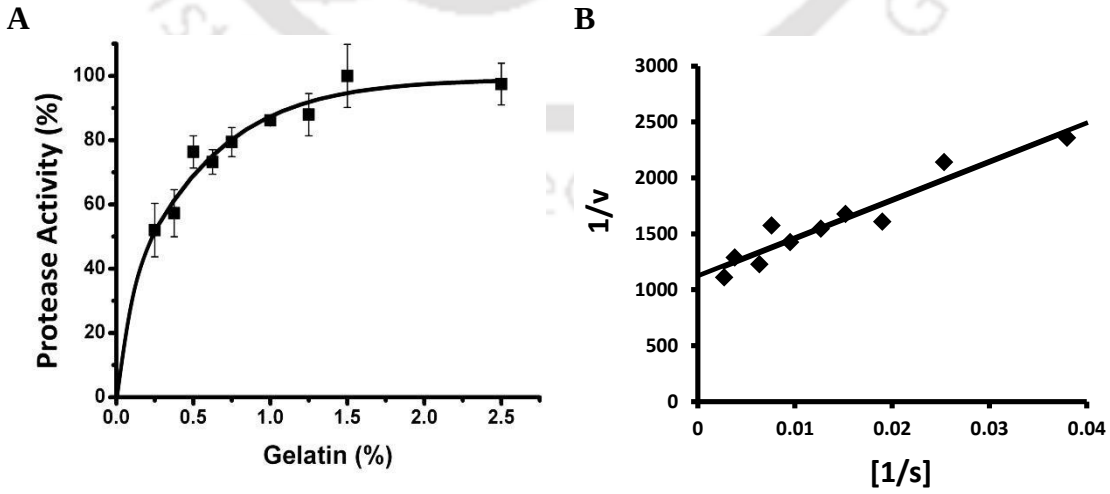


Figure 4.3(i) PFI1625c activity towards gelatin as substrate. **(A)** Michaelis-Menten curve of PFI1625c activity using different concentrations of gelatin . **(B)** Lineweaver Burk plot of PFI1625c with different concentration of gelatin.

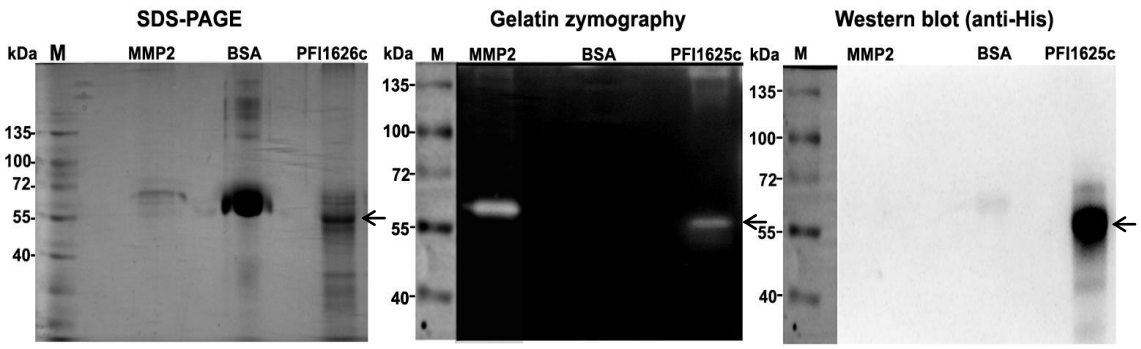


Figure 4.3(ii) Gelatin zymography of PFI1625c using 0.1% gelatin in 8% SDS-PAGE. Arrows indicate PFI1625c band. SDS-PAGE shows the pattern of protein bands used for gelatin zymography and western blot. Gelatin zymography shows a clearance zone in the region where PFI1625c is present as evident from SDS-PAGE and molecular weight marker. MMP2 shows an intense clearance zone whereas BSA does not show any clearance zone. Western blot shows intense signal only at the region were PFI1625c is present.

4.3.4 PFI1625c is a slow acting enzyme

Recombinant metalloproteases are slow acting and require more substrate and incubation time for a significant degradation to occur [165, 214]. Matrix metalloproteases require more amount of enzyme (1µM or higher) to obtain significant maximum velocity and cleavage of collagen type-II [215, 216]. To determine the optimum time point for catalytic reaction, 20µg of PFI1625c was incubated with 1% gelatin at different time points (0-12hrs) at 37°C. The reaction was stopped by precipitating with 10% TCA. The change in activity was calculated by considering highest activity (4hrs) as 100%. The enzymatic activity was calculated from the OD values obtained from reaction supernatant. As observed in the graph, the activity of PFI1625c became stable after 2.5hrs of incubation (Figure 4.4). Hence optimum reaction time for PFI1625c is 2.5hrs and it was followed in the subsequent assays.

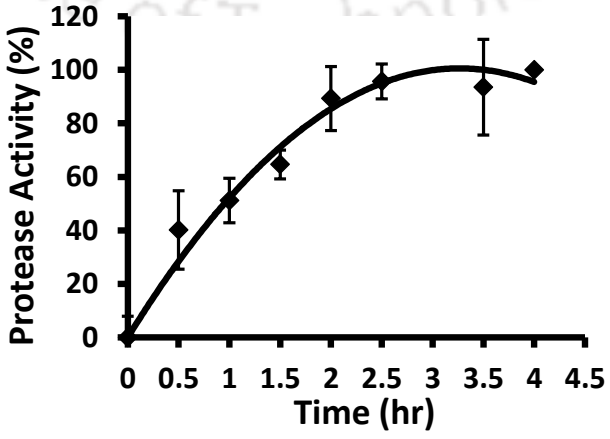


Figure 4.4 Time dependence activity of PFI1625c towards gelatin as substrate.

4.3.5 PFI1625c catalyzed reactions are sensitive to pH and temperature

Protease activity is highly sensitive to pH. Preliminary data by sequence analysis from plasmodium database and Expasy (http://www.expasy.org/) showed that PFI1625c has a pI of 6.78 and 6.34 respectively. To check the effect of PFI1625c activity at different pH, different buffers pH ranging from pH 3 to 12 was used to prepare enzyme and substrate. Gelatin substrate at a final concentration of 1% was added to 20µg of purified PFI1625c enzyme. The reaction mixture (0.5ml) was incubated for about 2hrs at 37˚C. The reaction was stopped by precipitating with 10% TCA. The change in activity was calculated by considering highest activity (pH 8.5) as 100%. Optimum activity of PFI1625c was observed at pH 8.5 (Figure 4.5A). It also corresponds to the localization studies that PFI1625c may be a cytosolic protease. M17 is a cytosolic metalloprotease with optimum activity at pH at 8.5 [165]. Another metalloprotease falcilysin which cleaves hemoglobin oligopeptides has dual activity in acidic and alkaline environment [151].

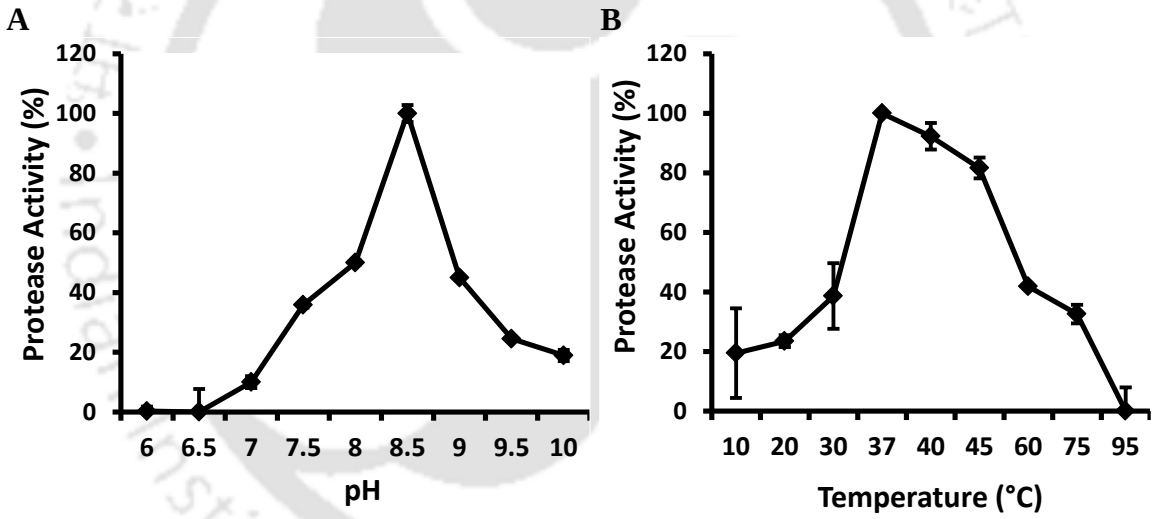


Figure 4.5 Effect of pH and temperature on PFI1625c protease activity. (A) pH effect showing optimum activity at pH 8.5. **(B)** Temperature effect showing optimum temperature at 37°C.

Temperature is a significant parameter in malaria proteases. Proteases which are active inside the invertebrate host have optimum activity at around 26°C whereas proteases which are active inside the human host have optimum activity at normal body temperature 37°C or during high fever at 39°C. Therefore the effect of temperature on PFI1625c was performed by incubating 20µg enzyme with 1% gelatin substrate at different temperature range from 4-100 ˚C for about 2hrs. The reaction was stopped by precipitating with 10% TCA. The change in activity was calculated by considering highest activity (37˚C) as 100%. The activity of PFI1625c was optimum at 37°C (Figure

4.5B). This shows that PFI1625c is active inside human host. PFI1625c has 90% activity up to 40°C (Figure 4.5B) indicating that it is active during high fever conditions. Some malarial metalloproteases are reported to have optimum temperature during high fever conditions such as M18 aminopeptidase [156].

4.3.6 PFI1625c catalyzed reactions are sensitive to metals

Members of the M16 family are metalloproteases and therefore require metal ions for their enzymatic activity and stability. Different divalent metal ions enhance enzyme activity. Transition metal such as Co^{2+}, Mn^{2+} and Zn^{2+} enhance activity of falcilysin and M17 aminopeptidase [154]. But they are zinc metalloproteases as Zn^{2+} exhibit dose response curve and activity decreases with further increase in metal concentration. To test whether PFI1625c requires metal ions for its activity, different metals (Zn^{2+}, Mn^{2+}, Mg^{2+}, Cu^{2+}, Ca^{2+}, Fe^{2+}, Hg^{2+}) at fixed concentration (100µM) were pre-incubated with 20µg of PFI1625c for 30mins at 37°C. After pre-incubation, 1% gelatin was added to the reaction mixture and was incubated for about 2hrs at 37˚C. The reaction was stopped by precipitating with 10% TCA. The change in activity was calculated by considering reaction without metal as 100%. From all the metals tested, Mn^{2+}, Cu^{2+}, Hg^{2+} and Zn^{2+}ions enhanced PFI1625c activity by 20-40% with respect to activity without addition of metal whereas Fe^{2+} and Ca^{2+} were observed to have inhibitory effect (Figure 4.6A). Fixed concentrations of metals could not give conclusive result until concentration curve was performed. Therefore, dose-dependent curve was drawn for different metal ions (Mn^{2+}, Cu^{2+}, Hg^{2+} and Zn^{2+}) by taking different concentrations of metals (0-150µM). Different concentrations of metal ions were pre-incubated with 20µg of PFI1625c. After pre-incubation, 1% gelatin was added to the reaction mixture and incubated for about 2hrs at 37˚C. The reaction was stopped by precipitating with 10% TCA. The change in activity was calculated by considering reaction without metal as 100%. Among the metals tested, only Zn^{2+} ion exhibited a dose dependent response on PFI1625c activity. The activity was optimum at 70µM of Zn^{2+}concentration with 80% increase in PFI1625c activity (Figure 4.6E). Other metals did not show dose dependent response (Figure 4.6B-D). This means Zn^{2+} stabilize and render proper conformation of enzyme active site by coordinating with the zinc binding motif and residues as discussed in the previous chapter (Figure 4.6F). Further increase in Zn^{2+} became inhibitory to PFI1625c activity which is typical characteristic of zinc metalloproteases [217].

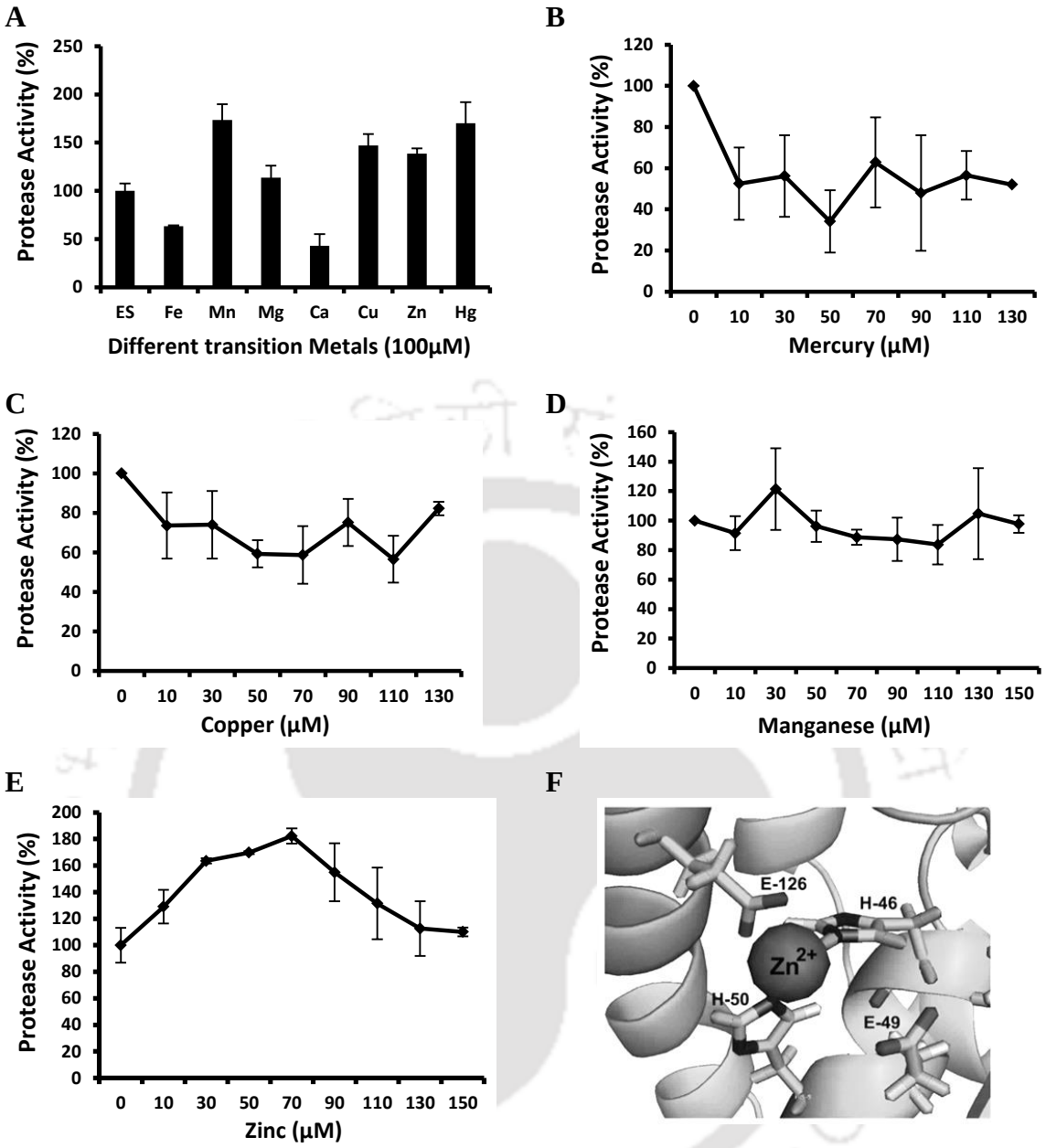


Figure 4.6 Effect of transition divalent metals on PFI1625c activity. (A) Effect of different metal on PFI1625c activity. ES=without metal. **(B-E)** Dose response curve of selected metal ions (Hg, Cu, Mn and Zn) using 0-150µM. **(F)** Predicted mechanism of Zn^{2+} binding on PFI1625c metal binding motif.

4.3.7 PFI1625c is inhibited by 1,10-Phenanthroline

Zinc dependent metalloproteases are sensitive to metal chelators particularly 1,10-Phenanthroline. Their activities are abolished completely with increased concentration of chelator [168, 218]. To test the dependence of PFI1625c activity on Zn^{2+}, zinc metal chelator (1,10-Phenanthroline) at different concentrations ranging from 0-10mM was pre-incubated with 20µg enzyme in presence of 10µM of Zn^{2+} for about 30mins at 37°C. After pre-incubation, 1% gelatin substrate was added to the reaction mixture (0.5ml) and

incubated for about 2hrs at 37°C. The reaction was stopped by precipitating with 10% TCA. The change in activity was calculated by considering reaction without chelator as 100%. From the graph, 1,10-Phenanthroline was shown to cause significant decrease in PFI1625c activity. Complete inhibition was obtained at 2mM while IC_{50} was found to be 0.8mM as calculated from the graph (Figure 4.7). Effect of EDTA on PFI1625c activity was also tested in similar manner but no inhibition was observed in EDTA treated assay (data not shown). The inhibition of PFI1625c activity only by 1,10-Phenanthroline shows the specificity towards PFI1625c. It also shows that PFI1625c requires Zn^{2+} for its activity which is in agreement from the *in silico* data that PFI1625c has a zinc binding motif and confirmation from biochemical data that is it a metalloprotease.

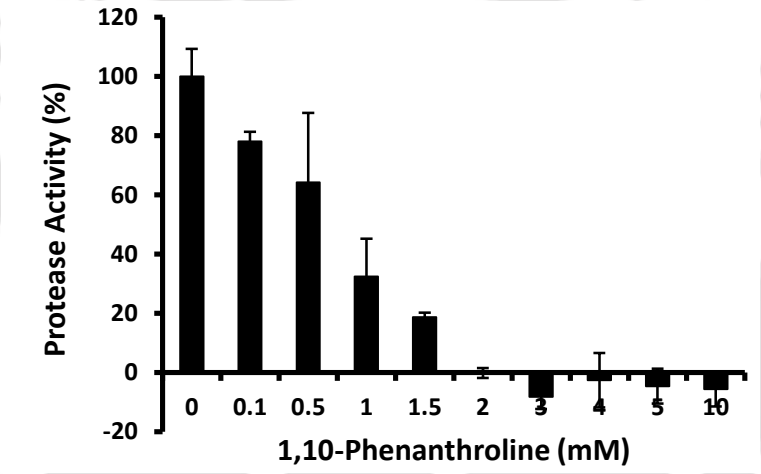


Figure 4.7 Effect of 1,10-Phenanthroline on PFI1625c protease activity. Activity decreases with increase in chelator concentration with a complete inhibition of enzyme activity is observed at 2mM.

4.3.8 PFI1625c catalyzed reactions are sensitive to detergents

The effects of detergents on metalloproteases activity have been studied. Detergents significantly enhance metalloproteases activity [219]. Some membrane proteases which have more exposed hydrophobic moiety require surfactants for enzyme solubility and activity. Understanding the effect of detergents on known metalloproteases, PFI1625c activity was checked for different detergent effect. For this, different detergents such as SDS, Tween-20, Tween-80, Triton x-100 and CTAB at 0.5% were incubated with 20µg of PFI1625c and 1% gelatin at 37°C for about 2hrs. The reaction was stopped by precipitating with 10% TCA. The change in activity was calculated by considering reaction without detergent as 100%. Interestingly detergents significantly enhance PFI1625c activity. Among the detergents used, SDS, tween-20 and CTAB enhances enzyme activity by 46%, 49% and 18% respectively (Figure 4.8). This shows that under

lipid or microenvironment PFI1625c activity is enhanced which is in consistence with matrix metalloproteases [220].

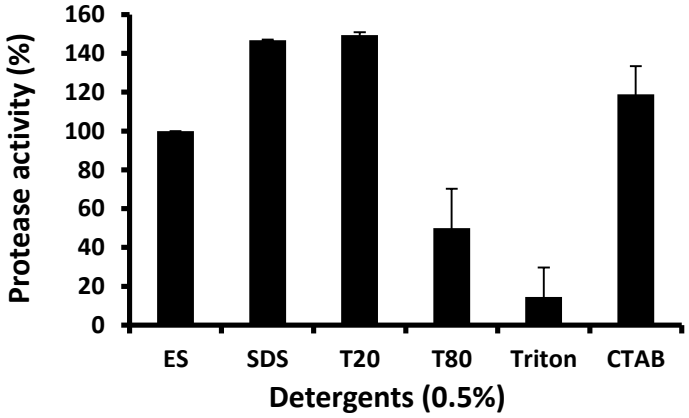


Figure 4.8 Effect of different detergents on PFI1625c protease activity. ES=enzyme substrate without detergent. SDS=Sodium dodecyl sulphate, T20=Tween-20, T80=Tween-80, Triton=Triton X-100, CTAB=Cetyl trimethylammonium bromide.

4.3.9 PFI1625c is secretory in nature

Dot blot is a simple but robust technique for detection of antigens or proteins present inside and outside the cell. Detection of secretory proteins is useful in malaria diagnostic system. Preparation of detection kits have been studied for malaria parasites using parasite proteins but potential proteins are non-secretory [221]. Immuno-staining technique is limited to staining of proteins present in stained cells but it cannot give clear information for proteins which are secreted outside the cells. Therefore to detect the secretion of PFI1626c outside the host cells, dot blot was performed and developed using DAB as substrate. The dot blot of parasite supernatant at different parasitemia showed signals in all the parasitemia which was highest in 2.5% followed by 10% and lowest in 5% parasitemia. BSA and culture media (RPMI) did not show any signal (Figure 4.9A). The dot blot, at different stages also showed intense signal in all the stages. The blot was developed by two peroxidase substrates (DAB and chemiluminescent). Chemiluminescent signals (grey scale) were used for analyzing the signals obtained from different stages using image analyzing tool (NIH-Image J). The grey scale of the dots was submitted to Image J tool (Figure 4.9B) and densities of the dots were measured and analyzed. The percent value of each blot was divided by the percent value of standard (in this case PFI1625c) and the resulting values gave measure of the relative density of each peak (Figure 4.9C). Based on the image J calculation, the concentration of PFI1625c in the parasite supernatant was 10.2, 14.5 and 13.9 ng/µl for ring, trophozoite and schizont stage respectively (Figure 4.9D). According to image analysis, trophozoite stage has the

highest amount of secreted PFI1625c which is also evident from visual observation. Secretory nature of protein strongly implies their potential role in the parasite biology. This will allow PFI1625c as a drug target candidate but also may be usable as detection kit for malaria parasites. PFI1625c was present in all the stages which become more peripheral as the stages progress and it is also detectable in outside the cell which is more obvious in the later stage. This shows the possible involvement of PFI1625c during invasion, egress and outside the RBC with other host cells.

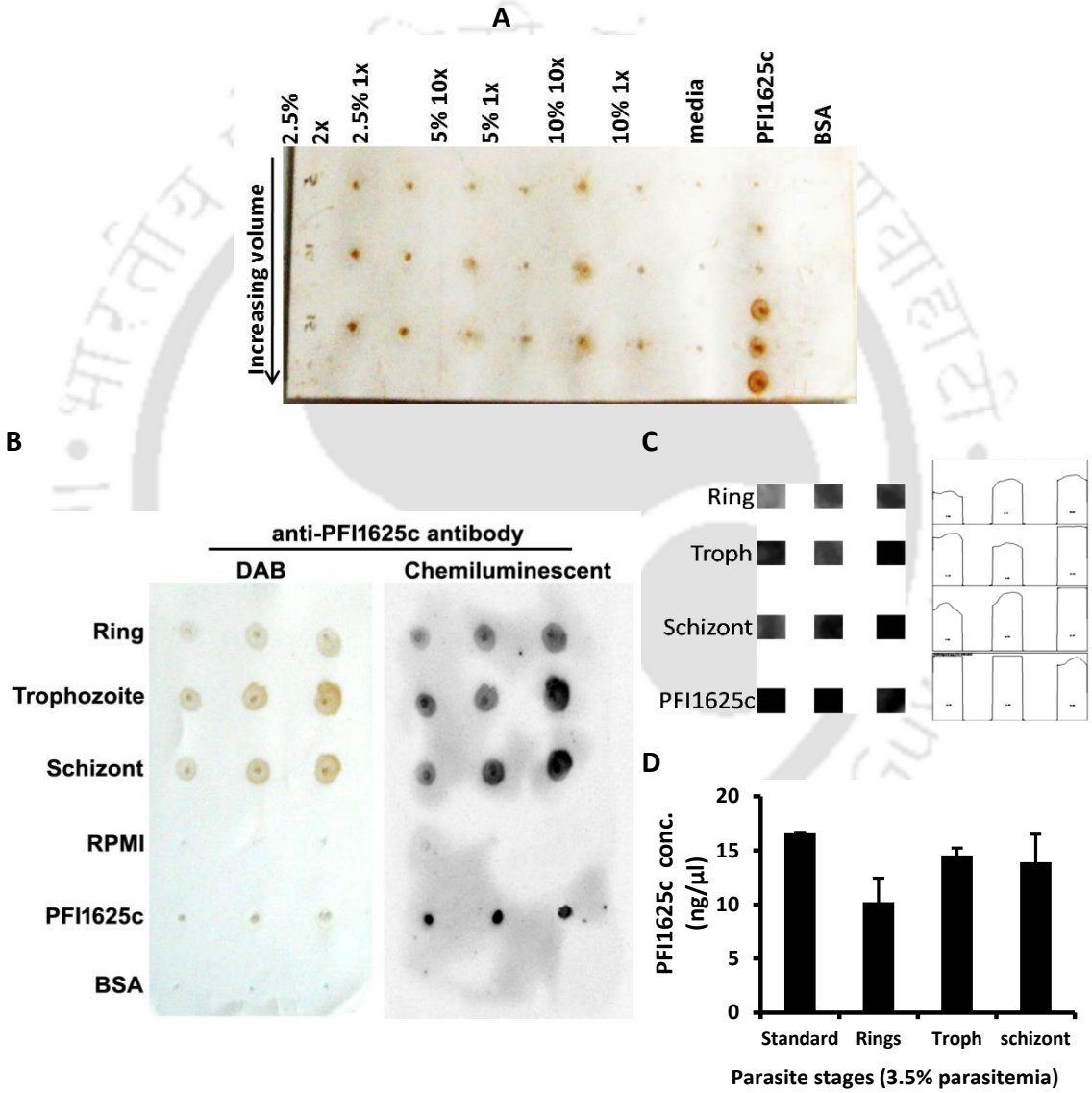


Figure 4.9 Detection of PFI1625c in parasite culture supernatant. **(A)** Dot blot at different parasitemia. Signal is observed in all the three parasitemia, +ve control (PFI1625c) but media (RPMI) and BSA do not show any signal **(B)** Dot blot at different stages showing both DAB and chemiluminescent signals. All the stages and +ve control show intense signal whereas RPMI media and BSA did not show any signal. **(C)** Chemiluminescent signal was used in Image J tool which generated density graph of different stages obtained from Image J tool. **(D)** Estimation of PFI1625c present in different stages using standard PFI1625c.

