Characterization of Apicoplast-targeted Isoprenoids Biosynthesis Pathway Enzymes from *Plasmodium vivax* and *P. falciparum*

THESIS

Submitted in partial fulfilment
of the requirements for the degree of
DOCTOR OF PHILOSOPHY

by

GAGANDEEP SINGH SAGGU

Under the Supervision of

Prof. Vishal Saxena

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI (RAJASTHAN) INDIA

2016

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI, RAJASTHAN, INDIA

CERTIFICATE

This is to certify that the thesis entitled **"Characterization of Apicoplast-targeted Isoprenoids Biosynthesis Pathway Enzymes from *Plasmodium vivax* and *P. falciparum*"** and submitted by Gagandeep Singh Saggu, ID. No. 2009PHXF412P for award of Ph.D. Degree of the Institute embodies original work done by him under my supervision.

Signature (Supervisor)	:
Name (Supervisor)	: Vishal Saxena, Ph.D.
Designation	: Associate Professor, Department of Biological Sciences, BITS Pilani, Pilani Campus.
Date	:

Dedicated to My Family
& to
Those who have lost their
Life due to Malaria

ACKNOWLEDGEMENTS

The sun shines with its splendor to enlighten the world; in the same way the grace of Lord Almighty was in my life, especially during the completion of this research work. There is a paradox in the growth of scientific knowledge. As information accumulates in ever more intimidating quantities, disconnected facts and impenetrable mysteries give rise to rational explanations, and simplicity emerges from chaos.

The endless effort and patience of all the scientists and teachers in this perspective is the most motivational aspect. In the urge of contributing the society their heritage of knowledge they have given very best. It is a great privilege to take opportunity to thank all those who have directly and indirectly helped me in the completion of this endeavor. However, I was fortunate enough in obtaining the help of a large number of experts, scientists and colleagues to make this work completed. I wish to acknowledge the cooperation and encouragement of these people during my research work. Gratitude, one feels from heart, cannot be described exactly.

I earnestly wish to express my sincere thanks to my supervisor ***Prof. Vishal Saxena****, (Associate Professor, Department of Biological Sciences), BITS Pilani, for his unflinching interest and supervision, thought provoking discussions and meticulous guidance which enhanced me for learning. He spared his valuable time and shared his precious treasure of knowledge for this work. It was my privilege to work under his supervision. Being my guide, he was the biggest motivational source for me to complete my research work.*

I express my gratitude to Prof. Souvik Bhattacharyya, Vice-Chancellor, Prof. V. S. Rao, Prof. Bijendra Nath Jain, Prof. L. K. Maheshwari, (Former Vice-Chancellors) and Prof. Ashoke Kumar Sarkar, Director, Prof. G. Raghurama, (Former Director, BITS Pilani), for providing me very pleasant and enjoyable environment and allowing me to pursue my research work successfully. I am immensely thankful to Deputy Director, Deans and Associate Deans of Birla Institute of Technology & Science (BITS), Pilani for providing me the necessary facilities and financial support.

I also express my sincere regards to Prof. S. K. Verma (Dean, Academic Research Division, BITS, Pilani) for his motivation, constant support and encouragement. I express my sincere thanks to Prof. A. K. Das (Former Dean, Academic Research Division, BITS, Pilani) for his invaluable support and timely help. I thank Prof. Hemant Jadhav,

(Associate Dean, Academic Research Division, BITS, Pilani, Pilani Campus), for his cooperation and constant guidance during each phase of my research work. I also express my gratitude to the office staff of ARD, Mahipal Ji and Raghuveer Ji whose secretarial assistance helped me in submitting the various evaluation documents in time.

I am grateful to the members of my Doctoral Advisory Committee, Prof. Shibasish Chowdhury and Prof. Sanjeev Kumar for taking a lot of pain in reading my entire thesis carefully and giving some constructive comments which guided and inspired me to improve this work. I am thankful to Dr. Shilpi Garg for her unconditional support and time spent for helping me in my experimental work. My sincere thanks also go to Prof. Jitendra Panwar and Prof. Rajesh Mehrotra for their help and guidance in taking decision in tough situations. I am thankful to all the respectable faculty members, Department of Biological Science, BITS, Pilani for their generous help and support along with fruitful discussions during the different stages of my doctoral study. Thanks to all the office staff of the Dept of Biological Sciences Subash ji, Kamlesh soni, Mukesh ji, Naresh, Prameshwar ji, Mahindra Ji for their help during my experimental or official work.

Collaborators Prof Dhanpat K. Kochar and Prof Sanjay K. Kochar, Dr. Sheetal & Jyoti for their warm wishes from Bikaner and for the help in samples processing and shipping. I am also thankful to Dr. R. P. Pareek for his timely help and providing with the facilities at the BITS medical centre and Mr. Leeladhar for aiding in sample collection. I would also like to thank Prof. John H. Adams and Dr. Richard from Department of Global Health, College of Public Health, University of South Florida, Tampa, FL, USA for their timely help in generating flow cytometry data.

I always had a hearty inspiration from my seniors and lab mates Dr. Deepak Palakpati, Dr. Narayan Jalan, Dr. Pradeep Kumar, Dr. Amit Subudhi, Dr. Prakash, Dr. Sachi Verma, Dr. Garima, Dr. Purva, Dr. Navin Jain and Mr. Boopathi. I would also like to thank all my friends who provided me friendly atmosphere and helpful attitude at each step of my work and I have no words to express my feelings for happy and cherished moments and sweet memories I shared with them.

Few people are always there to stand beside you in ups and down of your research career and it's very hard to recall all of them but few of them are very special and you are not complete without them. Simardeep, Aprit Bhargava, Kuldeep Gupta, Isha, Shilpi Kaushik, Kanchan Soni, Senthil, Panchsheela, Gurpreet, Parva, Rini, Vandana, Rajnish, Vidushi,

Jyothi, Ramandeep, Shubhra, Leena, Heena, Naveen Singh, Monica, Sakthi Sree, Uthra, Ritika, Akanksha, Dilip, Sandeep, Vikram, Vikas, Shraddha, Poonam, Ranita, and many others. I will always be short of words to express my thanks to Zarna Pala, the one who is not just a colleague or a lab-mate but an individual who always sat next to me with a positive attitude and ready to correlate with my pace. Her presence was always responsible for the pleasantries in the lab.

I am overwhelmed with gratitude to my research scholar colleagues from other departments with whom I spent my holidays on cricket field Ankur, Prashant, Manoj, Kashi, Mukund and Parvez. This work was not possible without the suggestion and guidance of Dr. Sushil Yadav and the Central Animal House facility of BITS Pilani. I would also like to pay my gratitude to the creation of god those sacrified their life for my experiment.

Words fail to express my humble gratitude and profound regards to my loving parents, Shri. Gurtej Singh Saggu and Gurmeet Kaur Saggu for their love, inspiration, support and never ending blessings. I express my heartfelt thanks to my beloved sister Mrs. Manpreet Kaur and my brother Amandeep Singh Saggu for their love; motivation and accepting my decision in letting me pursue things with enormous encouragement.

Lastly I would like to say thanks to my mother Mrs. Gurmeet Kaur Saggu for her untiring support, encouragement, courage, love and belief in me throughout this work and my life. Without her love, encouragement, co-operation and sacrifice, successful achievement of this work would have only remained a dream.

My thanks are duly acknowledged to DST, CSIR and UGC, New Delhi, not only for providing the financial support which allowed me to undertake this research, but also for giving me the opportunity to attend conferences and meet so many interesting people.

Gagandeep Singh Saggu

Abstract

Abstract

Isoprenoids are structurally and functionally the most diverse group of metabolites known to play a key role in all aspects of life; e.g. in regulation of gene expression (prenylation of proteins), as membrane constituents (prenyl lipids in archaebacteria and sterol in eubacteria and eukaryotes), as vitamins, plant hormones (gibberellins, brassinosteroids, abscisic acid), photosynthetic pigments (carotenoids, side chain of chlorophyll), quinone in ETC (electron transport chain) and plant defence compounds (monoterpenes, sesquiterpenes, diterpenes). In *Plasmodium* the sole requirement of isoprenoids seems to be fulfilled only by the prokaryotic type isoprenoids biosynthesis pathway functional in the apicoplast, which makes it indispensable for the parasite survival. Further, the absence of this pathway in humans, consequently raises great interest in targeting the enzymes of this pathway for antimalarial drug development. Despite this, most of its participating enzymes remain uncharacterized in *Plasmodium* species, especially in *P. vivax*.

In the present study, we have characterized two major enzymes 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (IspD) and 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (IspG) participating in this pathway from *P. vivax* and *P. falciparum*. To achieve this objective, parasite DNA and RNA were isolated from the blood samples collected from malaria patients. After the confirmation of the type of infection, *ispD* and *ispG* genes were amplified from these samples, cloned, sequenced and compared with sequences of homologues and orthologs available from different prokaryotes, apicomplexans (including *Plasmodia*) and organisms containing photosynthetic plastids. The presence of conserved domains and signature motifs suggests their function to be retained during evolution. Further, the purified recombinant *Pv*IspD and *Pv*IspG proteins were used to raise specific antibodies, with the help of which these respective enzymes were successfully co-localized to the apicoplast in *P. vivax* infected blood smears obtained from patients. The purified proteins were also used for different biochemical assays to prove the functionality of the protein as well as the cofactors required for the activity of the enzyme. While the enzyme kinetics and activity of *Pv*IspD enzyme were proved by the pyrophosphate release assay, the reconstitution of Fe-S clusters on *Pv*IspG was performed under *in-vitro* conditions suggesting their requirement for the activity of IspG enzyme.

To check the potential of IspD and IspG enzymes as probable drug targets, we performed *in-silico* and *in-vitro* drug inhibition assays. We generated homology model of the two enzymes and identified their active sites and substrate binding residues. Later, we docked different drug molecules to check their binding efficacy. Fosmidomycin, a known inhibitor of IspC enzyme involved in the initial steps of the pathway, showed binding to the similar residues as MECP substrate for IspG enzyme, thus suggesting its action at multiple sites in the pathway. Similarly, amongst the other drug molecules analysed, azolopyrimidines showed good binding with IspD and alkyl diphosphate analogues with IspG, utilizing similar residues as their respective substrates indicating possible inhibitory effect of these drugs. We also performed *in-vitro* inhibition assays for cefapime (an antibiotic) against *Pv*IspD enzyme which showed promising results. This study detailing the two major enzymes of the Isoprenoids pathway, IspD and IspG, and their probable inhibitory molecules will lead to better understanding of the pathway and thus may help in designing future intervention strategies to fight malaria.

Contents

Contents

List of Tables

List of Figures

Abbreviations

3PGA	3-Phospho-glyceraldehyde
6FMVA-PP	6-Fluoro-mevalonate-pyrophosphate
ATP	Adenosine triphosphate
ACD	Acid Citrate Dextrose
ACT	Acyl CoA acetyl transferase
BPNs	Bisphosphonates
BSA	Bovine Serum Albumin
CDP	Cytidine 5'-diphosphate
CDP-ME	4-diphosphocytidyl-2C-methyl-D-erythritol
CDP-ME2P	4-diphosphocytidyl-2C-methyl-D-erythritol -2-phosphate
CL	Cholesterol
CTP	Cytidine 5'-triphosphate
DAPI	4',6-diamidino-2-phenylindole
DC	Dolichols
DEPC	Diethyl Pyrocarbonate
DHAP	Dihydroxy acetone phosphate
DMAPP	Dimethylallyl pyrophosphate
DNA	Deoxy ribonucleic acid
dNTP's	Deoxy ribonucleoside triphosphates
DXS	1-deoxy-D-xylose-5-phosphate synthase
DOXP	1-deoxy-D-xylose-5-phosphate
DTT	Dithiothreitol
DXP	1-Deoxy-D-xylulose 5-phosphate
DXS	1-Deoxy-D-xylulose 5-phosphate synthase
EDTA	Ethylene Diamine tetra acetic acid
ELISA	Enzyme Linked Immuno-sorbent Assay
FITC	Fluorescein isothiocyanate
FOS	Fosmidomycin
FPP	Farnesyl diphosphate
FPPS	Farnesyl diphosphate synthase
GA3P	Glyceraldehyde-3-phosphate
GGPP	Geranyl geranyl pyrophosphate
GGPS	Geranyl geranyl pyrophosphate synthase
GPP	Geranyl pyrophosphate
HA	Heme A
HMBDP	4-Hydroxy-3-methyl-2-(E)-butenyl-4-diphosphate
HMG-CoAS	3-Hydroxyl-3-methylglutaryl-CoA synthase

HMG-CoA	3-Hydroxy-3-methylglutaryl-CoA
HMG-CoAR	3-Hydroxyl-3-methylglutaryl-CoA reductase
IPP	Isopentenyl-5-pyrophosphate
IPPI	Isopentenyl-5-pyrophosphate isomerise
IPTG	Isopropyl-β-D-thiogalactopyranoside
IspC/ dxr	2C-methyl-D-erythritol 4-phosphate synthase
IspD	4-Diphosphocytidyl-2C-methyl-D-erythritol synthase
IspE	4-Diphosphocytidyl-2C-methyl-D-erythritol kinase
IspF	2C-Methyl-D-erythritol 2,4-cyclodiphosphate synthase
IspG	2C-Methyl-D-erythritol 2,4-cyclodiphosphate reductase
IspH	1-Hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate reductase
KCZ	Ketoclomazone
K_I	Dissociation constant of an inhibitor
K_M	Michaelis constant
LB	Luria Bertani
M	Molar
MECP	2C-metyl-D-erythritol 2,4-cyclodiphosphate
MEP	2C-Methyl-D-erythritol 4-phosphate
Min	Minute
mM	milimolar
MOPS	3-[N-Morpholino]propanesulfonic acid
MVA	Mevalonic acid
MVA-5-P	Mevalonate-5-phosphate
MVA-5-PP	Mevalonate-5-pyrophosphate
MVA-5-PPDC	Mevalonate-5-pyrophosphate decarboxylase
MVAK	Mevalonate kinase
NADPH	Nicotinamide adenine dinucleotide phosphate
nm	nanometer
P.	*Plasmodium*
PAGE	Poly Acrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PEP	Phosphoenolpyruvate
PPT	Phosphoenolpyruvate transporter
PMVAK	Phosphomevalonate kinase
Pf	*Plasmodium falciparum*
Pv	*Plasmodium vivax*
q. s.	Quantity sufficient
nt	Nucleotides
O. D.	Optical Density

RDTs	Rapid diagnostic tests
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SDS	Sodium Dodecyl Sulphate
SQ	Squalene
ST	Sterol
TEMED	N, N, N’, N’-tetramethylethylenediamine
TPT	Triose-phosphate transporter
Tris	2-Hydroxy methylamine
UQ	Ubiquinones

Amino Acid

Alanine	ala	A
Arginine	arg	R
Asparagine	asn	N
Aspartic acid	asp	D
Cysteine	cys	C
Glutamine	gln	Q
Glutamic acid	glu	E
Glycine	gly	G
Histidine	his	H
Isoleucine	ile	I
Leucine	leu	L
Lysine	lys	K
Methionine	met	M
Phenylalanine	phe	F
Proline	pro	P
Serine	ser	S
Threonine	thr	T
Tryptophane	trp	W
Tyrosine	tyr	Y
Valine	val	V

Nucleotide bases

Adenine	A
Guanine	G
Cytosine	C
Thymine	T
Uracil	U

Introduction & Literature Review

Chapter I

Introduction & Literature Review

With an annual mortality of around a million humans, malaria remains a major problem of the developing world. Most of the drugs available against malaria are losing their efficacy due to the drug resistance acquired by the parasite. Failures of newly developed vaccines during the clinical trial phases is another frustrating effort for the scientific community. To treat this deadly disease new drugs are desperately required. The disease is caused by parasitic protozoa of the genus *Plasmodium*. Its life cycle is complex, alternating between sexual phase in an invertebrate (mosquito) host and asexual phase in a vertebrate (human) host.

Malaria parasite belongs to the phylum Apicomplexa and a characteristic feature of these apicomplexans parasite is the presence of a non-photosynthetic plastid organelle called apicoplast. Apicoplast houses four important metabolic pathways and amongst these, isoprenoids biosynthesis is shown to be essential for the parasite survival in different erythrocytic stages. The products synthesized by this pathway are important and involved in various cellular processes taking place inside the parasite. Inhibition of this pathway can cause immediate death of the parasite which shows the importance of the enzymes involved in this pathway as potential drug targets.

1.1 Malaria : An Introduction

Malaria has been a cause of enormous morbidity and mortality from decades and its control remains a farfetched dream despite numerous efforts and intervention strategies devised by researchers. According to the WHO malaria report (WHO, 2014), 198 million clinical infections of malaria occurred globally in 2013 leading to 0.58 million deaths. The malaria parasite belongs to genus *Plasmodium,* containing ~140 species, out of which only six species are known to infect humans, namely *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and recently reported *P. knowlesi* (Singh et al., 2004; White, 2008) and *P. cynomolgi* (Ta et al., 2014). 80% of all malaria cases are reported from Africa where majority of the infections are due to *P. falciparum* with no or very less cases of *P. vivax*. The absence of *Plasmodium vivax* in African population is attributed to the absence of duffy binding protein (DBP) in the population. Outside Africa, *P. vivax* accounts for most of the cases, with almost 95% of malaria cases in Asia (Carter and Mendis, 2002). The wider geographic distribution of *P. vivax* is due to its ability to develop in the *Anopheles* mosquito vector at lower temperatures, and its fitness to survive at higher altitudes. It also has a dormant liver stage known as hypnozoite that enables it to survive for many years as a potential reservoir of infection and can activate months later to cause a relapse. Both *P. falciparum* and *P. vivax* are reported to cause severe manifestations in humans like cerebral malaria, acute respiratory distress syndrome, hepatic dysfunction, severe anaemia, etc. (Kochar et al., 2005 & 2009; Alexandre et al., 2010; WHO, 2014) which has further complicated our fight against malaria.

1.2 Life cycle of *Plasmodium*

The malaria parasite completes its life cycle in two hosts: humans (asexual stages) and mosquitoes (sexual stages) (**Figure 1.1)**. The asexual life cycle in humans begins with the bite by a female *Anopheline* mosquito injecting around 100-200 sporozoites into the subcutaneous tissue (Medica and Sinnis, 2005; Jin et al., 2007). Some of these sporozoites which make it into the peripheral blood circulation infect hepatocytes (Vanderberg and Frevert, 2004; Yamauchi et al., 2007; Amino et al., 2008) initiating the pre-erythrocytic cycle and releasing thousands of merozoites into the circulation within 6-12 days. Although some of the merozoites are phagocytosed (Kumaratilake and Ferrante, 2000), the others enter the erythrocyte to begin the erythrocytic cycle (or erythrocytic schizogony).

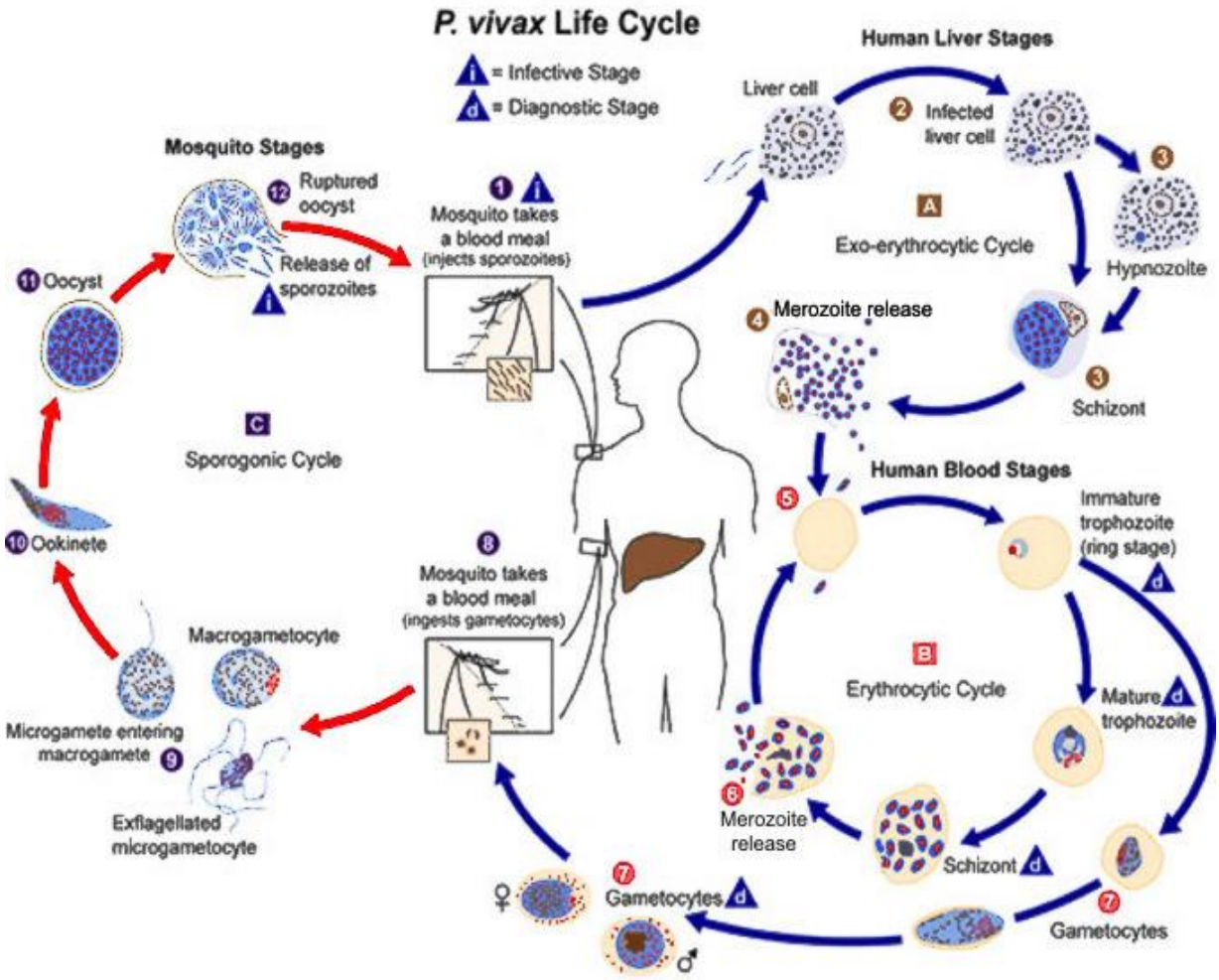


Figure 1.1: The *P. vivax* life cycle: It involves two hosts. During a blood meal, a malaria-infected female Anopheles mosquito inoculates sporozoites into the human host (1). Sporozoites infect liver cells (2) and either enter a dormant hypnozoite state or mature into schizonts (3), which rupture and release merozoites (4). After this initial replication in the liver (exo-erythrocytic schizogony A), the parasites undergo asexual multiplication in the erythrocytes (erythrocytic schizogony B). Merozoites infect red blood cells (5). The ring stage trophozoites mature into schizonts, which rupture, releasing merozoites (6). Some parasites differentiate into sexual erythrocytic stages (gametocytes) (7). Blood stage parasites are responsible for the clinical manifestations of the disease. The gametocytes, male (microgametocytes) and female (macrogametocytes), are ingested by an *Anopheles* mosquito during a blood meal (8). The parasite multiplication in the mosquito is known as the sporogonic cycle (C). While in the mosquito's stomach, the microgametes penetrate the macrogametes generating zygotes (9). The zygotes in turn become motile and elongated (ookinetes) (10) which invade the midgut wall of the mosquito where they develop into oocysts (11). The oocysts grow, rupture, and release sporozoites (12), which make their way to the mosquito's salivary glands. Inoculation of the sporozoites (1) into a new human host perpetuates the malaria life cycle. Source: Centers for Disease Control and Prevention (CDC, Atlanta)

Ref: http://www.cdc.gov/malaria/biology/life_cycle.htm

In RBCs (erythrocytes), each merozoite develops through ring and trophozoite stage into schizont containing multiple merozoites (erythrocytic schizogony). Further, with the rupture of schizont, these merozoites are released into the blood stream where they reinvade the fresh erythrocyte to perpetrate the erythrocytic cycle. Some of the schizont release sexually committed merozoites differentiating into gametocytes, all of one sex only (gametogony) (Baker, 2010; Bousema and Drakeley, 2011). The factors triggering gametogony are not clearly understood, and may include interplay between a number of host factors as well as parasite signalling pathways. In *P. vivax* and *P. ovale* some of the hepatic forms become dormant as hypnozoites and persist for a considerable period before they undergo pre-erythrocyte schizogony to liberate merozoites into the circulation causing relapse in these infections (Krotoski et al., 1980; Wells et al., 2010).

When the female *Anopheles* mosquito feeds on blood of an infected patient, it takes up these gametocytes from the human and further development of these gametocytes takes place in the mosquito mid-gut. Within a few minutes of the blood meal, each microgametocyte ex-flagellates in the mid-gut and forms eight micro gametes. The microgametes move quickly and fertilize a macrogamete to form the zygote. Within 18-24 hrs, the zygote elongates into a slowly motile ookinete. This ookinete transverses the peritrophic membrane and the epithelial cell lining of the midgut and transforms into an oocyst beneath the basement membrane. Between 7 to 15 days post infection, a single oocyst forms more than 10,000 sporozoites. These motile sporozoites migrate into the salivary glands and accumulate in the acinar cells. When an infected mosquito bites a susceptible vertebrate host, the *Plasmodium* life cycle begins again (Hall et al., 2005).

1.3 Treatment for Malaria

The fight against malaria dates over 125 years back. During this period, various drug molecules have been formulated to combat malaria, amongst which the blood schizonticides that act on symptomatic asexual intra-erythrocytic stage of *Plasmodium* development, constitute the major class of antimalarials in use. Chloroquine, a 4-aminoquinoline synthetic derivative of quinine, is a major blood schizonticide which remained the first line of therapy for many years due to its low cost and minimum side effects. However, with the development of resistance to this drug especially in *P. falciparum*, the drug therapy for the parasite shifted towards antifolates. The antifolates therapy consisted of a combination of

dihydrofolate reductase inhibitors like proguanil, chloroproguanil, pyrimethamine and trimethoprime and sulfa drugs like dapsone, sulfalene, sulfamethoxazole and sulfadoxine. While the *P. vivax* was reported to show inherent resistance to the antifolates, *P. falciparum* gained resistance to these drugs soon after their implementation (Wongsrichanalai et al., 2002). During the last decade, Artemisinin obtained from the plant *Artemisia annua* and its derivatives have gained a lot of attention as antimalarials. These drugs proved to be very effective due to the rapid parasite clearance rate and have been suggested to target multiple sites in the parasite (Ferreira et al., 2010). However, due to their very short half-life in the plasma, they are always prescribed in combination with long lasting drugs like sulfadoxine-pyrimethamine (SP), etc. Recent reports from Thai-Myanmar borders have shown that the parasite has developed resistance to these combination therapies as well (Noedl et al., 2008; Dondrop et al., 2009). Thus, the parasite has gained resistance to almost all the available chemotherapeutic antimalarial drugs. At present, in accordance to our National Drug Policy, to treat the disease Artesunate in combination with SP is administered for *P. falciparum* infection while chloroquine along with primaquine (hypnozonticide) remains the first choice for *P. vivax* (NVBDCP, 2010).

The other approach to tackle the disease had been to develop a potential vaccine against the parasite, for which the principal challenge lies in the ability of the parasite to change its form throughout the life cycle and its development in intracellular and extracellular niches in both the human host and the mosquito vector (Aly et al., 2009). Trials are still underway to look into the probable candidates, with one vaccine namely RTS, S/ AS01 reaching the third phase clinical trials and showing only moderate efficacy of 26-50%. However, as yet no vaccine is available with complete efficacy. Thus, the absence of an effective vaccine till date and the emergence of resistance to existing chemotherapeutic agents not only compromise our fight against malaria, but also highlight the need to develop new compounds to be used to treat malaria.

Scientists all over the world have adapted several approaches to combat this disease and studies are still going on to find reliable drug targets against malaria. The impact of selected antibiotics on combating malaria infections was discovered in the mid of last century. In the race for search of a suitable antimalarial, antibiotics like Azithromycin, Ciprofloxacin, Fosmidomycin, Rifampicin, Clindamycin etc. have been tried. While Rifampin inhibited cell division at clinically optimal concentrations quickly killing parasites, others remained

relatively inactive initially against parasites at desired pharmacological concentrations but exerted a delayed death effect, blocking the erythrocytic development of treated parasites progeny (Dahl and Rosenthal, 2007). Recently, studies on modes of action of these inhibitors in malaria parasites have been initiated, prompted by the discovery of the apicoplast, which because of its prokaryotic origin is being looked upon as a putative drug target.

1.4 Apicomplexan Parasite and Apicoplast

Plasmodium parasites belong to a group of intracellular parasites called apicomplexans. All apicomplexans possess an apical complex at the anterior end of the cell which consists of an assembly of organelles critical for the parasitic life cycle (Levine et al., 1985). Apicomplexans have evolved from free living photosynthetic red algae through the process of secondary endosymbiosis (**Figure 1.2**) (Wilson, 2005; Janouškovec et al., 2010) and harbours a plastid like organelle surrounded by three to four membranes called apicoplast, which has its own DNA (McFadden et al., 1996; Wilson et al., 1996).

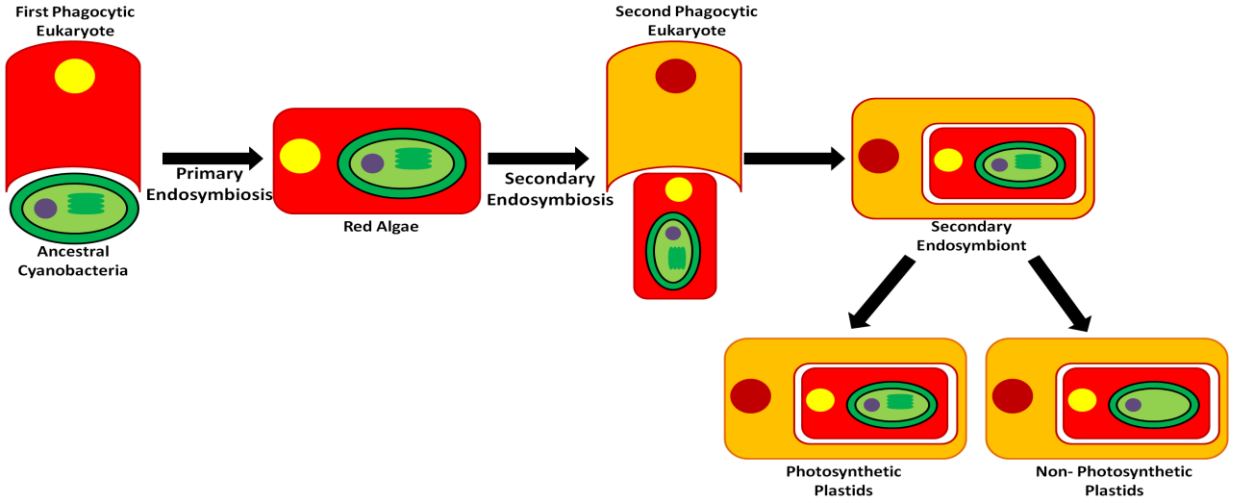


Figure 1.2: Diagrammatic representations for secondary endosymbiosis process.

Though the plastid has lost the photosynthetic genes and thus its ability of photosynthesis during the course of evolution as evident from the sequences of apicoplast genome from different *Plasmodium* species including *P. falciparum* and *P. vivax* (Wilson et al., 1996; Saxena et al., 2012), it can still perform several other important metabolic functions which include the biosynthesis of isoprenoids, Type – II fatty acids, heme and iron sulphur (Fe-S) clusters (Foth & McFadden, 2003). These biosynthetic pathways are indispensable for both intra-erythrocytic and intrahepatic development of the parasite (Dahl et al., 2006; Stanway et al., 2009) and are being looked upon as putative drug targets due to the absence of their

counterparts in humans (Janouškovec et al., 2010). Other than these metabolic pathways of prokaryotic origin, another common feature between apicoplast and prokaryotes is the presence of similar housekeeping processes. Inhibition of these processes can inhibit the replication of apicoplast genome and ultimately cause the disruption of the parasite cell by a process called as "delayed death effect" (Pfefferkorn et al., 1992; Fichera et al., 1995; He et al., 2001).

Apicoplast has been shown to have a dynamic morphology, tightly coordinated with the development of other organelles in apicomplexan parasites which is quite evident from the close association of apicoplast with mitochondria and endoplasmic reticulum during its development (van Dooren et al., 2005). Crosstalk has been observed between these organelles in certain cases like during heme biosynthesis (Dhanasekaran et al., 2004; Sato et al., 2004; Nagaraj et al., 2009, 2010) where the pathway and its enzymes are found to be shared between apicoplast, mitochondria and cytosol of the parasite. However, despite the fact that apicoplast and the metabolic pathways residing in it are indispensable for the parasite, as yet no therapeutic agent is known that target's the plastid effectively to treat acute and severe malaria. Development of new apicoplast inhibitors as antimalarials has been challenging due to gaps in our knowledge of apicoplast biology, specific pathways and proteins to target.

1.5 Metabolic Pathways in Apicoplast

The apicoplast in *Plasmodium* parasite is believed to have retained several characteristic features during the process of evolution indicating its prokaryotic origin. These include four major metabolic pathways specific to prokaryotes, including type – II fatty acid biosynthesis, non-mevalonate isoprenoids biosynthesis, heme biosynthesis and the Fe-S cluster biogenesis. Studies have suggested that the presence of this organelle and metabolic pathways functional in it are indispensable for the parasite survival and can be used as suitable drug targets without distressing the eukaryotic host (McFadden and Roos, 1999). Almost all the enzymes participating in these pathways are encoded by the parasite nuclear genome and targeted to the apicoplast (Nuclear Encoded Apicoplast Targeted, NEAT) *via* a bi-partite N-terminal leader sequence containing a signal and transit peptide (Altincicek et al., 2001; Ramya et al., 2006; Carlton et al., 2008). The signal peptide at the N-terminus is usually 20-30 residues and targets the protein to the secretory pathway in the endoplasmic reticulum whereas a

transit peptide ranges from 24 to 150 residues (Waller et al., 1998; Parsons et al., 2007) and is meant to cross the inner membranes of the apicoplast. Once the protein gets localized in the apicoplast, the signal peptide and the transit peptide are cleaved off within the apicoplast (He et al., 2001). Based on bacterial metabolism and *in-silico* targeting details, different enzymes participating in the metabolic pathways residing in apicoplast were shortlisted and the whole pathways were reconstructed and mapped out (Foth & McFadden, 2003; Ralph et al., 2004) (**Figure 1.3**).

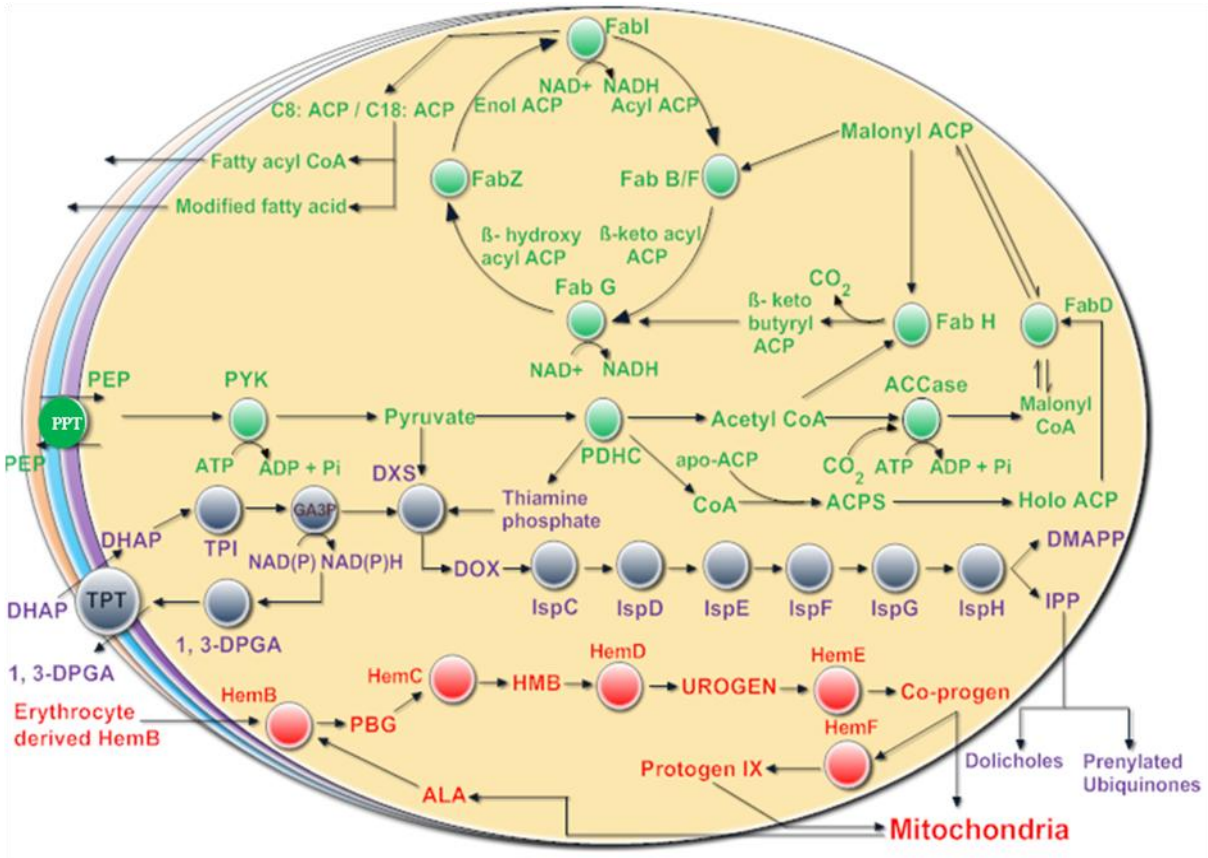


Figure 1.3: Overview of apicoplast metabolic pathways adapted from (Ralph et al., 2004).

1.5.1 Fatty Acid Biosynthesis Pathway

Fatty acid biosynthesis is fundamental to cell growth, differentiation as well as homeostasis. Almost all living organisms including apicomplexan parasites such as *P. falciparum* synthesize fatty acids. FAS pathway in apicomplexan parasites is different from that of their eukaryotic hosts in a number of aspects. Unlike their eukaryotic hosts, FAS pathway of apicomplexan parasites is catalyzed by discrete enzymes, and is known as type II pathway or "dissociative" pathway (Ralph et al., 2004). On the other hand, in mammalian FAS pathway (type I or associative pathway) all the steps are catalyzed by a single giant polypeptide enzyme called fatty acid synthase, (Stoops et al., 1975). In the parasite, fatty acid type II

pathway is the main source for de novo synthesis of fatty acids and lipoic acid (Mazumdar and Striepen, 2007), where lipoic acid is an essential cofactor for a number of oxidative decarboxylases including pyruvate dehydrogenase, α-ketogluterate dehydrogenase etc. Recent gene deletion experiments have proved that FAS II pathway of the apicoplast is only essential in the late liver stage of the erythrocytic life cycle of malaria parasite (Vaughan et al., 2009) and for the sporozoite development within the midgut oocyst during the mosquito stages (van Schaijk et al., 2014).

1.5.2 Heme Biosynthesis Pathway

Besides detoxifying heme that comes from host hemoglobin degradation, *P. falciparum* also synthesizes its own heme (Bonday et al., 1997). Heme is an important prosthetic group in many proteins such as cytochromes (Ferreira et al., 2008). It has been demonstrated that heme biosynthetic pathway in *P. falciparum* is shared among apicoplast, mitochondrion and cytosol of the parasite (Dhanasekaran et al., 2004; Sato et al., 2004; Nagaraj et al., 2009, 2010). After the initial synthesis of aminolevulinic acid in the mitochondrion, heme biosynthetic pathway relocates itself to the apicoplast after which it comes to the cytosol and finally gets back to the mitochondria. Hem H, the last enzyme in the heme biosynthetic pathway is also imported from the host to the cytosol of the parasite in a bigger proportion as compared to its de novo synthesis in the parasite mitochondrion (Nagaraj et al., 2010). Recent studies have suggested that the heme pathway is not required for the asexual blood stages but is needed during the sexual stages of the parasite in the mosquito (Nagaraj et al., 2013).

1.5.3 Fe-S Cluster Biogenesis Pathway

Iron-sulfur (Fe-S) clusters are ubiquitous and critical cofactors in diverse biochemical processes. Apicomplexan parasites, including *Plasmodium*, harbour two separate (Fe-S) cluster biogenesis pathways in the mitochondrion and the apicoplast (Gisselberg et al., 2013; Charan et al., 2014). The ISC pathway in *Plasmodium* supplies Fe-S clusters to mitochondrial proteins, for example to components of the electron transport chain (ETC). The maturation of cytosolic and nuclear Fe-S proteins is also dependent on the ISC pathway (Dellibovi-Ragheb et al., 2013). The SUF pathway on the other hand provides Fe-S clusters to the proteins targeted to the apicoplast like the IspG and IspH of the isoprenoids biosynthesis pathway, MiaB, LipA etc. All the genes involved in these pathways are encoded by the nuclear genome except SufB, which is encoded by the apicoplast genome (Kumar et

al., 2011, Pala et al., 2016). Recent studies have suggested that the Fe-S cluster biogenesis pathway in the parasite is essential for the vital activity of the parasite during the erythrocytic stages (Haussig et al., 2014). Literature suggests the role of SUF pathway during oxidative stress in bacteria, however the same in *Plasmodium* is yet to be investigated.

1.5.4 Isoprenoids Biosynthesis Pathway

Isoprenoids are structurally and functionally the most diverse group of metabolites known to play a key role in all aspects of life; eg. in regulation of gene expression (prenylation of proteins), as membrane constituents (prenyl lipids in archaebacteria and sterol in eubacteria and eukaryotes), as vitamins, plant hormones (gibberellins, brassinosteroids, abscisic acid), photosynthetic pigments (carotenoids, side chain of chlorophyll), quinones in ETC and plant defence compounds (monoterpenes, sesquiterpenes, diterpenes) (Sacchettini and Poulter, 1997; Bach et al., 1999; Hunter, 2007). They are synthesised from the same precursor molecules, Isopentenyl Pyrophosphate (IPP) and Dimethylallyl Pyrophosphate (DMAPP) in all the three domains of life i.e. eubacteria, archaebacteria and eukarya. Two distinct pathways have been reported for the synthesis of these precursor molecules - the Mevalonate dependent pathway (MVA) and the Methyl Erythritol Phosphate or 1-deoxy-D-xylulose-5-phosphate (MEP/DOXP) pathway. While the MVA pathway is predominantly present in archaea and most eukaryotes (including all mammals and higher plants), the MEP pathway is believed to be functional in bacteria, plant plastids, members of chlorophyta and pathogenic microorganisms (Rohmer, 1999). The two pathways differ from each other in terms of the preliminary substrate, the formation of mevalonate as an intermediate in MVA pathway and the quantity of final products (IPP and DMAPP).

Evolutionary divergence of the two isoprenoids biosynthesis pathway remains controversial. Various studies have examined the origin and evolution of the enzymes involved in isoprenoids biosynthesis process using phylogenetic analysis or by studying horizontal gene transfer events (Boucher and Doolittle, 2000). Finally, it was concluded that MVA pathway is likely to be an ancestral metabolic route in all the three domains of life, and was probably present in the last common ancestor of all organisms (Lombard and Moreira, 2011). However, it is believed that the MEP pathway is more efficient route than the MVA pathway for isoprenoids production in terms of energy consumption and production (Partow et al., 2012). In plants, both the MVA and the MEP pathways (**Figure 1.4**) are operational simultaneously, but they are compartmentalized in the cytosol and plastids, respectively

(Lichtenthaler et al., 1997). It is suggested that there exists a cross talk between the two pathways. The MVA-derived precursors are believed to be used for the synthesis of isoprenoids in the plastid, while MEP-derived precursors may be exported to the cytosol in some plants and tissues. However the exact mechanism and the transporters involved remains to be explored (Hemmerlin et al., 2003).

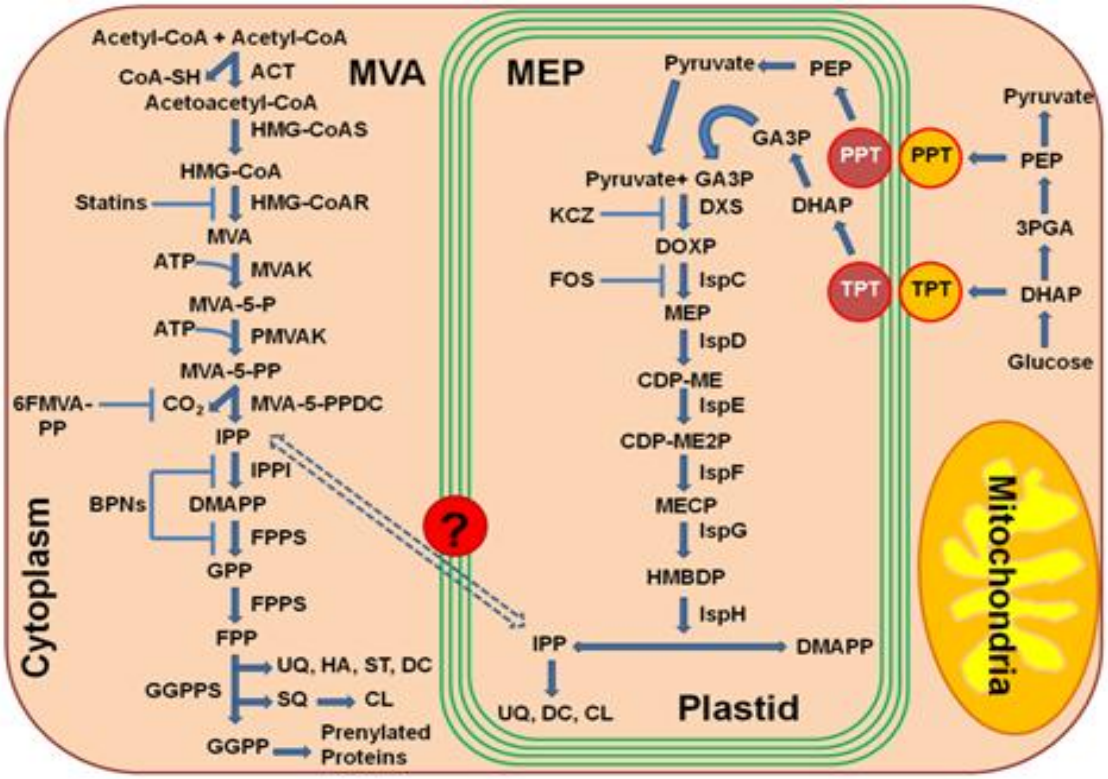


Figure 1.4: Isoprenoid biosynthesis by mevalonate dependent (MVA) and non mevalonate pathway (MEP) in a plant cell. Isoprenoids are derived from the basic 5-carbon isoprenoid building blocks IPP and its isomer, DMAPP. In MVA pathway IPP and DMAPP are synthesized by condensation of acetyl CoA. Enzymes participating in this pathway are ACT, HMGS, HMGR, MVK, PMK and MDS. Statins inhibits HMGR, the rate-limiting enzyme of the mevalonate pathway and 6FMVA-PP and BPNs inhibits the other steps of the pathway. In the MEP pathway, IPP and DMAPP are generated from pyruvate and GA3P. Enzymes of this pathway are named here according to their *E. coli* homologues (DXS, IspC, IspD, IspE, IspF, IspG and IspH). KCZ and FOS inhibit the rate limiting steps of this pathway and blocks isoprenoids biosynthesis**. Note:** The full form of enzymes and products is given in the list of abbre*via*tions.

The MVA pathway (**Figure 1.4**) starts with the condensation of acetyl-CoA and acetoacetyl-CoA in the presence of 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase and forms HMG-CoA. The first committed step of this pathway is the biosynthesis of mevalonic acid, where reductive deacylation of HMG-CoA takes place with the help of HMG-CoA reductase (HMGR) enzyme. Activity of this enzyme is tightly regulated at transcriptional and post-transcriptional levels in plant cytosol (Leivar et al., 2005; 2011). Studies on transgenic plants

have also shown up-regulation of this enzyme, which resulted in the increased level of sterols (Manzano et al., 2004). In the subsequent steps phosphorylation of mevalonic acid take place with the help of enzymes mevalonate kinase (MVAK) and phosphomevalonate kinase (PMVAK) which ultimately converts mevalonate to mevalonate-5-pyrophosphate (MVA-5PP). Decarboxylation of MVA-5PP in the presence of mevalonate-5-pyrophosphate decarboxylase (MVA-5-PPDC) enzyme gives the simplest isoprenoid unit IPP. These IPP units are further converted to DMAPP with the help of IPP isomerase (IPPI) enzyme or enter into the downstream process where they are converted to different complex molecules and performs various biochemical functions (Lichtenthaler et al., 1997).

The MEP pathway (**Figure 1.4**), in contrast, starts with the reaction of pyruvate and glyceraldehyde-3-phosphate (GA3P) in presence of the DOXP (or DXP) synthase enzyme (DXS). These initial substrates are imported to the apicoplast in the form of Dihydroxyacetone Phosphate (DHAP) and Phosphoenolpyruvate (PEP) *via* apicoplast membrane localized transporters 'triose phosphate transporter (TPT)' and 'phosphoenol pyruvate transporter (PPT)' respectively (Mullin et al., 2006). All the components of this pathway are believed to be conserved among Apicomplexans including *Plasmodium*. The end products of this pathway are IPP and DMAPP which are synthesized in larger quantities as compared to those produced in MVA pathway.

1.6 MEP/ DOXP Pathway in *Plasmodium*

The first evidence for the presence of isoprenoids biosynthesis pathway in apicomplexans was given by Jomaa et al. (1999) where after *P. falciparum* genome database mining, the gene for DOXP reductoisomerase (IspC), involved in second step of pathway was identified and localized to the apicoplast. With the availability of whole genome of *P. falciparum* in 2002, all the genes encoding the enzymes functional in this pathway were identified. All these enzymes are nuclear-encoded and translocated to the apicoplast through a secretory pathway utilizing a bipartite N-terminal signaling motif. (van Dooren et al., 2002).

1.6.1 MEP Pathway Enzymes and their role as a potential drug target

The synthesis of isoprenoids happens *via* a chain reaction and inhibition of any step can cause a lethal effect on the functionality of the pathway. Thus, the MEP pathway has long being considered as an important target of study in pathogenic organisms because of its

essentiality and ubiquity. The MEP pathway has been extensively studied as a drug target in prokaryotes, mainly against the human pathogens and this study was further extended to the chloroplast of herbicide plants, where inhibition of this pathway can help in removal of unwanted weeds from economically important crops (Rodriguez-Concepcion, 2004). All the enzymes for this pathway have been characterized in *E. coli* while very few have been elucidated from pathogenic organisms including *Plasmodium.*

In *Plasmodium*, the MEP pathway is reported to be indispensable for both hepatic and erythrocytic stages of the parasite asexual life cycle. Recent reports have suggested that the "apicoplast less" *Plasmodium* parasite can divide indefinitely in culture if supplemented exogenously with the end products IPP and DMAPP. This proves that during the erythrocytic stages, the only essential function of apicoplast is the synthesis of isoprene unit precursors, IPP and DMAPP (Yeh and DeRisi, 2011). Similar to the inhibition of erythrocytic stages of *P. falciparum* with Fosmidomycin (Jomaa et al., 1999), blockade of the parasite hepatic stages by peptide-tagged Fos was observed in *P. berghei* (Sparr et al., 2013). Recent reports have also proven that the products of MEP pathway are required for the early stages of parasite gamete development (Wiley et al., 2015). Thus, all these studies suggest the importance of this pathway at different phases of parasite's life cycle required for its survival. The importance of apicoplast being the functional site or compartment for the activity of the MEP pathway is suggested by the fact that even after the reconstruction of the complete MEP pathway with all its major and minor components in *Saccharomyces cerevisiae*, the pathway remained non-functional due to the lack of a reducing environment (Partow et al., 2012). However, the same fact was not true for MVA pathway, where insertion of the components involved in this pathway into the chloroplast genome resulted in the increased production of IPP (Kumar et al., 2012). Among other apicomplexans, the MEP pathway is also reported from *Toxoplasma gondii*, wherein inhibition of IspC and IspH enzymes was shown to be lethal for the parasite survival (Nair et al., 2011) confirming the importance of this pathway. Despite this, most of the enzymes involved in the MEP pathway of *Plasmodium* have not been experimentally characterized till date.

The MEP pathway functional in the apicoplast of *Plasmodium* involves 7 catalytic steps finally producing IPP and DMAPP. These enzymes are unique as drug targets because their function cannot be compensated with any other enzyme of the parasite and their absence from the human host cells ensures no side effect on the host physiology. The enzymes involved in the MEP pathway and their potential as a drug target is detailed:

DXS (DXP Synthase)

The initial substrates pyruvate and GA3P are acted upon by a thiamine pyrophosphate dependent enzyme DXS (DXP Synthase; EC 4.1.3.37) to generate 1-deoxy-D-xylose-5-phosphate (DOXP/ DXP) (Sprenger et al., 1997). This first step is believed to be rate limiting in some organisms and represent a branch point for Vitamin B and isoprenoids synthesis in bacteria (Hahn, 2001). In *Plasmodium*, the characterization of DXS from both *P. falciparum* and *P. vivax* suggests its existence as a homodimer (Handa et al., 2013) which contains one-bound Mg(II) per enzyme molecule, having both a catalytic and structural role in the enzyme. This enzyme consists of three functional domains: TPP (Thiamine Pyrophosphate binding domain), PYR (Pyrimidine binding domain) and transketolase C domain. Like other thiamine pyrophosphate dependent enzymes (ThDP), the activated acetyl group is condensed with an acceptor substrate glyceraldehyde-3-phosphate *via* nucleophilic addition to generate a new carbon-carbon bond yielding DXP. All ThDP-dependent enzymes catalyze two successive half reactions. The first step involves the attack of an activated ThDP ylide on the first substrate. The next step can occur *via* three distinct mechanisms: i) the most common classical ping-pong mechanism; ii) through an ordered sequential kinetic mechanism or iii) through an alternate random sequential mechanism (Brammer et al., 2011). Enzyme kinetics of DXS in *Plasmodium* point towards a random sequential kinetic mechanism, an unusual finding for ThDP-dependent enzymes similar to *Rhodobacter capsulatus* (Battistini et al., 2016).

A herbicide ketoclomazone has been found to inhibit the DXS enzyme of both *E. coli* and *Haemophilus influenzae* by binding to a site, which differs from both the pyruvate-binding site and the GAP-binding site on DXP synthase, thus suggesting non-competitive inhibition (Matsue et al., 2010). However the binding site of the drug has not been elucidated yet. Another inhibitor tried against DXS enzyme in *R. capsulatus* and *P. vivax* is an analogue of pyruvate; β-Fluoropyruvate (F-Pyr); which shows competitive inhibition pattern with respect to pyruvate (Handa et al., 2013).

IspC (DXP reductoisomerase)

IspC (EC 1.1.1.267) is the first enzyme committed to the isoprenoids biosynthesis in the MEP pathway. It is responsible for the conversion of DOXP to 2-C-methyl-D-erythritol-4-phosphate (MEP) by using the NADPH *pro-S* hydride. Its structure has been well defined in *E. coli* (Sweeney et al., 2005), *P. falciparum* (Umeda et al., 2015) and various other

pathogens including *Mycobacterium tuberculosis* and *Zymomonas mobilis* (Ricagno et al., 2004; Henriksson et al., 2006). It is a class B dehydrogenase enzyme which exists as a homodimer with each subunit composed of three domains. The N-terminal domain is a member of dinucleotide binding fold and serves for binding of NADPH. A central catalytic domain harbours the binding site for divalent metal ion like Mn^{2+}, Mg^{2+}, the phosphate binding site of the substrate and the catalytic loop. The C-terminal domain is connected to the catalytic domain by a linker region that spans entire monomer. The C-terminal domain appears to have a structural role in supporting the catalytic domain. Normally, the active site has two different conformations, open and closed. The open conformation allows the substrate, DXP, to enter and bind to the active site while in the closed conformation, a flap covers the active site and catalytic function is activated.

In *E. coli* and *P. falciparum*, fosmidomycin (Fos) has been characterized as an inhibitor of IspC enzyme. It behaves as a substrate analogue of DOXP and competes for its binding site on IspC (Jomaa et al., 1999; Steinbacher et al., 2003). In *P. falciparum*, fosmidomycin interacts with S270, N311, H293, W296, M298, D231, E233, and E315 of the catalytic domain, which are conserved in all human malaria parasites. The effect of Fos has been reported to be varied amongst different stages of *Plasmodium* development as well as within the apicomplexans. It was found that the Fos is only effective on the erythrocytic stages of *Plasmodium* due to the formation of new permeability pathways, but has no effect on the liver stages of *P. berghei*. Also Fos is not effective on many apicomplexans including *Toxoplasma, Eimeria* even at higher concentrations, probably due to its impermeability to the parasite plasma membrane (Baumeister et al., 2011).

In a study conducted by Zhang et al. (2011) concentration of isoprenoid pathway metabolites was checked in the presence of Fos, it was observed that after treatment of *E. coli* and *P. falciparum* cultures with Fos, there was accumulation of MEP (product of IspC and substrate for IspD) in treated samples when compared to controls. This indicates to partial or feeble effect of Fos on IspC inhibition in *in-vivo* conditions compared to *in-vitro* conditions where Fos readily inhibits recombinant purified IspC. Still, Fos treated parasites were unable to survive indicating to a second target of Fos in isoprenoid pathway. With the further investigations of metabolites they were able to conclude that IspD should be the second target of Fos as its downstream metabolites decreased considerably post Fos treatment. However, there still remains a major gap in this knowledge of IspD inhibition by Fos, as it

has been observed that a much higher concentration of Fos is required to inhibit *E. coli* recombinant purified IspD. In the same report, Zhang et al. also showed through *in-silico* studies that *Pf*IspD binds to its substrate MEP preferentially (binds with 5 residues) as compared to Fos (binds with two residues only). This indicates to a possibility that IspD in Fos treated bacteria or parasites is not inhibited directly but by accumulation of MEP or through some other feedback mechanism (Zhang et al., 2011). Thus, there is requirement of deeper investigation into this data.

IspD [2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (YgbP)]

IspD (EC 2.7.7.60) enzyme, catalyzes the cytidylation process and participates in the third step of the pathway where its activity is highly dependent upon divalent cations Mg^{2+} or Mn^{2+} (Richard et al., 2001). Nucleotide derivatives i.e. cytosine-5′-triphosphate (CTP) and phosphate groups i.e. 2-C-methyl-D-erythritol-4-phosphate (MEP) are introduced directly as substrates in the IspD catalyzed reaction and produces 4-diphosphocytidyl-2C-methyl-D-erythritol (CDP-ME) and pyrophosphate as a bi-product. Crystal structure of IspD protein is known from several organisms such as *M. tuberculosis* (Björkelid et al., 2011) *Thermatoga maritima*, *Neisseria gonorrhoeae* (Berman et al., 2000), *T. thermophilus*, *Listeria monocytogenes* and *Arabidopsis thaliana* (Gabrielsen et al., 2006). The presence of a conserved CDP-ME synthase domain (GTA superfamily) is important for the functionality of the protein. Other than this, the enzyme follows a Rossmann fold arrangement which only allows the binding of nucleotide derivatives to the protein.

Importance of non-mevalonate pathway has been also studied to identify new herbicides with novel mode of action; therefore, a program has been initiated to identify inhibitors of this pathway in plants (Rohdich et al., 2000). To identify the probable inhibitors, chemical compound from the BASF database were tested against IspD enzyme of *A. thaliana* and with the help of high through put screening (HTS) (Illarionova et al., 2006) inhibitors acting at low concentration were taken into further consideration. In this study, 7-hydroxy-[1,2,4] triazolo[1,5-a] pyrimidine, an Azolopyrimidine, was found to be most effective at low concentration with IC_{50} value of 140 ± 10nM (Witschel et al., 2011). Binding of this compound was analysed with crystal structure of *A. thaliana* IspD, but interestingly it binds to the allosteric site, thus modifying the size of enzyme's active site. This inhibitor has an IC_{50} of 35 ± 7nM and shows strong binding affinity with R157, I177, A202, I240, V263, I265 and V266 amino acid residues (Witschel et al., 2011). In order to identify and develop

new antimalarial, a database called Malaria box (Spangenberg et al., 2013) has been assembled jointly by the Medicines for Malaria Venture (MMV) and SCYNEXIS where 400 compounds are listed as potential antimalarial compounds. Several factors were considered for the selection of these compounds including their low toxicity, oral bioavailability, chemical diversity. These compounds were further screened against intra- erythrocytic stages of *P. falciparum*. Among these, MMV008138 compound has shown >95% growth inhibition and found to fulfil all the criteria. Initially, the mechanisms of action and efficacy of this compound against different development stages of the parasite was unknown. Their activity against MEP pathway was confirmed by reversing the parasite inhibition by IPP supplementation. The parasite showed 60 % growth recovery in the presence of 5µM of inhibitor and 200µM of IPP (Yao et al., 2015). Rescue of MMV008138 inhibited parasite with IPP suggests that it interferes with the function of the apicoplast and mainly acts against MEP pathway.

Further, to identify the target molecule of MMV008138, resistance strains of *P. falciparum* were developed for this drug where IspD enzyme was identified as the target of this drug. During this, two mutations E688Q and L244I in *Pf*IspD enzyme were identified responsible for the resistance against MMV008138 (Wu et al., 2015). Binding affinity of this drug has shown competitive inhibition with CTP substrate *in vitro*. To check the interacting residues between the drug and enzyme, Imlay and colleagues (2015) generated a homology model for the protein where they predicted the drug to interact with amino acid residue R208, K215 and T664. However, as these residues were different and falling away from the CTP binding site, thus the drug showing competitive inhibition with CTP raises concern. In addition, the activity of MMV008138 was higher in *P. falciparum* whereas in *P. vivax*, it is effective only at the lower concentration of CTP substrate (Imlay et al., 2015). Thus further investigation is needed for using this drug as a common target for both *P. falciparum* and *P. vivax*.

IspE [4-(cytidine-5-diphospho)-2-C-methyl-D-erythritol kinase (CMK)]

The fourth step of the pathway is catalyzed by IspE (EC 2.7.1.148) enzyme, an ATP and Mg^{2+} dependent enzyme belonging to the ATP – dependent GHMP kinase super-family which includes galactose kinase, homoserine kinase, mevalonate kinase, and phosphomevalonate kinase (Andreassi and Leyh, 2004). It catalyzes the conversion of CDP-ME to 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate (CDP-ME2P) in an ATP-dependent manner. Crystal structure of IspE has been elucidated from *E. coli*, *Thermus*

thermophilus HB8 (Wada et al., 2003). The IspE enzyme exists as a monomer and displays the characteristic two domain fold of the GHMP kinase superfamily, consisting of an ATP domain and a substrate-binding domain. There are three known conserved motifs in the GHMP kinase superfamily that are involved in creating the catalytic center (Eoh et al., 2009; Shan et al., 2011). Motif A (Lys13 to Leu18) involved in forming the substrate binding site; Motif B (Gly102 to Ser107), a glycine-rich phosphate binding loop that interacts with the triphosphate moiety of ATP and Motif C (Val254 to Gly258), which helps to stabilize the conformation of motifs A and B rather than interacting with ligands directly. Active site of IspE is enclosed in a deep cleft between these two domains where N-terminal domain folds into a phosphate binding loop for the binding of nucleotide (ATP) and C-terminal domain for the binding of substrate (Rohdich et al., 2000). In addition, an additional small, hydrophobic pocket lies adjacent to the CDP-binding site lined by amino acids Leu14, Ile27, Tyr175 and Leu208 (Hirsch et al., 2007).

Screening was done to find a suitable inhibitor for IspE enzyme in *E. coli*, where synthetically designed Ethyl {3-[4-amino-5-{3-[(cyclopropylsulfonyl) amino] prop-1-yn-1-yl}-2-oxopyrimidin-1(2H)-yl] oxetan-3-yl} acetate was found to be the most promising candidate with an IC_{50} of 590 ± 10nM (Hirsch et al., 2007, 2008). Further, the co-crystallization studies of IspE enzyme along with the inhibitor in *E. coli* and *A. aeolicus* confirmed that the inhibitor fits properly in the cytidine binding pocket of the enzyme where the cyclopropyl substituent of the sulfone moiety occupies the small cavity not used by the substrate. In *P. falciparum*, Leu28 is replaced by valine which can be considered to be of comparable steric demand and highly hydrophobic Phe185 in *E. coli* IspE is replacd with a more hydrophilic tyrosine residue in *M. tuberculosis*, *P. falciparum* and *A. aeolicus* affecting the binding characteristics of this pocket, by strongly reducing its hydrophobic character. Thus, the cyclopropyl ring may not locate to this subpocket and indeed, the cyclopropyl ring prevents any solvation of the phenolic hydroxyl group of Tyr175. This explains the reduced inhibitory activity of above compound against *A. aeolicus* and *P. falciparum* IspE. Another compound which has shown inhibitory activity against *Pf*IspE is 1,3-diiminoisoindoline carbohydrazide with an IC_{50} value <100nM in cell based assay. Synthesis of a variety of derivatives allowed an improvement of the initial antimalarial activity down to IC_{50} = 18nM for the most potent compound (Mombelli et al., 2012). However, till date none of the tested derivatives have shown any inhibitory activity against

P. falciparum IspE *in vitro* below 100µM (Masini and Hirsch, 2014), which demands further investigation into this enzyme.

IspF [2C-Methyl-D-erythritol-2, 4-cyclodiphosphate synthase (ygbB)]

IspF (EC 4.6.1.12) catalyzes the conversion of CDP-ME-2P into 2-*C*-methyl-D-erythritol-2,4-cyclodiphosphate (MECP), and like IspD, is dependent on divalent cations like Zn^{2+} or Mn^{2+} ions for its activity. Three dimensional structures for IspF from *E. coli* (Richard et al., 2001; Steinbacher et al., 2002), *A. thaliana* (Calisto et al., 2007), *H. influenzae* (Lehmann et al., 2002), *Mycobacterium smegmatis* (Buetow et al., 2007) and *P. falciparum* (O'Rourke et al., 2014) have been determined, showing that the proteins adopt a homotrimeric quaternary structure. The active sites are located at the interface of two monomer units where the pocket involved in binding the phosphate moiety of the substrate is capped with a flexible loop that becomes completely ordered when the reaction product is bound. In some of the IspF trimmers a central hydrophobic cavity is observed where the downstream products of the pathway like geranyl- or farnesyl-pyrophosphate have been observed which suggests a possible feedback role for the enzyme.

Despite the fact that the active site of IspF is considered the most druggable based on the presence of high apolar amino acid residues, very few inhibitors have been reported. In *E. coli*, similar to IspE, several chemical compounds have been synthesised and inhibition studies were performed against IspF enzyme, where a fluorescent inhibitor Diammonium 5'-O-{[({[2-({[5-(Dimethylamino) naphthalene-1- yl]sulfonyl}amino) ethyl] oxy} phosphinato)oxy] phosphinato} cytidene was found to be effective (Crane et al., 2006). The most successful inhibitors till date for IspF are the non-cytidine-like thiazolopyrimidines derivative (Geist et al., 2010), showing activity against both *P. falciparum* (IC_{50} = 9.6µM) and *M. tuberculosis* IspF (IC_{50} = 6.1µM) in micromolars and the aryl bis- sulphonamide inhibitors (Thelemann et al., 2015) showing inhibition of *Pf*IspF and *At*IspF with IC_{50} values as low as 1.4µM for *Pf*IspF and 240nM for *At*IspF. However, no information about their binding mode is available, and their optimization has not been reported to date.

