Chapter 1

Introduction

Malaria remains one of the most serious infectious diseases in the world, inflicting acute illness on more than 300 million people and leading to more than two million deaths annually. Human malaria is caused by four species of the apicomplexan parasite of genus *Plasmodium* viz., *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae* and *Plasmodium ovale*. Among these, *Plasmodium vivax* accounts for over half of all malaria infections outside Africa and an estimated 75 million acute episodes per year in Africa (Sina, 2002). *P. vivax* malaria was first identified by Golgi in 1885. In 1948, Short, Garnham, Covell and Shute identified tissue forms of *P. vivax* in the liver. In 1982, almost 100 years after the discovery of the parasite, Bray and Garnham proposed that some *P.vivax* sporozoites in the liver remain latent (hypnozoites), causing severe relapses later on (http://www.malariasite.com/malaria/History.htm). *P. vivax* malaria is considered to have originated in Southeast Asia and is believed to have diverged some 2 – 3 million years ago to different parts of the world including Africa and Europe (Carter, 2003). Outside tropical Africa, in Asia and Western Pacific and in Central and South America, *P. vivax* is most prevalent among all human malaria parasites (Mendis, 2002). *Plasmodium vivax* malaria constitutes at least 60 – 70% of total malaria cases in India, with pronounced morbidity particularly in the economically weaker sections of the society (Adak *et. al.*, 1998).

Plasmodium vivax Malaria

P. vivax produces a severe, fulminant illness that can relapse months after the original infection as dormant parasites are released from the liver. Many of the severe human malarial pathologies such as acute respiratory pathogenesis, severe anemia, cerebral malaria, etc., generally observed in *P. falciparum*, are now being associated with *P. vivax* malaria (Kochar *et. al.*, 2005). *P. vivax* induces anemia in young children even at low parasitemias and is also associated with an increased risk of low-birth-weight

babies (Sina, 2002). In India, *P. vivax* has been traditionally considered to be the dominant species with 1.2 – 1.5 million new cases reported annually, accounting for approximately 60 – 70% of total malaria cases (Dua *et. al.*, 1996). In recent years, reports on increase in *P. vivax* resistance towards malarial drugs have raised an issue of concern.

The *P. vivax* genome sequencing project is underway (Carlton, 2003) and various nuclear genes of the parasite are being unraveled. Despite this, the work done on *Plasmodium vivax* in varying aspects is not much as compared to other *Plasmodium* species. The main reason behind this is the ease of maintaining *in – vitro* or *in – vivo* cultures of other *Plasmodium* species such as *Plasmodium falciparum* (which requires only mature human red blood cells) or *P. berghei* (a rodent parasite). *P. vivax* preferentially invades human reticulocytes, immature red blood cells that are positive for the Duffy blood group antigen. The concentration of reticulocytes in normal, peripheral blood (0.5–1.5%) is not sufficient to sustain *P. vivax in vitro*, even with periodic replenishment. Therefore, to successfully recycle generations of *P. vivax in vitro*, it is necessary to add to cultures elevated concentrations of Duffy positive reticulocytes. Short-term cultures of *P. vivax* have been reported with the help of high haematocrit value (> 2% upto 10%) RBCs. As yet there are no reports stating a long term simplified procedure for *P. vivax in – vitro* culture (Golenda *et. al.*, 1997).

Drugs against Malaria

The fight against *P. vivax* malaria began some 100 years back. During this period, various drug molecules have been formulated in way of trials to combat malaria. Sadly the cheapest and most effective first line chemotherapies like chloroquine, pyrimethamine, etc. used to fight malaria for decades are now losing efficacy due to emergence of drug resistance in the parasite population. This has led to the resurgence of the disease, with malaria mortality rates redoubling in many areas. Even continuous trials to develop a successful vaccine against malaria have been impeded due to the antigenic variations shown by the parasite. Clearly, there is a need for new antimalarials or identification and exploitation of new therapeutic targets. Many of the more exciting new targets to be revealed by the *P. falciparum* genome project are

enzymes from the so-called Apicoplast — a relict plastid (or chloroplast) that is a legacy of the malaria parasite's distant photosynthetic ancestry. This cyanobacterial heritage of the apicoplast means that many of its bacteria-like enzymes are fundamentally different from the mammalian host equivalents, making them potential drug targets (Ralph *et. al*., 2004).