4.4 Discussion

PFI1625c shows optimum activity at pH 8.5 provides evidence that it is a mitochondrial or apicoplast protein. Although the actual pH of these two organelles are unknown, there are enzymes reported to localize in these organelles that work optimally at pH 8.5 [222]. Our present studies have proven that PFI1625c belong to metalloprotease class with the enhancement of activity with addition of zinc ion. This shows that PFI1625c is a zinc metaloprotease. The Km value of both Michaelis-Menten curve and Lineweaver Burk plot show a very close value of 25µM and 30µM respectively which is a good correlation. Plasmodium cytosolic metalloproteases M17aminopeptidase and PfMetAP1b are also seen to have high Km value of 12.12µM and 327.3µM respectively [165, 169]. Recombinant gelatinases are reported to show high Km values in the order of 10-100µM [223]. Incubation of enzyme for a longer period (2hrs) was required to obtain measurable readouts as substrate turnover may be very low. Detection of PFI1625c in parasite culture supernatant was performed in two ways. Firstly, to have initial evidence for presence of PFI1625c in culture supernatant regardless of stages, different parasitemia was taken for preliminary investigation. The second method was performed for different stages to obtain more specific and reliable result. Our present study showed the secretion of PFI1625c outside the parasite shows the possible involvement of PFI1625c during invasion step. Several secretory proteins have been studied for diagnostic tool but so far have not proven accuracy as false positive results are observed as in case of HRP-II [224]. The secretory nature of PFI1625c may be a good candidate for development of diagnostic or detection kit. Several crucial functions have been indirectly implicated for metalloproteases in the life cycle of *P.falciparum* besides its involvement in metabolic pathways. Metalloprotease inhibitor 1,10-phenanthroline was found to inhibit invasion of merozoites to new RBC [179]. 1,10-phenanthroline also inhibit the several proteases indirectly linked to invasion step [154, 155]. Besides asexual RBC stages, metalloprotease inhibitor 1,10-phenanthroline was also found to inhibit proteolysis during gametogenesis and exflagellation formation in *P.berghei* [225, 226]. Although metalloproteases inhibitor (1,10-Phenanthroline) inhibits all these pathways, the protease target is still unknown. It is plausible that PFI1625c may be a possible target as it is also effectively inhibited by 1,10-Phenanthroline. The ability of PFI1625c to cleave gelatin as obvious from gelatin zymography reveals that it may be a matrix metalloprotease. Recently there have been a number of evidences from patients with cerebral malaria that matrix metalloproteases are involved in the pathogenesis of severe malaria. Detergents

also significantly enhanced PFI1625c activity as shown by SDS, Tween-20 and CTAB (Figure 4.8). This is a characteristic feature exhibited by human matrix metalloproteases which play significant roles in development, inflammation, morphogenesis and cancer invasion [220]. Enhanced level of MMP-1 was observed in the brain of patients with severe malaria [227]. There is an increase level of matrix metalloprotease MMP-8 and its inhibitor TIMP-1 in severe malaria [228, 229]. Additionally, it has been found that activation of MMP-2, MMP-7, pro-MMP-9 and MMP-9 was observed in the non-human models of cerebral malaria i.e., in the brain of mice infected with *P.berghei* ANKA [230, 231]. So far there has been no report or information whether *P.falciparum* matrix metalloprotease contributes to the severity of severe malaria caused by matrix matalloproteases. All these indirect evidences paves a way that PFI1625c could be a powerful tool for investigating the role of *P.falciparum* matrix metalloprotease pathway and its role in particularly invasion step in pathogenesis of severe malaria. Therefore we can further characterize the inhibitory effects of PFI1625c*in vitro* and *in vivo* using different inhibitors and also investigating the presence of PFI1625c outside parasite which would increase our understanding towards PFI1625c in the pathogenesis of severe malaria and its potential as a drug target.

References: Please refer to bibliography section from page number 108-119.

Chapter-5: Designing and Screening Potent Inhibitor against PFI1625c and Understanding the Mode of Inhibition

Designing and Screening Potent Inhibitor against PFI1625c and Understanding the Mode of Inhibition

5.1 Introduction

Metalloproteases are a class of proteases present in the *P.falciparum* and they are mostly uncharacterized. Potent metalloprotease inhibitors block critical pathways of parasite and result in reduced growth or death of parasite. Several indirect evidences from metalloprotease inhibitors studies show the involvement of uncharacterized metalloproteases with multiple pathways essential for parasite survival and invasion [179]. PFI1625c is a zinc metalloprotease and is expressed throughout the developmental stages inside the RBC. In this chapter, I have discussed designing suitable inhibitors against PFI1625c from peptide library, bioactive compounds and heterocyclic compound library. Additionally, I discussed the effect of inhibitors on PFI1625c protease activity and mode of inhibition.

5.2 Experimental procedures

5.2.1 Molecular Modeling and generation of combinatorial peptide library:

Signal peptide (LSRVAKRA) of malate dehydrogenase shows binding in PFI1625c in a similar conformation as present in the crystal structure of yeast mitochondrial processing protease (PDB Accession Code 1HR9). The signal peptide of malate dehydrogenase (LSRVAKRA) was used as a template to prepare a peptide library to optimize the binding of peptide into the PFI1625c active site. A random single or double substitution into the template peptide sequence (LSRVAKRA) was done to prepare combinatorial peptide library of 733 peptides representing different types of probable substitution. The type of substitution was hydrophobic, neutral and long chain hydrophobic and charged residues. After critically analyzing complete peptide library and excluding similar substitution with weak interactions, 199 peptides were used for molecular modeling. 3-D structure modeling of peptide sequence from the library was done with the help of EasyModeller 2.0 [31].

5.2.2 In-Silico screening of peptides : The individual peptide from the combinatorial peptide library was fitted into the PFI1625c 3-D structure using Patchdock [232]. Fire Dock program is used initially to perform energy minimization of best 20 models to get stable PFI1625c-signal peptide complexes [233]. Atomic contact energy (ACE) and local proximity to the predicted binding sites in a PFI1625c active site of each

conformation was analyzed. Control Docking of signal peptide (LSRVAKRA) in yeast mitochondrial processing peptidase (PDB Code 1HR9) using Patchdock shows, peptide binds in the extended con-formation as given in co-crystal structure [200]. The root mean square deviation between generated model and crystallographic complex (1HR9) is 1.52Å. To validate the approach, docking was also performed by taking falcilysin and its peptide substrates obtained from literature [151].

5.2.3 Collection of bioactive peptides: Bioactive peptides were collected by thorough literature search and from bioactive peptide database (Bioactive Polypeptide Database-http://biopd.bjmu.edu.cn). They are mostly isolated from natural sources like plants, fungi and animals. In animals they are mostly present in the immune system. Those which have antimicrobial activity are further studied. Out of those hundreds of bioactive peptides, sequence similarity was performed using ClustalW2 tool (http://www.ebi.ac.uk/Tools/msa/clustalw2/) against the optimized hydrophobic peptide P550 [153]. Peptides which have above 50% sequence similarity were selected for the study. Among the 33 selected peptides, 20 were found to have antimalarial activities which were further used for docking against PFI1625c (Table 5.5). The antimalarial peptides and their sources are discussed below;

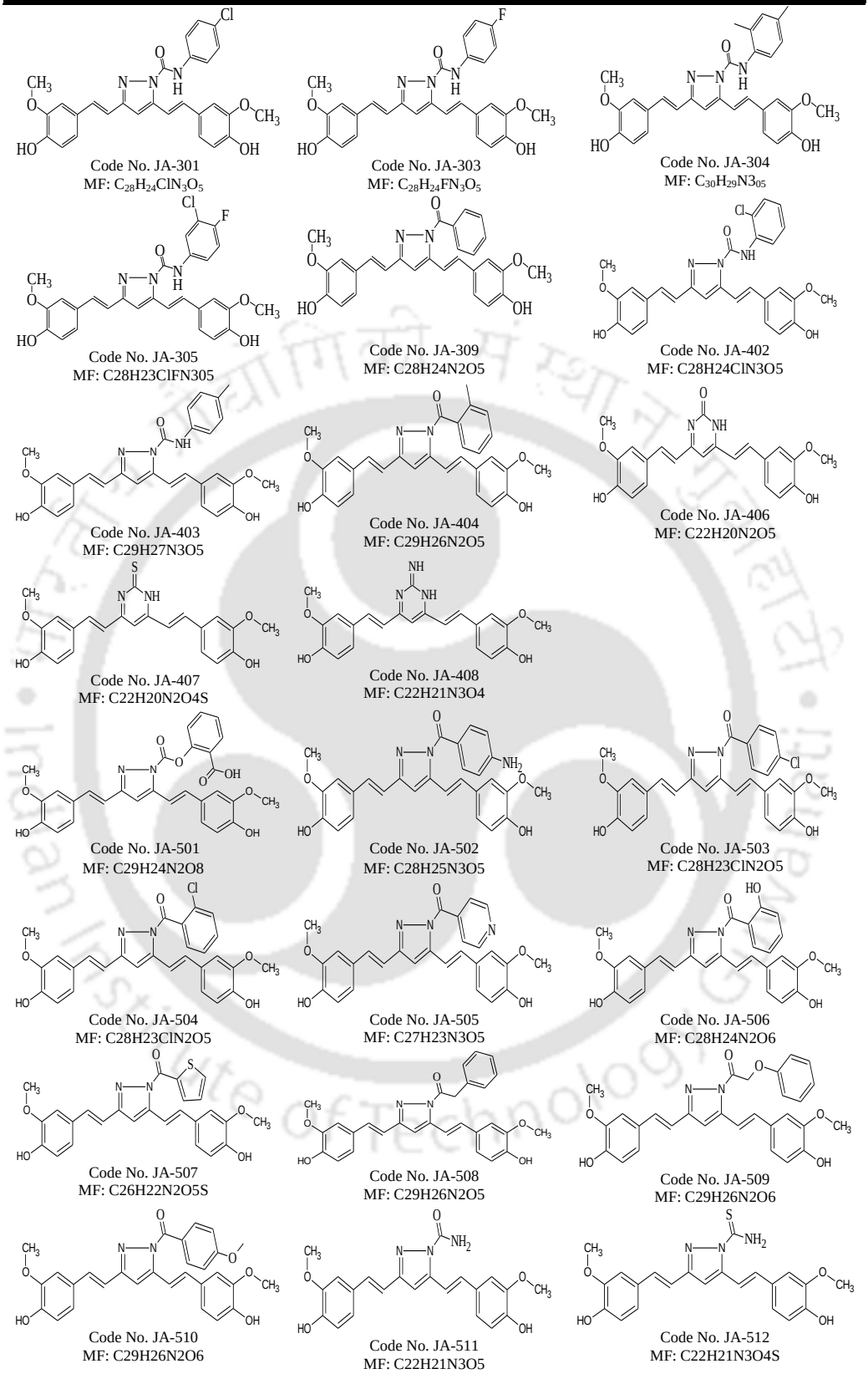
According to the source of origin, we classified the collected antimalarial peptides into the following categories. Peptides which are obtained from frog skin comprise Dermaseptin S3 and S4, and Magainin [234, 235]. Few peptides such as LVDAA and Dermaseptin derivative are chemically synthesized from peptide library or derivatives of natural peptides [236, 237]. Spider toxin is also a source of antimalarial peptides namely PcFK2 and PcFK1 which means *Psalmopoeus cambridgei falciparum killer* [238]. Peptides which are derived from insects consist of Melittin, Cecropin, Defensin. Melittin is derived from bee venom while Cecropin and Defensin are of silk moth and dragon fly immune system origin respectively [235, 239, 240]. Interestingly scorpion venom also produces antimalarial peptides such as Scorpine and Meucin 24/25 [241, 242]. Fungal metabolite also contains antimalarial peptides like Cyclosporin A and Efrapeptin [243, 244]. Few peptides are obtained from bacterial origin namely Defensin-like peptide, Gramicidin D and Thiostrepton and the bacteria which produce them are *Anaeromyxobacter dehalogenans* and *Bacillus brevis* and *Streptomyces* respectively [245-247]. Last but not the least, a human milk and liver also produce antimalarial peptides known as Lactoferrin and Angiotensin II [248, 249].

5.2.4 Molecular modeling of 3-D peptide models: 3-D model of several bioactive peptides with antimalarial activity were already available in the protein database (Protein Data Bank: (http://www.rcsb.org/pdb/home/home.do) but most of peptides did not have their 3-D models. After collection of bioactive peptide sequences, each peptide which has no 3-D model was blasted in NCBI BLASTP against PDB to obtain a template. The hits which showed highest sequence similarity were used as template for modeling of peptide using EasyModeler 2.0 [250] and energy minimization was performed in Swiss-Pdb viewer 4.0.3 (http://www.expasy.org/spdbv/). The modeled peptides were used for docking against PFI1625c using Patchdock.

5.2.5 Molecular modeling of PFI1625c-bioactive peptide models: The bioactive peptides with antimalarial effect were modeled with PFI1625c using Patchdock online server (bioinfo3d.cs.tau.ac.il/PatchDock/). This was performed to study the binding affinity of antimalarial peptides towards PFI1625c, for the targets of those antimalarial peptides were not specified. The top 20 hits from Patchdock result of each peptide were taken and the docking results were analyzed based on geometric shape complementarity score and Atomic Contact Energy (ACE).

5.2.6 Molecular modeling and in-silico screening of heterocyclic compounds against PFI1625c: Curcumin is a natural compound isolated from rhizome of *Curcuma longa* (turmeric). Traditionally, it has been used as medicine for its anti-inflammatory, anti-microbial, antioxidant properties etc. Curcumin and curcumin analogs have been reported to have anti-malarial activity towards chloroquine resistant strain of *P.falciparum* [251]. To isolate the suitable inhibitors against PFI1625c, 23 heterocyclic curcumin analogs were used for docking with PFI1625c. The molecular structure and molecular weight of these compounds are given in Table 5.2. The structures of compounds were drawn in ChemBioDraw (Perkin Elmer software) to firstly obtain the 2D structure from molecular structure. The 2D structures were converted into 3D structure using Chem3D software (http://www.cambridgesoft.com/Ensemble_for_Chemistry/ChemBio Office/). The 3D structures were energy minimized using Chem3D MM2 and MMF94 energy minimiser. After energy minimization the 3D structures were saved in PDB format. These curcumin analogs were docked with PFI1625c using AutoDock 4.0 program suite of MGL Tools 1.5.4 software (http://autodock.scripps.edu/resources /raccoon). A protocol was followed as described in [252]. The 3D co-ordinate of the curcumin analogs were checked for polar hydrogens

Table 5.1 List of synthetic curcumin analogs screened against PFI1625c



and atomic charges and flexible torsions were defined. The resulting PDBQT file was used as ligand for docking. For macromolecule PFI1625c, the polar hydrogen, Kollman charges and atomic solvation were assigned and a grid of 0.375Å resolution was centered

using an autogrid module of Autodock 4. Parameters such as number of evaluation, generation and GA run were set at 2500000, 27000 and 20 respectively. Finally, the hit compounds were selected based on binding energy and interaction.

5.2.7 Effect of Inhibitors on PFI1625c protease activity: To test the effect of inhibitors (from combinatorial peptide, bioactive compounds or heterocyclic compounds), inhibitors were dissolved in suitable solvent system at indicated concentrations were pre-incubated with PFI1625c at 37°C for 30mins. After pre-incubation, 1% gelatin prepared in Tris buffer (pH 8.5) was added to the reaction mixture (0.5ml reaction volume) and incubated at 37°C for 2hrs. The reaction was stopped by precipitating with 10% TCA. The precipitate was centrifuged at 13,000 rpm for 10mins and supernatant was collected for OD measurement at 280nm using UV-Visible spectrophotometer (Cary 100 Bio).

To test the nature of inhibition (reversible or irreversible) on PFI1625c activity, suitable concentration of inhibitor was pre incubated with PFI1625c at 37°C for 30mins in a reaction volume of 1ml. The reaction mixture was passed through NAP-10 desalting column (GE Healthcare Life Sciences). The fraction containing the enzyme was collected and activity assay was performed by incubating with 1% gelatin at 37°C for 2hrs. Control samples with and without inhibitor were also incubated under same conditions and same concentrations. The change in activity for each reaction was calculated by considering PFI1625c activity as 100%.

5.3 Results

5.3.1 Identification of inhibitors against PFI1625c

5.3.1.1 PFI1625c prefers hydrophobic substituted peptides from combinatorial peptide library

Peptide library was prepared from a malate dehydrogenase signal peptide (LSRVAKRA) co-crystallized with yeast mitochondrial processing peptidase (1HR9). As this yeast MPP was used a template for PFI1625c model building, a bound peptide would be an ideal base for preparation of peptide library to design inhibitors. Out of 733 peptides, 199 peptides were selected for docking with PFI1625c using Patchdock. The top 20 hits from Patchdock result of each peptide were taken and the docking results were analyzed based on geometric shape complementarity score and Atomic Contact Energy (ACE). The top 5 peptides (with –ve value ACE) from each substitution is highlighted in Table 5.2.

Table 5.2 Docking result of best interacting peptides selected from Combinatorial Peptide library

Peptide Code	Peptide sequence	Geometric Score	Atomic Contact Energy (kcal/mol)	Mode of interaction
Template	LSRVAKRA	7806	-69.85	Binds at catalytic site
Similar group substitution				
P6	GSRVAKRA	7468	-45.47	Close with polar contact
P25	LTRVAKRA	7838	-42.27	Away from catalytic site
P35	LSRLAKRA	6932	-96.96	Away from catalytic site
P73	FTRVAKRA	7678	-42.52	Close with polar contact
P91	1TRVAKRA	8194	-91.46	Close with polar contact
Negative charge substitution				
P29	LSDVAKRA	7464	-12.46	Away from catalytic site
P30	LSEVAKRA	6752	-27.62	Close without polar contact
P518	LDHVAKRA	7326	-31.83	Away from catalytic site
P520	LDDVAKRA	7518	-9.78	Away from catalytic site
P528	LEEVAKRA	7396	-6.81	Close with polar contact
Positive charge substitution				
P22	LRRVAKRA	7366	-51.98	Close with polar contact
P152	WHRVAKRA	7152	-114.38	Close without polar contact
P504	LKHVAKRA	6840	-56.8	Close with polar contact
P511	LRHVAKRA	7192	-70.38	Away from catalytic site
P556	LKRLAKRA	7868	-68.89	Away from catalytic site
Non-polar substitution				
P536	LIIVAKRA	6690	-144.42	Close with polar contact
P540	LLKVAKRA	6734	-174.58	Close but no polar contact
P543	LLIVAKRA	7136	-220.83	Close without polar contact
P550-r1	LVIVAKRA	7568	-262.47	Close with polar contact
P550-r11	LVIVAKRA	7058	-183.36	Close with polar contact
Polar substitution				
P19	YSRVAKRA	7136	-3.81	Away from catalytic site
P55	CTRVAKRA	6820	-68.23	Close with polar contact
P163	YTRVAKRA	8114	-69.44	Away from catalytic site
Negative + Positive charge substitution				
P499	LHDVAKRA	7812	-55.05	Away from catalytic site
P500	LHEVAKRA	8196	-64.72	Away from catalytic site
P507	LKEVAKRA	6924	-25.28	Close with polar contact
P513	LRDVAKRA	6850	-99.2	Close with polar contact
Negative charge + polar substitution				
P61	DTRVAKRA	7004	-9.5	Away from catalytic site
P300	YSEVAKRA	7338	-6.33	Close with polar contact
Positive charge + non polar substitution				
P502	LHLVAKRA	7146	-9.74	Close with polar contact
P503	LHVVAKRA	7000	-72.03	Close with polar contact
Positive + Polar (No negative ACE values)				No –ve ACE value
Negative charge + non-polar				
P523	LDLVAKRA	6688	-93.76	Away from catalytic site
P534	LIDVAKRA	6784	-224.67	Close but no polar contact
P542	LLEVAKRA	7230	-77.78	Away from catalytic site
P548-r19	LVDVAKRA	6850	-262.81	Close with polar contact
P548-r2	LVDVAKRA	7502	-147.11	Close without polar contact

The overall screening result showed that hydrophobic substituted peptides gave the best interacting peptides. All the top 10 peptides from 199 total peptides bind well at the active

site cavity of PFI1625c (Figure 5.1). Out of 21 peptides with hydrophobic substitution, 8 peptides gave significantly better interaction compared to template peptide with ACE -69.85 (Table 5.3). P550 (LVIVAKRA) shows the best interaction with ACE of -262.47 and Score of 7568. The scores and ACE values of top 10 best peptides from total of 199 peptides are shown in Table 5.3.

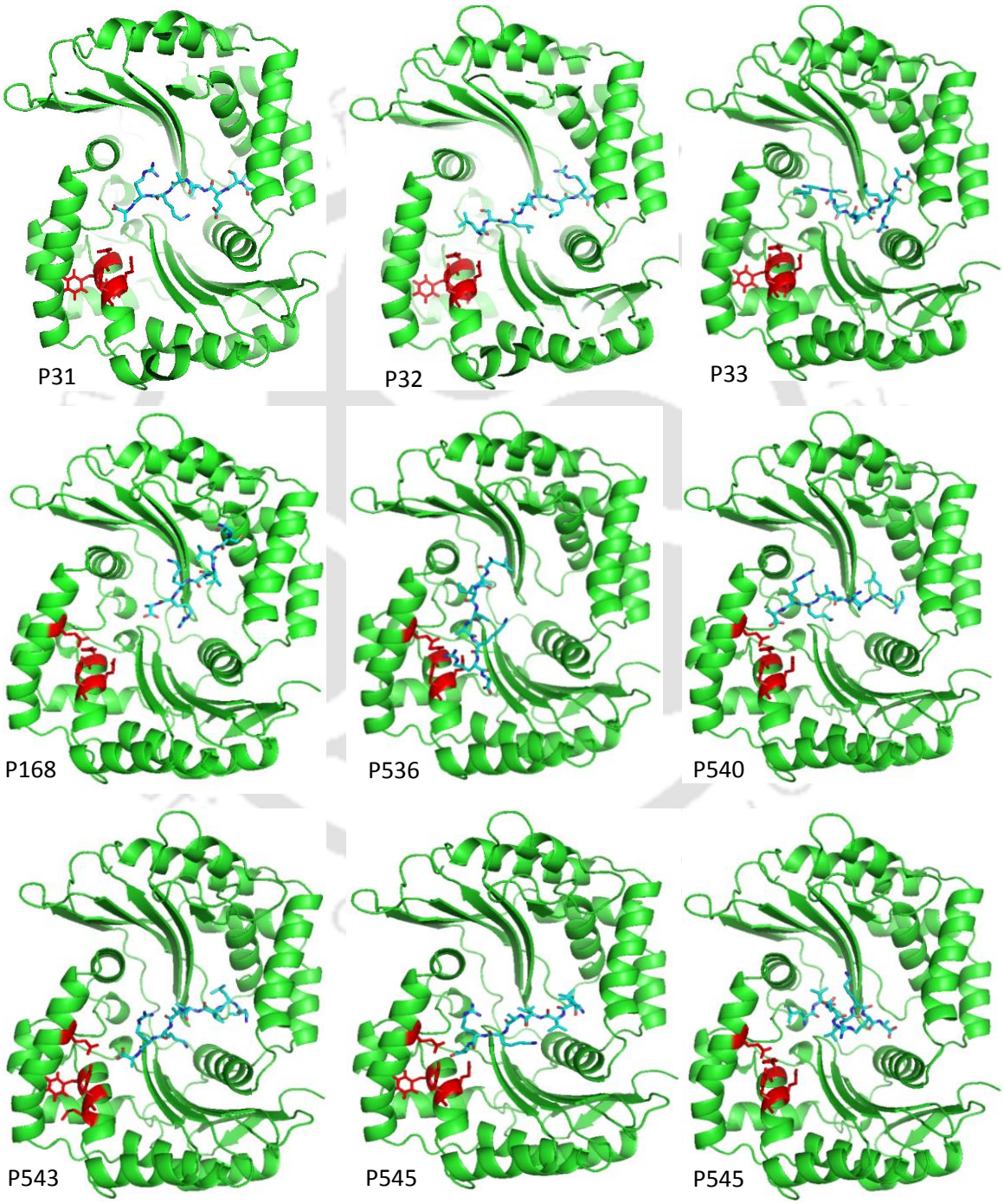


Figure 5.1: Interaction of PFI1625c with top hits peptides from Combinatorial Peptide library. Peptides and zinc binding residues are shown in cyan and red respectively.

Table 5.3 Analysis of top hit peptides from combinatorial peptide library

Peptide Code	Amino acid sequence	Geometrical Score	ACE (kCal/mol)
Template	LSRVAKRA	7806	-69.85
P548	LVDVAKRA	6850	-262.81
P550	LVIVAKRA	7568	-262.47
P543	LLIVAKRA	7136	-220.83
P540	LLKVAKRA	6734	-174.58
P536	LIIVAKRA	6690	-144.42
P33	LSVVAKRA	6602	-114.1
P32	LSLVAKRA	6928	-100.12
P31	LSIVAKRA	7174	-98.26
P168	ASIVAKRA	7304	-82.26
P545	LLVVAKRA	6688	-73.05

The best interacting peptide P550 showed high binding affinity and strong atomic contacts compared to the template. The hydrophobic part of the peptide lies at the hydrophobic interior region of PFI1625c which indicates a good interaction which is energetically favorable as shown in Figure 5.2B. The atoms involved in the interaction are given in Table 5.4 which shows better interaction as compared to template peptide.