IspG [4-Hydroxy-3-methyl-2-(E)-butenyl-4-diphosphate synthase (gcpE)] and IspH [4-Hydroxy-3-methyl-2-(E)-butenyl-4-diphosphate reductase (lytB)]

IspG and IspH enzymes participate in the last two steps of pathway respectively. Activity of both IspG and IspH enzymes is highly dependent on the redox environment and Fe-S clusters

(Rohrich et al., 2005; Partow et al., 2012). Initially, IspG catalyses the reduction of MECP through a multistep reaction and converts it into 4-hydroxy-3-methyl-2-(*E*)-butenyl-4-diphosphate (HMBDP) and then IspH converts it to the IPP and DMAPP (Rohdich et al., 2001). IspG enzyme contains two conserved domains: one is TIM barrel domain (for the binding of MECP substrate) and other is C-terminal domain (for Fe-S clusters), whereas IspH contains LytB domain (for the binding of HMBDP) and a Fe-S cluster binding domain similar to IspG. In *E. coli*, the electrons are required for the binding of MECP substrate to the IspG enzyme and provided by flavodoxin/ flavodoxin reductase/ NADPH system through [4Fe-4S] clusters. Similar to IspG, LytB of *E. coli* is catalytically active in presence of flavodoxin/ flavodoxin reductase/ NADPH system (Gräwert et al., 2004). However, in *P. falciparum*, instead of flavodoxin, ferredoxin/ ferredoxin $NADP^+$ reductase act as the physiological electron donor for LytB and as IspG is having similar requirements as IspH, it is assumed that IspG also use ferredoxin/ ferredoxin $NADP^+$ reductase system (Rohrich et al., 2005). This binding of the MECP substrate in the presence of [4Fe-4S] cluster, where the electrons are transferred is responsible for the formation of a double bond which converts cyclic form of the MECP molecule to aliphatic HMBDP. The HMBDP formed by IspG is then converted to the IPP and DMAPP by IspH enzyme and this conversion consists of three steps (i) removal of a hydroxyl group, (ii) transfer of two electrons from the [4Fe-4S] cluster, and (iii) the protonation of an intermediate allylic anion (Laupitz et al., 2004). Specific requirements for the activity of IspG and IspH enzyme was proven by a study in *Saccharomyces cerevisiae*, where even after incorporation of all the necessary components including all the enzymes, substrates and cofactors, the pathway remained non-functional. Finally, it was concluded that a particular reducing environment and compartmentalization is required for the functionality of these enzymes (Partow et al., 2012).

Most of the inhibitors developed so far against IspG and IspH target the Fe-S clusters, but as most of the metalloproteins in mammals including humans require these Fe-S clusters for their activity, these inhibitor have selectivity issues. Substrate analogoues that bind to IspG and IspH have been tried, but they could not turn out to be potent inhibitors. Replacing the diphosphate group with another moiety like carbamates or aminosulfonyl carbamates showed only very weak inhibitory activity (Van Hoof et al., 2008). Recently, the derivatives of diphosphonate such as alkyl phosphate have been identified as potential inhibitors for IspG and IspH enzymes (Guerra et al., 2014), however the inhibition mechanism is yet to be detailed.

The final products of MEP pathway in the apicoplast i.e. 5-Carbon Isoprenoids: IPP and DMAPP are believed to be exported out into the parasite cytosol where they are used as precursors for the synthesis of a wide variety of lipids having important functions in membrane structure, protein modification, metabolism, and hormone signalling. In apicomplexan parasites including *Plasmodium*, IPP units play an important role in the modification of tRNAs and translation processes in the apicoplast (Cassera et al., 2004). Other than these, there are evidences from pathogenic bacteria, that the MEP pathway may be involved in combating oxidative stress. In mycobacteria and corynebacteria, there is accumulation of 2C-methyl-D-erythritol-2,4-cyclopyrophosphate (MECP), an intermediate of the MEP pathway, in response to oxidative stress conditions (Ostrovsky et al., 1998). The mutants (DlytB and Ddxs) of the MEP pathway generated for *M. tuberculosis* have shown impaired ability to survive within the phagosome, thus suggesting the role of this pathway to combat intracellular stress in this bacterium (Pethe et al., 2004). Thus, isoprenoids or derivatives of isoprenoids help the pathogen to adapt to the host environment through enhanced stress resistance.

1.6.2 Products of Isoprenoid Biosynthesis Pathways: 10-, 15-, and 20-carbon isoprenoids

Sesquiterpenes and Diterpenes

In the cytosol of *Plasmodium*, two 5C isoprene units undergo condensation to form a 10-carbon isoprenoid geranyl pyrophosphate (GPP). This GPP then condenses with IPP to produce Farnesyl pyrophosphate (FPP) which further reacts with IPP and is converted to geranylgeranyl diphosphate (GGPP). This complete process is taken care of by a bi-functional enzyme FPP-GGPP synthase (Jordao et al., 2013), which is sensitive to Zoledronate and Risedronate (Jordao et al., 2011). Sesquiterpenes and diterpenes are the major compounds synthesized from Farnesyl Pyrophosphate (FPP) and Gernayl – Gernanyl Pyrophospahate (GGPP). These compounds are further utilized in protein prenylation process in *P. falciparum* which results in protein association with membranes. Prenylation is essential for the functionality of various membrane bound enzymes, such as the Ras, Rho, and Rab families of small GTPases. *P. falciparum* has been found sensitive to chemical inhibition of protein prenylation by inhibitors like Limonene where, on administration of the

drug, the parasites are unable to progress from the ring to the trophozoites stage (Moura et al., 2001).

Carotenoids

Carotenoids are 40-carbon chain isoprenoid compounds synthesized by plants, algae, as well as some bacteria and fungi. In *P. falciparum*, carotenoids are observed in the intra-erythrocytic stages where, its synthesis starts during ring stage and highest concentrations are achieved during the schizont stage (Tonhosolo et al., 2009). Carotenoids are synthesized in a sequential manner where first two GGPP molecules react to produce phytoene in a reaction catalysed by the enzyme phytoene synthase. The phytoene formed is finally converted to carotenoids by another enzyme called phytoene desaturase. A herbicide norflurazon which inhibits phytoene desaturase causes the accumulation of phytoene and decreases the level of carotenoids (Tonhosolo et al., 2009). Functions of carotenoids are well defined in plants, algae, bacteria, and fungi but no details are available for their role in *Plasmodium*. However, another apicomplexan *T. gondii* has been shown to produce abscisic acid from carotenoids to control calcium signalling for processes like protein secretion and parasite egress. While the abscisic acid response genes are conserved between *T. gondii* and *P. falciparum*, it is not clear whether *P. falciparum* synthesizes this isoprenoid product.

Coenzyme Q (Ubiquinone)

In most eukaryotes, Coenzyme Q serves as an electron acceptor in the electron transport chain for the ATP production in mitochondria. However, in *Plasmodium*, as the parasite relies mainly on glycolysis for its energy requirements, the role of coenzyme Q in ETC is mainly required during the mosquito stages. In asexual-stage of *Plasmodium* parasites, the electron transport chain operates to provide a continuous supply of reduced coenzyme Q as an electron acceptor for dihydroorotate dehydrogenase (DHODH), an enzyme required for pyrimidine synthesis. DHODH is essential for the parasite survival and DHODH inhibitors are shown to have potent antimalarial activity. Ubiquinone levels have been shown to peak at the beginning of schizogony and are sensitive to fosmidomycin (Cassera et al., 2004).

Coenzyme Q synthesis requires the addition of an isoprenyl side chain to a benzoquinone ring which is accomplished by the help of octaprenyl pyrophosphate synthase enzyme. This multifunctional enzyme produces different lengths of isoprenoid products ranging from 40-55 carbons (Tonhosolo et al., 2005; Tonhosolo et al., 2009). Nerolidol, a sesquiterpene

alcohol, was found to inhibit the synthesis of the isoprenyl side chain destined for coenzyme Q, likely because of its structural similarity to FPP. Treatment with nerolidol inhibits the intraerythrocytic development of *P. falciparum* (de Macedo et al., 2002; Tonhosolo et al., 2005).

Dolichols

Dolichols are long-chain hydrocarbon compounds made of various number of isoprene units and are essential for the transfer of sugars onto proteins, i.e., dolichylation, N-linked glycosylation and production of GPI anchors, essential for infection by the parasite (Cowman and Crab, 2006). In *P. falciparum*, 11 to 12 isoprene units long dolichols (formed from FPP and GPP) and their intermediates have been reported (Couto et al., 1999; D'Alexandri et al., 2006), which are synthesized by polyprenol reductase and dolichyldiphosphatase enzymes (D'Alexandri et al., 2006). The GPI anchors synthesized are utilized for the modifications of proteins like merozoite surface antigens, a serine protease and a heat shock 70 protein.

Other than these compounds, several other isoprenoids products like monoterpenes and sterols are synthesised in different organisms but they are not present in *Plasmodium*. Absence of sterols synthesis process in *P. falciparum* was confirmed by radioactive labelling (Vial et al., 1984; Mbaya et al., 1990) and homology search for squalene synthase enzyme (Rock, 1971). In this situation, *P. knowlesi* and *P. falciparum* import these compounds from the host to fulfil the requirement of sterols and cholesterol (Trigg, 1968; Labaied et al., 2011).

1.7 Scope of Isoprenoids Biosynthesis Pathway and Gaps in Existing Research

Rapid development of parasite resistance against the available chemotherapeutic agents and the antimalarial drugs is a major concern which poses a threat for the eradication of malaria. Till date, there is no effective vaccine against any human malarial parasites due to the significant genetic polymorphism exhibited by its surface antigens, which imposes a major hurdle in anti-malarial vaccine development.

In this situation, apicoplast an organelle of prokaryotic origin, and the metabolic pathways residing in it are being looked upon as a suitable target for the development of chemotherapeutic agents. Enzymes participating in these metabolic pathways are mostly encoded by the nuclear genome and targeted to the apicoplast. These enzymes are highly conserved across different *Plasmodium* species and apicomplexans and cannot be substituted by any other component, making these pathways indispensable for the parasite survival.

One of these pathways, the isoprenoids biosynthesis pathway is being considered as the most promising drug target to treat malaria as it is required for the survival of the parasite in both the hepatic and erythrocytic stages of asexual life cycle of the parasite. Thus, the enzymes involved in this pathway are being considered as attractive targets for researchers to develop new drugs or antibiotics against malaria. Therefore, a detailed understanding of this pathway and the enzymes participating in the pathway is necessary to develop an effective treatment strategy.

1.8 Aims and Objectives

- Molecular and biochemical characterization of isoprenoids biosynthesis pathway enzymes IspD and IspG from *Plasmodium* species.
- Expression profiling and immune-localization of *Pv*IspD and *Pv*IspG enzymes from field isolates.
- *In-silico* and *in-vitro* studies for structural validation of IspD and IspG enzymes as antimalarial drug targets.

1.9 Organization

This work aims to characterize novel enzymes IspD and IspG involved in Non-mevalonate isoprenoids biosynthesis pathway from *Plasmodium* and perform a comparative analysis with other *Plasmodium* species and apicomplexans to establish a common drug target. The thesis is organized according to the aims set as above. Chapter 1 is an introduction of the problem based on literature available till date, Chapter 2 gives general information related with various materials and methods used to achieve the proposed objectives, Chapter 3 & 4 details the various *in-vitro* and *in-silico* studies performed to characterize the IspD enzyme at the structural and functional level. Chapter 5 & 6 details the various *in-vitro* and *in-silico* studies performed to characterize the IspG enzyme at the structural and functional level. Chapter 7 lists the conclusions drawn out of the complete work.

Chapter II

Materials and Methods

To achieve the said objectives, various techniques were adopted based on the exhaustive literature survey. The protocols, materials (reagents and kits) and equipments used for characterizing the isoprenoids biosynthesis pathway enzymes IspD and IspG are detailed in this chapter. The chapter starts by detailing the approach for shortlisting of the proteins for the study. Further, the chapter is divided in two major subparts: the first part details various laboratory experiments performed, including the blood sample collection from the malaria infected patients, parasite DNA/ RNA isolation, confirmation of the type of parasite infection by PCR, amplification of isoprenoids biosynthesis pathway genes, their cloning and expression. Later in this subpart, we have detailed various biochemical and enzymatic assays along with the sub-cellular immune-localization studies. The results of these experiments are discussed at length in the later chapters (Chapter 3 and 5).

The second part of the chapter details about the *in-silico* characterization of *Pv*IspD and *Pv*IspG proteins using different bio-informatics tool. This includes the comparative modelling of the two enzymes and docking of their respective substrates and probable inhibitors to the obtained structure. All the structures were refined and validated with various freely available online softwares. The results are discussed in chapters 4 and 6.

2.1 Short listing of genes involved in Isoprenoids biosynthesis pathway

In *Plasmodium*, Isoprenoids biosynthesis pathway is believed to be functional inside the organelle known as apicoplast. Even when apicoplast possess its own DNA, most of the proteins involved in this pathway are synthesized by the parasite nucleus and targeted to the apicoplast by the help of a bi-partite N-terminal leader sequence containing a signal and transit peptide. Thus, the proteins involved in this pathway belong to the category known as NEAT proteins (Parsons et al., 2007). There are over 300 proteins in *Plasmodium vivax* (Carlton et al., 2008) and over 500 proteins in *Plasmodium falciparum* (Gardner et al., 2002) which are believed to be targeted to the apicoplast to participate in various housekeeping and metabolic functions.

In the present study, the proteins hypothesized to be functional in the Isoprenoids biosynthesis pathway were identified from the database based on the presence of conserved domains and signature motifs as predicted by **Conserved Domain Search** tool from **NCBI** (Marchler-Bauer et al., 2015) and **Prosite** from ExPASy (Sigrist et al., 2002), respectively. Their targeting to different organelles was further predicted using various online servers like **SignalP 4.1** (Petersen et al., 2011), **PlasmoAP** (Foth et al., 2003), **PATS** (Zuegge et al., 2001), **MitoProtII** (Claros and Vincens, 1996) and **PlasMit** (Bender et al., 2003).

SignalP 4.1 is a server which utilizes the neural networks to predict the presence and location of signal peptide in different organisms that destines the protein to the secretory pathway through endoplasmic reticulum. It gives three output scores for each position – *C-score* (raw cleavage site score), *S-score* (signal peptide score) and *Y-score* (combined cleavage site score). A combination (geometric average) of the *C-score* and the slope of the *S-score* helps in the cleavage site prediction whereas *Y-score* distinguishes between *C-score* peaks by choosing the one where the slope of the *S-score* is steep. Based on these scores a cumulative score is generated which helps in the signal sequence prediction.

In *Plasmodium* proteins, the signal peptide is followed by a transit peptide which directs the protein to either the apicoplast or mitochondria. To check for the presence of apicoplast transit peptide, PlasmoAP and PATS were used. PlasmoAP is based on an algorithm that uses amino-acid frequency and distribution to identify putative apicoplast-targeting peptides. PATS uses the neural network, which has been trained with known apicoplast targeting and non-targeting proteins in *P. falciparum* to determine the apicoplast transit peptide. Mitochondrial targeting was analysed using servers like MitoProtII and PlasMit, which calculates the N-terminal protein

region that can support a mitochondrial targeting sequence and the cleavage site. While the PlasMit server is specifically meant for the prediction of mitochondrial targeted proteins of *Plasmodium,* MitoProtII can be employed to the proteins from any organism. Targeting to mitochondria was also analysed for proteins having apicoplast transit peptides to further validate if the protein is dual targeted to both the organelles.

2.2 Study site and Malaria infected blood sample collection

P. vivax and *P. falciparum* infected patient's blood samples (3 – 6mL) were collected in acid citrate dextrose (ACD) anticoagulant by trained clinicians at P. B. M. and associated group of Hospitals, Bikaner, Rajasthan, India. The infections in these samples were clinically detected by standard Microscopy and Rapid Diagnostic Tests (OptiMAL® and Falcivax). A formal approval of participating Institute's Human Ethics Committees' (Approval No. IHEC-29/12-13) and consent of patients was obtained for further studies. The blood samples were pre-treated separately for DNA and RNA isolation. These samples were then shipped to our lab at BITS Pilani in cold chain where they were stored at -70°C till further use.

2.3 Parasite DNA isolation

Parasite DNA was extracted from blood samples collected from study site using standard protocols with slight modification (Sambrook and Russell, 2001; Saxena et al., 2007).

Reagents

1. Lysis buffer (1X): 10mM NaCl; 50mM Tris (pH=8.0); 10mM EDTA (pH=8.0); 1% SDS
2. Lysozyme (50mg/mL) (Ameresco)
3. Proteinase K (20mg/mL) (Merck Bioscience)
4. RNase A (20mg/mL) (Merck Bioscience)
5. 3M Sodium Acetate (pH 5.2)
6. Tris-saturated Phenol (pH 8.0)
7. Chloroform : Isoamyl alcohol (24:1)
8. Absolute alcohol (99.5%) and 70% Ethanol

The *Plasmodium* parasite infected frozen blood samples were thawed at room temperature, and transferred into Oakridge tube. The extraction procedure was based upon the lysis of cells with equal volume of 1X lysis buffer. Content was mixed properly by repeated inversion and pelleted down by centrifugation at 12,000 rpm for 10 min at 4°C (Eppendorf, Germany). Supernatant was decanted and the pellet was re-suspended in 2X lysis buffer. RNase A was added to the final

concentration of 20µg/mL and the sample was incubated at 37°C for 15 min. Following this lysozyme was added (final concentration of 100µg/mL) and the sample was incubated at 37°C for 45 min in a water bath. To digest the protein, Proteinase K was added to the final concentration of 100µg/mL and incubated at 50°C with frequent swirl up (water bath) for 2-2.5 hrs. A conventional phenol: chloroform: iso-amyl alcohol (25:24:1) extraction was performed to remove the proteins and the DNA was precipitated from the aqueous phase using chilled absolute ethanol in the presence of 0.3M sodium acetate at -20°C overnight. The precipitated DNA was pelleted by centrifuging the tubes at 8,000 rpm for 30 min at 4°C. The pellet was washed with 70% ethanol to remove the excess of salts, air dried and re-dissolved in 1X TE buffer (pH 8.0). Integrity of DNA was checked using agarose gel electrophoresis and the DNA fragments were viewed in UV gel documentation (Syngene USA) (**Figure 2.1**). Purity of DNA was checked by the Nano UV-vis Spectrophotometer (Simpli Nano GE, UK)

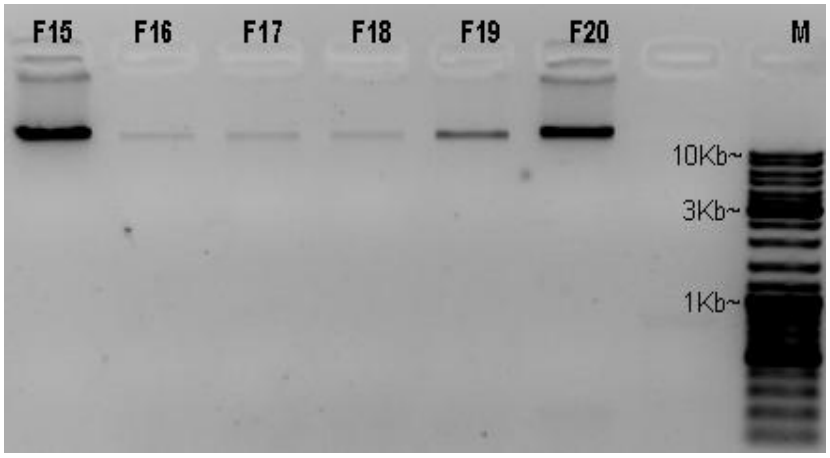


Figure 2.1: DNA isolation from the patient's blood infected with *Plasmodium* [M= Gene ruler DNA Ladder mix (SM0331 Fermentas); F15 to F20 are the field isolates].

2.4 Confirmation of parasite infection by PCR

In areas like Bikaner, where both *P. falciparum* and *P. vivax* infection coexists, there are chances of misdiagnosis of parasite infection. To confirm the type of infection as *P. falciparum* or *P. vivax* or a combination of both, a diagnostic PCR based on 18S rRNA gene (Das et al., 1995; Kochar et al., 2005) and 28S rRNA gene (Pakalapati et al., 2013) amplification was performed.

2.4.1 Diagnostic PCR based on 18S rRNA gene

It is a multiplex PCR, which involves a single forward universal primer for the genus *Plasmodium* and two reverse species specific primers for *P. vivax* and *P. falciparum* (**Table 2.1**). The reaction conditions utilized are shown in **Figure 2.2**. The presence of *P. falciparum* was

indicated by a ~1400bp amplicon whereas, *P. vivax* shows a ~500bp amplicon (**Figure 2.3**) (Das et al., 1995; Kochar et al., 2005).

Table 2.1 Diagnostic Primers for 18S rRNA gene amplification by Multiplex PCR

Primer sequence	Detection	Orientation
5' ATC AGC TTT TGA TGT TAG G GT ATT G 3'	Genus Specific	Forward
5' TAA CAA GGA CTT CCA AGC C 3'	*P. vivax*	Reverse
5' GCT CAA AGA TAC AAA TAT AAG C 3'	*P. falciparum*	Reverse

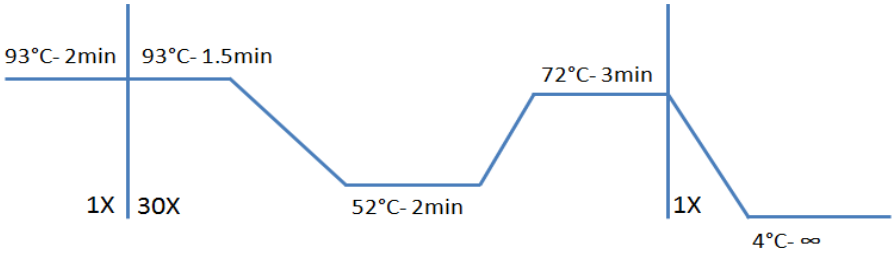


Figure 2.2: Reaction conditions used in multiplex PCR

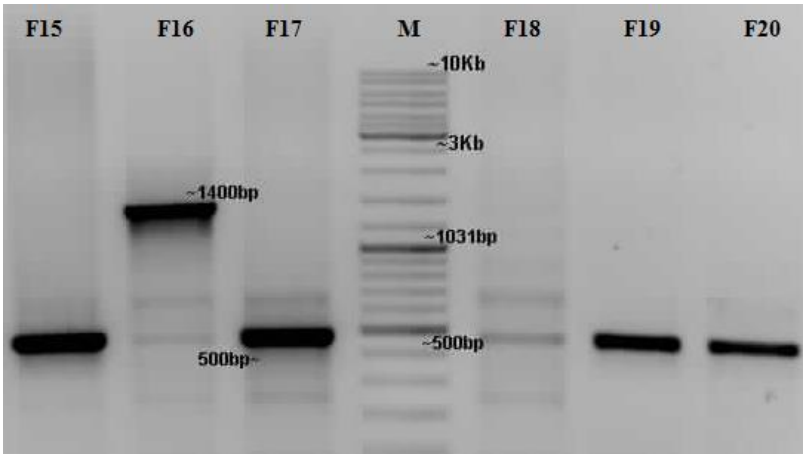


Figure 2.3: Parasite infection confirmed by multiplex PCR based on 18S rRNA gene amplification [M= Gene ruler DNA Ladder mix (SM0331 Fermentas). F15 to F20 are the field isolates; F15, F17, F19 & F20 = *P. vivax* infected samples; F16 = *P. falciparum* infected sample; F18 = not included in study].

2.4.2 Nested PCR based on 28S rRNA

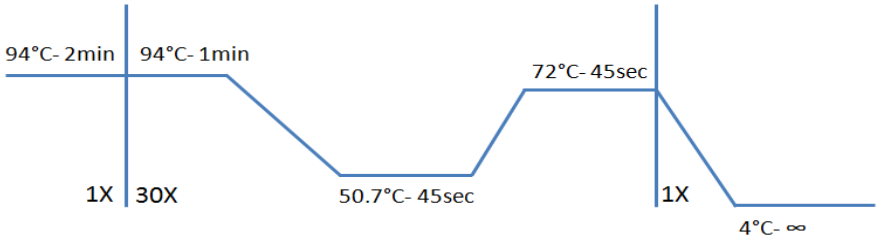
This PCR involves a primary reaction that serves as a positive control while the second nested PCR finally diagnoses the type of infection (Pakalapati et al., 2013). Primers used in the primary reaction are designed from a region common to 28S rRNA gene of human, *P. falciparum* and *P. vivax* and gives an amplification of ~790bp. For the nested amplification, 0.5µL of the amplicon from the above primary reaction is used as a template DNA. The nested PCR 1 using primers NPV2 and NPVR amplifies ~294bp region specific for *P. vivax* 28S rRNA gene, while the nested PCR 2 amplifies ~286bp region specific for *P. falciparum* 28S rRNA gene. The primers used for

these primary and nested PCR are given in **Table 2.2** while the reaction conditions utilized are shown in **Figure 2.4**. The samples which showed the band at the desired position were utilized for further studies (**Figure 2.5**).

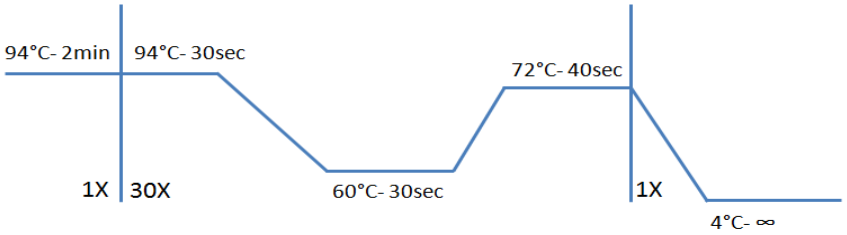
Table 2.2 Diagnostic Primers for 28S rRNA gene amplification by Nested PCR

Lab Id	Primer sequence	Detection	Orientation
NUF	5' GAT TTC TGC CCA GTG CTT TGA ATG T 3'	Human, Pf & Pv	Forward
NUR	5'AAT GAT AGG AAG AGC CGA CAT CGA A 3'	Human, Pf & Pv	Reverse
NPV2	5' TCG GTT CGC CGG GTA TTC ATA TT 3'	*P. vivax*	Forward
NPVR	5' CAC AGT AGG AAG ATA AAT TCC T 3'	*P. vivax*	Reverse
NPF1	5' TAT CCT TCG GGA AGG CAT TCT G 3'	*P. falciparum*	Forward
NPF2	5' CTA TAT GCA CAG TAG TAA GTA ATT TA 3'	*P. falciparum*	Reverse

a) Primary Reaction condition:-



b) Nested PCR 1 (*P. vivax* specific) Reaction conditions:-



c) Nested PCR 2 (*P. falciparum*) Reaction conditions: -

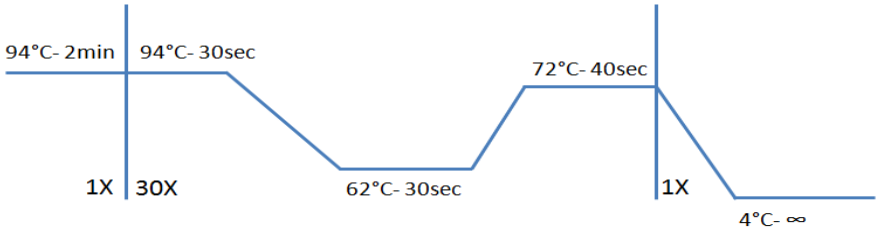


Figure 2.4: Reaction conditions used for primary and nested PCR

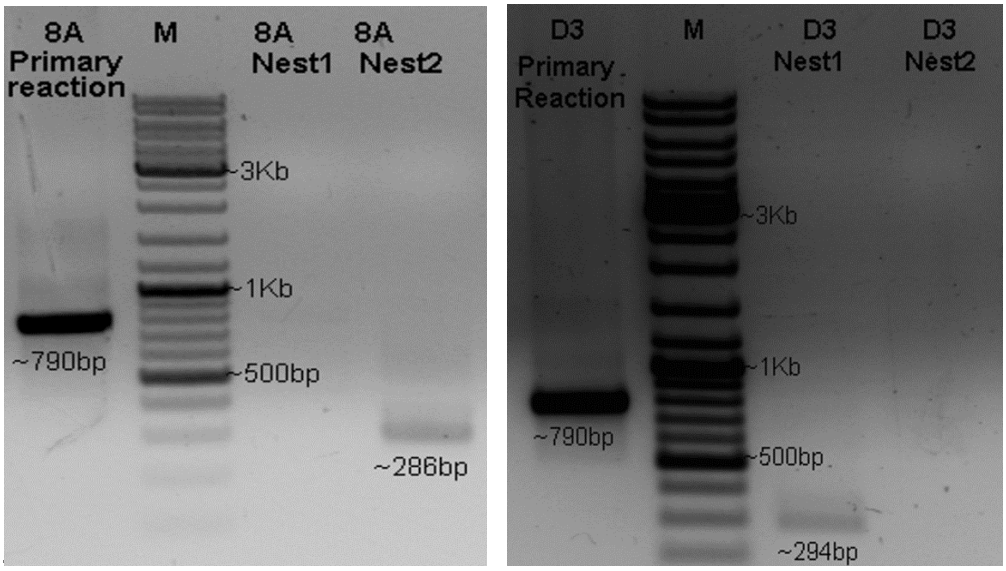


Figure 2.5: Parasite infection confirmed by Nested PCR based on 28S rRNA gene amplification [M= Gene ruler DNA Ladder mix (SM0331 Fermentas); 8A and D3 are parasite field isolates]

2.5 Amplification of genes involved in Isoprenoids biosynthesis pathway

2.5.1 Primer designing

Primers were designed for genes participating in isoprenoids biosynthesis pathway of *P. falciparum* and *P. vivax* using the sequences retrieved from online available database PlasmoDB (Bahl et al., 2003) and NCBI (Marchler-Bauer et al., 2015). For designing the primers, freely available online software's like Gene runner (Hasting Software's Inc., USA), Primer3 (Untergrasser et al., 2012) and Oligo (Rychlik, 2007) were used. The presence of any dimers and hairpin loop Tm were also taken into consideration. The primers were synthesized commercially and the Log Tm was calculated using the standard formula-

$$\text{Log Tm} = 81.5+16.6\,[\log_{10}(Na^{+})] + [41(G+C)/n] - 675/n$$

(n = number of bases)

2.5.2 Amplification and sequence analysis

Standard PCR was performed with minor modifications for amplification of genes encoding isoprenoids biosynthesis pathway enzymes: IspD and IspG. The primers and the PCR conditions used are given in the respective chapters in the thesis. The PCR amplicons obtained were commercially sequenced for both strands following Sanger's dideoxy method using ABI 3100 DNA sequencer version 5.1.1 (Applied Biosystems, USA). The obtained raw sequences were analyzed for contig alignments using DNAMAN (Lynnon Corp). To check the similarity between different orthologues and homologs, the sequences were aligned using CLUSTAL W (Larkin et al., 2007) and BioEdit (Hall, 1999).

2.6 Cloning of *ispG* and i*spD* genes in Prokaryotic expression system

To characterize the desired proteins in *P. vivax* Isoprenoids biosynthesis pathway, the full length gene amplicon obtained from *P. vivax* field isolates was cloned in different cloning and expression vectors. Most of the amplified fragments were having restriction sites at their ends as incorporated during primer designing for directional cloning. These restriction sites were incorporated based upon desired vector and respective multiple cloning site region. The complete cloning process is a multi-step process including competent cell preparation, plasmid isolation, processing of vector (plasmid) and insert (amplified gene) using restriction digestion, ligation, transformation and screening of recombinant clone.

2.6.1 Competent cell preparation

Desired strains of *E. coli* host cells were streaked on a LB Agar plate and grown overnight at 37°C. Thereafter, a single colony was picked and used to inoculate 5mL LB broth. The culture was allowed to grow overnight at 37°C with shaking at 150 rpm. The following day, 1 part of the overnight grown culture was used to inoculate 40 parts of 150mL fresh LB broth and again kept at 37°C, till an OD_{650} value of 0.5 was attained. After this, the culture was chilled on ice for 10 min and subsequently centrifuged at 8000 rpm, 4°C for 8 min. Supernatant was discarded, the pellet was re-suspended slowly in 80mL of ice-cold 0.1M calcium chloride ($CaCl_2$) and incubated on ice for 15 min after which it was again centrifuged at 8000 rpm, 4°C for 8 min. The supernatant was discarded and the pellet was re-suspended in 10mL of ice-cold $CaCl_2$, after which 1.4mL of 100% glycerol was added to the re-suspended cells while stirring gently. Cells were chilled on ice for 10 min and then distributed as 300µL aliquots in each micro centrifuge tubes. Immediately the tubes were closed and stored at -80°C for further use.

2.6.2 Plasmid DNA Isolation

Reagents

1. Solution I: 50mM Glucose; 25mM Tris Cl (pH=8.0); 10mM EDTA (pH=8.0)
2. Solution II: 0.2N NaOH; 1% SDS (should be prepared fresh each time).
3. Solution III: 3M Sodium acetate (pH=5.2; chilled)

Individual colonies from the bacterial culture agar plate were inoculated in 5mL LB broth containing appropriate antibiotic and allowed to grow for 16 hrs at 37°C, with shaking at 200 rpm. Plasmid DNA was prepared from these cultures following the Alkaline Lysis method given by Birnboim and Doly, 1979. Briefly, the overnight grown culture was centrifuged at 5000 rpm for 5 min. Supernatant was discarded and bacterial pellet was re-suspended in ice cold Solution I.

To this, freshly prepared Solution II was added and incubated at room temperature for a couple of minutes with intermittent mixing. To the above mix, chilled Solution III was added, mixed gently and incubated on ice for 5 min. The lysed material was then spun at 10,000 rpm for 10 min at 4°C and the supernatant was transferred to a fresh tube. A phenol: chloroform: isoamyl alcohol treatment was then given to remove all the proteins and the DNA was precipitated overnight by adding twice the volume of 100% ethanol in presence of Na – acetate. The precipitated DNA was pelleted by centrifuging the tubes at 10,000 rpm for 15 min at 4°C. The supernatant was discarded and the pellet was washed with chilled 70% ethanol. The pellet was air dried and suspended in adequate amount of 1X TE (Sambrook et al., 1989). The integrity of the plasmid DNA was checked by running on 1% agarose gel.

2.6.3 Preparation of vector and insert (gene) for cloning

The isolated plasmid vector was double digested with appropriate restriction enzymes and purified using QIAquick Gel Extraction kit (QIAGEN, Germany) following the manufacturer's protocol. The insert being a PCR product was purified using QIAquick PCR purification kit (QIAGEN, Germany) and double digested with same restriction enzymes (Sambrook et al., 1989). After this, digested insert was gel eluted and quantified for setting up a ligation.

Restriction digestion of vector and gene insert

Template DNA	=	500ng – 1µg
Restriction Enzyme	=	3 – 5U
Buffer (10X)	=	2µL
Sterile Millipore Water	=	Volume to make 20µL

The reaction was incubated at the desired temperature (mostly 37°C) for 1 – 2 hours.

Gel elution of vector and insert:

Purification of DNA fragments (PCR product and digested plasmid) was performed using QIAquick Gel extraction Kit (QIAGEN, Germany) following the manufacturer's protocol with minor modifications. DNA fragment was excised from gel using a clean scalpel and gel slices were melted in 3 volumes of QG buffer (w/v) by incubation at 50°C in water bath. One volume of isopropanol was added and incubated at RT for 2 min. To bind the DNA, sample was applied to QIAquick spin column and centrifuged at 13000 rpm for 1 min. Spin column was washed with 500µL QG buffer and 720µL PE buffer by centrifuging it at 13,000 rpm for 1 min. Spin column was transferred to fresh 1.5mL eppendorf and DNA was eluted by using 1X TE buffer by

centrifuging at 13,200 rpm for 1.5 min.

2.6.4 Ligation of insert with vector

Purified plasmid and insert after double digestion were ligated with the help of T4 DNA ligase enzyme. Equal quantity of the vector and insert were checked on gel prior to the ligation, to compare the concentration (**Figure 2.6**). Self-ligation of the vector only to check the frequency of false positives and ligation of double digested vector and insert was set as detailed below.

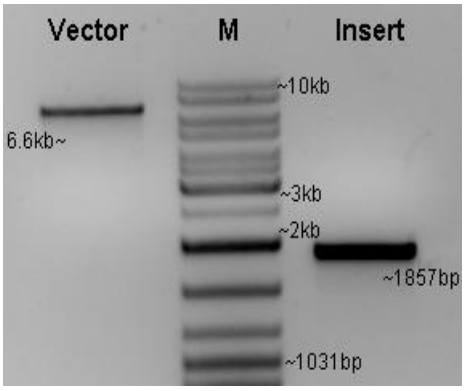


Figure 2.6: Vector and insert concentration check prior to ligation

M= Gene ruler DNA Ladder mix (SM0331 Fermentas).

Ligation mix:

Contents	Quantity
Vector (plasmid DNA)	100ng
Gene of interest	300ng
T4 DNA ligase buffer (10X)	2µL
T4 DNA ligase (5U/µL)	1µL
Sterile Deionized water q.s.	20µL

Reaction mix was incubated for 16 hrs at 16°C in a circulating water bath.

Following above reaction, the ligated product was transformed into *E. coli* DH5α competent cells. Briefly, competent cells (stored at -80°C) were thawed on ice and 10µL of ligated mix was added to 100µL of competent cells and mixed properly. Cells were incubated on ice for 25 min. The samples were thereafter given a heat shock in a water bath at 42°C for 90 seconds, and then immediately placed on ice, after which 800µL of LB broth medium, pre-warmed at 37°C, was added. This mixture was incubated at 37°C for 45 min followed by centrifugation at 3000 rpm for 5 min. 800µL of supernatant was removed and cells were resuspended in remaining 100µL of supernatant and plated on LB Agar plates containing the desired concentration of antibiotics.

Simultaneously, a digested vector without insert was also transformed in the same bacterial host to act as a control.

2.6.5 Identification of recombinant clones

After the transformation of digested vector and ligated product, the numbers of colonies obtained in above two sets were compared and checked for the recombinant molecules.

2.6.5.1 Gel shift assay

The colonies obtained after transformation of ligated product was initially analyzed using the gel shift assay (Sekar, 1987). Individual colonies were streaked on LB Agar-antibiotic plates and incubated overnight at 37°C. A small part of each colony was picked and mixed with lysis buffer (0.5mM EDTA; 10% Sucrose; 0.25% SDS; 100mM NaOH; 60mM KCl) in a micro-centrifuge tube. The tubes were further incubated stepwise, at 65°C (water bath) for 5 min, on ice for 5 min, at room temperature for 5 min and centrifuged at 8,000 rpm for 5 min. The obtained supernatant along with sample lysis buffer was loaded on the gel.

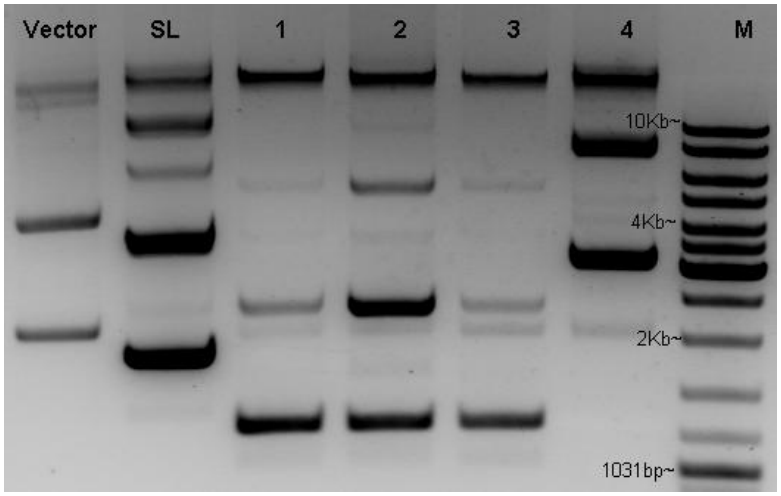


Figure 2.7: Gel shift pattern obtained after assay for clone identification [M= Gene ruler DNA Ladder mix (SM0331 Fermentas); SL: Pattern from colony having self ligated vector as a control; 1 to 4 different patterns from colonies obtained after the transformation of ligated product].

The presence of clone is identified by a shift in the plasmid bands due to the varied size of recombinant plasmid as compared to others or the standards (**Figure 2.7**). Colonies showing a shift in gel shift assays were grown in LB broth and recombinant plasmid was isolated by using the routine alkaline lysis method. The obtained plasmid was further analyzed by colony PCR and restriction digestion analysis using the multiple cloning site of vector or restriction sites present in the insert.

2.6.5.2 Colony PCR

Colonies shortlisted after Gel shift assay were screened by PCR using gene specific primers to select positive clones. Colonies were picked from the plate under sterile conditions, suspended in 50µL nuclease free water in a 0.5mL micro-centrifuge tube. Vortexing was done for proper suspension, followed by heating at 100°C for 4 min. After centrifugation 10µL of supernatant was used as a template for each reaction mixture. Colonies obtained from self-ligation reaction were used as a negative control. PCR reaction products were analyzed on a 1% agarose gel.

2.6.5.3 Restriction Digestion analysis

Further to confirm the desired clone, the recombinant plasmids isolated using Alkaline Lysis method (as discussed previously) were digested using one or more restriction enzymes to confirm the desired size of the clone and presence of single insert. Following this, confirmation of clone and chances of frame-shift were detected by sequencing of the recombinant plasmid using T7 promoter primers. These confirmed recombinant molecules were taken further for protein expression trials.

2.7 Protein Expression

Reagents

1. 1.5M Tris (SIGMA) (pH=8.8)
2. 0.5M Tris (pH=6.8)
3. 30% Acrylamide solution (29.2g Acrylamide and 0.8g N,N'-Methylene Bis-Acrylamide) (SIGMA)
4. 10% Ammonium Persulfate (SRL)
5. Sodium Dodecyl Sulfate (1% SDS)
6. Tetra methyl ethylene diamine (TEMED; SIGMA)
7. 2X Sample lysis buffer (Merck Bioscience, Germany)
8. Electrophoresis Running Buffer: 25mM Tris; 192mM Glycine; 0.1% SDS
9. Protein Molecular Weight Marker (Merck Bioscience, Germany)
10. Staining solution: 0.2% Coomassie Brilliant Blue R-250 (HiMedia); 5% Methanol; 10% Acetic acid
11. Destaining solution: 40% Methanol; 7% Acetic acid in deionized water

Competent cells prepared for *E. coli* expression host system were freshly transformed with recombinant plasmid containing the desired gene using heat shock method. A single colony was

picked from this transformed plate, inoculated in LB broth and culture was grown overnight at 37°C with shaking. After 14 hrs, 1 part of this primary culture was inoculated to 20 parts of fresh LB broth and incubated at 37°C with shaking. When this culture reaches to an OD of 0.5-0.6 at 600nm (~3hrs), 1mL aliquot was withdrawn as a control sample from the culture (marked as 0 hr) and remaining culture was induced with IPTG (final concentration 0.5mM) and again incubated at 37°C. Samples were collected from this induced culture at regular intervals (2hrs, 4hrs and 6hrs) and pelleted by centrifugation at 3,000 rpm for 5 min at 4°C. The obtained pellet was resuspended in 2X sample lysis buffer, heated at 100°C for 10 min and stored at -20°C till further use. Preheated samples were loaded on a denaturing polyacrylamide gel. The obtained gel was stained with Coomassie Brilliant Blue solution and protein bands were viewed after destaining the gel. Analysis of the gel was done where vector and clone samples were compared before and after induction.

2.8 Western blotting

To confirm the expression of desired protein, the protein samples were transferred to nitrocellulose membrane and the membrane was subjected to antibodies specific to these proteins or to the His-tag or MBP-tag fused to the protein. If the protein is present, the antibodies will bind to it and this can be checked by the addition of an enzyme conjugated secondary antibody (Horse-radish Peroxidase or Alkaline Phosphatase) specific against the primary antibody.

Reagents & Material:

1. Tris buffered saline (TBS): 50mM Tris; 150mM Sodium chloride (pH=7.5).
2. Transfer buffer: 48mM Tris; 39mM Glycine; 10-20% Methanol
3. Washing buffer: (1X TBS with 0.05% Tween20)
4. Blocking buffer: (5% Fat free milk powder (HiMedia) in 1X TBST)
   Anti-His antibodies (QIAGEN) / Anti-MBP monoclonal antibodies (NEB), Antibodies raised against specific proteins
6. Goat anti-mouse IgG HRP labeled antibodies (Merck Bioscience, Germany)
7. Substrate solution: TMB/ H_2O_2 (SIGMA)
8. Nitrocellulose membrane (Schleicher & Schuell, Germany)
9. Mini Trans-Blot® Cell apparatus (Bio-Rad, USA)

A pre-run and unstained SDS polyacrylamide gel was used for the transfer of protein on the nitrocellulose membrane. The nitrocellulose membrane, SDS PAGE gels and filter pads were soaked in the pre-cooled transfer buffer. Assembly was set with the gel at the negative end and the membrane at the positive end and proteins were transferred to the membrane by electro

transfer. After the transfer process, the membrane was stained with Ponceau-S (Merck Bioscience, Germany) to check the transfer efficiency and after confirmation of transfer of desired band; the Ponceau-S stain was removed by excessive washing with deionized water. To prevent the non-specific binding, the membrane was kept overnight in blocking buffer (5% fat free milk powder in 1X TBST) followed by washing thrice with washing buffer (1X TBST). After washing, the membrane was left soaked in primary antibody (dilution of antibodies varies according to the protein targeted) for 1½ hours at 37°C. The membrane was again washed thrice (as previously done) and incubated with secondary HRP conjugated antibody. Finally membrane was washed thrice with washing buffer and incubated for 15-20 min with TMB/H_2O_2 substrate solution in dark. The reaction was stopped by adding distilled water and bands obtained at membrane were analyzed.

2.9 Protein Purification

Once the protein was confirmed by western blotting, to perform further experiments like enzymatic assays, localization studies, etc., the protein was purified from other proteins using affinity chromatography. As the fusion tags used for our enzymes were either His-tag or MBP-tag, we used appropriate pre-packed/ manually packed columns for these purifications.

2.9.1 His-tagged protein purification

Reagents

1. Equilibration buffer: (25mM Tris Cl; 0.3M NaCl; 8M Urea; 20mM β-mercaptoethanol (pH 8.0)
2. Phenyl methyl sulphonyl fluoride: (PMSF; 100mM stock in Isopropanol) (SIGMA)
3. Lysozyme (50mg/mL) (Ameresco)
4. Wash buffer: [Equilibration buffer (pH 6.3)]
5. Elution buffer: [Equilibration buffer (pH 4.5)]

His-tagged protein was purified manually using Ni-NTA agarose resin (QIAGEN, Germany) or with the help of pre-packed column (His Trap FF, 5mL column) in Acta pure FPLC system (GE Healthcare, USA). After 6 hrs of IPTG induction, bacterial culture was harvested in conical bottom centrifuge tube by centrifugation (VS-550, Vision Scientific, Korea) at 3800xg for 15 min at 4°C. PMSF (final concentration 1mM) was added to the wet pellet to inhibit any protease activity so as to prevent the digestion of the target protein during protein solubilization. After this, the pellet was resuspended and maintained in equilibration buffer. Lysozyme (final

concentration of 1mg/mL) was added and sonication (MISONIX, USA) was performed on the cell lysate for further release of proteins from the bacterial cells, where sonication parameter includes output of 10 watts with 10 sec pulse thrice on ice. 1mL aliquot of prepared lysate was stored for total protein analysis and the rest of the sample was centrifuged at 9300g for 15 min at 4°C. Supernatant (soluble fraction) and pellet (insoluble fraction) were separated and the presence of desired protein was confirmed by running them on SDS PAGE.

Protein molecules present in soluble fraction were purified using Ni-NTA agarose resin (QIAGEN, Germany) affinity chromatography. Purification is based on electrostatic interaction between the 6X-His peptide tag attached to the solubilized protein and Ni^{2+} cation attached to the agarose resin. 2-3mg/mL soluble fraction of protein was mixed with agarose resin, applied to equilibrated polypropylene column and allowed to settle. Flow through of the sample was collected and the column was washed with wash buffer to remove unbound protein. Specific 6X-His tagged protein was eluted with the help of elution buffer in separate fractions and the fractions were run on a SDS PAGE to confirm the presence of desired protein.

To obtain the pure protein in higher concentrations, ACTA Pure FPLC His-Trap columns were used. Briefly, ACTA Pure system was washed once with deionized water followed by equilibration with 5 column volumes of equilibration buffer (pH 8.0). 10mL of protein lysate in the form of soluble fraction was applied to the column. The column was washed with 10 column volumes of wash buffer (pH 6.3) and protein was eluted using 10mL elution buffer (pH 4.5). *Pv*IspG protein concentration was determined by Bradford assay (Bradford, 1976) where known concentration of BSA was used as a standard.

2.9.2 MBP-tagged protein purification

Reagents

1. Column buffer: [20mM Tris Cl (pH 7.4); 0.2M NaCl; 1mM EDTA; 10mM β-mercaptoethanol]
2. 1M Maltose solution

The proteins expressed as fusion proteins with MBP were purified using the amylose resin provided with the pMAL™ Protein Fusion & Purification System (NEB, USA). For this the polypropylene columns provided in the kit were packed with the amylose resin and the standard protocol with slight modifications was followed. Briefly, protein from the crude lysate was

released in the column buffer by employing the sonication process and after centrifugation, supernatant was applied to the column containing amylose resin. Protein was eluted by using column buffer with varying concentration of maltose.

2.9.3 Protein purification by Gel elution

To elute the protein from the gel, the protocol given by Kurien and Scofield, 2012 was followed. For this, the SDS-PAGE of required porosity was casted. To load the sample with protein in excess, the wells of this gel (depending upon sample volume) were clubbed leaving three wells intact for control sample and marker. Electrophoresis was performed, after which wells with control protein and protein marker were cut and kept for normal staining and de-staining procedure while the remaining gel portion having protein for elution was preserved in running buffer at 4°C. Taking an estimate from the stained gel, the desired protein band was cut from the unstained gel and was directly placed in 1mL protein isolation buffer (50mM Tris-Cl (pH 7.5); 150mM NaCl; 1mM EDTA). Tubes were incubated at 30°C for 14-16 hrs with shaking at 800 rpm in thermo-shaker. After overnight incubation, tubes were spun at 12000 rpm for 20 min at 4°C and the supernatant was collected. Protein concentration was estimated by Bradford assay and samples were further concentrated using vacuum concentrator to achieve the desired concentration.

2.9.4 Protein dialysis

The presence of salt impurities in the above purified protein samples was removed by dialysis. Protein samples were precipitated by ethanol and resuspended in 1X PBS. Dialysis membrane of desired cut-off molecular weight (50 kDa) was cut according to the volume of the solution and pretreatment was given as prescribed by the manufacturer. Protein sample was added to the tubing, air bubbles were removed manually and tubing was closed with the help of closures. This tubing was placed in a container filled with dialysis buffer (1X PBS) where the volume of the buffer was 10-20 times to the volume of protein sample. Dialysis buffer was changed at regular interval of 4 hrs for 24-30 hrs. After dialysis, protein sample was removed from the tubing and analyzed by SDS-PAGE.

2.9.5 Protein concentration determination

Several methods are available for protein concentration determination, however the most famous Bradford method (Bradford, 1976) was selected because of its simplicity. This method relies on the electrostatic and hydrophobic interaction of sulfonic acid groups in Coomassie Brilliant Blue

(CBB) G250 dye to positively charged, aromatic, basic amino acids that occur in low pH value. After the binding of protein, the absorbance maxima of CBB dye shifts from 465 to 495nm, which was assessed using UV-visible spectrophotometer (Jasco 630, Japan). BSA at different concentrations ranging from 6.5µg to 200µg was used for plotting the standard curve for protein estimation. By extrapolating the absorbance obtained for the unknown protein sample on this standard curve, the concentration of the protein was determined.

2.10 MALDI-TOF-Mass Spectrometry Analysis

The utility of MALDI-TOF-Mass Spectrometry for protein and peptide analyses lies in its ability to provide highly accurate molecular weight information for intact molecules and sequence information for peptides. The ability to generate such accurate information can be extremely useful for protein identification and characterization. In our study, molecular mass of the protein and peptide sequence was identified commercially using 4800 plus MALDI TOF/TOF Analyzer (AB SCIEX, USA) equipped with nitrogen laser and operated in reflector and delay extraction mode. Purified proteins in the gel were digested with trypsin supplied with In-gel Tryptic digestion kit (Thermo Scientific, USA) using the manufacturer's protocol. 6µL of this eluted peptide solution was mixed with saturated α-cyano-4-hydroxycinnamic acid matrix (10mg/mL in 50% acetonitrile and 0.05% trifluoroacetic acid) for each spot. The samples were immediately vortexed to make a homogenous solution and 1µL of the above mixture was dropped onto the metal target plate. The peak was detected after calibrating the plate with 4700 proteomics analyzer calibration mix (AB SCIEX, USA). The obtained data was analyzed using MASCOT.

2.11 Enzyme activity Assay

Activity of purified proteins was checked using their respective assays based on colorimetry. For this, required substrate and cofactors were provided externally in oxidized or reduced conditions as required for the functionality of the enzymes. The UV-visible spectrophotometer (Jasco 630, Japan) was used for detecting the absorbance of products formed as a result of enzyme substrate reaction. Electron paramagnetic resonance (EPR) JES-FA200 spectrophotometer was used for detecting the electron state (oxidation/ reduction) of Fe indicating its probability of binding with the enzymes to act as a co-factor. The details of the specific assays used for the characterization of the enzymes are given in the respective chapters.

2.12 Protein sub-cellular localization and Antibody binding studies

The purified proteins were used to raise antibodies in mice to co-localize the protein in parasite

infected thin blood smears prepared from clinical patients. These antibodies were also used to detect their specificity and binding efficiency with the parasite infected blood.

2.12.1 Antibody raising and ELISA

To raise antibodies against proteins involved in the isoprenoids biosynthesis pathway, the purified recombinant proteins were injected intra-peritoneal in mice along with Freund's adjuvant. For each protein, 10 female Swiss albino (Out bred) mice of 4 to 5 weeks, bred under specific- pathogen free conditions and supplied by Animal House Facility, BITS Pilani, Rajasthan, were used, where 5 mice constituted the experimental group and 5 mice were considered as controls. Before injecting the mice with the desired recombinant protein, all the mice from both the groups were pre-bled to check any inherent cross reactive antibodies present. Following this, for the primary injection in the experimental group, an emulsion of purified protein (25-30µg) was prepared by mixing it with equal quantity of complete Freund's adjuvant (FCA) and this emulsion was injected per mice. For control set, 1X PBS was emulsified with FCA (primary injection) and injected in mice. For booster doses, the protein/ PBS was emulsified with Freund's incomplete adjuvant (FIA) in similar concentrations as above and injected after regular intervals of 21 days. All the protocols were approved by the Institute's Animal Ethics Committee (Approval No. IAEC/RES/18/10). After 9 days of each booster dose, blood samples were withdrawn from mice by retro orbital method. Sera was separated from these samples by incubating the sample at 37°C for 1 hour, followed by overnight incubation at 4°C and centrifugation at 10,000 rpm at 4°C for 10 min. The supernatant (serum) was collected separately in a fresh tube and stored at -20°C till further use (Harlow and Lane, 1988).

Further, to check the IgG antibody titer specific against our desired recombinant proteins, ELISA was performed using different protein and serum dilutions. For this, 100µL of the recombinant protein samples (diluted to 1ng/µL in sodium carbonate bicarbonate buffer, pH=9.4) was coated in each well of 96-well flat bottom plate. The plates were kept overnight at 4°C followed by blocking with 5% fat free milk powder prepared in 1X PBS (pH=7.4) at 37°C. After incubation of 2 hrs, the plates were washed thrice with 1X PBS containing 0.2% Tween 20. Serum samples with varying dilutions (1:100 to 1:1000) were used as primary antibodies against specific proteins and the plates were incubated for 2 hrs at 37°C. Antibodies specific to recombinant protein were detected using HRP conjugate goat anti-mouse IgG (Fc fragment specific). TMB/H_2O_2 was used as a substrate and absorbance was recorded at 450nm on an ELISA plate reader (STAT FAX 2100). All tests were performed in duplicate and antibody levels were expressed as average

concentration units. Serum samples collected from mice injected with 1X PBS were used as controls. The specificity of these antibodies was further confirmed by western blot analysis.

2.12.2 Immuno-localization studies

To co-localize the proteins in the parasite, Immuno-fluorescence microscopy was performed on thin smears prepared from blood collected from *P. vivax* infected patients. The slides were fixed with 100% cold methanol. Fixed cells were washed with 1X PBS, permeabilised with 0.05% Saponin (30 mins) followed by 0.1% Triton X-100 (5 mins), and blocked for overnight using 3% BSA. After blocking, the cells were incubated with primary polyclonal antibodies raised in mice against the specific protein for 4 hrs at 37°C. For this antibody dilutions varying from 1:100 to 1:2000 were tried. After washing the cells with 1X PBS, Goat anti-mouse IgG FITC conjugated antibodies with a dilution of 1:2000 (Merck, Germany) was used as secondary antibodies and cells were incubated for one and a half hour at 37°C. Counterstaining of the parasite nucleus was done using DAPI (Life Technologies, USA) for 10 min and apicoplast membrane using Qdot® 585 Streptavidin conjugate (Life Technologies, USA) for 1 hr at room temperature (Jelenska et al., 2001). Intermittent washing with PBS was performed in between each step. The cells were finally mounted with VECTASHIELD (Vector Laboratories, USA) and viewed in a confocal laser scanning microscope (Leica TCS SP5) under a 63X oil immersion lens. RBCs infected with parasite were identified in bright field and protein of interest with green fluorescence emitted by FITC. DAPI was viewed as a blue dot while Qdot® 585 Streptavidin conjugate appeared as a red fluorescence under the microscope (**Table 2.3**). Exact location of the proteins was defined by the overlapping of the images of the same field.

Table 2.3 Fluorescent dyes and their excitation and emission wavelength peaks

Filter Set Description	Excitation (nm)	Emission (nm)	Barrier filter (nm)	Remarks
DAPI	359	461	450-465	Violet EX / Blue EM
FITC antibody conjugate	495	519	505-535	Blue EX / Green EM
Qdot® 585 Streptavidin conjugate	565	585	580-620	Green EX / Orange-Red EM

2.13 Flow cytometry

To check the binding of raised antibodies specifically to the parasitized RBCs, Flow Cytometry was performed. The study was carried out at the lab of Prof. John H. Adams, Department of

Global Health, College of Public Health, University of South Florida, Tampa, FL, USA. The antibodies were tested against previously collected and cryopreserved *P. vivax* infected blood samples. These samples were already verified for parasite infection.

2.13.1 Samples Preparation for Flow Cytometry

Cryopreserved *P. vivax* field isolates were first thawed, revived and enriched for ring stages using Percoll (Sigma, USA) gradient. Following this, the sample was fixed, permeabilized, and treated with antibodies of interest.

2.13.1.1 Thawing of Cryopreserved *P. vivax* field isolates

Reagents

(1) 1X Incomplete Media: McCoy's 5A media (filter sterilized), 0.18% Glucose, 0.2% $NaHCO_3$, 0.6% HEPES and 0.01% Gentamicin

(2) NaCl solutions: 12% and 1.6% (filter sterilized)

(3) Washing solution A: 1X Incomplete Media: McCoy's 5A media (filter sterilized)

(4) Washing solution B (PBS-B): 1 X PBS + 0.05% BSA

(5) 60% Percoll solution (5mL): 3mL 100% Percoll, 1.5mL 1XPBS, 0.5mL 10X Incomplete media

Cryopreserved blood samples were thawed to room temperature following the protocol published in Methods in Malaria Research (Moll et al., 2008) with mild modifications. Briefly, the sample was transferred to a 50mL sterile conical bottom centrifuge tube and treated drop wise with 12% NaCl (1/10 sample volume) while shaking followed by incubation for 2 min at RT. The sample was then centrifuged at RT for 3 min at 400g and supernatant was aspirated. To the pellet, 1.6% (10 volumes) of NaCl was added drop wise while shaking. After spinning the sample at 400g for 3 mins at RT, the pellet was washed twice with incomplete McCoy's 5A media (filter sterilized). The pellet was re-suspended in 600µL incomplete media and carefully overlaid on 60% Percoll solution in a fresh conical bottom tube. The sample was centrifuged at 1200 rpm for 15 min at RT to form a gradient where the top most layer is of serum followed by a layer of WBC (if present in original sample), followed by a layer of parasitized RBC and the bottom most layer of normal RBC with Percoll. The parasitized RBC layer was collected in a fresh tube, carefully without disturbing other layers. It was washed twice by spinning with PBS-B (500µL at 1250 rpm for 3 min) and finally left with 100µL PBS-B solution for the next step of the experiment.

2.13.1.2 Treatment of Red Blood cells with antibodies

The uninfected RBCs (uRBC) pellet isolated from fresh whole blood and the purified parasitized RBC (iRBC) pellet obtained from cryopreserved field isolates (from above) were suspended in 100µL of wash buffer PBS-B and fixed with 100µL of 0.05% glutaraldehyde for 15 min at 4^{o}C. Post fixation, cells were washed in PBS-B, treated with 100µL of 0.3% Triton X-100 for 10 min at RT, washed twice in PBS-B, and blocked with 3% BSA for 20-30 min at 4^{o}C. After washing in PBS-B, samples were processed for antibody staining at 4°C for 20-30 min in dark, using 100µL of the desired antibodies in a stepwise manner; where anti-*Pv*IspG (mouse antiserum prepared against bacterial expressed protein) was used as primary antibody; Goat anti-Mouse IgG Alexa Fluor 488 (1:100; Life Technologies, USA) was used as secondary antibody; Mouse anti-*Pv*BiP conjugated with AlexaFlour 594/ 647 (1:100; Life Technologies, USA) and CD235 APC (1:100; Life Technologies, USA) were other antibodies used for staining *P. vivax* and erythrocyte respectively. After washing with 100µL of PBS-B, the sample was suspended in 200µL of PBS-B. Multiple tubes with varying combination of cells and antibodies were set to assure binding of desired antibodies with *P. vivax* parasitized RBC. *P. falciparum* infected culture and fresh enriched uRBC were used as control. The samples were kept in the dark at 4°C until FC analysis.

2.13.2 Sample analysis using Flow Cytometry & Gating Strategy

Accuri C6 (BD Biosciences, USA) system with standard optic configuration (488nm blue laser and 633nm red laser) was used for flow cytometry analysis. The sample was transferred to 96 well flat bottom plate (NUNC, USA). The threshold was set at 80,000 on forward scatter (FSC-A) and about 10,000 events were acquired for each sample. As a control, uninfected RBC, and parasite infected RBC (both *Pf* and *Pv* infected) were used. Uninfected and infected RBC were also separately stained with Alexa Fluor 488 for use as a negative and positive control respectively to define cell populations with non specific or background reactivity. Data analysis was performed with CFlow Sampler version 1.0.227.4 (BD Biosciences, USA). Gating strategy was adopted based on the FSC-A/SSC-A profile. Unstained samples were used to set the threshold for positive/negative signal. Individual antibodies were first used alone to determine the background following which signals from two antibodies were analyzed to confirm the proper binding of the antibodies.

2.14 RNA extraction and Real time PCR analysis

Total RNA was extracted from the RBC fraction of parasite infected blood samples preserved in Tri Reagent (Sigma-Aldrich). The RNA extraction was performed using manufacturer's

protocols based on use of the Tri - reagent which gives feasibility to isolate RNA, DNA and protein from the same sample. The isolated RNA was used for Semi-quantitative and Quantitative analysis using gene specific primers.

2.14.1 RNA extraction

Reagents

1. Di-ethyl Pyrocarbonate (DEPC) (SIGMA)
2. Back Extraction Buffer: 4M Guanidine thiocyanate, 50mM Sodium citrate, 1M Tris (pH= 8.0).
3. Bromo-chloro propane (BCP) (SIGMA)
4. Histopaque (SIGMA)
5. Tri Reagent (SIGMA)
6. Isopropanol
7. 70% ethanol
8. 3% Hydrogen Peroxide solution
9. 10X MOPS electrophoresis buffer: 0.2M MOPS; 20mM Sodium acetate; 10mM EDTA (pH=8.0).
10. 10X Formaldehyde gel loading buffer: 10mM EDTA (pH=8.0); 50% glycerol; 0.25% Bromophenol blue; 0.25% Xylene cyanol.

Parasite infected erythrocytes were separated from Peripheral blood mononuclear cells (PBMCs) using density gradient centrifugation (Histopaque 1077, Sigma Aldrich, USA) according to manufacturer's instructions. Erythrocytes were washed with phosphate buffered saline (1X PBS), lysed using Tri-Reagent and preserved immediately at −80°C. All the samples were then transported in dry ice to BITS, Pilani for further processing. Before proceeding for RNA isolation, all plastic-wares and glasswares were treated with DEPC and autoclaved. Reagents were prepared in DEPC treated and autoclaved distilled water, followed by filter sterilization. Samples were processed for isolation of RNA, DNA and Protein following manufacturer's (Sigma Aldrich, USA) protocol.

Briefly the samples were thawed and BCP was added to the sample ($1/10^{th}$ volume of Tri reagent) following vigorous shaking for 15 sec, followed by incubation at RT for 15 min. The sample was then centrifuged at 12,000 rpm for 15 min at 4°C. Different phases were obtained after centrifugation where pink phase contains DNA and protein while transparent phase contains RNA. The transparent phase was separated, precipitated with isopropanol and incubated for 20

mins at RT. RNA pellet was obtained by centrifugation at 12,000 rpm for 10 min at 4°C. Supernatant was discarded and the pellet was washed twice with 70% ethanol. The pellet was dried and then resuspended in 80-90µL DEPC treated water and stored at -70°C. RNA integrity was assessed using denaturing agarose gels and RNA purity and concentration was checked by the Nano UV-vis Spectrophotometer (Simpli Nano GE Healthcare, UK).

For DNA isolation, pink phase obtained initially was mixed with back extraction buffer and shaken vigorously for 15 sec followed by incubation at RT for 15 min. After centrifugation DNA was obtained in aqueous phase, which was precipitated with isopropanol. DNA pellet was washed with 70% ethanol and resuspended in 100µL of 1X TE buffer. The DNA was used to confirm the infection using diagnostic multiplex and nested PCR as discussed above.

2.14.2 cDNA synthesis

First Strand cDNA synthesis was carried out from total RNA obtained from each sample using Quantitect Reverse Transcription kit (QIAGEN, Germany) according to the manufacturer's recommendations.

Genomic DNA elimination reaction

Template RNA	1µg
gDNA Wipeout buffer (7X)	2µL
RNA free water q. s.	14µL

Incubate at 42°C for 2 min and immediately transfer it to ice.

Reverse Transcription master mix:

Above mix	14µL
RT Buffer (5X)	4µL
RT Primer mix	1µL
Reverse transcriptase	1µL
Total	20µL

The reaction mixture was incubated at 42°C for 30 min and the enzyme was inactivated at 95°C for 3 min. After the completion of the reaction, cDNA sample was diluted to 50µL by adding 30µL of TE buffer and stored at -80°C till further use.

2.14.3 Quantitative PCR

Quantitative PCR using iQ SYBR green PCR super mix (Bio-Rad, USA) was performed on CFX connect Real-Time System (Bio-Rad, USA) using Seryl-tRNA synthetase gene as an internal reference. Fold change expression was calculated using ΔCt method.

Reaction mix:

2X *Taq* SYBR green mix	10µL
Forward primer	100ng
Reverse primer	100ng
cDNA template	50-100ng
Deionized water q. s.	20µL

The cycling parameter used were: 1 cycle of 95°C for 10 min, 35 cycles of 95° C for 0.5 min, 55°C for 0.5 min, and 72°C for 1 min followed by 1 cycle of 72°C for 3 min. Melt curve was generated for range of 55°C to 95°C with increment of 0.5°C per 5 sec. The real time primers used are listed in the respective chapters.

2.15 Studies on evolutionary position

2.15.1 Multiple sequence alignment

The PCR products of desired genes under study were sequenced commercially and the obtained gene sequences were thoroughly cross checked with the chromatogram and standard sequences available in database, to avoid any errors. The nucleic acid sequences were translated to protein sequences which were aligned against all available homologues and orthologs from different *Plasmodium* species, apicomplexans and prokaryotes using Clustal W2 (Larkin et al., 2007) and Clustal Omega (Sievers et al., 2011).

2.15.2 Phylogenetic Analysis

The evolutionary history of proteins involved in the isoprenoids biosynthesis pathway from field isolates of *Plasmodium* parasites was inferred by Maximum Likelihood method using MEGA 6.0 software (Jones et al., 1992; Tamura et al., 2013). The MSA obtained from Clustal Omega was used to generate the phylogenetic tree where Neighbor-Join and BioNJ algorithms were applied to a matrix of pairwise distances estimated using a JTT model, and then the topology with superior log likelihood value was selected (Tamura et al., 2013). The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analysed

(Felsenstein, 1985) and all positions containing gaps and missing data were eliminated. The obtained tree was analysed for the evolutionary position of the respective proteins in study to infer their evolutionary distance from different *Plasmodium* species as well as other apicomplexans and prokaryotes.

2.16 Protein Structure prediction

2.16.1 Secondary structure prediction

Three dimensional crystal structure of most of the enzymes involved in the isoprenoids biosynthesis pathway of *Plasmodium* parasite is not available. In the absence of three dimensional structure, prediction of secondary structure is a useful intermediate step which can facilitate the process of 3D structure prediction and helps to identify conserved domains. Secondary structures of proteins in our study were generated using PSI PRED server to find out the total number of α-helices and β-sheets present, to depict the conformation of amino acid residues that are seen repeatedly in the proteins. The obtained structures were compared with secondary structure of various other microorganisms with the help of Geneious II software (http://www.geneious.com, Kearse et al., 2012).

2.16.2 Prediction of 3D structure using Homology Modeling

To predict the 3D model of the proteins, template model was selected using HHpred server (Söding et al., 2005). For this, the translated amino acid sequences of enzymes involved in the isoprenoids biosynthesis pathway of *P. vivax* and *P. falciparum* were submitted to HHpred online server. Based on the sequence similarity of target with the template and resolution of template structure, the best structure was selected. This selected crystal structure was used as a template for protein structure prediction using MODELLER 9v11 (Sali and Blundell, 1993; Eswar et al., 2006). Discrete optimized protein energy (DOPE) was calculated for each structure and one with minimum DOPE score was selected.

Energy Minimization (EM) and Molecular Dynamics (MD) is the next step in generating a protein structure, where attempts are made to obtain a perfect folding of the protein. Molecular simulation was carried for modelled proteins using the GROMACS 4.6 software package (Van Der Spoel et al., 2005) following the manual developed by the GROMACS team. As suggested, GROMOS 96 force field and flexible SPC (Simple point charge) water was employed. OPLS-AA/L force field and tip3p water model was used to generate topology files and parameters stored in topology and force field files were used to solvate the protein. Decahedron box of 0.7Å

was selected to predict the size of simulation box and water was added as a solvent by using small pre-equilibrated system of water coordinates from SPC. Initially short energy minimization was done to check the presence of any broken hydrogen bond network in water which might lead to structure distortion. The standard cutoff of 1.0 nm was used for both neighbour list generation and Coulomb and Lennard-Jones interactions. Particle-Mesh-Ewald (PME) summation method was used for calculating electrostatic interactions and electrostatic energy, where energies and other statistical data are stored after every 10 steps (nst-energy). A pre-processing program "grompp" was run as a single input file (em.tpr) which collects all parameters, topology and coordinates, after which the simulation starts.

To evaluate the reliability of three dimensional structures of proteins, structure validation was accomplished by using different online available servers, where quality evaluation of the models for the environment profile was done with PROCHECK (Laskowski et al., 1993), VERIFY3D (Eisenberg et al., 1997) and ERRAT (Colovos and Yeates, 1993). Atomic contacts and packing of side chains in final refined models were evaluated using WHAT IF program (Vriend, 1990). Complete assessment and evaluation of the generated models were performed by Ramachandran plot analysis using RAMPAGE (Lovell et al., 2002). RMSD (root-mean-square deviation) value was calculated by superimposition of template and modelled protein structure using Chimera (Pettersen et al., 2004).

2.16.3 Substrate docking and *In-silico* drug inhibition study

Presence of binding site for substrates is a characteristic feature of a protein participating in different enzymatic reaction. Binding pocket made up of conserved amino acid residues were identified using BioLip (Yang et al., 2013) server. Based on the confidence score (C-score ranging from 0 to 1) obtained for this site, docking studies were performed by using Molegro Virtual Docker (Thomsen and Christensen, 2006). In the first step, coordinates for the proteins, substrate molecules and drugs were stored as a PDB file using Discovery studio 3.5 Client (BIOVIA, US) and uploaded to the workspace of MVD. Here in MVD, the workspace acts like a central component and represents all the information available to the user in terms of molecules (proteins, ligand, cofactors, water molecules, and poses), user-defined constraints (visualized as small spheres), cavities (visualized as a grid mesh), and various graphical objects (molecular surfaces, backbone visualizations, labels, etc.). Binding cavities were located on these proteins and selected as constraints and used as a site of docking. Docking process starts and eventually times out on its own after 2000 iteration (or about 100, 000 evaluation) where generally docking

engine should find a good solution in 800 iterations. Once the simulation is over, docking results were imported from the output directory together with pose organizer. In pose organizer all the poses were analyzed manually to check the interaction between protein and ligand molecule. During analysis, affinity of conserved amino acid residues towards substrate ligand molecules was studied where hydrogen bond, electrostatic interaction and steric interaction was mainly observed. Simultaneously, to check the inhibition of these *Plasmodium* proteins, *in-silico* High throughput screening (HTS) was performed where the structure of antimicrobial drug molecules (showing inhibition of prokaryotic isoprenoids biosynthesis) was obtained from the FDA approved drug library and other available sources. Ligand or drug molecules were modified to produce structural variations, perform corrections, exclude undesirable structures, add hydrogen atoms, charged groups neutralization and structures optimization for a pH range from 6 to 9. Glide (Grid-based ligand docking energetics) software (Schrodinger, Mannheim, Germany) was used for this screening, and was run from the graphical interface of Maestro (Schrodinger, Mannheim, Germany) where ligand grid box was set directly on the catalytic site of protein. Drug molecules identified with higher affinity were docked with molecular docking algorithm MolDock using the Molegro Virtual Docker version 5.5 software (Molegro ApS, Aarhus, Denmark; http://www.molegro.com) according to instructions. Interaction of conserved amino acid residues was studied and compared with the amino acids interacting with substrate molecules. Based on the level of interactions, the effect of different drug molecules was assessed.

2.16.4 Molecular Dynamics for ligand docked proteins

The obtained ligand protein complex was subjected to energy minimization and molecular dynamics simulation. This is a complicated, multistep process, and divided into following subparts.

1. Generate topology
2. Define box and solvate
3. Add ions to the system
4. Energy minimization
5. Equilibration
6. Molecular dynamics
7. Analysis

In first step coordinates of ligand were separated from the protein-ligand complex and topology was generated for protein and ligand molecules separately by using GROMACS 87 force field. Once the topology was generated, topology complex was prepared by adding the ligand topology

to the protein topology file manually. Next step was to determine the type and size of the box required for simulation process. Three kinds of boxes can be used in simulation studies, namely triclinic, cubic and octahedron where box size varies with the size of the protein. Protein-ligand complex defined in a most favourable box was solvated by using SPC water model system. The positive or negative charge present on the protein molecule in a solvated system was neutralized by using Na^+ or Cl^- ions. Usually, during the molecular simulation study hydrogen bonds are added to the atoms which may decrease the distance between the atoms, causing collision between them. To prevent this, Energy Minimization (EM) was performed where the required parameter was added in the form of "em. mdp" (energy minimization molecular dynamics parameter) file specifying steps and cut-off scheme. After this, the system was equilibrated at optimum temperature and pressure by keeping a high control on both parameters, by using the specific coupling group treatment in parameter files. Upon the completion of the two equilibration phases, the system is well-equilibrated and ready for running molecular dynamics for data collection. The process of molecular simulation is similar to EM where required parameters are added in the form of "md.mdp" file and simulation can be run for 1 to 5 ns. Finally the results obtained were analysed and the protein structure before and after simulation was compared. The binding affinity of drug molecule was evaluated by comparing the binding energies between the substrate and enzyme with respect to that of drug molecule and enzyme using MVD.

Materials used for various reactions

Vector and Bacterial strains used

pRSET A	Invitrogen, USA
pMAL-c2X	New England Biolabs, USA
RIG plasmid	Kind gift from W.G.M. Hol USA.
E. coli strain DH5α	Thermo Scientific, USA
E. coli strain BL21 (DE3)	Novagen, Germany
E. coli strain BL21 (DE3) pLysS	Novagen, Germany
E. coli strain Rosetta (DE3)	Novagen, Germany
E. coli strain Rosetta (DE3) pLysS	Novagen, Germany
E. coli strain ER2523	New England Biolabs, USA

Various kits used in this study

Anti-His Antibody kit	QIAGEN®, Germany
Gel extraction kit	QIAGEN®, Germany
Ni-NTA agarose Resin	QIAGEN®, Germany
PCR purification kit	QIAGEN®, Germany
pMAL protein fusion and purification system	New England Biolabs, USA
Plasmid miniprep kit	QIAGEN®, Germany
Quantitect Reverse Transcriptase kit	QIAGEN®, Germany

Antibodies and Fluorescence dyes

Anti-His antibody	QIAGEN, Germany
Goat anti-Mouse IgG HRP conjugate	Merck, Germany
Goat anti-Mouse IgG FITC conjugate	Merck, Germany
Qdot® 585 Streptavidin conjugate	Life Technologies, USA
DAPI	Life Technologies, USA
Goat anti-Mouse IgG Alexa Fluor 488	Life Technologies, USA
Goat anti-*Pv*BiP AlexaFlour 594/ 647	Life Technologies, USA
Anti-human CD235a APC	Life Technologies, USA

Composition of various buffer and solution used in the study

Acid Citrate Dextrose (1000mL)	Sodium citrate 13.2gm
	Citric acid: 4.8gm

	Dextrose: 14.7gm
Coating buffer for ELISA (10X; 100mL)	Sodium carbonate 1.59gm
	Sodium bicarbonate 2.93gm
DNA loading buffer, 6X	9mg Bromophenol blue
	9mg Xylene cyanol FF
	Dissolve in 8.8mL of 60% Glycerol and add
	1.2mL of 0.5M EDTA
LB broth (1000mL)	10gm Tryptone
	5gm Yeast extract
	5gm NaCl
	Autoclave
Phosphate Buffered Saline pH 7.4 (1X)	137mM NaCl
	2.7mM KCl
	4.3mM Na_2HPO_4
	1.4mM KH_2PO_4
RNA gel running buffer, 20X	20mM MOPS
	2mM Sodium acetate
	0.25mM EDTA
TE buffer (1X; pH 8.0)	10mM Tris-Cl (pH 8.0)
	1mM EDTA (pH=8.0)
TAE buffer (50X; 1000mL)	40mM Tris-acetate
	1mM EDTA (pH 8.0)
	57.1mL Glacial acetic acid

Buffers for restriction endonucleases, T4 DNA Ligase and *Taq* DNA Polymerase

The various buffer used were provided by the manufacturer for these enzymes and used as per the instruction provided.

Accession number for IspG and IspD sequences (NCBI) used in various analyses:

Table 2.4 Details of sequences from NCBI database of different organisms used for analysis (NA* not applicable)

S. No.	Organism	Database sequence Annotation	
		IspD	IspG
1.	*P. vivax*	PVX_081425	PVX_111575
2.	*P. inui*	XP_008818697	EUD66257
3.	*P. cynomolgi*	XP_004220569	XP_004221386
4.	*P. knowlesi*	XP_002257737	XP_002261743
5.	*P. vinckei*	XP_008626375	EUD69889
6.	*P. chabaudi*	XP_742791	XP_743307
7.	*P. berghei*	XP_670908	XP_680310
8.	*P. yoelii*	XP_727858	ETB58995
9.	*P. falciparum*	XP_001351000	XP_001347505
10.	*Babesia bovis*	XP_001611481	XP_001610211
11.	*Toxoplasma gondii*	XP_002370388	XP_002365345
12.	*Theileria annulata*	XP_954773	XP_952048
13.	*Theileria parva*	TP03_0057	XP_765233
14.	*Babesia equi*	NA*	XP_004829307
15.	*Aquifex aeolicus*	NA*	O67496
16.	*Thermus thermophiles*	YP_006057937	Q84GJ3
17.	*Bacillus subtilis*	NP_387971	WP_010886567
18.	*Escherichia coli*	YP_003000333	WP_052929324
19.	*Chromera velia*	Cvel_20751	Cvel_23671
20.	*A. protothecoides*	XP_011402101	NA*
21.	*Arabidopsis thaliana*	NP_565286	NP_200868
22.	*Chlamydomonas reinhardtii*	NA*	XP_001690937

Functional and Structural Characterization of IspD Enzyme from *P. vivax* and *P. falciparum*

Chapter III

Characterization of Isoprenoids biosynthesis pathway enzyme IspD from *P. vivax* and *P. falciparum*

This chapter focuses on the characterization of CDP-ME synthase (IspD) enzyme involved in the third step of the isoprenoids biosynthesis pathway. In this chapter, we have detailed the amplification, cloning, expression and purification of IspD protein from Indian *P. vivax* field isolates. Various biochemical and enzyme inhibition assays were set to detail the catalytic activity of IspD enzyme and to look for compounds that can act as inhibitors of IspD respectively. Real time PCR was performed to check the expression profile of IspD during the erythrocytic stages of the parasite obtained from clinical isolates. Immuno-localization studies were performed to confirm its site of action and phylogenetic tree was constructed to assess its evolutionary position.

3.1 Introduction

Non-mevalonate isoprenoids biosynthesis pathway houses many probable drug targets and amongst them, IspD [2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (YgbP)] is one of the most promising and attractive drug targets. IspD enzyme participates in the third step of the DOXP pathway, where it catalyses the cytidylation process and exclusively shows activity only in the presence of divalent ions Mg^{2+} or Mn^{2+}. These divalent cations form coordination bonds with α-, β-, and γ-phosphate oxygen of CTP or the α-phosphate oxygen of CDP-ME, and stabilize their binding (Richard et al., 2001). During the catalytic reaction, nucleotide derivatives i.e. cytosine 5′-triphosphate (CTP) and phosphate groups i.e. 2-C-methyl-D-erythritol-4-phosphate (MEP) are directly introduced as substrates and produces 4-diphosphocytidyl-2C-methyl-D-erythritol (CDP-ME) and pyrophosphate as a bi-product (**Figure 3.1**). IspD enzyme's catalytic domain is highly basic and has a high degree of specificity for CTP molecules where pyrimidine base forms hydrogen bonds with the backbone of conserved amino acids residues. In fact, mutations in these conserved amino acids residues abolished or reduced the catalytic activity in *E. coli* IspD.

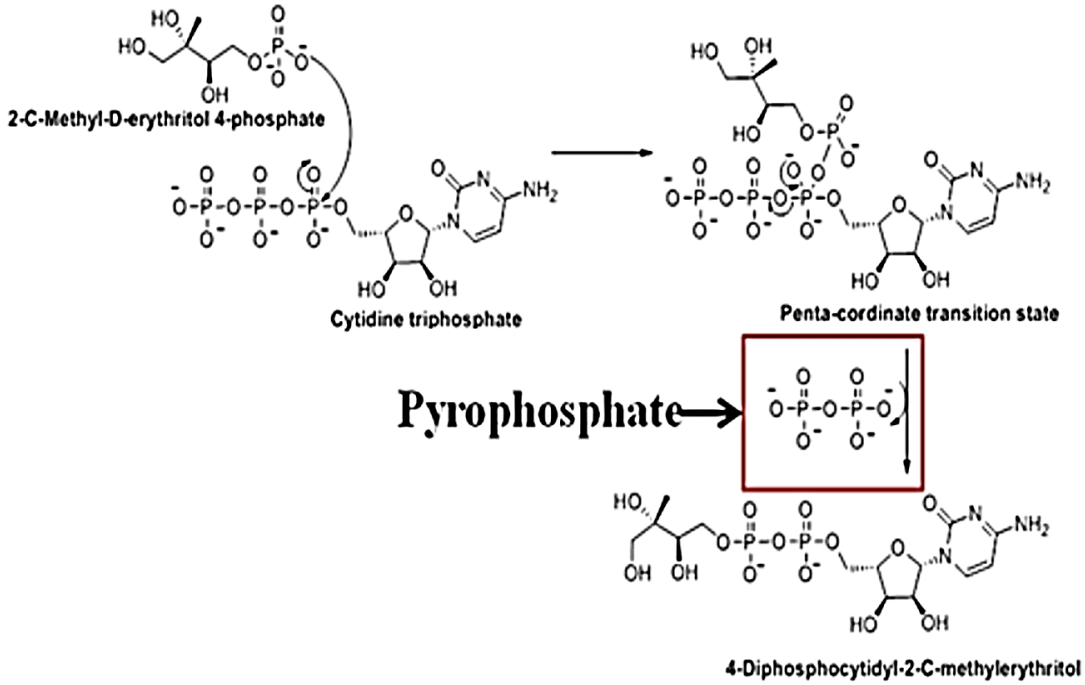


Figure 3.1: Associative mechanisms to represent the catalytic activity of IspD enzyme

The catalytic mechanism for IspD enzyme is unique and comes under the category of associative mechanism. According to this process, when there is a nucleophile attack on the α-phosphate of CTP by the 4-phosphate of MEP, a negatively charged penta-coordinate transition state is formed. The collapse of this charged transition state finally leads to

pyrophosphate release and CDP-ME formation (**Figure 3.1**). Two lysine residues (Lys27 and Lys213 in *E. coli* IspD) are critical to stabilize this pentavalent transition state. Amino acid alignments showed that the basic residues in the IspD active site; in particular the *E. coli* IspD Lys27/Lys213 are conserved in *P. falciparum* also (Hunter, 2011). In *P. falciparum*, by performing chemical and genetic validation experiments and generating disrupted and control vectors with modified *Pf*IspD locus, Imlay et al. (2015) gave evidence that complete IspD disruption is lethal for the parasite, suggesting its importance for asexual growth of malaria parasites. This validates IspD as an essential, parasite-specific target for antimalarial drug development and thus anticipates the inhibitors of this enzyme to possess potential for the parasites demise. Further, to screen various compounds available at Malaria Box, they prepared a codon optimized synthetic IspD gene from *Plasmodium* and expressed it in a prokaryotic system. They have shown MMV008138 as a potential inhibitor for parasite growth amongst the screened compounds (Imlay et al., 2015). However, there are still no reports detailing this enzyme from any field isolates.

In 2011, when we started this work, there were no reports on characterization of this enzyme from any *Plasmodium* species. The PlasmoDB database also had designated this enzyme as putative in *P. falciparum* (PF3D7_106900), however, its *P. vivax* orthologue was designated as hypothetical. Thus, we chose this enzyme for further characterization to elucidate its function in *Plasmodium* especially in the samples collected from patients infected with either *P. falciparum* or *P. vivax*.

3.2 Results and Discussion

Based on various alignments and conserved domain analysis, as discussed in Chapter 2, we shortlisted a hypothetical protein (Gene ID: PVX_081425) as *P. vivax* IspD. Further, we predicted the targeting of this gene as well as PF3D7_106900 by checking the presence of N-terminal bipartite leader sequences using various available online servers as described in methodology section 2.1.

3.2.1 Apicoplast targeting

The amino acid sequence of putative *Pv*IspD (PVX_081425) and *Pf*IspD (PF3D7_106900) when submitted to online server PlasmoAP, showed three out of total four tests positive for the signal peptide. Based on different parameters considered by PlasmoAP like the presence of basic amino acids and KN enriched regions at the N terminal, IspD in *P. falciparum*

seems targeted to apicoplast with all the tests positive for the presence of transit peptide. However, surprisingly for *Pv*IspD only 3/5 tests were positive, indicating lesser probability of targeting of this protein to the apicoplast (**Figure 3.2**). To confirm the PlasmoAP results, we used the PATS server and obtained similar results, which showed high probability for presence of apicoplast targeting peptide in *Pf*IspD with a score of 0.965, but a low score of 0.558 for apicoplast targeting in *P. vivax.* As the apicoplast targeting was not conclusive for *Pv*IspD enzyme by the above software's, we went ahead to check if mitochondrial signals are present in these sequences. We submitted the sequences to MitoProtII and PlasMit online servers. However, we got contradictory results where MitoprotII showed a high probability of targeting of both *Pv*IspD and *Pf*IspD to mitochondria (0.98 for both *P. falciparum* and *P. vivax*), while PlasMit predicted the proteins as non-mitochondrial (99% non-mitochondrial for both). The doubt created by these mixed-up results thus demands a deeper investigation into targeting of these proteins, possibly using localization studies.

	Plasmodium vivax		*Plasmodium falciparum*	
Criterion	**Value**	**Decision**	**Value**	**Decision**
Signalpeptide	3 of 4 tests positive	+	3 of 4 tests positive	+
apicoplast-targeting peptide	3 of 5 tests positive	-	5 of 5 tests positive	++
Ruleset 1				
Ratio acidic/basic residues in first 22 amino acids <=0.7	0.000	yes	0.167	yes
Does a KN-enriched region exist (40 AA with min. 9 K or N) with a ratio acidic/basic <=0.9	0.000	no	0.083	yes
Ruleset 2				
number of acidic residues in first 15 amino acids (<=2)	0	yes	0	yes
Does a KN-enriched region exist (40 AA with min. 9 K or N) ? Ratio acidic/basic residues in this region <0.6	0.000	no	0.083	yes
Is the first charged amino acid basic ?		yes		yes

Note: The **final decision** is indicated by "++, +, 0 or -", where apicoplast-localisation for a given sequence is considered as "++ very likely; + likely; 0 undecided; - unlikely

Figure 3.2: Prediction of bi-partite N-terminal leader sequence with PlasmoAP for *Pv*IspD and *Pf*IspD.

3.2.2 Amplification of *ispD* gene

3.2.2.1 Primer designing and amplification of *PvispD*

The *ispD* gene (putative) of *P. vivax* Salvador I (Gene ID: PVX_081425) is 1857bp long and expands from 332,972 – 334,828 position on chromosome no. 2. It is an intron-less gene encoding for 1857 bases mRNA. To amplify the complete *ispD* gene from different *P. vivax* field isolates, two sets of primers were designed from genomic sequence of PVX_081425 (**Figure 3.3 & Table 3.1)**. Amplification of the full-length gene was obtained from 5 field isolates using primers VS5A and VS8B following conditions given in **Table 3.2** (**Figure 3.4**). Partial regions of *PvispD* were also amplified using additional set of primers where primers VS5A and VS6 amplified the upper 1220bp region and primers VS7 and VS8B amplified the lower 942bp region. The conditions used are as given in **Table 3.2**. The obtained amplicons (**Figure 3.4**) were purified using QIAquick Gel Extraction kit, commercially sequenced and analysed by different software as described in materials and methods.

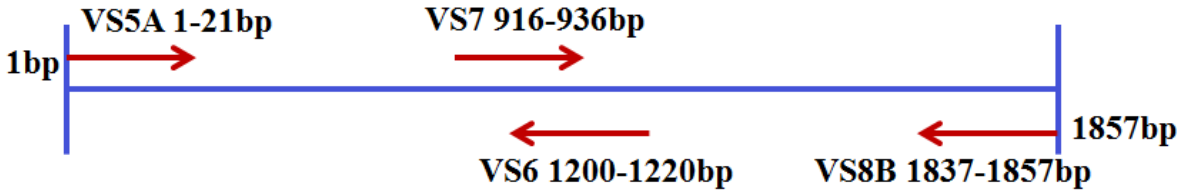


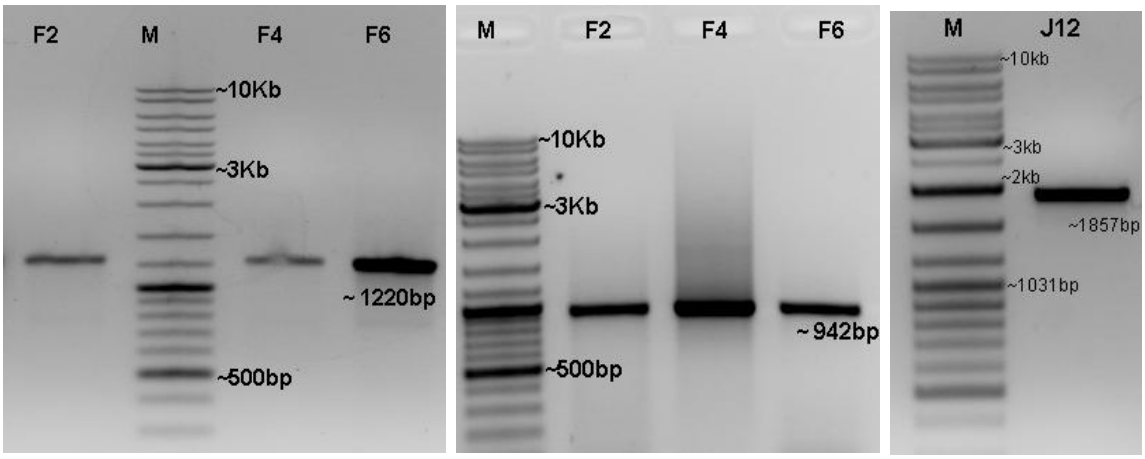
Figure 3.3: Position of primers designed for amplification of *ispD* gene from *P. vivax*.

Table 3.1 Primer sequences for the amplification of *ispD* gene from *P. vivax* genome

Lab Id	Primer sequence	Restriction enzyme site
VS5A	5'GCG **GGA TCC** ATG AAC ACC CTG CTG AGC GTT 3'	BamHI
VS6	5'GAT TTA TCC CCT TTA TCC GCG 3'	NA
VS7	5'GAT AAT GAG AAG GGG AAG TGC 3'	NA
VS8B	5'GCG **CTG CAG** CTA CTC GTA GTA AAA GTG GCG 3'	PstI

Table 3.2 Reaction conditions employed for the amplification of *PvispD* gene

Reaction Steps	Primers used			
	VS5A and VS8B (1857bp)	VS5A and VS6 (1220bp)	VS7 and VS8B (942bp)	
Pre-denaturation	94 °C for 3 min	94 °C for 3 min	94 °C for 3 min	
Denaturation	94 °C for 1 min	94 °C for 1 min	94 °C for 1 min	
Annealing	48.2 °C for 1.5 min	49 °C for 1 min	49 °C for 1 min	x35
Extension	72 °C for 2.5 min	72 °C for 2.5 min	72 °C for 2.5 min	
Post Extension	72 °C for 4min	72 °C for 4min	72 °C for 4min	



a) **b)** **c)**

Figure 3.4: *PvispD* gene amplification (Parasite DNA samples: F2, F4, F6 and J12; M: Gene Ruler DNA Ladder mix Fermentas SM0331). (a) Upper gene fragment (1220bp) (b) lower region (942bp) (c) Complete *PvispD* gene (1857bp)

3.2.2.2 Primer designing and amplification of *PfispD*

The *P. falciparum ispD* gene; putative is 2205bp in size and expands from position 290,386-292,590 on chromosome 1 in *P. falciparum* genome. Based on similar approach as used for *Pv*ispD, two sets of primers were designed to amplify the *PfispD* gene, (**Figure 3.5; Table 3.3**) using gene sequence PF3D7_106900. To amplify the full-length *PfispD* gene from different *P. falciparum* field isolates, we used VS1 and VS4 as per the conditions mentioned in **Table 3.4**. Partial fragments of the gene were amplified using primers VS1 and VS2 that gave amplification of upper 1306bp region and VS3 and VS4, which gave an amplification of lower 1424bp region. Amplicons obtained (**Figure 3.6**) from 3 field isolates were commercially sequenced and analysed.

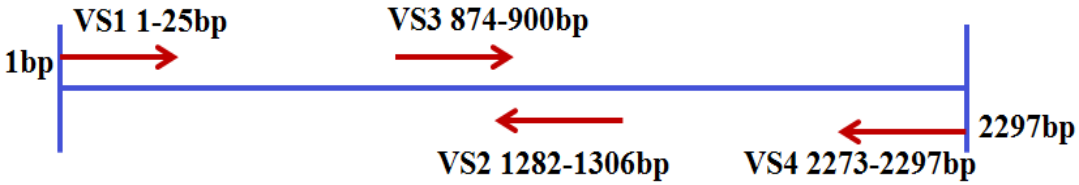


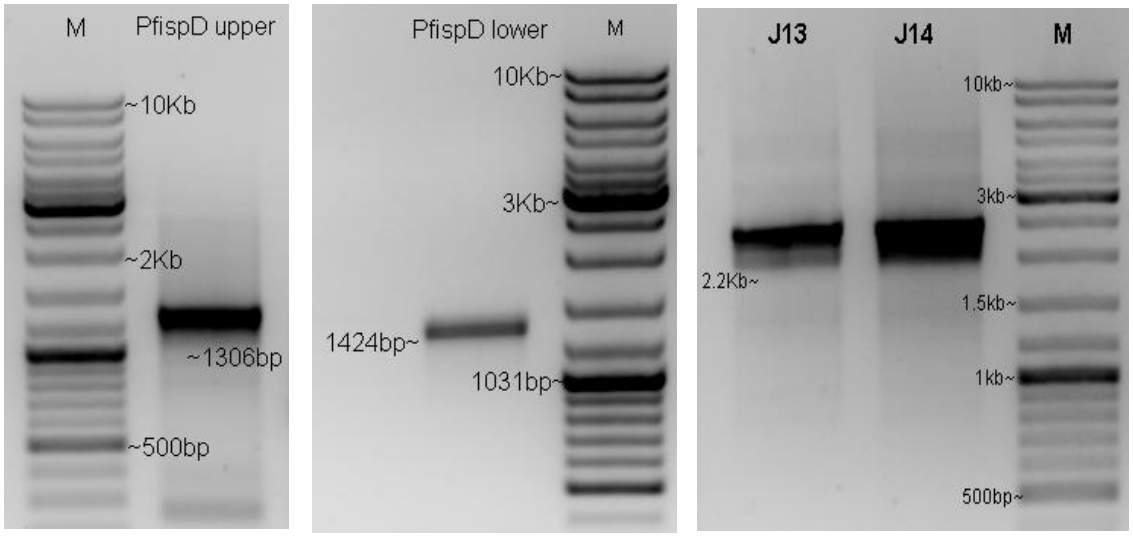
Figure 3.5: Position of primers designed for amplification of *ispD* gene from *P. falciparum*

Table 3.3 Primer sequences for the amplification of *ispD* gene from *P. falciparum* genome

Lab ID	Primer Sequence
VS1	5'ATG CAT TTT GTT CAT ACG TTT ATT C 3'
VS2	5'CAT TTA AAA AGG AAT CTA ATC GTT C 3'
VS3	5'GAT GAA AAA GAC AAC ATA AAA GAT TTC 3'
VS4	5'CGA ATT TAA GAA TGC GAT GTA TAT 3'

Table 3.4 Reaction conditions employed for the amplification of *PfispD* gene

Reaction Steps	Primers used			
	VS1 and VS4 (2205bp)	VS1 and VS2 (1306bp)	VS3 and VS4 (1424bp)	
Pre Denaturation	94 °C for 3 min	94 °C for 3 min	94 °C for 3 min	
Denaturation	94 °C for 1.5 min	94 °C for 1 min	94 °C for 1 min	x35
Annealing	45.7 °C for 2 min	50 °C for 1 min	50 °C for 1 min	
Extension	72 °C for 2.5 min	72 °C for 1.5 min	72 °C for 1.5 min	
Post Extension	72 °C for 4min	72 °C for 4min	72 °C for 4min	



(a) **(b)** **(c)**

Figure 3.6: *PfispD* gene amplification Parasite DNA samples: J13 and J14; M: Gene Ruler DNA Ladder mix Fermentas SM0331). (a) Upper region (1306bp) (b) Lower region (1424bp) (c) Complete *PfispD* gene amplification (2205bp).

3.2.3 Sequence Analysis and Multiple Sequence Alignment

The obtained sequences of Indian *PvispD* (BankIt ID: KX176847-51) and *PfispD* (BankIt ID: 1935442) were analysed at the nucleotide and amino acid level, amongst different Indian isolates as well as with Salvador I and *Pf*3D7 respectively, and were found to have 99.9% similarity. These sequences were also analysed with homologues/orthologues from other organisms including prokaryotes, apicomplexans, chromalveolate, chlorophyta and plants (Acc. No. as given in **Table 2.4**). In the multiple sequence alignment, the Indian *P. vivax* sequences showed percent identity ranging from 19 to 77%. Within the *Plasmodium* genus, the *Pv*IspD showed percent identity of 75.7% with *P. cynomolgi*, 69.4% with *P. inui*, 61.4% with *P. knowlesi* and 38.5% with *P. falciparum*. Alignment between IspD protein of prokaryotes and apicomplexans showed a long gap at the N-terminal region because of the

absence of N-terminal targeting sequence (**Figure 3.7**). In comparison, the sequence obtained for IspD protein of *P. falciparum* showed a match of 49.2% with *P. vinckei*, 61.3% with *P. chabaudi* and 43% with *P. yoelii*.

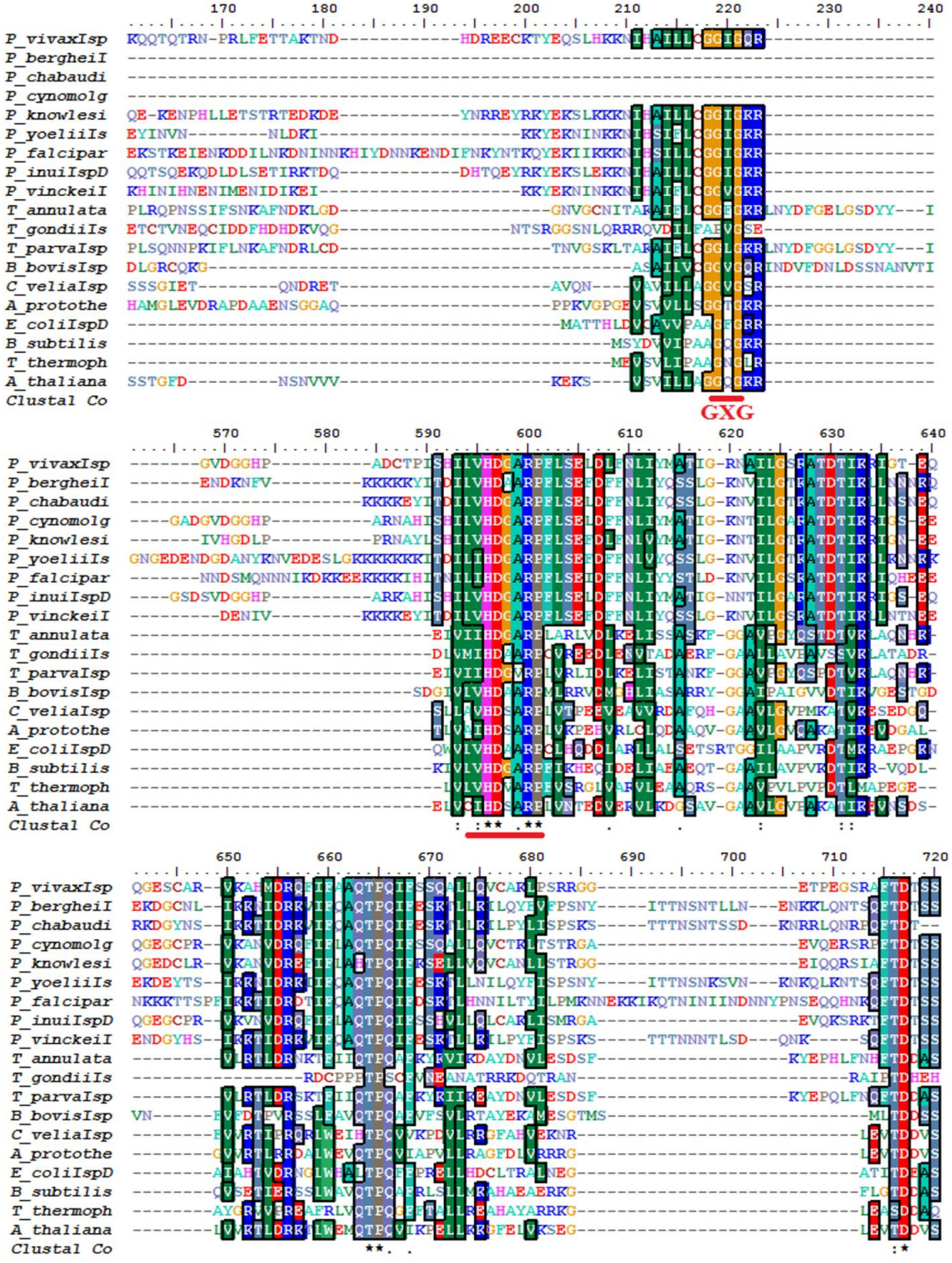


Figure 3.7: Multiple sequence alignment (Clustal Omega) of Indian *P. vivax* IspD sequence with IspD orthologues from different organisms. The marked lines depict the conserved residues in the signature motifs.

The similarity score of Indian *Pv*IspD and *Pf*IspD with other *Plasmodium* species, apicomplexans and prokaryotes at the nucleotide and amino acid level is as shown in **Table 3.5**.

Table 3.5 Percent identity of *Pv*IspD protein sequence from Indian field isolates with IspD sequence from other a) *Plasmodium* species (b) prokaryotes (c) apicomplexans

(a)

Nucleotide/ Amino acid	***P. vivax***	***P. falciparum***	***P. berghei***	***P. chabaudi***	***P. cynomolgi***	***P. knowlesi***	***P. yoelii***	***P. inui***	***P. vinckei***
P. vivax		45.0	50.9	47.2	81.3	72.5	44.5	78.6	46.7
P. falciparum	38.5		68.1	62.6	51.3	48.0	59.8	48.3	63.0
P. berghei	45.7	54.9		80.1	52.3	53.5	82.4	50.0	87.0
P. chabaudi	48.7	61.3	80.7		49.6	50.2	78.5	47.2	85.8
P. cynomolgi	75.7	43.8	49.8	53.2		81.0	47.2	86.0	52.6
P. knowlesi	61.4	36.9	47.1	48.7	73.9		47.7	78.9	49.9
P. yoelii	37.1	43.0	76.4	77.4	42.9	38.9		46.0	79.2
P. inui	69.4	39.9	46.1	54.1	81.2	68.3	39.3		48.9
P. vinckei	41.5	49.2	82.7	89.1	51.0	42.1	68.2	45.3	

(b)

Nucleotide/ Amino acid	***P. vivax***	***P. falciparum***	***C. velia***	***A. proto-thecoides***	***E. coli***	***B. subtilis***	***T. thermophilus***	***A. thaliana***
P. vivax		47.2	31.2	32.9	36.0	35.8	32.9	37.1
P. falciparum	37.5		30.8	27.2	31.8	37.9	22.4	43.5
C. velia	20.7	21.5		48.4	40.5	44.8	45.2	46.1
A. protothecoides	21.5	21.3	43.5		40.3	43.9	50.4	45.7
E. coli	24.6	23.8	29.1	27.8		46.0	42.0	39.7
B. subtilis	29.4	25.9	35.1	35.5	33.8		41.9	47.7
T. thermophilus	22.9	18.3	33.8	34.5	33.2	37.5		37.1
A. thaliana	25.4	26.3	40.9	52.7	30.3	38.0	31.9	

(c)

Nucleotide/ Amino acid	***P. vivax***	***P. falciparum***	***T. annulata***	***T. gondii***	***T. parva***	***B. bovis***
P. vivax		47.2	39.3	39.6	39.7	40.8
P. falciparum	37.5		46.6	38.7	47.1	41.5
T. annulata	24.7	24.9		36.3	82.1	46.1
T. gondii	20.7	19.1	26.1		35.5	35.5
T. parva	25.5	23.9	78.2	26.7		45.1
B. bovis	25.0	22.1	34.3	20.9	33.6	

3.2.4 Phylogenetic analysis

The evolutionary position of *Pv*IspD and *Pf*IspD was predicted by constructing a phylogenetic tree using MEGA 6.0 (Tamura et al., 2013). The MSA generated for

phylogeny tree construction included 20 amino acid sequences including major apicomplexans (*Plasmodium, Toxoplasma, Babesia, Theileria*), prokaryotes (*E. coli, B. subtilis, A. aeolicus, T. thermophilus*), plants (*Arabidopsis),* chlorophyta (*Chlamydomonas)* and a Chromerida (*Chromera velia* – known as the photosynthetic bridging link between algae and apicomplexans) (Acc. Nos. as given in **Table 2.4**). The phylogenetic tree for *Pv*IspD protein was constructed using Maximum Likelihood method based on JJT matrix-based model (**Figure 3.8**).

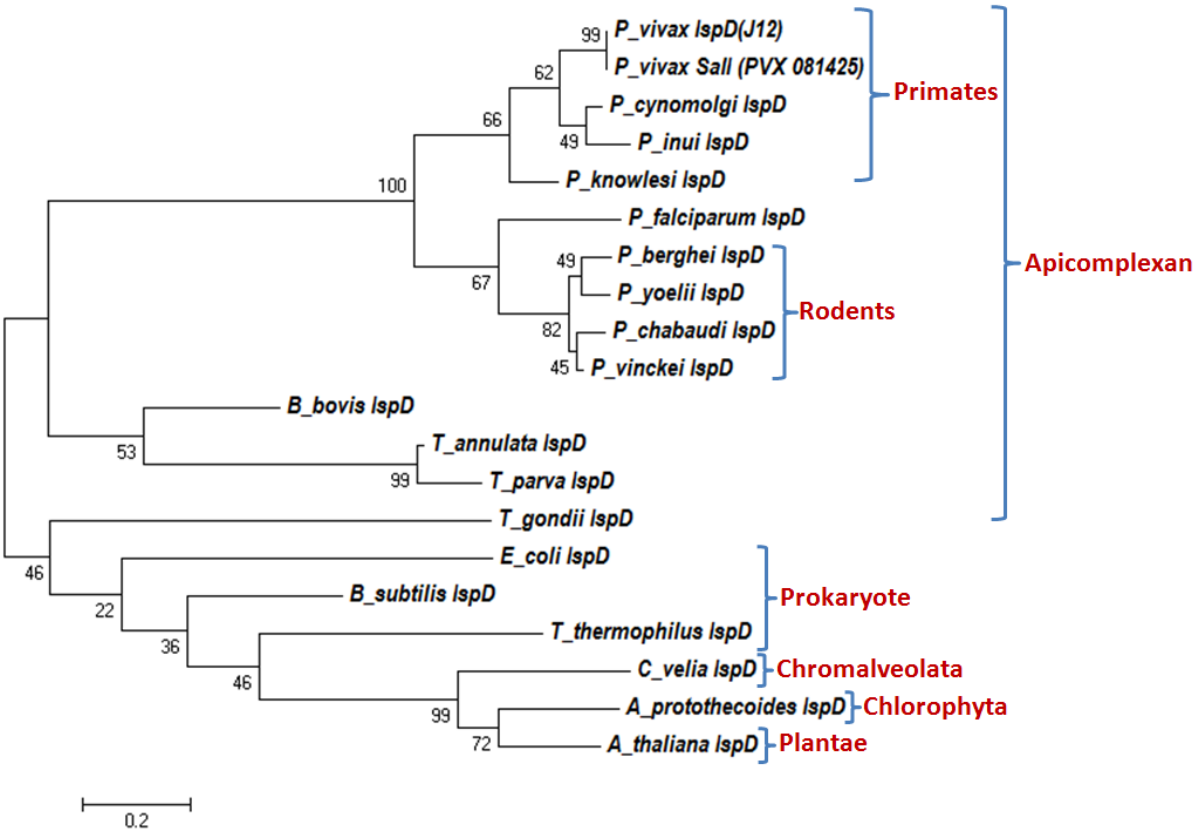


Figure 3.8: Phylogenetic analysis of IspD protein sequences. [IspD protein sequence of Indian *P. vivax* with orthologues]. Boot strapping done for 1000 iteration and percentage of trees in which the associated taxa clustered together shown next to the branches.

On analysis of different clades, *Pv*IspD protein grouped along with other primate parasites *P. knowlesi* and *P. cynomolgi* while rodent malaria parasite *P. berghei*, *P. chabaudi*, *P. yoelii* and *P. vinckei* were present in different clade. These two clades seems to diverge out of *P. falciparum* clade indicating a separate origin of *P. falciparum*. This evolutionary analysis of *Pv*IspD protein is in agreement to other *Plasmodium* genome members reported in various studies (Rathore et al., 2001; Escalante et al., 2005; Saxena et al., 2007 & 2012; Martinson et al., 2008) detailing the origin of *P. vivax* as primate parasite and a major divergence of *P. vivax* from *P. falciparum*. All the apicomplexans apart from the

Plasmodium species formed a separate clade, where *T. gondii* was in close proximity to prokaryotic clade, away from other apicomplexans. All the organisms with photosynthetic plastid formed a single clade showing their close association with each other, but divergence from other non-photosynthetic plastid containing organisms. The average evolutionary distance calculated amongst these enzymes was 0.595.

3.2.5 Conserved domain and Signature motif analysis

A Glycosyl transferase family A (GT-A) domain with a CDP-ME synthetase activity is the characteristic feature of all IspD proteins. This domain participates in the synthesis of oligosaccharides, polysaccharides, and glycol conjugates by transferring the sugar moiety from an activated nucleotide sugar donor to an acceptor molecule (Lairson et al., 2008)). The Conserved Domain Detection (CDD) tool available at the NCBI website confirmed the presence of GT-A domain in Indian *Pv*IspD and *Pf*IspD proteins (**Figure 3.9 and 3.10**) towards the C- terminal of the protein spanning 191-558 amino acid residues in *P. vivax* and 191-677 in *P. falciparum*.

The detection of signature motifs in *Pv*IspD and *Pf*IspD proteins using PROSITE showed the presence of two signature motifs, viz., GXG motif and [IVT]-[LIVMC]-[IVT]-[HS]-D-[SGAV]-[AV]-R motif spanning the amino acid position 199-201 and 424-431 in *P. vivax* and 204-206 and 512-519 in *P. falciparum* (**Figure 3.7**). The sequence of second signature motif in *Pv*IspD protein was ILVHDGAR whereas in *Pf*IspD it was ILIHDGAR. This pattern was present throughout the IspD orthologues. The presence of these conserved domain and signature motifs indicates that this protein might function as IspD enzyme that participates in the MEP isoprenoids biosynthesis pathway.

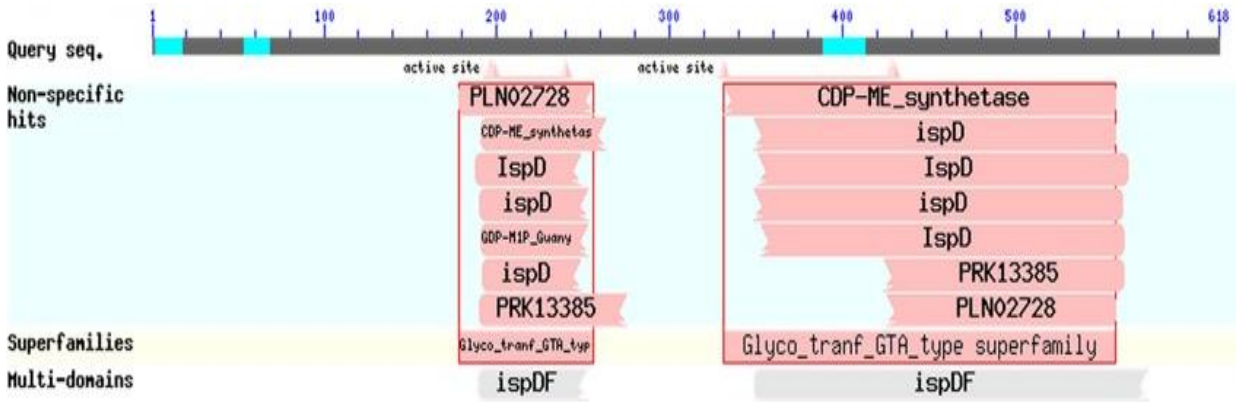


Figure 3.9: Conserved domain detection (CDD) of *Pv*IspD protein The region spanning the amino acid position 191-558 of the query protein sequence showed a match with GT-A type super families.

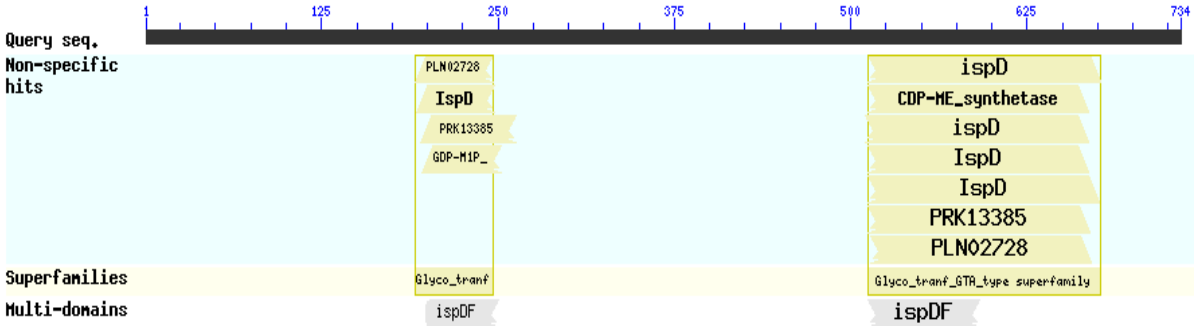


Figure 3.10: Conserved domain detection (CDD) of *Pf*IspD protein. The region spanning the amino acid position 191-677 of the query protein sequence showed a match with GT-A type super families.

3.2.6 Cloning and expression of *PvispD* gene

After the confirmation of the amplified product as *PvispD* gene by sequencing, we proceeded with the cloning of amplicons in pMAL-c2X expression vector at BamHI and PstI restriction enzyme sites (**Figure 3.11**) Further to screen the recombinant colonies containing desired gene, Gel Shift assays, and colony PCR were performed (**Figure 3.12**). The correct orientation of *PvispD* gene was confirmed using restriction digestion (**Figure 3.13**). Sequencing using M13 and gene specific primers confirmed the integrity of the gene in the recombinant clone and the presence of start codon in-frame.

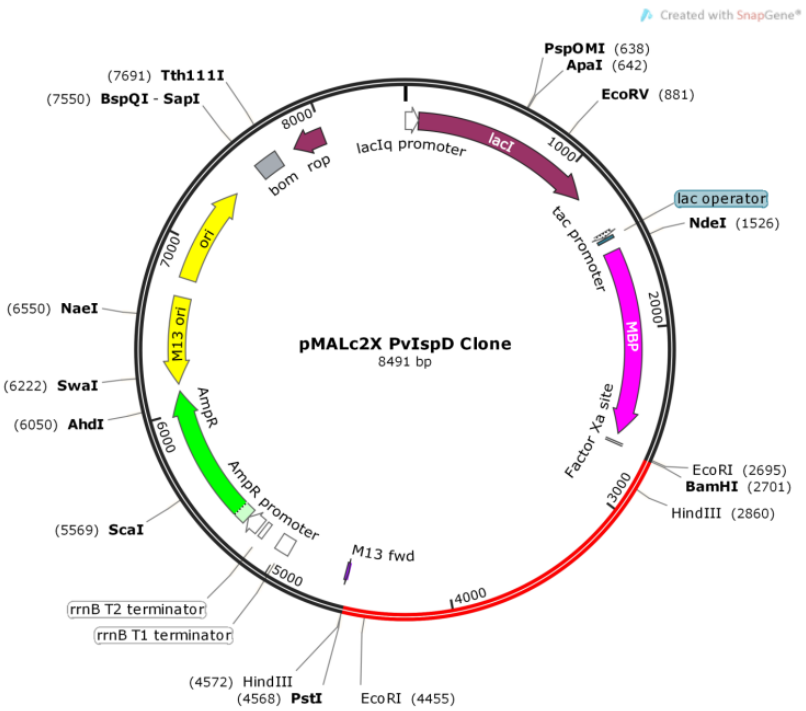


Figure 3.11: The pMALc2X – *PvispD* clone map: Red part depicts *PvispD* gene and black part shows the regions of pMALc2X vector. Restriction sites marked in the gene were used for restriction analysis of the clone.

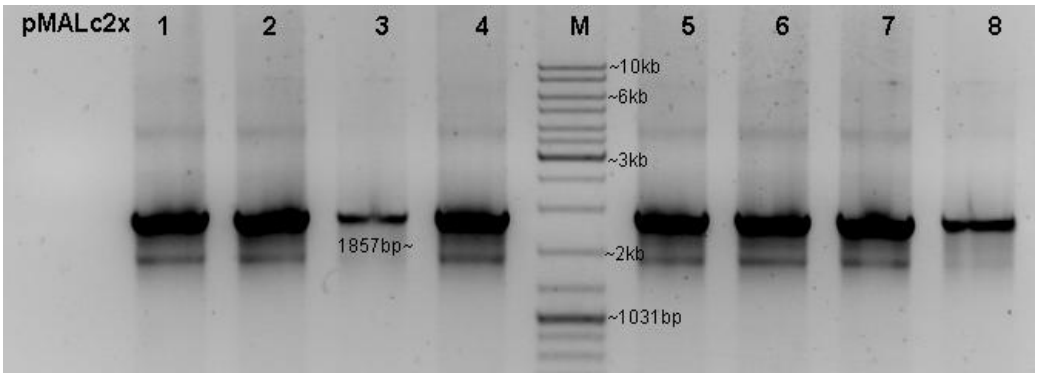


Figure 3.12: Colony PCR to check for the colonies showing recombinant construct (pMALc2X: Vector; 1-8: colony number after vector insert ligation and transformation; M: Gene Ruler DNA Ladder Mix)

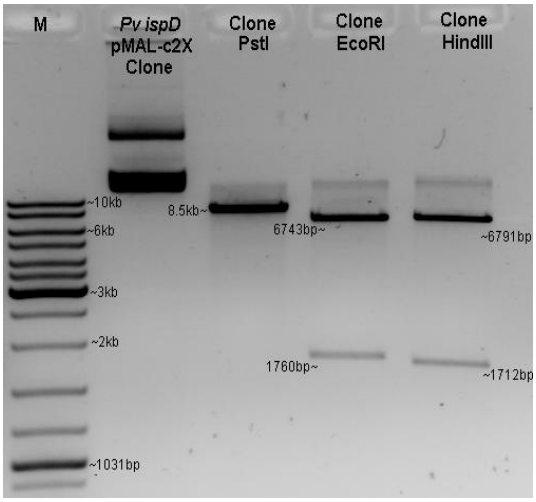


Figure 3.13: Restriction analysis of *PvispD* recombinant constructs. M: Gene Ruler DNA Ladder Mix (Thermo Scientific, USA); Lane 2: uncut pMALc2X *PvispD* clone plasmid; lane 3-5: Recombinant construct digested with PstI, EcoRI and HindIII restriction enzymes

After successful validation, the obtained recombinant clone was transformed into *E. coli* ER2523 host cells and protein expression studies (Chapter 2, 2.7) were performed. However, after repeated trials, protein did not expressed, which might be due to the presence of rare codons in the *PvispD* gene sequence (**Figure 3.14**). Thus, for protein expression pMAL-c2X *PvispD* construct was co-transformed with RIG plasmid (Baca & Hol, 2000) in *E. coli* ER2523 host cells. RIG plasmid consists of tRNA genes for Arginine (R), Isoleucine (I) and Glycine (G). The recombinant strain along with RIG plasmid was cultured in similar manner where bacterial cells were harvested by centrifugation, resuspended in lysis buffer and checked on SDS PAGE after staining with Coomassie brilliant blue R250 (**Figure 3.15**). The recombinant *Pv*IspD protein expressed as a fusion protein of ~118 kDa carrying a maltose binding protein of ~50 kDa.

atg aac acc ctg ctg agc gtt cgc ctg gtc act cgc ctg atc act ctc ctc ctt gcc cgc
ttc CCC ccg tgg tgc ctt tta att aac tat ttg ctt tcc ctg caa ggg tgc ctc CCC cat
ATA ctg cac AGG aag gac gca gca aca tgt gcc aag ctt AGA AGG gta cac cac gtg AGG
gta ttc cac gga cac CCC ggg ggg gaa gtt gcc tct tcc ctg cac gac atc aaa ggg ggg
gtc cat tta cca gct gcc aat cgg AGG aag gtg agc gat atg cag ttt att gtc acc agt
tgg acg ggt acc aaa cgg ggc gct act tac aaa gtg cgt ttt ttt aca cct gag cag ggg
AGG cag ttc gag ggg ggg gaa agc aaa cgg ggg cag ctg cat gcc acg gag ggg ggc gaa
aat ggg cgc gga gcg ggt gcc act gag cgg cac acg ccg aag cag cag acg cag acg cgg
aac CCC cgc ctc ttc gag aca acc gca aag aca aac gac cac gac CGA gag gaa tgc aaa
acc tat gaa cag tcg ctg cac aaa aaa aac atc cac gcc att ttg ctc tgc gga ggg atc
ggg caa AGA acc gag ttg gct agt CGA aag caa ttc CTA aag ctg aat gat gtc CCC ctc
ttc gtc tac tcc ttt aat ttg ttt gtg aaa tgc aat tta ATA aaa gct atc acc ctt gtg
tgc gac ccg aac tat ttc cag cac gtc atc gag agc att aac AGG cat aac gct tct ctt
ctg AGG AGG aag cac cag AGG ggg ttc ctg cgc AGG ggt cac tcg aag aat gtg ttg gcc
tcg tcg gga gag caa tcc gag ggg gat gcc agc gga gct ctc cat ttt tta aaa aaa aat
aaa tat atc ctg tat gat aat gag aag ggg aag tgc gtc acc aac ttg gat gag ctg ctc
agc gat gtg acg gcc acg aag ggg caa tac gtt tca gag gtt gac gag gtg gac gag gtg
aag cca ggc gat gtt aac tca aat cgg tac aag ctc ATA cgg ctg gtg gag agc ggc cgc
gag AGG gct gac tcc ctg ttg aat gcg ctg cgg ggg ctg gac ttg cgg gtg caa acc ggg
gag tac atc agc cag ctg ttg cgg ggg ggt ggc acg gac gaa gcg gga aaa ggg gat ggc
gcg gat aaa ggg gat aaa tcg gat ggt gtg gat ggg ggc cac CCC gct gac tgc aca CCC
atc tcg cac atc ctg gtg cac gac ggg gcg cgg CCC ttc ctg tcc gag ctc gac ctc ttc
aac ctg att tac atg gcc acc atc ggg cgc aat gcc att ctg gga tcg cgc gcc acg gac
acc att aag cgg ATA ggc acc gag cag caa ggg gag agt tgc CCC AGG gtg aag gcc cac
atg gat cgg caa ttc ATA ttc gca gct caa act cca cag ATA ttc agc agc caa gcg ttg
CTA caa gtg tgt gcg aag tta cct tcc AGA AGA gga ggg gaa acc cca gaa ggg agc AGA
gcc ttc acc gat acc tcc tcc CTA ttt caa cac gtt acc aaa aag aag gtg ttc gct ttg
cag gca aaa ttt CCC aat ttt aaa att aca acg ccg acg gat gtc ttc ctc gcc atc ttt
ctg atg gga tgc att ttt aaa act tct cat tcc gac gtg gac atg ggg atg ttt aaa gaa
acg ttc gtg aat tcc cca tct agc tgc gtg cct gcc aat cag ctg aac gac cat ttc ttt
tac cac tcc ctg ggg ggg aag cag cgc gtc ctg tac cgc cac ttt tac tac gag tag

Figure 3.14: Rare codon analysis of *Pv*IspD gene: Red = rare Arg codons AGG, AGA, CGA (19 in number); Blue: Rare Ile codon ATA (6 in number)

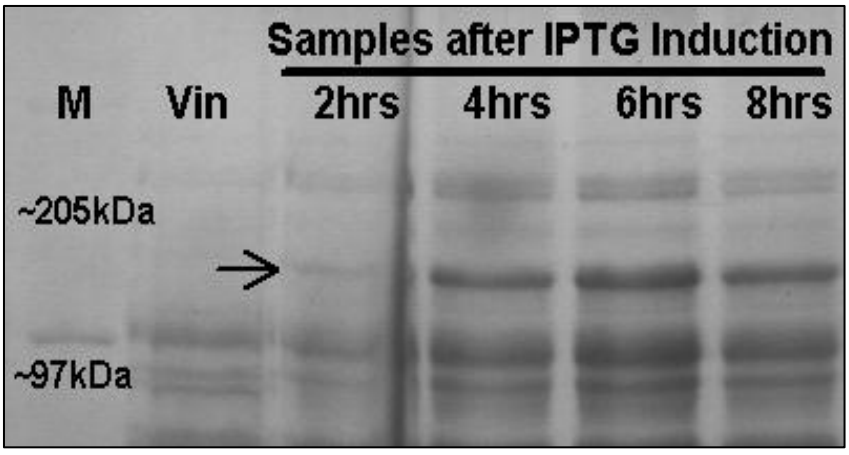


Figure 3.15: Over expression of full-length *Pv*IspD protein in *E. coli*. PMWH: Protein High Molecular Weight Marker (Merck Bio Science). (Vin: pMALc2X induced; pMALc2X *PvispD* clone induced samples 2-8 hrs).

The obtained expression for the recombinant protein was confirmed by western blotting (**Figure 3.16**) using anti-MBP antibody (NEB, USA) as per the manufacturer's protocol

where a band at ~50 kDa in the vector induced sample shows the presence of maltose binding protein.

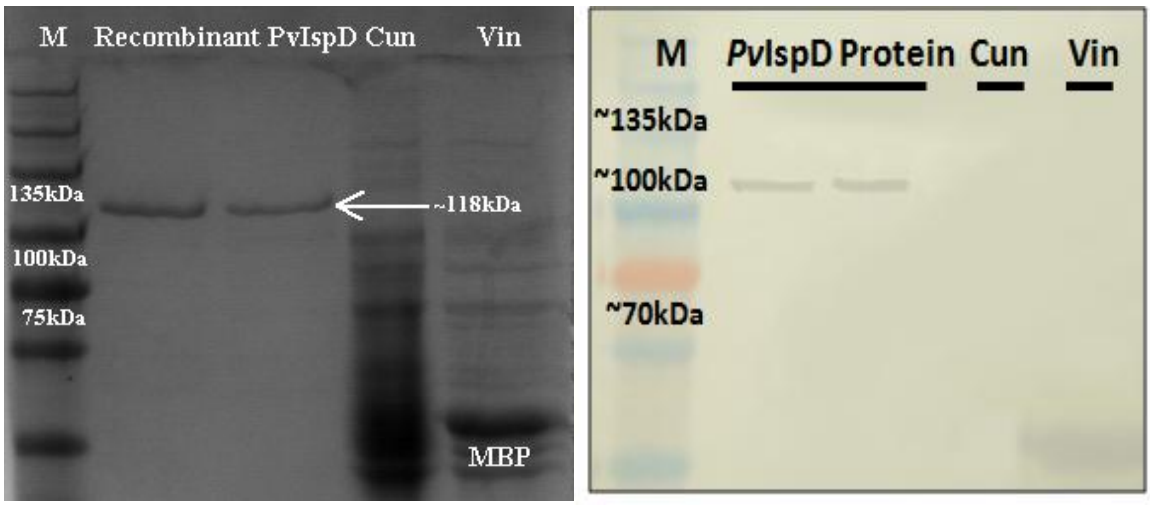


Figure 3.16: Expression of MBP tagged *Pv*IspD protein confirmation by western blotting. (M; BLUelf Prestained Protein Ladder, BM008-0500; BR BIOCHEM); Cun; Clone uninduced; Vin; Vector induced)

3.2.7 Protein purification

To perform enzyme kinetics studies and subcellular localization, pure protein is required. Thus, the *Pv*IspD recombinant protein was purified using amylose resin following the protocol discussed in materials and methods (Chapter 2, 2.9). Multiple fractions (**Figure 3.17**) of purified protein were analysed. The fractions with maximum intensity for the desired band and minimum non-specific bands were included in further experiments.

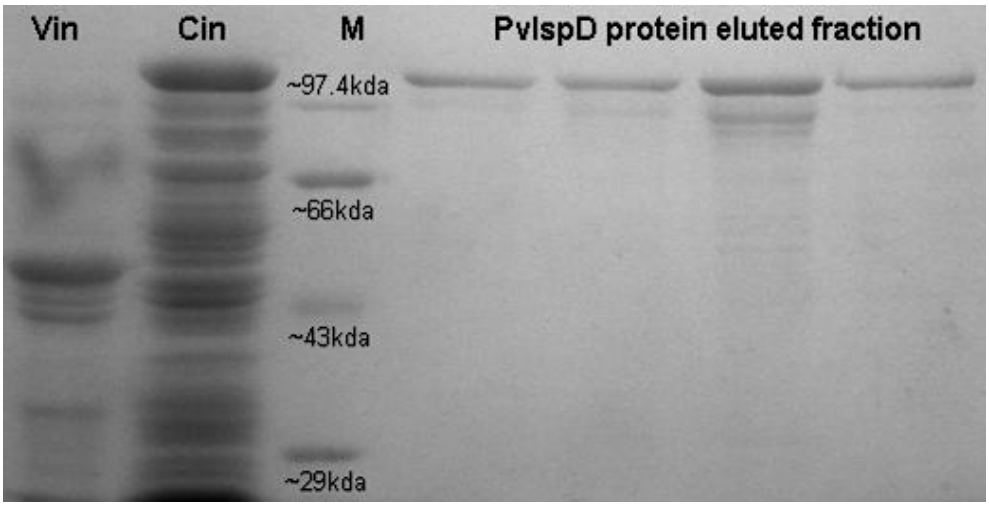


Figure 3.17: Purification of MBP tagged *Pv*IspD protein. PMWH: Protein High Molecular Weight Marker (Merck Bio Science). (Vin: pMALc2X induced; Cin: pMALc2X *PvispD* induced; *Pv*IspD protein eluted fractions).