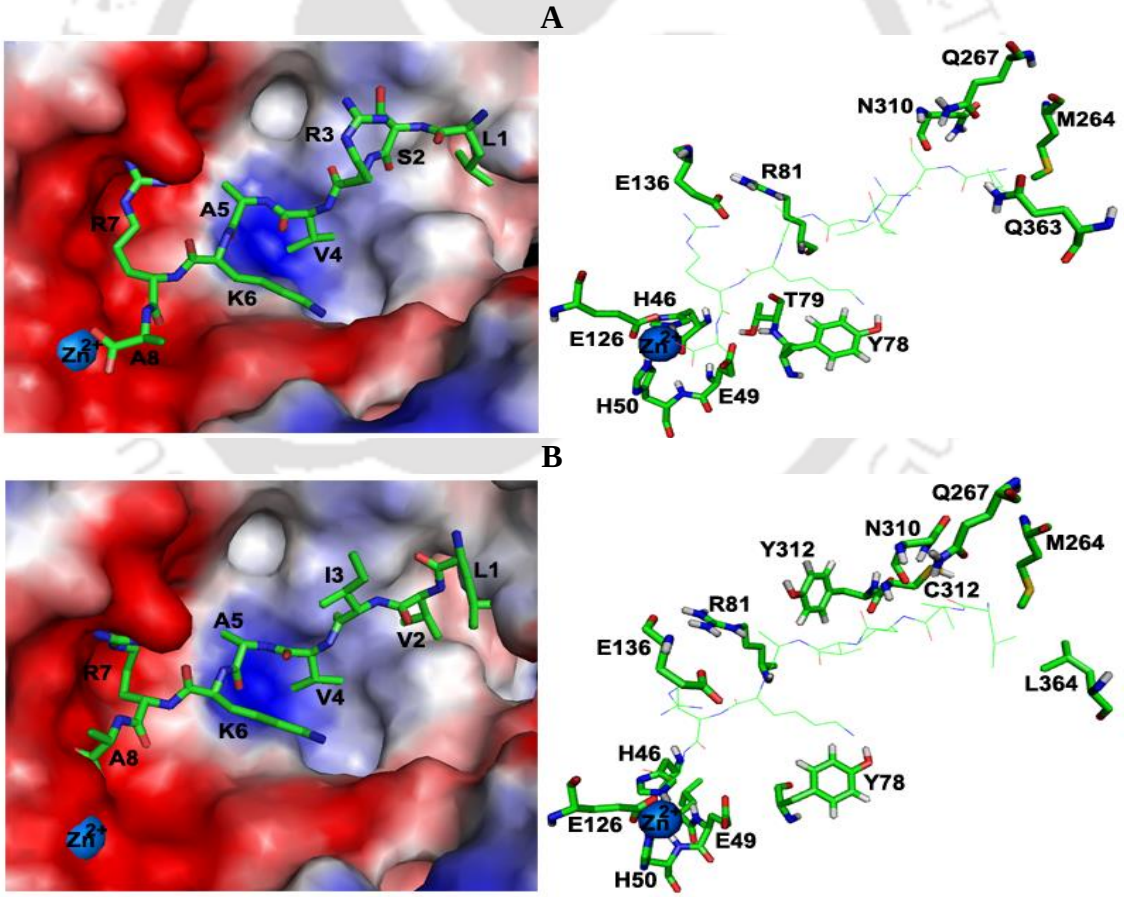


Figure 5.2 Comparative analysis of PFI1625c-peptide molecular model. (A) PFI1625c-template peptide model in electrostatic representation and stick form **(B)** PFI1625c-P550 peptide model in electrostatic representation and stick form. P550 shows better fitting into the active site exploiting opposite charge pockets and hydrophobic patches. Red represents −ve charge, blue represents +ve charge.

Table 5.4 Interaction of the peptide within the active site of PFI1625c

Template peptide	Protein Residue	Atom	Distance (Å)	Peptide P550	Protein residue	Atom	Distance (Å)
L1	M264	CB	2.04	L1	Q267	OE1	3.18
L1	M264	CD2	2.46	L1	L364	CD2	2.61
S2	N310	O	2.9	L1	L364	CD1	2.84
S2	N310	H	1.82	L1	M264	CE	2.01
S2	Q267	NE2	2.98	V2	C312	CB	2.25
R3	Q363	NE2	4.88	V2	C312	N	2.55
V4	R81	N	4.27	I3	C312	N	3.07
A5	R81	NE	4.34	I3	N310	O	2.42
K6	Y78	OH	2.34	V4	C312	O	2.93
R7	E136	OE2	1.64	V4	Y313	CE1	3.34
A8	E49	OE2	2.34	A5	R81	NE	3.06
A8	H46	HD1	2.50	K6	Y78	CE1	3.71
				R7	E136	OE2	1.29
				A8	E126	OE2	2.15
				A8	H46	NE2	2.19

To further validate the approach, docking was also performed by taking protease with multiple peptide substrates or multiple proteases with single peptide substrate. Falcilysin and its multiple substrates were obtained from literature [151]. The docking experiments showed a strong correlation (r^2 = 0.941) between the affinity of the peptide substrate (Km) to the atomic contact energy (Figure 5.3). Correlation of experimentally determined biochemical data with docking score proves the reliability of Patchdock as docking tool to generate PFI1625c-peptide molecular models and correlates the docking score with suitability of the bound peptide as potential substrate.

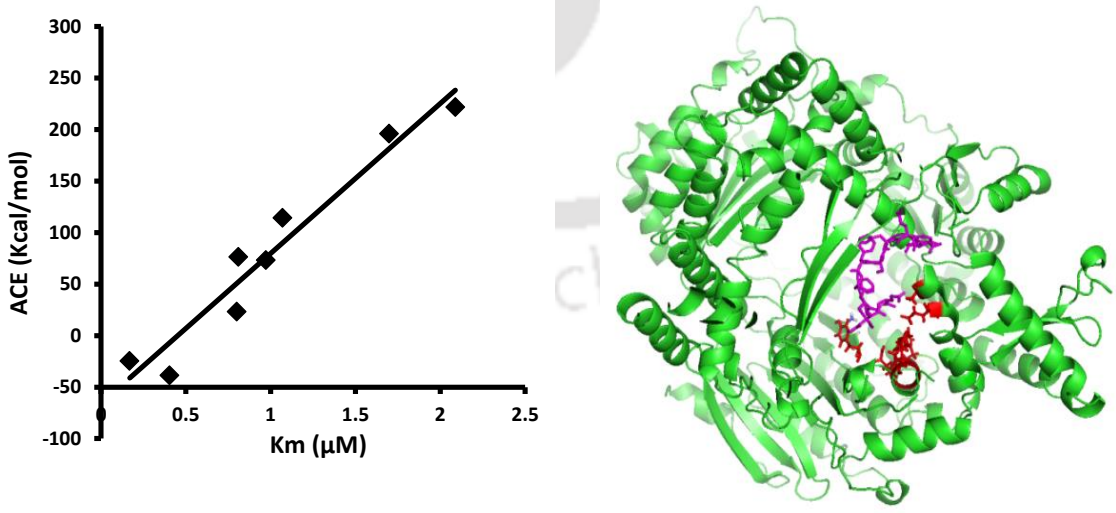


Figure 5.3 Validation of Patchdock as docking tool for docking PFI1625c with peptides. A well-known metalloprotease falcilysin (green cartoon form) was docked with multiple peptide substrates. **(A)**The Km values of falcilysin activity were correlated with the ACE values obtained from Patchdock. A strong correlation with r^2=0.941 was obtained. **(B)** The binding mode of peptide (pink stick form) with the active site of falcilysin (red stick form).

5.3.1.2 Anti-malarial bioactive peptides fit well into the PFI1625c active site

The active site of PFI1625c is well organized forming a large cavity. The left region of the outer part is lined by hydrophilic residues where the catalytic zinc binding residues are present. The rest part of the active site cavity is neutral or hydrophobic where substrate binding scaffolds are present. Most of the high affinity binding peptides have their hydrophobic residues lying in the hydrophobic core of the enzyme as in the case of P550 [153]. Out of 20 peptides selected based on their antimalarial properties, 9 peptides bind to catalytic site and fit well into the active site cavity namely Scorpine, Dermaseptin S3, PcFK2, Angiotensin II, Cyclosporin A, Efrapeptin, Defensin, Thiostrepton and Gramicidin D (Table 5.5). 4 peptides such as Magainin, Cecropin, PcFK1 and Defensin-like peptide bind only to catalytic site but away from active site cavity (Table 5.5). On the other hand, 2 peptides namely Dermaseptin S4 and Dermaseptin derivative bind into the active site cavity but not close to the catalytic site. There are 5 peptides which do not bind at all to the active site cavity and catalytic site which comprises Lactoferrin, Meucin 24, Meucin 25, LVDAA and Melittin-Cecropin hybrid (Table 5.5). The antimalarial peptides namely, Dermaseptin S3, Angiotensin II, Efrapeptin, PcFK2 and Cyclosporin A, Thiostrepton and Gramicidin D showed electrostatically favorable high binding affinity despite their large sized peptides (Figure 5.4A-D). These peptides interact with the active site residues of PFI1625c. The hydrophilic region of PFI1625c or the zinc binding site is occupied by hydrophilic groups of the bioactive peptides giving energetically favorable binding. The large sized peptides with high binding affinity fitted the active site pocket of the putative enzyme (PFI1625c) resembling the lock and key interaction of enzyme and substrate. On the other hand, peptides with low affinity such as Scorpine, Magainin and Cecropin bind to the catalytic region but not fitted completely in the active site cavity (Table 5.5). Their charge distribution is also incompatible for interaction as in the case of Cecropin which is a cationic peptide but distributed over neutral region of PFI1625c. Some peptides such as Efrapeptin, Gramicidin D and Thiostrepton bind and fit well in the active site cavity with low ACE. However their corresponding IC_{50} values do not correlate each other thus having no importance to exploit PFI1625c to exhibit their antimalarial activities (Table 5.5).

Table 5.5 Interaction of known antimalarial bioactive peptides with PFI1625c

Sl. No.	Antimalarial peptide	ACE (kCal /mol)	Antimalarial IC50 (µM)	Binding site in PFI1625c
1	Dermaseptin S4	22.46	0.27-2.2	Only at active site but not at catalytic site
2	Scorpine	109	10	Active site and Catalytic site
3	Dermaseptin S3	-275.62	34	Active site and catalytic site
4	Magainin	172.55	0.39	Neither at active nor catalytic site
5	Cecropin	149.41	0.23	Only at catalytic site
6	PcFK1	-1.95	1.59	Only at catalytic site
7	PcFK2	-42.12	1.15	Active site and catalytic site
8	Dermaseptin derivative	-70.88	3.4-7.7	Only at active site but not at catalytic site
9	Cyclosporin A	-215.33	16	Active site and catalytic site
10	Lactoferrin	-170.99	3.2	Neither at active nor catalytic site
11	Efrapeptin	-192.4	1.37	Active site and catalytic site
12	Angiotensin II	-234.33	34	Active site and catalytic site
13	Meucin 24	271.43	20	Neither at active nor catalytic site
14	Meucin 25	-210	20	Neither at active nor catalytic site
15	LVDAA	-210.81	29	Neither at active nor catalytic site
16	Melittin+cecropin A	NA	10	Neither at active nor catalytic site
17	Defensin-like	-89.62	5	Only at catalytic site
18	Defensin	-32.69	1	Active site and catalytic site
19	Gramicidin D	-250.09	1.854^{-05}	Active site and catalytic site
20	Thiostrepton	-370.49	8.9	Active site and catalytic site

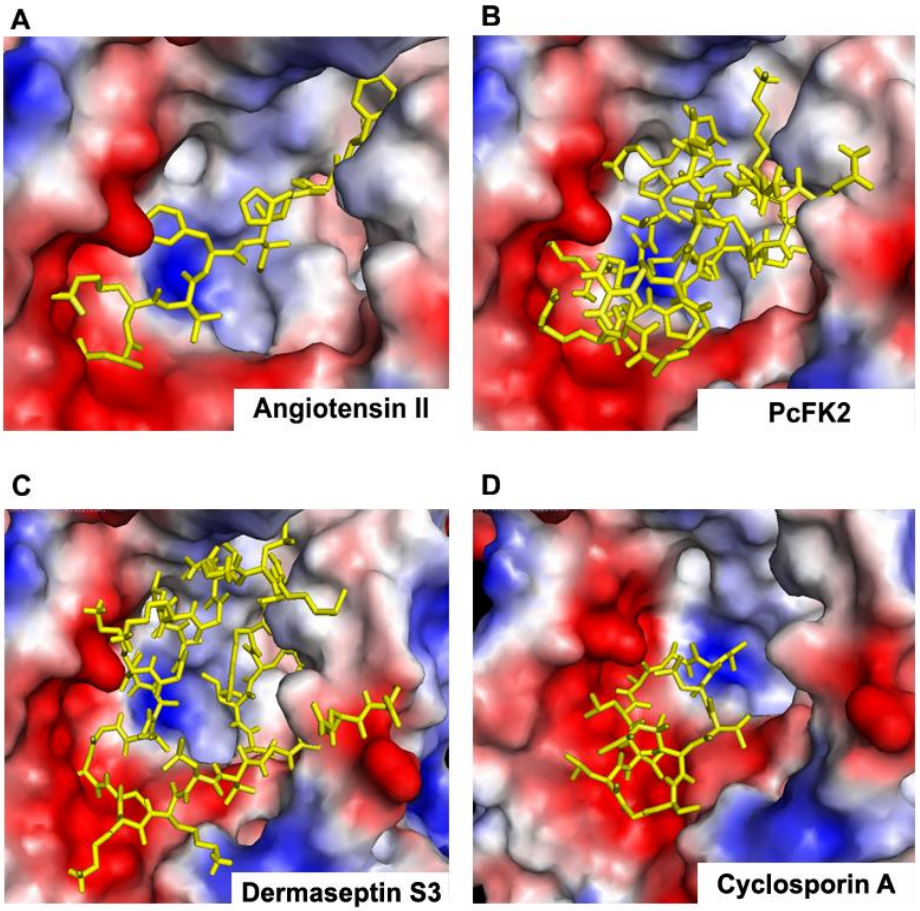


Figure 5.4 Interaction of PFI1625c with selected bioactive peptides which are shown to have correlation between their anti-malarial IC_{50} and ACE values.

5.3.1.3 Antimalarial Bioactive peptides might exploit PFI1625c as a target.

The antimalarial peptides inhibit the growth of parasites in different modes. Most of the inhibition is studied up to the cellular level like gametocytocidal activity, ookinette inhibition and infected RBC lysis. The exact mechanism of these peptides is still not fully investigated up to the specific target or signaling pathway. Knowing the exact mechanism or pathway will provide a way for developing better and specific inhibitor for a particular target. To verify whether PFI1625c is behind the antimalarial activities of these peptides, we performed a correlation curve between the ACE values obtained from PFI1625c-peptide interactions and the antimalarial IC_{50} of peptides obtained from literature (section 5.2.3). The complete set of peptides show a correlation factor R = 0.52 (Figure 5.5A). However, a subset of selected peptides having high affinity with PFI1625c and good antimalarial activity show a correlation with R=0.921 (Figure 5.5B). The selected peptides of subset 2 include Cecropin, Dermaseptin S3, Dermaseptin S4, PcFK1, PcFK2, Dermaseptin derivative, Defensin-like peptide, Cyclosporin A and Meucin 25 (Figure 5.5B). A strong correlation of selected antimalarial bioactive peptides indirectly consolidates the idea that PFI1625c might be responsible for the antimalarial action of these peptides.

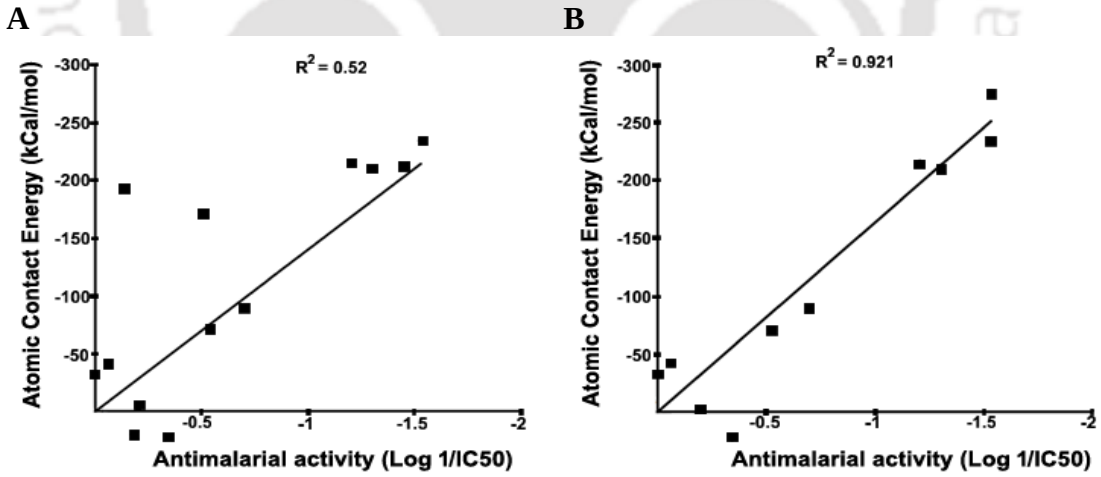


Figure 5.5 Correlation analysis of antimalarial IC_{50} values and ACE value of bioactive peptides with PFI1625c. (A) Complete set of bioactive peptides showing a correlation R^2 value of 0.52. (B) Subset of selected bioactive peptides which falls in the correlation curve showing R^2 value of 0.921.

5.3.1.4 PFI1625c binds well with synthetic curcumin analogs

Docking of PFI1625c with 23 curcumin analogs showed good interactions (Table 5.6). Based on binding energy values, 6 compounds viz., JA-304, JA-305, JA-403, JA-408, JA-501 and JA-506 showed significantly better binding energy than other compounds.

Interaction analysis showed that these 6 compounds interacted with the active site residues of PFI1625c and exhibit strong polar contacts with active site residues. JA-305 formed polar contacts with the catalytic residue Glu49and substrate binding residue Thr79 (Figure 5.6A). JA-408 formed polar contact with zinc binding residue Glu126 and catalytic residue Glu49 (Figure 5.6B). JA-506 formed polar contact with zinc binding residue His50 and substrate binding residue A77 (Figure 5.6C).

Table 5.6: Docking score of heterocyclic compounds with PFI1625c

Ligand code	Binding energy	Ligand code	Binding energy
JA-301	-5.46	JA-502	-6.05
JA-303	-5.93	JA-503	-6.17
JA-304	-6.03	JA-504	-6.01
JA-305	-6.69	JA-505	-6.04
JA-309	-6.16	JA-506	-6.7
JA-402	-5.51	JA-507	-6.05
JA-403	-6.78	JA-508	-5.99
JA-404	-5.65	JA-509	-5.96
JA-406	-6.27	JA-510	-5.86
JA407	-5.3	JA-511	-5.93
JA-408	-6.83	JA-512	-5.69
JA-501	-6.28		

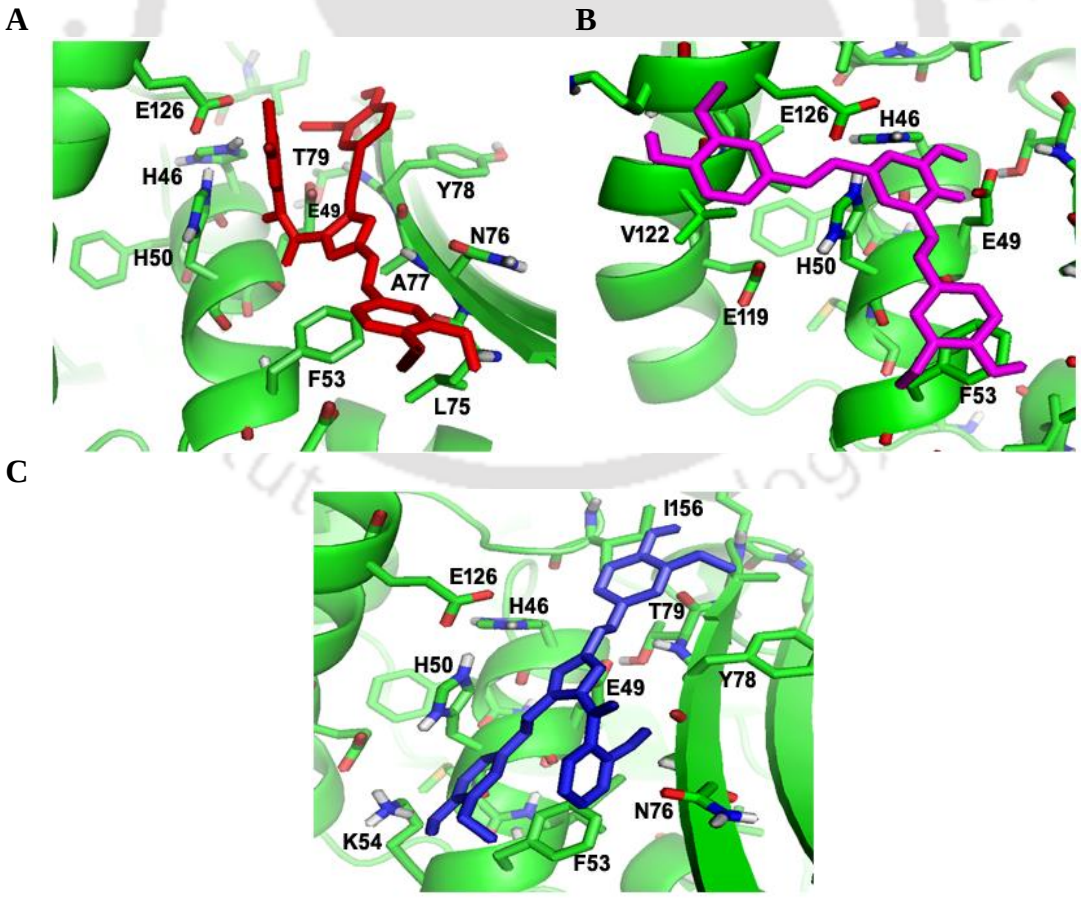


Figure 5.6 Interaction of top hits curcumin analogs with PFI1625c active site. Important active site residues are labeled. **(A)** JA-305 in red, **(B)** JA-408 in magenta and **(C)** JA-506 in blue stick form.

5.3.2 Effect of selected inhibitors on PFI1625c activity

5.3.2.1 Synthetic peptides exhibited inhibitory effect on PFI1625c activity

Peptide P536, P548 and P550 were synthesized by solid phase peptide synthesis. The methodology as well as validation of peptide synthesis is given in the Appendix (section 5.5). There was a decrease in activity of PFI1625c observed in all the three peptides (Figure 5.7). P550 showed highest inhibition and IC_{50} of inhibition was found to be 1.5µM. P536 inhibited PFI1625c activity with an IC_{50} of 4.2µM. P548 showed least inhibition on PFI1625c activity with maximal inhibition of 30%. The stronger inhibition in P550 and P536 is in agreement with the molecular modeling studies of PFI1625c with hydrophobic substituted peptides (Figure 5.2 and Table 5.4) whereas P548 is a mixed substituted peptide (negative + non-polar). The mode of interaction for P550 and P536 is closer to the catalytic site compared to P548 which binds to the active site but not on the catalytic site (Figure 5.1).

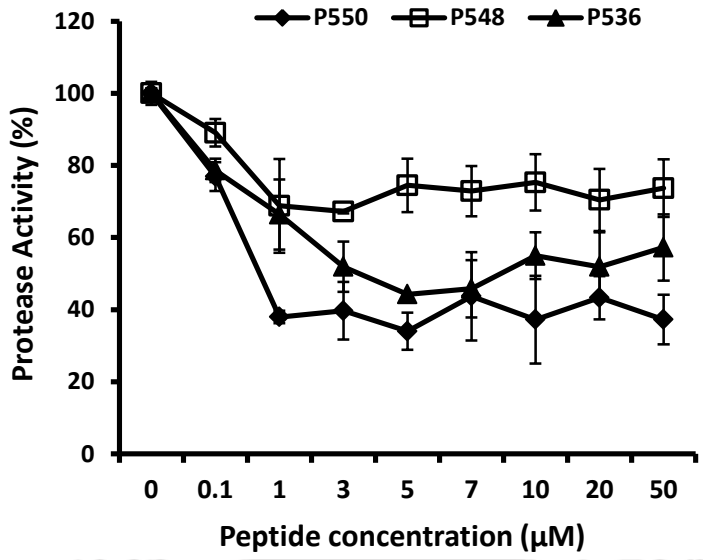


Figure 5.7 Effect of synthetic peptides selected from Combinatorial Peptide library on PFI1625c protease activity. The three peptides show inhibitory effect on PFI1625c activity. P550 shows highest inhibition, followed by P536 and least inhibition is observed in P548.

5.3.2.2 Antimalarial bioactive Cyclosporin A irreversibly inhibits PFI1625c activity

Due to limited resources in bioactive peptides procurement and synthesis, we tested commercially available cyclosporin A. on PFI1625c protease activity. Cyclosporin A is one of the best interacting bioactive peptide with PFI1625c with ACE value of -215 and an antimalarial IC_{50} of 16µM. It has 65% sequence similarity with optimized peptide P550. Cyclosporin A is found to inhibit PFI1625c activity completely. The IC_{50} value as calculated from the graph was ~2.6µM (Figure 5.8). The low IC_{50} value of enzyme

inhibition indicates that cyclosporin A effectively binds and inhibits PFI1625c activity. The type of inhibition whether it is reversible or irreversible, was performed by passing the enzyme-inhibitor complex (10µM of inhibitor) through NAP-10 desalting column. The eluted enzyme was used for activity assay and it was found that the enzyme inhibition was irreversible as there was no activity regained by the enzyme (Figure 5.8B). The enzyme inhibition and associated cell death indicates that PFI1625c is involved in the mechanism.

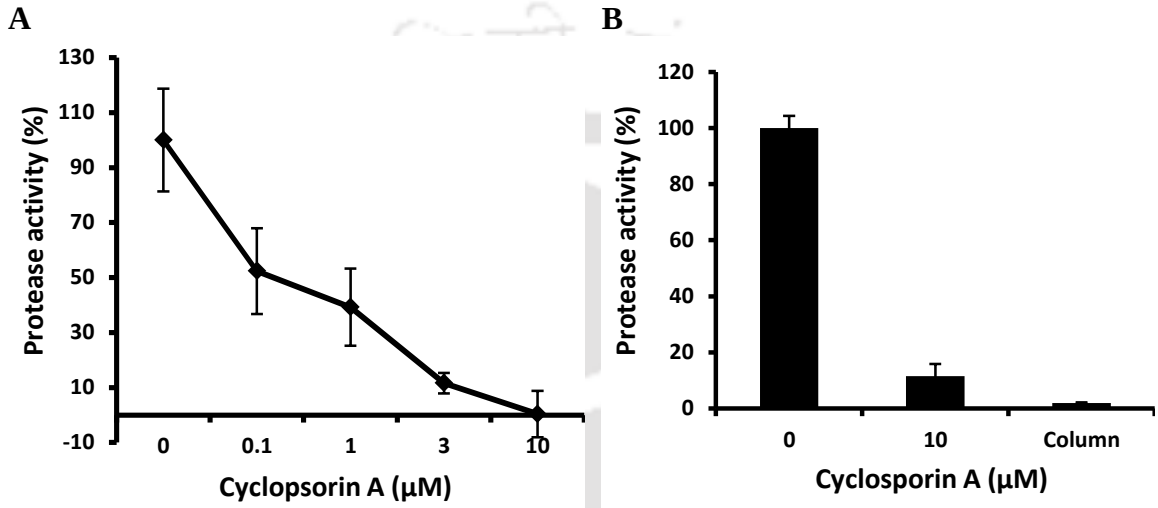


Figure 5.8 Irreversible inhibition of Cyclosporin A on PFI1625c protease activity. **(A)** Cyclosporin A completely inhibits PFI1625c activity with an IC_{50} value of 2.7µM. **(B)** Cyclosporin A irreversibly inhibits PFI1625c activity as evident from the activity assay after column elution of enzyme inhibitor complex (column) from NAP-10 desalting column.