3.2.8 Enzyme Kinetics studies for *Pv*IspD protein

Activity of the *Pv*IspD enzyme was determined by an assay based on the colorimetric determination of one of the reaction products, inorganic pyrophosphate, with malachite green. The standard CDP-ME synthase activity reaction mixture contained 50mM Tris–HCl (pH 7.9), 10mM $MgCl_2$, 0.5mM MEP, 0.5mM CTP, 1mM DTT, 20mM Na_2F and varying concentrations of purified recombinant IspD protein (1-20µg) in a final volume of 100µL. The above reaction mix was incubated at 37°C for 40 min. Reactions were terminated by the addition of 10µL of β-mercaptoethanol, and absorption was taken using a STAT FAX®2200 micro plate reader after adding 40µL coloration reagent; containing malachite green dye in ammonium molybdate solution. To find the concentration of enzyme at which maximum reaction takes place, the absorbance observed for different concentrations was plotted against the concentration of the enzyme. It was found that the enzyme shows maximum activity at a concentration of 67.8pM (**Figure 3.18**)

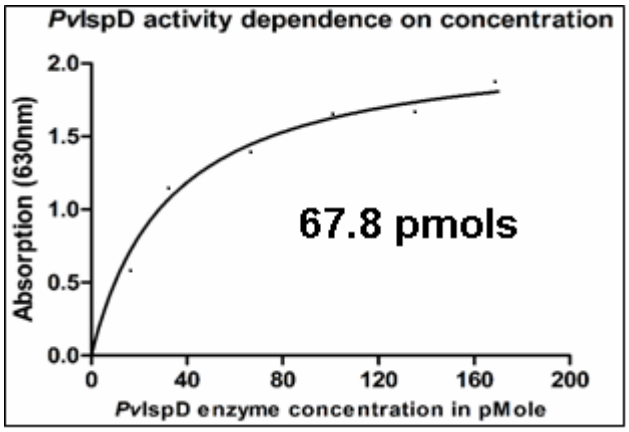


Figure 3.18: *Pv*IspD enzyme activity and its dependence on concentration

Activity of *Pv*IspD enzyme is highly dependent upon its two substrates; MEP and CTP. To measure the effect of MEP concentration on *Pv*IspD activity, the concentration of CTP was kept constant at 1mM and concentrations of MEP varied from 1mM to 3.9µM (through serial, twofold dilutions), maintaining the concentration of enzyme as 67.8pM; the concentration at which it showed maximum activity. Similarly, effect of CTP was determined by using MEP at a constant concentration of 1mM and varying concentration of CTP from 1mM to 15.63mM. The K_M values were determined by plotting the rate as a function of concentration of the variable substrate at the same enzyme concentration using nonlinear regression analysis (Graph Pad, USA). With this study, K_M^{MEP} and K_M^{CTP} calculated for *Pv*IspD were 25.13 ± 5.34µM and 23.05 ± 8.57µM, respectively (**Figure 3.19a & 3.19b**).

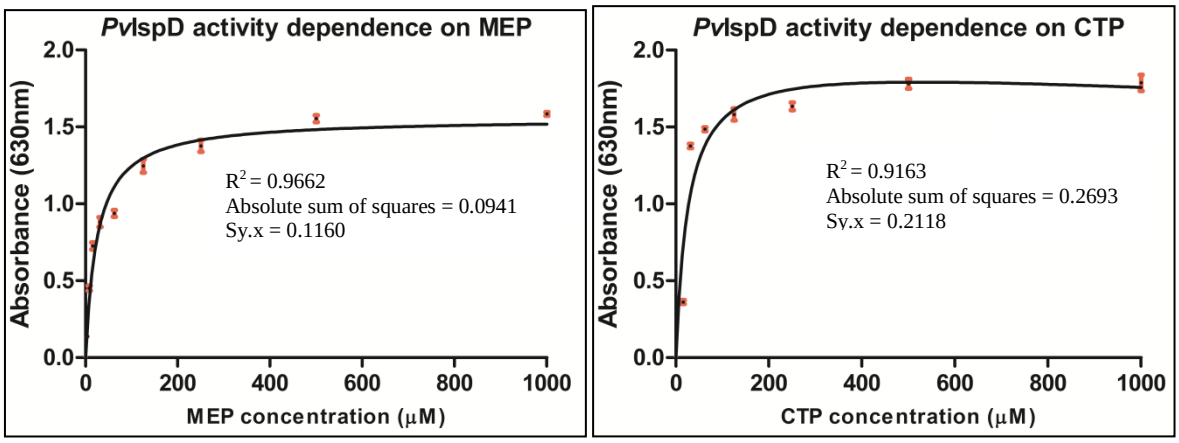


Figure 3.19: Enzyme kinetics studies for *Pv*IspD protein. (a) MEP (b) CTP

3.2.9 Inhibition studies on *Pv*IspD

As *Pv*IspD is been looked upon as a potential drug target, using HTS (as discussed in Chapter 2) we screened certain molecules, especially antibiotics that can act as an inhibitor of this enzyme. To shortlist the molecules, first we generated the three-dimensional model of *Pv*IspD and after validation and energy minimization studies performed docking of these antibiotics to this structure (the modelling and docking details are discussed in Chapter 4). Few molecules that showed similar binding residues as the two substrates were finally shortlisted for the biochemical inhibition assays.

Based on the common interacting residues between the drug molecule and substrate as predicted by structural and docking studies, three molecules were selected for the i*n-vitro* inhibition studies: Domiphen bromide (DMB; tried against bacterial IspD), Cefepime (first time on IspD enzyme), Rifampicin (shown to inhibit parasite *in vitro*, but with no study on IspD) and Fosmidomycin (reported as *Pf*IspC inhibitor). For the reaction, those concentrations of the *Pv*IspD enzyme were used at which the enzyme showed maximum activity. The reaction mix consisted of 67.8pM IspD enzyme, 500µM CTP, 250µM MEP, 50mM Tris–HCl (pH 7.9), 10mM $MgCl_2$, 1mM DTT, 20mM Na_2F with varying concentrations of selected inhibitors. The concentration of inhibiting compounds used are as follows:

- DMB (6.25, 12.5, 25, 50, 100, 200, and 400 µg/mL)
- Cefepime (0.25, 0.5, 0.75, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0 µg/mL)
- Rifampicin (0, 25, 50, 75, 100, 250 and 500 nM)
- Fosmidomycin (0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 µM)

The IC_{50} value for DMB, Cefepime, Rifampicin and Fosmidomycin were calculated as 169 ± 5nM, 2.33 ± 0.139nM, 75.88 ± 1.19nM and 1.5µM respectively at 67.8 pmols concentration of enzyme as per the data shown in **Figure 3.20a, 3.20b, 3.20c & 3.20d**. This study shows that

*Pv*IspD enzyme is sensitive to all the four compounds i.e. DMB, Cefepime, Rifampicin and Fosmidomycin, and thus, these drugs may block the catalytic activity of *Pv*IspD.

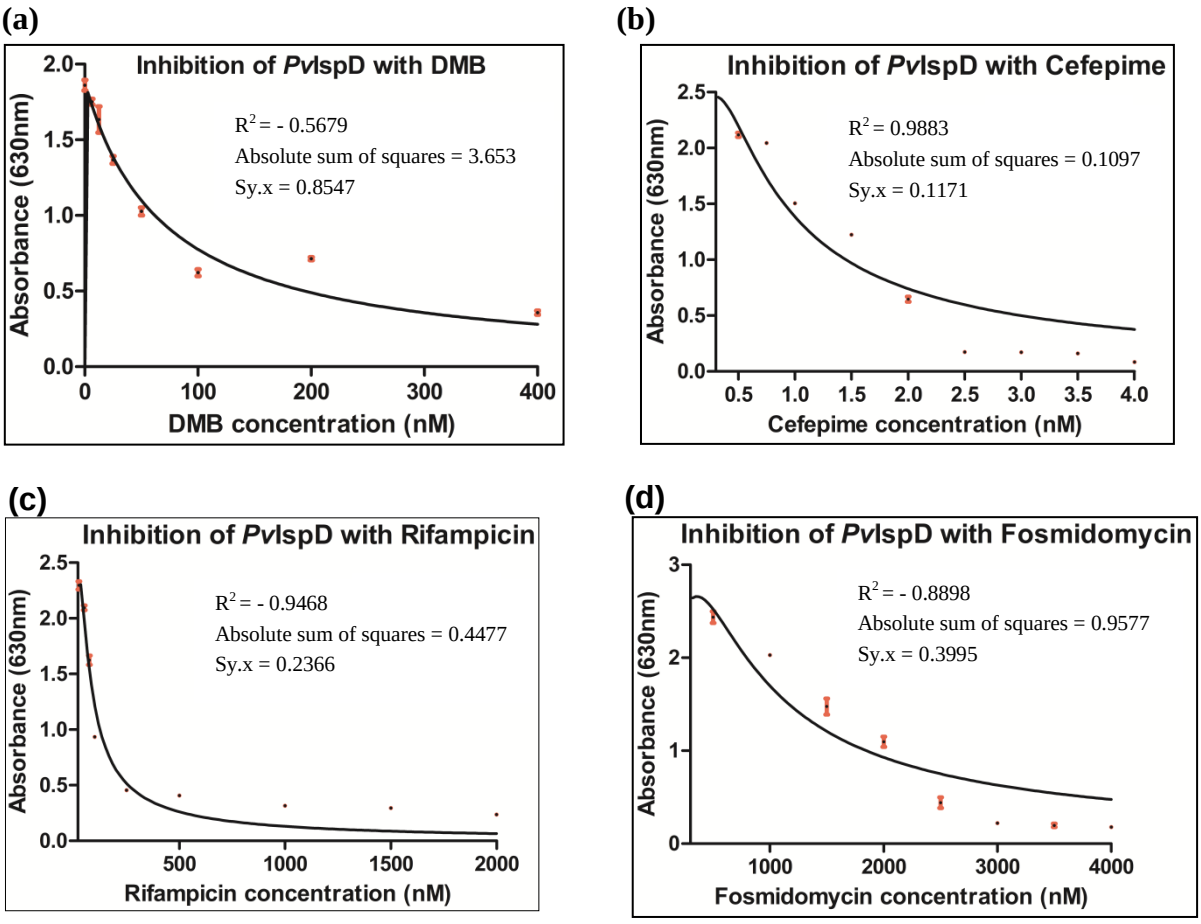


Figure: 3.20 Inhibitory effects of different drug on molecules on the activity of *Pv*IspD protein.

(a) Domiphen bromide; (b) Cefepime; (c) Rifampicin; (d) Fosmidomycin

3.2.10 Antibody raising and Immuno-localization of *Pv*IspD protein

In-silico prediction studies to analyse the targeting of *Pv*IspD gave inconclusive results, thus to confirm the subcellular localization of *Pv*IspD, mice antiserum was produced against full length recombinant *Pv*IspD protein following the procedure as detailed in Chapter 2 (2.12). Antibody titer was calculated (Harlow and Lane, 1988) for the sera collected after different booster doses of *Pv*IspD protein using standard ELISA where serum collected from mice injected with 1X PBS was used as a control. Pre-immune sera was collected to check if any cross-reactive antibodies were present before immunization. Maximum antibody titre for *Pv*IspD antiserum was obtained after the third booster dose (**Figure 3.21**). Specific band obtained at the desired position in western blotting also confirms the specificity of the antibodies towards *Pv*IspD protein (**Figure 3.22**).

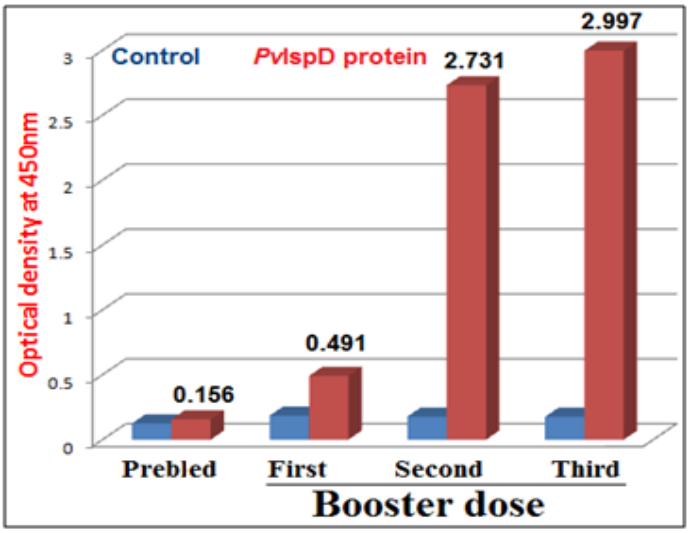


Figure 3.21: Antibody titre for *Pv*IspD protein by performing ELISA.

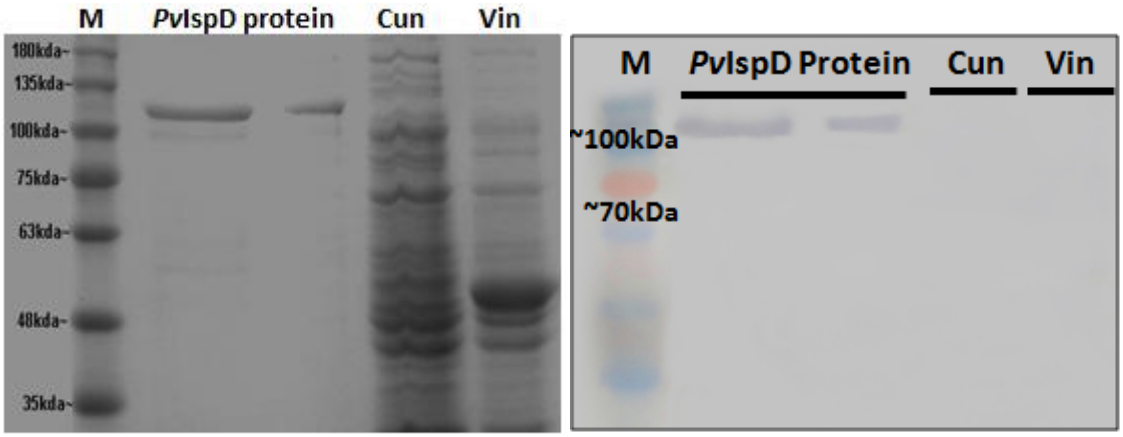


Figure 3.22: Expression of MBP tagged *Pv*IspD protein confirmation by western blotting (M; BLUelf Prestained Protein Ladder, BM008-0500; BR BIOCHEM; Cun; Clone uninduced; Vin; Vector induced).

To perform sub-cellular localization on the blood smears prepared from malaria-infected patients, the parasite cells were fixed, permeabilised and sequentially treated with mouse polyclonal anti-*Pv*IspD sera (1:500 dilution), following the protocol as discussed in Chapter 2. A green fluorescing spot towards the apical end of the parasite with an overlapping red spot indicated the possible presence of *Pv*IspD protein (green) along apicoplast (red), respectively. Nucleus was evident as a blue spot adjacent to this merged yellowish spot. In one of the fields analyzed, we could identify a green spot in the trophozoite as confirmed by bright field, while in another field there were a number of green fluorescing spots indicating a schizont, as confirmed in a bright field (**Figure 3.23 & 3.24**). The overlapping red spot to all these green spots confirms that *Pv*IspD protein is functionally active in the apicoplast.

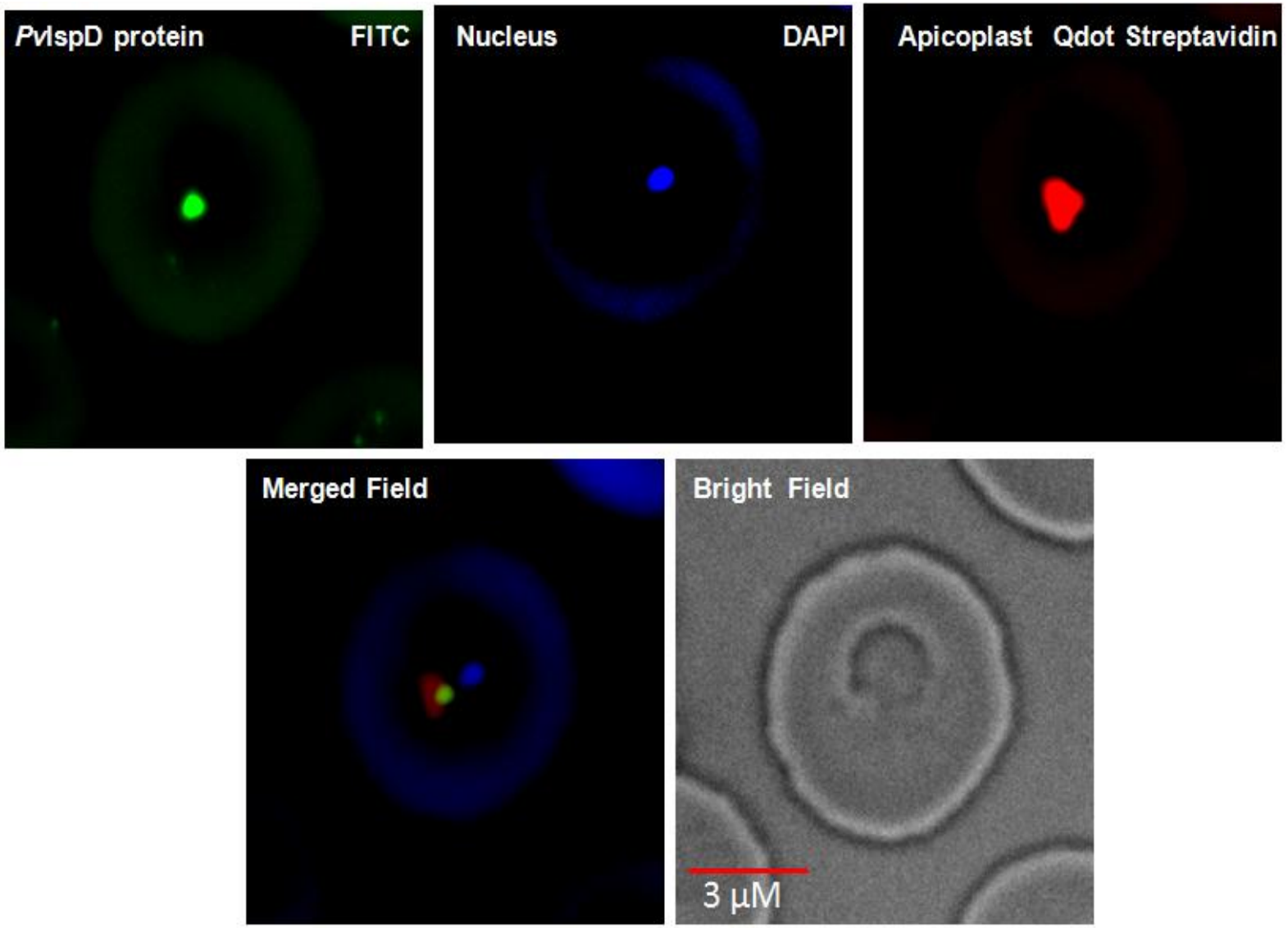


Figure 3.23: Sub cellular localization of *Pv*IspD protein in *P. vivax* Indian field isolates. The anti-*Pv*IspD antibodies (raised from recombinant protein) tagged with FITC (visualized as a green fluorescence). The parasite nucleus was stained with DAPI and apicoplast with Qdot® 585 Streptavidin conjugate. Image z-stacks were deconvolved and then presented as a single combined image. The merged image shows the co-localization of *Pv*IspD to the apicoplast (yellow) of the parasite.

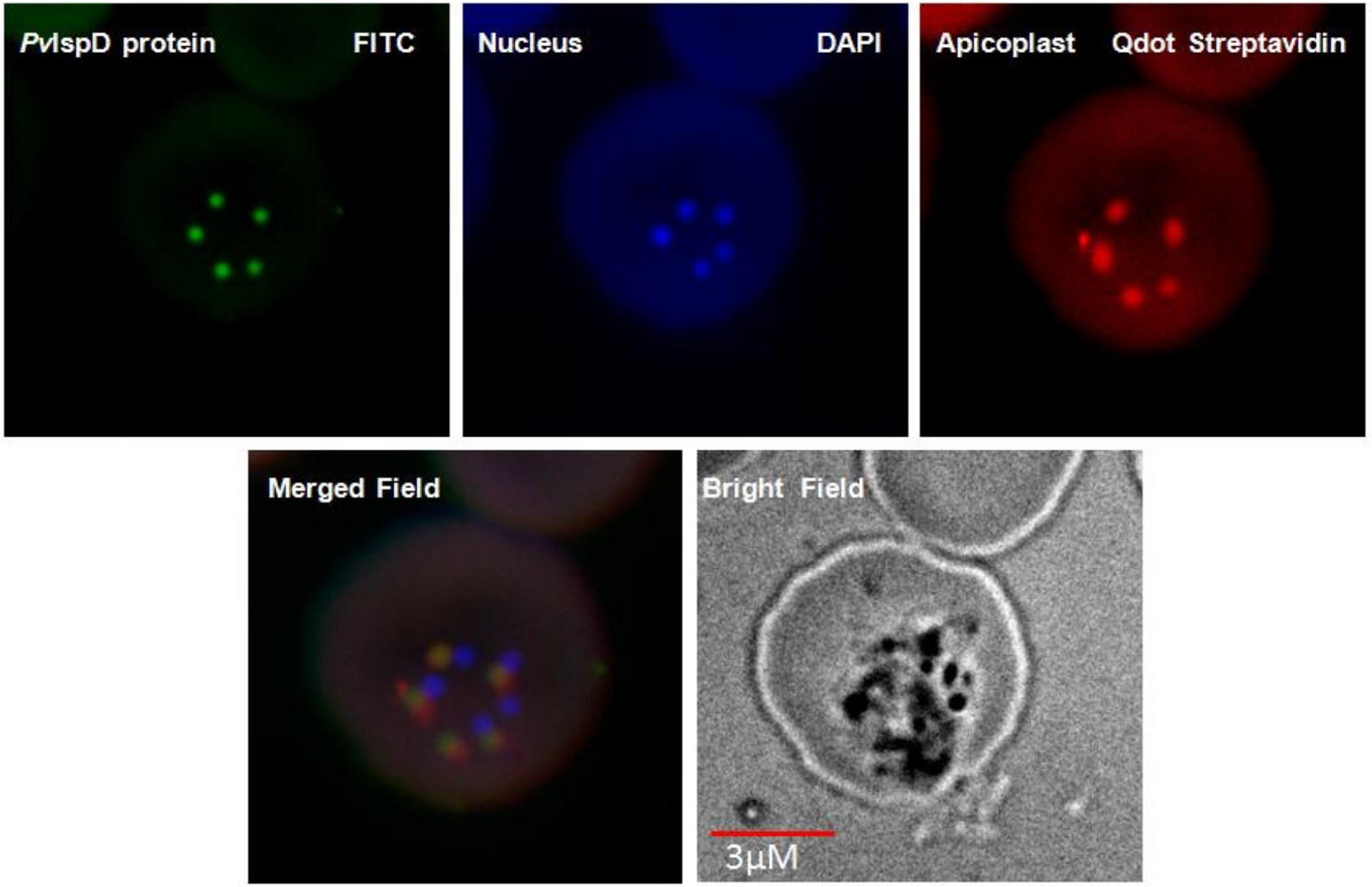


Figure 3.24: Sub cellular localization of *Pv*IspD protein in *P. vivax* Indian field isolates. The anti-*Pv*IspD antibodies (raised from recombinant protein) tagged with FITC and visualized as a green fluorescence. The parasite's nucleus was stained with DAPI and apicoplast with Qdot® 585 Streptavidin conjugate. Image z-stacks were deconvolved and then presented as a single combined image. The merged image shows the co-localization of *Pv*IspD to the apicoplast (yellow) of the parasite.

3.2.11 Real Time PCR analysis

The mRNA abundance of genes involved in isoprenoids biosynthesis pathway varies at different blood stages of *P. vivax* (Bozdech et al., 2008). IspD gene is shown to be involved in the metabolic regulation of the isoprenoids biosynthesis pathway and currently is looked upon as a potential drug target. However, till date whatever studies have been reported for this gene involves *in vitro* culture only. There are no reports on this gene detailing its expression in clinical isolates. Even Bodzech et al., who has performed the expression analysis of three field isolates of *P. vivax* has cultured these isolates for short term to synchronize the stages. Thus, we have analyzed the transcript level of the IspD gene by qPCR to analyze its expression in different field isolates of *P. vivax*, specific to the study region Bikaner. For this, RNA was isolated from parasite-infected blood collected from clinical isolates having different erythrocytic stages (**Table 3.6**). This RNA was used for cDNA preparation as described before (Chapter 2). The sequences of the primers used is as given in Table 3.6. Absolute expression of *Pv*IspD gene was calculated using ΔCT method by using Seryl tRNA Synthetase as a normalizing factor.

Table 3.6 Primers for quantitative PCR of *ispD* gene from *P. vivax* genome

Lab Id	Primer sequence	Orientation
VS7	5’ GAT AAT GAG AAG GGG AAG TGC 3’	Forward primer
VS6	5’ GAT TTA TCC CCT TTA TCC GCG 3’	Reverse primer
*Pv*STSF	5’ CAT CTC CTG AAC GGC ACC AT 3’	Forward primer
*Pv*STSR	5’ GGT GAA GGG AAT AAA CTC GAC G 3’	Reverse primer

Table 3.7 Parasites samples with their erythrocytic stages and ΔCT values

S. No.	Field Isolates	Stages as per slide (approx.)	ΔCT values
1	A	80% trophozoite, 20% schizont	5.29
2	B	70% rings, 30% trophozoites	1.52
3	C	20% rings, 50% trophozoites, 30% schizont	2.38
4	D	80% Schizont, 20% ring	0.21
5	E	70% Schizont, 30% ring	0.41
6	F	90% trophozoites, 10% schizonts	8.14

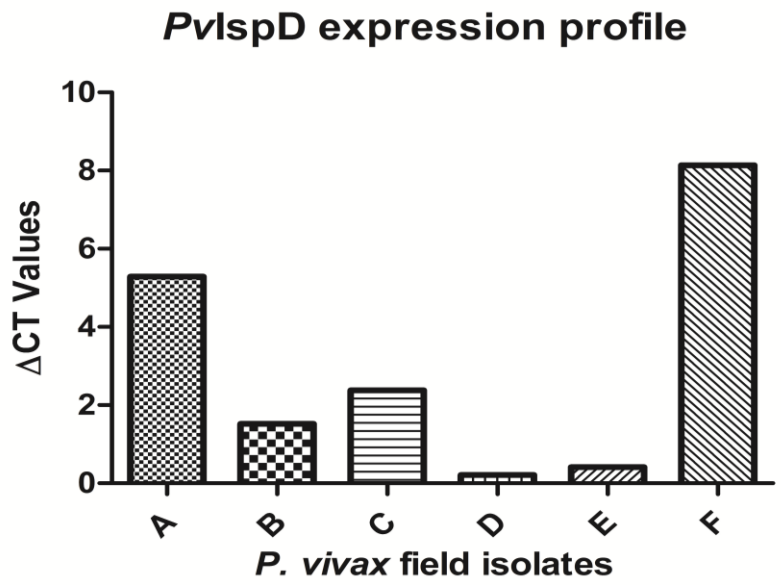


Figure 3.25: Quantitative PCR analyses for *Pv*IspD protein (X axis; *P. vivax* field isolates with different erythrocytic stages; Y axis: ΔCT values of *Pv*IspD w.r.t. housekeeping gene Seryl-tRNA synthetase).

Based on the qPCR data, we have observed the expression of *Pv*IspD gene in all the samples that have shown the mixed erythrocytic stages as per the PBF examination (Table 3.7). However, the expression was found higher in sample (A, F) than (B and C), while there was just basal level expression in D and E. Based on the slide data A and F has shown maximum of trophozoite stages with few schizonts, while samples D and E has shown maximum of Schizont stages with few early rings. Overall combining the qPCR and slide data, our results indicate that the expression of *Pv*IspD gene is maximum at the trophozoite stages and lowest at the schizont and early ring stages. Study by Bozdech et al., 2008 have shown that the expression of other genes of the isoprenoids biogenesis pathway (IspD gene is not analysed in this study) peaks at trophozoite stages in all the three isolates (smru 1, 2 and 3) and low level expression at early ring or late schizont stages in any of the three isolates. A separate microarray study done by Boopathi et al., (communicated, 2016) has shown the expression of *Pv*IspD gene in majority of the samples (12 out of 14 samples). These samples when compared with Bozdech data have shown the presence of mixed stages. This indicates that the gene is expressed at all the stages of this parasite. Our study agrees with Bozdech data showing maximum expression at trophozoite stages as well as the microarray data suggesting the expression of this gene in all the isolates. However, adding more no of samples with varying percentage of blood stages might help to reveals the requirement of the gene in stage specific manner.

3.3 Conclusion

The IspD enzyme is one of the major catalytic enzyme functional in the MEP pathway, where the catalytic site of IspD enzyme is highly specific for pyrimidine nucleotide tri-phosphates. Recent studies have gained a lot of attention towards this enzyme due to its potential as a drug target and probable regulatory role in the isoprenoids biosynthesis pathway. In this study, we have characterized the *Pv*IspD protein from the major human malaria parasites *P. vivax* and *P. falciparum.* During the course of this study, there had been discrete reports on synthetic *PvispD* and *PfispD*, where the inhibition of *Pf*IspD by MMV008138, a compound from Malaria Box was shown *in vitro*. However, this compound has very low inhibitory effect on *P. vivax*.

The current study is the first report of IspD enzymes from the *Plasmodium* field isolates. The conservation of the CDP-ME synthetase domain belonging to a GT-A superfamily and the two signature motifs GXG and [IVT]-[LIVMC]-[IVT]-[HS]-D-[SGAV]-[AV]-R from all *Plasmodium* species, apicomplexans and prokaryotes indicate a similar function for this enzyme in all these organisms. Even with a lot of difference between the sequences of this enzyme from two major human *Plasmodium* species i.e. *P. falciparum* and *P. vivax,* these domains were found conserved. The purified protein retained the catalytic activity as was evident from our results, which are in-line with the enzymatic studies done on synthetic modified gene and assays reported on prokaryotic orthologues.

As the isoprenoids biosynthesis pathway is functional in the apicoplast, we initially predicted the targeting of IspD using available bio-informatics software. However, the results obtained by different software's gave contradictory results. Thus, we proceeded with the immune-localization studies. We were able to co-localize the *Pv*IspD antibodies to the apicoplast on the *P. vivax* infected blood smears obtained from the field, indicating possible presence of this protein. The real-time expression profiling also suggested that the protein is transcriptionally active throughout the erythrocytic stages.

The essential nature of IspD in *Plasmodium* parasites with no orthologues in humans indicates the potential of this enzyme as a future drug target after the first two members DXS and DXR (IspC). To strengthen this view, we tried inhibition of *Pv*IspD enzyme with different antimicrobials like DMB, Cefepime and Rifampicin. There was strong inhibiton of the enzyme indicating that these compounds or their derivatives may be used to target this

enzyme. However, this result needs further confirmation by drug assays in *Plasmodium in-vitro* cultures, which is difficult in case of *Plasmodium vivax.* Thus, overall the present study suggests that IspD is functional in *Plasmodium vivax* and is a potential candidate for designing new antimalarial.

Chapter IV

In-silico studies of Isoprenoids biosynthesis pathway enzyme IspD from *P. vivax* and *P. falciparum*

Functional importance of isoprenoids biosynthesis pathway and its enzymes is discussed in previous sections. This chapter focuses on the *in-silico* characterization of *P. vivax* and *P. falciparum* IspD enzymes detailing their structure, binding to natural substrates and inhibitors including various chemical compounds or antimicrobial agents.

IspD enzyme is being looked upon as a putative drug target. Till date two inhibitors have been detailed against IspD, Azolopyrimidine in *A. thaliana* and MMV001838 (compound from Malaria Box) in *Plasmodium*. However, MMV001838 has been shown less effective against *P. vivax.* Thus, to extend our knowledge on these inhibitors in both *Plasmodium* species as well as to look into new probable inhibitors, we have studied the inhibitory mechanism of these molecules and different anti-microbial agents previously reported against other pathogenic microorganisms, but with unknown/ different target. Our results suggest the importance of these compounds as potential inhibitors of *Plasmodium* IspD.

4.1 Introduction

The enzymes of the non-mevalonate pathway are interesting targets for anti-parasitic drugs, as their homologs are absent in mammals and therefore target related toxicological side-effects by inhibitors of this pathway are not expected. IspD is the third enzyme participating in the MEP pathway that catalyzes the cytidylation of MEP and can act as an alternate or additional drug target to its predecessor IspC. Relatively few IspD homologs have been isolated and characterized including bacteria such as *E. coli* (Rohdich et al., 1999) and *M. tuberculosis* (Eoh et al., 2007). Some reports detail the bifunctional IspDF enzyme in organisms such as *C. jejuni* (Gabrielsen et al., 2004) and *H. pylori* (Perez-Gil et al., 2010) in which both proteins are encoded in the same transcription unit, forming a single enzyme. In this unit the N-terminal domain has cytidyltransferase activity (IspD) and C-terminal has MECP synthase activity (IspF). Investigation of *C. jejuni* IspDF crystal structure revealed the presence of IspDF in the form of a tetramer whereas IspD and IspF enzymes in other organisms are present separately in the form of dimer and trimer respectively (Richard et al., 2001 & 2002). In *E. coli* and *M. tuberculosis*, the genes encoding IspD and IspF are in close proximity to each other on the bacterial chromosome, but they are not expressed as a single unit. In *E. coli*, IspD and IspF can interact only in the presence of IspE enzyme. Thus, it was suggested that there may be a thermodynamic benefit, responsible for the fusion of IspD and IspF enzyme where this fusion would reduce the entropy of dissociation thereby reducing the association free energy.

The essentiality of IspD was proven initially in plants by antisense experiments, thus, inhibitors of this enzyme were anticipated to be potentially herbicidal. To look into the inhibitors, after high throughput screening molecular library of natural products was constructed and with the help of LC-MS, binding of these compounds with IspD enzyme was analyzed. Of all the molecules tested, only one molecule; 7-hydroxy-[1,2,4] triazolo [1,5-a] pyrimidine was identified as the most active inhibitor with an IC_{50} value of 140 ± 10 nM in *A. thaliana* (Witschel et al., 2011). In prokaryotes, L-erythritol-4-phosphate (Ki = 240 ± 17mM) and D-erythritol 4-phosphate (IC_{50} = 1.36mM) are reported as weak inhibitors against IspD enzyme (Lillo et al., 2003). Recently, in *Plasmodium* using codon optimized synthetic genes of *Pf*IspD and *Pv*IspD, MMV008138, shortlisted compound from Malaria Box (Spangenberg et al., 2013), had shown inhibition of above enzymes under *in-vitro* conditions (Imlay et al., 2015). However, till date there is lack of any study detailing the structure and active sites of above enzymes, especially *Pv*IspD.

In the current study, we have generated the three dimensional structure of *Pf*IspD and *Pv*IspD molecules using homology modelling. These structures were used for docking of different drug molecules and analysed for their effect on substrate binding to shortlist potential candidates for IspD enzyme inhibition.

4.2 Results and Discussion

4.2.1 Secondary structure prediction

To generate the model of *Pv*IspD and *Pf*IspD proteins, the first step performed was secondary structure prediction using PSIPRED. In *Pv*IspD protein, presence of 16 α-helices and 23 β-strands were observed which were joined with the help of 41 coils; whereas in *Pf*IspD 24 α-helices, 18 β-strands and 42 coils were present (**Figure 4.1 & 4.2**). The intrinsic disordered profile was analysed to verify the stability of the structure and their interactions with environment. Information related to characteristic features and chemical properties of *Pv*IspD and *Pf*IspD proteins provided by PSIPRED are enlisted in **Table 4.1.**

Table 4.1: Characteristics features of *Pv*IspD and *Pf*IspD protein analysis based on Secondary structure

Characteristic features	Value	
	***Pv*IspD**	***Pf*IspD**
Molar extinction coefficient	30580	62850
Molecular weight	69.05	88.24
Isoelectric point	9.80	9.91
Percent negative residues	0.10	0.13
Percent positive residues	0.14	0.19
Charge	35.56	52.19
Signal Peptide	1-13	Not detected
Transmembrane topology	16-34	223-238
Aliphatic index (0-100)	80.16	81.93
Hydrophobicity (-4 to 3)	-0.43	-0.92

Secondary structure of *Pv*IspD and *Pf*IspD protein was also compared with orthologues of IspD available from different organisms including apicomplexans, plants and prokaryotes where high level of variation was observed (http://www.geneious.com, Kearse et al.,

2012). This deviation was mainly because of the high variation in sequence length of the IspD protein obtained from different organisms. A conserved signature motif was observed at the C-terminal of *Pv*IspD and *Pf*IspD proteins (**Figure 4.3**).

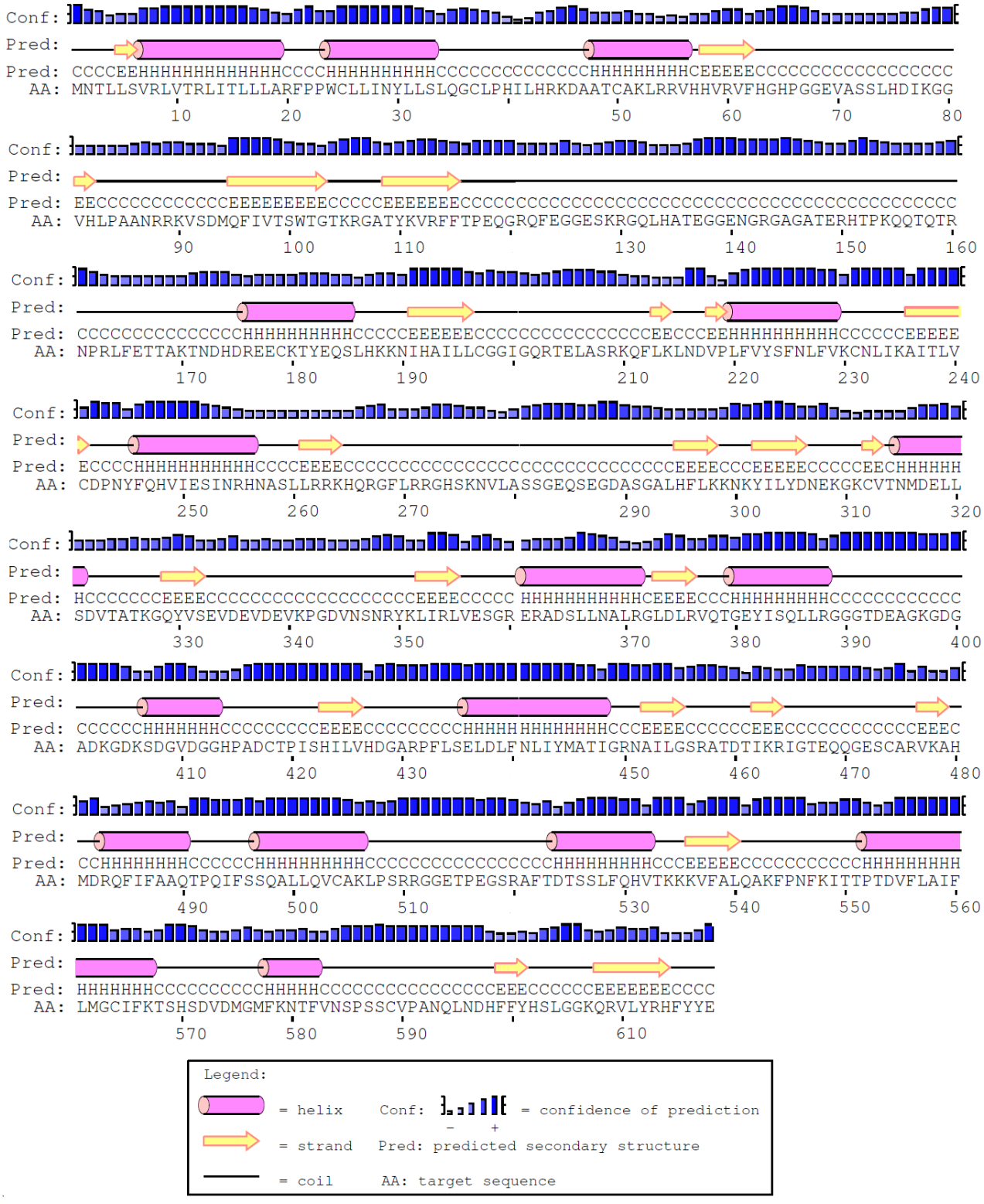


Figure 4.1: Secondary Structure predictions for *Pv*IspD protein using PSIPRED.

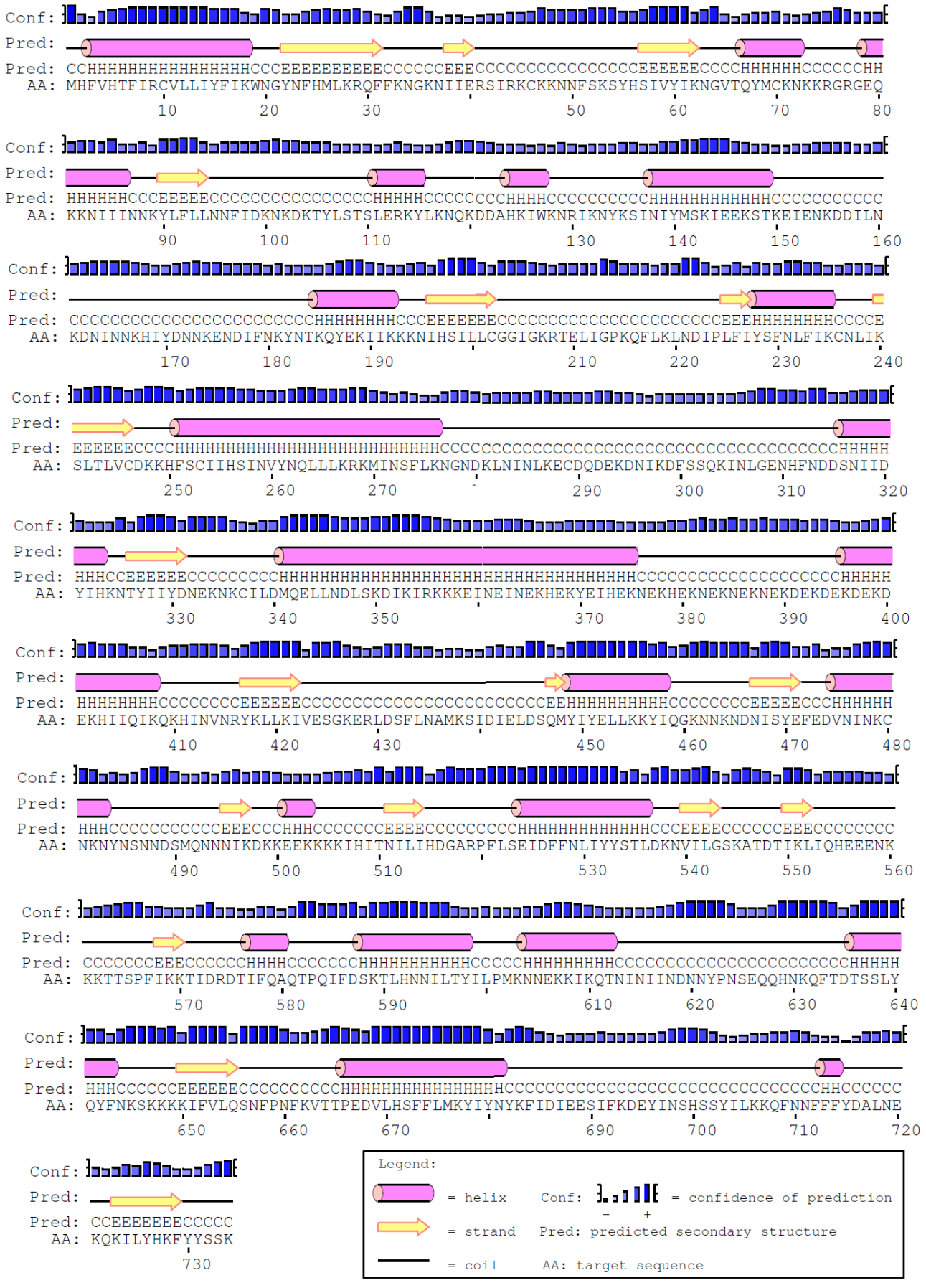


Figure 4.2: Secondary Structure predictions for *Pf*IspD protein by PSIPRED.

Figure 4.3: Comparison of IspD protein secondary structure of *P. vivax* with other organisms. [Consensus (multiple coloured bar), mean hydrophobicity (red bar) and identity (green bar) between IspD amino acid sequences obtained from different organisms. The secondary structures visible include pink cylinders as alpha helix, yellow arrows as beta sheets and blue arrows as turns in beta strands. The red block denotes the conserved domain present in IspD proteins. Figure generated with the help of GeneiousII.

IspD is suggested to regulate the metabolic flux in the MEP pathway by phosphorylation of Thr141 in *Francisella tularensis* (bacteria), which has been shown as the key component for cytidylation reaction. Its phosphorylation may lead to negative effect on the activity of the enzyme, disrupting the enzyme-substrate interaction. Additionally, mutation at this threonine leads to reduce or loss of enzyme activity (Tsang et al., 2011). During multiple sequence alignment, this Threonine was found conserved in apicomplexans like *Plasmodium* and *Babesia*, prokaryotes like *E. coli, T. thermophilus* and *B. subtilis*, and was present near the interface of the β-arm. However, in *T. gondii* there was a cysteine present in place of threonine.

4.2.2 Comparative modelling and Molecular simulation

Three-dimensional (3D) model of *Pv*IspD and *Pf*IspD proteins were built by homology modelling using the program MODELLER 9v11 (Sali and Blundell, 1993; Eswar et al., 2006). The IspD amino acid sequences obtained from *Plasmodium* Indian field isolates were submitted to HHpred online server for template search. Crystal structure of *E. coli* IspD (PDB; A chain of 1I52) protein with 1.5 Å resolution was selected as a template. However, as the cyrstallization of this enzyme was performed in the presence of substrate, coordinates for the substrate molecule were removed before using this structure as a template. The A chain of *E. coli* IspD (1I52) protein is of 236 amino acids and comparatively smaller to sequences of *Pv*IspD (618 aa) and *Pf*IspD (734 aa). IspD sequences of *Plasmodium* parasite shares only 26% amino acid identity with *E. coli* IspD, mainly in the conserved domain region. The major difference in the template and target sequences was because of the presence of apicoplast targeting sequences on the N-terminal of the target. This sequence is cleaved as soon as these proteins reach their site of action and thus, this region might not be important for the catalytic activity of these enzymes. Taking care of these points, after the removal of targeting sequences selected region was submitted for protein modeling, which yielded 3D models of the targets with all non-hydrogen atoms without any user intervention.

The best model with minimum DOPE score was selected and quality of the structure was further improved by energy minimization and molecular simulation using GROMACS 4.6 software package (Van Der Spoel et al., 2005). The structure was evaluated by using different servers where PROCHECK showed a G-score 'BETTER' while an average profile was obtained with VERIFY3D. Identification of bad packaging of side chain atoms or unusual residue contact in the IspD structure when checked using WHAT IF program,

gave a Z-score indicating the good quality of the structure. The main chain conformations for 99% of amino acid residues were within the favoured or allowed regions for both *Pv*IspD and *Pf*IspD protein in the Ramachandran plot analysis (**Figure 4.4)**. The ERRAT score for *Pv*IspD and *Pf*IspD protein obtained was 97 and 90.5, which indicated that the molecular geometry of the models is of good quality (**Figure 4.5**) Root-mean-square deviation (RMSD) value of 0.583 Å and 0.824 Å between the backbone atoms of the template and *Pv*IspD and *Pf*IspD models indicated the significant homology between the structures.

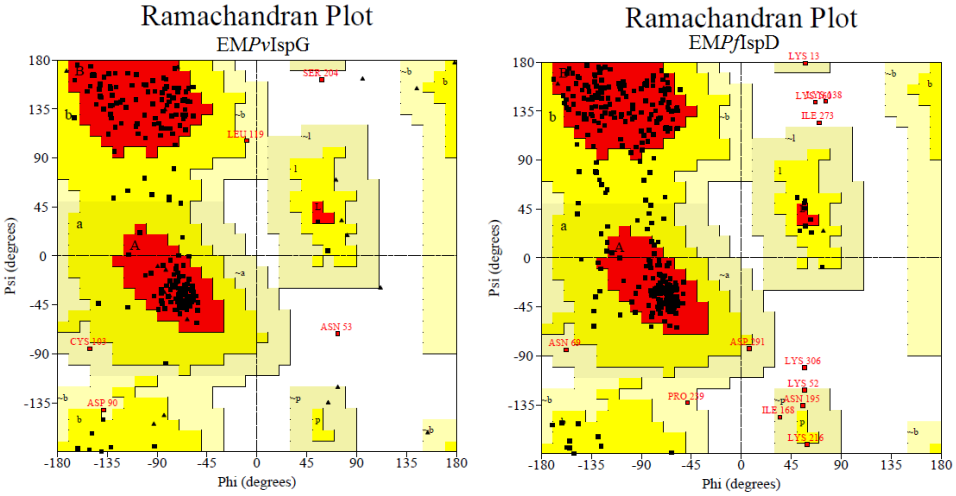


Figure 4.4: Validation of 3-dimensional structures of *Pv*IspD and *Pf*IspD using Ramachandran Plot

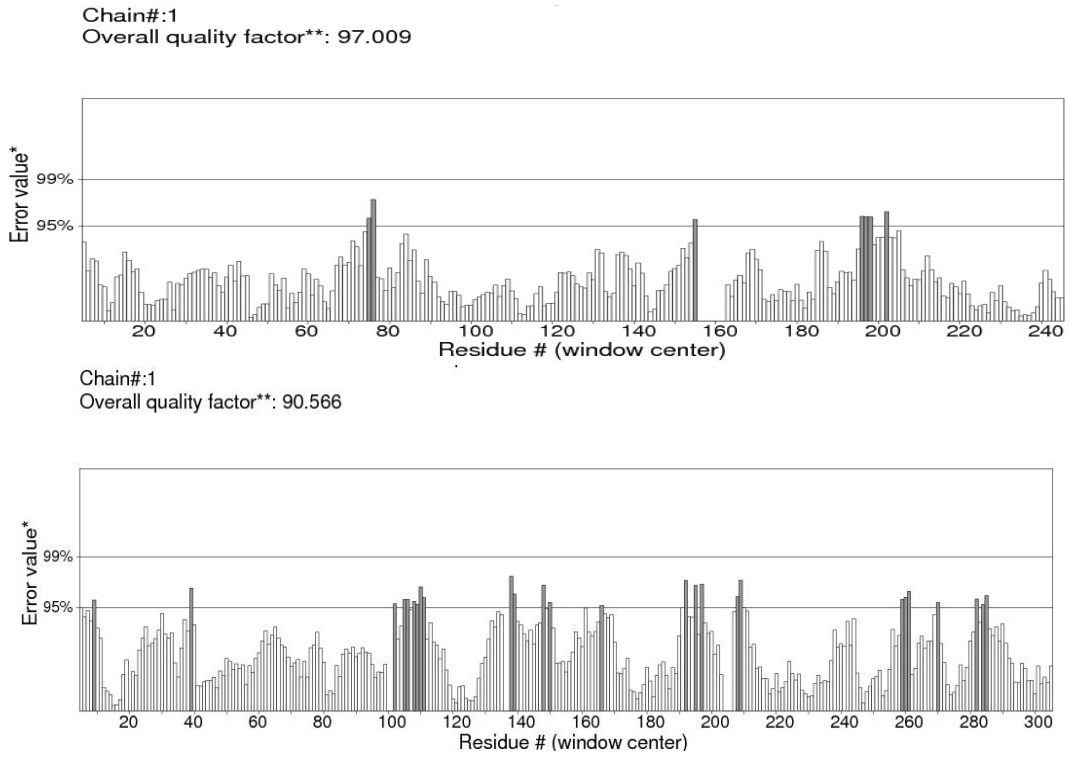


Figure 4.5: Validation of three dimensional structures of *Pv*IspD (upper) and *Pf*IspD (lower) proteins using ERRAT score.

Analysis of the modelled structures shows that these proteins follow a Rossmann fold arrangement specific to the protein that binds to nucleotides or enzyme cofactors such as FAD, NAD, and NADP. The structure is composed of seven parallel beta strands containing conserved signature motifs where first two strands are connected with α- helix (**Figure 4.6**).

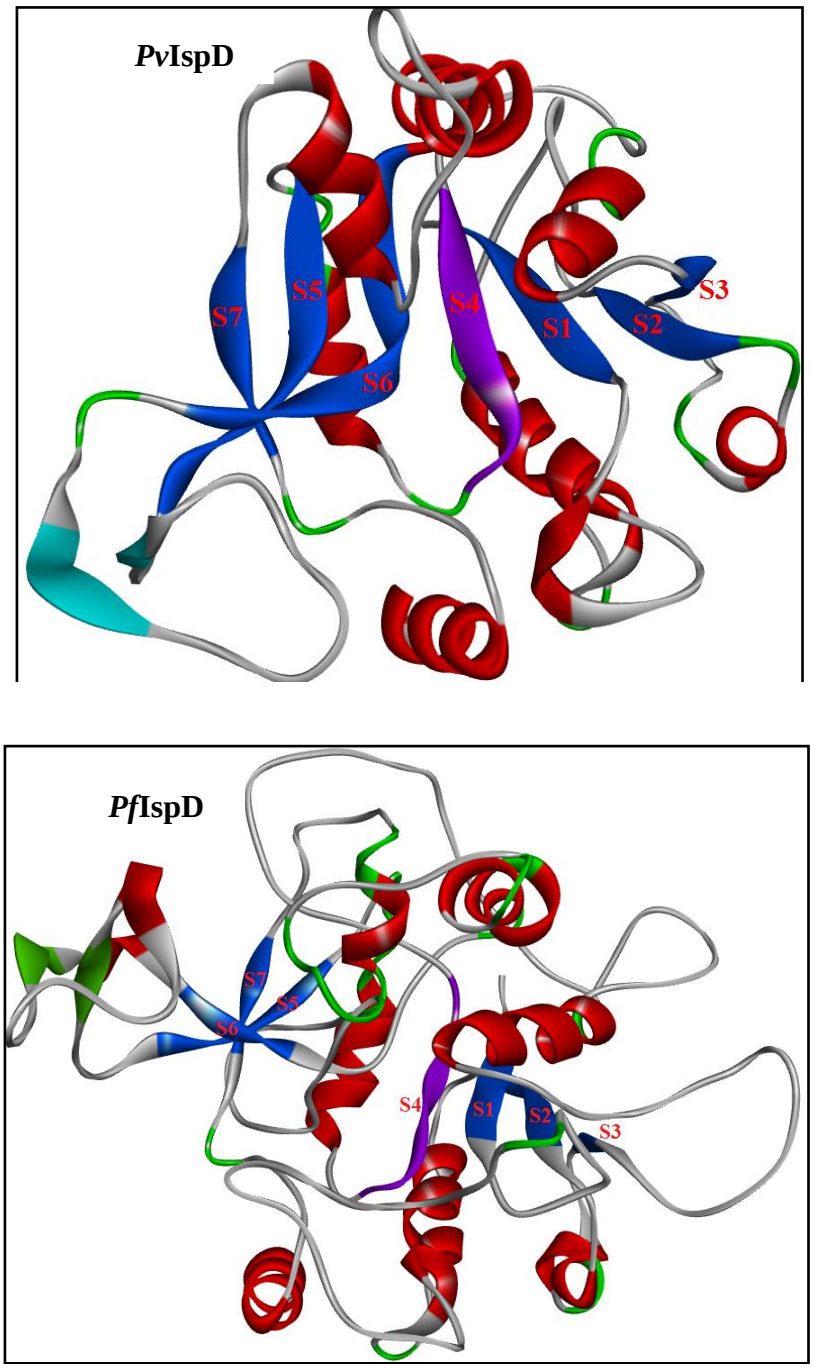


Figure 4.6: *Plasmodium* IspD protein structure predictions *Pv*IspD (upper) and *Pf*IspD (lower) structure prediction by homology modelling. *E. coli* IspD (1I52) crystal structure was used as a template. The Rossmann folds structure shown is made of 7 sheets S1 to S7; Signature motif residues are present in S4 sheet.

4.2.3 Active site architecture and docking with substrate

To predict the binding pocket with conserved amino acid residues, BioLip software was used. Once the binding cavity was identified (**Table 4.2)**, the ligands like CTP and MEP were sequestered in the active site. CTP showed interactions with the amino acids V323, T324, T326, E365, K406, R431, Q493, K548, I549 (**Figure 4.7a**) in the binding pocket while MEP interacted with the amino acid residues V426, D428, R431, T523, T525, K548 (**Figure 4.7b**) of *Pv*IspD. As the product is formed in the same binding cavity, we docked the product CDP-ME also to the obtained *Pv*IspD structure, where it showed interaction with specifically A325, T326, K327, G328, Q329, Q378, A401, K406, D428 (**Figure 4.7c**),

Table 4.2 *Pv*IspD active site prediction

Rank	C-score	Cluster size	PDB Hit	Ligand Name	Ligand Binding Site Residues
1	0.45	50	1iniA	CDP-ME	196-199, 203, 210, 403-406, 409, 428-431, 489, 491, 523-524, 548
2	0.08	10	4kqlA	CTP	196-199, 210-211, 402, 405, 407, 409, 428, 430, 456, 489-490, 546-548

*(**The C - score** is the confidence score of the prediction. C - score ranges [0-1], where a higher score indicates a more reliable prediction; **Cluster size** is the total number of templates in a cluster**)**.

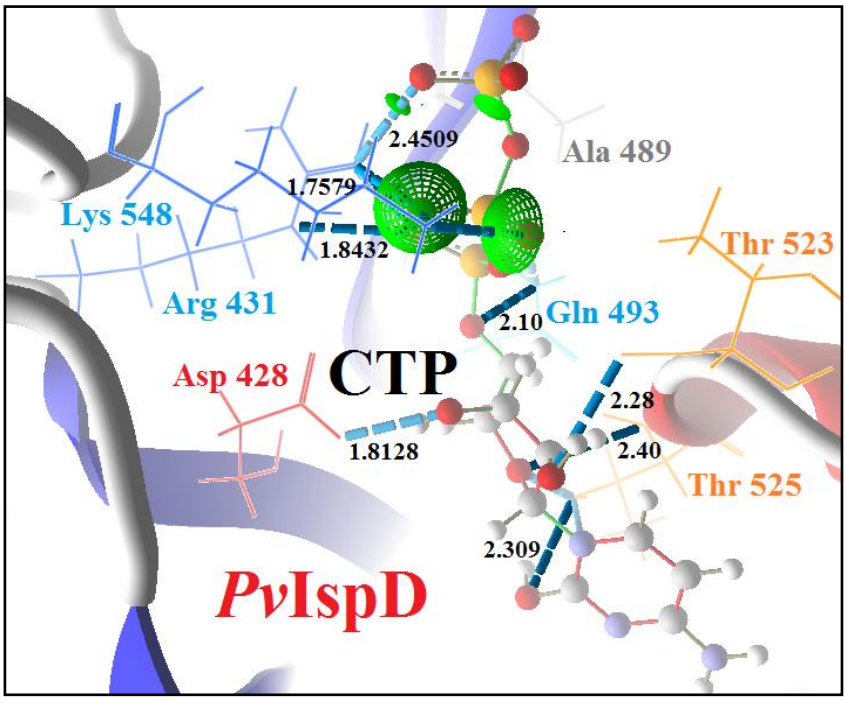


Figure 4.7a: Interaction of *Pv*IspD protein with CTP (shown with blue dotted lines**)**

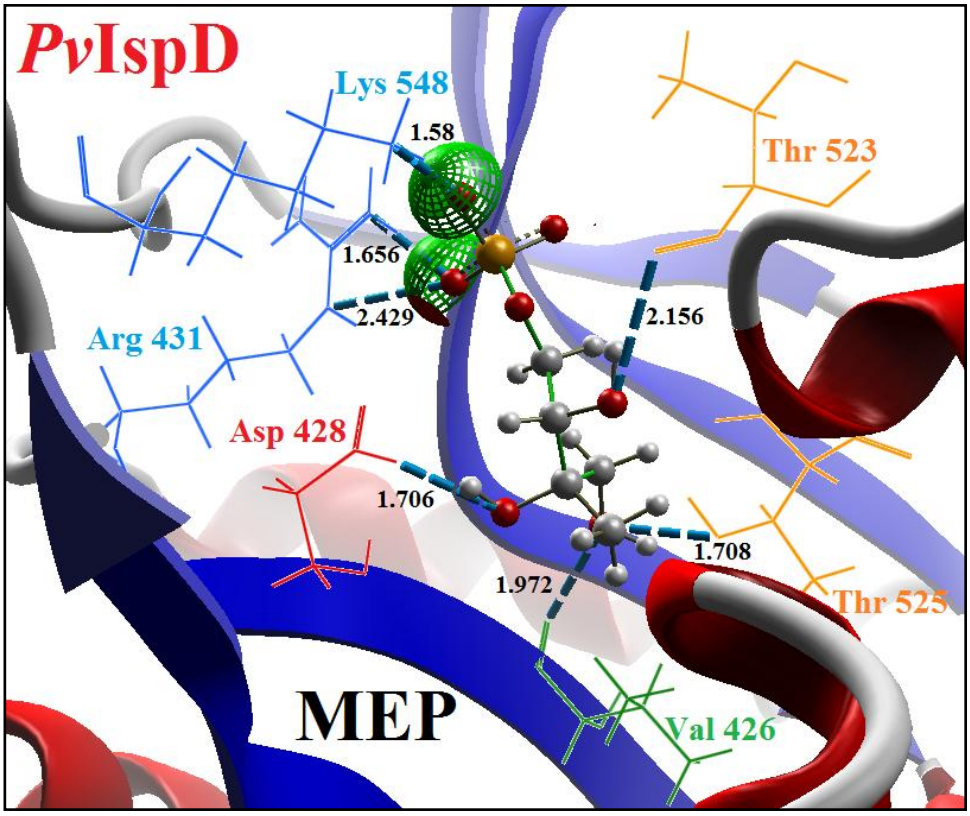


Figure 4.7b: Interaction of *Pv*IspD protein with MEP (shown with blue dotted lines)

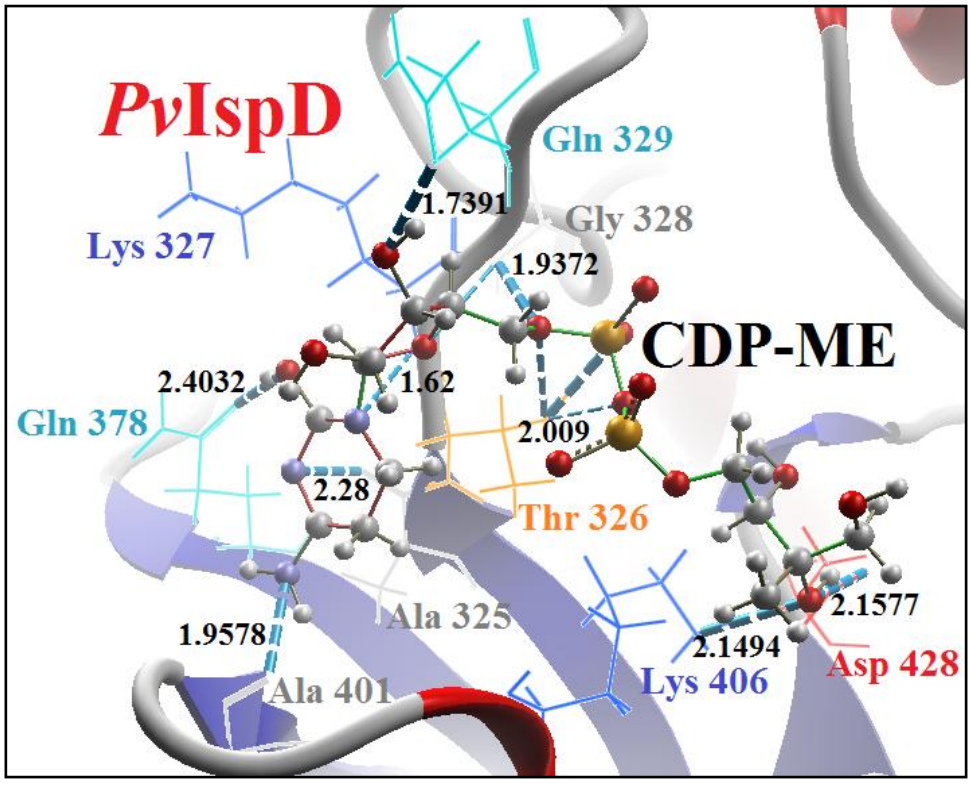


Figure 4.7c: Interaction of *Pv*IspD protein with CDP-ME (shown with blue dotted lines)

Compared to *Pv*IspD, in *Pf*IspD protein **(Table 4.3)** it was observed that CDP-ME molecule shows interaction with amino acids E371, E380, K464, A580, N618, D619, N620, N621 and K662 **(Figure 4.8a**). In comparison to this, the CTP and MEP molecules interacted with the amino acid residues K464, F578, A580, N618, N620, S637, S638, K662 (**Figure 4.8b**) and K610, T612, N618, N620, N621, D635, S638 (**Figure 4.8c**) respectively present in the binding cavity. Interaction studies of *Pv*IspD and *Pf*IspD proteins confirms that amino acid residues present in conserved domain are playing major role in their interaction with various substrates.

Table 4.3 *Pf*IspD active site prediction

Rank	C-score	Cluster size	PDB Hit	Ligand Name	Ligand Binding Site Residues
1	0.07	4	1sumB	FE	494,495,498
2	0.07	4	3is8T	MEP	458,461,462,465,484,491
3	0.04	2	5rebH	MEP	452,455
4	0.04	2	4p6vB	RBF	553,556
5	0.04	2	2xwnA	CTP	422-429, 436, 484-487, 490, 516-518, 662

(The C - score is the confidence score of the prediction. C - score ranges [0-1], where a higher score indicates a more reliable prediction; **Cluster size** is the total number of templates in a cluster**)**

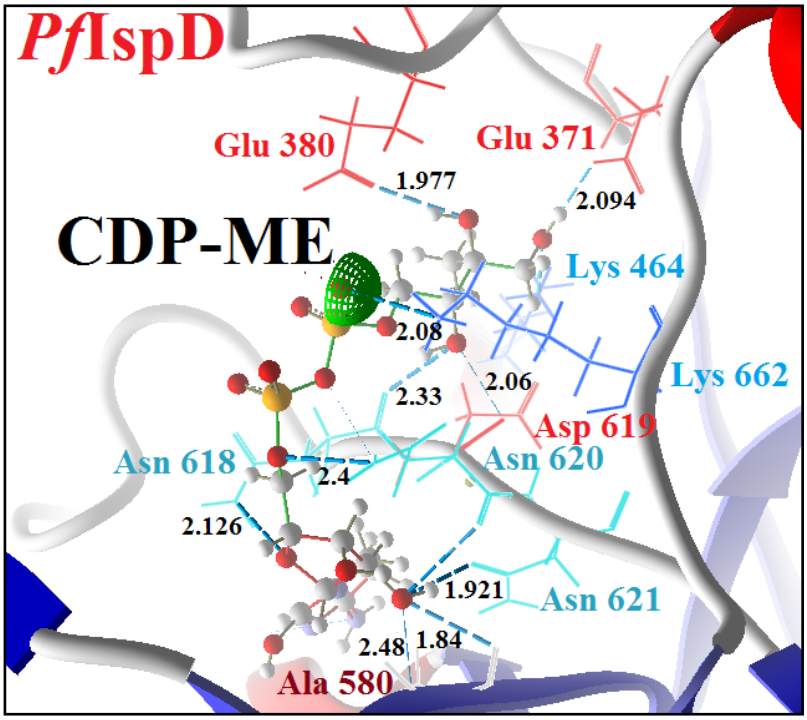


Figure 4.8a: Interaction of *Pf*IspD protein with CDP-ME (Interaction between *Pf*IspD enzyme's conserved residues and molecules shown with blue dotted lines)

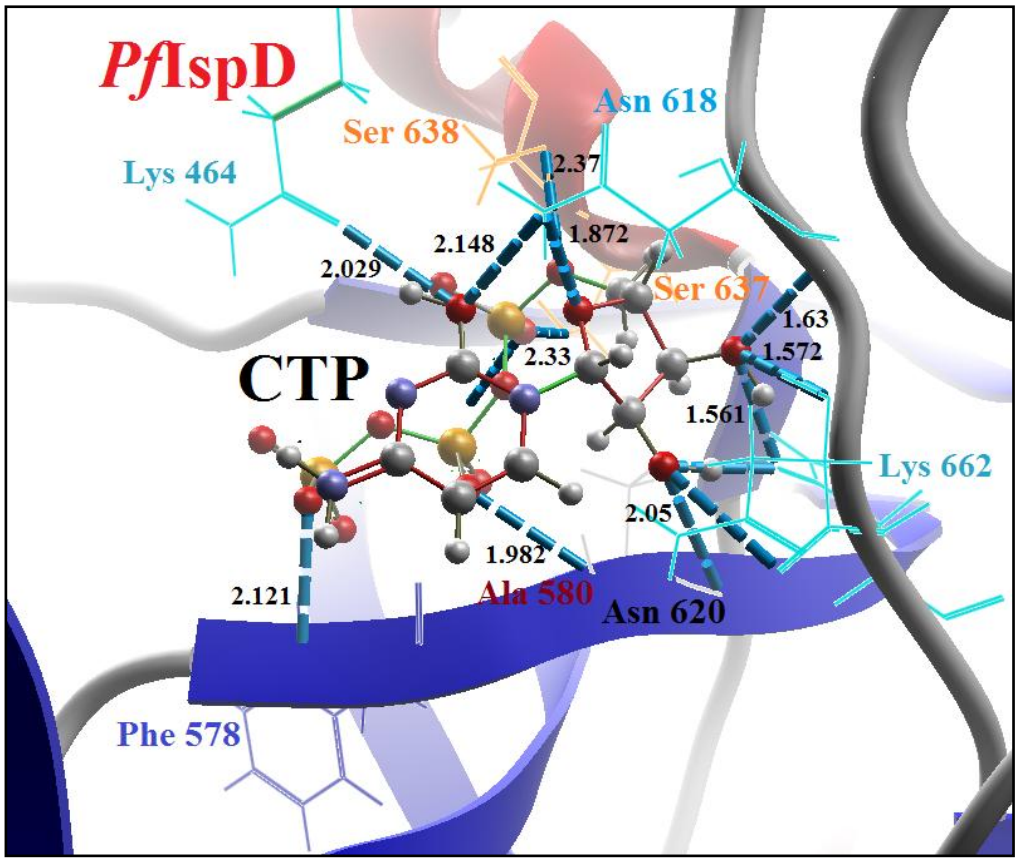


Figure 4.8b: Interaction of *Pf*IspD protein with CTP (Interaction between *Pf*IspD enzyme's conserved residues and molecules shown with blue dotted lines)

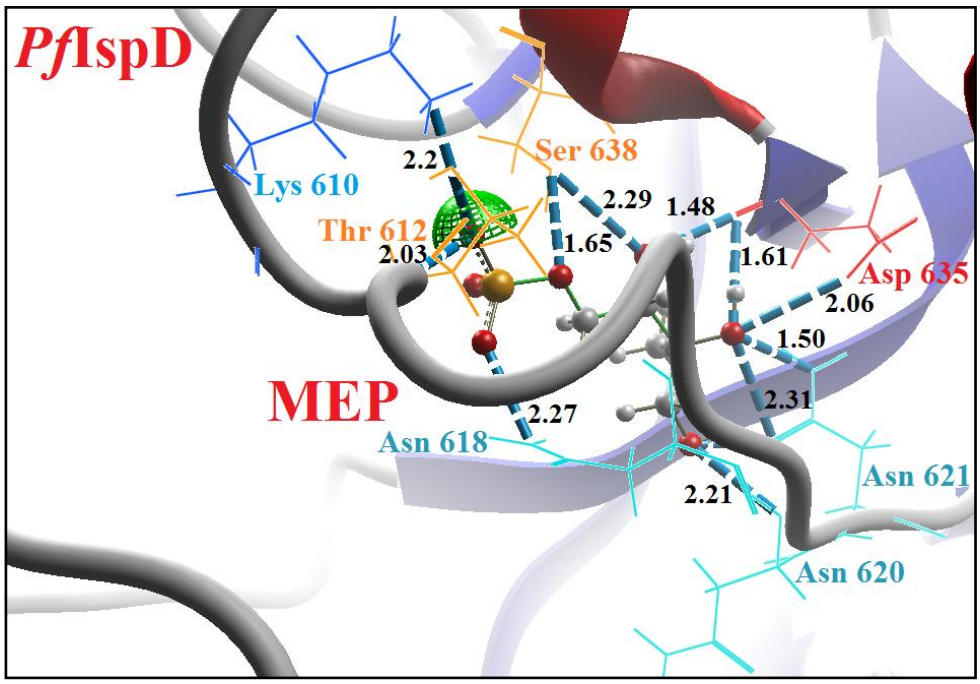


Figure 4.8c: Interaction of *Pf*IspD protein with MEP (Interaction between *Pf*IspD enzyme's conserved residues and molecules shown with blue dotted lines)

Presence of Rossmann fold signature motif in IspD proteins also confirms its ability to bind with different nucleotide derivatives. It was observed that binding pocket of IspD enzyme makes it very selective towards CTP molecules and allows only the sequestration of pyrimidine bases and not the binding of larger purine bases. Reports from various homologous molecules have shown very little activity with other dNTPs. Binding of CTP to IspD enzymes induces conformational changes in the active site, allowing MEP to attain the proper orientation for attack. Interaction between CTP molecule and amino acids of *Pv*IspD and *Pf*IspD enzymes when examined showed similar results as obtained for that of *E. coli*.

4.2.4 Catalytic mechanism and *in-silico* drug binding studies

Presence of conserved residues for the binding of substrates, indicates presence of a unique catalytic mechanism for IspD enzyme which comes under "associative mechanism". In this process a negatively charged penta-coordinate transition state forms upon a nucleophilic attack on the α-phosphate of CTP by the 4-phosphate of MEP and the collapse of this charged state leads to pyrophosphate release and CDP-ME formation (**Figure 3.1**). Binding of negatively charged CTP molecules with IspD protein is highly dependent on the positively charged catalytic residues. In *Pv*IspD protein, the positively charged residues K406, K548 and R431 were present that stabilizes the binding of CTP by the formation of salt bridge with α, β and γ-phosphate groups, respectively (**Figure 4.9**). This interaction provides a proper positioning for α -phosphate group for the nucleophilic attack.

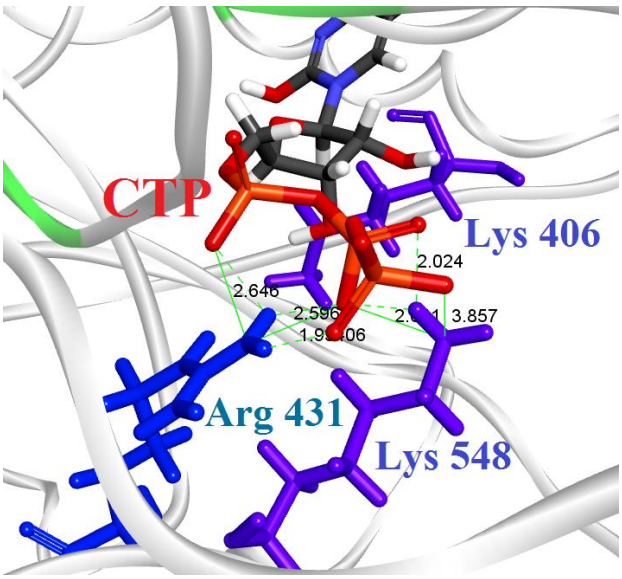


Figure 4.9: *Pv*IspD proteins binding with the CTP and the formation of salt bridge (catalytic residues K406, K548 and R431 which stabilize the binding of CTP by salt bridge formation).

Similar to this K464, K662, N618 and N620 residues present in *Pf*IspD protein performs an identical role in the substrate interaction. Amino (NH^{2+}) group present in these amino acid residues is of prime importance and mutation in these residues cause destabilization of binding. Further, to address the issue of IspD inhibition, various drug molecules were screened for their binding affinity towards *Pv*IspD and *Pf*IspD proteins.

4.2.4.1 IspD inhibition by Fosmidomycin (Fos)

Study performed by Zhang et al. (2011) detailed the IspD enzyme as a second target of Fos. To elucidate this data further and to identify how Fos may bind to or inhibit *P. vivax* IspD enzyme, we performed *in-silico* inhibition assays on *Plasmodium* IspD proteins. During analysis hydrogen bonds were observed between Fos and amino acid residues D428, R431, Q493, T523, D524, T525 and K548 of *Pv*IspD protein where key residues R431 and K548 forms salt bridge with phosphate group present in Fos (**Figure 4.10**). In case of *Pf*IspD protein, Fos showed interaction with amino acid residues A580, N618, N620, D635, S637 and S638, where the salt bridge was observed with A580 and N618 (**Figure 4.11**). These residues constitute the active site of the protein. The interaction between Fos and *Pv*IspD protein was stabilized with Molecular Dynamics studies which suggest the binding affinity of Fos with *Pv*IspD protein.

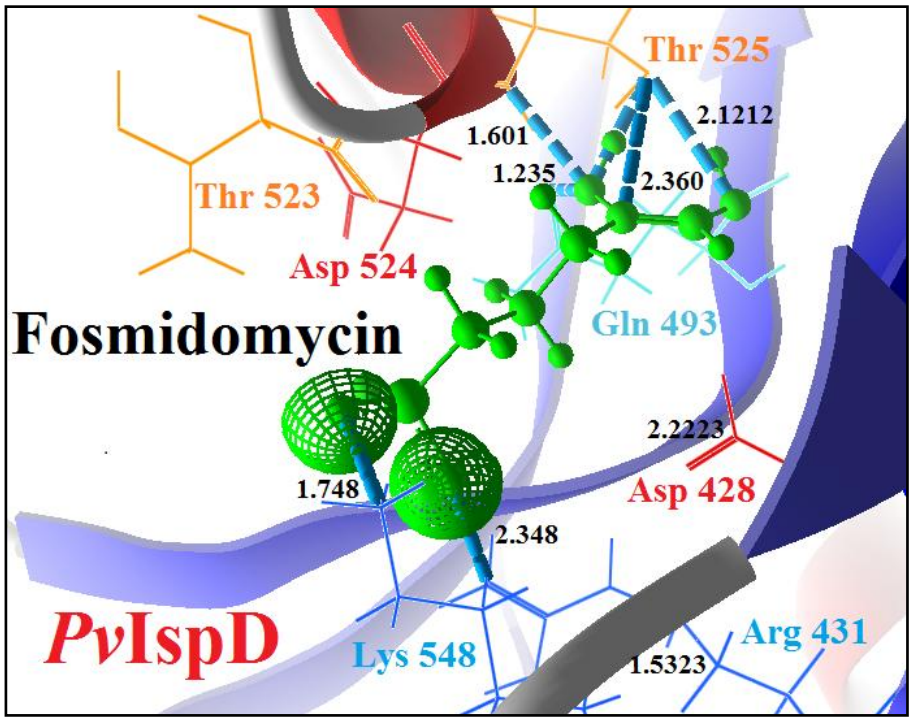


Figure 4.10: *Pv*IspD protein interaction with Fosmidomycin (shown in green colour). The conserved residues of *Pv*IspD enzyme are shown with blue dotted lines.

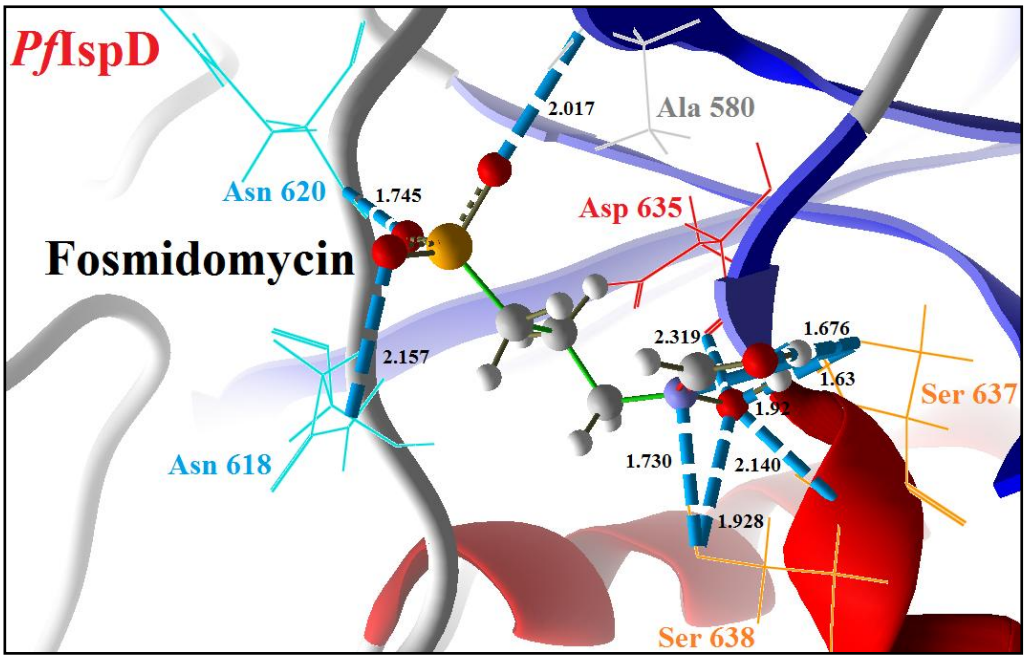


Figure 4.11: *Pf*IspD protein interaction with Fosmidomycin (shown in grey colour). The conserved residues of *Pf*IspD enzyme are shown with blue dotted lines.

4.2.4.2 IspD inhibition by MMV008138

MMV008138 has been shown to act as inhibitor of *Pf*IspD and synthetic/ codon optimized *Pv*IspD. To elucidate its binding, we performed *in-silico* binding assay of this drug with IspD enzyme. The structure of MMV008138 was obtained from Malaria box, its rigid chemical bonds were made flexible and converted to PDB format. MMV008138 and IspD protein structures (after removing HETATM and water molecules manually) were uploaded as an input in the Molegro Virtual Docker (Thomsen and Christensen, 2006). Interaction between MMV008138 and IspD proteins residues were confirmed by Energy Minimization and Molecular Dynamics studies. Stable confirmation was achieved for this interaction with the stabilization of total energy (Kinetic energy and Potential energy), RMSD, temperature and pressure. Constant values were attained for all these parameter and coordinates in protein ligand complex, before and after simulation were compared.

Different type of interactions were observed between *Pv*IspD & MMV008138 viz; hydrogen bond formation with D428, T525, K548; steric interaction with K406, D428, A489, D524, T523, T525, K548 and electrostatic interaction with D428 amino acid residues. Among these interacting residues, K548 and K406 are the common residues between CTP substrate and MMV008138, which are important for stabilizing the binding of CTP to the enzyme. These residues were found to be located at the binding cavity of

CTP substrate and few of the interacting residues were also shared by MEP substrate (**Figure 4.12**).

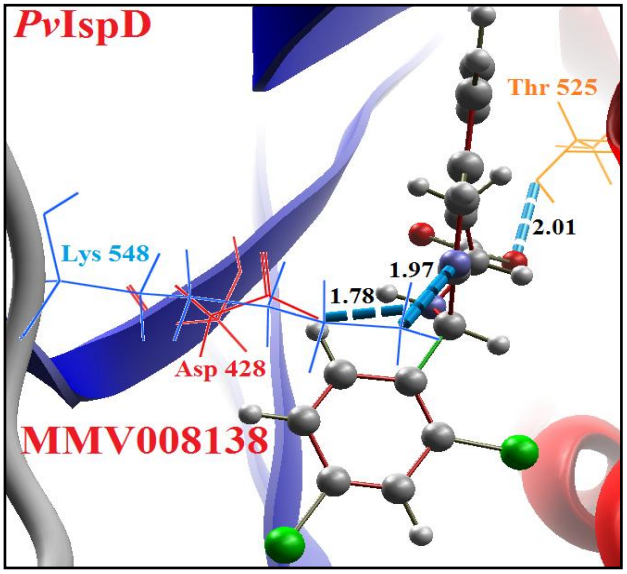


Figure 4.12: *Pv*IspD protein interaction with MMV008138 (shown in grey colour)**.** The conserved residues of *Pv*IspD enzyme are shown with blue dotted lines.

In comparison to this, *Pf*IspD enzyme shows interaction in the form of hydrogen bonds with I614, N615 and K662 and steric interaction with N376, K378, K381, I614, N615 and N618 amino acid residues (**Figure 4.13**). Similar to *Pv*IspD, K662 and N618 residues involved in stabilization of the CTP binding to *Pf*IspD protein are involved in binding to MMV008138. Other than these there were common residues between MEP binding and drug binding too.

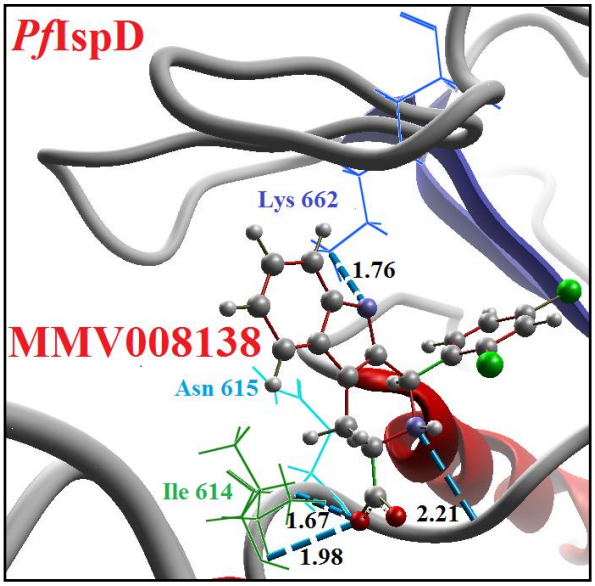


Figure 4.13: *Pf*IspD protein interaction with MMV008138 (shown in grey colour). The conserved residues of *Pf*IspD enzyme are shown with blue dotted lines.

These docking experiments suggest that MMV008138 interferes with the binding of IspD enzyme substrates indicating to its inhibition. Since the profile of both *Pf*IspD and *Pv*IspD against MMV008138 *in silico* showed almost similar results, in contrast to the experimental *in-vitro* biochemical assays where different affinity of the drug with the two enzymes were seen, this demands further investigation into the compound before using it as an futuristic antimalarial.

4.2.4.3 IspD inhibition by Azolopyrimidines

Inhibition of IspD enzyme by herbicides like Azolopyrimidine has been reported for *Arabidopsis thaliana.* Its mode of action against *Plasmodium* IspD enzymes was also studied to check whether binding of this drug disturbs the active site of the enzyme. We propose that Azolopyrimidines may be used as an inhibitor of *Plasmodium* IspD enzyme. We performed the docking of Azolopyrimidine with *Plasmodium* IspD enzyme. The amino acid residues T326, K406, D428, R431, A 489, T525 and K548 of *Pv*IspD protein were observed to interact with Azolopyrimidines in the form of strong hydrogen bonds whereas salt bridge was observed with K406 and D428 residues (**Figure 4.14**). Most of these amino acid residues are common with the two substrates MEP and CTP, where the amino acid residues D428, R431, T525 and K548 are common with MEP and amino acid residues T326, K406, R431 and K548 are common with CTP binding. Also the two residues involved in the formation of salt bridge for stabilization of CTP binding are common with Azolopyrimidines salt bridge residues. As most of these residues are located in the active site of the IspD protein where MEP and CTP molecules comes and bind to initiate the enzymatic reaction, there is a very high possibility that the binding of this molecule may affect catalytic activity of the protein.

Binding affinity of Azolopyrimidines was also checked with *Pf*IspD protein where it showed interaction with E368, K381, N463, N465, I616 and N618 amino acid residues (**Figure 4.15**). As discussed above, except N618, rest of the residues were different from the substrate binding residues. However, the exact binding nature of Azolopyrimidines with *Plasmodium* IspD enzymes can be concluded only after *in-vitro* biochemical assays or parasite culture assays.

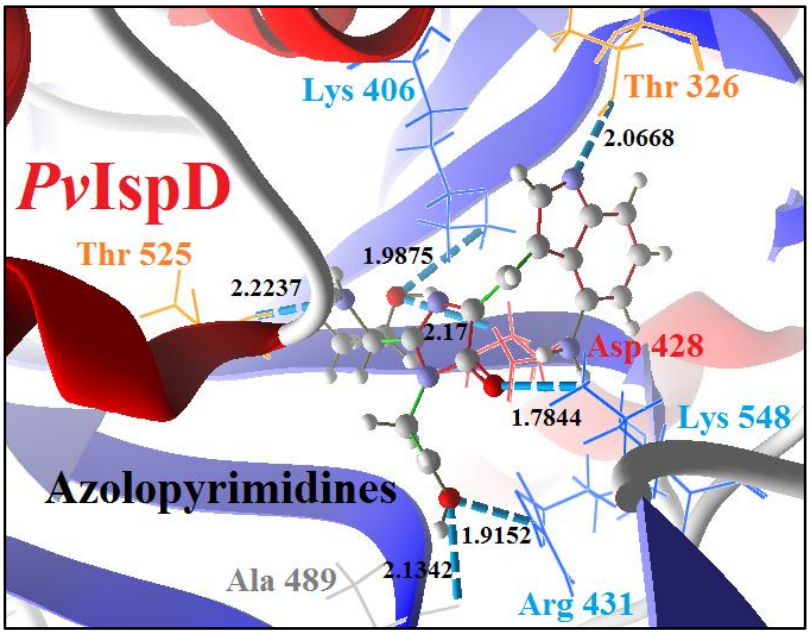


Figure 4.14: *Pv*IspD protein interaction with Azolopyrimidines (shown in grey colour). The conserved residues of *Pv*IspD enzyme are shown with blue dotted lines.

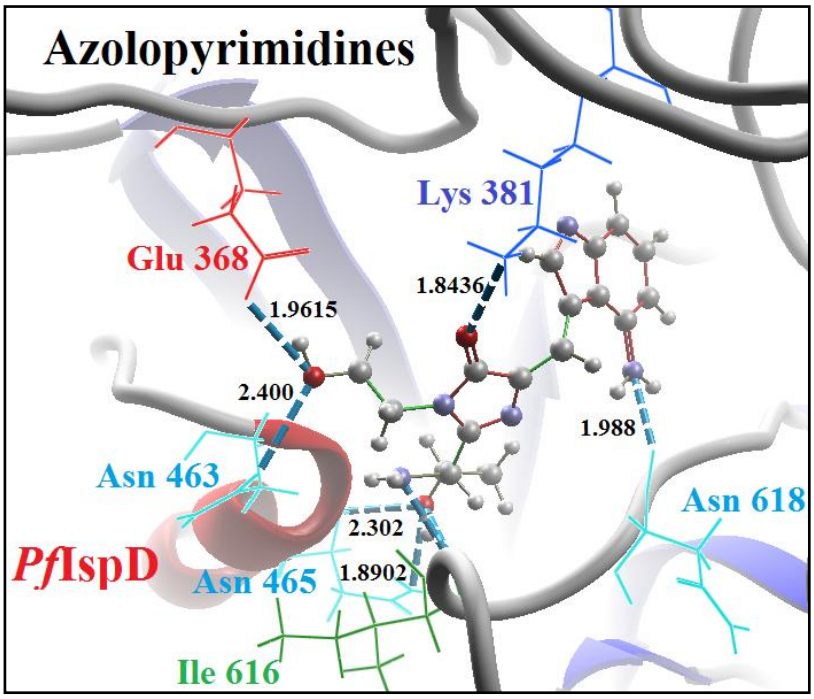


Figure 4.15: *Pf*IspD protein interaction with Azolopyrimidines (shown in grey colour). The conserved residues of *Pf*IspD enzyme are shown with blue dotted lines.

The binding energies for Fos, MMV008138 and Azolopyrimidines were calculated and compared with the already reported data where lower binding energy for Azolopyrimidines show its efficient binding. Significant values were obtained for the other inhibitor Fos and MMV008138 as listed in **Table 4.4**.

Table 4.4 Binding energies for IspD enzyme interaction with substrates and different inhibitors.

Proteins	Ligand molecule	Binding energy ΔG (Kj/mol)
E. coli IspD (1I52)	CTP	-289.658
M. tuberculosis IspD (3Q80)	CDP-ME	-289.435
A. thaliana IspD (2YC3)	Azolopyrimidines	-122.082
P. vivax IspD	CTP	-145.158
	MEP	-79.572
	CDP-ME	-151.675
	Azolopyrimidines	-144.603
	MMV008138	-116.078
	Fosmidomycin	-108.979
P. falciparum IspD	CTP	-157.537
	MEP	-77.255
	CDP-ME	-150.985
	Fosmidomycin	-87.784
	MMV008138	-128.078
	Azolopyrimidines	-151.686

4.2.4.4 IspD inhibition with regular anti-microbials

During drug inhibition assays, we observed that active site of the IspD protein is specific for the binding of drug molecules containing phosphate group, which have the affinity towards the positively charged amino acid residues present in the catalytic site of the protein. Following this observation, we analyzed the chemical structure of various anti-microbial drugs being used at clinical level. Presence of active group was the prime concern in this analysis and selected drugs were further docked against IspD protein. In the drug binding studies, few drug molecules like Cefepime and Rifampicin were found to be very effective and showed high binding affinity towards IspD protein that might be responsible for competitive inhibition (**Table 4.5**). These drugs were tested for their IC_{50} values to elucidate their inhibitory efficiency, as described in Chapter 3.

Table 4.5 Inhibition of *Pv*IspD protein with commercial available molecules

Binding Pocket	**V323-V329**		**K406**	**R431**	**Q493**	**K548**	**I549**		
CTP	V323	T 324	T326	E 365	K 406	R 431	Q 493	K 548	I549
MEP	V426	D428	R431	T523	T525	K548			
CDPME	A325	T326	K327	G328	Q329	Q378	A401	K406	D428
Antimycin A	T 326	K 406	K 548	F 522					
Cefapime	K 406	R 431	T 491	T 525	K548				
Ceftriaxone	S378	S394	D417	I421					
Ciprofloxacin	K525	K548							
Clarithromycin	R431	S 526	T 525	K 548	T550				
Doxycyclin	T326	K 406	G429	A430					
Isoniazid	R 431	S456	Q490	T491					
Monensin	T523	T525	K 548						
Rifampin	T326	R 431	K 548	T525					
Tetracycline	T326	G328	K 406	G429	A430	T550			
Tunicamycin	V323	T324	T326	D338	L381	L384	K406	T491	
Atormentin	R431	S456	Q490	K548					
Azolopyrimidines	T326	K406	D428	R431	A489	T525	K548		
Fosmidomycin	D428	R431	Q493	T523	D524	T525	K548		

4.3 Conclusion

The need for novel chemotherapeutic agents as antimalarials has grown due to the extensive spread of resistance to the available drugs. Many pathogenic bacteria as well as apicomplexans including *Plasmodium*, are known to solely utilize the MEP pathway for isoprenoids biosynthesis, which humans do not use, thus making it a highly attractive target. IspD, the third enzyme in the MEP pathway, is proposed as one of the potential antimalarial target. The *Plasmodium* protein, though longer in sequence possessed highly conserved domains and residues as its homologues/ orthologues from prokaryotes including some major human pathogens and apicomplexans. It has been detailed as the second major target to Fosmidomycin but there is a major knowledge gap that needs to be fulfilled to use Fos as major anti-malarial.

Till date, only few drugs have been tested against IspD in different organisms. Thus, there is a need to identify novel inhibitors that may act upon IspD enzymes leading to highly effective therapy in combination with Fos as an antimalarial. Our analyses confirm the binding affinity of drugs like MMV008138, Cefepime, Rifampicin and Azolopyrimidines with IspD protein possibly leading to inhibition of protein activity in *Plasmodium*. A major finding of our study includes the analyses of anti-microbial agents, Cefepime and Rifampicin. These two drugs are administered clinically for the treatment of various anti-bacterial infections. Our analyses show that these drugs bind efficiently to the active site of IspD enzyme, thus acting as potential anti-malarials. This *in silico* data of Cefepime and Rifampicin was used for *in-vitro* analysis on purified *Pv*IspD enzyme as detailed in chapter 3.

The herbicidal compound Azolopyrimidines, showed different binding affinities for *P. vivax* and *P. falciparum in silico*. While most of the residues for binding of Azolopyrimidines were similar to the two substrates CTP and MEP in *P. vivax*, only few of the residues were common in case of *P. falciparum*. This suggests that while Azolopyrimidines, can act as competitive inhibitors of *Pv*IspD enzyme, similar efficacy may not be observed for other human malaria parasite *P. falciparum* and thus, Azolopyrimidines can be used only as specific inhibitors to target *P. vivax*.

In the present study, we have tried to gain deeper insight into the structural and functional features of IspD enzyme of *Plasmodium*, to provide possible hints in the discovery of the anti-malarial compounds using IspD as a target. However, major challenge in developing new effective MEP pathway inhibitors is the optimization of the drug specificity, its binding affinity and delivery of the drug molecule. The overall analysis performed in the given study is based on only *in-silico* studies to detail the various molecules that can be used as inhibitors for IspD enzyme. To validate these compounds, further experimentation has to be performed, before categorizing them as potential anti-malarials.

Functional and Structural Characterization of IspG Enzyme from *P. vivax* and *P. falciparum*

Chapter V

Characterization of Isoprenoids biosynthesis pathway enzyme IspG from *P. vivax* and *P. falciparum*

The current chapter focuses on the characterization of IspG enzyme involved in the second last step of isoprenoids biosynthesis pathway from *Plasmodium*. In this section, the expression, purification and characterization of IspG enzyme from Indian field isolates of *P. vivax* and *P. falciparum* has been detailed. Various biochemical assays were performed to elucidate the functions of IspG enzyme along with sub-cellular immune-localization to prove the site of action of this protein. This is the first study on this enzyme from any *Plasmodium* species. Our results confirm that IspG is an important component of the prokaryotic type MEP pathway functional in the apicoplast of *Plasmodium*. We propose that since enzyme is able to bind Fe-S cluster, thus, it may be functional in the apicoplast. However, final conclusion can be drawn with direct activity assay in the presence of substrate.

5.1 Introduction

IspG enzyme [4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (also named GcpE)] participates in the penultimate step of the MEP pathway where it catalyses the reduction of 2-C-methyl-D-erythritol-2,4-cyclodiphosphate (MECP) to 4-hydroxy-3-methyl-2-(E)-butenyl-4-di phosphate (HMBDP) (**Figure 5.1**). This conversion is a step wise reaction (**Figure 5.1**) where IspG protein provides a site for the binding of MECP. This binding of the MECP substrate to IspG enzyme is aided by a redox reaction where the electrons required for the binding are hypothesized to be provided by the flavodoxin/flavodoxin reductase/NADPH system through [4Fe-4S] clusters in *E. coli* and through ferredoxin/ ferredoxin $NADP^+$ reductase system in a cyanobacterium *Thermosynechococcus elongatus* and *Plasmodium* which lacks the flavodoxin system (similar to IspH as reported by Rohrich et al., 2005). Transfer of electron to the MECP substrate in turn oxidizes the [4Fe-4S] cluster which can be later recharged (reduced) again by a ferredoxin molecule (Okada and Hase, 2005; Seemann et al., 2006). As a result of the electron transferred during this process there is formation of a double bond which converts the cyclic form of MECP molecule to aliphatic HMBDP.

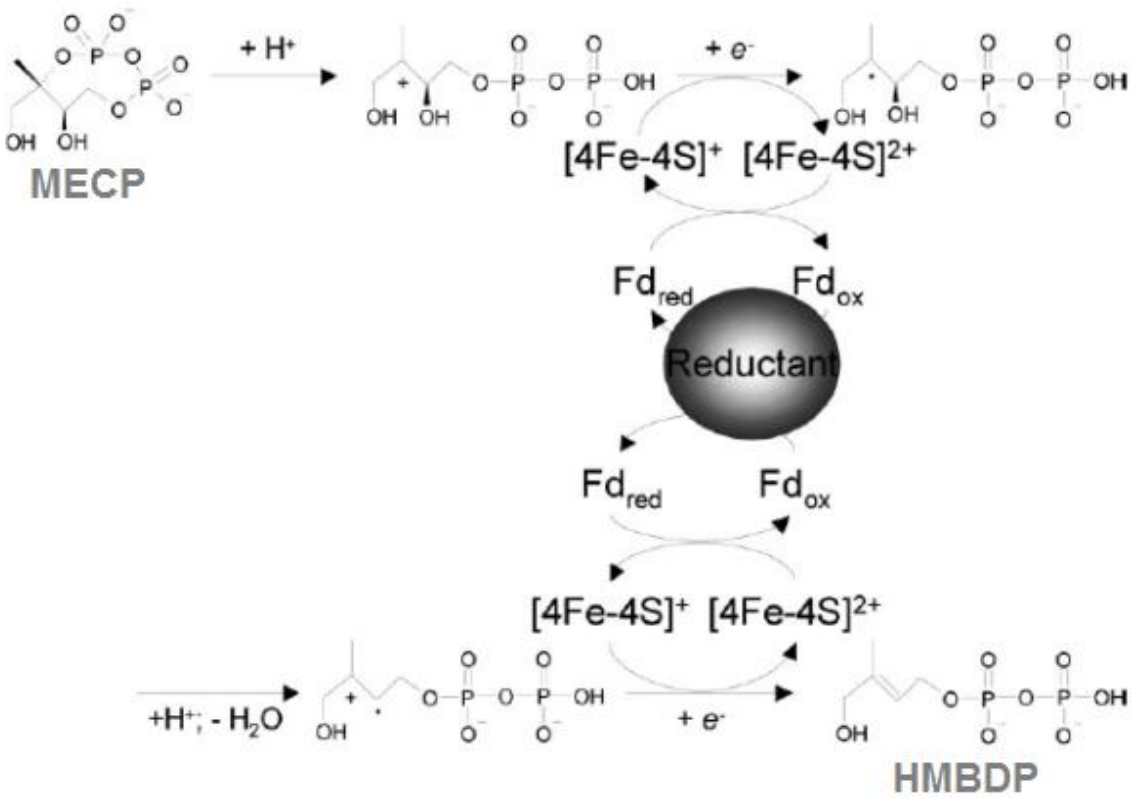


Figure 5.1: Schematic diagram representation of the proposed mechanism of action for the activity of IspG enzyme (adapted from Okada and Hase, 2005).

Presence of Fe-S cluster, electron transfer system and reducing environment comes under the major requirements for the activity of the IspG enzyme. IspG enzyme remains active in the absence of added metal ions, suggesting that metals are not necessary for its activity. However, some metal ions do have the ability to enhance or inhibit its activity. For example, Mn^{2+} increases the activity of IspG enzyme two folds while Zn^{2+} have a negative impact on IspG activity. Addition of a chelating agent like EDTA, inhibits the activity of IspG enzyme, however, the activity can be restored by using a higher concentration of flavodoxin and flavodoxin reductase (Zepeck et al., 2005).

Functionally Important Residues for IspG activity

In case of GcpE, studies have proposed that the C-terminal region of the protein might be key for the activity. In *E. coli*, the presence of three conserved cysteine residues at positions 269, 272, and 305 suggests the binding of a transition metal such as Fe (Hecht, 2001). More specifically, they form a $CPXCXRX_{27\text{-}33}GC$ motif which is involved in the binding of the Fe-S clusters (Zepeck et al., 2005). To check the involvement of other residues in the enzyme activity, mutation studies were performed in *Salmonella enterica* where four missense mutations were identified as R133H, G206R, R237C and G309R. Among these, R133, R237, and G206 were found to be fully conserved within bacteria, plants, and the malaria parasites. Mutation in these residues was found to be lethal for the microorganism's survival and it was speculated that the polar charged nature of Arginine residues might help in the binding of substrate and catalysis process whereas non-polar aliphatic Glycine can have a role in the structural stability (Cornish et al., 2006).

Structure and Importance of Iron Sulphur Clusters

As already stated that activity of IspG protein is highly dependent on Fe-S cluster and these clusters can exist naturally on different proteins or they can be synthesised by chemical methods. These clusters have the tendency to interact with Cysteine residues but in some cases they can also bind to Histidine residues. In *E. coli*, *M. tuberculosis* and *P. falciparum*, the presence of three fully conserved Cysteine residues has been reported which may help in the binding of iron atoms. Mutation studies in *E. coli*, where these cysteine residues were replaced with serine residues, resulted in the loss of Fe and S content by 6% and 23% respectively in comparison to the wild type enzyme, leading to apparent loss of enzyme activity in the mutant strains. These findings have confirmed the

role of Fe-S cofactors for the activity of IspG enzyme (Zepeck et al., 2005). Fe-S clusters on IspG proteins are naturally present and come under the category of Cubane or D type clusters (Guerra et al., 2014) and they are shown to be highly susceptible for molecular oxygen and other environmental conditions like pH, oxidation-reduction which can cause destruction of the clusters, resulting in the loss of catalytic function of the proteins carrying these cluster. In biological systems these Fe-S clusters mediate electron transport process because of the redox potential acquired by the atoms present in the cluster. This redox potential has a voltage range which depends upon the number of iron and sulphur atoms present in the cluster, where HiPIPs (high potential iron sulphur proteins) have potential ranging from 350 to 450 mV. IspG proteins belong to the HiPIPs family of Fe-S cluster containing protein and carried a high potential to accomplish electron transfer process (Wang and Oldfield, 2014).

The Fe-S cluster proteins also possess different magnetic and electronic properties which depend on the oxidation states, spin states, orbital contribution and coordination numbers of electrons. While the diamagnetic species with closed shells of electron does not have any magnetic properties, the paramagnetic ones, with unpaired electron respond to a magnetic field where the iron ion can be in +2 or +3 oxidation states with high spin configuration when coordinated with Cysteine residues. Paramagnetic types of [4Fe-4S] clusters are present in IspG protein with +2 oxidation state where they have binding affinity with three Cysteine residues (Seemann et al., 2005).

Reduced conditions are required for the activity of IspG enzyme

IspG enzyme requires a specific requirements for activity. This was proven by a study in *Saccharomyces cerevisiae* where attempts were made to reconstruct complete MEP pathway by integrating the genes for all the eight enzymatic steps of the MEP pathway into the chromosome of *S. cerevisiae*. Despite the fact that machinery for Fe-S cluster assembly is present in the yeast (ISC system in mitochondria and CIA in cytosol) (Lill and Mühlenhoff, 2005; Sharma et al., 2010), IspG enzyme was found to be non-functional and thus served as the potential bottleneck for the pathway. To overcome this problem, all the components responsible for the formation of Fe-S clusters and electron transfer system were incorporated into the yeast cytosol. While the Fe-S cluster requirement was fulfilled by the addition of *erpA* and *iscA* gene of *E. coli*, the electron

transfer system was provided by the flavodoxin and flavodoxin reductase of *A. thaliana*. However, addition of all these components also did not result in a functional MEP pathway. Further investigations suggested that other than these clusters, the IspG enzyme requires a suitable reducing environment for its activity and that may be compartmentalization required for the activity of IspG enzyme (Partow et al., 2012).

In *Plasmodium*, as the pathway is functional in the apicoplast, there is a possibility that the apicoplast is providing a suitable reducing environment for the activity of IspG enzyme. The Fe-S clusters required for the activity of IspG enzyme are believed to be provided by the SUF system (an Fe-S cluster biogenesis pathway) reported to be functional in the apicoplast of *Plasmodium* (Xiao et al., 2009, Pala et al., 2016).

Importance of IspG enzyme and impact of its inhibition

Several reports have detailed the importance of IspG enzyme in different organisms. In *E. coli*, deletion of this enzyme resulted in total loss of isoprenoids biosynthesis and thus impairment of various house-keeping functions like cell wall formation, electron transport etc. (Altincicek et al., 2001). Apart from this, inhibition studies of IspG enzyme have also revealed about its indirect involvement in many cellular processes. In plants and other microorganisms inactivation of IspG enzyme is responsible for the accumulation of MECP substrate, which results in the retrograde signalling, where it can alter the nuclear architecture and functional dynamics which lead to the expression of stress responsive genes (Xiao et al., 2012). Recent literature (Quitterer et al., 2015) have suggested the importance of IspG enzyme in various prokaryotic pathogenic organisms, however no study has been done for IspG protein from *Plasmodium* parasite. Thus, this enzyme was selected for further exploration.

5.2 Results and Discussion

5.2.1 Apicoplast targeting

The amino acid sequence of *Pv*IspG, putative (PVX_111575) and *Pf*IspG (Pf3D7_1022800) retrieved from the PlasmoDB was subjected to various online servers like PlasmoAP and PATS, which predicts apicoplast targeting, PlasMit and MitoProt II for mitochondrial targeting prediction and SignalP 4.1 for checking the presence of signal peptide. Based on different parameters considered by PlasmoAP like the presence of

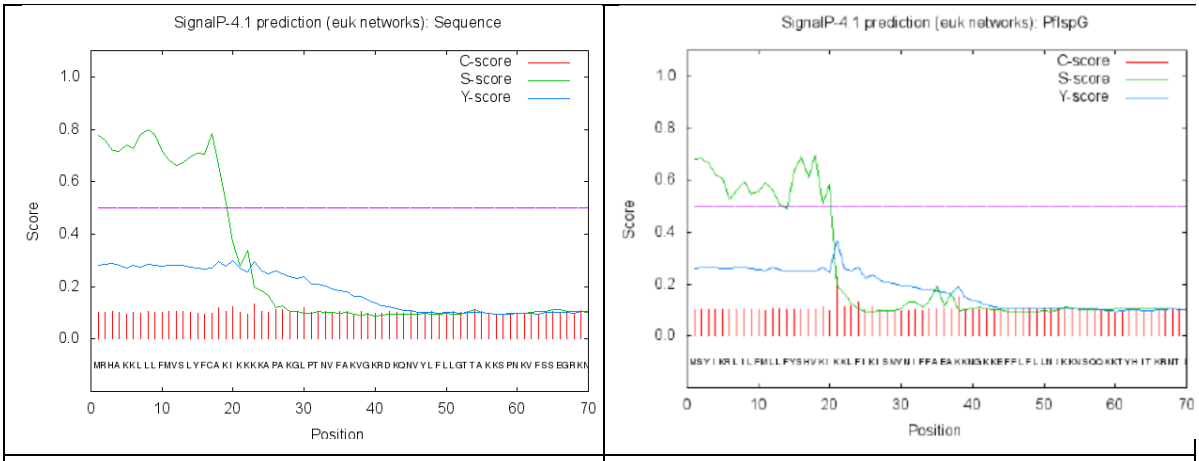
basic amino acids and KN enriched regions at N terminal, IspG in both *P. vivax* and *P. falciparum* showed a very high possibility of apicoplast targeting as inferred by all 5/5 tests positive (**Figure 5.2**). The results of PlasmoAP were further confirmed using PATS which gave a high score of 0.859 for apicoplast targeting in *P. vivax* and 0.950 in *P. falciparum*.

	Plasmodium vivax		***Plasmodium falciparum***	
Criterion	**Value**	**Decision**	**Value**	**Decision**
Signalpeptide	2 of 4 tests positive	**0**	3 of 4 tests positive	+
apicoplast-targeting peptide	5 of 5 tests positive	++	5 of 5 tests positive	++
Ruleset 1				
Ratio acidic/basic residues in first 22 amino acids <=0.7	0.000	yes	0.167	yes
Does a KN-enriched region exist (40 AA with min. 9 K or N) with a ratio acidic/basic <=0.9	0.000	yes	0.182	yes
Ruleset 2				
number of acidic residues in first 15 amino acids (<=2)	0	yes	0	yes
Does a KN-enriched region exist (40 AA with min. 9 K or N) ? Ratio acidic/basic residues in this region <0.6	0.000	yes	0.182	yes
Is the first charged amino acid basic ?		yes		yes

Note: The **final decision** is indicated by "++, +, 0 or -", where apicoplast-localisation for a given sequence is considered as "++ very likely; + likely; 0 undecided; - unlikely

Figure 5.2: Prediction of bi-partite N-terminal leader sequence with PlasmoAP for *Pv*IspG and *Pf*IspG

However, the results for the presence of signal peptide were inconclusive for *P. vivax* IspG in PlasmoAP. While the *Pf*IspG showed 3/4 tests positive for signal peptide indicating presence of signal peptide, *Pv*IspG showed only 2 out of 4 tests positive. Thus, to further check the presence of signal peptide, the sequences were submitted to SignalP 4.1, where using default parameters, presence of signal peptide was observed in both *Pf*IspG and *Pv*IspG, with a cleavage site between amino acid 19 and 20 in *P. vivax* and between amino acid 20 and 21 in *P. falciparum* (**Figure 5.3**).



*Pv*IspG | *Pf*IspG

Measure	*Plasmodium vivax*				*Plasmodium falciparum*			
Position	**Value**	**Cut off**	**Signal peptide**		**Value**	**Cut off**	**Signal Peptide**	
Max. C	23	0.133			21	0.226		
Max. Y	20	0.298			21	0.369		
Max. S	8	0.798			18	0.696		
Mean S	1-19	0.716			1-20	0.596		
D	1-19	0.524	0.450	Yes	1-20	0.491	0.450	Yes
	Name =*Pv*IspG; SP= Yes; Cleavage site between pos. 19 & 20; VKI-KK D= 0.524; D-cutoff =0.450 Networks = SignalP – no TM				Name =*Pf*IspG; SP= Yes; Cleavage site between pos. 20 & 21; VKI-KK D= 0.491; D-cutoff =0.450 Networks = SignalP – no TM			

Figure 5.3: Prediction of signal sequences and cleavage site with SignalP 4.1 for *Pv*IspG and *Pf*IspG.

Although the above software suggested the presence of apicoplast targeting peptide, for validation purpose targeting of *Pv*IspG and *Pf*IspG protein to the mitochondria was also checked using MitoProtII and PlasMit servers. To our surprise, we found a high possibility of both *Pf*IspG and *Pv*IspG to be targeted to the mitochondria by MitoProtII (0.89 for both *P. falciparum* and *P. vivax*) as well as PlasMit (91% poss. Mitochondrial *for both*), which suggested that further localization studies are required for confirmation of the functional site.

5.2.2 Amplification of *ispG* gene

5.2.2.1 Primer designing and amplification of *Pv*IspG

The *ispG* gene, putative (PVX_111575) reported from *P. vivax* Salvador I available at PlasmoDB, is 2463bp in length and expands from 1010100-1012562 position on chromosome number 6. Using this sequence, three pairs of primers were designed for the amplification of complete *ispG* gene (**Table 5.1**) as depicted in **Figure 5.4**. Amplification

of full length gene was attempted from different field isolates using primers VS56 and VS59 (**Figure 5.5a**), while the partial amplifications of the gene was performed using nested PCR where fragments of 956bp (VS56 and VS60), 880bp (VS57 and VS61) and 872bp (VS58 and VS59) of *PvispG* gene were obtained **(Figure 5.5b).** The reaction conditions used for the above PCR are given in **Table 5.2**. Amplicons obtained from different field isolates were purified using QIAquick Gel Extraction kit and sequenced commercially.

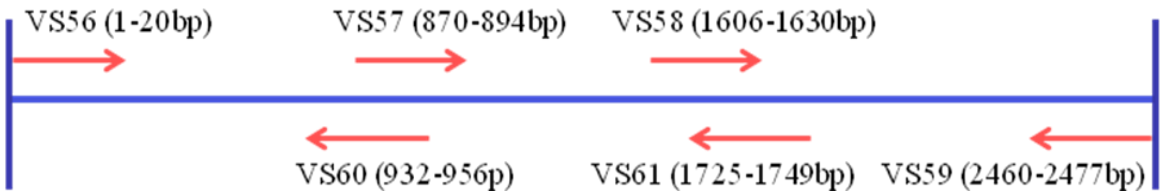


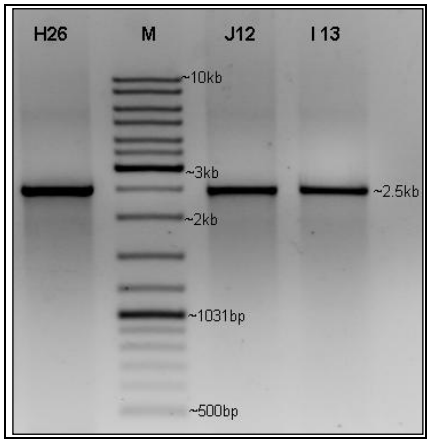
Figure 5.4: Position of primers designed for *PvispG* gene

Table 5.1: Primers used for the amplification of *ispG* gene from *P. vivax* nuclear genome (Restriction sites are highlighted in bold)

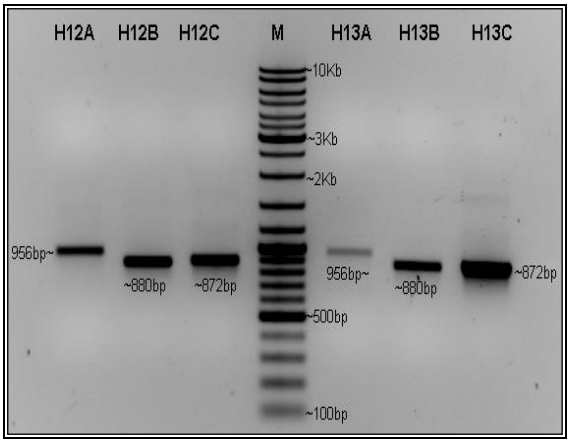
Lab Id	**Primer sequence**	**Restriction enzyme site incorporated**
VS56	5' AGC **CTC GAG** ATG GTA TCC CTA TAT TTT TG 3'	XhoI
VS57	5' GGT TTT CTC CAT GAA GGC TTC CAAC 3'	NA
VS58	5' GCC AAC TTG AAT CAC ATG CAG ACC C 3'	NA
VS59	5' AGC **AAG CTT** ATA CCC GTA GAA GGC TAC 3'	HindIII
VS60	5' TTG CCC ATC CTT TCA TAT TGC CTC G 3'	NA
VS61	5' GGT TAA ACA CTT GCT CCA CCA GGT G 3'	NA

Table 5.2: Reaction conditions employed for the amplification of *PvispG* gene

Reaction Steps	**Primers used**				
	VS56 and VS59 (2463bp)	VS56 and VS60 (956bp)	VS57 and VS61 (880bp)	VS58 and VS59 (872bp)	
Pre-Denaturation	94°C for 3 min	94°C for 3 min	94°C for 3 min	94°C for 3 min	
Denaturation	94°C for 1 min	94°C for 1 min	94°C for 1 min	94°C for 1 min	X35
Annealing	56.6°C for 1 min	56.6°C for 1 min	54.6°C for 1 min	56.2°C for 1 min	
Extension	72°C for 1.5 min	72°C for 1.5 min	72°C for 1.5 min	72°C for 1.5 min	
Post-Extension	72°C for 4min	72°C for 4min	72°C for 4min	72°C for 4min	



a) Complete *PvispG* gene amplification (2463bp)



b) Amplification of *PvispG* gene fragments: upper (956bp), middle (880bp) and lower (872bp) region.

Figure 5.5: *PvispG* gene amplification (Parasite DNA samples: H12, H13, H26, J12 and I3; M: Gene Ruler DNA Ladder mix Fermentas SM0331).

5.2.2.2 Primer designing and amplification of *Pf*IspG gene

P. falciparum ispG gene (PF3D7_1022800; PlasmoDB) is 2475bp in size and expands from position 955883-958357 on chromosome 10 in *P. falciparum* genome sequence. Primers for the full length gene amplification as well as partial fragments were designed (**Figure 5.6**; **Table 5.3**). Amplifications were performed following the conditions as given in **Table 5.4**.

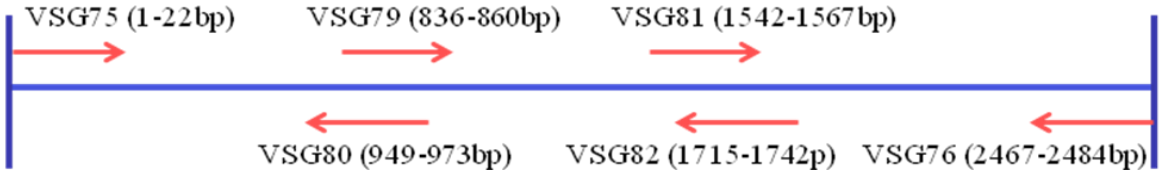


Figure 5.6: Position of primers designed for *P. falciparum ispG* gene.

Table 5.3: Primers for the amplification of *ispG* gene from *P. falciparum* nuclear genome (Restriction sites are highlighted in bold)

Lab Id	**Primer sequence**	**Restriction enzyme site incorporated**
VSG 75	5’ CGG **CAT ATG** AGT TAT ATA AAA AGA CTG A 3’	NdeI
VSG 79	5’ CTC GAG TAT TAT CAT ATT ATG GAG A 3’	NA
VSG 81	5’ GCA AGT GGG AAT TAA TAT AGT AGT TC 3’	NA
VSG 76	5’ GGT **GTC GAC** ATA TTC AAT TTA TGG ATC 3’	SalI
VSG 80	5’ GTA TCA TAA CAT AAG CAT TAG AAG C 3’	NA
VSG 82	5’ CCG TTT AAA ATT AAT CCA TAT CCT TTA G 3’	NA

Table 5.4 Reaction conditions employed for the amplification of *PfispG* gene

Reaction Steps	Primers used			
	VSG75 and VSG76 (2475bp)	VSG75 and VSG80 (973bp)	VSG79 and VSG82 (906bp)	VSG81 and VSG76 (942bp)
Pre -Denaturation	94°C for 3 min	94°C for 3 min	94°C for 3 min	94°C for 3 min
Denaturation	94°C for 1.5 min	94°C for 1 min	94°C for 1 min	94°C for 1 min
Annealing	63.3°C for 2 min	50°C for 1 min	50°C for 1 min	52°C for 1 min
Extension	72°C for 2.5 min	72°C for 1.5 min	72°C for 1.5 min	72°C for 1.5 min
Post-Extension	72°C for 4min	72°C for 4min	72°C for 4min	72°C for 4min

X35

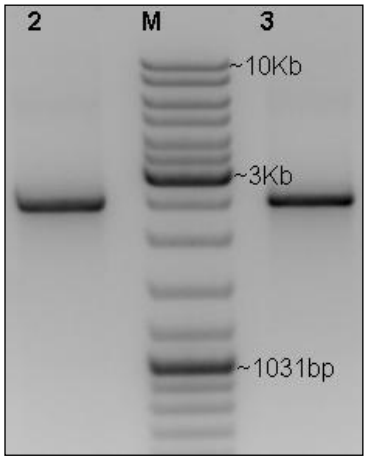


Figure 5.7: Complete *PfispG* gene amplification (2484 bp).

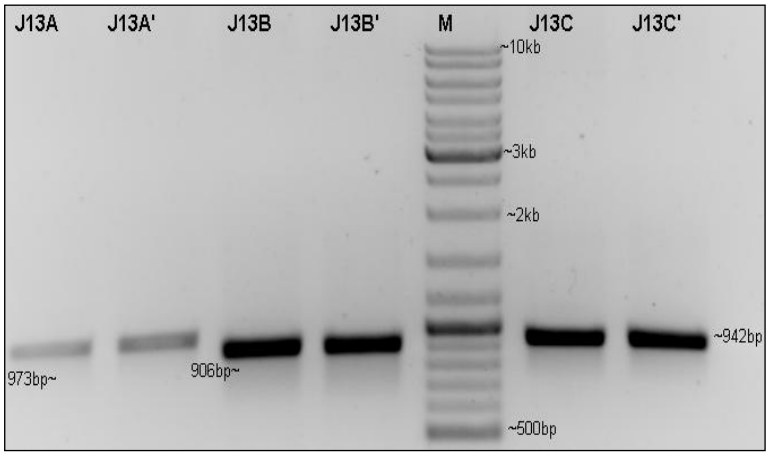


Figure 5.8: Amplification of *PfispG* gene fragment: upper (973bp), middle (906bp) and lower (942bp) region. (J13 - Parasite DNA sample; M: Gene Ruler DNA Ladder mix Fermentas SM0331)

Amplicons obtained from different field isolates (**Figure 5.7 & 5.8**) were gel purified and commercially sequenced. The obtained sequences of Indian *Pv*IspG (BankIt ID: 1761743) and *Pf*IspG (BankIt ID: KX176850) were analysed at the genomic and proteomic level within various Indian isolates and were found to be similar without any variations.

5.2.3 Conserved domain and Signature motif analysis

IspG enzyme of *P. falciparum* and *P. vivax* in PlasmoDB is enlisted as a putative enzyme. A GcpE super family domain is a characteristic feature of all IspG proteins and the presence of this domain in the obtained *Pv*IspG and *Pf*IspG protein sequences from Indian isolates was confirmed with the help of Conserved Domain Detection (CDD) tool available at the NCBI website (**Figure 5.9 & 5.10)**.

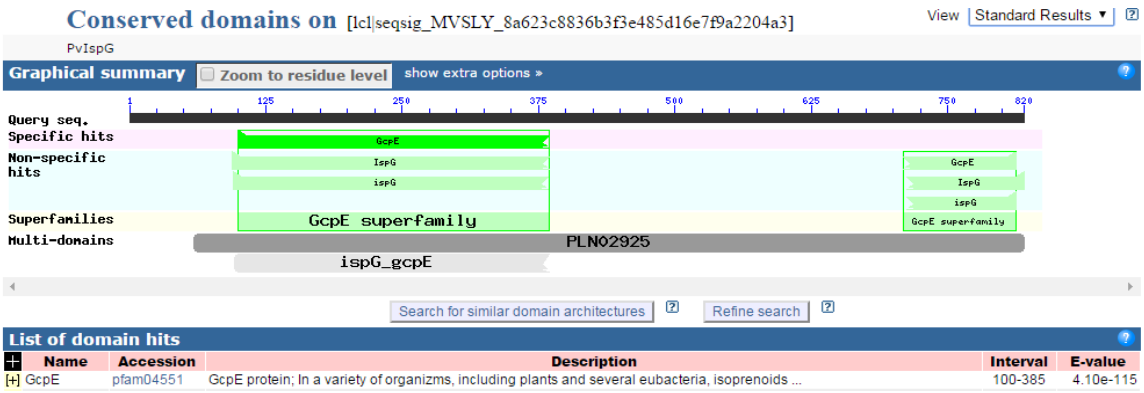


Figure 5.9: Conserved domain detection (CDD) of Indian *Pv*IspG protein.

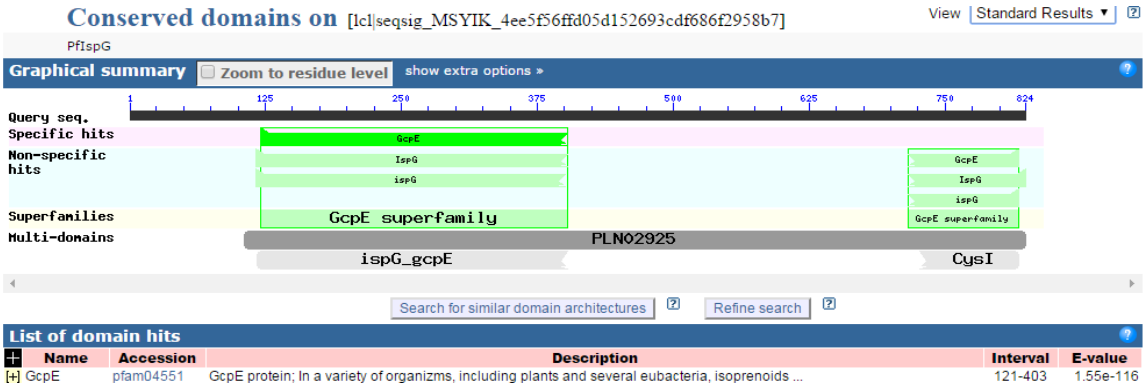


Figure 5.10: Conserved domain detection (CDD) of Indian *Pf*IspG protein.

CDD analysis showed that GcpE domain of *Pv*IspG and *Pf*IspG protein spans in two segments; one present on N-terminal (amino acid residues 100 - 385 in *P. vivax* and 121-401 in *P. falciparum*) known as 'A domain' and other on C-terminal (amino acid residues 708 – 814 in *P. vivax* and 715-818 in *P. falciparum*) known as 'B domain'. Both these domains were found to be conserved among all human, primate and rodent malaria parasites. The detection of signature motifs in *Pv*IspG and *Pf*IspG proteins using PROSITE showed the presence of multiple conserved residues in these domains. The conserved residues identified in *P. vivax* IspG were Q119; M121; R148, 198, 248, 359; D179; N200, 252; S293, 361; K295; H326; T330; E331, 769 and G768 (**Figure 5.11**) along with the three cysteine residues at positions 728, 731 and 762. These residues were also found to be conserved in IspG proteins obtained from other organisms including *P. falciparum* (**Table 2.2**), but the position of these residues varies within these organisms. In *Pf*IspG protein, the signature motif consists of similar conserved residues at positions Q140; M142; R169, 218, 267 and 376; D200; N220, 272; S314, 379; K316; H343; T348; Glu 349, 773; C732, 735 and 766 and G772 **(Figure 5.11)**. The presence of these conserved domain and signature motifs confirms these proteins as an IspG enzyme.

5.2.4 Sequence Analysis and Multiple Sequence Alignment

The obtained sequences of Indian *Pv*IspG (BankIt ID: 1761743) and *Pf*IspG (BankIt ID: KX176850) were analysed at the genomic and proteomic level with homologues/ orthologues from other organisms including prokaryotes, apicomplexans and plants (Acc. No. given in **Table 2.2**) by performing multiple sequence alignment using Clustal W2.

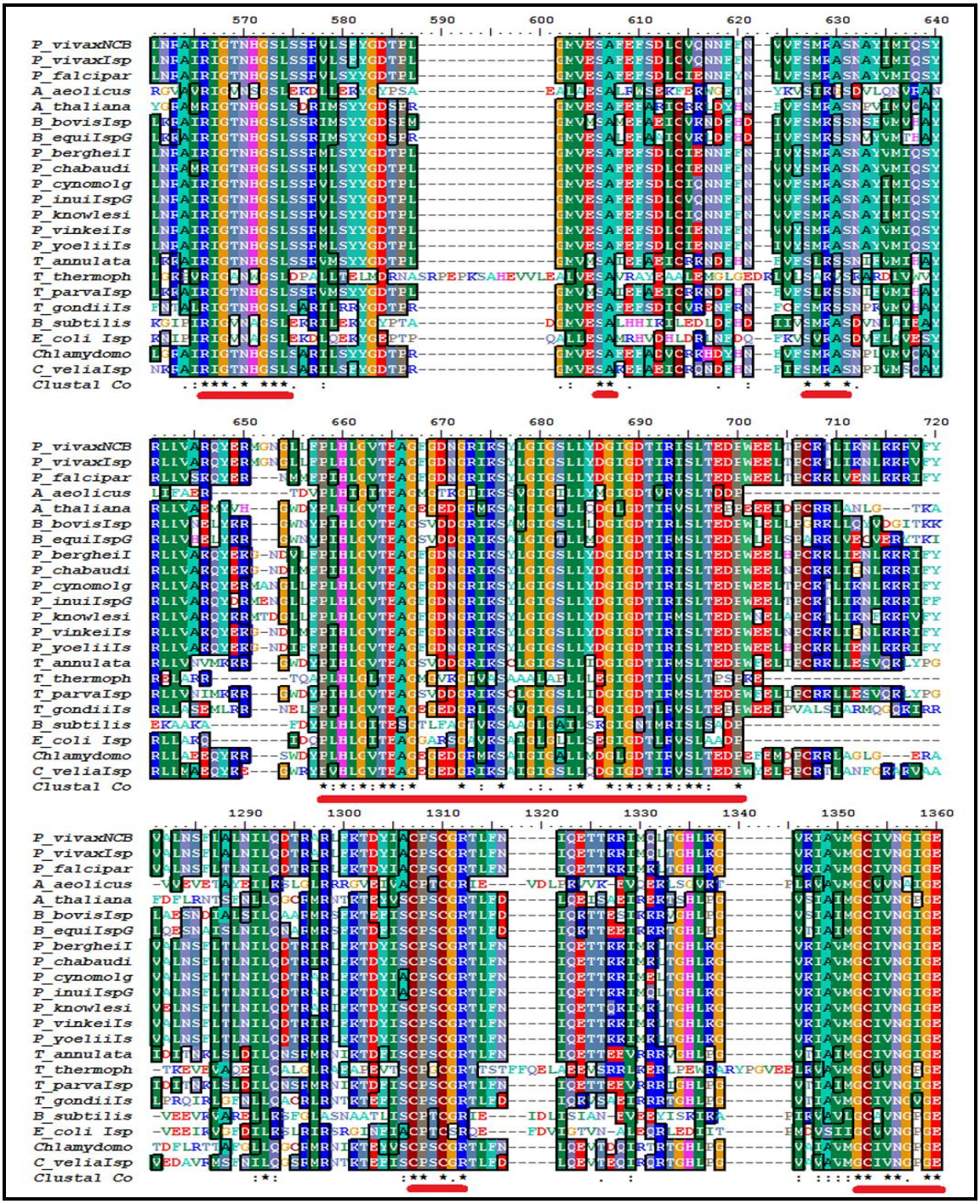


Figure 5.11: Multiple sequence alignment of *P. vivax* IspG sequence with other apicomplexans and bacterial homologues. Conserved residues of the signature motifs are depicted as red lines

In MSA, the *P. vivax* sequences obtained from Indian field isolates show 100% similarity among each other while their percent identity ranges from 30 to 84% with other homologues/ orthologues. Within the *Plasmodium* genus, at the amino acid level a match of 82.94% with *P. cynomolgi*, 81.92% with *P. knowlesi* and 63.05% with *P. falciparum* was obtained, suggesting *P.vivax* IspG to be more closely related to the primate parasites. Sequence identity between these sequences was also checked at nucleotide level where slight variation in similarity score was observed. In comparison, the sequence obtained for IspG protein of *P. falciparum* showed a match of 70.63% with *P. vinckei*, 69.86% with *P. chabaudi* and 69.66% with *P. yoelii*, showing more closeness of *P. falciparum* to the rodent parasites. The similarity score of Indian *P. vivax* IspG and *P. falciparum* IspG with other *Plasmodium* species, apicomplexans and prokaryotes at both the nucleotide and amino acid levels are shown in **Table 5.5**.