5.3.2.3 Synthetic curcumin analogs have inhibitory effect on PFI1625c activity

Synthetic curcumin analogs were screened for inhibitory effect against PFI1625c activity. Firstly, the interaction study was performed between the curcumin analogs and PFI1625c using *in silico* models. To validate the *in silico* data, biochemical experiment of proteases assay was performed for PFI1625c against the five top hits compounds (JA-304, JA-305, JA-403, JA-408, JA-501 and JA-506). Experimental result showed that inhibition was observed in JA-305, JA-408 and JA-506 while no inhibition was observed in JA-304, JA-403 and JA-501 (Figure 5.9). The 50% inhibition concentration or IC_{50} was calculated for the three inhibitory compounds from the graph and it was found that, IC_{50} values for JA-305, JA-408 and JA-506 were 4.35µM, 198µM and 6.6µM respectively. Among them JA-305 and JA-506 showed effective inhibition. Biochemical inhibition of PFI1625c by three compounds in correlation with docking data suggests that selected curcumin analogs target PFI1625c for their inhibitory actions.

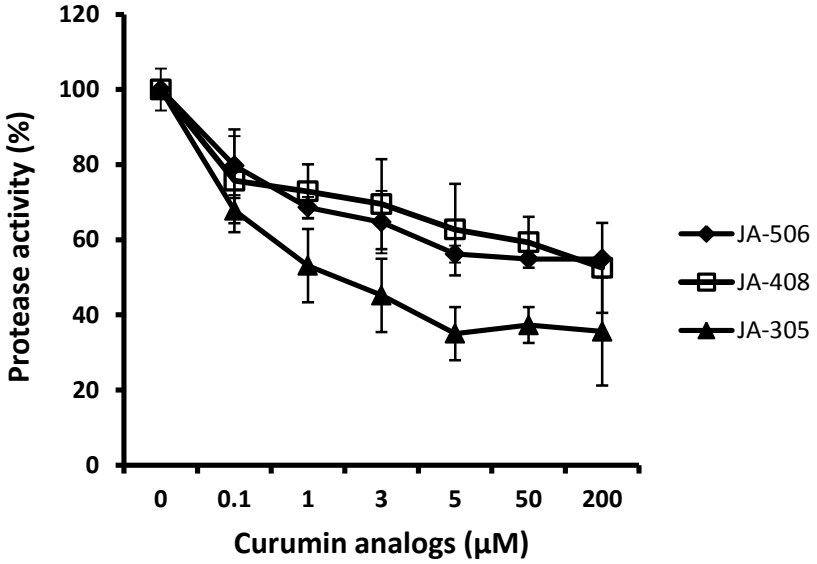


Figure 5.9 Inhibitory effects of synthetic curcumin analogs on PFI1625c protease activity. Three compounds show inhibitory effect. Compound JA-305 has highest inhibitory effect followed by JA-506 and lowest in JA-408.

5.3.2.4 Synthetic curcumin analogs exhibits irreversible inhibition against PFI1625c activity

Synthetic curcumin analogs JA-305, JA-408 and JA-506 are shown to have inhibitory effect on PFI1625c activity. To test the nature of inhibition (reversible or irreversible) of these inhibitors on PFI1625c activity, the enzyme-inhibitor reaction mixture was passed through NAP-10 desalting column and activity assay was performed from eluted PFI1625c fraction. From the graph result of all three inhibitors (JA-305, JA-408 and JA-506), it was observed that activity of PFI1625c was inhibited in both column eluted PFI1625c as well as PFI1625c in presence of inhibitors (Figure 5.10). This indicates that PFI1625c is irreversibly inactivated by JA-305, JA-408 and JA-506.

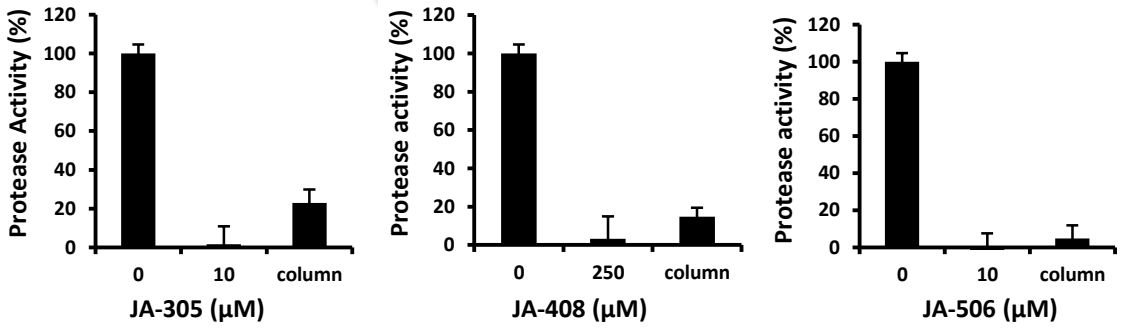


Figure 5.10 Irreversible inhibition of synthetic curcumin analogs against PFI1625 activity. Activity is not recovered after passing the enzyme-inhibitor complex through NAP-10 desalting column as compared to enzyme activity without inhibitor.

5.4 Discussion

Characterizing and validating proteins to examine their potential as a drug target is the need of the hour to develop novel malarial chemotherapy. Protein-peptide docking is a promising ultimate tool to decipher substrate or ligand specificity. Computational approaches to determine specificity in a peptide substrate toward HIV-1 protease indicate a correlation between substrate specificity and its binding energy with the protease [42]. In another study, protein-protein complexes are used to identify the determinants that are crucial for complex formation [43]. Patchdock was used as a tool to generate PFI1625c-peptide molecular models. Control docking of signal peptide (LSRVAKRA) in yeast mitochondrial processing peptidase (PDB Code 1HR9) gives complex (yeast MPP-peptide) closer to known structure [25]. Correlation analysis of PFI1625c-peptide complexes indicates a preference of PFI1625c toward hydrophobic residues in the substrate due to the local environment. All peptides fitting well into the PFI1625c active site and stereochemistry indicate that most likely cleavage site is a peptide bond between A8 and positively charged R7, as positive charge can be stabilized by negatively charged patch within the catalytic site. Substitution of amino acids in the peptide occupying the hydrophobic cavity with non-polar residues increases the affinity of bound peptides within the PFI1625c active site due to the presence of a hydrophobic pocket.

Protein-peptide docking is a powerful tool to predict the binding affinities, specificity and such a study has immense potential in drug discovery by utilizing bioactive peptides [47,48]. A number of bioactive peptides against malaria are under study but the targets for these peptides are not known. Several bioactive peptides have been studied for their inhibition towards PFI1625c. Some of the effective peptides bind with high affinity as in the case of Dermaseptin s3. This peptide is shown to have hemolytic activity and cytotoxic effect on *P.falciparum* which binds to RBC membrane and kills the parasite [234]. Many of the antimalarial peptides including Cyclosporin A induce hemolysis or membrane permeability (Angiotensin) resulting killing of parasites. However the exact mechanism or the target proteins of the antimalarial peptides under study are still not known. Metalloproteases are involved in different biological activity imparting to parasitic growth as in the invasion step, organelle processing activity, hemoglobin digestion etc during RBC stages of malaria. They also process different proteins to reach their target organelles like apicoplast and mitochondria. All these functions and antimalarial peptides activities implicates PFI1625c to fulfill the criteria to be one of the targets inhibited by antimalarial activities. Structural and functional studies further

reveals that it is a metalloprotease similar to M16 family involved in hemoglobin digestion. The binding mechanism of PFI1625c with antimalarial peptide Dermaseptin S3 also resembles the mechanism of action proposed in the paper published earlier where alanine is at the position where cleavage occurs [234].

Supportive to the above implications and evidences, there has been correlation observed between the IC_{50} values and binding affinity of some of the antimalarial peptides against PFI1625c which are mentioned earlier. It is therefore plausible that PFI1625c may be the target for these antimalarial peptides to eventually exhibit killing of the parasite. It is evident that PFI1625c is involved during the invasion steps as Cyclosporin A (known to inhibit invasion) effectively and irreversibly inhibits PFI1625c resulting in disruption of merozoites invasion to new RBCs.

Out of 23 synthetic curcumin analogs used for interaction study with PFI1625c, 6 compounds (JA-304, JA-305, JA-403, JA-408, JA-501 and JA-506) have been shown to have strong interaction. Among them, 3 compounds (JA-305, JA-408 and JA-506) have irreversible inhibitory effect on PFI1625c activity. It is evident from the enzymatic inhibition assay that enzyme did not regain its original activity even after passing the enzyme-inhibition complex through NAP-10 desalting column. The selected 6 compounds were also tried for anti-parasitic and schizonticidal effect (data not shown). Interestingly, 3 compounds which were to inhibit PFI1625c activity also showed anti-parasitic and schizonticidal effect whereas the same 3 compounds which did not inhibit PFI1625c activity did not have any effect on parasite growth (data not shown). It is strongly justifiable that anti-parasitic and schizonticidal effect are related to PFI1625c inhibition. This selective response to 3 curcumin analogs (JA-305, JA-408 and JA-506) from *in silico* study to parasiticidal effect strongly supports the involvement of PFI1625c in the curcurmin analogs induced parasite death. The schizonticidal effect of PFI1625c-inhibiting compounds showed that PFI1625c is active during schizont development thus indicating its role in schizont stage such as egress of merozoites and re-invasion.

References: Please refer to bibliography section from page 108-119.

5.5 Appendix

Solid Phase synthesis of top hits peptides from peptide library: The top hits peptides which interact best with PFI1625c were selected for chemical synthesis as described [253]. The synthesized peptides include P536 (LIIVAKRA) and P550

(LVIVAKRA) and P548 (LVDVAKRA). Required amino acids with Fmoc-protecting group and all reagents (synthesis grade) were purchased from Merck Millipore NovaBiochem. 4-Hydroxymethylphenoxyacetyl (HMPA) resin (NovaBiochem) was used according to the desired peptide product. The resin was soaked in dimethylformamide (DMF) overnight. 10x molar excess of Fmoc-amino acid was dissolved in minimum volume of dried dichloromethane (DCM). Then 10x molar excess of dicyclohexylcarbodiimide (DCC) was then added to solution and kept for stirring for 30min till white precipitate was obtained. The DCM was then evaporated completely using nitrogen gas. The dried anhydride was then re-dissolved in minimum volume of DMF and added to the resin (where Fmoc was already removed with 20% piperidine in DMF). 0.1 equivalents of 4-(N,N-dimethylamino) pyridine (DMAP) was added to the suspension to obtain viscous suspension and agitated for 3hrs. The suspension was transferred to synthesis vial and washed sequentially with DMF, isoamylalcohol, glacial acetic acid, isoamylalcohol and diethyl ether for storage. The resin was then soaked in DMF for 30mins followed by Fmoc-removal of first amino acid using 20% piperidine. The piperidine was then washed completely with DMF until neutral pH was obtained. The next coupling mixture consisted of 3x, 6x, 3x and 3x molar excess of amino acid, diisopropylethylamine (DIPEA), 1-hydroxybenzotriazole (HOBt) and O-(Benzotriazol-1-yl)-N,N,N’,N’-tetramethyluroniumhexafluorophosphate (HBTU) respectively which was added to resin and degassed. The reaction was allowed under agitation for 1hr followed by 30mins in the second attachment. Then Fmoc was removed and cycle continued. After completion of coupling and removal of Fmoc, deprotection was performed using deprotection mixture containing m-cresol, thioanisol, ethanedithiol and trifluoroacetic acid (TFA) and kept overnight. Precipitation of the peptide was performed by filtering the mixture in glass filter into ice cold diethyl ether. The solution was washed 3-4 times with diethyl ether for complete removal of deprotection reagents. The diethyl ether was then dried off completely and the peptide was ready to be used. The purity of the synthesized peptides was analyzed using Reverse Phase HPLC on C18 reverse phase column (Prominence, Shimadzu), using a linear elution gradient from 10 to 100% of buffer A (0.1% TFA in 10% acetonitrile) and buffer B (0.1% TFA in 100% acetonitrile) and run for 60mins at a flow rate of 0.5ml/min. The mass of the peptides was assessed by electrospray analysis using a system mass spectrometer ESI-TOF (Waters, Q-Tof Premier). The top hits peptides viz, P536 (LIIVAKRA), P550 (LVIVAKRA) and P548 (LVDVAKRA) were synthesized chemically by solid phase chemical synthesis. Using

HPLC, the purity of the peptides was checked and single peaks were observed in all the three peptides. The percent purity was estimated by peak integration (peak with absorption at 210 nm). The purity was calculated and was found to be 93.3% for P536, 86.3% for P550 and 88% for P548 (Figure 5.11 HPLC). The m/z ratio of mass spectroscopy showed a correct molecular weight for the three peptides. A corresponding peak for P536 (MW=884 dalton) in the form of $[M+2H]^{++}$ ionization state giving a 442.34 dalton peak (Figure 5.11A). A corresponding peak for P550 (MW=869.11 dalton) is present in the form of $[M+H+Na]^{++}$ giving a peak of 445.36 (Figure 5.11B). A corresponding peak for P548 (MW=871.04 dalton) is present in the form of $[M+3H]^{+++}$ giving a peak of 293.23 (Figure 5.11C).

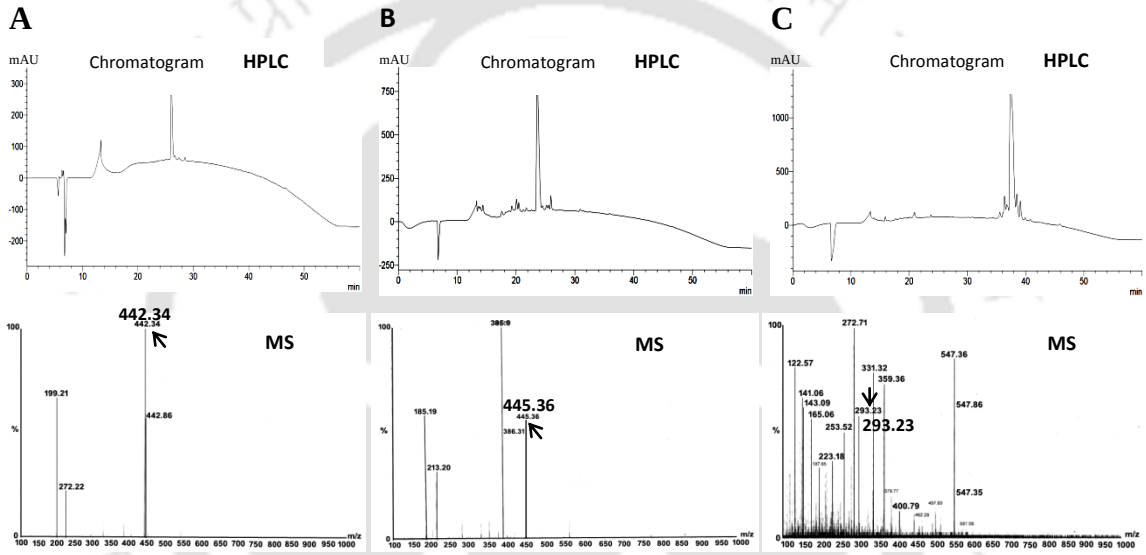


Figure 5.11 Chromatograms of Reverse phase HPLC and ESI-Mass Spectrometry (MS) of synthesized peptides. HPLC of all samples show prominent single peak with negligible impurities. (A) P536 peptide (MW=884dalton) in $[M+2H]^{++}$state showing a peak at 442.34 dalton. (B) P550 peptide (MW=869.11) in $[M+H+Na]^{++}$ state showing a peak at 445.36 dalton. (C) P548 peptide (MW=871.04 dalton) in $[M+3H]^{+++}$ state showing a peak at 293.23 dalton.

Chapter-6 Summary, Conclusion and Future Prospects of the Work

Summary, Conclusion and Future Prospects of the Work

Summary

1. PFI1625c is a protease present in *P.falciparum* genome. Amplification of PFI1625c was performed from *P.falciparum* genomic DNA using site specific primers with BamHI and XhoI site. The amplified product was cloned in pTZ57R/T cloning vector followed by sub cloning in expression vectors pET22b and pET23a. The gene was over-expressed in *E.coli* expressing system by transforming the cells with pET22b containing the gene insert. Optimization of expressed PFI1625c was performed by standardizing the bacterial strains, IPTG concentration, time and temperature. The optimum expression was observed when BL21(DE3) was used at 0.8mM IPTG concentration, 25°C (or lower up to 18°C) and 11hrs induction (or more up to 24hrs). The over-expression was confirmed by western blotting using anti-his antibody. Solubility test of over-expressed PFI1625c was performed using lysis buffer of different pH, different bacterial strains, varying inducer concentrations, time, temperature etc but it was present in inclusion bodies which was confirmed by western blot. PFI1625c was processed under denaturing. The denatured protein was refolded by dialysis. Purification of PFI1625c was performed using Ni-IDA affinity chromatography and purified PFI1625c was obtained as a single band. Oligomeric status of PFI1625c was determined using gel filtration column and is a monomeric protein which was confirmed by comparing the molecular weight of denatured PFI1625c (in SDS PAGE) and gel filtration eluted PFI1625c. The yield of PFI1625c was ~1mg/L.

2. Antibody generation for PFI1625c was performed by electro-elution of PFI1625c from SDS PAGE. The emulsion containing pure PFI1625c along with Freund's complete adjuvant was injected into rabbit. The next booster dose was given by injecting emulsion containing PFI1625c with Freund's incomplete adjuvant. The antibody titer was estimated by ELISA using PFI1625c samples and blood was collected from the rabbit. Serum containing the anti-PFI1625c antibody was collected from the blood. Localization of PFI1625c in *P.falciparum* was performed using generated anti-PFI1625c antibody and FITC-labeled secondary antibody. DAPI was used to stain the nucleus. Preliminary Localization of PFI1625c shows the presence of PFI1625c outside the nucleus (may be cytosol) and small vesicles appeared in the later stages. PFI1625c was not observed in nucleus and food vacuole.

3. Structural characterization of PFI1625c was performed by modeling of PFI1625c using yeast mitochondrial processing peptidase (1HR6) as template. Sequence similarity was performed by blasting PFI1625c amino acid sequence with PDB database. Top hit blast result showed similarity of PFI1625c with metalloproteases of yeast, bacillus and bc1 core protein of bovine and chicken. Multiple sequence alignment showed the presence of conserved zinc binding motif HXXEH in yeast, bacillus and PFI1625c. Surface representation of PFI1625c showed an organized active site cavity with hydrophilic exterior and hydrophobic interior. Structural study showed the presence of zinc binding motif at the catalytic site of active site cavity. Substrate binding scaffold similar to other metalloprotease was also present in PFI1625c. Hydropathy plot showed that PFI1625c is not a transmembrane protein.

4. Biochemical characterization of PFI1625c indicates that PFI1625c is a protease. Testing of different substrate showed the preference of PFI1625c towards gelatin and hemoglobin with Km value of 30µM and 48.4µM respectively. Effect of pH, temperature and time on PFI1625c activity showed the optimal activity of PFI1625c at pH 8.5, 37°C and 2.5hrs incubation period. Testing transition metals (Zn^{2+}, Mn^{2+}, Mg^{2+}, Cu^{2+}, Ca^{2+}, Fe^{2+}, Hg^{2+}) effect on PFI1625c activity showed a dose response curve by zinc. PFI1625c activity was enhanced by 80% in presence of 70µM zinc which confirms PFI1625c as a metalloprotease. Detergents such as SDS, Tween-20 and CTAB significantly enhanced PFI1625c activity by 40%, 49% and 18% respectively. Matrix metalloprotease nature of PFI1625c was tested by gelatin zymography. Zone of clearance at PFI1625c site indicated the probability of PFI1625c as a metal dependent protease. Detection of PFI1625c in parasite supernatant was performed by dot blot using anti-PFI1625c antibody. Signals observed in ring, trophozoite and schizont stages showed that PFI1625c was secretory and secretion was more in trophozoite and schizont stage.

5. PFI1625c inhibitors were screened using different approach. Combinatorial peptide library was prepared by different substitutions using template peptide (LSRVAKRA). Docking of PFI1625c with peptides from combinatorial peptide library showed the preference of PFI1625c towards hydrophobic substituted peptides. Selected peptides (P536, P548 and P550) were synthesized chemically and were checked for effects on PFI1625c activity. P536 and P550 have significant inhibitory effects on PFI1625c with IC_{50} value of 4.2µM and 1.5µM respectively. Selective antimalarial bioactive peptides bind strongly to PFI1625c such as Angiotensin II, Cyclosporin A,

Dermaseptin S3 and PcFK2 with a correlation value of 0.92 between ACE and antimalarial IC_{50} values. Cyclosporin A has completely inhibited PFI1625c activity with an IC_{50} value of ~2.6µM. Curcumin analogs were tested for inhibitory effect against PFI1625c. Interaction analysis performed by docking showed strong interaction of PFI1625c with 6 curcumin analogs (JA-304, JA-305, JA-403, JA-408, JA-501 and JA-506). Out of 6 compounds, 3 compounds (JA-305, JA-408 and JA-506) inhibited PFI1625c activity with an IC50 value of 4.35µM, 198µM and 6.6µM respectively.

Conclusion of the work

From the above summary, two relevant conclusions are made;

1. Probable role of PFI1625c in parasite physiology and host pathology:

PFI1625c is a metalloprotease present in *P.falciparum* genome as revealed by enhanced activity in presence of zinc ion and inhibition of protease activity by metal chelator 1,10-phenanthroline. Presence of PFI1625c in cytosol and its substrate preference for hemoglobin reveals that it takes part in hemoglobin dipeptide degradation to release free amino acids in the cytosol. Presence of PFI1625c towards membrane periphery during late stage (schizont) as transported by small vesicles reveals its probable role(s) in egress and invasion. Optimal activity of PFI1625c at pH 8.5 shows the involvement in apicoplast and mitochondria as proteases localized to these organelles have optimal activity in the alkaline pH [222]. Moreover docking of PFI1625c with organelle peptides shows high affinity with apicoplast targeting signal peptides (data not shown). Activity of PFI1625c from 37°C up to 40°C indicates its active role during high fever of malaria.

2. Potentials of PFI1625c in drug development

Cyclosporin A is an antimalarial bioactive peptide which blocks invasion of RBC by merozoites. Effective inhibition of cyclosporin A against PFI1625c activity gives indirect evidence for involvement of PFI1625c during invasion which is clear by the fact that PFI1625c is exported towards membrane periphery during late stage. PFI1625c involvement in invasion is supported by the fact that it strongly interacts with signal peptides of rhoptry proteins (data not shown). PFI1625c inhibition by curcumin analogs is associated with parasiticidal and schizonticidal effect (data not shown). These effects are induced by increase in oxidative stress such as lipid peroxidation and protein carbonyl level. Dysregulation of hemozoin formation is also induced by PFI1625c-specific curcumin analogs (data not shown). PFI1625c inhibition induces parasitic death, increase hemozoin dysregulation and disruption of invasion can be a novel approach for targeting

PFI1625c for anti-malarial drug development. In addition, PFI1625c preference towards hydrophobic peptides support the involvement of PFI1625c in apicoplast protein processing as these type of proteins are prevalent in apicoplast.

Future prospects

1. To study the inhibitory effect on PFI1625c activity by antimalarial bioactive peptides such as Dermaseptin S3, Angiotensin II and PcFK2.
2. To study the effect on hemoglobin digestion, merozoite egress and invasion by blocking PFI1625c in vivo using specific inhibitors
3. To study the down regulation or up regulation of cytokines or other factors leading to cerebral malaria by blocking PFI1625c using specific inhibitors.
4. To detect the presence of PFI1625c in body fluid for development of malaria detection kit.

References: Please refer to bibliography section from page number 108-119.

Bibliography

1. World Malaria Report, *World Health Organization.* 2010.
2. Organization, W.H., 2012.
3. World Malaria Report, *World Health Organization*. 2013.
4. Pamplona A, et al., *Cerebral malaria and the hemolysis/methemoglobin/heme hypothesis: Shedding new light on an old disease.* Int J Biochem Cell Biol., 2009. **41**(4): p. 711-6.
5. Trampuz A, et al., *Clinical review: Severe malaria.* Crit Care, 2003. **7**(4): p. 315-23.
6. Viriyavejakul, et al., *Liver changes in severe Plasmodium falciparum malaria: histopathology, apoptosis and nuclear factor kappa B expression.* Malar J, 2014. **13**(1): p. 106.
7. Kerlin, et al., *Preferential invasion by Plasmodium merozoites and the self-regulation of parasite burden.* PLoS One, 2013. **8**(2): p. e57434.
8. J., L., *J. B. S. Haldane (1949) on infectious disease and evolution.* Genetics, 1999. **153**(1): p. 1-3.
9. Michalakis Y, et al., *Malaria: Evolution in vector control.* Nature, 2009. **462**(7271): p. 298-300.
10. Cowman AF, et al., *Invasion of red blood cells by malaria parasites.* Cell, 2006. **124**(4): p. 755-66.
11. Wright, et al., *Plasmodium falciparum Erythrocyte Invasion: Combining Function with Immune Evasion.* PLoS Pathog, 2014. **10**(3): p. e1003943.
12. Liu J, et al., *Plasmodium falciparum ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems.* Proc Natl Acad Sci U S A, 2006. **103**(23): p. 8840-5.
13. Yap, A., et al., *Conditional expression of apical membrane antigen 1 in Plasmodium falciparum shows it is required for erythrocyte invasion by merozoites.* Cell Microbiol, 2014.
14. Gao, X., et al., *Triggers of key calcium signals during erythrocyte invasion by Plasmodium falciparum.* Nat Commun, 2013. **4**: p. 2862.
15. Lew VL, et al., *Excess hemoglobin digestion and the osmotic stability of Plasmodium falciparum-infected red blood cells.* Blood, 2003. **101**(10): p. 4189-94.
16. Esposito, et al., *FRET imaging of hemoglobin concentration in Plasmodium falciparum-infected red cells.* PLoS One, 2008. **3**(11): p. e3780.
17. Baton, et al., *Spreading the seeds of million-murdering death: metamorphoses of malaria in the mosquito.* Trends Parasitol, 2005. **21**(12): p. 573-80.
18. Elliott DA, et al., *Four distinct pathways of hemoglobin uptake in the malaria parasite Plasmodium falciparum.* Proc Natl Acad Sci U S A, 2008. **105**(7): p. 2463-8.
19. Goldberg, D.E., et al., *Hemoglobin degradation in the malaria parasite Plasmodium falciparum: an ordered process in a unique organelle.* Proc Natl Acad Sci U S A, 1990. **87**(8): p. 2931-5.
20. Banerjee R, et al., *Food vacuole plasmepsins are processed at a conserved site by an acidic convertase activity in Plasmodium falciparum.* Mol Biochem Parasitol, 2003. **129**(2): p. 157-65.
21. Coronado, et al., *Malarial hemozoin: From target to tool.* Biochim Biophys Acta, 2014. **1840**(6): p. 2032-2041.
22. Wickham ME, et al., *Selective Inhibition of a Two-step Egress of Malaria Parasites from the Host Erythrocyte.* J Biol Chem, 2003. **278**(39): p. 37658-63.
23. Callan-Jones, et al., *Red blood cell membrane dynamics during malaria parasite egress.* Biophys J, 2012. **103**(12): p. 2475-83.