Table 5.5: Similarity index of *P. vivax* IspG nucleotide (brown) and protein (blue) sequence from Indian field isolates with a) other *Plasmodium* species b) apicomplexans and *Chromera velia* c) prokaryotes

a)

Nucleotide/ Protein	***P. vivax***	***P. falciparum***	***P. berghei***	***P. chabaudi***	***P. cynomolgi***	***P. inui***	***P. knowlesi***	***P. vinckei***	***P. yoelii***
P. vivax		63.3	62.4	62.8	82.5	38.7	83.4	62.8	37.3
P. falciparum	65.7		77.0	76.4	63.9	38.7	66.3	76.6	42.5
P. berghei	66.7	71.9		93.5	63.1	37.4	66.0	94.0	41.4
P. chabaudi	67.6	71.9	91.2		63.4	37.2	65.8	96.9	40.4
P. cynomolgi	83.5	65.3	66.4	66.1		38.8	85.2	63.5	36.3
P. inui	85.5	67.3	67.1	68.1	86.3		39.2	36.7	66.0
P. knowlesi	83.4	66.8	68.1	68.7	81.9	85.5		65.8	38.1
P. vinckei	67.9	72.2	92.1	96.9	66.6	69.0	69.1		40.3
P. yoelii	66.5	71.5	95.0	91.7	65.8	66.8	67.3	92.6	

b)

Nucleotide/ Protein	***P. vivax***	***P. falciparum***	***A. aeolicus***	***A. thaliana***	***T. thermophiles***	***B. subtilis***	***E. coli***	***C. reinhardtii***	***C. velia***
P. vivax		64.1	39.1	46.3	43.5	47.5	46.3	49.1	48.8
P. falciparum	65.7		40.2	47.5	36.5	48.3	45.4	40.5	45.5
A. aeolicus	35.0	35.3		37.8	28.3	36.9	36.9	36.8	41.4
A. thaliana	38.9	38.7	36.1		41.5	47.5	46.0	54.6	53.3
T. thermophiles	36.2	37.1	40.9	34.6		44.5	47.1	54.3	46.6
B. subtilis	35.7	35.1	49.6	37.4	38.1		53.6	47.2	43.1
E. coli	35.7	36.2	46.2	35.2	38.1	46.0		48.6	43.8
C. reinhardtii	43.8	44.4	36.1	62.9	35.7	36.0	35.4		58.4
C. velia	42.8	42.4	35.5	55.3	36.0	36.3	35.2	60.4	

c)

Nucleotide/ Protein	***P. vivax***	***P. falciparum***	***B. bovis***	***B. equi***	***T. annulata***	***T. parva***	***T. gondii***
P. vivax		64.1	48.7	48.4	46.9	47.4	44.7
P. falciparum	65.7		50.0	53.7	52.3	53.2	42.3
B. bovis	43.6	43.9		61.2	55.8	56.1	45.9
B. equi	43.7	45.8	62.8		61.3	61.8	46.3
T. annulata	43.4	45.6	55.7	62.6		86.5	46.2
T. parva	43.0	44.9	55.2	62.6	90.8		46.3
T. gondii	40.5	39.7	42.3	44.2	38.5	38.5	

5.2.5 Phylogenetic analysis

The evolutionary position of *Pv*IspG protein and *Pf*IspG was predicted by constructing a phylogenetic tree using MEGA 6.0 (Tamura et al., 2013). The MSA generated for phylogenetic tree construction included 22 amino acid sequences consisting of major apicomplexans (*Plasmodium, Toxoplasma, Babesia, Theileria*), prokaryotes (*E. coli, B. subtilis, A. aeolicus, T. thermophilus*), *Arabidopsis,* chlorophyta (*Chlamydomonas)* and *Chromera velia* (Acc. No. of all the sequences are given in **Table 2.2**). The phylogenetic tree for *Pv*IspG protein was constructed (**Figure 5.12**) using the Maximum Likelihood method based on JJT matrix-based model.

As expected, *Pv*IspG protein grouped along with other primate parasite orthologues while rodent malaria parasites *P. berghei*, *P. chabaudi*, *P. yoelii* and *P. vinckei* formed a different clade. These two clades seemed to diverge out of *P. falciparum* clade indicating a separate origin for *P. falciparum*. This evolutionary analysis of *Pv*IspG protein is in agreement to other reports in various studies (Rathore et al., 2001; Escalante et al., 2005; Saxena et al., 2007 & 2012, Martinson et al., 2008) detailing the origin of *P. vivax* as a primate parasite. *Chromera velia,* recently identified bridging link between algae and apicomplexans (Weatherby and Carter, 2013) chlorophyta and higher plants (all containing photosynthetic plastid), formed a single separate clade showing their close association with each other but evolutionary divergence from other apicomplexans and *Plasmodium* species (harbouring non photosynthetic plastids). An average evolutionary distance of 0.309 was obtained among these proteins.

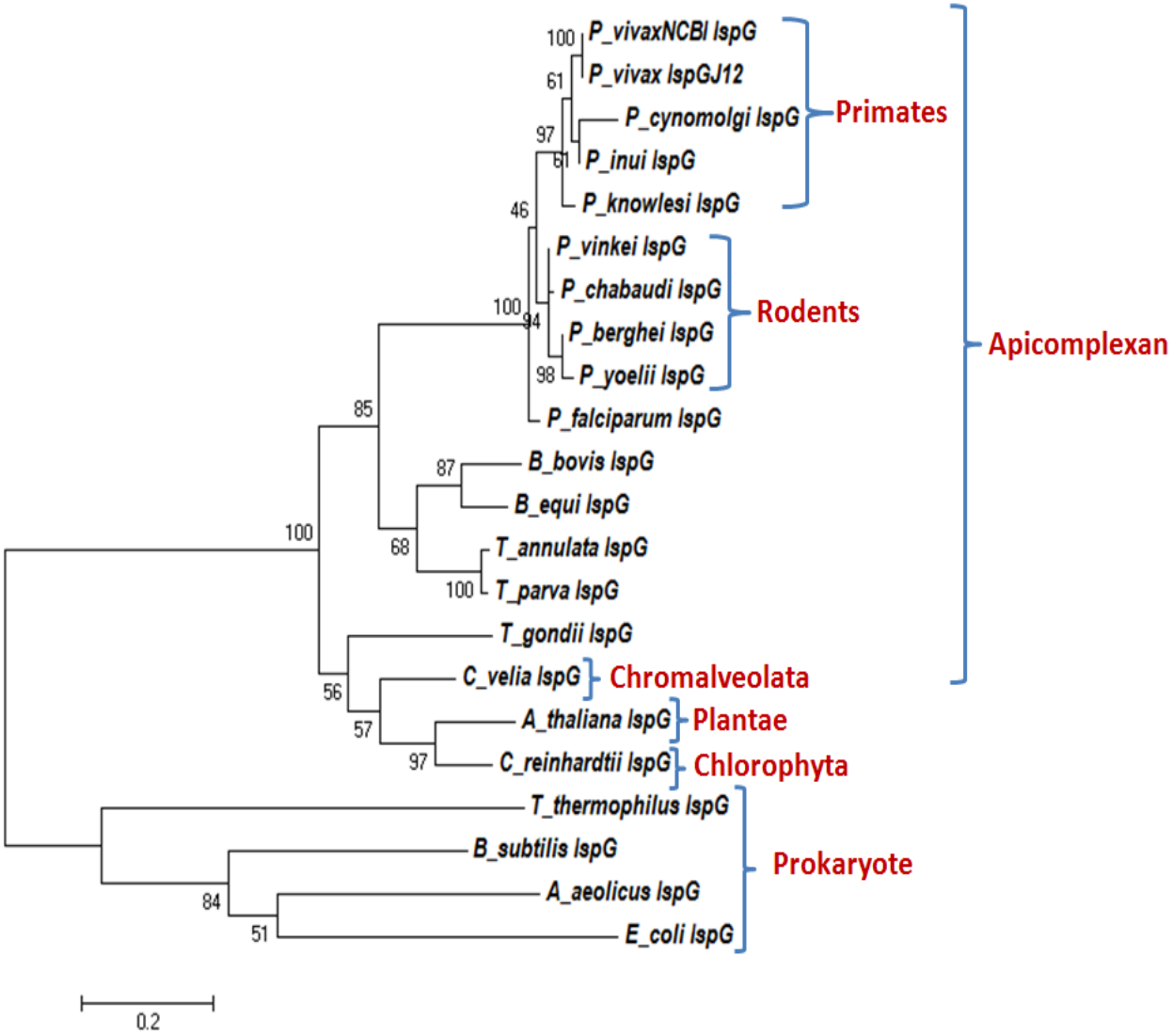


Figure 5.12: Phylogenetic analysis of Indian *P. vivax* IspG protein with the IspG protein sequences of different apicomplexans, chromalveolate, plants and prokaryotes. Boot strapping was done for 1000 iteration and percentage of trees in which the associated taxa clustered together is shown next to branches.

5.2.6 Cloning and expression of *PvispG* gene

After initial confirmation of the protein as *Pv*IspG and its targeting to the apicoplast using the bioinformatic tools, the amplified *PvispG* gene was cloned in pRSET A vector using XhoI and HindIII restriction sites (**Figure 5.13**) incorporated in the primer. The colonies obtained were first screened through the gel shift assay (**Figure 5.14**) and the shortlisted colonies were analysed using restriction digestion to check the correct orientation of *PvispG* gene (**Figure 5.15**). The integrity of the gene sequence in the clone and the presence of start codon in frame was also confirmed by sequencing using T7 and gene specific primers.

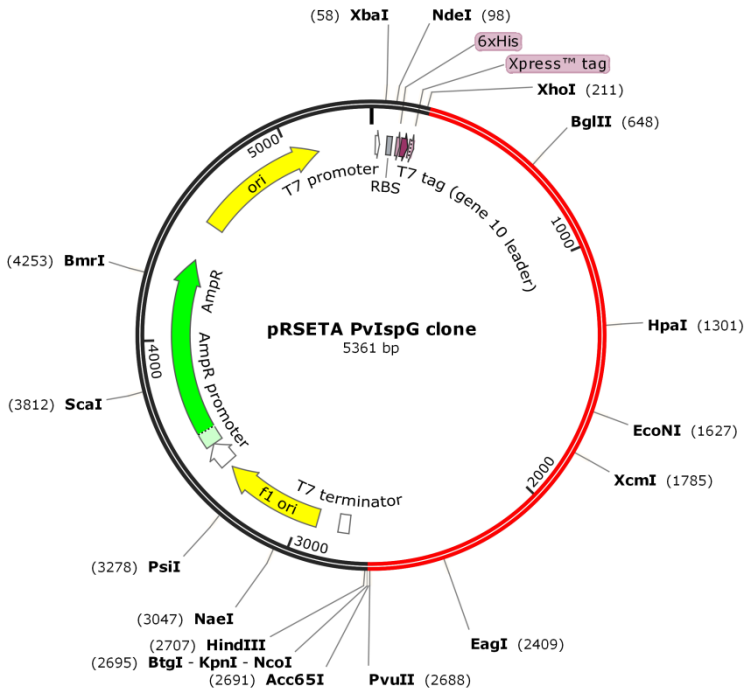


Figure 5.13: The pRSET A-*Pv*IspG clone map: Red part depicts *Pv*IspG gene and black part shows the regions of pRSET A vector backbone. Restriction sites marked in the gene were used for restriction analysis of the clone.

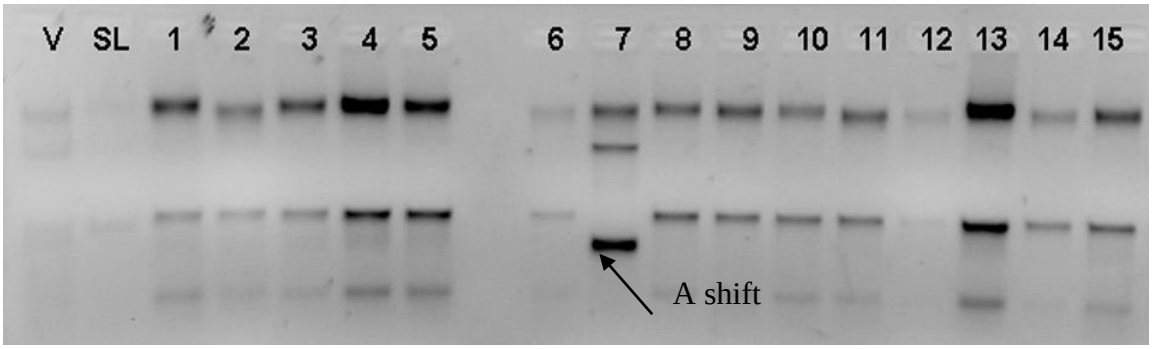


Figure 5.14: Gel shift assay to check for the colonies showing recombinant construct (V: Vector; SL: Self ligation control; 1-15: colony number after vector insert ligation and transformation)

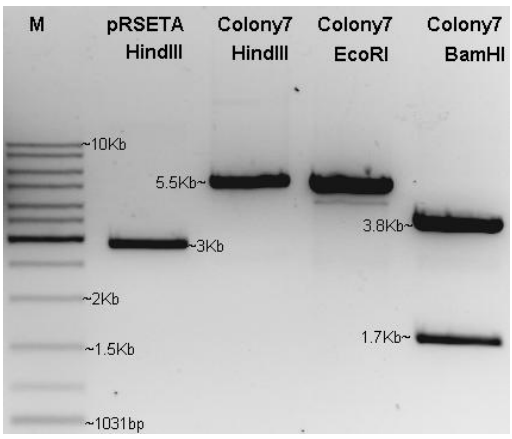


Figure 5.15: Restriction analysis of pRSET A *PvispG* recombinant constructs. M: Gene Ruler DNA Ladder Mix (Thermo Scientific, USA). lane2: pRSETA plasmid Hind III digested; Lane 3, 4, and 5: Recombinant construct digested with HindIII, EcoRI and BamHI restriction enzymes respectively.

After successful validation of the obtained clone, to express the recombinant protein, pRSET A-*PvispG* clone was transformed into *E. coli* BL21 (DE3) pLysS host cells. Recombinant strains were grown in LB broth containing 100 µg/mL ampicillin and 35µg/mL chloramphenicol at 37°C overnight in an orbital incubator shaker, as discussed in Chapter 2. After inducing the culture with 0.5mM IPTG, cell lysate samples were collected at different time intervals (0 hr, 2 hr, 4 hr, 6 hr etc.) and the expression was checked on 10% SDS PAGE. Proteins were visualized after staining with Coomassie brilliant blue R250 (**Figure 5.16**). The recombinant *Pv*IspG protein was successfully expressed as a His-tagged protein as confirmed by western blotting (**Figure 5.17**) using anti-His antibody.

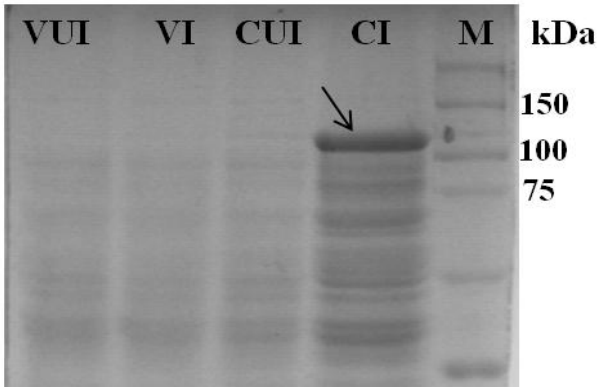


Figure 5.16: Expression of full length *Pv*IspG protein in *E. coli*. M: Western Blotting Protein Standards (BioRad). VUI: pRSET A uninduced; VI: pRSET A induced; CUI: pRSET A *PvispG* uninduced; CI: pRSET A *PvispG* induced (the arrow indicates the induced band of *Pv*IspG).

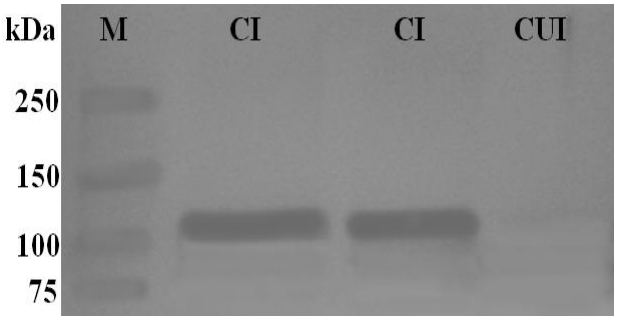


Figure 5.17: Western blot analysis of Recombinant *Pv*IspG protein using anti-His antibodies. M: Western Blotting Protein Standards (BioRad). CI: pRSET A *PvispG* induced; CUI: pRSET A *PvispG* uninduced.

5.2.7 Protein isolation and purification

After standardizing the conditions for protein expression, for further analyses like activity assays and immune-localization studies, the *Pv*IspG recombinant protein was purified using Ni-NTA agarose resin column following the denaturing protocol 2.9.1. The specific his tagged protein was eluted in different fractions by using elution buffer (pH 4.5) (**Figure 5.18**). *Pv*IspG protein concentration was determined by Bradford assay where known concentration of BSA was used as a standard.

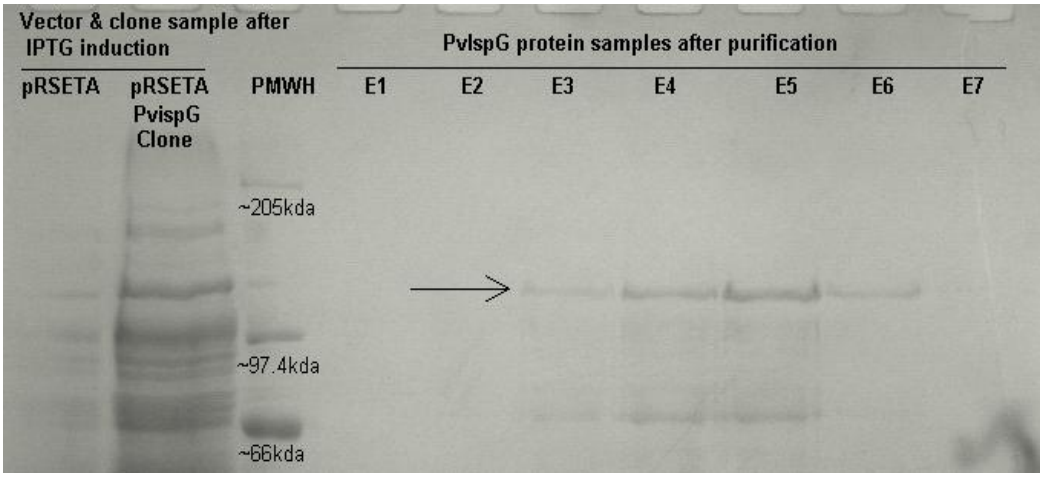


Figure 5.18: Profile of eluted fractions after Ni-NTA of His tagged recombinant *Pv*IspG protein. PMWH: Protein High Molecular Weight Marker (Merck Bio Science); E1 to E7: eluted fractions of *Pv*IspG

5.2.8 MALDI-TOF-MS analysis of the *Pv*IspG

As the induced protein band was obtained slightly above 100 kDa as compared to expected 98.7 kDa, even after repeated trials, the purified *Pv*IspG protein was subjected to MS-MS and retrieved data was analysed by MASCOT server (AB SCIEX, USA). Using mass spectroscopy analysis, the mass/ charge ratio of the intact *Pv*IspG protein size was obtained as 98.7 kDa (**Figure 5.19**). Unavailability of the data for *Pv*IspG protein in the MASCOT server database was a major obstacle in our analysis, however, similarity of our protein with GcpE domain of *Desulfotalea psychrophila*, a sulfate reducing bacteria confirmed the protein as IspG.

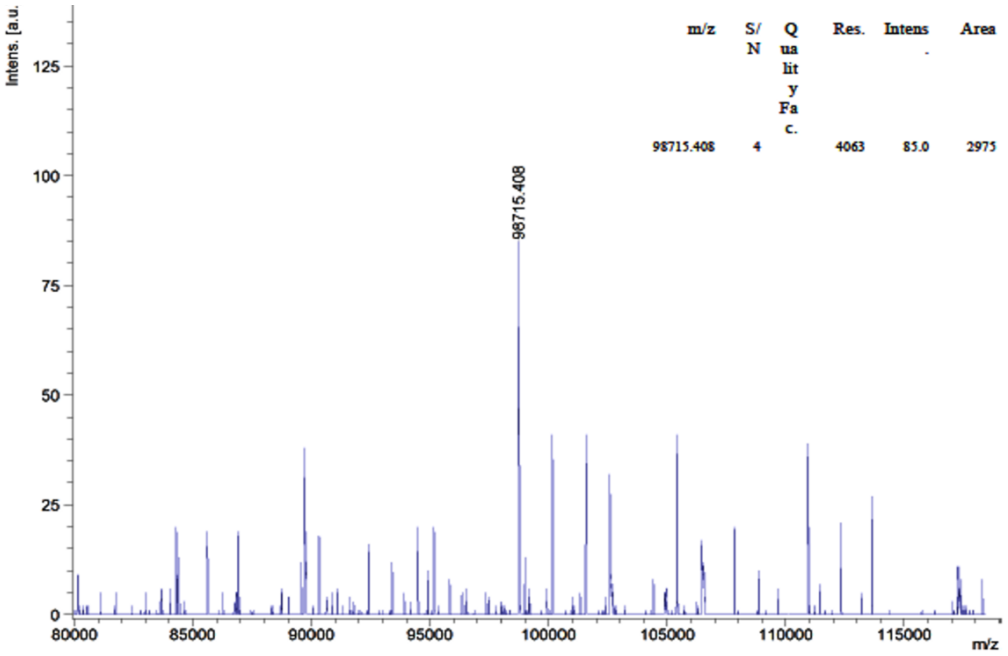


Figure 5.19: MS-MS analysis of the purified *Pv*IspG protein to determine the size of the protein.

5.2.9 Analysis of Fe-S clusters on *Pv*IspG protein

5.2.9.1 Reconstitution of Fe-S clusters on *Pv*IspG protein

The activity of *Pv*IspG protein is highly dependent on the Fe-S clusters. However, as the purification of *Pv*IspG protein was done in aerobic conditions, it resulted in the oxidation of this protein which ultimately led to the removal of the Fe-S cluster from the protein. To reconstitute these clusters on the protein, the whole reaction setup was done under anaerobic environmental conditions. Initially, 5µM of purified *Pv*IspG protein was treated with 50-fold molar excess of dithiothreitol (dTT) for 10 min on ice. 5 folds molar excess of ammonium iron (II) sulfate was then added, followed by addition of Na_2S (5 folds molar excess) drop wise over 10 min. Reaction mix was incubated at 25°C for 4 hr at 100 rpm followed by desalting of reaction mix with Amicon® Ultra-4 filters (Merck Millipore, UK) already equilibrated with 50mM HEPES-NaOH buffer (pH 7.5). To record the UV-visible absorption spectrum, a fraction of the reconstituted protein was anaerobiacally transferred into a cuvette, and UV- Vis spectra of originally purified protein and reconstituted protein were analysed at 350-450 nm. Maximum absorption was obtained at ~395 nm for the reconstituted protein only, which is typical for protein containing Fe-S clusters (as discussed by Okada and Hase, 2005), thus confirming the reconstitution of these clusters.

5.2.9.2 Rapid Colorimetric Assay: Iron and Sulphur determination

After reconstitution, to check the amount of Fe and S bound to IspG, the Ferrozine and methylene blue assay was performed according to Fish, 1988 and Siegel, 1965 respectively. For the Ferrozine assay, the iron standard solutions of varying concentration (1µg to 6µg/mL) were prepared by dissolving required concentration of ferrous ethylene diammonium sulfate tetra hydrate salt (SIGMA, USA) in 0.01M HCl. To these standard solutions and the protein sample, 0.5mL of freshly prepared iron releasing reagent containing 0.6M HCl and 0.142M potassium permanganate ($KMnO_4$) was added. The digested mixture was incubated in dark for 2 hr at 60°C followed by the addition of 0.1mL of reducing, iron chelating reagent containing 6.5mM ferrozine (disodium 3-(2-pyridyl)-5,6-bis(4-phenylsulfonate)-1,2,4-triazine), 13.1mM neocuprine (2,9-dimethyl-1,10-phenanthroline), 2M ascorbic acid and 5M ammonium acetate. The solution was left to stand at room temperature for at least 30 min. After this, the absorbance of standards and protein samples were measured at 562 nm (Fish, 1988).

For sulfur determination, sodium sulfide standards were prepared in sodium carbonate buffer (pH 9.0) with concentration varying from 0.01mM to 0.1mM. Standards and protein sample were further treated with 0.02M N, N-Dimethyl-p-phenylene-diamine sulfate (DPD) and 0.03M Ferric chloride ($FeCl_3$). After incubation at room temperature for 10 min, absorbance was taken at 650 nm (Siegel, 1965).

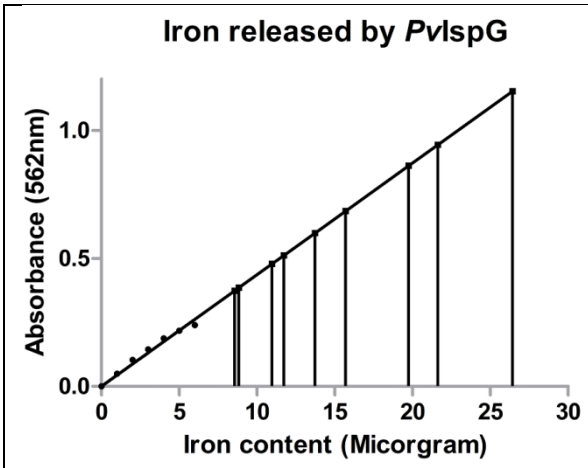


Figure 5.20a: A regular standard curve to determine the iron content by Ferrozine assay

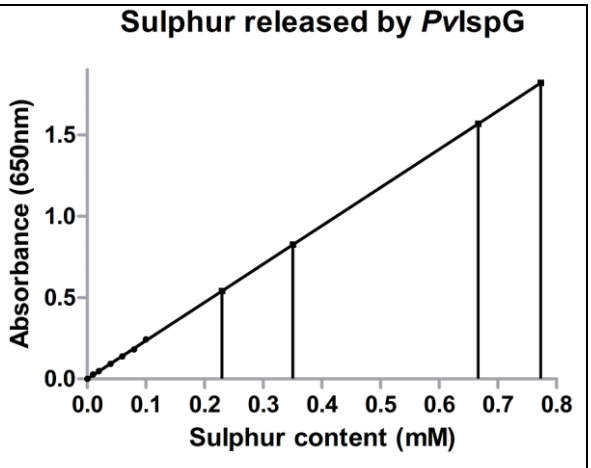


Figure 5.20b: A regular standard curve to determine the sulfide content by methylene blue assay

The calibration curves were prepared separately for iron and sulfur assay for standards and proteins. Quantity of iron and sulfur released by *Pv*IspG protein was calculated by extrapolating the absorbance of these calibration curves (**Figure 5.20 (a) & (b)).** The total iron content of reconstituted *Pv*IspG was calculated as 8.4µM (µmole protein)$^{-1}$ by the Ferrozine assay while the sulphur content was found to be 12.8µM (µmole protein)$^{-1}$ in methylene blue assay, indicating binding of [4Fe-4S] clusters (Okada and Hase, 2005).

5.2.9.3 Electron Paramagnetic Resonance (EPR) Study

Further the above reconstituted protein complexed with Fe-S clusters was analyzed by EPR Spectroscopy to check for the oxidation state of bound Fe-S cluster (Wang et al., 2011). Briefly, the samples were diluted to a concentration of 5µM in 20mM phosphate buffer containing 0.5M NaCl (pH 7.4) followed by incubation with 5mM ammonium persulfate for 30 min to oxidize *Pv*IspG for spectra analysis. X-band EPR spectra were recorded at 100K on a JES-FA200 spectrometer. Parameters for recording the EPR spectra were typically 15–30mT/ min sweep rate, 0.63mT modulation amplitude, 9.44GHz frequency, and 4mW incident microwave power with a sweep time of 2 min.

The purified *Pv*IspG protein in the reduced state does not exhibited any EPR activity, however, the oxidized *Pv*IspG gave an S=5/2 signal which shows the presence of Fe^{3+} oxidation state of iron (**Figure 5.21**). The g value was estimated as 2.18. The results are in agreement with previous reports for Fe-S containing proteins and indicate that the [4Fe-4S] clusters were successfully incorporated in *Pv*IspG protein upon reconstitution.

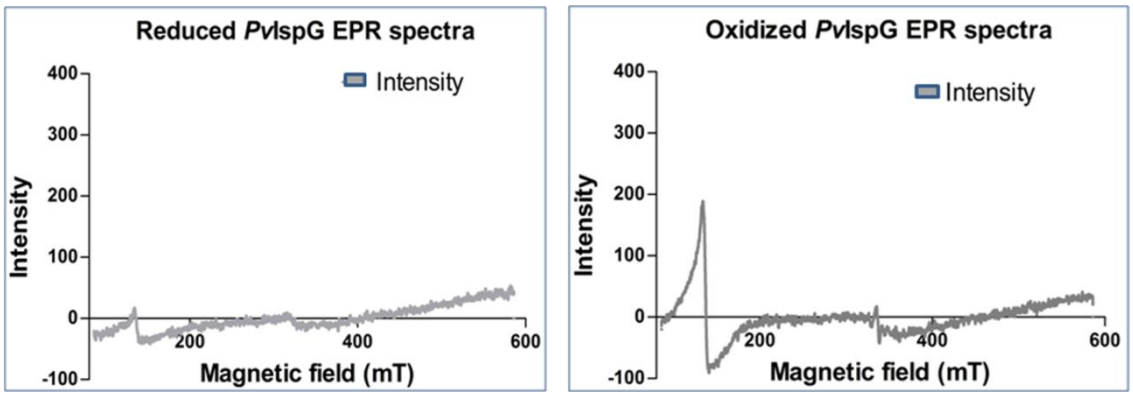


Figure 5.21: Electron paramagnetic resonance spectra of reduced and oxidized recombinant *Pv*IspG.

5.2.10 Antibody raising and Immuno-localization

The isoprenoids biosynthesis pathway is said to be localized to the apicoplast in *Plasmodium*. While the PlasmoAP showed the presence of apicoplast transit peptide in *Pv*IspG and PATS supported its apicoplast localization, the MitoProt and PlasMit showed this protein to be targeted to mitochondria as well, thus suggesting its dual targeting. Thus to confirm the subcellular localization of *Pv*IspG, antibodies was raised against full length recombinant *Pv*IspG in mice following the procedure as detailed in Chapter 2. Antibody titer was calculated (Harlow and Lane, 1988) for the sera collected after different booster doses of *Pv*IspG protein using standard ELISA where serum collected from mice injected with 1X PBS was used as a control. Pre-immune sera was collected to check presence of any cross-reactive antibodies. Best antibody titre was observed after third booster dose (**Figure 5.22**). Specific band obtained at the desired position in western blotting confirms specificity of antibodies towards *Pv*IspG protein (**Figure 5.23**).

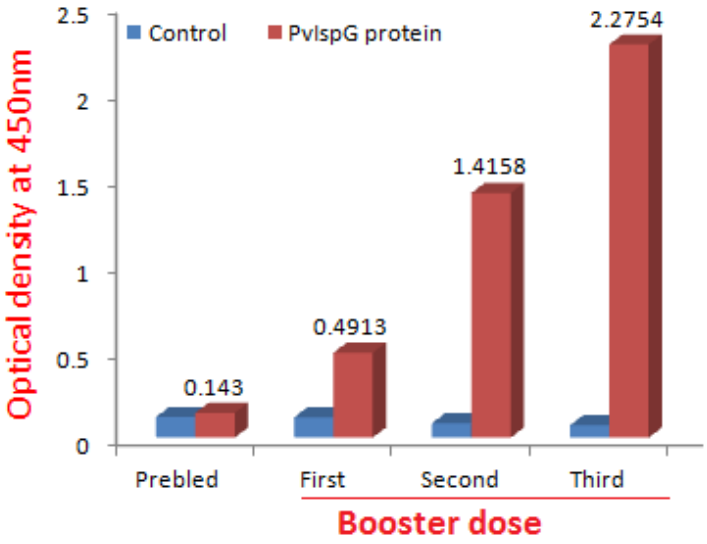


Figure 5.22: Antibody titre for *Pv*IspG protein by performing ELISA.

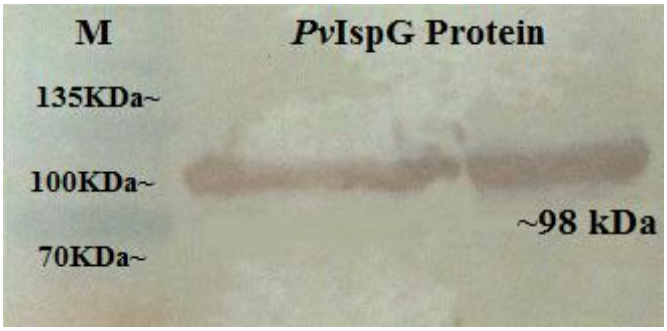


Figure 5.23: Western blot of anti- *Pv*IspG antibodies and *Pv*IspG.

To perform sub-cellular localization on blood smears prepared from *P. vivax* infected patients, parasite cells were fixed, permeabilised and sequentially treated with mouse polyclonal anti-*Pv*IspG sera followed by goat anti-mouse monoclonal IgG FITC antibodies. DAPI and Qdot® 585 Streptavidin conjugate counter staining was done to observe the parasite's nucleus and apicoplast respectively. Slides were observed using the

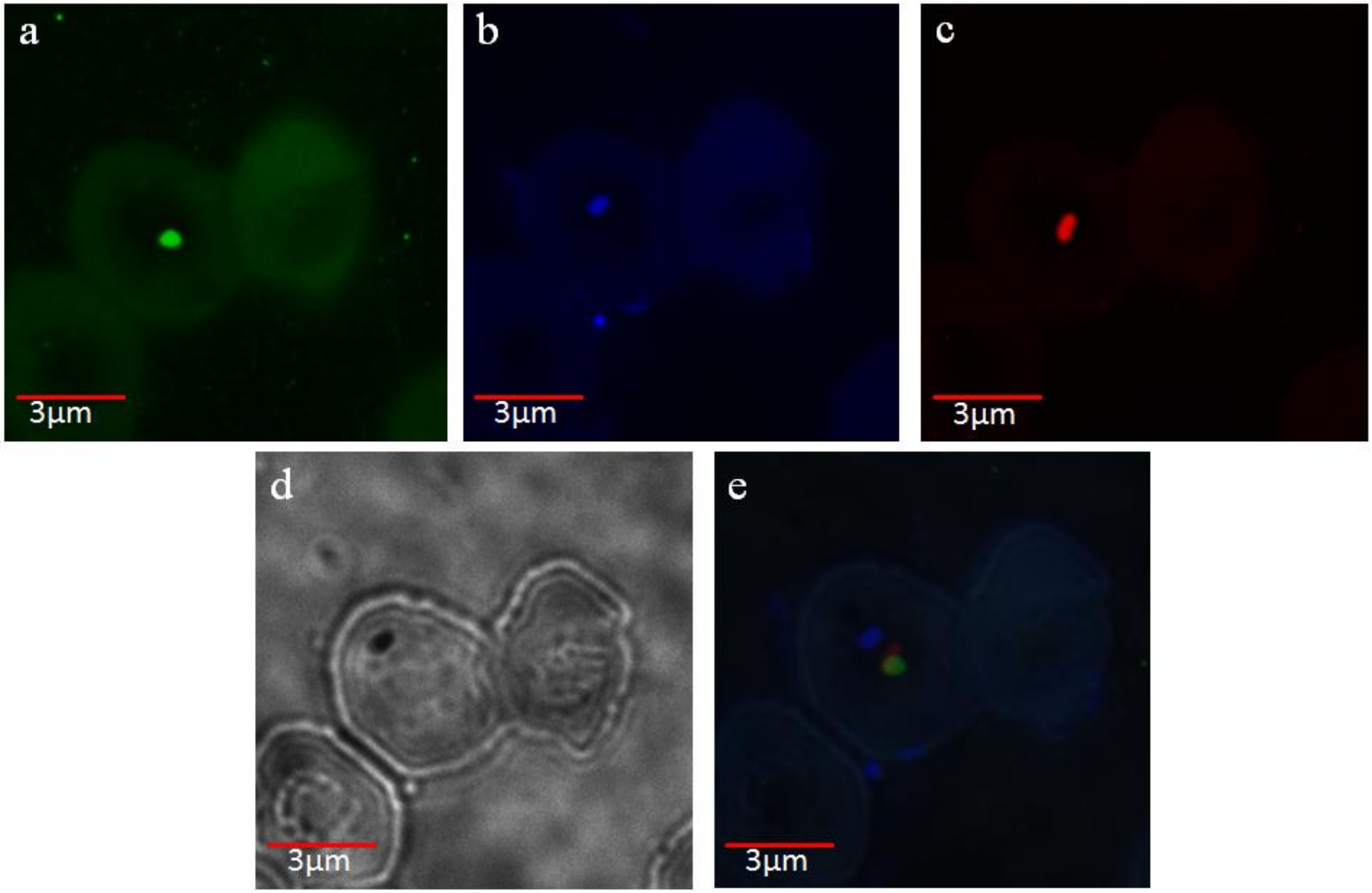


Figure 5.24: Sub cellular localization of *Pv*IspG protein in *P. vivax* Indian field isolates. a. Staining with FITC, b. DAPI, c. Qdot ® 585 Streptavidin, d. Bright field, e Merged

respective filters, where a bright field was used to identify the RBCs infected with parasite, green fluorescence to localize the *Pv*IspG protein, blue and red fluorescence for nucleus and apicoplast respectively (**Figure 5.24**). The *Pv*IspG protein was localized as a green fluorescing spot with an overlapping red spot of apicoplast and a blue spot staining nucleus adjacent to this merged yellowish spot. This study confirms that *Pv*IspG protein is functionally active in the apicoplast.

5.2.11 Analysis of anti-*Pv*IspG antibodies using Flow Cytometry

To confirm that the anti-*Pv*IspG antibodies are able to detect the *P. vivax* apicoplast and to further elucidate their specificity to the parasite, Flow Cytometry analysis of antibodies against the field isolates was performed. The uRBC and iRBC were treated with antibodies as discussed in **2.13**. Differrent tubes with varying combinations of antibodies were set **(Table 5.6).**

Table 5.6: Sample sets for FC analysis

Reaction	Fresh RBC	Parasite (*P. vivax*)	Parasite (*P. falciparum*)	Anti - *Pv*IspG	AlexaF 488	Anti-*Pv*BiP AF647
Tube 1	+					
Tube 2	+			+		
Tube 3	+				+	
Tube 4	+			+	+	
Tube 5		+				
Tube 6		+		+		
Tube 7		+			+	
Tube 8		+		+	+	
Tube 9		+		+	+	+
Tube 10			+			
Tube 11			+	+		
Tube 12			+		+	
Tube 13			+	+	+	
Tube 14			+	+	+	+

Fresh RBC = fresh enriched RBC obtained from blood as control; Parasite (*P. vivax/ P. falciparum)* = untreated parasitized RBCs; Anti – *Pv*IspG= Lab raised *Pv*IspG antibodies (1° Ab) from diluted serum (1:500 – 1: 2000); AlexaF 488 = Goat anti- mouse IgG conjugated with Alexa Fluor 488 as secondary antibody; Anti-*Pv*BiP AF647 = anti – *Pv*BiP antibodies conjugated with Alexa Fluor (AF) 647 used to counter stain *P. vivax* BiP protein in parasitized RBCs as control. Tube 1 – 4 = Fresh uRBC used as control; Tubes 5 – 9 = *P. vivax* iRBCs used for detection; Tubes 10 – 14 = *P. falciparum* iRBC controls.

The fresh uRBCs were also mixed with pure mouse serum and with above antibodies to rule out any possibility of non-specific binding of test and other standard antibodies. Further the uRBC and iRBC were treated separately with anti-*Pv*IspG (primary Abs), and counterstained with goat anti-mouse IgG Alexa Fluor 488 (Secondary Abs). Signals

were obtained from iRBCs only **(Figure 5.25b)** indicating that the anti-*Pv*IspG antibodies raised in our lab are recognizing *P. vivax* iRBCs while there was no binding of these to the control i.e. uRBCs **(Figure 5.25a).** Different dilutions of anti-*Pv*IspG antibodies were tried (1:100, 1:500, 1:1000 and 1:2000) and specific binding was observed in all. The anti – *Pv*IspG antibodies show reactivity with *Pv*iRBC even at high dilutions (1:2000). However, since the dilutions were prepared from the unpurified serum sample, there are chances that due to the polyclonal nature of antibodies, such high reactivity was visible.

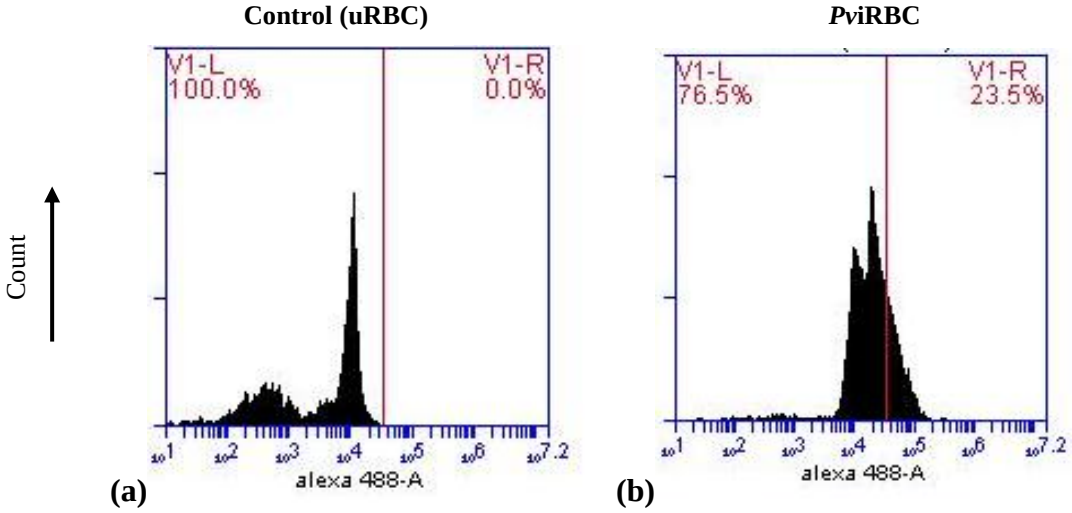


Figure 5.25: Fluorescence observed for (a) uRBC b) *Pv*iRBC stained with anti-*Pv*IspG + Goat anti-mouse IgG AlexaFlour 488 (tube 4and tube 8 respectively))

To determine that the signal for anti-*Pv*IspG was due to binding with *P. vivax* parasite only and to verify the cross reactivity of our lab anti-*Pv*IspG antibodies with *P. falciparum*, we used *P. falciparum* iRBC pelleted from *in-vitro* cultures with parasitemia of at least 2% as another control. The results **(Figure 5.26a)** indicated that the anti-*Pv*IspG antibodies raised in our lab are specific to *P. vivax*, and they do not show any binding to *P. falciparum*, which is evident from the increased fluorescence of antibody binding in *P. vivax* **(Figure 5.26b)** compared to the negligible binding to *P. falciparum*. Finally to crosscheck that the flouroscence is obtained only from *P. vivax* iRBC, we counterstained both *Pf*iRBC and *Pv*iRBC with anti-*Pv*BiP 647 along with our anti *Pv*IspG and AF488. The *P. vivax* iRBC stained with both AF488 and AF647 **(Figure 5.27b)** show more fluorescence as compared to *P. falciparum* iRBC **(Figure 5.27a)**. The fluorescence was observed with AF 647 for anti-*Pv*BiP as well in almost equal

proportion as with AF 488 for anti-*Pv*IspG (**Figure 5.26b, 5.27b & 5.28**). These results indicate that anti – *Pv*IspG antibodies raised in our lab are highly specific to *P. vivax*.

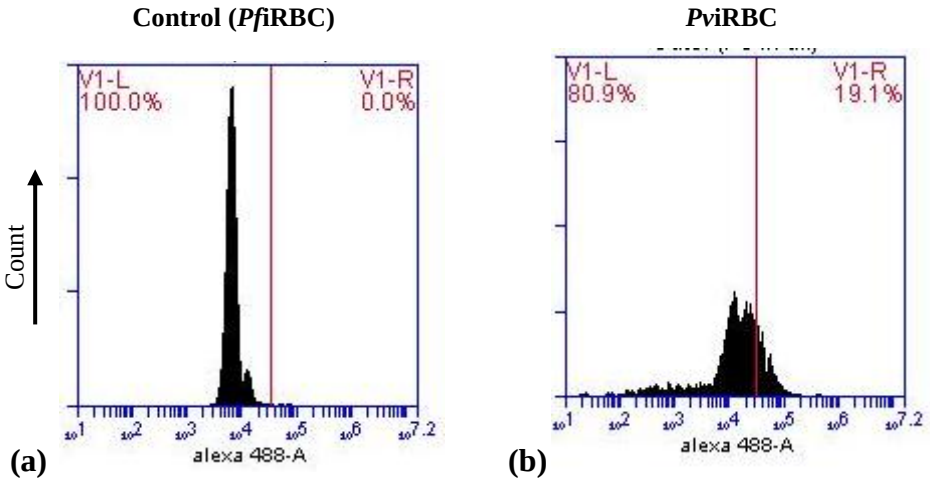


Figure 5.26: Fluorescence observed for (a) *Pf*iRBC (b) *Pv*iRBC stained with anti-*Pv*IspG + Goat anti-mouse IgG AlexaFlour 488 (tube 13 and tube 8 respectively)

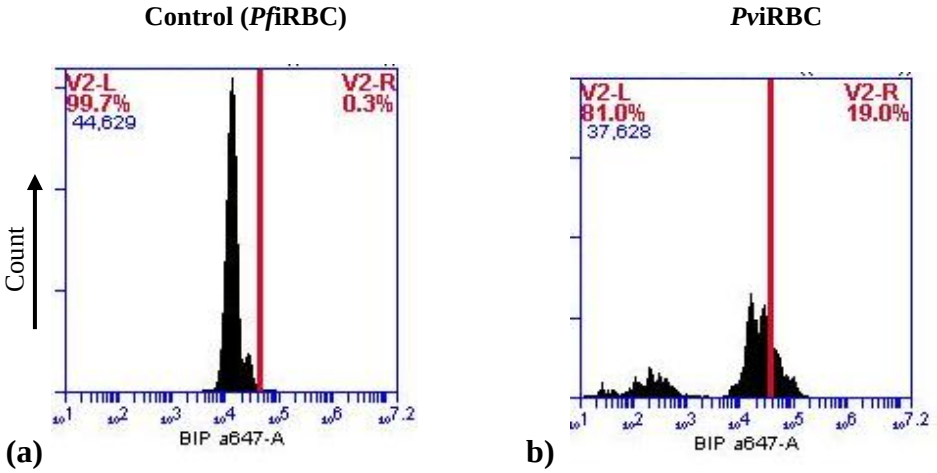


Figure 5.27: Fluorescence observed for (a) *Pf*iRBC (b) *Pv*iRBC stained with anti-*Pv*IspG + Goat anti-mouse IgG AlexaFlur 488 + anti-*Pv*BiP AF 647 (tube 14 and tube 9 respectively)

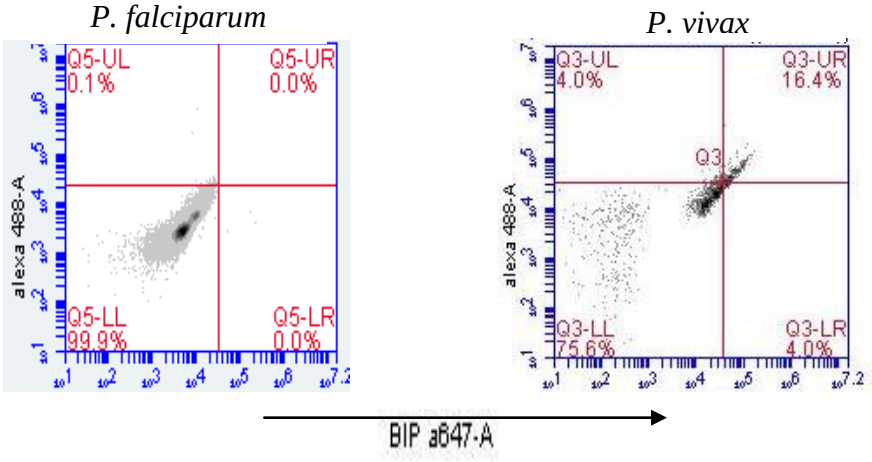


Figure 5.28: **Comparative plot showing fluorescence to *P. falciparum* and *P. vivax* iRBC**

5.2.12 Real Time PCR analysis

Level of isoprenoids synthesis varies in different blood stages of the parasite which attributes to the differential expression of various enzymes involved in this pathway (Bozdech et al., 2008). IspG which acts in the penultimate step of the pathway has been reported to be important for the functionality of the isoprenoids biosynthesis pathway and its deletion was found lethal in *E. coli*. Thus, we performed qPCR to analyse the transcript level of the IspG gene so as to analyze its expression in different field isolates of *P. vivax*, specific to the study region Bikaner. For this, RNA was isolated from parasite infected blood collected from clinical isolates having different erythrocytic stages. This RNA was used for cDNA preparation as described before (Chapter 2). Absolute expression of *Pv*IspD gene (primers sequences used are given in **Table 5.7**) was calculated using ΔCT method (**Table 5.8**) by using Seryl tRNA Synthetase as a normalizing factor (**Figure 5.29)**.

Table 5.7: Primers for quantitative PCR of *ispG* gene from *P. vivax* genome

Lab Id	Primer sequence	Orientation
VS57	5’ GGT TTT CTC CAT GAA GGC TTC CAAC 3’	Forward primer
VS60	5’ TTG CCC ATC CTT TCA TAT TGC CTC G 3’	Reverse primer
*Pv*STSF	5’ CAT CTC CTG AAC GGC ACC AT 3’	Forward primer
*Pv*STSR	5’ GGT GAA GGG AAT AAA CTC GAC G 3’	Reverse primer

Table 5.8: Parasites samples with their erythrocytic stages and ΔCT values

S. No.	Field Isolates	Stages as per slide (approx.)	ΔCT values
1	A	80% trophozoite, 20% schizont	7.63
2	B	70% rings, 30% trophozoites	2.77
3	C	20% rings, 50% trophozoites, 30% schizont	3.88
4	D	80% Schizont, 20% ring	0.14
5	E	70% Schizont, 30% ring	0.58
6	F	90% trophozoites, 10% schizonts	7.72

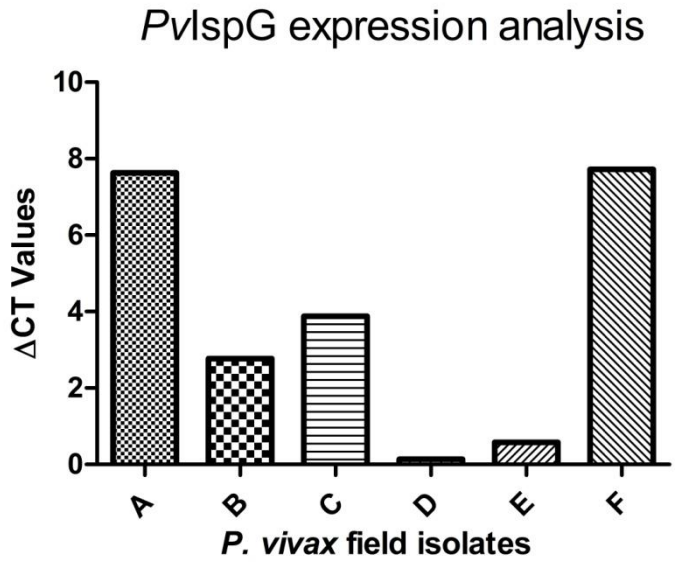


Figure 5.29: Quantitative PCR analyses for *Pv*IspG protein (X axis; *P. vivax* field isolates with different erythrocytic stages; Y axis: ΔCT value for *Pv*IspG in comparison to housekeeping gene Seryl-tRNA synthetase).

The qPCR data obtained showed similar results as obtained for *Pv*IspD. Expression of *Pv*IspG gene was observed in all the samples that have shown the mixed erythrocytic stages as per the PBF examination **(Table 5.7).** Similar to IspD, the expression was found higher in sample (A, F) than B and C, while there was just basal level expression in D and E. Overall after combining the qPCR and slide data, our results indicate that the expression of *Pv*IspG gene is maximum at the trophozoite stages and lowest at the schizont and early ring stages. Study by Bozdech et al., 2008 have shown that the expression of IspG gene involved in the isoprenoids biogenesis pathway peaks at trophozoite stages in all the three isolates (smru 1, 2 and 3) and low level expression is seen at early ring or late schizont stages. Our study agrees with Bozdech data that this gene is expressed maximum at trophozoite stages and expressed minimum at schizont stages. However, adding more no of samples with varying percentage of blood stages might help to reveals the requirement of IspG gene in stage specific manner.

5.3 Conclusion

4-hydroxy-3-methyl-but-2-en-1-yl diphosphate synthase (IspG) is characterized by the GcpE domain and participates in the concluding steps of isoprenoids biosynthesis pathway. It is believed to function in the reduced conditions and is highly dependent on

[4Fe-4S] cluster (which acts as prosthetic group for this enzyme), and cofactors like flavodoxin, flavodoxin reductase, and NADPH (Seemann et al., 2002) for its activity. Although the comparative sequence analysis of the Indian *Pv*IspG (BankIt ID: 1761743) and PVX_111575 sequences with other *Plasmodium*, apicomplexans and bacterial orthologues show similarity of only 30-84%, a high level of similarity was observed in conserved domain and signature motifs suggesting that the function has been retained during the course of evolution. Between *P. falciparum* and *P. vivax* IspG, an identity of approx. 63% was observed. The *P. vivax* IspG shows a high similarity with the IspG protein of other primate malaria parasite like *P. knowlesi, P. cynomolgi* and *P. inui*. The presence of three conserved Cysteine residues suggest that they can coordinate with a [4Fe-4S] cluster and the unique iron atom present in this cluster could function in binding of MECP substrate during the reaction cycle.

*Pv*IspG protein was successfully expressed in prokaryotic expression system but as the protein was purified in presence of oxygen which can inhibit the catalytic activity by removing Fe-S cluster from *Pv*IspG protein, these clusters were reassembled on the protein by reconstituting the protein in the presence of excess iron and sulphur under exclusion of molecular oxygen. Formation of Fe-S clusters was confirmed by spectra analysis, which was further confirmed by Ferrozine and methylene blue assay. The bound iron was found to be present in +3 oxidation state as suggested by electro-paramagnetic resonance studies.

Further to elucidate the expression profile of *Pv*IspG in *P. vivax* parasite, co-localization of the protein was performed with anti-*Pv*IspG antibodies raised against recombinant *Pv*IspG protein. Localization of the protein to *P. vivax* apicoplast was confirmed in thin blood smears made from malaria parasite infected patient thus, proving its functionality in the isoprenoids biosynthesis pathway. The immune-flouroscence data was well supported with our flow cytometry data where the anti-*Pv*IspG antibodies show high specificity towards *P. vivax*. Real time PCR analysis was performed to check the transcript level of *Pv*IspG protein in different field isolates, which suggested that the protein is transcriptionally active throughout the erythrocytic stages. However, there is a need of a synchronized *in-vitro* culture of *P. vivax* to specifically elucidate the detailed time line-profiling of the activity and stage-specificity of the enzyme in the parasite.

Chapter VI

In-silico studies of Isoprenoids biosynthesis pathway enzyme IspG from *P. vivax* and *P. falciparum*

This study focuses on the *in-silico* analysis of IspG enzyme from *Plasmodium* parasites and its inhibition by different antimicrobial agents. In the previous chapter, the molecular and functional characterization of this enzyme is detailed. The objective of this *in-silico* study is to establish the model of the enzyme and further elucidate the substrate and the drug binding sites.

In this study, initially *Pv*IspG and *Pf*IspG enzymes structures were generated by homology modelling and probable substrate binding cavities and active site residues were identified. Based on amino acids present in active sites, docking of respective substrates and drug molecules were performed which were validated with Molecular Dynamics. This study confirms the presence of conserved residues in IspG enzyme from different apicomplexans and their interaction with different drug molecules that present it as a probable drug target for future drug designing.

6.1 Introduction

In the previous chapter, we have detailed the IspG protein from a major apicomplexan parasite, *Plasmodium*. This is the first study on this protein from any *Plasmodium* species. We have found that in *P. vivax*, the protein is transcriptionally active in the major erythrocytic stages of parasite and localizes to the apicoplast, the functional site of Isoprenoids biosynthesis pathway. The conservation of the functional domains across the *Plasmodium* species, apicomplexans and major prokaryotes, makes it an important putative drug target. As different drugs like Fosmidomycin inhibit the initial steps by targeting the first committed enzyme of the pathway IspC, targeting an enzyme like IspG involved in the penultimate step of the pathway will result in cumulative effect and thus may result in retarded parasite growth.

To initiate the process of shortlisting of the molecules that can inhibit this enzyme and thus may act as an inhibitor of Isoprenoids biosynthesis pathway, first the structures of the *Pv*IspG and *Pf*IspG enzymes were generated using homology modelling. The obtained structures after refinement and energy minimizations were analysed for the binding affinity of substrate and finally used for docking various antimicrobial drug molecules shortlisted from the literature based upon their use against IspG in prokaryotes or reported MEP pathway.

6.2 Results & Discussion

6.2.1 Secondary structure prediction

The secondary structure of *Pv*IspG and *Pf*IspG protein was generated using online server PSIPRED (Buchan et al., 2013). The *Pv*IspG protein consists of 22 α-helices and 21 β-sheets joined with the help of 41 coils (**Figure 6.1a**) while the *Pf*IspG protein consists of 26 α-helices and 24 β-strands joined with the help of 50 coils (**Figure 6.1b**). These α-helices and β-sheets are arranged in an alternate manner over the peptide backbone (also known as α/ β protein folds) which indicates that IspG protein belongs to the TIM barrel super family (Lee et al., 2010). A TIM barrel typically consists of the helices and strands (usually 8 of each) forming a solenoid that curves around to close on itself in a doughnut shape, topologically known as a toroid. The parallel β-strands form the inner wall of the doughnut (hence, a β-barrel), whereas the α-helices form the outer wall of the doughnut. The protein's core in a TIM barrel is tightly packed, mostly with bulky hydrophobic amino acid residues with a few glycine residues to allow wiggle room for the highly constrained

center of eight approximate repeats to fit together. In *Pv*IspG, the TIM barrel domain was observed to span the amino acid residues 100 – 385 where the protein core consisted mainly of the hydrophobic amino acid residues (48.1%). 36% of the total residues in the β-strands were branched aliphatic residues namely valine, leucine and isoleucine.

(a) *Pv*IspG

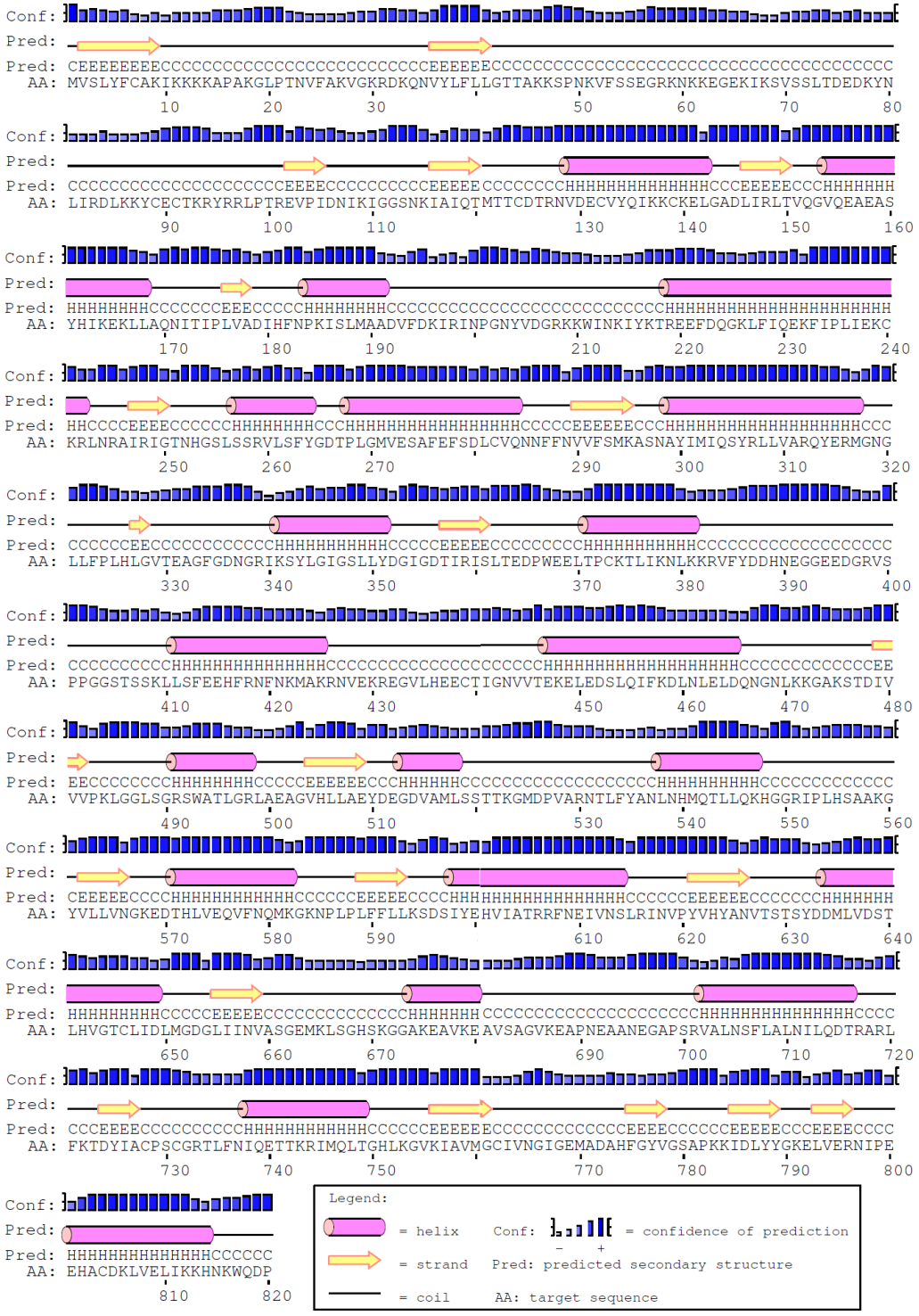


Figure 6.1a: Secondary Structure predictions of (a) *P. vivax* IspG by PSIPRED (α helices are shown as pink colour, β-sheets in yellow colour and coil as black line).

Similarly, in *P. falciparum* the TIM barrel domain spanning the amino acid residues 121 – 401 was rich in hydrophobic residues (52.3%) with 37% of the total residues in the β-strands as branched aliphatic residues namely valine, leucine and isoleucine.

*Pf*IspG

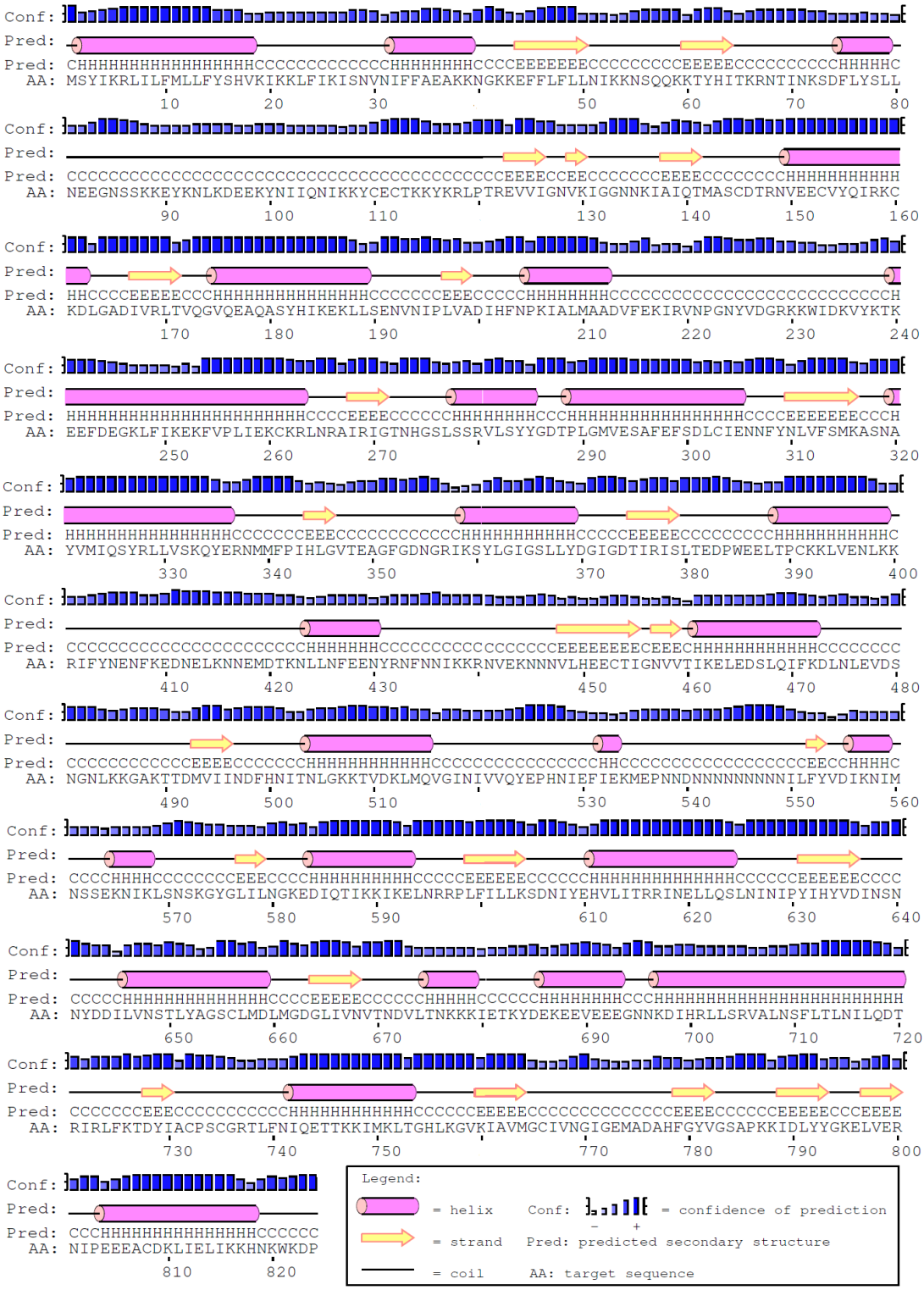


Figure 6.1b: Secondary Structure predictions of *P. falciparum* IspG protein by PSIPRED (α helices are shown with pink colour, β-sheets in yellow colour and coil as black line).

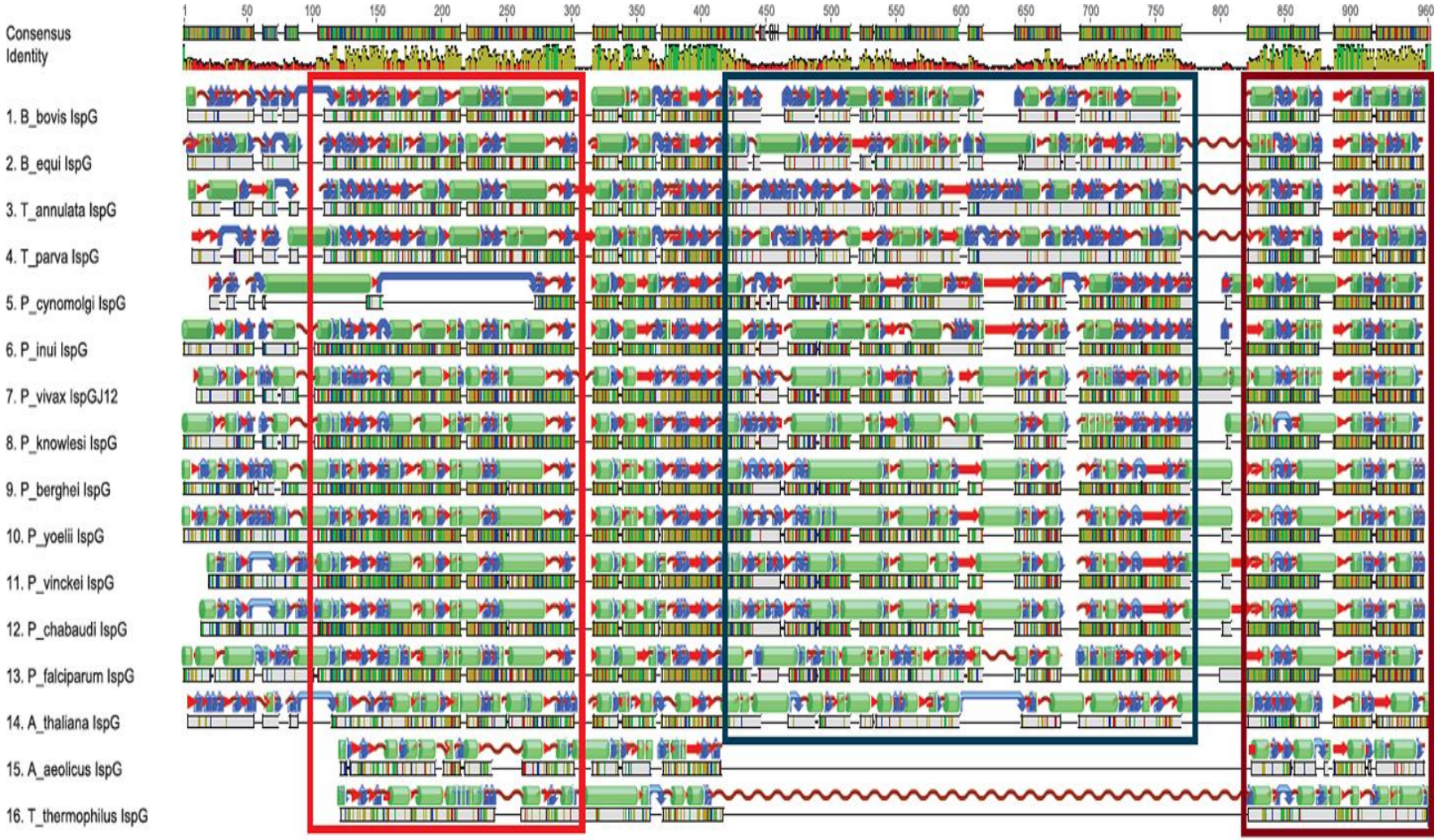


Figure 6.2: Comparison of IspG protein secondary structure of *P. vivax* with other organisms [Consensus region and identity between IspG amino acid sequences obtained from different organisms. The secondary structures visible include green cylinders as alpha helix, red arrows as beta sheets and blue arrows as turns in beta strands. The red block denotes the conserved N-terminal TIM barrel domain; brown block denotes the C-terminal conserved region for Fe-S cluster binding and black block denotes the extra domain present in apicomplexans and plants but absent in prokaryotes. Figure generated with the help of GeneiousII (www.geneious.com).

A comparative analysis between the obtained secondary structures of *Pv*IspG and *Pf*IspG proteins with their homologues/ orthologues from *Plasmodium* and other apicomplexans, plants and prokaryotes show a high level of conservation (**Figure 6.2**). Interestingly, the secondary structure of apicomplexans IspG protein showed similarity with that from prokaryotes and plant protein. This indicates conservation of functional aspects of IspG protein across different organisms. CDD analysis of these proteins (**Figure 5.9 & 5.10**) show presence of two conserved domains at N and C-terminals that are designated domains A and B respectively, which show high level of similarity with those reported from prokaryotes. The N-terminal region (A domain) forms a TIM barrel like structure having closest structural homology to dihydropteroate synthase. It consists of ten β-sheets and nine α-helices involved in binding the diphosphate group of MEcPP as well as in providing H^+ ions. The C-terminal B domain is a [4Fe-4S] cluster domain involved in reduction and consists of a fold that is similar to that seen in sulfite reductase and the ferredoxin domains of nitrite reductase.

While only these two domains are present in prokaryotes, in *Pv*IspG protein sequence, an additional region is present next to "A" domain. This third domain of IspG protein was designated as A* from plants like *A. thaliana*, many apicomplexans like *P. falciparum, B. bovis, T. parva* and certain bacteria like *Chlamydia trachomatis* and *Leptospira interrogans* (**Figure 6.3**) (Liu et al., 2012). Computational studies suggest that this extra domain is similar to "A" domain and forms a structure similar to TIM barrel (Wang and Oldfield, 2014). However, there are few conserved residues in the A* domain between plant, malarial, and bacterial 3-domain IspGs (**Figure 6.4**); while in the 2-domain proteins, there are large number of conserved residues that are essential for catalytic activity. Thus, this extra domain is hypothesized to have only the structural aspect rather than a functional role.

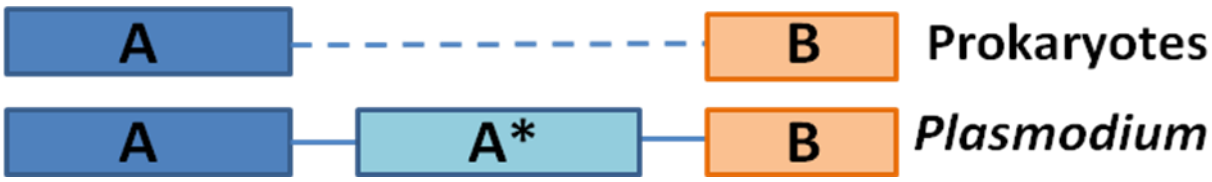


Figure 6.3: Schematic representation of IspG protein in prokaryotes and *Plasmodium*.

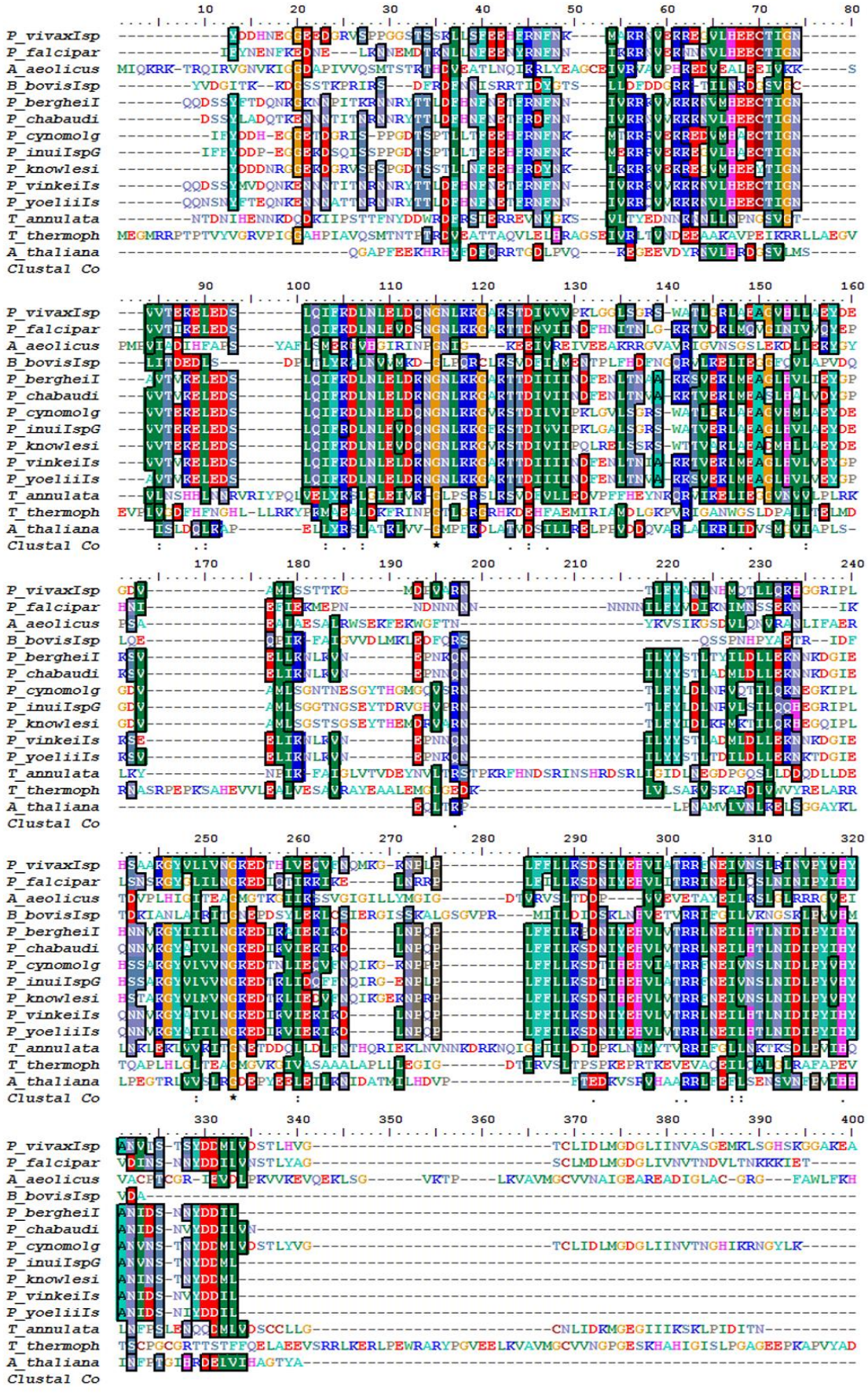


Figure 6.4: Multiple sequence alignment of A* domain from different Apicomplexans including *Plasmodium* and plants

Further, using PSIPRED other details related to the characteristic features of *Pv*IspG and *Pf*IspG protein sequences were identified (**Table 6.1).**

Table 6.1: Comparative Biochemical analysis of *Pv*IspG and *Pf*IspG proteins

Characteristic features	***Pv*IspG Value**	***Pf*IspG Value**
Molar extinction coefficient	57320.00	59320.00
Molecular weight	91610.40	95225.93
Isoelectric point	9.25	8.91
Percent negative residues	11.83	12.83
Percent positive residues	14.15	13.15
Charge	29.16	29.16
Signal Peptide	1-13	1-22
Transmembrane topology	639-654	642-658
Aliphatic index (0-100)	90.23	89.37
Hydrophobicity (-4 -3)	-0.32	-0.42

6.2.2 Comparative modelling and Molecular simulation of *Pv*IspG and *Pf*IspG

To generate the tertiary structures of *Pv*IspG and *Pf*IspG proteins, we used comparative homology modelling approach. For this, the nucleotide sequences of *ispG* genes obtained from *P. vivax* and *P. falciparum* Indian field isolates (BankIt ID: 1761743 and BankIt ID: KX176850) were translated and submitted to HHpred online server for template search. Crystal structure of thermophile *Aquifex aeolicus* (Protein Data Bank ID: 3NOY; A chain) with resolution of 2.7Å was shortlisted as a template. The A chain of *A. aeolicus* IspG (3NOY) protein is about 357 amino acids long when compared to 820 amino acid sequence of *Pv*IspG and 824 amino acid long sequence of *Pf*IspG and shared 36% amino acid identity to the target *Pv*IspG sequence and 35% match to *Pf*IspG, mainly matching with the conserved GcpE domain.

The non-similar regions in IspG sequences were removed manually which consisted mainly of the signal and transit peptide and the insert A* domain. After this, we submitted

these processed sequences to MODELLER 9v11, which generated 3D models of the targets with all non-hydrogen atoms without any user intervention.

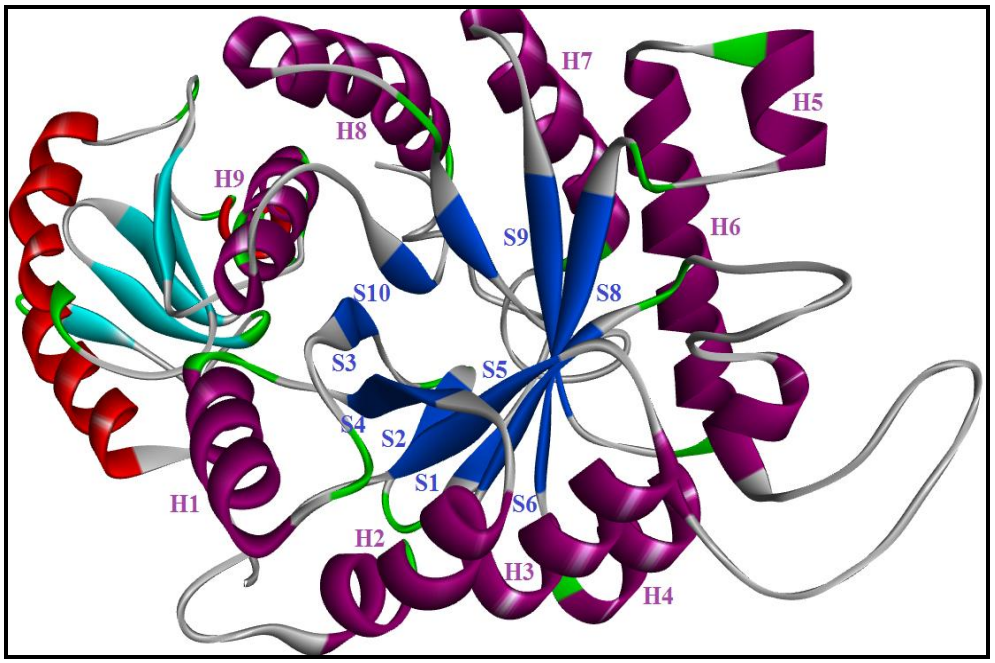


Figure 6.5 (a): *Pv*IspG protein structure predictions using *A. aeolicus* IspG (3NOY) crystal structure as template

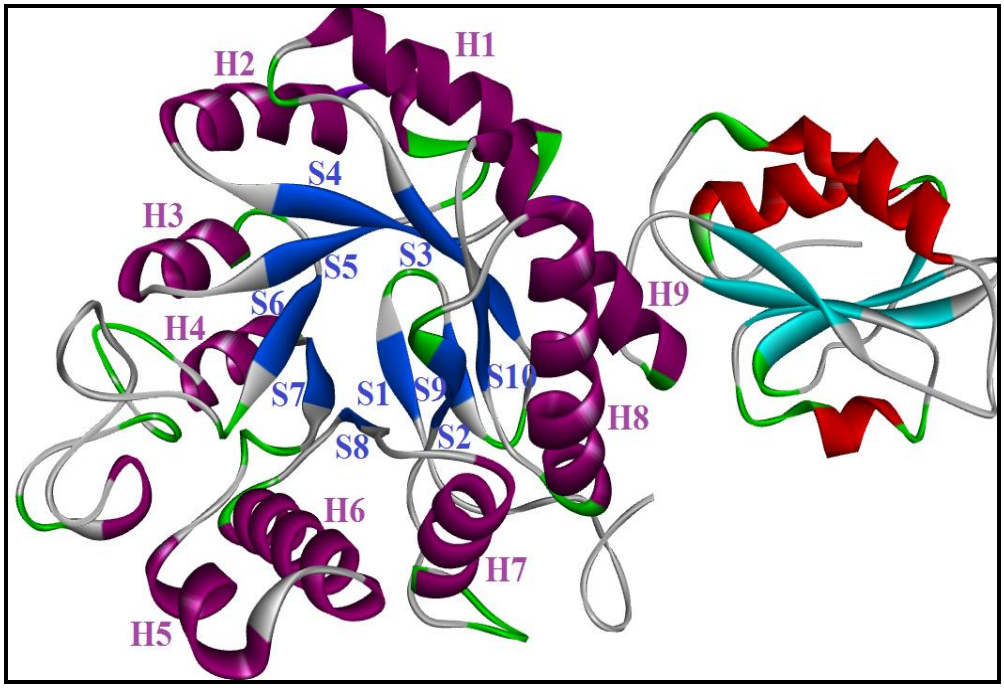
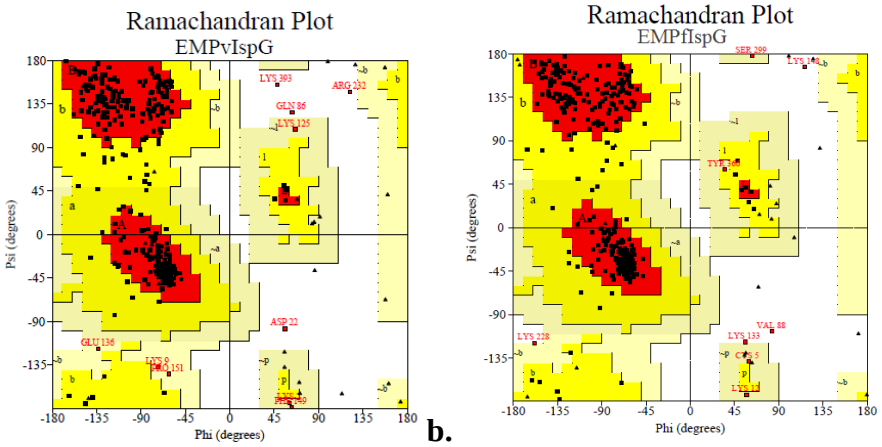


Figure 6.5 (b): *Pf*IspG protein structure predictions (*A. aeolicus* IspG (3NOY) crystal structure was used as a template for structure prediction)].

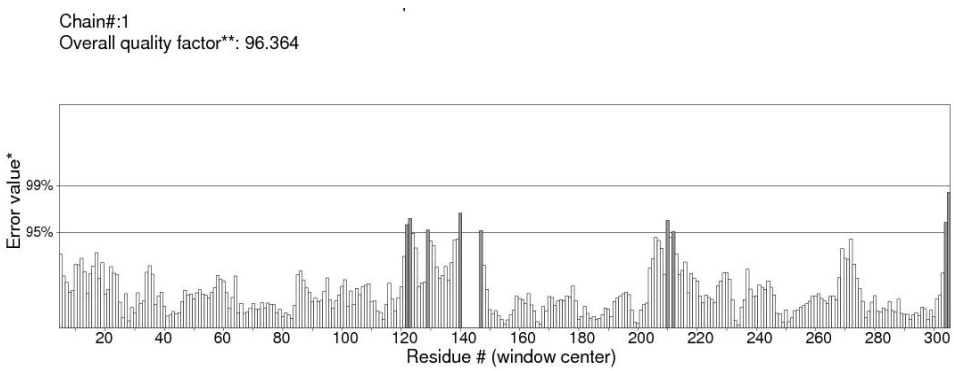
Further, we refined these models by minimizing violations of many distances and dihedral angle restraints extracted from the template structures. After selecting the best models of

*Pv*IspG and *Pf*IspG proteins with minimum DOPE score, we further refined the structure by energy minimization and molecular simulation using GROMACS 4.6. Refined structures (**Figure 6.5a & 6.5b**) were evaluated using PROCHECK which indicated G score 'BETTER'. The evaluation with VERIFY3D gave a score of 0.66 (range of 0.0 to 1.0) for *Pv*IspG structure and 0.67 for *Pf*IspG. Ramachandran plot analysis with RAMPAGE shows main chain conformations for 97.5% of amino acid residues of *Pv*IspG and 99% for *Pf*IspG within the favoured or allowed regions of the Ramachandran plot (**Figure 6.6a & 6.6b**). Quality evaluation of the model for the environment profile was predicted using ERRAT structure evaluation server where a score of 96.36 (max. 100) for *Pv*IspG and a score of 97.15 for *Pf*IspG indicated that the molecular geometry of the models is of good quality (**Figure 6.7a & 6.7b**). Identification of bad packaging of side chain atoms or unusual residue contact was checked in the refined models using WHAT IF program where Z-score calculated by the server confirmed the good quality of the structure. Root-mean-square deviation (RMSD) value of 0.972 Å for *Pv*IspG and 0.824 Å for *Pf*IspG between the backbone atoms of the template and models indicates significant homology between the structures.

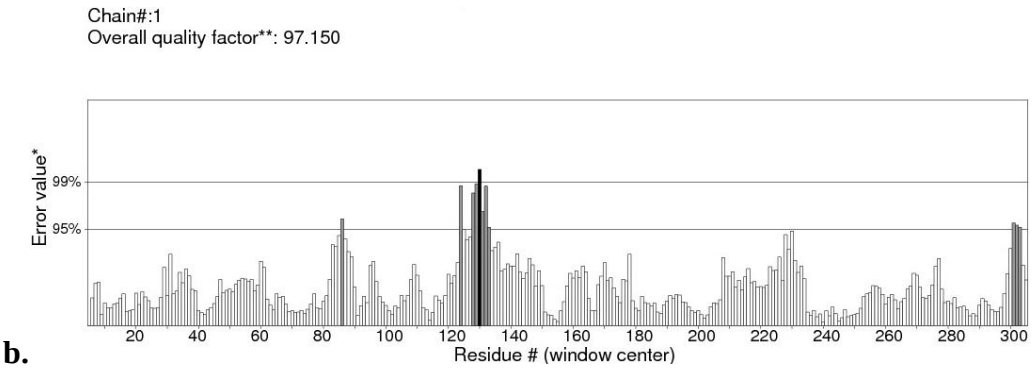


a. **b.**

Figure 6.6: Ramachandran Plot assessments of (a) EM*Pv*IspG and (b) EM*Pf*IspG (Amino acids away from favoured region labelled with position).



a.



b.

Figure 6.7: ERRAT validations of three-dimensional structures of (a) *Pv*IspG and (b) *Pf*IspG proteins. The error value for most of the residues is less except for a few which did not had any match in the template structure. These residues are involved in forming loops.

The *Pv*IspG protein is 820 amino acid long wherein the N terminal domain spans from 85-381 and the C terminal domain spans from 723-820. A linker consisting of amino acids at positions 544-549 connects these domains. The N - terminal domain consists of conserved residues (R148, R198, R248 and R359), Q119, D179, S293 and H326 located in the β sheets of TIM barrel and M121, N200, N252, K295 T330, E331 and S361 located in the loop connecting the β strand to the helix. The C-terminal domain consists of three α-helices and five β-sheets along with three conserved Cysteine residues (C728, C731 and C762) present in the form of a loop and E769 present in the α-helix to interact with the Fe-S clusters. The Fe, which interacts with the E769 residue, also interacts with MECP substrate during the transition from open state to the closed conformation. This glutamate is important for the activity of enzyme, as shown in *A. aeolicus* (Lee et al., 2010).

Similar to *Pv*IspG, the *Pf*IspG protein sequence (824 amino acids long) also show a conserved GcpE domain. The N-terminal domain spanning from 121-401 amino acid residues forms a TIM barrel with alternate organisation of α-helices and β-sheets. The conserved arginine residues (R169, R218, R267 and R376), Q140, D200, S314 and H343 are located in the β-sheets of TIM barrel while M142, N220, N272, K316 T348, E349 and S379 are located in the loop connecting β sheets to α-helix. Similarly, the C-terminal domain spanning 718-818 amino acid residues consisted of three α-helices and five β-sheets along with conserved Cysteine residues (C732, C735 and C766) present in the form of a loop between them and E773 present in the α-helix.

In *A. aeolicus*, IspG acts as a dimer and hydrophobic conserved residues present in α-helices H8 and H9 of N-terminal domain initiate this dimerization (Lee et al., 2010). In a

dimer, the 'B' domain [4Fe-4S] cluster present at C terminus of first molecule interacts with the open mouth of TIM barrel of 'A' domain of the second molecule in a head to tail fashion. This conformation is responsible for trapping MEcPP substrate between the two domains, and allows it to interact with both domains at the same time (Lee et al., 2010; Rekiittke et al., 2011). Although similar hydrophobic conserved residues are present in the three- domain *Pv*IspG, however, its existence as a monomer or dimer is still unclear. It is proposed that the plant and other three-domain IspGs may have originated via an initial gene fusion to form a bacterial homo-dimer, followed by gene duplication, exon loss, and recombination, eliminating one of the [4Fe-4S] clusters to form the (AA*B) proteins.

6.2.3 *Pv*IspG and *Pf*IspG activity and Interaction with substrate

After identifying the binding cavities in the obtained three-dimensional models of *Pv*IspG enzyme, docking of MECP substrate to these identified cavities was performed (**Table 6.2**). Based on visible interactive H-bonds, electrostatic and steric interaction, the affinity of MECP substrate and [4Fe-4S] clusters towards these binding sites was calculated.

Table 6.2: *Pv*IspG active site prediction

Rank	C-score	Cluster size	PDB Hit	Ligand Name	*Ligand Binding Site Residues*
1	0.29	12	4g9pA	MECP	121,148,179,181,198,248,252,295, 330,331,359,361
2	0.28	13	4s23A	[4Fe-4S]	728,729,730,731,733,734,760,761,7 62,765
3	0.03	2	2y0fC	SF4	200,201,202,250,251,274,275,278

*(The C - score is the confidence score of the prediction ranging from [0-1], where a higher score indicates more reliable prediction; **Cluster size** is the total number of templates in a cluster)

The MECP substrate show binding affinity with four conserved arginine residues (R148, R198, R248 and R359), S293, K295 and T330 present towards the N-terminal of the protein (**Figure 6.8**), while the [4Fe-4S] cluster show the binding affinity towards the three conserved cysteine residues (C728, C731, C762) and Glu769 present at the C-terminal of the *Pv*IspG protein **(Figure 6.9)**. In addition to the H- bonds present between the substrate and the enzyme, steric interactions were also observed between [4Fe-4S] cluster and conserved residues because of its atomic structure. Assessment of the *Pv*IspG structure elucidated that in absence of substrate, the protein resides in an open conformation but once the binding of MECP substrate takes place at the N-terminal,

flexible amino acids present in the linker region (residue 544-549) help the movement of C-terminal domain having [4Fe-4S] towards the N-terminal, forming a cap like appearance. This movement allows the direct contact between the (Fe-S) cluster and the bound substrate and thus inhibits the release of MECP substrate prior to enzyme activity. The reaction takes place inside the closed cavity, where cyclic MECP molecule converts to aliphatic HMBDP and allows the release of the reaction product.

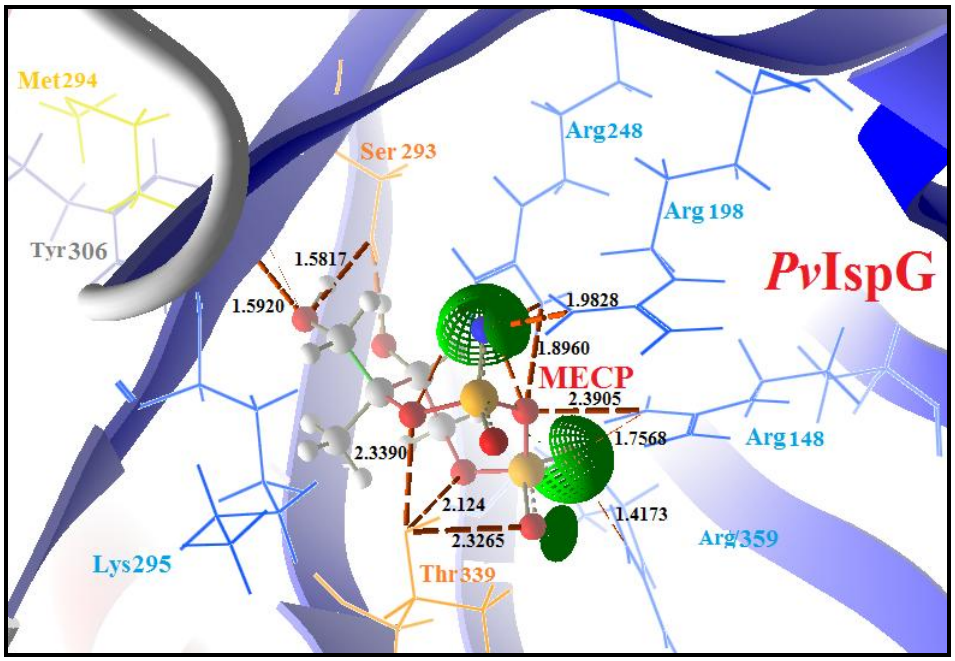


Figure 6.8: Active site prediction and docking analysis of MECP to the *Pv*IspG protein (The interaction of MECP substrate (in green) with *Pv*IspG enzyme is shown by red dotted lines.

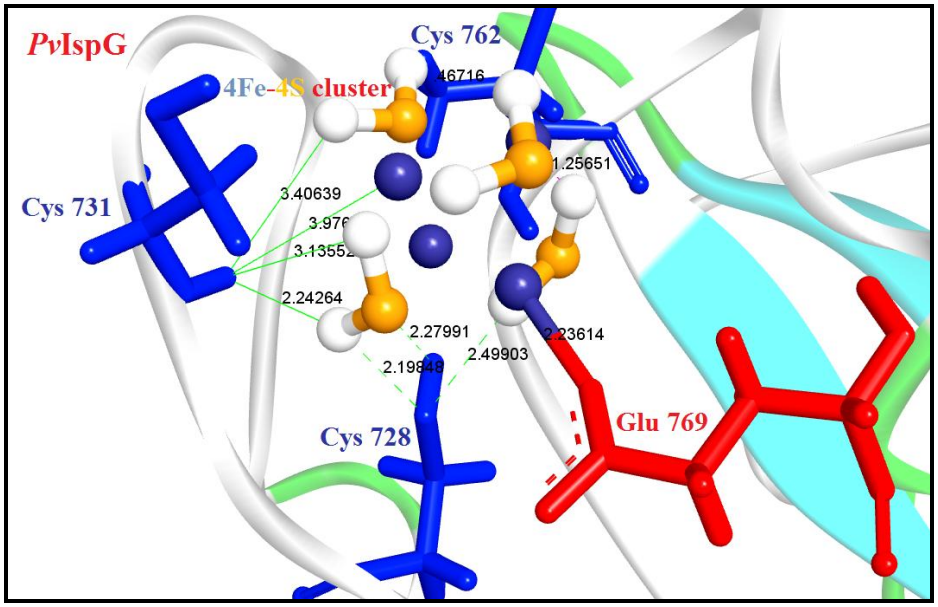


Figure 6.9: Active site prediction and docking analysis of [4Fe-4S] cluster to *Pv*IspG protein **(**Interaction of iron (in blue) and sulfur (in yellow) atoms with *Pv*IspG shown with green lines)

In *P. falciparum* also, the binding cavity was identified **(Table 6.3)**, where the residues were found conserved. The [4Fe-4S] cluster binding coordinated at Cysteine residues 732, 735, 766 and E773 (**Figure 6.10**) and MECP substrate shows affinity towards amino acid residues R169, 219 and 377; S314, K316; T348 and H344 (**Figure 6.11**).

Table 6.3 *Pf*IspG active site prediction

Rank	C-score	Cluster size	PDB Hit	Ligand Name	Ligand Binding Site Residues
1	0.20	11	4s23A	[4Fe-4S]	732, 733, 734, 735, 764, 765, 766
2	0.18	8	4s23A	MECP	142, 169, 200, 218, 220, 267, 272, 314, 316, 343, 348, 349, 376, 379
3	0.05	4	4s39A	H6P	142, 148, 179, 181, 198, 220, 248, 252, 272, 295
4	0.03	2	4iibA	NAG	242, 243, 244, 245, 287, 288, 289

(The C - score is the confidence score of the prediction. C-score ranges [0-1], where a higher score indicates a more reliable prediction; **Cluster size** is the total number of templates in a cluster**)**

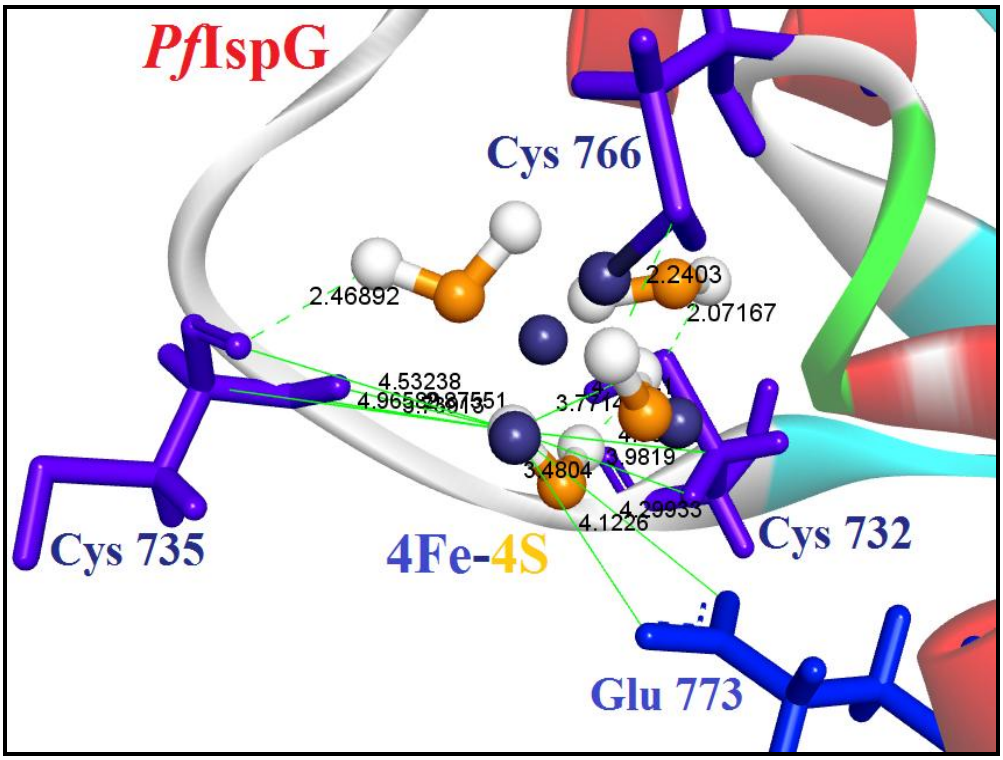


Figure 6.10: *Pf*IspG protein structure prediction and 4Fe-4S cluster docking analysis ([4Fe-4S] cluster is shown in blue and yellow colours. Amino acid residues with distinguished colour with their respective position in the structure shown by green lines).

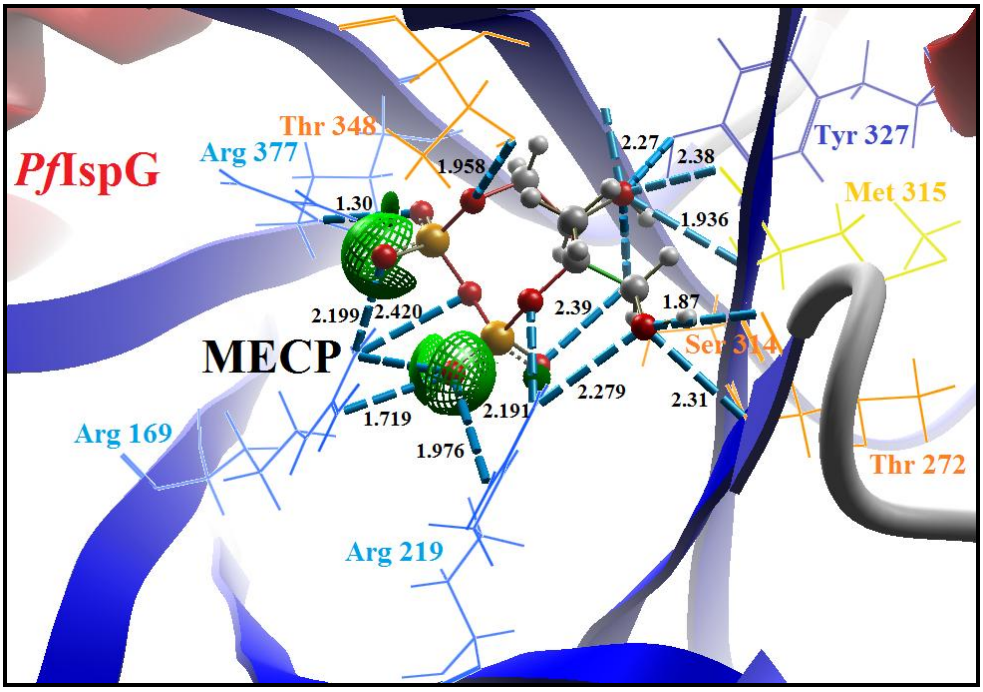


Figure 6.11: *Pf*IspG protein structure prediction and MECP docking analysis (MECP substrate shown in red, interaction shown with a blue dotted line).

These IspG substrate and Fe-S cluster interaction when compared with that reported in *A. aeolicus* showed the interaction of similar conserved residues suggesting that this substrate and cluster reaction mechanism with the enzyme is conserved (Lee et al., 2010). When the conserved residues were mutated *in-silico*, a decrease in the binding affinity of substrate and cluster was observed, similar to study reported by Lee et al. (2010).

6.2.4 *Pv*IspG and *Pf*IspG Interaction with drug molecules

The analysed active sites and binding cavities in the models of *Pv*IspG and *Pf*IspG enzymes were also used for the docking of three drug molecules: Atorvastatin, Fosmidomycin and Alkyne diphosphate to check the *in-silico* inhibition of IspG.

6.2.4.1 Interaction with Atorvastatin

Atorvastatin belongs to the statins class of drugs and administered for lowering bad cholesterol in the patients' suffering from cardiovascular disorders (Rosenson, 2001). It acts upon the HMG-CoA reductase enzyme participating in the initial steps of MVA pathway (the eukaryotic isoprenoid biosynthesis pathway) and decrease the synthesis of cholesterol by decreasing the basic isoprene units required for synthesis of cholesterol. It also acts against *T. gondii* where it acts upon the Farnesyl Diphosphate/ Geranyl Geranyl diphosphate synthase functional at the downstream of MEP pathway. However, this drug

has not been tested against any other enzyme of the pathway, especially in *Plasmodium*. We performed docking analysis of this drug against all isoprenoid pathway enzyme and we could see interaction possibility of this drug against IspG at the TIM barrel domain.

Structure of Atorvastatin ($C_{33}H_{35}FN_2O_5$) drug molecule was retrieved from PubChem Open Chemistry database. After refinement of Atorvastain by Discovery studio (Dassault Systèmes, BIOVIA), it was converted to PDB format to achieve flexibility in structural bonds to get a better picture of its binding affinity towards the *Pv*IspG protein. Finally, the protein molecule without any HETATM and refined drug molecule were uploaded in the Molegro Virtual Docker. Binding cavities based on the integrity of the protein were identified and these sites were selected for docking.

Atorvastatin shows interaction with amino acid residues K115, K374, K720, D145, G143, G754 and L752 of *Pv*IspG protein (**Figure 6.12a**) and amino acid residues I106, Y116, K117, K136, K438, K399, L400, D454 and Y546 of *Pf*IspG protein (**Figure 6.13**). Interaction of atorvastatin and amino acid residues was confirmed by Molecular dynamics studies where binding of ligand and respective coordinates were stabilized with respect to the total energy (Kinetic energy and Potential energy), RMSD, temperature and pressure **(Figure 6.12b)**. Simulation was performed until these parameters reached a constant value and the best protein-ligand complex was chosen for final study.

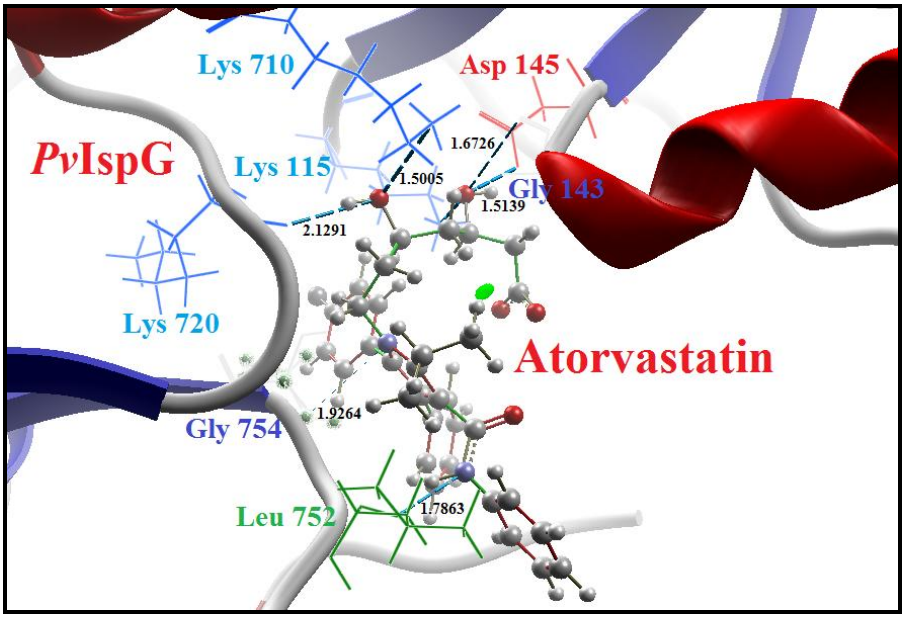
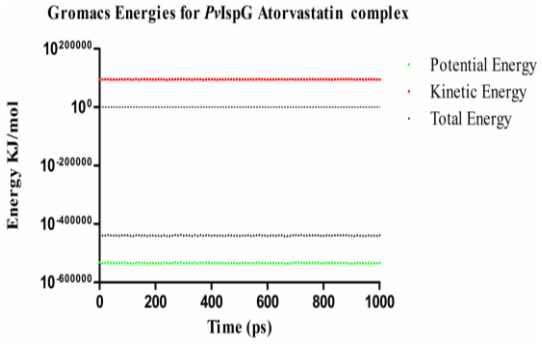
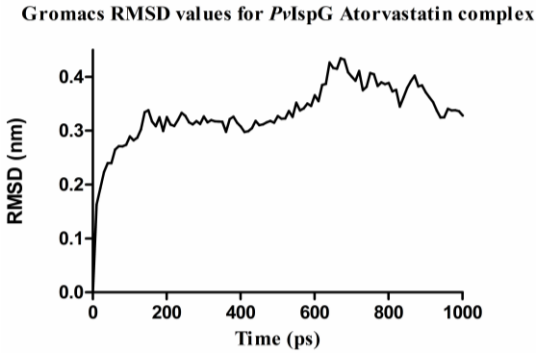


Figure 6.12a: *In-silico* drug inhibition studies for Atorvastatin (Atorvastatin drug was shown in grey colour) against ***Pv*IspG protein** (interactions shown with blue dotted lines)**.**

a.



b.



c.

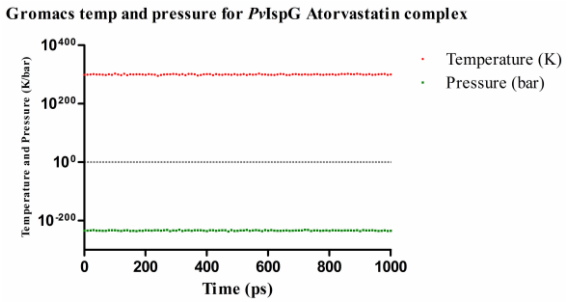


Figure 6.12b: *Pv*IspG-Atorvastatin complex validations by Molecular Dynamics studies. (a) Gromacs energies; (b) RMSD calculation.; (c) Gromacs temp and pressure

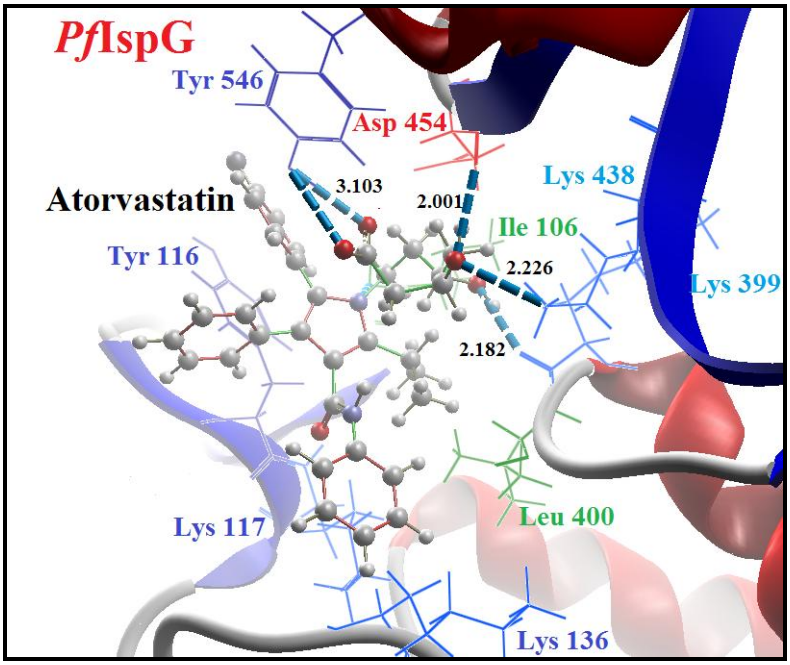


Figure 6.13: *In-silico* drug inhibition study of *Pf*IspG protein with Atorvastatin (Atorvastatin drug is shown in grey color and amino acid residues with their respective colour and interaction shown with blue dotted lines).

Since most of these residues are present in the same binding cavity as MECP substrate with some residues at the C terminal near the binding site of Fe-S clusters, there is a possibility that this drug may affect the binding of MECP or may disturb the formation of closed conformation required for the activity of the enzyme.

6.2.4.2 Interaction with Fosmidomycin

Fosmidomycin (Fos) has been characterized as an inhibitor of IspC, the second enzyme of the MEP pathway, where it shows a competitive binding with DOXP substrate (Jomaa et al., 1999; Steinbacher et al., 2003). Fos was first used by Jomaa et al. (1999) for inhibiting *P. falciparum in-vitro* culture to prove its activity against IspC enzyme. It also decreased the levels of intermediates of ubiquinone and dolichols biosynthesis in blood stages of *P. falciparum* (Cassera et al., 2007), thus affecting pathways outside the apicoplast involved in parasite development, which may result in protein mislocalization and organelle disruption. Following these results, Fos was included in the therapies to treat human malaria. However, due to less absorption, shorter half-life of the drug leading to incidences of recurrence and its toxicity at higher doses, this drug alone could not act as potential antimalarial (Kuemmerle et al., 1987). Thus, Fos is being used in combination with other drugs (Lell et al., 2003) for treating uncomplicated malaria in humans. The answer to low absorption of Fos came from reports detailing the increased permeability of Fos and FR900098 (methyl derivative of Fos) in the iRBCs, due to presence of New Permeability Pathways (NPP) (Baumeister et al., 2011). Fos inhibits the development of apicoplast in hepatocytes as well, ultimately reducing the number of merosomes which shows the importance of MEP pathway in the growth and development of the parasite in hepatic stages also (Nair et al., 2011).

Zhang et al. (2011) had raised the possibility of presence of multiple targets of Fos in the MEP pathway. Keeping this in mind, we checked its inhibitory effect on IspD (Chapter 4) and here with IspG. Fos shows binding affinity with amino acid residue R148, R248, R359, S293, M294, Y306, H326, G328 and T330 of *Pv*IspG protein (**Figure 6.14**) and R169, R219, T272, S314, R377, and T348 of *Pf*IspG protein (**Figure 6.15**). These residues are shared with the MECP substrate, indicating the inhibitory action of this drug on the activity of both *Pv*IspG and *Pf*IspG enzymes. Interaction between *Pv*IspG and *Pf*IspG enzyme with Fos was validated with molecular dynamics studies similarly as performed for atorvastatin. Stable configuration obtained for Fos confirms efficient binding of the drug which indicates its ability to be used as an inhibitor for both *Pv*IspG and *Pf*IspG.

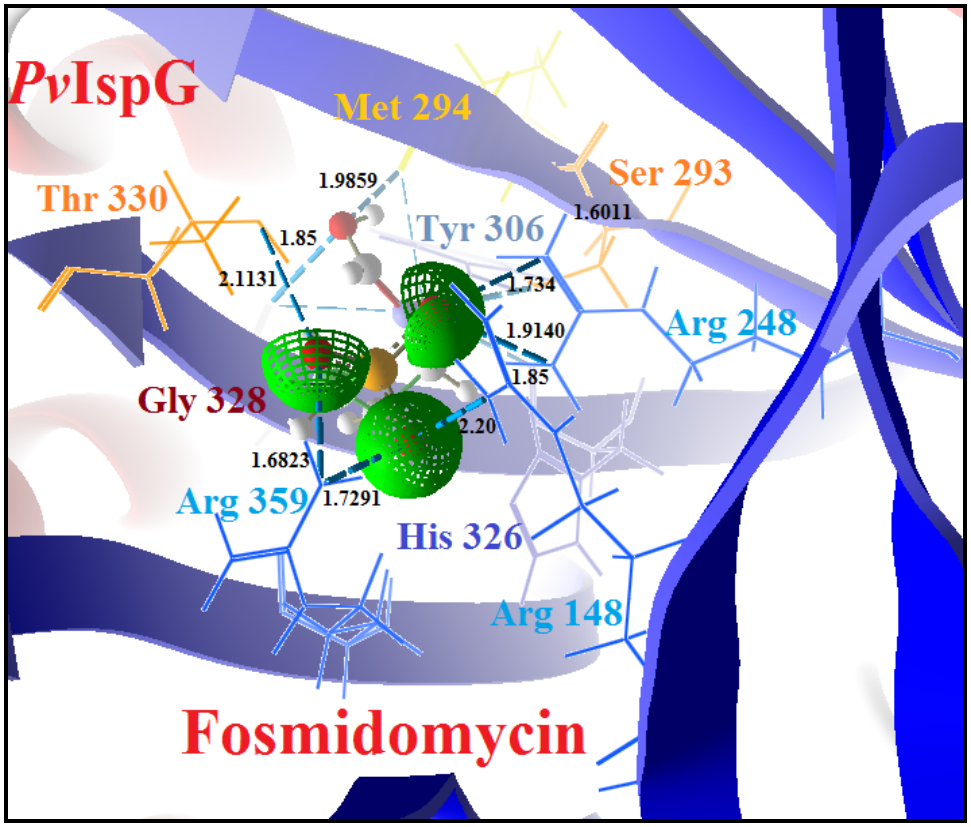


Figure: 6.14 *In-silico* drug inhibition studies for Fosmidomycin against *Pv*IspG protein. (drug is shown in green colour) and interactions shown with blue dotted lines.

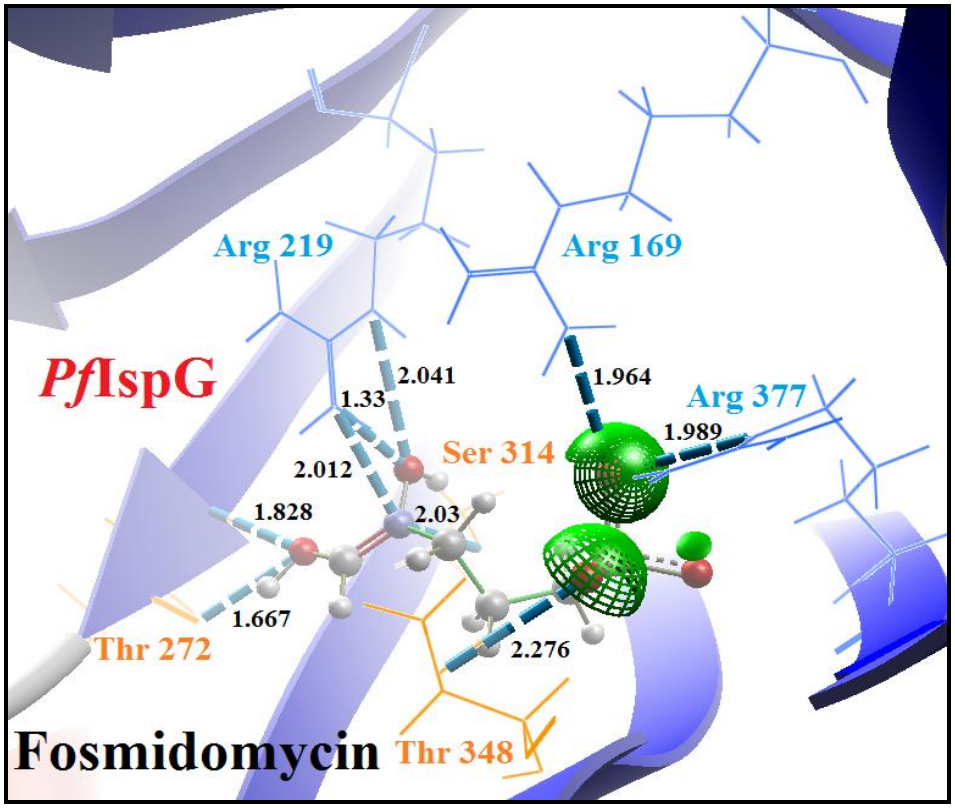


Figure: 6.15 *In-silico* drug inhibition study of *Pf*IspG protein with Fosmidomycin (Fos is shown in grey color and interaction shown with blue dotted lines).

6.2.4.3 Interaction with bacterial inhibitor Alkyne Diphosphate

Alkyne diphosphate derivatives have been recently reported as inhibitors of IspG enzyme in *E. coli*, *A. aeolicus* and *T. thermophilus* (Quitterer et al., 2015). IspG enzyme contains a binding site in its TIM barrel domain and this site is specific for the diphosphate group present in MECP substrate. Based on this observation, we reason that the alkyne diphosphate groups should also be having binding affinity for this active site and should be potent IspG inhibitors.

To study the inhibitory effect of these molecules, various derivative of alkyne diphosphates when docked against *Pv*IspG and *Pf*IspG enzymes validated, prop-2-yn-1-yl trihydrogen phosphate, an analogue of alkyl diphosphate, as a potent inhibitor. Prop-2-yn-1-yl trihydrogen phosphate has the binding affinity for the amino acid residues R148, R248, R359, S293, T330 and H326 of *P. vivax* IspG and amino acid residues R169, R218, R267, R376 and T348 of *Pf*IspG. These residues are common with the TIM barrel domain and the binding residues of MECP substrate, thus suggesting an inhibitory role of this derivative (**Figure 6.16 & 6.17**). Interaction of alkyne diphosphate was validated with molecular dynamics, confirming its affinity for *Pv*IspG and *Pf*IspG enzymes.

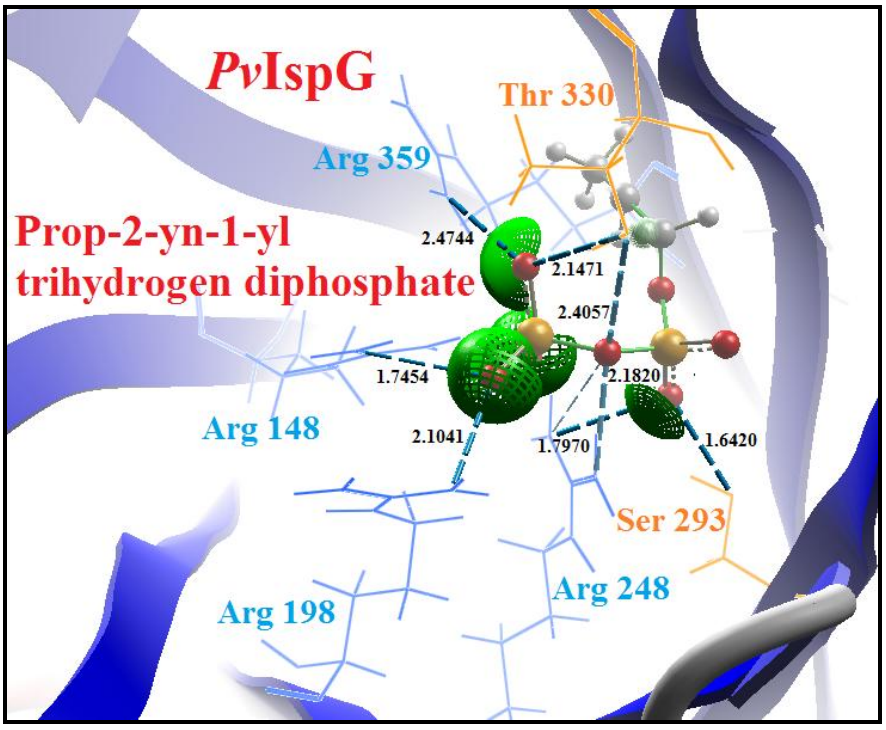


Figure: 6.16 *In-silico* drug inhibition studies for Prop-2-yn-1-yl trihydrogen diphosphate against *Pv*IspG protein. (Prop-2-yn-1-yl trihydrogen diphosphate molecule shown in green colour with conserved residues of *Pv*IspG enzyme shown with blue dotted lines)

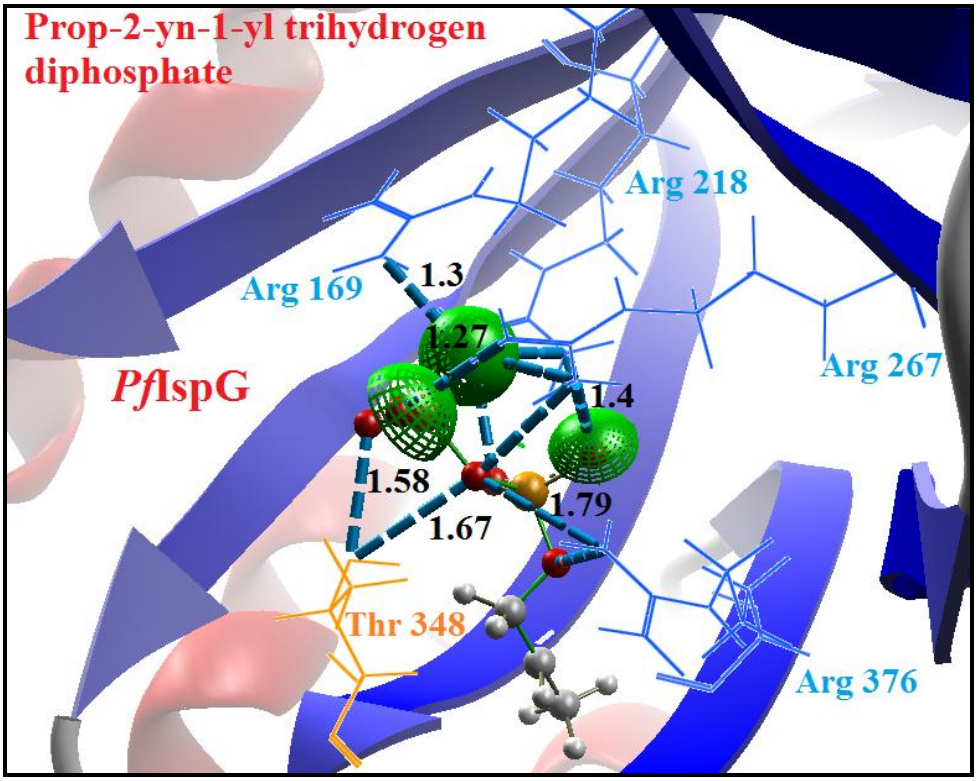


Figure: 6.17 *In-silico* drug inhibition studies for Prop-2-yn-1-yl trihydrogen diphosphate against *Pf*IspG protein. (Prop-2-yn-1-yl trihydrogen diphosphate molecule shown in green colour with conserved residues of *Pv*IspG enzyme shown with blue dotted lines)

Binding energies for Atorvastatin, Fos, and Alkyne diphosphate were calculated and compared with the binding energy of enzyme substrate complex **(Table 6.4 & 6.5)**.

Table 6.4: Binding energies for *Pv*IspG enzyme interaction with its substrates and different inhibitors.

Proteins	Ligand molecule	Binding energy ΔG (Kj/mol)
T. thermophilus IspG (4G9P)	MECP	-131.369
A. aeolicus IspG (3NOY)	[4Fe-4S] cluster	-44.873
P. vivax IspG	MECP	-107.815
	[4Fe-4S] cluster	-42.145
	Atorvastatin	-173.816
	Fosmidomycin	-115.103
	Alkyne diphosphate	-140.135

Table 6.5: Binding energies for *Pf*IspG enzyme interaction with its substrates and different inhibitors.

Proteins	Ligand molecule	Binding energy ΔG (Kj/mol)
T. thermophilus IspG (4G9P)	MECP	-131.369
A. aeolicus IspG (3NOY)	[4Fe-4S] cluster	-44.873
P. falciparum IspG	MECP	-107.228
	[4Fe-4S] cluster	-42.136
	Atorvastatin	-167.325
	Fosmidomycin	-101.341
	Alkyne diphosphate	-140.101

For both *P. falciparum* and *P. vivax*, Atorvastatin and Alkyne diphosphate have lower binding energy w.r.t. the natural substrate, which show that they can be used as a potent inhibitor whereas Fos with slightly higher binding energy can act as a weak inhibitor.

6.3 Conclusion

In the quest to identify a common novel target for action of existing or new antimalarial, we have concentrated on the IspG enzyme, catalysing the penultimate step of the MEP pathway. IspG is present in most of the bacteria as well as in malaria parasites and plants. Because of its essentiality for pathogen survival and absence in humans, it creates interest as a drug target. As this enzyme has not been characterized from any of the *Plasmodium* species or apicomplexans till date and there is no structure available for the enzyme from these organisms, we have generated a homology model for *Pv*IspG and *Pf*IspG of Indian isolates based on a prokaryotic organism *A. aeolicus*. The refined model was analysed to elucidate the active site for the binding of the substrate and the Fe-S clusters.

In comparison to the two domain bacterial IspG, the IspG enzyme of both *P. falciparum* and *P. vivax* consisted of an extra domain A* which has also been reported from certain bacteria, plants and apicomplexans. However, its function in protection of parasites from host ROS or as a supportive part for protein structure needs deeper analysis possibly by generating an X-ray crystal structure of the protein or mutagenesis studies using modified IspG with deleted A* domain.

To elucidate the probability of IspG enzyme as a drug target, we tested it against three drug molecules Fosmidomycin, Atorvastatin and Alkyne diphosphate. Among these, Fosmidomycin due to its ability to inhibit the IspC enzyme, is used often in *in-vitro* conditions to block *Plasmodium* growth and also as a part of antimalarial drug combination therapy used to treat malaria in patients. Our analysis shows that Fos can also bind with IspG in place of MECP substrate and thus can inhibit this enzyme. Similarly, Alkyne diphosphate appears to have perfect binding with the IspG enzyme and seems to show strong inhibition for the binding of MECP substrate. However, this drug has not been tested against *Plasmodium* as yet. The third molecule, Atorvastatin binds to the same cavity as for MECP substrate but utilizing different residues and thus may affect the catalytic activity by disturbing the formation of the closed cavity required for trapping the MECP substrate essential for proper functioning of the enzyme. However, final conclusion for the above study can be drawn only after certain detailed analysis using X-ray crystal structure of the IspG enzyme along with the drug molecules and *in-vitro* drug inhibition culture studies. Overall, in the drug binding studies, it was observed that IspG enzyme of *P. falciparum* shows interaction with similar conserved residues as *Pv*IspG and thus, can serve as a common molecule to target both *P. falciparum* and *P. vivax*.

Chapter VII

Conclusions and Future Prospectives

Previous chapters detailed about the various studies performed on the characterization of Isoprenoid biosynthesis pathway enzymes from *P. vivax* and *P. falciparum* Indian field isolates. These studies explored the significance of these enzymes in the pathway by detailing their structural, functional and evolutionary aspects. Based on these findings and various enzymatic activity and inhibition assays, the potential of these enzymes as drug candidates was assessed. This chapter mainly discuss the various conclusions drawn from the study and future perspectives suggesting the experimentation needed for further characterization of these proteins in the human malaria parasite, *Plasmodium*.

7.1 Specific Conclusions

1. The *Plasmodium ispD* and *ispG* genes are larger than their prokaryotic orthologues. The identity of these genes from *P. vivax* and *P. falciparum* varied from 30-84% with different *Plasmodium*, apicomplexans, and plants, still the active sites and domains remained conserved.

2. Recombinant clones of *Pv*IspD and *Pv*IspG were expressed as fusion proteins, further purified to perform enzyme kinetics and various activity assays:

   a) Enzyme kinetics studies performed for recombinant *Pv*IspD protein using commercially obtained substrates show maximum activity of enzyme at 67.8pM whereas K_M^{MEP} and K_M^{CTP} were calculated as 25.13 ± 5.344µM and 23.05 ± 8.575µM respectively. These assays indicate conservation of the functional domains in the *Pv*IspD enzyme.

   b) The cysteine residues in *Pv*IspG enzyme act as the site for formation of Fe-S clusters required for the activity of the enzyme. Our analysis show that Fe-S clusters can form on the enzyme only under reduced conditions, which is also required for the activity of the enzyme, proving the hypothesis stated by Partow et al. (2012) in *S. cerevisiae* studies. The Fe-S clusters were shown to have iron in +3 oxidation state, thus showing the formation of [4Fe-4S] clusters.

3. The purified recombinant *Pv*IspD and *Pv*IspG proteins were injected to raise antibodies in mice. Immuno-localization studies performed using these antibodies on *P. vivax* infected blood smears collected from field, indicated to possible localization of these proteins in the apicoplast only, the predicted functional site of the Isoprenoids biosynthesis pathway. This also suggests that the online targeting prediction servers need more standardization with the *P. vivax* database, as currently they are based on the sequences of *P. falciparum* only. The binding of anti-*Pv*IspG antibodies specifically with the *P. vivax* infected RBCs isolated from field isolates with no/ negligible binding with

normal uninfected RBCs or *P. falciparum* infected cells further support the immune-localization data.

4. Tertiary structures of *P. vivax* and *P. falciparum* IspD and IspG proteins show conservation of domain structures with presence of a typical Rossmann fold arrangement in IspD and TIM barrel domain in IspG. These structures indicate the conservation of the substrate binding sites throughout evolution, which suggests similar function of these enzymes within different organisms.

5. *In-silico* drug binding studies with IspD show the interaction of catalytic site residues with MMV008138 and Fosmidomycin indicating their inhibitory effect. Screening of various antibiotics against IspD suggested two molecules, Cefepime and Rifampicin as probable inhibitors, utilizing similar residues as the substrate. *In-vitro* inhibition studies performed for these two molecules against *Pv*IspD enzyme show their inhibitory effect where the IC_{50} values obtained for Cefepime and Rifampicin was 2.33 ± 0.13nM and 75.88 ± 1.19nM. This is the first study using Cefepime as an inhibitory molecule against *Plasmodium*. Rifampicin has been reported to cause parasite death *in vitro*, however, for the first time we are suggesting the inhibitory role of Rifampicin on any of the isoprenoids biosynthesis pathway enzyme. Further studies are needed to evaluate its role as an antimalarial.