24. Garg S, et al., *Calcium-dependent permeabilization of erythrocytes by a perforin-like protein during egress of malaria parasites.* Nat Commun. , 2013. **4:**1736. doi: 10.1038/ncomms2725.
25. Collins CR, et al., *Malaria Parasite cGMP-dependent Protein Kinase Regulates Blood Stage Merozoite Secretory Organelle Discharge and Egress.* PLoS Pathog, 2013. **9**(5): p. doi: 10.1371/journal.ppat.1003344. Epub 2013 May 9.
26. Lew VL., *Malaria: surprising mechanism of merozoite egress revealed.* Curr Biol, 2011. **21**(9): p. R314-6.
27. Arora, U., et al., *Emergence of drug resistance in India.* J Indian Med Assoc, 2008. **106**(10): p. 678-81, 683.
28. Farooq U, et al., *Drug resistance in malaria.* J Vector Borne Dis, 2004. **41**(3-4): p. 45-53.
29. Arama, C. et al., *The path of malaria vaccine development: challenges and perspectives.* J Intern Med, 2014.
30. Phyo AP, et al., *Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study.* Lancet. 2012 May 26;379(9830):1960-6. doi: 10.1016/S0140-6736(12)60484-X. Epub 2012 Apr 5., 2012. **369**(9830):1960-6. doi: 10.1016/S0140-6736(12)60484-X. Epub 2012 Apr 5.
31. Dev, V. et al., *The Dominant Mosquito Vectors of Human Malaria in India*. Anopheles mosquitoes - New insights into malaria vectors. 2013.
32. Bloland PB, *Drug resistance in malaria.* World Helath Organization, 2001.
33. Martin SK, et al., *Reversal of chloroquine resistance in Plasmodium falciparum by verapamil.* Science, 1987. **235**(4791): p. 899-901.
34. Mehlotra RK, et al., *Evolution of a unique Plasmodium falciparum chloroquine-resistance phenotype in association with pfcrt polymorphism in Papua New Guinea and South America.* Proc Natl Acad Sci U S A, 1987. **98**(22): p. 12689-94.
35. Lewis, I.A., et al., *Metabolic QTL analysis links chloroquine resistance in Plasmodium falciparum to impaired hemoglobin catabolism.* PLoS Genet, 2014. **10**(1): p. e1004085.
36. Plowe CV, et al., *P. falciparum dihydrofolate reductase and dihydropteroate synthase mutations: epidemiology and role in clinical resistance to antifolates.* Drug Resist Updat. 1998;1(6):389-96., 1998. **1**(16): p. 389-96.
37. McCollum, A.M., et al., *Antifolate resistance in Plasmodium falciparum: multiple origins and identification of novel dhfr alleles.* J Infect Dis, 2006. **194**(2): p. 189-97.
38. Fisher, N., et al., *Cytochrome b mutation Y268S conferring atovaquone resistance phenotype in malaria parasite results in reduced parasite bc1 catalytic turnover and protein expression.* J Biol Chem, 2012. **287**(13): p. 9731-41.
39. Na-Bangchang, K., et al., *Identification of resistance of Plasmodium falciparum to artesunate-mefloquine combination in an area along the Thai-Myanmar border: integration of clinico-parasitological response, systemic drug exposure, and in vitro parasite sensitivity.* Malar J, 2013. **12**: p. 263.
40. Tall, A., et al., *Efficacy of artesunate plus amodiaquine, artesunate plus sulfadoxine-pyrimethamine, and chloroquine plus sulfadoxine-pyrimethamine in patients with uncomplicated Plasmodium falciparum in the Comoros Union.* Acta Trop, 2007. **102**(3): p. 176-81.
41. Savini, H., et al., *First case of emergence of atovaquone-proguanil resistance in Plasmodium falciparum during treatment in a traveler in Comoros.* Antimicrob Agents Chemother, 2008. **52**(6): p. 2283-4.
42. Charman, S.A., et al., *Synthetic ozonide drug candidate OZ439 offers new hope for a single-dose cure of uncomplicated malaria.* Proc Natl Acad Sci U S A, 2011. **108**(11): p. 4400-5.
43. Moehrle, J.J., et al., *First-in-man safety and pharmacokinetics of synthetic ozonide OZ439 demonstrates an improved exposure profile relative to other peroxide antimalarials.* Br J Clin Pharmacol, 2013. **75**(2): p. 524-37.

44. Perry TL, et al., *Severe atovaquone-resistant Plasmodium falciparum malaria in a Canadian traveller returned from the Indian subcontinent.* Open Med, 2009. **3**(1): p. e10-6.
45. Pradines B, et al., *Absence of association between pyronaridine in vitro responses and polymorphisms in genes involved in quinoline resistance in Plasmodium falciparum.* Malar J. 2010 Nov 25;9:339. doi: 10.1186/1475-2875-9-339., 2010: p.
46. Taylor, S.M., et al., *Artemisinin combination therapies and malaria parasite drug resistance: The game is afoot.* J Infect Dis, 2014.
47. O'Neill PM, et al., *The molecular mechanism of action of artemisinin--the debate continues.* Molecules, 2010. **15**(3): p. 1705-21. 1705-21. doi: 10.3390/molecules 15031705.
48. Klonis, N., et al., *Artemisinin activity against Plasmodium falciparum requires hemoglobin uptake and digestion.* Proc Natl Acad Sci U S A, 2011. **108**(28): p. 11405-10.
49. Ariey F, et al., *A molecular marker of artemisinin-resistant Plasmodium falciparum malaria.* Nature, 2014. **505**(7481): p. 50-5. doi: 10.1038/nature12876. Epub 2013 Dec 18.
50. Deitsch, K.W., *Malaria virulence genes controlling expression through chromatin modification.* Cell, 2005. **121**(1): p. 1-2.
51. Rosenthal, P.J., *Cysteine proteases of malaria parasites.* Int J Parasitol, 2004. **34**(13-14): p. 1489-99.
52. Lucet, I.S., et al., *Plasmodium kinases as targets for new-generation antimalarials.* Future Med Chem, 2012. **4**(18): p. 2295-310.
53. Przyborski, J. et al., *Parasitology. The malarial secretome.* Science, 2004. **306**(5703): p. 1897-8.
54. Bautista, J.M., et al., *Malaria proteomics: insights into the parasite-host interactions in the pathogenic space.* J Proteomics, 2014. **97**: p. 107-25.
55. Perez, B., et al., *Development of Plasmodium falciparum protease inhibitors in the past decade (2002-2012).* Curr Med Chem, 2013. **20**(25): p. 3049-68.
56. Chugh M, et al., *Protein complex directs hemoglobin-to-hemozoin formation in Plasmodium falciparum.* Proc Natl Acad Sci U S A, 2013. **110**(14): p. 5392-7. doi: 10.1073/pnas.1218412110. .
57. Goldberg, D.E., *Complex nature of malaria parasite hemoglobin degradation [corrected].* Proc Natl Acad Sci U S A, 2013. **110**(14): p. 5283-4.
58. AK, G., *Aspartic Acid Proteases as Therapeutic Targets.* WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31811-7, 2010.
59. McKerrow JH, et al., *Proteases in parasitic diseases.* Annu Rev Pathol., 2006. **1**: p. 497-536.
60. Wu Y, et al., *Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite.* Genome Res., 2003. **13**(4): p. 601-16.
61. Tang J, W.R., *Evolution in the structure and function of aspartic proteases.* J Cell Biochem., 1987. **33**(1): p. 53-63.
62. Walker, W.G., et al., *Relation between blood pressure and renin, renin substrate, angiotensin II, aldosterone and urinary sodium and potassium in 574 ambulatory subjects.* Hypertension, 1979. **1**(3): p. 287-91.
63. Bhaumik, P., A. Gustchina, and A. Wlodawer, *Structural studies of vacuolar plasmepsins.* Biochim Biophys Acta, 2012. **1824**(1): p. 207-23.
64. Klemba M, et al., *Trafficking of plasmepsin II to the food vacuole of the malaria parasite Plasmodium falciparum.* J Cell Biol, 2004. **164**(1): p. 47-56.
65. Coombs GH, et al., *Aspartic proteases of Plasmodium falciparum and other parasitic protozoa as drug targets.* Trends Parasitol, 2001. **17**(11): p. 532-7.
66. Boddey JA, et al., *An aspartyl protease directs malaria effector proteins to the host cell.* Nature, 2010. **463**(7281):627-31. doi: 10.1038/nature08728.).

67. Rawat, M., et al., *Imperfect duplicate insertions type of mutations in plasmepsin V modulates binding properties of PEXEL motifs of export proteins in Indian Plasmodium vivax.* PLoS One, 2013. **8**(3): p. e60077.
68. Boddey, J.A., et al., *Role of plasmepsin V in export of diverse protein families from the Plasmodium falciparum exportome.* Traffic, 2013. **14**(5): p. 532-50.
69. C., B., *Plasmepsins as antimalarial targets.* Curr Opin Drug Discov Devel, 2000. **3**(5): p. 624-9.
70. Miura, T., et al., *Optimization of plasmepsin inhibitor by focusing on similar structural feature with chloroquine to avoid drug-resistant mechanism of Plasmodium falciparum.* Bioorg Med Chem Lett, 2014.
71. Francis, S.E., et al., *Molecular characterization and inhibition of a Plasmodium falciparum aspartic hemoglobinase.* Embo j, 1994. **13**(2): p. 306-17.
72. Nezami, A., et al., *Identification and characterization of allophenylnorstatine-based inhibitors of plasmepsin II, an antimalarial target.* Biochemistry, 2002. **41**(7): p. 2273-80.
73. Moura, P.A., et al., *Role of Plasmodium falciparum digestive vacuole plasmepsins in the specificity and antimalarial mode of action of cysteine and aspartic protease inhibitors.* Antimicrob Agents Chemother, 2009. **53**(12): p. 4968-78.
74. Bhaumik, P., et al., *Crystal structures of the histo-aspartic protease (HAP) from Plasmodium falciparum.* J Mol Biol, 2009. **388**(3): p. 520-40.
75. Aureggi, V., et al., *Potent inhibitors of malarial aspartic proteases, the plasmepsins, by hydroformylation of substituted 7-azanorbornenes.* Chemistry, 2013. **19**(1): p. 155-64.
76. Jin, H., et al., *Dietary flavonoids fisetin and myricetin: Dual inhibitors of Plasmodium falciparum falcipain-2 and plasmepsin II.* Fitoterapia, 2014. **94c**: p. 55-61.
77. Saify, Z.S., et al., *New benzimidazole derivatives as antiplasmodial agents and plasmepsin inhibitors: synthesis and analysis of structure-activity relationships.* Bioorg Med Chem Lett, 2012. **22**(2): p. 1282-6.
78. Bonilla JA, et al., *Critical roles for the digestive vacuole plasmepsins of Plasmodium falciparum in vacuolar function.* Mol Microbiol, 2007. **65**(1): p. 64-75.
79. Luksch, T., et al., *Pyrrolidine derivatives as plasmepsin inhibitors: binding mode analysis assisted by molecular dynamics simulations of a highly flexible protein.* ChemMedChem, 2010. **5**(3): p. 443-54.
80. Chapman HA, et al., *Emerging roles for cysteine proteases in human biology.* Annu Rev Physiol, 1997. **59**: p. 63-88.
81. McCoubrie, J.E., et al., *Evidence for a common role for the serine-type Plasmodium falciparum serine repeat antigen proteases: implications for vaccine and drug design.* Infect Immun, 2007. **75**(12): p. 5565-74.
82. Rawlings, N.D., et al., *MEROPS: the database of proteolytic enzymes, their substrates and inhibitors.* Nucleic Acids Res, 2014. **42**(Database issue): p. D503-9.
83. Rawlings ND, et al., *Families of cysteine peptidases.* Methods Enzymol, 1994. **244**: p. 461-86.
84. Barrett AJ., *Proteases.* Curr Protoc Protein Sci, 2001. **Chapter 21:**Unit 21.1. doi: 10.1002/0471140864.ps2101s21.
85. Wang SX, et al., *Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease.* Proc Natl Acad Sci U S A., 2006. **103**(31): p. 11503-8. Epub 2006 Jul 24.
86. Salmon BL, et al., *Malaria parasite exit from the host erythrocyte: a two-step process requiring extraerythrocytic proteolysis.* Proc Natl Acad Sci U S A., 2001. **98**(1): p. 271-6.
87. Rosenthal PJ., *Cysteine proteases of malaria parasites.* Int J Parasitol, 2004. **34**(13-14): p. 1489-99.
88. Sijwali PS, et al., *Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of Plasmodium falciparum.* Mol Biochem Parasitol., 2006. **150**(1): p. 96-106. Epub 2006 Jul 20.

89. Kumar A, et al., *Falcipain-1, a Plasmodium falciparum cysteine protease with vaccine potential.* Infect Immun. **75**(4): p. 2026-2034.
90. Wang, F., et al., *Biochemical characterization of Plasmodium falciparum dipeptidyl aminopeptidase 1.* Mol Biochem Parasitol, 2011. **175**(1): p. 10-20.
91. Moura PA, et al., *Role of Plasmodium falciparum digestive vacuole plasmepsins in the specificity and antimalarial mode of action of cysteine and aspartic protease inhibitors.* Antimicrob Agents Chemother., 2009. **53**(12): p. 4968-78.doi: 10.1128/AAC.00882-09. Epub 2009 Sep 14.
92. Arastu-Kapur S, et al., *Identification of proteases that regulate erythrocyte rupture by the malaria parasite Plasmodium falciparum.* Nat Chem Biol., 2008. **4**(3): p. 203-13. doi: 10.1038/nchembio.70. Epub 2008 Feb 3.
93. Fairlie WD, et al., *Inhibition of malaria parasite development by a cyclic peptide that targets the vital parasite protein SERA5.* Infect Immun., 2008. **76**(9): p. 4332-44.doi: 10.1128/IAI.00278-08. Epub 2008 Jun 30.
94. Hodder, A.N., et al., *Enzymic, phylogenetic, and structural characterization of the unusual papain-like protease domain of Plasmodium falciparum SERA5.* J Biol Chem, 2003. **278**(48): p. 48169-77.
95. Hodder, A.N., et al., *Structural insights into the protease-like antigen Plasmodium falciparum SERA5 and its noncanonical active-site serine.* J Mol Biol, 2009. **392**(1): p. 154-65.
96. Dasaradhi PV, et al., *A role of falcipain-2, principal cysteine proteases of Plasmodium falciparum in merozoite egression.* Biochem Biophys Res Commun., 2005. **336**(4): p. 1062-8.
97. Olson JE, et al., *Antimalarial effects in mice of orally administered peptidyl cysteine protease inhibitors.* Bioorg Med Chem., 1999. **7**(4): p. 633-8.
98. Shenai BR, et al., *Structure-activity relationships for inhibition of cysteine protease activity and development of Plasmodium falciparum by peptidyl vinyl sulfones.* Antimicrob Agents Chemother., 2003. **47**(1): p. 154-60.
99. Lee BJ, et al., *Antimalarial activities of novel synthetic cysteine protease inhibitors.* Antimicrob Agents Chemother., 2003. **47**(12): p. 3810-4.
100. Ring CS, et al., *Structure-based inhibitor design by using protein models for the development of antiparasitic agents.* Proc Natl Acad Sci U S A., 1993. **90**(8): p. 3583-7.
101. Rathi, B., et al., *Functionalized hydroxyethylamine based peptide nanostructures as potential inhibitors of falcipain-3, an essential proteases of Plasmodium falciparum.* Bioorg Med Chem, 2013. **21**(17): p. 5503-9.
102. Kesharwani, R.K., et al., *Computation-based virtual screening for designing novel antimalarial drugs by targeting falcipain-III: a structure-based drug designing approach.* J Vector Borne Dis, 2013. **50**(2): p. 93-102.
103. Wang, J., et al., *Structural features of falcipain-3 inhibitors: an in silico study.* Mol Biosyst, 2013. **9**(9): p. 2296-310.
104. Bova, F., et al., *Constrained peptidomimetics as antiplasmodial falcipain-2 inhibitors.* Bioorg Med Chem, 2010. **18**(14): p. 4928-38.
105. Huang, H., et al., *Design and synthesis of small molecular dual inhibitor of falcipain-2 and dihydrofolate reductase as antimalarial agent.* Bioorg Med Chem Lett, 2012. **22**(2): p. 958-62.
106. Marques, A.F., et al., *Falcipain-2 inhibition by suramin and suramin analogues.* Bioorg Med Chem, 2013. **21**(13): p. 3667-73.
107. Stolze, S.C., et al., *The antimalarial natural product symplostatin 4 is a nanomolar inhibitor of the food vacuole falcipains.* Chem Biol, 2012. **19**(12): p. 1546-55.
108. Deu, E., et al., *Functional studies of Plasmodium falciparum dipeptidyl aminopeptidase I using small molecule inhibitors and active site probes.* Chem Biol, 2010. **17**(8): p. 808-19.

109. Ch'ng JH, et al., *A programmed cell death pathway in the malaria parasite Plasmodium falciparum has general features of mammalian apoptosis but is mediated by clan CA cysteine proteases.* Cell Death Dis., 2010. **1:e26.** doi: 10.1038/cddis.2010.2.
110. Koeppe RE 2nd, et al., *Mechanism of hydrolysis by serine proteases: direct determination of the pKa's of aspartyl-102 and aspartyl-194 in bovine trypsin using difference infrared spectroscopy.* Biochemistry., 1976. **15**(16): p. 3450-8.
111. Chapman, H.A., et al., *Emerging roles for cysteine proteases in human biology.* Annu Rev Physiol, 1997. **59**: p. 63-88.
112. Silmon de Monerri, N.C., et al., *Global identification of multiple substrates for Plasmodium falciparum SUB1, an essential malarial processing protease.* Infect Immun, 2011. **79**(3): p. 1086-97.
113. Yeoh S, et al., *Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes.* Cell., 2007. **131**(6): p. 1072-83.
114. Withers-Martinez, C., et al., *Plasmodium subtilisin-like protease 1 (SUB1): insights into the active-site structure, specificity and function of a pan-malaria drug target.* Int J Parasitol, 2012. **42**(6): p. 597-612.
115. Janse CJ, et al., *The exoneme helps malaria parasites to break out of blood cells.* Cell., 2007. **131**(16): p. 1036-8.
116. Silmon de Monerri, N.C., et al., *Global identification of multiple substrates for Plasmodium falciparum SUB1, an essential malarial processing protease.* Infect Immun, 2011. **79**(3): p. 1086-97.
117. Koussis, K., et al., *A multifunctional serine protease primes the malaria parasite for red blood cell invasion.* Embo j, 2009. **28**(6): p. 725-35.
118. Suarez C, et al., *The malarial serine protease SUB1 plays an essential role in parasite liver stage development.* PLoS Pathog. 2013 Dec;9(12, 2013. **9**(12).
119. Harris PK, et al., *Molecular identification of a malaria merozoite surface sheddase.* PLoS Pathog., 2005. **1**(3): p. 241-251.
120. Green, J.L., et al., *Plasmodium thrombospondin related apical merozoite protein (PTRAMP) is shed from the surface of merozoites by PfSUB2 upon invasion of erythrocytes.* Mol Biochem Parasitol, 2006. **150**(1): p. 114-7.
121. Child, M.A., et al., *Molecular determinants for subcellular trafficking of the malarial sheddase PfSUB2.* Traffic, 2013. **14**(10): p. 1053-64.
122. Barale, J.C., et al., *Plasmodium falciparum subtilisin-like protease 2, a merozoite candidate for the merozoite surface protein 1-42 maturase.* Proc Natl Acad Sci U S A, 1999. **96**(11): p. 6445-50.
123. Roggwiller E, et al., *A role for erythrocyte band 3 degradation by the parasite gp76 serine protease in the formation of the parasitophorous vacuole during invasion of erythrocytes by Plasmodium falciparum.* Mol Biochem Parasitol., 1996. **82**(1): p. 13-24.
124. Sibley, L.D., *The roles of intramembrane proteases in protozoan parasites.* Biochim Biophys Acta, 2013. **1828**(12): p. 2908-15.
125. Ejigiri, I., et al., *Shedding of TRAP by a rhomboid protease from the malaria sporozoite surface is essential for gliding motility and sporozoite infectivity.* PLoS Pathog, 2012. **8**(7): p. e1002725.
126. Singh S, et al., *Mononeme: a new secretory organelle in Plasmodium falciparum merozoites identified by localization of rhomboid-1 protease.* Proc Natl Acad Sci U S A., 2007. **104**(50): p. 20043-8.
127. Baker RP, et al., *Two Plasmodium rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria.* PLoS Pathog., 2006. **2**(10).
128. Vera, I.M., et al., *Plasmodium protease ROM1 is important for proper formation of the parasitophorous vacuole.* PLoS Pathog, 2011. **7**(9): p. e1002197.
129. Tawk, L., et al., *A key role for Plasmodium subtilisin-like SUB1 protease in egress of malaria parasites from host hepatocytes.* J Biol Chem, 2013. **288**(46): p. 33336-46.

130. Agarwal, S., et al., *Ca(2+) -mediated exocytosis of subtilisin-like protease 1: a key step in egress of Plasmodium falciparum merozoites.* Cell Microbiol, 2013. **15**(6): p. 910-21.
131. Gemma, S., et al., *Quinolylhydrazones as novel inhibitors of Plasmodium falciparum serine protease PfSUB1.* Bioorg Med Chem Lett, 2012. **22**(16): p. 5317-21.
132. Alam, A., et al., *Proteolytic activity of Plasmodium falciparum subtilisin-like protease 3 on parasite profilin, a multifunctional protein.* Mol Biochem Parasitol, 2013. **191**(2): p. 58-62.
133. Alam, A., et al., *Expression and characterization of catalytic domain of Plasmodium falciparum subtilisin-like protease 3.* Mol Biochem Parasitol, 2012. **183**(1): p. 84-9.
134. Vera IM, et al., *Plasmodium protease ROM1 is important for proper formation of the parasitophorous vacuole.* PLoS Pathog., 2011. **7**(9).
135. Seemuller E, et al., *Autocatalytic processing of the 20S proteasome.* Nature., 2004. **382**(6590): p. 468-71.
136. Lecker SH, et al., *Protein degradation by the ubiquitin-proteasome pathway in normal and disease states.* J Am Soc Nephrol., 2006. **17**(7): p. 1807-19.
137. Mordmuller, B., et al., *Plasmodia express two threonine-peptidase complexes during asexual development.* Mol Biochem Parasitol, 2006. **148**(1): p. 79-85.
138. Mordmuller, B., et al., *Plasmodia express two threonine-peptidase complexes during asexual development.* Mol Biochem Parasitol, 2006. **148**(1): p. 79-85.
139. Chung, D.W., et al., *Characterization of the ubiquitylating components of the human malaria parasite's protein degradation pathway.* PLoS One, 2012. **7**(8): p. e43477.
140. Jain, S., et al., *The prokaryotic ClpQ protease plays a key role in growth and development of mitochondria in Plasmodium falciparum.* Cell Microbiol, 2013. **15**(10): p. 1660-73.
141. Koyama, F.C., et al., *Ubiquitin proteasome system and the atypical kinase PfPK7 are involved in melatonin signaling in Plasmodium falciparum.* J Pineal Res, 2012. **53**(2): p. 147-53.
142. Aminake, M.N., et al., *The proteasome of malaria parasites: A multi-stage drug target for chemotherapeutic intervention?* Int J Parasitol Drugs Drug Resist, 2012. **2**: p. 1-10.
143. Chung, D.W. and K.G. Le Roch, *Targeting the Plasmodium ubiquitin/proteasome system with anti-malarial compounds: promises for the future.* Infect Disord Drug Targets, 2010. **10**(3): p. 158-64.
144. Ramasamy G, et al., *Characterization and localization of Plasmodium falciparum homolog of prokaryotic ClpQ/HslV protease.* Mol Biochem Parasitol., 2007. **152**(2): p. 139-48.
145. Tschan S, et al., *Threonine peptidases as drug targets against malaria.* Expert Opin Ther Targets., 2011. **15**(4): p. 365-78. doi: 10.1517/14728222.2011.555399. Epub 2011 Feb 1.
146. Prasad, R., et al., *Blocking Plasmodium falciparum development via dual inhibition of hemoglobin degradation and the ubiquitin proteasome system by MG132.* PLoS One, 2013. **8**(9): p. e73530.
147. Tschan, S., et al., *Broad-spectrum antimalarial activity of peptido sulfonyl fluorides, a new class of proteasome inhibitors.* Antimicrob Agents Chemother, 2013. **57**(8): p. 3576-84.
148. Mansfeld J, et al., *Probing the unfolding region in a thermolysin-like protease by site-specific immobilization.* Biochemistry., 1999. **38**(26): p. 8240-5.
149. Rawlings ND, et al., *MEROPS: the database of proteolytic enzymes, their substrates and inhibitors.* Nucleic Acids Res., 2014. **42**(Database issue):D503-9. doi: 10.1093/nar/gkt953. Epub 2013 Oct 23.
150. Hooper NM., *Families of zinc metalloproteases.* FEBS Lett., 1994. **354**(1): p. 1-6.
151. Murata CE, et al., *Plasmodium falciparum falcilysin: a metalloprotease with dual specificity.* J Biol Chem., 2003. **278**(39): p. 38022-8.
152. Amata O, et al., *Human insulin-degrading enzyme working mechanism.* J Am Chem Soc., 2009. **131**((41):14804-11. doi: 10.1021/ja9037142.).

153. Lhouvum, K., et al., *Insight into structural and biochemical determinants of substrate specificity of PFI1625c: correlation analysis of protein-peptide molecular models.* J Mol Graph Model, 2013. **43**: p. 21-30.
154. Eggleson KK, et al., *Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite Plasmodium falciparum.* J Biol Chem., 1999. **274**(45): p. 32411-7.
155. Lauterbach SB, et al., *The M18 aspartyl aminopeptidase of Plasmodium falciparum binds to human erythrocyte spectrin in vitro.* Malar J., 2008. **7**(161. doi: 10.1186/1475-2875-7-161.).
156. Teuscher F, et al., *The M18 aspartyl aminopeptidase of the human malaria parasite Plasmodium falciparum.* J Biol Chem., 2007. **282**(42): p. 30817-26.
157. Dalal S, et al., *Roles for two aminopeptidases in vacuolar hemoglobin catabolism in Plasmodium falciparum.* J Biol Chem., 2007. **282**(49): p. 35978-87.
158. Allary M, et al., *Properties, stage-dependent expression and localization of Plasmodium falciparum M1 family zinc-aminopeptidase.* Parasitology., 2002. **125**(Pt 1): p. 1-10.
159. McGowan S, et al., *Structural basis for the inhibition of the essential Plasmodium falciparum M1 neutral aminopeptidase.* Proc Natl Acad Sci U S A., 2009. **106**(8): p. 2537-42. doi: 10.1073/pnas.0807398106. Epub 2009 Feb 5.
160. Trenholme, K.R., et al., *Aminopeptidases of malaria parasites: new targets for chemotherapy.* Infect Disord Drug Targets, 2010. **10**(3): p. 217-25.
161. Dalal, S., et al., *Engagement of the S1, S1' and S2' subsites drives efficient catalysis of peptide bond hydrolysis by the M1-family aminopeptidase from Plasmodium falciparum.* Mol Biochem Parasitol, 2012. **183**(1): p. 70-7.
162. Ragheb, D., et al., *Distribution and biochemical properties of an M1-family aminopeptidase in Plasmodium falciparum indicate a role in vacuolar hemoglobin catabolism.* J Biol Chem, 2011. **286**(31): p. 27255-65.
163. Ragheb, D., et al., *Evidence for catalytic roles for Plasmodium falciparum aminopeptidase P in the food vacuole and cytosol.* J Biol Chem, 2009. **284**(37): p. 24806-15.
164. Harbut, M.B., et al., *Bestatin-based chemical biology strategy reveals distinct roles for malaria M1- and M17-family aminopeptidases.* Proc Natl Acad Sci U S A, 2011. **108**(34): p. E526-34.
165. Stack CM, et al., *Characterization of the Plasmodium falciparum M17 leucyl aminopeptidase. A protease involved in amino acid regulation with potential for antimalarial drug development.* J Biol Chem., 2007. **282**(3): p. 2069-80.
166. Dalal, S., et al., *A naturally variable residue in the S1 subsite of M1 family aminopeptidases modulates catalytic properties and promotes functional specialization.* J Biol Chem, 2013. **288**(36): p. 26004-12.
167. Azimzadeh O, et al., *Plasmodium falciparum PfA-M1 aminopeptidase is trafficked via the parasitophorous vacuole and marginally delivered to the food vacuole.* Malar J. , 2010. **9**(189. doi: 10.1186/1475-2875-9-189.).
168. Ponpuak, M., et al., *A role for falcilysin in transit peptide degradation in the Plasmodium falciparum apicoplast.* Mol Microbiol, 2007. **63**(2): p. 314-34.
169. Chen, X., et al., *Inhibitors of Plasmodium falciparum methionine aminopeptidase 1b possess antimalarial activity.* Proc Natl Acad Sci U S A, 2006. **103**(39): p. 14548-53.
170. Chen, X., et al., *Fumagillin and fumarranol interact with P. falciparum methionine aminopeptidase 2 and inhibit malaria parasite growth in vitro and in vivo.* Chem Biol, 2009. **16**(2): p. 193-202.
171. Tanveer, A., et al., *An FtsH protease is recruited to the mitochondrion of Plasmodium falciparum.* PLoS One, 2013. **8**(9): p. e74408.
172. Kannan Sivaraman, K., et al., *Synthesis and structure-activity relationships of phosphonic arginine mimetics as inhibitors of the M1 and M17 aminopeptidases from Plasmodium falciparum.* J Med Chem, 2013. **56**(12): p. 5213-7.

173. Velmourougane, G., et al., *Synthesis of new (-)-bestatin-based inhibitor libraries reveals a novel binding mode in the S1 pocket of the essential malaria M1 metalloaminopeptidase.* J Med Chem, 2011. **54**(6): p. 1655-66.
174. Gardiner, D.L., et al., *Overexpression of leucyl aminopeptidase in Plasmodium falciparum parasites. Target for the antimalarial activity of bestatin.* J Biol Chem, 2006. **281**(3): p. 1741-5.
175. Cunningham, E., et al., *Chemical target validation studies of aminopeptidase in malaria parasites using alpha-aminoalkylphosphonate and phosphonopeptide inhibitors.* Antimicrob Agents Chemother, 2008. **52**(9): p. 3221-8.
176. Skinner-Adams, T.S., et al., *Identification of phosphinate dipeptide analog inhibitors directed against the Plasmodium falciparum M17 leucine aminopeptidase as lead antimalarial compounds.* J Med Chem, 2007. **50**(24): p. 6024-31.
177. Skinner-Adams, T.S., et al., *Plasmodium falciparum neutral aminopeptidases: new targets for anti-malarials.* Trends Biochem Sci, 2010. **35**(1): p. 53-61.
178. Zhang, P., et al., *Angiogenesis inhibitors specific for methionine aminopeptidase 2 as drugs for malaria and leishmaniasis.* J Biomed Sci, 2002. **9**(1): p. 34-40.
179. Kitjaroentham, A., et al., *Effect of metalloprotease inhibitors on invasion of red blood cell by Plasmodium falciparum.* Acta Trop, 2006. **97**(1): p. 5-9.
180. Prato, M., et al., *Matrix Metalloproteinase-9 and Haemozoin: Wedding Rings for Human Host and Plasmodium falciparum Parasite in Complicated Malaria.* J Trop Med, 2011. **2011**: p. 628435.
181. Trager, W. and J.B. Jenson, *Cultivation of malarial parasites.* Nature, 1978. **273**(5664): p. 621-2.
182. Zhou, M.Y. and C.E. Gomez-Sanchez, *Universal TA cloning.* Curr Issues Mol Biol, 2000. **2**(1): p. 1-7.
183. Osborne, C. and S.A. Brooks, *SDS-PAGE and Western blotting to detect proteins and glycoproteins of interest in breast cancer research.* Methods Mol Med, 2006. **120**: p. 217-29.
184. Kumar, A., et al., *Falcipain-1, a Plasmodium falciparum cysteine protease with vaccine potential.* Infect Immun, 2007. **75**(4): p. 2026-34.
185. Hornbeck, P., *Enzyme-linked immunosorbent assays.* Curr Protoc Immunol, 2001. **Chapter 2**: p. Unit 2.1.
186. Flick, K., et al., *Optimized expression of Plasmodium falciparum erythrocyte membrane protein 1 domains in Escherichia coli.* Malar J, 2004. **3**: p. 50.
187. Strandberg, L. and S.O. Enfors, *Factors influencing inclusion body formation in the production of a fused protein in Escherichia coli.* Appl Environ Microbiol, 1991. **57**(6): p. 1669-74.
188. Coutard, B., et al., *Single pH buffer refolding screen for protein from inclusion bodies.* Protein Expr Purif, 2012. **82**(2): p. 352-9.
189. Oganesyan, N., et al., *Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization.* Protein Expr Purif, 2007. **52**(2): p. 280-5.
190. Prasad, S., et al., *Effect of chemical chaperones in improving the solubility of recombinant proteins in Escherichia coli.* Appl Environ Microbiol, 2011. **77**(13): p. 4603-9.
191. Van den Berg, B., et al., *Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell.* Embo j, 2000. **19**(15): p. 3870-5.
192. Alibolandi, M. and H. Mirzahoseini, *Chemical assistance in refolding of bacterial inclusion bodies.* Biochem Res Int, 2011. **2011**: p. 631607.
193. Bushmarina, N.A., et al., *Cofactor effects on the protein folding reaction: acceleration of alpha-lactalbumin refolding by metal ions.* Protein Sci, 2006. **15**(4): p. 659-71.
194. Yadava, A. and C.F. Ockenhouse, *Effect of codon optimization on expression levels of a functionally folded malaria vaccine candidate in prokaryotic and eukaryotic expression systems.* Infect Immun, 2003. **71**(9): p. 4961-9.

195. Miller, S.K., et al., *A subset of Plasmodium falciparum SERA genes are expressed and appear to play an important role in the erythrocytic cycle.* J Biol Chem, 2002. **277**(49): p. 47524-32.
196. Chung, C.T., et al., *One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution.* Proc Natl Acad Sci U S A, 1989. **86**(7): p. 2172-5.
197. Engebrecht, J., et al., *Minipreps of plasmid DNA.* Curr Protoc Mol Biol, 2001. **Chapter 1**: p. Unit1.6.
198. Seidman, C.E., et al., *Introduction of plasmid DNA into cells.* Curr Protoc Mol Biol, 2001. **Chapter 1**: p. Unit1.8.
199. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.
200. Taylor, A.B., et al., *Crystal structures of mitochondrial processing peptidase reveal the mode for specific cleavage of import signal sequences.* Structure, 2001. **9**(7): p. 615-25.
201. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.
202. Ramachandran, G.N., C. et al., *Stereochemistry of polypeptide chain configurations.* J Mol Biol, 1963. **7**: p. 95-9.
203. Colovos, C. and T.O. Yeates, *Verification of protein structures: patterns of nonbonded atomic interactions.* Protein Sci, 1993. **2**(9): p. 1511-9.
204. Luthy, R., J.U. Bowie, and D. Eisenberg, *Assessment of protein models with three-dimensional profiles.* Nature, 1992. **356**(6364): p. 83-5.
205. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein.* J Mol Biol, 1982. **157**(1): p. 105-32.
206. Dereeper, A., et al., *Phylogeny.fr: robust phylogenetic analysis for the non-specialist.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W465-9.
207. Chou, K.C. and H.B. Shen, *ProtIdent: a web server for identifying proteases and their types by fusing functional domain and sequential evolution information.* Biochem Biophys Res Commun, 2008. **376**(2): p. 321-5.
208. Becker, A.B. and R.A. Roth, *Identification of glutamate-169 as the third zinc-binding residue in proteinase III, a member of the family of insulin-degrading enzymes.* Biochem J, 1993. **292 (Pt 1)**: p. 137-42.
209. Laskowski, R.A. and M.B. Swindells, *LigPlot+: multiple ligand-protein interaction diagrams for drug discovery.* J Chem Inf Model, 2011. **51**(10): p. 2778-86.
210. Pelmenschikov, V., M.R. Blomberg, and P.E. Siegbahn, *A theoretical study of the mechanism for peptide hydrolysis by thermolysin.* J Biol Inorg Chem, 2002. **7**(3): p. 284-98.
211. Yokoyama, M., et al., *Structural basis for specific recognition of substrates by sapovirus protease.* Front Microbiol, 2012. **3**: p. 312.
212. Kiszewski, A.E., *Blocking Plasmodium falciparum Malaria Transmission with Drugs: The Gametocytocidal and Sporontocidal Properties of Current and Prospective Antimalarials.* Pharmaceuticals, 2010. **4**(1): p. 44-68.
213. Aimes, R.T. and J.P. Quigley, *Matrix metalloproteinase-2 is an interstitial collagenase. Inhibitor-free enzyme catalyzes the cleavage of collagen fibrils and soluble native type I collagen generating the specific 3/4- and 1/4-length fragments.* J Biol Chem, 1995. **270**(11): p. 5872-6.
214. Kuhnel, B., et al., *Precise and efficient cleavage of recombinant fusion proteins using mammalian aspartic proteases.* Protein Eng, 2003. **16**(10): p. 777-83.
215. Yasmin, H., et al., *Amplified and selective assay of collagens by enzymatic and fluorescent reactions.* Sci Rep, 2014. **4**: p. 4950.
216. Hasty, K.A., et al., *The collagen substrate specificity of human neutrophil collagenase.* J Biol Chem, 1987. **262**(21): p. 10048-52.

217. Auld, D.S., *Removal and replacement of metal ions in metallopeptidases.* Methods Enzymol, 1995. **248**: p. 228-42.
218. Cheung, P.Y., et al., *Matrix metalloproteinase-2 contributes to ischemia-reperfusion injury in the heart.* Circulation, 2000. **101**(15): p. 1833-9.
219. Doddapaneni, K.K., et al., *Purification and characterization of a solvent and detergent-stable novel protease from Bacillus cereus.* Microbiol Res, 2009. **164**(4): p. 383-90.
220. Park, H.I., et al., *Effects of detergents on catalytic activity of human endometase/matrilysin 2, a putative cancer biomarker.* Anal Biochem, 2010. **396**(2): p. 262-8.
221. Iqbal, J., et al., *Persistent histidine-rich protein 2, parasite lactate dehydrogenase, and panmalarial antigen reactivity after clearance of Plasmodium falciparum monoinfection.* J Clin Microbiol, 2004. **42**(9): p. 4237-41.
222. Saito, T., et al., *A novel GDP-dependent pyruvate kinase isozyme from Toxoplasma gondii localizes to both the apicoplast and the mitochondrion.* J Biol Chem, 2008. **283**(20): p. 14041-52.
223. Fasciglione, G.F., et al., *pH- and temperature-dependence of functional modulation in metalloproteinases. A comparison between neutrophil collagenase and gelatinases A and B.* Biophys J, 2000. **79**(4): p. 2138-49.
224. Iqbal, J., et al., *Plasmodium falciparum histidine-rich protein 2-based immunocapture diagnostic assay for malaria: cross-reactivity with rheumatoid factors.* J Clin Microbiol, 2000. **38**(3): p. 1184-6.
225. Brooks, S.R. and K.C. Williamson, *Proteolysis of Plasmodium falciparum surface antigen, Pfs230, during gametogenesis.* Mol Biochem Parasitol, 2000. **106**(1): p. 77-82.
226. Torres, J.A., et al., *Plasmodium berghei: effect of protease inhibitors during gametogenesis and early zygote development.* Exp Parasitol, 2005. **111**(4): p. 255-9.
227. Deininger, M.H., et al., *Angiogenic proteins in brains of patients who died with cerebral malaria.* J Neuroimmunol, 2003. **142**(1-2): p. 101-11.
228. Dietmann, A., et al., *Matrix metalloproteinases and their tissue inhibitors (TIMPs) in Plasmodium falciparum malaria: serum levels of TIMP-1 are associated with disease severity.* J Infect Dis, 2008. **197**(11): p. 1614-20.
229. D'Alessandro, S., N. Basilico, and M. Prato, *Effects of Plasmodium falciparum-infected erythrocytes on matrix metalloproteinase-9 regulation in human microvascular endothelial cells.* Asian Pac J Trop Med, 2013. **6**(3): p. 195-9.
230. Van den Steen, P.E., et al., *Matrix metalloproteinases, tissue inhibitors of MMPs and TACE in experimental cerebral malaria.* Lab Invest, 2006. **86**(9): p. 873-88.
231. Szklarczyk, A., et al., *Glial activation and matrix metalloproteinase release in cerebral malaria.* J Neurovirol, 2007. **13**(1): p. 2-10.
232. Schneidman-Duhovny, D., et al., *PatchDock and SymmDock: servers for rigid and symmetric docking.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W363-7.
233. Andrusier, N., R. Nussinov, and H.J. Wolfson, *FireDock: fast interaction refinement in molecular docking.* Proteins, 2007. **69**(1): p. 139-59.
234. Ghosh, J.K., et al., *Selective cytotoxicity of dermaseptin S3 toward intraerythrocytic Plasmodium falciparum and the underlying molecular basis.* J Biol Chem, 1997. **272**(50): p. 31609-16.
235. Gwadz, R.W., et al., *Effects of magainins and cecropins on the sporogonic development of malaria parasites in mosquitoes.* Infect Immun, 1989. **57**(9): p. 2628-33.
236. Eda, K., S. Eda, and I.W. Sherman, *Identification of peptides targeting the surface of Plasmodium falciparum-infected erythrocytes using a phage display peptide library.* Am J Trop Med Hyg, 2004. **71**(2): p. 190-5.
237. Efron, L., et al., *Direct interaction of dermaseptin S4 aminoheptanoyl derivative with intraerythrocytic malaria parasite leading to increased specific antiparasitic activity in culture.* J Biol Chem, 2002. **277**(27): p. 24067-72.

238. Choi, S.J., et al., *Isolation and characterization of Psalmopeotoxin I and II: two novel antimalarial peptides from the venom of the tarantula Psalmopoeus cambridgei.* FEBS Lett, 2004. **572**(1-3): p. 109-17.
239. Boman, H.G., et al., *Antibacterial and antimalarial properties of peptides that are cecropin-melittin hybrids.* FEBS Lett, 1989. **259**(1): p. 103-6.
240. Shahabuddin, M., et al., *Plasmodium gallinaceum: differential killing of some mosquito stages of the parasite by insect defensin.* Exp Parasitol, 1998. **89**(1): p. 103-12.
241. Conde, R., et al., *Scorpine, an anti-malaria and anti-bacterial agent purified from scorpion venom.* FEBS Lett, 2000. **471**(2-3): p. 165-8.
242. Gao, B., et al., *Characterization of two linear cationic antimalarial peptides in the scorpion Mesobuthus eupeus.* Biochimie, 2010. **92**(4): p. 350-9.
243. Nickell, S.P., et al., *Inhibition by cyclosporin A of rodent malaria in vivo and human malaria in vitro.* Infect Immun, 1982. **37**(3): p. 1093-100.
244. Nagaraj, G., et al., *Antimalarial activities of peptide antibiotics isolated from fungi.* Antimicrob Agents Chemother, 2001. **45**(1): p. 145-9.
245. Turani, F., et al., *[Validity of indirect calorimetry in critical patients undergoing mechanical ventilation].* Minerva Anestesiol, 1991. **57**(10): p. 880-9.
246. Gumila, C., et al., *Characterization of the potent in vitro and in vivo antimalarial activities of ionophore compounds.* Antimicrob Agents Chemother, 1997. **41**(3): p. 523-9.
247. Aminake, M.N., et al., *Thiostrepton and derivatives exhibit antimalarial and gametocytocidal activity by dually targeting parasite proteasome and apicoplast.* Antimicrob Agents Chemother, 2011. **55**(4): p. 1338-48.
248. Sinnis, P., et al., *Remnant lipoproteins inhibit malaria sporozoite invasion of hepatocytes.* J Exp Med, 1996. **184**(3): p. 945-54.
249. Maciel, C., et al., *Anti-plasmodium activity of angiotensin II and related synthetic peptides.* PLoS One, 2008. **3**(9): p. e3296.
250. Kuntal, B.K., et al., *EasyModeller: A graphical interface to MODELLER.* BMC Res Notes, 2010. **3**: p. 226.
251. Manohar, S., et al., *Synthesis, antimalarial activity and cytotoxic potential of new monocarbonyl analogues of curcumin.* Bioorg Med Chem Lett, 2013. **23**(1): p. 112-6.
252. Morris, G.M., et al., *AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.* J Comput Chem, 2009. **30**(16): p. 2785-91.
253. Abraham, J.D., et al., *Cerebrospinal Abeta11-x and 17-x levels as indicators of mild cognitive impairment and patients' stratification in Alzheimer's disease.* Transl Psychiatry, 2013. **3**: p. e281.
254. Polimeni, M. and M. Prato, *Host matrix metalloproteinases in cerebral malaria: new kids on the block against blood-brain barrier integrity?* Fluids Barriers CNS, 2014. **11**(1): p. 1.
255. Prato, M., et al., *Phagocytosis of hemozoin enhances matrix metalloproteinase-9 activity and TNF-alpha production in human monocytes: role of matrix metalloproteinases in the pathogenesis of falciparum malaria.* J Immunol, 2005. **175**(10): p. 6436-42.

List of Publications

- **Kimjolly Lhouvum**, Vibin Ramakrishnan, and Vishal Trivedi, *Insight into structural and biochemical determinants of substrate specificity of PFI1625c: correlation analysis of protein-peptide molecular models.* J Mol Graph Model, 2013. **43**: p. 21-30.
- Swagata Nag, Devendra Kumar Chouhan, S. N. Balaji, Arnish Chakraborty, **Kimjolly Lhouvum**, Chandralata Bal, Ashoke Sharon, Vishal Trivedi, *Comprehensive screening of heterocyclic compound libraries to identify novel inhibitors for PfRIO-2 kinase through docking and substrate competition studies*. Med Chem Res (2013) 22:4737–4744. DOI 10.1007/s00044-013-0483-x.
- **Kimjolly Lhouvum**, Bhayur Kshitij Shekhar, Vishal Trivedi. *Molecular modeling and correlation analysis of PFI1625c-peptide models explain anti-malarial mystery of bioactive peptides*. Med Chem Res. (Communicated).
- **Kimjolly Lhouvum**, SN Balaji, Vishal Trivedi. Irreversible inhibition of Cyclosporin A on PFI1625c inhibition; Insight into action of inhibition. (Manuscript in preparation).
- **Kimjolly Lhouvum**, SN Balaji, Vishal Trivedi. Biochemical characterization of PFI1625c from *P.falciparum* 3D7; Probing of antimalarial drugs against PFI1625c. (Manuscript in preparation).

List of Conferences

- **Kimjolly Lhouvum** and Vishal Trivedi. Potentials of structural and biochemical characterization of PFI1625c from *Plasmodium falciparum* in drug development against malaria. 80th Annaual Meeting of the Society of Biological Chemist (India). CSIR-CIMAP, Lucknow 12-15, November 2011.
- **Kimjolly Lhouvum**. International Symposium on “BIOENGINEERING 2012” (ISBE 2012) organized by Biotech Hub, Centre for the Environment, Indian Institute of Technology Guwahati held on December 10, 2012.
- Rohitas Deshmukh, Arnish Chakraborty, **Kimjolly Lhouvum**, S.N. Balaji, Sourav Layek and Vishal Trivedi. Methemoglobin mediated immuno-toxicity modulation of pro-oxidant molecules towards macrophages during malaria. 82th Annual Meeting of The Society of Biological Chemists (India) held At School of Life Sciences, University of Hyderabad, India 2-5 Dec 2013.
- Swagata Nag, KMN Prasad, Arnish C., Rohitas Deshmukh , **Kimjolly Lhouvum,** Suman Joyti Deka, Ananya Bhowmick, S.N. Balaji, and Vishal Trivedi. Comprehensive screening of heterocyclic phytochemical and kinase database against P*f* RIO-2 kinase from plasmodium *falciparum*. 81st Annual Meeting of Society of Biological Chemists, India held at, Kolkata, India, 8-11 Nov, 2012.

Journal of Molecular Graphics and Modelling 43 (2013) 21–30

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Insight into structural and biochemical determinants of substrate specificity of PFI1625c: Correlation analysis of protein-peptide molecular models

Kimjolly Lhouvum[a], Vibin Ramakrishnan[b], Vishal Trivedi[a,*]

[a] Malaria Research Group, Department of Biotechnology, Indian Institute of Technology-Guwahati, Guwahati 781039, Assam, India
[b] Molecular Informatics & Design Laboratory, Department of Biotechnology, Indian Institute of Technology-Guwahati, Guwahati 781039, Assam, India

ARTICLE INFO

Article history:
Accepted 28 March 2013
Available online 8 April 2013

Keywords:
Malaria
Drugs
Resistance
Modeling
Peptides
Protein sorting
Metalloprotease
Mitochondria
Apicoplast

ABSTRACT

Bioinformatics and sequence comparison indicate PFI1625c as a putative metalloprotease present in plasmodium genome. The structure of PFI1625c consists of two domains with nearly identical folding topology. The active site of PFI1625c is located in a large central cavity between the two domains. Substrate binding regions of PFI1625c are lined by E-136, D-140 which provides negatively charged patches whereas F-53 facilitates binding of bulky hydrophobic residues of substrates. Probing PFI1625c active site with 199 different peptides from a combinatorial peptide library indicates preference of PFI1626c toward hydrophobic residue substituted peptides. Correlation analysis of each position of the peptide indicates that Ser 2 is the most crucial residue and no significant improvement was observed until it is mutated to a hydrophobic residue. The peptide P550 (LVIVAKRA) exhibits significantly better interaction within the active site than a template peptide (LSRVAKRA). The molecular dynamic's simulation studies confirms integrity of the complex, with all structures well within the qualitative limit of compactness and stability during the simulation time. There are structural and biochemical differences between PFI1625c with human metalloprotease and these are sufficient enough to allow us to exploit PFI1625c as drug targets. These computationally obtained insights provided clues about substrate selectivity in PFI1625c and it can be used to exploit PFI1625c as a target for future anti-malarial development.

1. Introduction

Malaria causes ~1 million deaths every year and with the development of drug-resistant parasite strains situation is quite alarming in tropical and sub-tropical area [1,2]. Malarial parasite depends extensively on proteolysis events for survival throughout its life cycle [3]. One of the most conspicuous events of the invasion is the removal or "shedding" of the fuzzy coat proteins covering the merozoite surface by serine protease SUB2 [4]. This process is also aided by rhomboid protease [5]. Serine protease gp76 degrades band 3 and glycophorin A to weaken the RBC cytoskeleton to facilitate a merozoite invasion [6]. Hemoglobin digestion, the most important metabolic pathway is completely controlled by co-ordinated action of different proteolytic events inside the acidic food vacuole of parasite [7–11].

Comparative Genome Analysis has identified 92 putative proteases in *P. falciparum* genome [12]. Potentially important proteases which remain uncharacterized consist of the calpain type proteases, metacaspase, primary processing proteases or signal peptidase [12–14]. In *P. falciparum*, a calpain, yet unidentified, was believed to be essential in merozoite invasion, based on the observation that calpain inhibitors I and II strongly blocked invasion [15]. Parasite exports proteins to different organelles to carry out essential functions. Transport of proteins requires processing of signal sequence by signal peptidase for delivery of protein to reach their destinations [16]. Transport of proteins outside the parasite interacts with proteins of the erythrocyte membrane and remodels the host cell morphologically to induce disease associated pathology, responsible for deaths during malaria [17,18]. Proteins targeted to mitochondria and apicoplast require pre-processing of nuclear coded proteins by mitochondrial or apicoplast processing peptidases or signal peptidases [19]. These organelles are prokaryotic in origin and were endo-symbiotic within the parasite body to perform essential functions for parasite growth. A number of proteins transported into mitochondria and apicoplast play an important role in the invasion of merozoite to start the new infection cycle in the RBC [20]. Therefore current efforts have been made to unveil the signal processing machinery of the mitochondria/apicoplast for identification of a new drug target [21]. PFI1625c is present in the *P. falciparum* genome (http://www.plasmodb.org) and is

* Corresponding author. Tel.: +91 361 2582217; fax: +91 361 258 2249.
E-mail addresses: vtrivedi@iitg.ernet.in, Vishalash_1999@yahoo.com (V. Trivedi).

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.03.008

proposed to play a role as a putative organelle processing peptidase. The constitutive expression of PFI1625c was found throughout the erythrocytic stages of the parasite life cycle [22]. In the current study an effort was made to perform an in-depth structural and functional characterization of PFI1625c. Multiple sequence alignment analysis and homology modeling provide evidence that PFI1625c belongs to metallo-protease M16 family. A combinatorial peptide library was prepared based on the signal peptide sequence LSRVAKRA present in yeast malate dehydrogenase (a template peptide used in the current study to model PFI1625c 3-D structure). Analysis of the molecular model of PFI1625c-peptide indicates a preference of PFI1625c toward hydrophobic groups substituted peptides. PFI1625c is evolutionarily very distant from human matrix processing peptidase. Hence our structural and peptide modeling studies characterize the substrate preference of PFI1625c and highlight its importance as an excellent drug target for anti-malarial drug development due to no potential cross reactivity with host proteases.

2. Materials and methods

2.1. Sequence analysis and alignment

Amino Acid Sequence of bc1 core protein of chicken, bovine, yeast mitochondrial processing peptidase and bacillus peptidase was retrieved from NCBI database and multiple sequence alignment was performed using Clustal W 2.0.11 [23]. A phylogenetic tree was generated by blasting PFI1625c into NCBI non-redundant protein database (nr) using the neighbor joining method of Clustal W2 phylogeny program [24].