6. Similarly, Fosmidomycin and prop-2-yn-1-yl trihydrogen phosphate (an analogue of alkyl diphosphate) show interaction with common residues as substrate in IspG, suggesting their inhibitory role. Based on the interacting residues, it seems that these alkyl diphosphate derivatives may act as potent inhibitors against both *P. vivax* and *P. falciparum*.

7. *Pv*IspG and *Pv*IspD protein are transcriptionally active throughout the erythrocytic stages as analysed in different clinical isolates using qPCR. This is in accordance with Bozdech et al, suggesting the varied expression of isoprenoids pathway genes in the erythrocytic stages with maximum expression in trophozoite stages.

8. The IspD and IspG proteins from *P. vivax* grouped with other primate parasites and diverged from the *P. falciparum* clade correlating with previous studies indicating a common origin of *P. vivax* and all primate parasites and separate origin for *P. falciparum*. *Chromera velia*, the bridging link between red algae and apicomplexans grouped together with the higher plants possessing photosynthetic plastids and diverged away from the apicomplexans with non-photosynthetic plastids suggesting an early divergence of the two groups during evolution.

Overall, in this study, two major enzymes of the isoprenoid biosynthesis pathway were characterized from *Plasmodium*, and their potential as probable drug target was assessed. Few molecules have been shortlisted as probable candidates for future antimalarials.

7.2 Direction for future research (future perspectives)

- Isoprene units synthesized by this pathway are exported to the cytosol from apicoplast, but the exact mechanism and transporter molecules involved in this process is yet to be explored.
- Specific environmental conditions required for IspG activity hindered the *in-vitro* activity assay with its substrates. Further standardization is required to perform the enzyme kinetics and inhibition assays.
- Regulation of metabolic flux through the MEP pathway in *Plasmodium* parasite is poorly understood, and enzymes participating in this process needs to be further explored.

References

References

Alexandre, M. A., Ferreira, C. O., Siqueira, A. M., Magalhaes, B. L., Mourao, M. P., Lacerda, M. V., and Alecrim, M. G. (2010). Severe *Plasmodium vivax* malaria, Brazilian Amazon. *Emerg Infect Dis* 16(10), 1611-4.

Altincicek, B., Kollas, A., Sanderbrand, S., Wiesner, J., Hintz, M., Beck, E., and Jomaa, H. (2001). GcpE is involved in the 2-*C*-methyl-Derythritol-4-phosphate pathway of isoprenoid biosynthesis in *Escherichia coli*. *J Bacteriol* 183, 2411-2416.

Aly, A. S., Vaughan, A. M., and Kappe, S. H. (2009). Malaria Parasite Development in the Mosquito Infection of the Mammalian Host. *Annu rev Microbio* 63, 195-221.

Amino, R., Giovannini, D., Thiberge, S., Gueirard, P., Boisson, B., Dubremetz, J. F., and Ménard, R. (2008). Host cell traversal is important for progression of the malaria parasite through the dermis to the liver. *Cell Host Microbe* 3(2), 88-96.

Andreassi, J. L., and Leyh, T. S. (2004). Molecular functions of conserved aspects of the GHMP kinase family. *Biochemistry* 43(46), 14594-14601.

Baca, A. M., and Hol, W. G. (2000). Overcoming codon bias: a method for high-level overexpression of *Plasmodium* and other AT-rich parasite genes in *Escherichia coli*. *Int J Parasitol* 30 (2), 113-118.

Bach, T. J., Boronat, A., Campos, N., Ferrer, A., and Vollack, K. U. (1999). Mevalonate biosynthesis in plants. *Crit Rev Biochem Mol* 34, 107-122.

Bahl, A., Brunk, B., Crabtree, J., Fraunholz, M. J., Gajria, B., Grant, G. R., Ginsburg, H., Gupta, D., Kissinger, J. C., and Labo, P. (2003). PlasmoDB: the *Plasmodium* genome resource. A database integrating experimental and computational data. *Nuc Acids Res* 31, 212-215. doi:10.1093/nar/gkg081

Baker, D. A. (2010). Malaria gametocytogenesis. *Mol Biochem Parasitol* 172(2), 57-65.

Battistini, M. R., Shoji, C., Handa, S., Breydo, L., and Merkler, D. J. (2016). Mechanistic binding insights for 1-deoxy-d-Xylulose-5-Phosphate synthase, the enzyme catalyzing the first reaction of isoprenoid biosynthesis in the malaria-causing protists, *Plasmodium falciparum* and *Plasmodium vivax*. *Protein Express Purif* 120, 16-27.

Baumeister, S., Wiesner, J., Reichenberg, A., Hintz, M., Bietz, S., Harb, O. S., Roos, D. S., Kordes, M., Friesen, J., et al. (2011). Fosmidomycin Uptake into *Plasmodium* and *Babesia*-Infected Erythrocytes Is Facilitated by Parasite-Induced New Permeability Pathways. *PLoS ONE* 6, 19334.

Bender, A., van Dooren, G. G., Ralph, S. A., McFadden, G. I., and Schneider, G. (2003). Properties and prediction of mitochondrial transit peptides from *Plasmodium falciparum*. *Mol Biochem Parasit* 132, 59-66. doi:10.1016/j.molbiopara.2003.07.001

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000). The Protein Data Bank. *Nuc Acids Res* 28, 235-242.

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nuc Acids Res* 7(6), 1513-1523.

Björkelid, C., Bergfors, T., Henriksson, L. M., Stern, A. L., Unge, T., Mowbray, S. L., and Jones, T. A. (2011). Structural and functional studies of mycobacterial IspD enzymes. *Acta Crystallogr Sect D Biol Crystallogr* 67(5), 403-414.

Bonday, Z. Q., Taketani, S., Gupta, P. D., and Padmanaban, G. (1997). Heme Biosynthesis by the Malarial Parasite. *J Biol Chem* 272, 21839-46.

Boucher, Y., and Doolittle, W. F. (2000). The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Mol Microbiol* 37(4), 703-716.

Bousema, T., and Drakeley, C. (2011). Epidemiology and infectivity of *Plasmodium falciparum* and *Plasmodium vivax* gametocytes in relation to malaria control and elimination. *Clin Microbiol Rev* 24(2), 377-410.

Bozdech, Z., Mok, S., Hu, G., Imwong, M., Jaidee, A., Russell, B., and Carlton, J. M. (2008). The transcriptome of *Plasmodium vivax* reveals divergence and diversity of transcriptional regulation in malaria parasites. *Proc Natl Acad Sci USA* 105(42), 16290-16295.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal biochem* 72(1-2), 248-254.

Brammer, L. A., Smith, J. M., Wade, H., and Meyers, C. F. (2011). 1-Deoxy-D-xylulose-5-phosphate synthase catalyzes a novel random sequential mechanism. *J Biol Chem* 286(42), 36522-36531. doi:10.1074/jbc.M111.259747.

Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., and Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nuc Acids Res* 41(W1), W349-W357.

Buetow, L., Brown, A. C., Parish, T., and Hunter, W. N. (2007). The structure of Mycobacteria 2 C-methyl-D-erythritol-2, 4-cyclodiphosphate synthase, an essential enzyme, provides a platform for drug discovery. *BMC Struct Biol* 7(1), 1.

Calisto, B. M., Perez-Gil, J., Bergua, M., Querol-Audi, J., Fita, I., and Imperial, S. (2007). Biosynthesis of isoprenoids in plants: Structure of the 2C-methyl-d-erithrytol 2, 4-cyclodiphosphate synthase from *Arabidopsis thaliana*. Comparison with the bacterial enzymes. *Protein Sci* 16(9), 2082-2088.

Carlton, J. M., Adams, J. H., Silva, J. C., Bidwell, S. L., Lorenzi, H., Caler, E., Crabtree, J., Angiuoli, S. V., Merino, E. F., et al. (2008). Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* 455(9), 757-763.

Carter, R., and Mendis, K. N. (2002). Evolutionary and historical aspects of the burden of malaria. *Clin Microbiol Rev* 15(4), 564-594.

Cassera, M. B., Gozzo, F. C., Alexandri, F. L. D., Merino, E. F., del Portillo, H. A., Peres, V. J., Almeida, I. C., Eberlin, M. N., Wunderlich, G., et al. (2004). The methylerythritol phosphate pathway is functionally active in all intraerythrocytic stages of *Plasmodium falciparum*. *J Biol Chem* 279, 749-759.

Cassera, M. B., Merino, E. F., Peres, V. J., Kimura, E. A., Wunderlich, G., and Katzin, A. M. (2007). Effect of fosmidomycin on metabolic and transcript profiles of the methylerythritol phosphate pathway in *Plasmodium falciparum*. *Mem Inst Oswaldo Cruz* 102(3), 377-384.

Charan, M., Singh, N., Kumar, B., Srivastava, K., Siddiqi, M. I., and Habib, S. (2014). Sulfur Mobilization for Fe-S cluster assembly by the essential SUF pathway in the *Plasmodium falciparum* apicoplast and its inhibition. *Antimicrob Agents Chemother* 58(6), 3389-98.

Claros, M. G., and Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* 241, 779-786. doi:10.1111/j.1432-1033.1996.00779.x

Colovos, C., and Yeates, T.O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Prot Sci* 2, 1511-1519.

Cornish, R. M., Roth, J. R., and Poulter, C. D. (2006). Lethal Mutations in the Isoprenoid Pathway of *Salmonella enteric*. *J Baterol* 188(4), 1444-1450. doi:10.1128/JB.188.4.1444-1450.2006

Couto, A. S., Kimura, E. A., Peres, V. J., Uhrig, M. L., and Katzin, A. M. (1999). Active isoprenoid pathway in the intra-erythrocytic stages of *Plasmodium falciparum* presence of dolichols of 11 and 12 isoprene units. *Biochem J* 341, 629-637.

Cowman, A. F., and Crabb, B. S. (2006). Invasion of red blood cells by malaria parasites. *Cell* 124, 755-766.

Crane, C. M., Kaiser, J., Ramsden, N. L., Lauw, S., Rohdich, F., Eisenreich, W., Hunter, W. N, Bacher, A., and Diederich, F. (2006). Fluorescent Inhibitors for IspF, an Enzyme in the Non-Mevalonate Pathway for Isoprenoid Biosynthesis and a Potential Target for Antimalarial Therapy. *Angew Chem* 118, 1082-1087.

Dahl, E. L., and Rosenthal, P. J. (2007). Multiple antibiotics exert delayed effects against the *Plasmodium falciparum* apicoplast. *Antimicrob Agents Chemother* 51(10), 3485-3490.

Dahl, E. L., Shock, J. L., Shenai, B. R., Gut, J., DeRisi, J. L., and Rosenthal, P. J. (2006). Tetracyclines specifically target the apicoplast of the malaria parasite *Plasmodium falciparum*. *Antimicrob Agents Chemother* 50, 3124-3131.

Das, A., Holloway, B., Collins, W. E., Shama, V., Ghosh, S. K., Sinha, S., Hasnain, S. E., Talwar, G. P., and Lal, A. A. (1995). Species-specific 18S rRNA gene amplification for the detection of *P. falciparum* and *P. vivax* malaria parasites. *Mol Cell Probe* 9, 161-165. doi:10.1006/mcpr.1995.0025

Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 4.5, San Diego: Dassault Systèmes, 2015.

D'Alexandri, F. L., Kimura, E. A., Peres, V. J., and Katzin, A. M. (2006). Protein dolichylation in *Plasmodium falciparum*. *FEBS Lett* 580, 6343-6348.

Dellibovi-Ragheb, T. A., Gisselberg, J. E., and Prigge, S. T. (2013). Parasites FeS up: iron-sulfur cluster biogenesis in eukaryotic pathogens. *PLoS Pathog* 9(4), e1003227.

de Macedo, C. S., Uhrig, M. L., Kimura, E. A., and Katzin, A. M. (2002). Characterization of the isoprenoid chain of coenzyme Q in *Plasmodium falciparum*. *FEMS Microbiol Lett* 207, 13-20.

Dhanasekaran, S., Chandra, N. R., Sagar, B. C., Rangarajan, P. N., and Padmanaban, G. (2004). δ-Aminolevulinic acid dehydratase from *Plasmodium falciparum* indigenous versus imported. *J Biol Chem* 279(8), 6934-6942.

Dondorp, A. M., Nosten, F., Yi, P., Das, D., Phyo, A. P., Tarning, J., and Ringwald, P. (2009). Artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med* 361(5), 455-467.

Eisenberg, D., Lüthy, R., and Bowie, J. U. (1997). VERIFY3D: Assessment of protein models with three-dimensional profiles. *Method Enzymol* 277, 396-404. doi:10.1016/S0076-6879(97)77022-8

Eoh, H., Narayanasamy, P., Brown, A. C., Parish, T., Brennan, P. J., and Crick, D. C. (2009). Expression and characterization of soluble 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase from bacterial pathogens. *Chemistry & Biology* 16(12), 1230-1239.

Escalante, A. A., Cornejo, O. E., Freeland, D. E., Poe, A. C., Durrego, E., Collins, W. E., and Lal, A. A. (2005). A monkey's tale: the origin of *Plasmodium vivax* as a human malaria parasite. *Proc Natl Acad Sci USA* 102, 1980-1985. doi:10.1073/pnas.0409652102

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2006). Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* 5(6), 1-30. doi:10.1002/0471140864.ps0209s50

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 783-791.

Ferreira, A., Balla, J., Jeney, V., Balla, G., and Soares, M. P. (2008). A Central Role for Free Heme in the Pathogenesis of Severe Malaria: The Missing Link? *J Mol Med* 86 (10), 1097-1111.

Ferreira, J. F., Luthria, D. L., Sasaki, T., and Heyerick, A. (2010). Flavonoids from Artemisia annua L. as antioxidants and their potential synergism with artemisinin against malaria and cancer. *Molecules* 15(5), 3135-3170.

Fichera, M. E., Bhopale, M. K., and Roos, D. S. (1995). *In-vitro* assays elucidate peculiar kinetics of clindamycin action against *Toxoplasma gondii*. *Antimicrob Agents Chemother* 39(7), 1530-1537.

Fish, W. W. (1988). Rapid Colorimetric Micro method for the Quantitation of Complexed Iron in Biological Samples. *Methods Enzymol* 158, 357-364.

Foth, B. J., and McFadden, G. I. (2003). The apicoplast: a plastid in *Plasmodium falciparum* and other Apicomplexan parasites. *Int Rev Cytol* 224, 57-110.

Foth, B. J., Ralph, S. A., Tonkin, C. J., Struck, N. S., Fraunholz, M., Roos, D. S., Cowman, A. F., and McFadden, G. I. (2003). Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. *Science* 299, 705-708. doi:10.1126/science.1078599

Gabrielsen, M., Bond, C. S., Hallyburton, I., Hecht, S., Bacher, A., Eisenreich, W., and Hunter, W. N. (2004). Hexameric assembly of the bifunctional methylerythritol 2, 4-cyclodiphosphate synthase and protein-protein associations in the deoxy-xylulose-dependent pathway of isoprenoid precursor biosynthesis. *J Biol Chem* 279(50), 52753-52761.

Gabrielsen, M., Kaiser, J., Rohdich, F., Eisenreich, W., Laupitz, R., Bacher, A., and Hunter, W. N. (2006). The crystal structure of a plant 2C-methyl-D-erythritol 4-phosphate cytidylyltransferase exhibits a distinct quaternary structure compared to bacterial homologues and a possible role in feedback regulation for cytidine monophosphate. *FEBS J* 273(5), 1065-1073.

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., and Paulsen, I. T. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419(6906), 498-511.

Geist, J. G., Lauw, S., Illarionova, V., Illarionov, B., Fischer, M., Gräwert, T., Rohdich, F. Eisenreich, W. Kaiser, J., et al. (2010). Thiazolopyrimidine inhibitors of 2-methylerythritol 2,4-cyclodiphosphate synthase (IspF) from *Mycobacterium tuberculosis* and *Plasmodium falciparum*. *Chem Med Chem* 5, 1092-1101.

Gisselberg, J. E., Dellibovi-Ragheb, T. A., Matthews, K. A., Bosch, G., and Prigge, S. T. (2013). The Suf iron-sulfur cluster synthesis pathway is required for apicoplast maintenance in malaria parasites. *PLoS Pathog* 9(9), e1003655.

Gräwert, T., Kaiser, J., Zepeck, F., Laupitz, R., Hecht, S., Amslinger, S., and Buchner, J. (2004). IspH Protein of *Escherichia coli*: Studies on Iron-Sulfur Cluster Implementation and Catalysis. *J Am Chem Soc* 126(40), 12847-12855.

Guerra, F., Wang, K., Li, J., Wang, W., Liu, Y. L., Amin, S., and Oldfield, E. (2014). Inhibition of the 4Fe–4S proteins IspG and IspH: an EPR, ENDOR and HYSCORE investigation. *Chem Sci* 5(4), 1642-1649.

Hahn, F. M., Eubanks, L. M., Testa, C. A., Blagg, B. S., Baker, J. A., and Poulter, C. D. (2001). 1-Deoxy-d-Xylulose 5-Phosphate synthase, the gene product of open reading frame (ORF) 2816 and ORF 2895 in *Rhodobacter capsulatus*. *J Bacteriol* 183(1), 1-11.

Hall, N., Karras, M., Raine, J. D., Carlton, J. M., Kooij, T. W., Berriman, M., and James, K. (2005). A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 307(5706), 82-86.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41, 95-98.

Handa, S., Ramamoorthy, D., Spradling, T. J., Guida, W. C., Adams, J. H., Bendinskas, K. G., and Merkler, D. J. (2013). Production of recombinant 1-deoxy-d-xylulose-5-phosphate synthase from *Plasmodium vivax* in *Escherichia coli*. *FEBS Open Bio* 3(1), 124-129.

Harlow, E., and Lane, D. (1988). Antibodies - A Laboratry Manual, Cold Spring Harbor Laboratory Press, NY, USA.

Haussig, J. M., Matuschewski, K., and Kooij, T. W. A. (2014). Identification of vital and dispensable sulfur utilization factors in the *Plasmodium* apicoplast. *PLoS ONE* 9(2), e89718.

Hemmerlin, A., Hoeffler, J. F., Meyer, O., Tritsch, D., Kagan, I. A., Grosdemange-Billiard, C., Rohmer, M., and Bach, T. J. (2003). Cross-talk between the Cytosolic Mevalonate and the Plastidial Methylerythritol Phosphate Pathways in Tobacco Bright Yellow-2 Cells. *J Biol Chem* 278(29), 26666-26676.

He, C. Y., Striepen, B., Pletcher, C. H., Murray, J. M., and Roos, D. S. (2001). Targeting and processing of Nuclear encoded apicoplast proteins in plastid segregation mutants of *Toxoplasma Gondii*. *J Biol Chem* 276, 28436–28442.

Hecht, S., Eisenreich, W., Adam, P., Amslinger, S., Kis, K., Bacher, A., and Rohdich, F. (2001). Studies on the nonmevalonate pathway to terpenes: the role of the GcpE (IspG) protein. *Proc Natl Acad Sci USA* 98(26), 14837-14842.

Henriksson, L. M., Bjorkelid, C., Mowbray, S. L., and Unge, T. (2006). The 1.9Å resolution structure of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose 5-phosphate reductoisomerase, a potential drug target. *Acta Crystallo Sect D Biol Crystallo* 62, 807-813.

Hirsch, A. K., Alphey, M. S., Lauw, S., Seet, M., Barandun, L., Eisenreich, W., and Diederich, F. (2008). Inhibitors of the kinase IspE: structure–activity relationships and co-crystal structure analysis. *Org Biomol Chem* 6(15), 2719-2730.

Hirsch, A. K. H., Lauw, S., Gersbach, P., Schweizer, W. B., Rohdich, F., Eisenreich, W., Bacher, A., and Diederich, F. (2007). Nonphosphate inhibitors of IspE protein, a kinase in the non-mevalonate pathway for isoprenoid biosynthesis and a potential target for antimalarial therapy. *Chem Med Chem* 2, 806-810.

Hunter, W. N. (2007). The non-mevalonate pathway of isoprenoid precursor biosynthesis. *J Biol Chem* 282(30), 21573-21577.

Hunter, W. N. (2011). Isoprenoid precursor biosynthesis offers potential targets for drug discovery against diseases caused by apicomplexan parasites. *Curr Top Med Chem* 11(16), 2048-2059.

Illarionova, V., Kaiser, J., Ostrozhenkova, E., Bacher, A., Fischer, M., Eisenreich, W., and Rohdich, F. (2006). Nonmevalonate terpene biosynthesis enzymes as antiinfective drug targets: substrate synthesis and high-throughput screening methods. *J Org Chem* 71(23), 8824-8834.

Imlay, L. S., Armstrong, C. M., Masters M. C., Li, T., Price, K. E., Edwards, R. L., Mann, K. M., Li, L. X., Stallings, C. L., et al. (2015). *Plasmodium* IspD (2-C-Methyl-d-

erythritol-4-Phosphate Cytidyltransferase), an Essential and Druggable Antimalarial Target. *ACS Infect Dis* 1(4), 157-167. doi**:** 10.1021/id500047s

Jelenska, J., Crawford, M. J., Harb, O. S., Zuther, E., Haselkorn, R., Roos, D. S., and Gornicki, P. (2001). Subcellular localization of acetyl-CoA carboxylase in the apicomplexan parasite *Toxoplasma gondii*. *Proc Natl Acad Sci USA* 98(5), 2723-2728.

Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Turbachova, I., Eberl, M., Zeidler, J., et al. (1999). Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* 285, 1573-1575.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer applications in the biosciences: CABIOS* 8(3), 275-282.

Jordão, F. M., Saito, A. Y., Miguel, D. C., de Jesus Peres, V., Kimura, E. A., and Katzin, A. M. (2011). *In vitro* and *in vivo* antiplasmodial activities of risedronate and its interference with protein prenylation in *Plasmodium falciparum*. *Antimicrob Agents Chemother* 55:2026-2031. http://dx.doi.org/10.1128/AAC.01820-10.

Jordão, F. M., Gabriel, H. B., Alves, J. M., Angeli, C. B., Bifano, T. D., Breda, A., de Azevedo, M. F., Basso, L. A., Wunderlich, G., Kimura, E. A., and Katzin, A. M. (2013). Cloning and characterization of bifunctional enzyme farnesyl diphosphate/geranylgeranyl diphosphate synthase from *Plasmodium falciparum*. *Malar J* 12, 184. http://dx.doi.org/10.1186/1475-2875-12-184.

Jin, Y., Kebaier, C., and Vanderberg, J. (2007). Direct microscopic quantification of dynamics of *Plasmodium berghei* sporozoite transmission from mosquitoes to mice. *Infect Immun* 75, 5532-5539.

Janouškovec, J., Horák, A., Oborník, M., Lukeš, J., and Keeling, P. J. (2010). A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc Natl Acad Sci USA* 107(24), 10949-10954.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., and Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12), 1647-1649.

Kochar, D. K., Das, A., Kochar, S. K., Saxena, V., Sirohi, P., Garg, S., and Gupta, V. (2009). Severe *Plasmodium vivax* malaria: a report on serial cases from Bikaner in north western India. *Am J Trop Med Hyg* 80(2), 194-198.

Kochar, D. K., Saxena, V., Singh, N., Kochar, S. K., and Kumar, S. V. (2005). *Plasmodium vivax* malaria. *Emerg Infect Dis* 11(1), 132-134. doi:10.3201/eid1101.040519

Krotoski, W. A., Krotoski, D. M., Garnham, P. C., Bray, R. S., Killick-Kendrick, R., Draper, C. C., and Guy, M. W. (1980). Relapses in primate malaria: discovery of two populations of exoerythrocytic stages. Preliminary note. *Br Med J* 280(6208), 153-154.

Kumaratilake, L. M., and Ferrante, A. (2000). Opsonization and phagocytosis of *Plasmodium falciparum* merozoites measured by flow cytometry. *Clin Diagn Lab Immunol* 7(1), 9-13.

Kumar, B., Chaubey, S., Shah, P., Tanveer, A., Charan, M., Siddiqi, M. I., et al. (2011). Interaction between sulphur mobilisation proteins SufB and SufC: Evidence for an iron–sulphur cluster biogenesis pathway in the apicoplast of *Plasmodium falciparum*. *Intl J Parasitol* 41, 991-99.

Kumar, S., Hahn, F. M., Baidoo, E., Kahlon, T. S., Wood, D. F., McMahan, C. M., Cornish, K., Keasling, J. D., Daniell, H., and Whalen, M. C. (2012). Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts. *Metab Eng* 14(1), 19-28. http://dx.doi.org/10.1016/j.ymben.2011.11.005.

Kuemmerle, H. P., Murakawa, T., and De Santis, F. (1987). Pharmacokinetic evaluation of fosmidomycin, a new phosphonic acid antibiotic. *Chemioterapia: international journal of the Mediterranean Society of Chemotherapy* 6(2), 113-119.

Labaied, M., Jayabalasingham, B., Bano, N., Cha, S. J., Sandoval, J., Guan, G., and Coppens, I. (2011). *Plasmodium* salvages cholesterol internalized by LDL and synthesized de novo in the liver. *Cell Microbiol* 13, 569-586. http://dx.doi.org/10.1111/j.1462-5822.2010.01555.x.

Larkin, M. A., Blackshields, G., Brown, N., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., and Lopez, R. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948. doi:10.1093/bioinformatics/btm404

Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Biochemistry* 77(1), 521.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK-a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26, 283-291.

Laupitz, R., Grawert, T., Rieder, C., Zepeck, F., Bacher, A., Arigoni, D., Rohdich, F., and Eisenreich, W. (2004). Stereochemical Studies on the Making and Unmaking of

Isopentenyl Diphosphate in Different Biological Systems. *Chem Biodiversity* 1, 1367-1376.

Lee, M., Gräwert, T., Quitterer, F., Rohdich, F., Eppinger, J., Eisenreich, W., Bacher, A., and Groll, M. (2010). Biosynthesis of Isoprenoids: Crystal Structure of the [4Fe–4S] Cluster Protein IspG. *J Mol Biol* 404, 600-610. doi:10.1016/j.jmb.2010.09.050

Lehmann, C., Lim, K., Toedt, J., Krajewski, W., Howard, A., Eisenstein, E., and Herzberg, O. (2002). Structure of 2C-methyl-D-erythrol-2, 4-cyclodiphosphate synthase from *Haemophilus influenzae*: Activation by conformational transition. *Proteins: Struct Func Bioinfo* 49(1), 135-138.

Leivar, P., Antolín-Llovera, M., Ferrero, S., Closa, M., Arró, M., Ferrer, A., and Campos, N. (2011). Multilevel control of *Arabidopsis* 3-hydroxy-3-methylglutaryl coenzyme A reductase by protein phosphatase 2A. *Plant Cell* 23(4), 1494-1511.

Leivar, P., González, V. M., Castel, S., Trelease, R. N., López-Iglesias, C., Arró, M., and Fernàndez-Busquets, X. (2005). Subcellular localization of Arabidopsis 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Plant physiol* 137(1), 57-69.

Lell, B., Ruangweerayut, R., Wiesner, J., Missinou, M. A., Schindler, A., Baranek, T., and Kremsner, P. G. (2003). Fosmidomycin, a novel chemotherapeutic agent for malaria. *Antimicrob Agents Chemother* 47(2), 735-738.

Levine, N. D., Lee, J. J., Hunter, S. H., and Bovee, E. C. (1985). Phylum Apicomplexa An Illustrated Guide to the Protozoa. *Society of Protozoologists, Allen Press, Lawrence*. 322-374.

Lichtenthaler H. K., Schwender J., Disch A., and Rohmer M. (1997). Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway. *FEBS Lett* 400, 271–274.

Lill, R., and Mühlenhoff, U. (2005). Iron–sulfur protein biogenesis in eukaryotes. *Trends Biochem Sci* 30(3), 133-141.

Lillo, A. M., Tetzlaff, C. N., Sangari, F. J., and Cane, D. E. (2003). Functional expression and characterization of EryA, the erythritol kinase of Brucella abortus, and enzymatic synthesis of L-erythritol-4-phosphate. *Bioorg Med Chem Lett* 13(4), 737-739.

Liu, Y. L., Guerra, F., Wang, K., Wang, W., Li, J., Huang, C., Zhu, W., Houlihan, K., Li, Z., et al. (2012). Structure, function and inhibition of the two- and three-domain 4Fe-4S IspG proteins. *Proc Natl Acad Sci USA* 109, 8558-8563. doi:10.1073/pnas.1121107109

Lombard, J., and Moreira, D. (2011). Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. *Mol Biol Evol* 28, 87-99.

Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003). Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins: Struct Func Genet* 50, 437-450. doi:10.1002/prot.10286

Mac Sweeney, A., Lange, R., Fernandes, R. P., Schulz, H., Dale, G. E., Douangamath, A., and Oefner, C. (2005). The crystal structure of *E. coli* 1-deoxy-D-xylulose-5-phosphate reductoisomerase in a ternary complex with the antimalarial compound fosmidomycin and NADPH reveals a tight-binding closed enzyme conformation. *J Mol Biol* 345(1), 115-127.

Manzano, D., Fernandez-Busquets, X., Schaller, H., González, V., Boronat, A., Arró, M., and Ferrer, A. (2004). The metabolic imbalance underlying lesion formation in *Arabidopsis thaliana* overexpressing farnesyl diphosphate synthase (isoform 1S) leads to oxidative stress and is triggered by the developmental decline of endogenous HMGR activity. *Planta* 219(6), 982-992.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., et al. (2015). CDD: NCBI's conserved domain database. *Nuc Acids Res* 28, 43. doi: 10.1093/nar/gku1221

Martinsen, E. S., Perkins, S. L., and Schall, J. J. (2008). A three-genome phylogeny of malaria parasites (*Plasmodium* and closely related genera): evolution of life-history traits and host switches. *Mol Phylogenet Evol* 47(1), 261-273.

Masini, T., and Hirsch, A. K. H. (2014). Development of Inhibitors of the 2C‑Methyl‑D‑ erythritol 4‑Phosphate (MEP) Pathway Enzymes as Potential Anti-Infective Agents. *J Med Chem* 57, 9740-63.

Matsue, Y., Mizuno, H., Tomita, T., Asami, T., Nishiyama, M., and Kuzuyama, T. (2010). The herbicide ketoclomazone inhibits 1-deoxy-D-xylulose 5-phosphate synthase in the 2-C-methyl-D-erythritol 4-phosphate pathway and shows antibacterial activity against *Haemophilus influenzae*. *J Antibiot* 63, 583-588.

Mazumdar, J., and Striepen, B. (2007). Make It or Take It: Fatty Acid Metabolism of Apicomplexan Parasites. *Eukaryotic Cell* 6 (10), 1727-35.

Mbaya, B., Rigomier, D., Edorh, G. G., Karst, F., and Schrevel, J. (1990). Isoprenoid metabolism in *Plasmodium falciparum* during the intraerythrocytic phase of malaria. *Biochem Biophys Res Commun* 173, 849–854. http://dx.doi.org/10.1016/S0006-291X(05)80864-2.

McFadden, G. I., Reith, M. E., Munholland, J., and Lang-Unnasch, N. (1996). Plastid in human parasites. *Nature* 381, 482.

McFadden, G. I., and Roos, D. S. (1999). Apicomplexan plastids as drug targets. *Trends Microbiol* 7(8), 328-333.

Medica, D. L., and Sinnis, P. (2005). Quantitative dynamics of *Plasmodium yoelii* sporozoites transmission by infected anopheline mosquitoes. *Infect Immun* 73, 4363-4369.

Mombelli, P., Witschel, M. C., van Zijl, A. W., Geist, J. G., Rottmann, M., Freymond, C., and Siepe, I. (2012). Identification of 1, 3-Diiminoisoindoline Carbohydrazides as Potential Antimalarial Candidates. *Chem Med Chem* 7(1), 151-158.

Moll, K., Ljungström, I., Perlmann, H., Scherf, A., Wahlgren, M., and Manassas, V. (2008). MR4/ATCC (Manassas, Virginia and BioMalPar (Paris, France)). Methods in malaria research.

Moura, I. C., Wunderlich, G., Uhrig, M. L., Couto, A. S., and Peres, V. J. (2001). Limonene arrests parasite development and inhibits isoprenylation of proteins in *Plasmodium falciparum*. *Antimicrob Agents Chemother* 45, 2553–2558.

Mullin, K. A., Lim, I., and Ralph, S. A. (2006). Membrane transporters in the relict plastid of malaria parasites. *Proc Natl Acad Sci USA* 103(25), 9572-9577.

Nagaraj, V. A., Arumugam, R., Prasad, D., Rangarajan, P. N, and Padmanaban, G. (2010). Protoporphyrinogen IX Oxidase from *Plasmodium falciparum* is Anaerobic and is Localized to the Mitochondrion. *Mol Biochem Parasitol* 174(1), 44-52.

Nagaraj, V. A., Prasad, D., Rangarajan, P. N, and Padmanaban, G. (2009). Mitochondrial localization of functional ferrochelatase from *Plasmodium falciparum. Mol Biochem Parasitol* 168, 109-112.

Nagaraj, V. A., Sundaram, B., Varadarajan, N. M., Subramani, P. A., Kalappa, D. M., Ghosh, S. K., and Padmanaban, G. (2013). Malaria parasite-synthesized heme is essential in the mosquito and liver stages and complements host heme in the blood stages of infection. *PLoS Pathog* 9(8), e1003522.

Nair, S. C., Brooks, C. F., Goodman, C. D., Sturm, A., McFadden, G. I., Sundriyal, S., Anglin, J. L., Song, Y., and Moreno, S. N. J. (2011). Apicoplast isoprenoid precursor synthesis and the molecular basis of fosmidomycin resistance in *Toxoplasma gondii*. *J Exp Med* 208, 1547-1559. http://dx.doi.org/10.1084/jem.20110039.

National Vector Borne Disease Control Programme. "National drug policy on malaria. National Vector Borne Disease Control Program 2010", Directorate General of Health Services Ministry of Health and Family Welfare, Delhi, India.

Noedl, H., Se, Y., Schaecher, K., Smith, B. L., Socheat, D., and Fukuda, M. M. (2008). Evidence of artemisinin-resistant malaria in western Cambodia. *N Engl J Med 359*(24), 2619-2620.

Okada, K., and Hase, T. (2005). Cyanobacterial Non-mevalonate Pathway: (*E*)-4-Hydroxy-3-methylbut-2-enyl diphosphate synthase interacts with ferredoxin in *Thermosynechococcus elongatus* BP-1. *J Biol Chem* 280 (21), 20672-20679.

O'Rourke, P. E., Kalinowska-Tłuścik, J., Fyfe, P. K., Dawson, A., and Hunter, W. N. (2014). Crystal structures of IspF from *Plasmodium falciparum* and *Burkholderia cenocepacia*: comparisons inform antimicrobial drug target assessment. *BMC Struct Biol* 14(1), 1.

Ostrovsky, D., Diomina, G., Lysak, E., Matveeva, E., Ogrel, O., and Trutko S. (1998). Effect of oxidative stress on the biosynthesis of 2-*C*-methyl-D-erythritol-2,4-cyclopyrophosphate and isoprenoids by several bacterial strains. *Arch Microbiol* 171, 69-72.

Pakalapati, D., Garg, S., Middha, S., Acharya, J., Subudhi, A. K., Boopathi, A. P., Saxena, V., Kochar, S. K., Kochar, D. K., and Das, A. (2013). Development and evaluation of a 28S rRNA gene-based nested PCR assay for *P. falciparum* and *P. vivax*. *Pathog Glob Health* 107, 180-188. doi:10.1179/2047773213Y.0000000090

Pala, Z. R., Saxena, V., Saggu, G. S., Yadav, S. K., Pareek, R. P., Kochar, S. K., and Garg, S. (2016). Structural and functional characterization of an iron–sulfur cluster assembly scaffold protein-SufA from *Plasmodium vivax*. *Gene* 585(1), 159-165.

Parsons, M., Karnataki, A., Feagin, J. E., and DeRocher, A. (2007). Protein trafficking to the apicoplast: deciphering the apicomplexan solution to secondary endosymbiosis. *Eukaryot Cell* 6, 1081–1088.

Partow, S., Siewers, V., Daviet, L., Schalk, M., and Nielsen, J. (2012). Reconstruction and Evaluation of the Synthetic Bacterial MEP Pathway in *Saccharomyces cerevisiae*. *PLoS ONE* 7(12), e52498. doi:10.1371/journal.pone.0052498

Pfefferkorn, E. R., Nothagel, R. F., and Borotz, S. E. (1992). Parasiticidal effect of clindamycin on *Toxoplasma gondii* grown in cultured cells and selection of a drug resistant mutant. *Antimicrob Agents Chemother* 36, 1091–1096.

Pérez-Gil, J., Bergua, M., Boronat, A., and Imperial, S. (2010). Cloning and functional characterization of an enzyme from *Helicobacter pylori* that catalyzes two steps of the methylerythritol phosphate pathway for isoprenoid biosynthesis. *Biochim Biophy Acta* 1800, 919-928. doi:10.1016/j.bbagen.2010.06.008

Pethe, K., Swenson, D. L., Alonso, S., Anderson, J., Wang, C. and Russell, D. G. (2004). Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. *Proc Natl Acad Sci USA* 101, 13642–13647.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera-a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612. doi:10.1002/jcc.20084.

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8, 785-786.

Quitterer, F., Frank, A., Wang, K., Rao, G., O'Dowd, B., and Li, J. (2015). Atomic-Resolution Structures of Discrete Stages on the Reaction Coordinate of the [Fe_4S_4] Enzyme IspG (GcpE). *J Mol Biol* 427, 2220-2228. doi: org/10.1016/j.jmb.2015.04.002

Ralph, S. A., van Dooren, G. G., Waller, R. F., Crawford, M. J., Fraunholz, M. J., Foth, B. J., Tonkin , C. J., Roos, D. S., and McFadden, G. I. (2004). Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nature Rev Microbiol* 2, 203-216.

Ramya, T. N. C., Surolia, N., and Surolia, A. (2006). 15-Deoxyspergualin modulates *Plasmodium falciparum* heat shock protein function. *Biochem Biophys Res Commun* 348, 585-592.

Rathore, D., Wahl, A. M., Sullivan, M., and McCutchan, T. F. (2001). A phylogenetic comparison of gene trees constructed from plastid, mitochondrial and genomic DNA of *Plasmodium* species. *Mol Biochem Parasitol* 114(1), 89-94.

Rekittke, I., Nonaka, T., Wiesner, J., Demmer, U., Warkentin, E., Jomaa, H., and Ermler, U. (2011). Structure of the E-1-hydroxy-2-methyl-but-2-enyl-4-diphosphate synthase (GcpE) from *Thermus thermophilus*. *FEBS lett* 585(3), 447-451.

Ricagno, S., Grolle, S., Bringer-Meyer, S., Sahm, H., and Lindqvist, Y. (2004). Crystal structure of 1-deoxy-D-xylulose-5-phosphate reductoisomerase from *Zymomonas mobilis* at 1.9Å resolution. *Biochim Biophys Acta* 1698, 37-44.

Richard, S. B., Bowman, M. E., Kwiatkowski, W., Kang, I., Chow, C., Lillo, A. M., Cane, D. E., and Noel, J. P. (2001). Structure of 4-diphosphocytidyl-2-C-methylerythritol synthetase involved in mevalonate- independent isoprenoid biosynthesis. *Nat Struct Biol* 8, 641-648.

Richard, S. B., Ferrer, J. L., Bowman, M. E., Lillo, A. M., Tetzlaff, C. N., Cane, D. E., and Noel, J. P. (2002). Structure and mechanism of 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase an enzyme in the mevalonate-independent isoprenoid biosynthetic pathway. *J Biol Chem* 277, 8667-8672

Rock, R. C. (1971). Incorporation of 14 C-labelled fatty acids into lipids of rhesus erythrocytes and *Plasmodium knowlesi in vitro*. *Comp Biochem Physiol* 40, 893–906. http://dx.doi.org/10.1016/0300-9629(71)90278-7.

Rodriguez-Concepcion, M. (2004). The MEP pathway: a new target for the development of herbicides, antibiotics and antimalarial drugs. *Curr Pharm Des* 10(19), 2391-2400.

Rohdich, F., Wungsintaweekul, J., Fellermeier, M., Sagner, S., Herz, S., Kis, K., and Zenk, M. H. (1999). Cytidine 5′-triphosphate-dependent biosynthesis of isoprenoids: YgbP protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-C-methylerythritol. *Proc Natl Acad Sci USA* 96(21), 11758-11763.

Rohdich, F., Wungsintaweekul, J., Luttgen, H., Fischer, M., Eisenreich, W., Schuhr, C. A., Fellermeier, M., Schramek, N., Zenk, M. H., and Bacher, A. (2000). Biosynthesis of terpenoids: 4-Diphosphocytidyl-2-C-methyl-D-erythritol kinase from tomato. *Proc Natl Acad Sci USA* 97, 8251-8256.

Rohdich, F., Eisenreich, W., Wungsintaweekul, J., Hecht, S., Schuhr, C. A., and Bacher, A. (2001). Biosynthesis of terpenoids. 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspF) from *Plasmodium falciparum*. *Eur J Biochem* 268, 3190-3197.

Rohmer, M. (1999). The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat Product Rep* 16(5), 565-574.

Röhrich, R. C., Englert, N., Troschke, K., Reichenberg, A., Hintz, M., Seeber, F., Balconi, E., Aliverti, A., Zanetti, G., and Köhler, U. (2005). Reconstitution of an apicoplast-localised electron transfer pathway involved in the isoprenoid biosynthesis of *Plasmodium falciparum*. *FEBS Lett* 579(28), 6433-6438.

Rosenson, R. S. (2001). Pluripotential mechanisms of cardioprotection with HMG-CoA reductase inhibitor therapy. *Am J Cardiovasc Drugs* 1(6), 411-420.

Rychlik, W. (2007). OLIGO 7 Primer Analysis Software, in Methods in Molecular Biology Vol. 402: PCR Primer Design; Ed. A. Yuryev; Humana Press Inc., Totowa, NJ. pp. 35-59.

Sacchettini, J. C., and Poulter, C. D. (1997). Creating isoprenoid diversity. *Science* 277, 1788-1789.

Šali, A., and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234, 779-815. doi:10.1006/jmbi.1993.1626

Sambrook, J., and Russell, D. W. (2001). Molecular Cloning A Laboratory manual 3rd ed., Cold Spring Harbor Laboratory Press, NY, USA.

Sambrook, J., Fritschi, E. F., and Maniatis, T. (1989). Molecular cloning: a laboratorymanual, Cold Spring Harbor Laboratory Press, NY, USA.

Sato, S., Clough, B., Coates, L., and Wilson, R. J. M. (2004). Enzymes for heme biosynthesis are found in both the mitochondrion and plastid of the malaria parasite *Plasmodium falciparum*. *Protist* 155(1), 117-125.

Saxena, V., Garg, S., Ranjan, S., Kochar, D., Ranjan, A., and Das, A. (2007). Analysis of elongation factor Tu (tuf A) of apicoplast from Indian *Plasmodium vivax* isolates. *Infect Genet Evol* 7, 618-626. doi:10.1016/j.meegid.2007.05.012

Saxena, V., Garg, S., Tripathi, J., Sharma, S., Pakalapati, D., Subudhi, A. K., Boopathi, P. A., Saggu, G. S., Kochar, S. K., Kochar, D. K., Das, A. (2012). *Plasmodium vivax* apicoplast genome: A comparative analysis of major genes from Indian field isolates. *Acta Tropica* 122, 138-149. doi:10.1016/j.actatropica.2012.01.007

Seemann, M., Bui, B. T. S., Wolff, M., Tritsch, D., Campos, N., Boronat, A., Marquet, A., and Rohmer, M. (2002). Isoprenoid Biosynthesis through the Methylerythritol Phosphate Pathway: The (E)-4-Hydroxy-3-methylbut-2-enyl Diphosphate Synthase (GcpE) is a [4Fe–4S] Protein. *Angew Chem Int Ed* 114, 4513-4515. doi:10.1002/1521-3773

Seemann, M., Wegner, P., Schunemann, V., Bui, B. T., Wolff, M., et al. (2005). Isoprenoid biosynthesis in chloroplasts via the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) from *Arabidopsis thaliana* is a [4Fe-4S] protein. *J Biol Inorg Chem* 10, 131-37

Seemann, M., Bui, B. T. S., Wolff, M., Maslow, M. M., and Rohmer, M. (2006). Isoprenoid biosynthesis in plant chloroplasts via the MEP pathway: Direct thylakoid/ferredoxin-dependent photoreduction of GcpE/IspG. *FEBS Lett* 580, 1547-1552.

Sekar, V., (1987). A Rapid Screening Procedure for the Identification of Recombinant Bacterial Clone. *BioTech* 5(1), 11-13.

Shan, S., Chen, X., Liu, T., Zhao, H., Rao, Z., and Lou, Z. (2011). Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from *Mycobacterium tuberculosis*. *FASEB J* 25(5), 1577-1584.

Sharma, A. K., Pallesen, L. J., Spang, R. J., and Walden, W. E. (2010). Cytosolic iron-sulfur cluster assembly (CIA) system: factors, mechanism, and relevance to cellular iron regulation. *J Biol Chem* 285(35), 26745-26751.

Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A., and Bucher, P. (2002). PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform* 3, 265-274. doi:10.1093/bib/3.3.265

Siegel, S. M. (1965). A Direct Microdetermination for Sulfide. *Anal Biochem* 11, 126-132.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., and Thompson, J. D. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Sys Biol* 7(1) 539.

Singh, B., Sung, L. K., Matusop, A., Radhakrishnan, A., Shamsul, S. S., Cox-Singh, J., ansd Conway, D. J. (2004). A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *The Lancet* 363(9414), 1017-1024.

Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nuc Acids Res* 33, W244-W248. doi:10.1093/nar/gki408

Sparr, C., Purkayastha, N., Kolesinska, B., Gengenbacher, M., Amulic, B., and Matuschewski, K. (2013). Improved efficacy of fosmidomycin against *Plasmodium* and *Mycobacterium* species by combination with the cell-penetrating peptide octaarginine. *Antimicrob Agents Chemother* 57, 4689-4698.

Spangenberg, T., Burrows, J. N., Kowalczyk, P., McDonald, S., Wells, T. N., and Willis, P. (2013). The open access malaria box: a drug discovery catalyst for neglected diseases. *PloS ONE* 8(6), e62906.

Sprenger, G. A., Schorken, U., Wiegert, T., Grolle, S., Graff, A. A. D., Taylor, S. V., Begley, T. P., Bringer-Meyer, S., and Sahm, H. (1997). Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proc Natl Acad Sci USA* 94, 12857-12862.

Steinbacher, S., Kaiser, J., Eisenreich, W., Huber, R., Bacher, A., and Rohdich, F. (2003). Structural Basis of Fosmidomycin Action Revealed by the Complex with 2-*C*-Methyl-D-erythritol 4-phosphate Synthase (IspC). *J Biol Chem* 278, 18401-18407.

Steinbacher, S., Kaiser, J., Wungsintaweekul, J., Hecht, S., Eisenreich, W., Gerhardt, S., and Rohdich, F. (2002). Structure of 2C-methyl-D-erythritol-2, 4-cyclodiphosphate synthase involved in mevalonate-independent biosynthesis of isoprenoids. *J Mol Biol* 316(1), 79-88.

Stanway, R. R., Witt, T., Zobiak, B., Aepfelbacher, M., and Heussler, V. T. (2009). GFP-targeting allows visualization of the apicoplast throughout the life cycle of live malaria parasites. *Biology of the Cell* 101(7), 415-435.

Stoops, J. K., Arslanian, M. J., Oh, Y. H., Aune, K. C., Vanaman, T. C., and Wakil, S. J. (1975). Presence of two polypeptide chains comprising fatty acid synthetase. *Proc Natl Acad Sci USA* 72(5), 1940-1944.

Ta, T. H., Hisam, S., Lanza, M., Jiram, A. I., Ismail, N., and Rubio, J. M. (2014). First case of a naturally acquired human infection with *Plasmodium cynomolgi*. *Malar J* 13(1), 68.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30, 2725-2729. doi:10.1093/molbev/mst197

Thomsen, R., and Christensen, M. H. (2006). MolDock: a new technique for high-accuracy molecular docking. *J Med Chem* 49, 3315-3321. doi:10.1021/jm051197e

Thelemann, J., Illarionov, B., Barylyuk, K., Geist, J., Kirchmair, J., Schneider, P., and Witschel, M. (2015). Aryl Bis-Sulfonamide Inhibitors of IspF from *Arabidopsis thaliana* and *Plasmodium falciparum*. *Chem Med Chem* 10(12), 2090-2098.

Trigg, P. I. (1968). Sterol metabolism of *Plasmodium knowlesi in vitro*. *Ann Trop Med Parasitol* 62:481–487.

Tonhosolo, R., D'Alexandri, F. L., de Rosso, V. V., Gazarini, M. L., Matsumura M. Y., Peres, V. J., Merino, E. F., Carlton, J. M., Wunderlich, G., Mercadante, A. Z., et al. (2009). Carotenoid biosynthesis in intraerythrocytic stages of *Plasmodium falciparum*. *J Biol Chem* 284:9974–9985.

Tonhosolo, R., D'Alexandri, F. L., Genta, F. A., Wunderlich, G., Gozzo, F. C., Eberlin, M. N., Peres, V. J., Kimura, E. A., and Katzin, A. M. (2005). Identification, molecular cloning and functional characterization of an octaprenyl pyrophosphate synthase in intra-erythrocytic stages of *Plasmodium falciparum*. *Biochem J* 392, 117–126.

Tsang, A., Seidle, H., Jawaid, S., Zhou, W., Smith, C., and Couch, R. D. (2011). Francisella tularensis 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase: kinetic characterization and phosphoregulation. *PloS ONE* 6(6), e20884.

Umeda, T., Kusakabe, Y., Sakamoto, Y., Odanaka, Y., Matsubayashi, S., Kitagawa, Y., and Tanaka, N. Crystal structures of 1-deoxy-D-xylulose 5-phosphate reductoisomerase from *Plasmodium falciparum* complexed with fosmidomycin analogs. Photon Factory activity Report 2015, 33.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., and Rozen, S. G. (2012). Primer3- new capabilities and interfaces. *Nuc Acids Res* 40(15), e115-e115.

Vanderberg, J. P., and Frevert, U. (2004). Intravital microscopy demonstrating antibody-mediated immobilisation of *Plasmodium berghei* sporozoites injected into skin by mosquitoes. *Int J Parasitol* 34(9), 991-996.

van Dooren, G. G., Marti, M., Tonkin, C. J., Stimmler, L. M., Cowman, A. F., and McFadden, G. I. (2005). Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of *Plasmodium falciparum*. *Mol Microbiol* 57, 405-419.

van Dooren, G. G., Su, V., DiOmbrain, M. C., and McFadden, G. I. (2002). Processing of an apicoplast leader sequence in *Plasmodium falciparum*, and the identification of a putative leader cleavage enzyme. *J Biol Chem* 277, 23612-23619.

Van Hoof, S., Lacey, C. J., Röhrich, R. C., Wiesner, J., Jomaa, H., and Van Calenbergh, S. (2008). Synthesis of analogues of (E)-1-hydroxy-2-methylbut-2-enyl 4-diphosphate, an isoprenoid precursor and human γδ T cell activator. *J Org Chem* 73(4), 1365-1370.

Van Schaijk, B. C., Kumar, T. S., Vos, M. W., Richman, A., van Gemert, G. J., Li, T., Eappen, A. G., Williamson, K. C., Morahan, B. J., et al. (2014). Type II fatty acid biosynthesis is essential for *Plasmodium falciparum* sporozoite development in the midgut of Anopheles mosquitoes. *Eukaryot Cell* 13(5), 550-559.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *J Comput Chem* 26, 1701-1718. doi:10.1002/jcc.20291

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J Mol Graph* 8, 52-56.

Vaughan, A. M., O'Neill, M. T., Tarun, A. S., Camargo, N., Phuong, T. M., Aly, A. S., Cowman, A. F., and Kappe, S. H. (2009). Type II Fatty Acid Synthesis is Essential only for Malaria Parasite Late Liver Stage Development. *Cell Microbiol* 11(3), 506-20.

Vial, H. J., Philippot, J. R, and Wallach, D. F. H. (1984). A reevaluation of the status of cholesterol in erythrocytes infected by *Plasmodium knowlesi* and *P. falciparum*. *Mol Biochem Parasitol* 13:53-65.

Wada, T., Kuzuyama, T., Satoh, S., Kuramitsu, S., Yokoyama, S., Unzai, S., and Park, S. Y. (2003). Crystal structure of 4-(cytidine 5′-diphospho)-2-C-methyl-d-erythritol kinase, an enzyme in the non-mevalonate pathway of isoprenoid synthesis. *J Biol Chem* 278, 30022-30027.

Waller, R. F., Kelling, P. J., Donald, R. G. K., Striepen, B., Handman, E., et al. (1998) Nuclear encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. *Proc Natl Acad Sci USA* 95, 12352–12357.

Wang, W., and Oldfield, E. (2014). Bioorganometallic Chemistry with IspG and IspH: Structure, Function, and Inhibition of the [Fe_4S_4] Proteins Involved in Isoprenoid Biosynthesis. *Angew Chem Int Ed* 53, 4294-4310. doi:10.1002/anie.201306712

Weatherby, K., and Carter, D. (2013). Chromera velia: The Missing in the Evolution of Parasitism. *Adv Appl Microbiol* 85, 119.

Wells, T. N., Burrows, J. N., and Baird, J. K. (2010). Targeting the hypnozoite reservoir of *Plasmodium vivax*: the hidden obstacle to malaria elimination. *Trends Parasitol* 26(3), 145-151.

White, N. J. (2008). *Plasmodium knowlesi*: the fifth human malaria parasite. *Clin Infec Dis* 46(2), 172-173.

Wiley, J. D., Merino, E. F., Krai, P. M., McLean, K. J., Tripathi, A. K., Vega-Rodríguez, J., Jacobs-Lorena, M., Klemba, M., and Cassera, M. B. (2015). Isoprenoid precursor biosynthesis is the essential metabolic role of the apicoplast during gametocytogenesis in *Plasmodium falciparum*. *Eukaryot Cell* 14, 128-139. doi:10.1128/EC.00198-14.

Wilson, R. J. (2005). Parasite plastids: approaching the endgame. *Biol Rev Camb Philos Soc* 80(1): 129-153

Wilson, R. J. M., Denny, P. W., Preiser, P. R., Rangachari, K., Roberts, K., Roy, A., Whyte, A., and Williamson, D. H. (1996). Complete Gene Map of the Plastid-like DNA of the Malaria Parasite *Plasmodium falciparum*. *J Mol Biol* 261(2), 155-172.

Witschel, M. C., Hçffken, H. W., Seet, M., Parra, L., Mietzner, T., Thater, F., Niggeweg, R., Rçhl, F., Illarionov, B., et al. (2011). Inhibitors of the Herbicidal Target IspD: Allosteric Site Binding. *Angew Chem* 50, 7931-7935.

Wongsrichanalai, C., Pickard, A. L., Wernsdorfer, W. H., and Meshnick, S. R. (2002). Epidemiology of drug-resistant malaria. *Lancet Infec Dis* 2(4), 209-218.

World Health Organization, World Malaria Report, 2014.

Wu, W., Herrera, Z., Ebert, D., Baska, K., Cho, S. H., DeRisi, J. L., and Yeh, E. (2015). A chemical rescue screen identifies a *Plasmodium falciparum* apicoplast inhibitor targeting MEP isoprenoid precursor biosynthesis. *Antimicrob Agents Chemother* 59(1), 356-364.

Xiao, Y., Savchenko, T., Baidoo, E. E., Chehab, W. E., Hayden, D. M., Tolstikov, V., and Dehesh, K. (2012). Retrograde signaling by the plastidial metabolite MEcPP regulates expression of nuclear stress-response genes. *Cell* 149(7), 1525-1535.

Xiao, Y., Zahariou, G., Sanakis, Y., and Liu, P. (2009). IspG Enzyme activity in the deoxyxylulose phosphate pathway: Roles of the iron-sulfur cluster. *Biochemistry* 48(44), 10483-10485.

Yamauchi, L. M., Coppi, A., Snounou, G., and Sinnis, P. (2007). *Plasmodium* sporozoites trickle out of the injection site. *Cell Microbiol* 9(5), 1215-1222.

Yang, J., Roy, A., and Zhang, Y. (2013). BioLiP: a semi-manually curated database for biologically relevant ligand–protein interactions. *Nuc Acids Res* 41, 1096-1103.

Yao, Z. K., Krai, P. M., Merino, E. F., Simpson, M. E., Slebodnick, C., Cassera, M. B., and Carlier, P. R. (2015). Determination of the active stereoisomer of the MEP pathway-targeting antimalarial agent MMV008138, and initial structure-activity studies. *Bioorg Med Chem Lett* 25(7), 1515-1519.

Yeh, E., and DeRisi, J. L. (2011). Chemical Rescue of Malaria Parasites Lacking an Apicoplast Defines Organelle Function in Blood-Stage *Plasmodium falciparum*. *PLoS Biol* 9, e1001138.

Zhang, B., Watts, K. M., Hodge, D., Kemp, L. M., Hunstad, D. A., Hicks, L. M., and Odom, A. R. (2011). A second target of the antimalarial and antibacterial agent fosmidomycin revealed by cellular metabolic profiling. *Biochemistry* 50, 3570-3577.

Zepeck, F., Gräwert, T., Kaiser, J., Schramek, N., Eisenreich, W., Bacher, A., and Rohdich, F. (2005). Biosynthesis of isoprenoids. Purification and properties of IspG protein from *Escherichia coli*. *J Org Chem* 70(23), 9168-9174.

Zuegge, J., Ralph, S., Schmuker, M., McFadden, G. I., and Schneider, G. (2001). Deciphering apicoplast targeting signals - feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. *Gene* 280, 19-26.

Appendix

List of Publications

1. **Gagandeep Singh Saggu**, Shilpi Garg, Zarna Rajeshkumar Pala, Sanjay Kumar Kochar, Dhanpat Kumar Kochar, Vishal Saxena. Characterization of 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (GcpE) from *Plasmodium vivax* and its interaction with fosmidomycin. International Journal of Biological Macromolecules. (2017) 96: 466-473.
2. **Gagandeep Singh Saggu**, Zarna Rajeshkumar Pala, Shilpi Garg, Vishal Saxena. New insight in isoprenoids biosynthesis process and future prospects for drug designing in *Plasmodium*. Frontiers in Microbiology. (2016) **7**:1421. http://dx.doi:10.3389/fmicb.2016.01421
3. Zarna R Pala, Vishal Saxena, **Gagandeep S Saggu**, Sushil K Yadav, Rajendra P Pareek, Sanjay K Kochar, Dhanpat K Kochar, Shilpi Garg. Structural and Functional Characterization of an Iron-Sulfur Cluster Assembly Scaffold Protein-SufA from *Plasmodium vivax*. *Gene* 585 (2016) 159-165. http://dx.doi.org/10.1016/j.gene.2016.03.041
4. **Gagandeep Singh Saggu**, Divya Shrivastava. Study on Thermophiles reveals the presence of *Actinobacillus lignieresii* in cattle compost. *Inter J Appl Sci Biotechnol* 1 (2013) 33-41. http://dx.doi.org/10.3126/ijasbt.v1i2.8200
5. Megha Sharma, Kanchan Soni, **Gagandeep Singh Saggu**. Extra Cellular Enzyme Production by *Actinobacillus*: a thermophile present in cattle compost. *Inter J Curr Microbiol Appl Sci*. (2014) 3(8) 1081-1094. http://www.ijcmas.com/vol-3-10/Megha%20Sharma,%20et%20al.pdf
6. Vishal Saxena, Shilpi Garg, Jyotsna Thripathi, Sonal Sharma, Deepal Palakpati, Amit K. Subudhi, P.A. Boopathi, **Gagandeep S. Saggu**, Dhanpat K. Kochar, Sanjay K. Kochar, Ashis Das. *Plasmodium vivax* apicoplast genome: A comparitive analysis of major genes from Indian field isolates. *Acta Tropica* 122 (2012) 138-149. http://dx.doi.org/10.1016/j.actatropica.2012.01.007

Details of Conferences Attended

Oral Presentations:

1. **Gagandeep S**. **Saggu**, Shilpi Garg, Zarna Pala, Sanjay K. Kochar, Vishal Saxena. 2014. "Characterization of 4-Hydroxy-3-Methylbut-2En-1Yl Diphosphate Synthase (GcpE) enzyme from *Plasmodium vivax*" in 8th National Congress of Indian Academy of Tropical Parasitology (TROPACON 2014) at Regional Medical Research Centre (RMRC), Dibrugarh, from November 20 – 22.
2. **Gagandeep S.** Saggu, Shilpi Garg, Zarna Pala, Sanjay K. Kochar, Vishal Saxena. 2014. "Characterization of Isoprenoids biosynthesis pathway enzymes from Indian *Plasmodium vivax* field isolates". in Recent Trends and Future Prospects of Microbiology and Biotechnology organised by Shri J. J. T. University, Jhunjhunu, Rajasthan, from March 03 – 04. (**Received Young Scientist Award for Best Presentation)**

Poster Presentations:

3. **Gagandeep S. Saggu**, Shilpi Garg, Zarna Pala, Sushil K. Yadav, Sanjay K. Kochar, Dhanpat K. Kochar, Vishal Saxena. 2016 "Characterization of 4-Hydroxy-3-Methylbut-2-en-1-yl diphosphate Synthase (GcpE) enzyme from *Plasmodium vivax*" in BITS conference on Gene and Genome Regulation (BCGGR 2016) at Department of Biological Sciences, Birla Institute of Technology and Sciences, Pilani, India from Feb 18-20.
4. Zarna Pala, Shilpi Garg, **Gagandeep S. Saggu**, Sanjay K. Kochar, Dhanpat K. Kochar, Vishal Saxena. 2014 "Characterization of a nuclear encoded hypothetical protein involved in Fe-S cluster biogenesis in *Plasmodium vivax*" in 25th National Conference of Parasitology on "Global Challenges in the Management of Parasitic Diseases" organised jointly by CSIR-CDRI and Indian Society of Parasitologists at CDRI, Lucknow from October 16 – 18.
5. Zarna R. Pala, Shilpi Garg, **Gagandeep S. Saggu**, Vishal Saxena. 2014. "Characterization of nuclear encoded hypothetical proteins involved in *Plasmodium vivax* apicoplast metabolic pathways" in Recent Trends and Future Prospects of Microbiology and Biotechnology organized by Shri J. J. T. University, Jhunjhunu, Rajasthan, from March 03 – 04.
6. **Gagandeep S. Saggu**, Zarna Pala, Shilpi Garg, Sanjay K. Kochar, Vishal Saxena. 2013. "Characterization of IspD enzyme from Indian *Plasmodium vivax* field isolates" in 24th National Congress of Parasitology held at Regional Medical Research Centre for Tribal's, Jabalpur from April 27 – 29.
7. Zarna Pala, Ritika Lakhotia, Shilpi Garg, **Gagandeep S. Saggu**, Vishal Saxena. 2013. "Characterization and analysis of the RimM protein encoding gene in *Plasmodium vivax*"

in National Conference on Fight Against Malaria: Prospects and Perspectives held at Centre for Biotechnology & Centre for Medical Biotechnology, M. D. University, Rohtak, Haryana on March 9th.

8. Ritika Lakhotia, Shilpi Garg, Zarna Pala, **Gagandeep S. Saggu**, Akanksha Gupta, Vishal Saxena. 2012. Characterization of the RimM Protein encoding gene in *P. vivax* in International Conference on Microbial, Plant and Research at M. I. T. S., Sikar, Rajasthan, from March 29 – 31.
9. **Gagandeep S. Saggu**, Ritika Lakhotia, Akansha Gupta, Shilpi Garg, **Vishal Saxena**. 2011. "Analysis of apicoplast targeted hypothetical proteins in *Plasmodium vivax*" in Contemporary Trends in Biological and Pharmaceutical Research, CTBPR– 2011 held at Birla Institute of Technology and Science, Pilani, India from March 12-13.

Conference & Workshop Attended

1. Molecular Modeling and Drug Design (2012), National workshop conducted by Birla Institute of Scientific Research, Jaipur, Rajasthan, India from Dec 14-16.
2. Search for Antimalarial: Mechanism Based Approach (2012), National conference conducted by Jawaharlal Nehru University Delhi, India from April 27-29.
3. Introductory Bioinformatics, its Applications and Genome Analysis (2012), National Workshop conducted by National Institute of Cholera and Enteric Diseases, ICMR, India from April 9-10.

Biography

Brief biography of Prof. Vishal Saxena, Ph.D.

Prof. Vishal Saxena is currently working as an Associate Professor in Molecular Parasitology & Systems Biology Lab, Department of Biological Sciences, Birla Institute of Technology and Science, Pilani, Rajasthan, India. He joined the Dept as an Assistant Professor in 2006. Prior to this, Prof. Saxena did B. Pharmacy from University of Rajasthan, Jaipur (August 1999) followed by M. E. Biotechnology (June 2001) and Ph. D. (October 2006) from Biological Sciences Group, BITS, Pilani. He has over 16 years (6 year's Pre-doctoral experiences and about 10 year's job tenure) of teaching and research experience. His major thrust area of research is Molecular Biology and Immunology with special emphasis to Genomics and Proteomics of Malaria parasites, *Plasmodium vivax* and *P. falciparum*. His group is focusing on various aspects related to *P. vivax* infections in humans, apicoplast and its genome, metabolic pathways functional in the apicoplast, hypothetical proteins encoded by *P. vivax* nuclear genome. He was a Visiting Scholar at Department of Public Health, College of Global Health, University of South Florida, Tampa, Florida, USA from June-July 2015. He is actively involved in teaching and research, has handled projects from various funding agencies, has supervised many graduate and post-graduate student thesis and is currently supervising 5 PhD students. He was a recipient of **Young Scientist Award** for Best Poster Presentation at International Conference on Molecular Epidemiology & Immunology of Malaria and other vector Borne Diseases at RMRCT, Jabalpur, M. P., INDIA, 2007. He has published research papers in international journals of repute with good impact and has authored a book on Genetic Engineering.

Biography of Mr. Gagandeep Singh Saggu

Gagandeep Singh Saggu completed his B.Sc. from S.G.N. Khalsa PG College, Sri Ganganagar affiliated to University of Bikaner, Bikaner Rajasthan and M.Sc. in Microbiology from Jaipur National University (JNU), Jaipur, Rajasthan. During his Masters, he had successfully completed a project entitled "Biochemical and Molecular characterization of Thermophiles from compost samples" under the supervision of Prof. Divya Shrivastava, Deputy Director, JNU, Jaipur. During his M.Sc., he qualified Graduate Aptitude Test in Engineering (GATE) in 2008 (92.5 percentile), State Level Eligibility Test for Lectureship (SET) in 2009 and ICMR Project Assistant Fellowship in 2009. Following his masters, he worked as a Process Associate Manager in Lead processing unit of Hindustan Zinc Ltd. (Vedanta Group) Udaipur, Rajasthan. Subsequently, he joined the Department of Biological Sciences, BITS, Pilani, Pilani Campus as a Project Assistant in a DST funded project of Prof. Vishal Saxena and was also enrolled as a PhD student (2010). He was awarded Junior Research Fellowship (Dec 2010) from University Grant Commission, Govt. of India to pursue Ph.D. He has published research articles in journals of international repute and has presented his research work in national and international conferences and workshops. He was a recipient of **Young Scientist Award** for Best Oral Presentation at Sri JJT University, Jhunjhunu. His current research interest includes functional characterization of Isoprenoids and other biosynthesis pathways functional in the apicoplast of human malaria parasites.