2.2. Molecular modeling and structure validation

The amino acid sequence of PFI1625c was retrieved from the Plasmodium Database (http://www.plasmodb.org). PFI1625c 3D structure modeling was performed in the following sequential steps: template selection, sequence-template alignment, model building, refinement and validation. The suitable template was identified by searching PFI1625c into the protein data bank of NCBI using PSI-blast (http://blast.ncbi.nlm.nih.gov/Blast). The top hits were further analyzed for least number of gaps and highest sequence identity. Crystal structure of mitochondrial processing peptidase from yeast (PDB code 1HR6, β-subunit) was found to be suitable [25] with 38% identity over 435 residues. The 3D model of PFI1625c was generated using MODELLER 9v9 [26]. Auto model module of modeler 9v9 was used to generate 100 initial models with a ga431 score of 1 and these models were ranked based on their DOPE scores and mol pdf. Top ten models having the lowest DOPE scores and mol pdf were selected. The stereo-chemical quality of each model was confirmed by PROCHECK and ramchandran plot [27,28]. The statistics of non-bonded interactions between different atom types were analyzed by ERRAT program which gives a measure of the structural error at each residue in the protein [29]. The compatibility of the atomic model (3D) with its own amino acid sequence was determined by Verify_3D [30]. The model with the least number of residues in the disallowed region was selected and energy minimized using Steepest Descent (SD) algorithm with GROMOS96 43a1 force fields in GROMACS 4.0.7 package (http://www.gromacs.org/). This process was repeated until most of the residues are not below 95% in ERRAT plot. The final model was checked by Verify-3D, Procheck and Ramchandran plot. The PROCHECK and ramchandran plot analysis indicate that final model is of good quality with more than 95% residues in the allowed region. The ERRAT plot gives most of the residues below 95% with overall quality factor of 90.071. Overall no short contact was observed in the final model (Fig. S1).

2.3. Generation of combinatorial peptide library

Signal peptide (LSRVAKRA) of malate dehydrogenase shows binding in PFI1625c in a similar conformation as present in the crystal structure of yeast mitochondrial processing protease (PDB Code 1HR9). The signal peptide of malate dehydrogenase (LSRVAKRA) was used as a template to prepare a peptide library to optimize the binding of peptide into the PFI1625c active site. A random single or double substitution into the template peptide sequence (LSRVAKRA) was done to prepare a combinatorial peptide library of 733 peptides representing different types of probable substitution (Table S1). The type of substitution was hydrophobic, neutral and long chain hydrophobic and charged residues. After critically analyzing a complete peptide library and excluding similar substitution with weak interactions, 199 peptides were used for molecular modeling. 3-D structure modeling of peptide sequence from the library was done with the help of EasyModeller 2.0 [31]. The structural quality of each modeled peptide was checked by Procheck and Ramchandran plot and most of the peptide models were of high quality.

2.4. Molecular Modeling of peptides from the combinatorial peptide library into PFI1625c active site

The individual peptide from the combinatorial peptide library was fitted into the PFI1625c 3-D structure using patchdock [32]. 'Fire-Dock' program is used initially to perform energy minimization of best 20 models to get stable PFI1625c-signal peptide complexes [33]. Atomic contact energy (ACE) and local proximity to the predicted binding sites in a PFI1625c active site of each conformation was analyzed. Control Docking of signal peptide (LSRVAKRA) in yeast mitochondrial processing peptidase (PDB Code 1HR9) using patchdock shows, peptide binds in the extended conformation as given in co-crystal structure [25]. The root mean square deviation between generated model and crystallographic complex (1HR9) is 1.52 Å. To validate the approach, docking was also performed by taking matrilysin and its peptide substrates from MEROPS database [34] and through literature search [35]. The docking experiments showed a strong correlation ($r = 0.941$) between the affinity of the peptide substrate (Km) to the atomic contact energy (Fig. S2). Correlation of experimentally determined biochemical data with docking score proves the reliability of using Patch-dock to generate PFI1625c-peptide molecular models and correlates the docking score with suitability of the bound peptide as potential substrate.

2.5. Molecular dynamics simulation

Molecular dynamics (MD) simulations and analysis were performed in an Intel Xeon Workstation using GROMACS program suite [36] with GROMOS96 43A1 force field. The dynamics of PFI1625C-peptide complex were studied at 300 K under NVT conditions with periodic boundary. Each of the ten complexes was energy minimized in vacuum using steepest descent algorithm in 500 steps. Solvent molecules (SPC water) were then added to the rectangular box with the distance between solute and the box set at 1 nm. The system was once again energy minimized in water before 1 ns production run with an integration step of 2 fs. Throughout the simulations, Bond lengths were constrained with Shake algorithm to a geometric accuracy of 10^{-4}. Initial velocities were taken from a Maxwellian distribution at the chosen temperature of 300 K temperature bath with a coupling relaxation time of 0.1 ps. The non bonded list cutoff of 1.4 nm was used. The nonbonded interactions

(VdW and electrostatics) were shifted to zero between 0.8 nm to 1.1 nm cutoff.

3. Result

3.1. PFI1625c is a metalloprotease

Multiple sequence alignment implicates that PFI1625c belongs to the metalloprotease family. The characteristic conserved sequences pattern present in PFI1625c matches with other known metalloprotease (Fig. 1). To further validate PFI1625c as metalloprotease, it was analyzed by ProtIdent web server [37] which is developed by fusing the functional domain and sequential evolution information. The first layer identifies the query protein as protease or nonprotease while the second layer identifies the protease class. ProIdent server indicates that PFI1625c is a protease belonging to the metalloprotease family. A characteristic zinc binding motif and active site residues specific for a metalloprotease class are also present in PFI1625c. Phylogenetic analysis of PFI1625c against non-redundant database shows that PFI1625c forms a separate cluster with putative proteases of other apicomplexan species and it is evolving very distantly from human proteases (Fig. 2A).

3.2. The overall structure of PFI1625c

The complete structure of PFI1625c can be divided into domains I and II. There is a sequence difference between two domains but the α-helix and β-sheet folds in each domain to give similar folding topology of a non-superimposable image (Fig. 2B). A loop of 25 residues connects α-helix 11 of a domain I and β-sheet 7 of domain II. The α-helices and β-sheets are equally distributed in both the two domains. β-Sheets of each domain are projected into the interior core of the PFI1625c and consist of two hairpin structures and Ψ-loop motif. The Ψ loop structures from each domain are composed of two anti parallel β strands. The α-helices are present toward the exterior portion of the PFI1625c to shield the interior β sheets (Fig. 2C). These arrangements give rise to a complete protease structure with an organized active site cavity and intact catalytic site present in other metalloproteases.

3.3. Zinc binding site

A characteristic feature of metalloprotease is the presence of the metal binding motif [38]. Multiple Sequence alignment indicates that a conserved zinc binding motif (HXXEH) with H-46, E-49, H-50 residues are present in PFI1625c where zinc co-ordinates with E-126 (Fig. 3A). The structure of the protein implies that α-helix 6 and 2 from a domain I come closer to each other with a unique folding pattern to align the zinc binding residues (H-46, E-49, H-50) to form a catalytic triad (Fig. 3A and B). The E-126, another zinc binding residue is also coming closer with the H-46 and H-50 to form the perfect geometry that fits the zinc metal.

3.4. Substrate binding site

Substrate binding region of PFI1625c is constituted by E-136, D-140 and F-53 (Fig. 3B). The active site is present inside a well organized large cavity in the center of the structure. The active site cavity closer to domain I is negatively charged whereas rest cavity is of mixed environment (Fig. 3C). The domain I region of the active site is negatively charged to facilitate binding of positively charged residues of the substrate while the remaining portion of the cavity consists of mixed charges with a hydrophobic environment toward the interior to provide docking sites for hydrophobic residues of a peptide substrate. The cavity extends toward the domain II which may assist in binding the large substrates to allow access of the cleavage site by the catalytic residues present in the domain I (Fig. 3D). Substrate binding scaffolds comprised of 2 β-sheets, contributed from both domains. The β-sheet from domain I has conserved HLNAY residues while the β-sheet from domain II has SFNTCY residues.

Table 1
Analysis of top hit peptides from combinatorial peptide library.

Peptide code	Amino acid sequence	Atomic contact energy (kcal/mol)
Template	LSRVAKRA	−69.85
P548	LVDVAKRA	−262.81
P550	LVIVAKRA	−262.47
P543	LLIVAKRA	−220.83
P536	LIIVAKRA	−144.42
P540	LLKVAKRA	−174.58
P33	LSVVAKRA	−114.1
P32	LSLVAKRA	−100.12
P31	LSIVAKRA	−98.26
P168	ASIVAKRA	−82.26
P545	LLVVAKRA	−73.05
P537	LILVAKRA	−57.47
P544	LLLVAKRA	−45.29
P546	LVHVAKRA	−33.53
P683	LSIVHKRA	−31.82
P170	ASVVAKRA	−23.61
P551	LVLVAKRA	−22.01
P538	LIVVAKRA	−13.35

A random combinatorial peptide library was prepared from template peptide LSRVAKRA present in co-crystallized structure of mitochondrial processing peptidase (PDB code 1HR9). 3-D model of peptides were prepared as described in Section 2. The peptides from combinatorial peptide library were modeled into the PFI1625c active site using patch-dock and atomic contact energy was calculated. The peptides with hydrophobic substitutions fit well into the PFI1625c active site.

Table 2
Comparison of peptide interaction within the active site of PFI1625c.

Template peptide	Protein residue	Atom	Distance (Å)
L1	M264	CB	2.04
L1	M264	CD2	2.46
S2	Q267	NE2	2.98
S2	N310	H	1.82
S2	N310	O	2.9
R3	Q363	NE2	4.88
V4	R81	N	4.27
A5	R81	NE	4.34
K6	Y78	OH	2.34
R7	T79	O	3.47
R7	E136	OE2	1.64
A8	H46	HD1	2.5
A8	E49	OE2	2.34
Peptide P550	**Protein residue**	**Atom**	**Distance (Å)**
L1	Q267	OE1	3.18
L1	L364	CD2	2.61
L1	L364	CD1	2.84
L1	M264	CE	2.01
V2	C312	CB	2.25
V2	C312	N	2.55
I3	C312	N	3.07
I3	N310	O	2.42
V4	C312	O	2.93
V4	Y313	CE1	3.34
A5	R81	NE	3.06
K6	Y78	CE1	3.71
R7	E136	OE2	1.29
A8	E126	OE2	2.15
A8	H46	NE2	2.19

Interactions of the peptide with active site protein residues in PFI1625c-template peptide and PFI1625c-peptide P550 molecular model. The optimized peptide sequence P550 fits well into the PFI1625c active site as evidenced from a number of strong interactions as compared to template peptide.

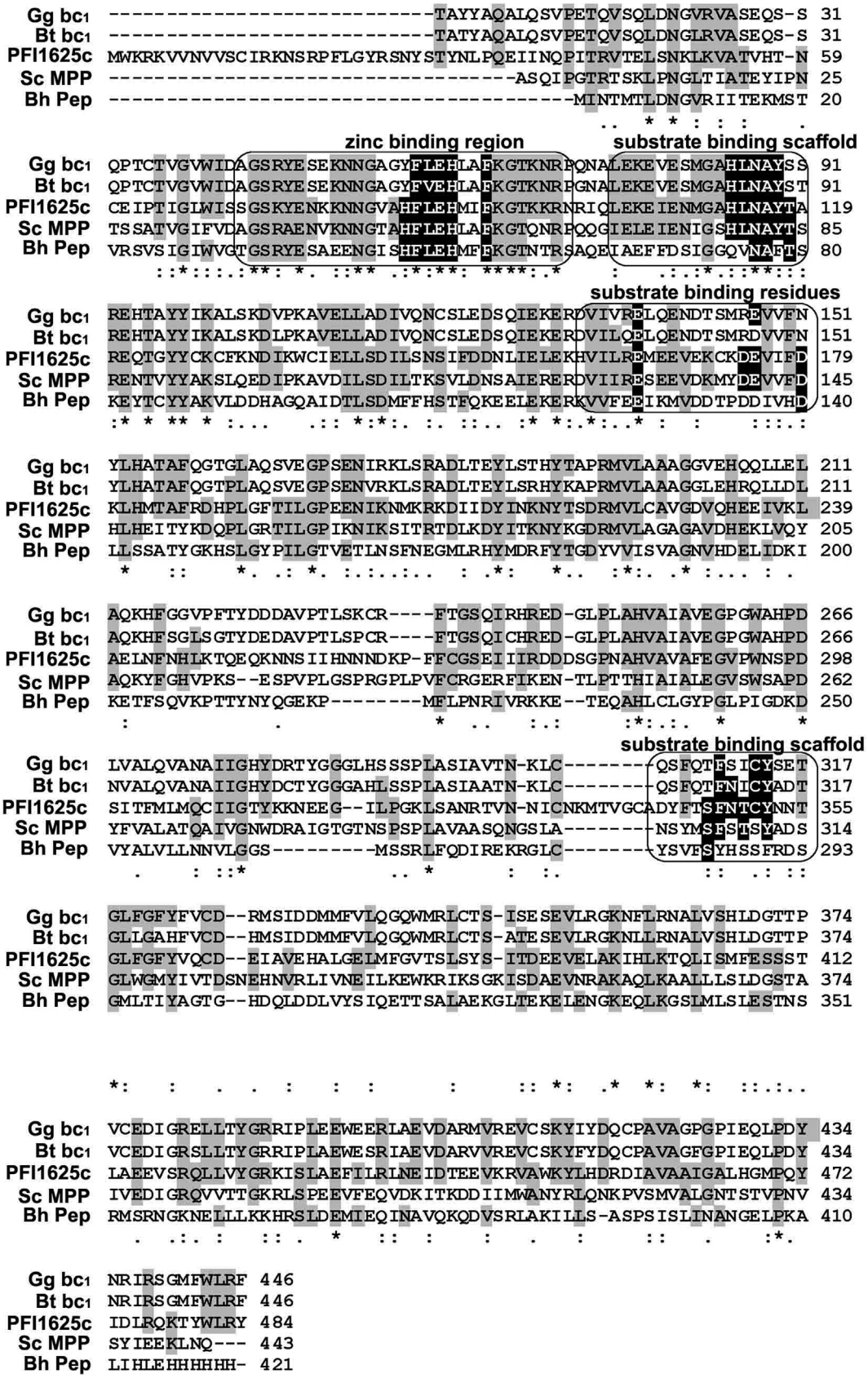


Fig. 1. Characterization of different structural elements in PFI1625c. Multiple sequence alignments of PFI1625c from *P. falciparum* with bc1 core protein of *G. gallus* (Gg bc1), *B. taurus* (Bt bc1), metalloprotease of *S. cerevisiae* (Sc MPP) and protease from *B. halodurans* (Bh Pep). Similar and identical residues are indicated by standard annotation. The zinc binding region and substrate binding scaffold are enclosed within the rectangular box to highlight the presence of these regions in PFI1625c.

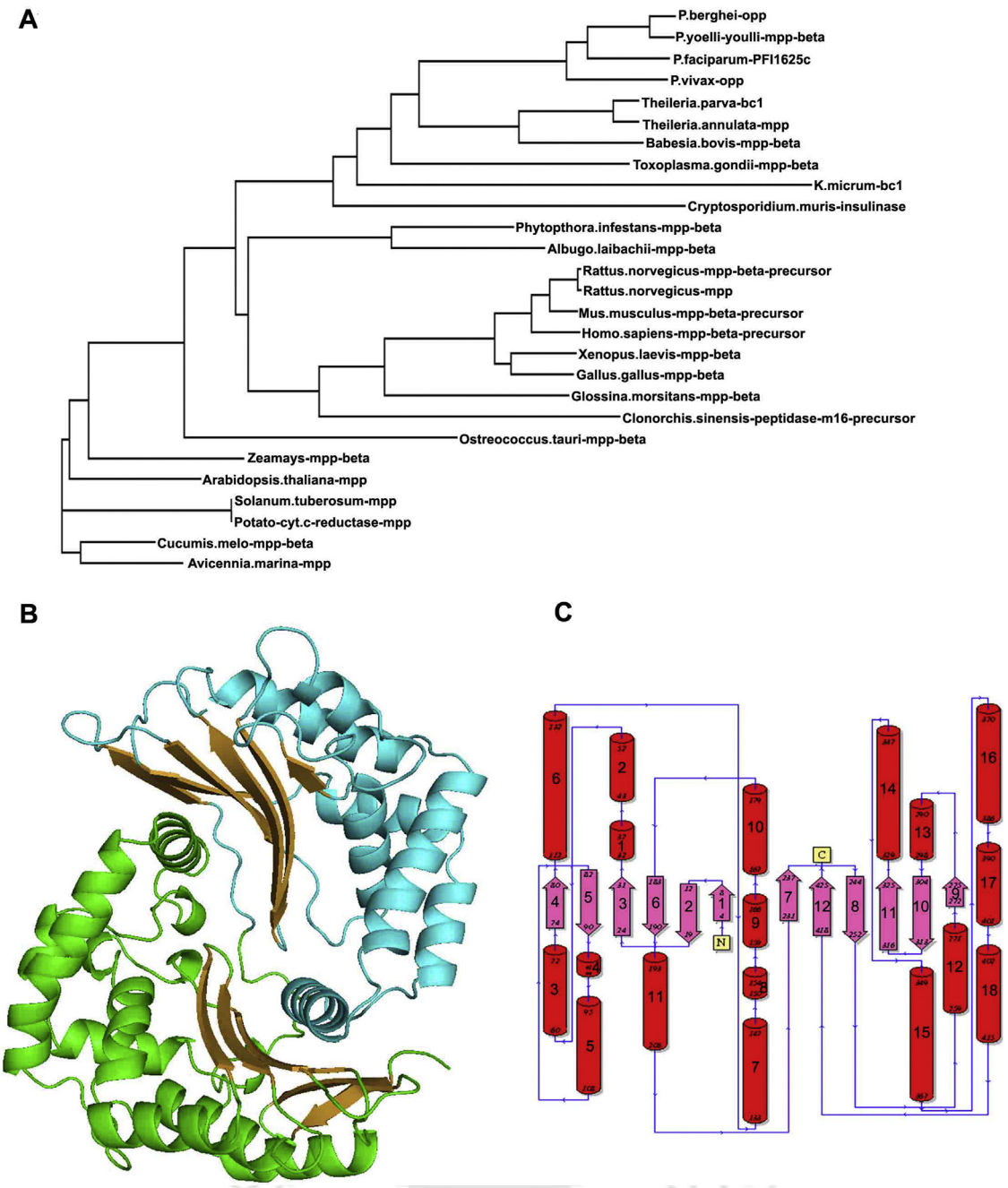


Fig. 2. Overall structure of PFI1625c. (A) Evolutionary relationship of PFI1625c with proteases present in other organism. Human metalloprotease is evolutionarily diverse from PFI1625c and form a distinct cluster. (B) 3-D structure of PFI1625c consists of two domains. Domain I is colored green and domain 2 is colored cyan. The β-sheets present in respective domains are colored in brown. (C) Topology diagram of PFI1625c to depict the organization of structural elements. The α-helices are colored red while β-sheets are colored pink and structural elements are connected by blue colored lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.5. PFI1625c prefers hydrophobic residues in substrate

The interior of the substrate binding cavity is more hydrophobic in PFI1625c indicating that binding of hydrophobic residues in the cavity will be energetically more favorable than binding a hydrophilic residue. A template peptide (LSRVAKRA) binds PFI1625c with an ACE of −69.85 (Table 1). It is present in an extended conformation as predicted in the co-crystal structure of mitochondrial processing peptidase bound to signal peptide (PDB code 1HR9). The LSRVA occupy the hydrophobic region of the active site pocket while KRA occupy the more negatively charged region (Fig. 4A). In the catalytic site, a template peptide makes polar contact, hydrogen bonding, van der waals with the residues surrounding substrate binding pocket (Fig. 4B). In the present binding mode residues of template peptide makes either weak interactions or no interaction at all (Table 1). A combinatorial peptide library of ∼733 peptides based on the peptide sequence (LSRVAKRA) was generated to represent all possible substitutions to map active site environments (Table S1). Individual peptides from the combinatorial peptide library (∼199 peptides) were modeled into the binding pocket to study the peptide substrate interaction and binding within binding pocket. An in-depth interactions

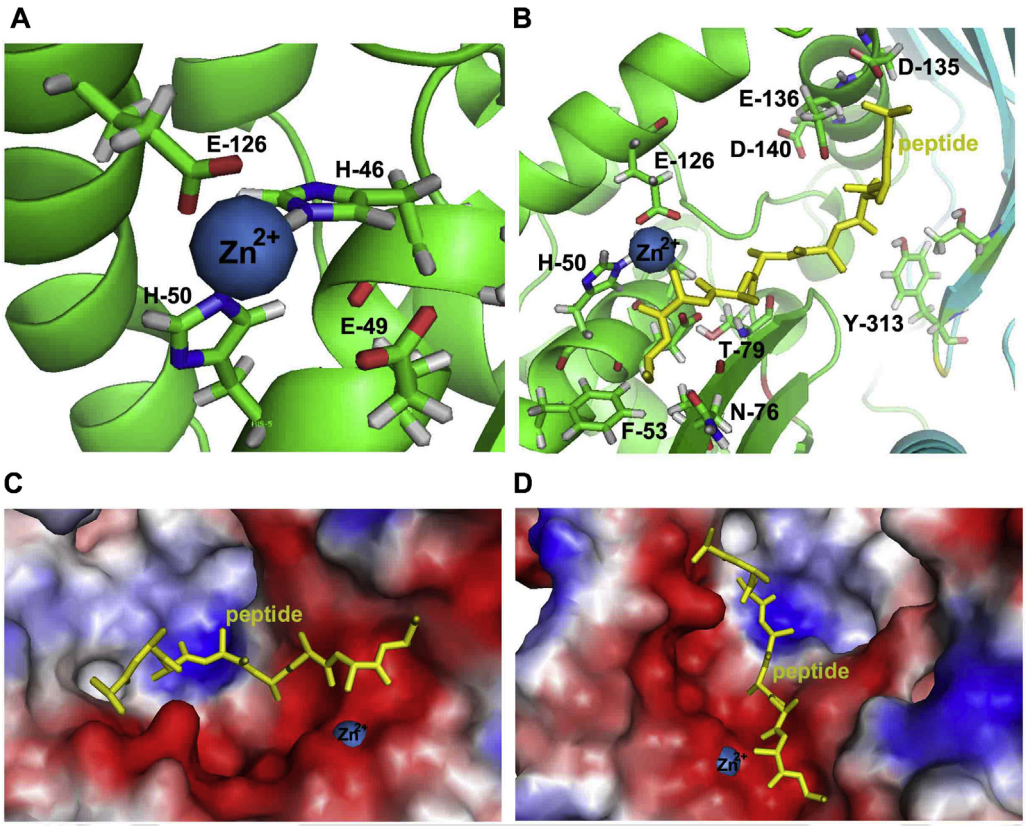


Fig. 3. PFI1625c is a metalloprotease with a well defined active site. (A) Zinc binding motif of PFI1625c present in the domain I. The zinc ion is shown in blue and the residues interacting with Zn^{2+} to stabilize it within the catalytic site are shown with corresponding label. (B) Active site of PFI1625c with bound peptide substrate. The template peptide (LSRVAKRA) was taken from yeast mitochondrial processing peptidase (PDB ID 1HR9) and modeled into the active site. The residues interacting with peptide substrate are labeled and shown. PFI1625c surface representation in (C) horizontal and (D) vertical view with bound template peptide in yellow. Red represents −ve charge, blue represents +ve charge. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

analysis indicates a number of potential peptide substrates with significantly high binding score and ACE values (Table 1).

Correlating the substitution at each position in a template peptide with other amino acid residues allow us to understand the catalytic preference of PFI1625c for substrate peptide. 17 peptides with similar group substitutions at position 1 or 2 did not give better interaction than template peptide except peptide P91 where Leu 1 is replaced by Ile and Ser 2 by Thr (Table S2). Substitutions at 1, 2 or 3 positions with acidic amino acid alone didn't show any improvement but substitution with acidic and hydrophobic amino acid has improved the peptide binding than a template peptide. 48 peptides with amino acid substitution at position 1 or 2 such as in peptide P152 where Leu 1 is substituted with Trp and Ser 2 is replaced with His provide better interaction with ACE of −114.38 (Table S2). Out of 21 peptides with hydrophobic substitution, 8 peptides gave significantly better interaction than a template peptide. The overall result showed that hydrophobic amino acid substituted peptides have the better interaction than a template peptide. The best interaction was found when the residue Ser 2 is replaced by Leu, Ile or Val and Arg 3 are replaced by Ile (Table 1). Correlation analysis of each position of the peptide indicates that Ser 2 is the most crucial residue and no significant improvement was observed until it is mutated to a hydrophobic residue. Peptide P550 (LVIVAKRA) fits well and shows the best interaction within the catalytic site of PFI1625c. In the catalytic site, P550 makes polar contact, hydrogen bonding, van der waals with the residues surrounding substrate binding pocket (Fig. 4C). In the present binding mode residues of P550 make several strong interactions as well as Π-Π stacking interactions with the hydrophobic groups present in binding pocket and hydrophobic groups substituted in P550 (Fig. 4D). In comparison to a template peptide, P550 makes several and very strong interactions within the binding pocket of PFI1625c (Table 2). Overall analyses of different peptide substrates indicate that peptide with hydrophobic substitution increases the interaction between peptide and active site residues especially residues present within a hydrophobic pocket of the active site.

3.6. Molecular dynamics analysis of PFI1625c-peptide complexes

Three programs *g_gyrate*, *g_hbond* and *g_rms* in GROMACS have been used for analysis. "Compactness" of the protein throughout the simulation was monitored using *g_gyrate*, which measures the radius of gyration. Radius of gyration values further underline the compactness of the complex with all structures with almost all complexes showing downward trend, with a negative slope (Fig. 5A). P31 and p543 were the most compact ones among the ensemble of structures sampled during 1 ns simulation whereas P168 was found to be comparatively loose. Root mean square deviation from starting structure was monitored using g_rms and hydrogen bonding interaction between the peptide and the receptor was mapped with the help of g_bond program. RMS deviation from the starting structure of the complex molecule were observed to be in the range of 0.2–0.3 nm with P168 being closest to the starting structure and P31 deviating maximum in one ns trajectory (Fig. 5B). Overall, simulation results underline integrity of the complex, with all structures well within the qualitative limit of compactness and stability during the simulation time.

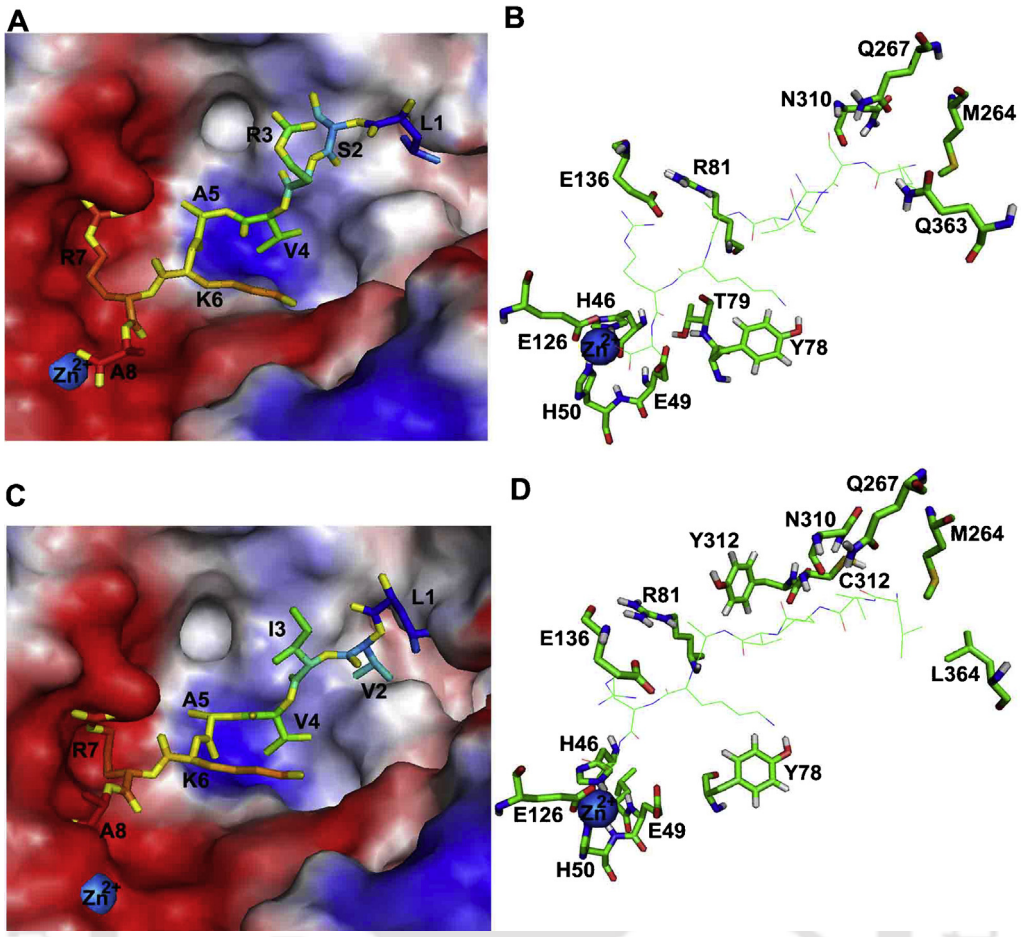


Fig. 4. Correlation analysis of PFI1625c-peptide molecular model. (A) PFI1625c-template peptide model and (C) PFI1625c-P550 peptide model binds peptide in an extended conformation to the large surface area within the active site of PFI1625c. P550 shows better fitting into the active site exploiting opposite charge pockets and hydrophobic patches. Red represents −ve charge, blue represents +ve charge. Interaction of (B) template peptide and (D) P550 with the residues present within the active site of PFI1625c.

Simulation results of top ten structures hence confirm the stability of interaction, though the simulation time may not be sufficient to create a large enough conformational ensemble to rank them accordingly.

3.7. *PFI1625c is a candidate drug target*

Blasting PFI1625c against non-redundant database shows that PFI1625c forms a separate cluster and it is evolving very distant from host proteases. The closest human metalloprotease (mpp β precursor) has diversified into a distinct group from the PFI1625c cluster (Fig. 2A). The closest available 3-D structure of human metalloprotease to PFI1625c is insulin degrading enzyme (IDE) and it is used to test the suitability of PFI1625c as a drug target. PFI1625c and IDE Exhibit 22% sequence identity with conserved catalytic residues but both are very distinct at structural level. The active site cavity of human IDE is buried inside a closed pocket and entry of a peptide is restricted and substrates need to enter through a

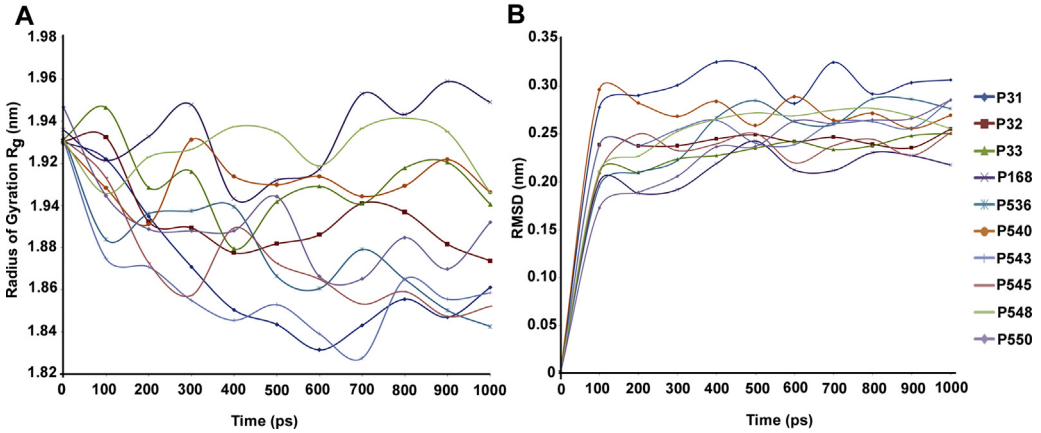


Fig. 5. Molecular dynamics simulation of top ten PFI1625c-peptide complexes. (A) Radius of Gyration and (B) RMSD of all complexes. MD analysis shows that the PFI1625c-peptide complexes are energetically stable under the simulation conditions.

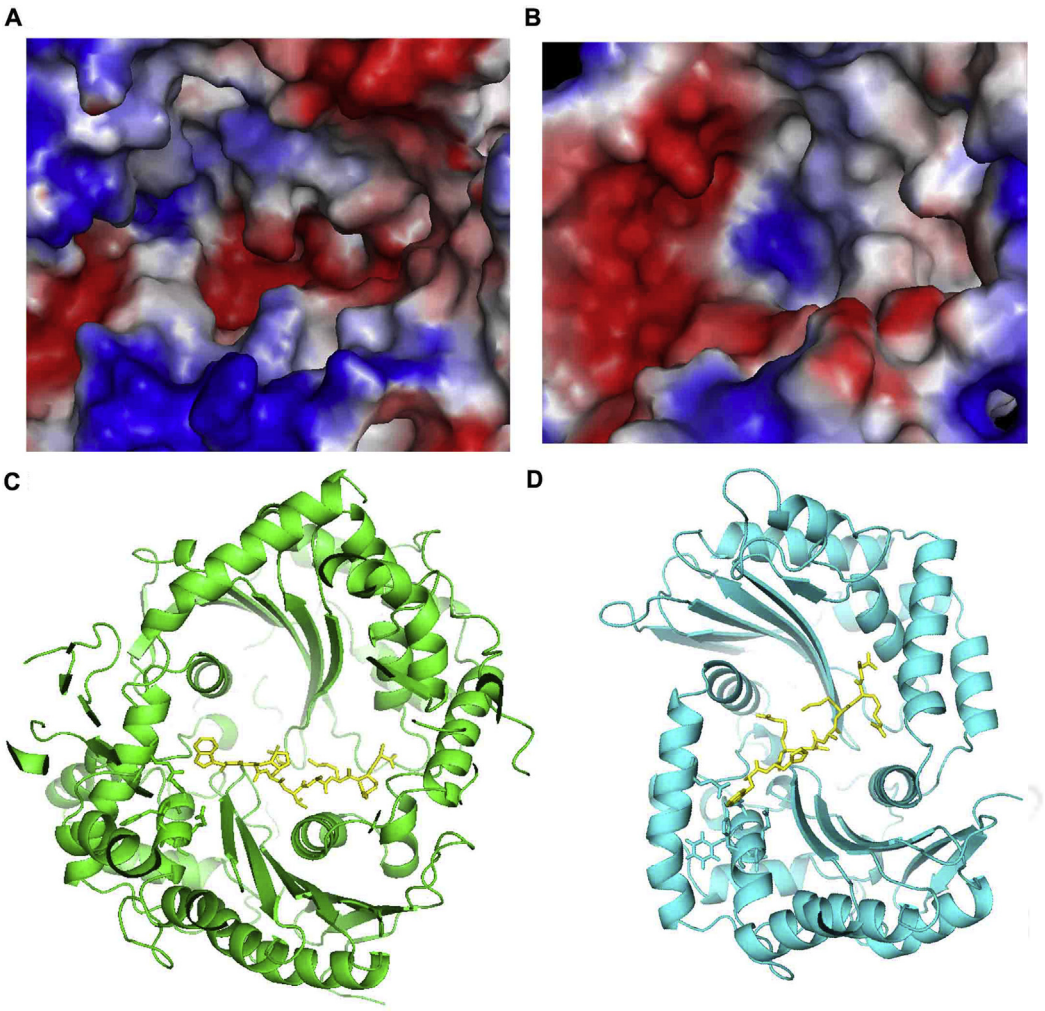


Fig. 6. The surface structure of the active site and interaction of bound peptide within human insulin degrading enzyme (IDE) and modeled structure of PFI1625c. (A) and (B) Surface Structure of the peptide binding pocket of IDE and PFI1625c-peptide model. Blue is positive, red is negative and peptide (yellow) is present in an extended conformation. (C) and (D) Interaction of bound peptide with key residues present within the binding pocket of IDE and PFI1625c. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

narrow tunnel (Fig. 6A) where as PFI1625c is present in an open groove with complete access of an incoming peptide (Fig. 6B). The active site of both proteases also exhibits significant differences in active site environment and charge distribution. The catalytic region of IDE is more positively charged or neutral except the zinc binding motif (Fig. 6C), PFI1625c is highly negatively charged (Fig. 6D). The validation of PFI1625c as a drug target is further affirmed by comparative studies with similar and characterized metalloprotease group in *P. falciparum* such as falcilysin, PfM18AAP, Pf-AM1 and PfLAP. Their functions in the parasite biology are characterized and their effect in the growth of the parasite is studied. Falcilysin has a role in parasite development during asexual blood stages [39]. Absence of PfM18AAP is shown to impart lethality in parasites using knockdown experiments. Inhibition of Pf-AM1 and PfLAP with specific inhibitors also has antimalarial effect [40,41]. All these comparative studies strongly support our vision toward PFI1625c as a drug target. Hence, PFI1625c exhibits a number of differences from host protease in terms of active site environment, substrate specificity and these properties are sufficient to exploit PFI1625c as drug targets.

4. Discussion

Characterizing and validating proteins to examine their potential as a drug target is the need of the hour to develop novel malarial chemotherapy. Protein-peptide docking is promising ultimate tool to decipher substrate or ligand specificity. Computational approaches to determine specificity in a peptide substrate toward HIV-1 protease indicate a correlation between substrate specificity and its binding energy with the protease [42]. In another study, protein-protein complexes are used to identify the determinants that are crucial for complex formation [43].

Proteases played a significant role in the malaria parasite life cycle and required for survival throughout its life cycle. Proteases present inside the food vacuole (FV) are involved in hemoglobin digestion to provide amino acid for protein synthesis [44]. Besides basic metabolic pathways, protease in the parasite is known to regulate invasion, egress and other molecular events [8,11,16,20,44]. A recent Bioinformatics analysis explored 92 proteases within plasmodium genome and a number of these uncharacterized proteases have potential to be utilized as an excellent drug target [12]. Potentially important proteases which remain uncharacterized consist of the calpain type proteases, metacaspase, primary processing proteases or signal peptidase. PFI1625c, an organeller processing peptidase is proposed to play an important role in protein targeting. Transport of proteins requires processing of signal sequence by signal peptidase for delivery of protein to reach their destinations [16]. PFI1625c has high similarity with other known metalloproteases with a conserved metal binding motif HXXEH [38]. A typical

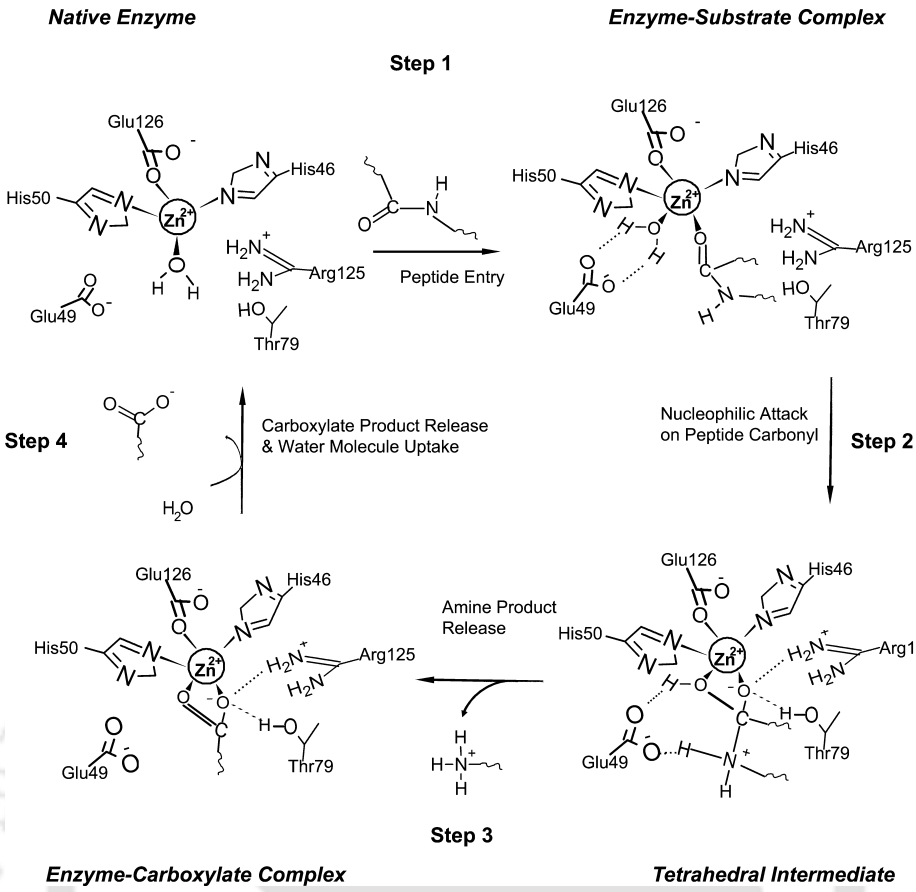


Fig. 7. Proposed catalytic mechanism of PFI1625c mediated peptide cleavage. Based on the stereospecific positions of different residues within the PFI1625c active site, a schematic cleavage mechanism is proposed.

Zn^{2+} dependent metalloprotease follows a mechanism involving different distinct reaction intermediate [45]. Structural determinants play a crucial role in substrate recognition and specificity [46].

Patch-dock was used as a tool to generate PFI1625c-peptide molecular models, as control docking of signal peptide (LSRVAKRA) in yeast mitochondrial processing peptidase (PDB Code 1HR9) gives complex (yeastMPP-peptide) close to known structure [25]. The root mean square deviation between generated model and crystallographic complex (1HR9) is 1.52 Å. To further validate the approach, docking was also performed by taking matrilysin and its peptide substrates from MEROPS database [34] and through literature search [35]. The docking experiments showed a strong correlation ($r = 0.941$) between the affinity of the peptide substrate (Km) to the atomic contact energy (Fig. S2). Correlation analysis of PFI1625c-peptide complexes indicates a preference of PFI1625c toward hydrophobic residues in the substrate due to the local environment. Protein-peptide docking is a powerful tool to predict the binding affinities, specificity and such a study has immense potential in drug discovery by utilizing bioactive peptides [47,48].

Based on the stereospecific positions of different residues within the PFI1625c active site, a schematic cleavage mechanism is given in Fig. 7. Comparing PFI1625c with metalloprotease shows that E-136 and D-140 present near the zinc binding site to form the S_2 and S_3 sites. The negatively charged sites form a salt-bridge with an arginine present in an incoming peptide to facilitate strong binding (Fig. 3). The F-53 lies at the S_1' site and is believed to interact with the aromatic residues found in the substrate. The residue E-49 which lies within hydrogen bonding distance to a water molecule coordinated to zinc, is predicted to polarize this water molecule, thereby aligning it for nucleophilic attack on the carbonyl carbon of the peptide bond of the substrate (Fig. 7). The active site of PFI1625c or other metalloproteases differs from thermolysin in the absence of analogous residues Y-157 and H-231 which are proposed to stabilize the oxyanion of a tetrahedral intermediate. Solvent molecules or similar charged amino acids T-79 and R-125 could be involved in hydrogen bonding to stabilize oxyanion of a tetrahedral intermediate in PFI1625c (Fig. 7).

All peptides fitting well into the PFI1625c active site and stereochemistry indicate that most likely cleavage site is a peptide bond between A8 and positively charged R7, as positive charge can be stabilized by negatively charged patch within the catalytic site. Substitution of amino acids in the peptide occupying the hydrophobic cavity with non-polar residues increases the affinity of bound peptides within the PFI1625c active site due to the presence of a hydrophobic pocket.

Acknowledgements

This work was partially supported by the Department of Biotechnology; Govt of India Grants (BT/PR13436/MED/12/450/2009) to V.T. KL acknowledges the financial support in the form of fellowship from Indian Institute of Technology-Guwahati, Assam, India.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2013.03.008.

References

[1] S.I. Hay, E.A. Okiro, P.W. Gething, A.P. Patil, A.J. Tatem, C.A. Guerra, et al., Estimating the global clinical burden of *Plasmodium falciparum* malaria in 2007, PLOS Medicine 7 (2010) e1000290.

[2] S. Mok, M. Imwong, M.J. Mackinnon, J. Sim, R. Ramadoss, P. Yi, et al., Artemisinin resistance in *Plasmodium falciparum* is associated with an altered temporal pattern of transcription, BMC Genomics 12 (2011) 391.

[3] H. Cai, R. Kuang, J. Gu, Y. Wang, Proteases in malaria parasites—a phylogenomic perspective, Curr Genomics 12 (2011) 417–427.

[4] P.K. Harris, S. Yeoh, A.R. Dluzewski, R.A. O'Donnell, C. Withers-Martinez, F. Hackett, et al., Molecular identification of a malaria merozoite surface sheddase, PLoS Pathogens 1 (2005) 241–251.

[5] S. Singh, M. Plassmeyer, D. Gaur, L.H. Miller, Mononeme: a new secretory organelle in *Plasmodium falciparum* merozoites identified by localization of rhomboid-1 protease, Proceedings of the National Academy of Science of the United States of America 104 (2007) 20043–20048.

[6] E. Roggwiller, M.E. Betoulle, T. Blisnick, C. Braun Breton, A role for erythrocyte band 3 degradation by the parasite gp76 serine protease in the formation of the parasitophorous vacuole during invasion of erythrocytes by *Plasmodium falciparum*, Molecular and Biochemical Parasitology 82 (1996) 13–24.

[7] M. Allary, J. Schrevel, I. Florent, Properties, stage-dependent expression and localization of *Plasmodium falciparum* M1 family zinc-aminopeptidase, Parasitology 125 (2002) 1–10.

[8] R. Banerjee, J. Liu, W. Beatty, L. Pelosof, M. Klemba, D.E. Goldberg, Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine, Proceedings of the National Academy of Science of the United States of America 99 (2002) 990–995.

[9] K.K. Eggleson, K.L. Duffin, D.E. Goldberg, Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*, Journal of Biological Chemistry 274 (1999) 32411–32417.

[10] M. Klemba, I. Gluzman, D.E. Goldberg, A Plasmodium falciparum dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation, Journal of Biological Chemistry 279 (2004) 43000–43007.

[11] P.S. Sijwali, P.J. Rosenthal, Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*, Proceedings of the National Academy of Science of the United States of America 101 (2004) 4384–4389.

[12] Y. Wu, X. Wang, X. Liu, Y. Wang, Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite, Genome Research 13 (2003) 601–616.

[13] P. Prabhu, V. Patravale, Novel targets for malaria therapy, Current Drug Targets 12 (2011) 2129–2143.

[14] B. Meslin, H. Zalila, N. Fasel, S. Picot, A.L. Bienvenu, Are protozoan metacaspases potential parasite killers? Parasites & Vectors 4 (2011) 26.

[15] P. Olaya, M. Wasserman, Effect of calpain inhibitors on the invasion of human erythrocytes by the parasite *Plasmodium falciparum*, Biochimica et Biophysica Acta 1096 (1991) 217–221.

[16] J.A. Boddey, A.N. Hodder, S. Gunther, P.R. Gilson, H. Patsiouras, E.A. Kapp, et al., An aspartyl protease directs malaria effector proteins to the host cell, Nature 463 (2010) 627–631.

[17] F.S. Boretti, P.W. Buehler, F. D'Agnillo, K. Kluge, T. Glaus, O.I. Butt, et al., Sequestration of extracellular hemoglobin within a haptoglobin complex decreases its hypertensive and oxidative effects in dogs and guinea pigs, Journal of Clinical Investigation 119 (2009) 2271–2280.

[18] A. Pamplona, T. Hanscheid, S. Epiphanio, M.M. Mota, A.M. Vigario, Cerebral malaria and the hemolysis/methemoglobin/heme hypothesis: shedding new light on an old disease, International Journal of Biochemistry & Cell Biology 41 (2009) 711–716.

[19] R. Tuteja, A. Pradhan, S. Sharma, *Plasmodium falciparum* signal peptidase is regulated by phosphorylation and required for intra-erythrocytic growth, Molecular and Biochemical Parasitology 157 (2008) 137–147.

[20] C.Y. He, M.K. Shaw, C.H. Pletcher, B. Striepen, L.G. Tilney, D.S. Roos, A plastid segregation defect in the protozoan parasite *Toxoplasma gondii*, EMBO Journal 20 (2001) 330–339.

[21] G.G. van Dooren, V. Su, M.C. D'Ombrain, G.I. McFadden, Processing of an apicoplast leader sequence in *Plasmodium falciparum* and the identification of a putative leader cleavage enzyme, Journal of Biological Chemistry 277 (2002) 23612–23619.

[22] Z. Bozdech, J. Zhu, M.P. Joachimiak, F.E. Cohen, B. Pulliam, J.L. DeRisi, Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray, Genome Biology 4 (2003) R9.

[23] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Research 22 (1994) 4673–4680.

[24] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Molecular Biology and Evolution 4 (1987) 406–425.

[25] A.B. Taylor, B.S. Smith, S. Kitada, K. Kojima, H. Miyaura, Z. Otwinowski, et al., Crystal structures of mitochondrial processing peptidase reveal the mode for specific cleavage of import signal sequences, Structure 9 (2001) 615–625.

[26] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, Journal of Molecular Biology 234 (1993) 779–815.

[27] G.N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, Stereochemistry of polypeptide chain configurations, Journal of Molecular Biology 7 (1963) 95–99.

[28] R.A. Laskowski, D.S. Moss, J.M. Thornton, Main-chain bond lengths and bond angles in protein structures, Journal of Molecular Biology 231 (1993) 1049–1067.

[29] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Science 2 (1993) 1511–1519.

[30] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83–85.

[31] Kuntal, B.K. Aparoy, P. Reddanna, P. EasyModeller., A graphical interface to MODELLER, BMC Research Notes 3 (2010) 226.

[32] Schneidman-Duhovny, D. Inbar, Y. Nussinov, R. Wolfson, H.J. PatchDock, SymmDock: servers for rigid and symmetric docking, Nucleic Acids Research 33 (2005) W363–W367.

[33] N. Andrusier, R. Nussinov, H.J. Wolfson, FireDock: fast interaction refinement in molecular docking, Proteins 69 (2007) 139–159.

[34] N.D. Rawlings, D.P. Tolle, A.J. Barrett, MEROPS: the peptidase database, Nucleic Acids Research 32 (2004) D160–D164.

[35] H.I. Park, B.E. Turk, F.E. Gerkema, L.C. Cantley, Q.X. Sang, Peptide substrate specificities and protein cleavage sites of human endometase/matrilysin-2/matrix metalloproteinase-26, Journal of Molecular Biology 277 (2002) 35168–35175.

[36] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, Journal of Computational Chemistry 26 (2005) 1701–1718.

[37] K.C. Chou, H.B. Shen, ProtIdent: a web server for identifying proteases and their types by fusing functional domain and sequential evolution information, Biochemical and Biophysical Research Communications 376 (2008) 321–325.

[38] A.B. Becker, R.A. Roth, Identification of glutamate-169 as the third zinc-binding residue in proteinase III, a member of the family of insulin-degrading enzymes, Biochemistry Journal 292 (Pt 1) (1993) 137–142.

[39] M. Ponpuak, M. Klemba, M. Park, I.Y. Gluzman, G.K. Lamppa, D.E. Goldberg, A role for falcilysin in transit peptide degradation in the *Plasmodium falciparum* apicoplast, Molecular Microbiology 63 (2007) 314–334.

[40] S. McGowan, C.J. Porter, J. Lowther, C.M. Stack, S.J. Golding, T.S. Skinner-Adams, et al., Structural basis for the inhibition of the essential Plasmodium falciparum M1 neutral aminopeptidase, Proceedings of the National Academy of Science of the United States of America 106 (2009) 2537–2542.

[41] S. Maric, S.M. Donnelly, M.W. Robinson, T. Skinner-Adams, K.R. Trenholme, D.L. Gardiner, et al., The M17 leucine aminopeptidase of the malaria parasite *Plasmodium falciparum*: importance of active site metal ions in the binding of substrates and inhibitors, Biochemistry 48 (2009) 5435–5439.

[42] S. Chaudhury, J.J. Gray, Identification of structural mechanisms of HIV-1 protease specificity using computational peptide docking: implications for drug resistance, Structure 17 (2009) 1636–1648.

[43] I.H. Moal, R. Agius, P.A. Bates, Protein-protein binding affinity prediction on a diverse set of structures, Bioinformatics 27 (2011) 3002–3009.

[44] I.Y. Gluzman, S.E. Francis, A. Oksman, C.E. Smith, K.L. Duffin, D.E. Goldberg, Order and specificity of the *Plasmodium falciparum* hemoglobin degradation pathway, Journal of Clinical Investigation 93 (1994) 1602–1608.

[45] V. Pelmenschikov, M.R. Blomberg, P.E. Siegbahn, A theoretical study of the mechanism for peptide hydrolysis by thermolysin, Journal of Biological Inorganic Chemistry 7 (2002) 284–298.

[46] M. Yokoyama, T. Oka, H. Kojima, T. Nagano, T. Okabe, K. Katayama, et al., Structural basis for specific recognition of substrates by sapovirus protease, Frontiers in Microbiology 3 (2012) 312.

[47] J. Audie, J. Swanson, Advances in the prediction of protein-Peptide binding affinities: implications for Peptide-based drug discovery, Chemical Biology & Drug Design 81 (2012) 50–60.

[48] J. Audie, J. Swanson, Recent work in the development and application of protein-peptide docking, Future Medicinal Chemistry 4 (2012) 1619–1644.