Apicoplast

A three – four membrane organelle was reported some 30 years back in the genus *Plasmodium*. The organelle was named Apicoplast, due to its presence in Apicomplexan parasites and possession of a genome similar to the chloroplast DNA. The organelle is believed to have a blue – green algal origin and is postulated to have appeared in the parasites due to secondary endosymbiosis (Figure 1.1).

According to the latest hypothesis on the origin of Apicoplast, it has been proposed that a single endocytobiotic event approximately 1300 million years ago, initiated the formation of monophyletic group of chromalveolate protists including the cryptophytes, haptophytes, heterokonts and apicomplexan parasites. This generally well-supported argument proposes that a red algal cell engulfed by an ancestral alveolate protist was enslaved after the transfer of many of its genes to the host cell nucleus and their subsequent inheritance (Figure 1.2). Following these ancient secondary endosymbiotic events in the red and green lineages, various algal structures and functions became incorporated into chimaeric 'meta-algal' cells that underwent considerable diversification as they evolved. A nucleomorph, the remnant of a red algal nucleus still persists in cryptophytes whereas chlorarachniophytes have a green algal version. In apicomplexans, the visible remains of the secondarily acquired algal cell are restricted apparently to a multi-membraned plastid organelle. This is no longer photosynthetic and is bereft of much of its former genic content. Some apicomplexans (Cryptosporidia) may even have lost the plastid organelle altogether, although a store of laterally transferred genes from the algal symbiont has yet to be discounted in such cases (Wilson, 2005).

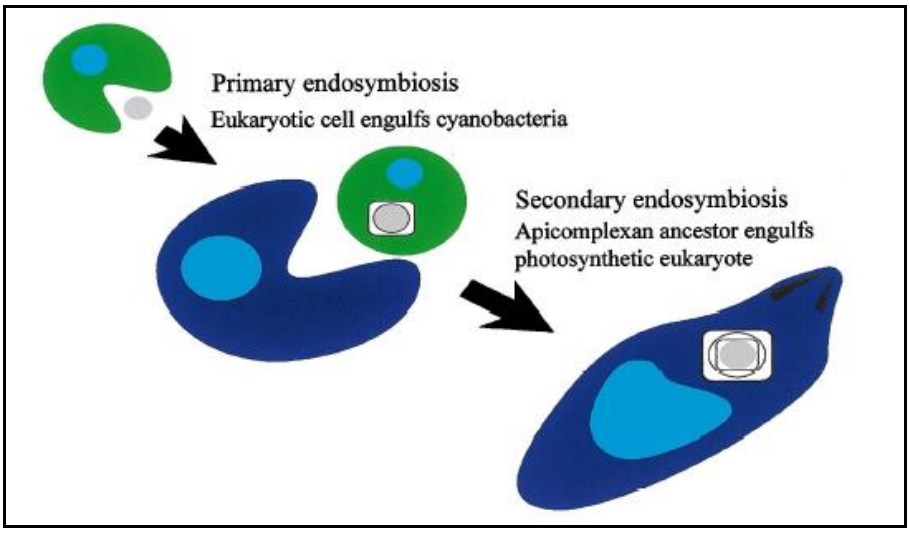

Figure 1.1: Diagrammatic representation of secondary endosymbiotic appearance of Apicoplast. (Ref: Gleeson, M. T., 2000; The plastid in Apicomplexa: what use is it?; Int. J. Parasitol. 30: 1053 – 1070)

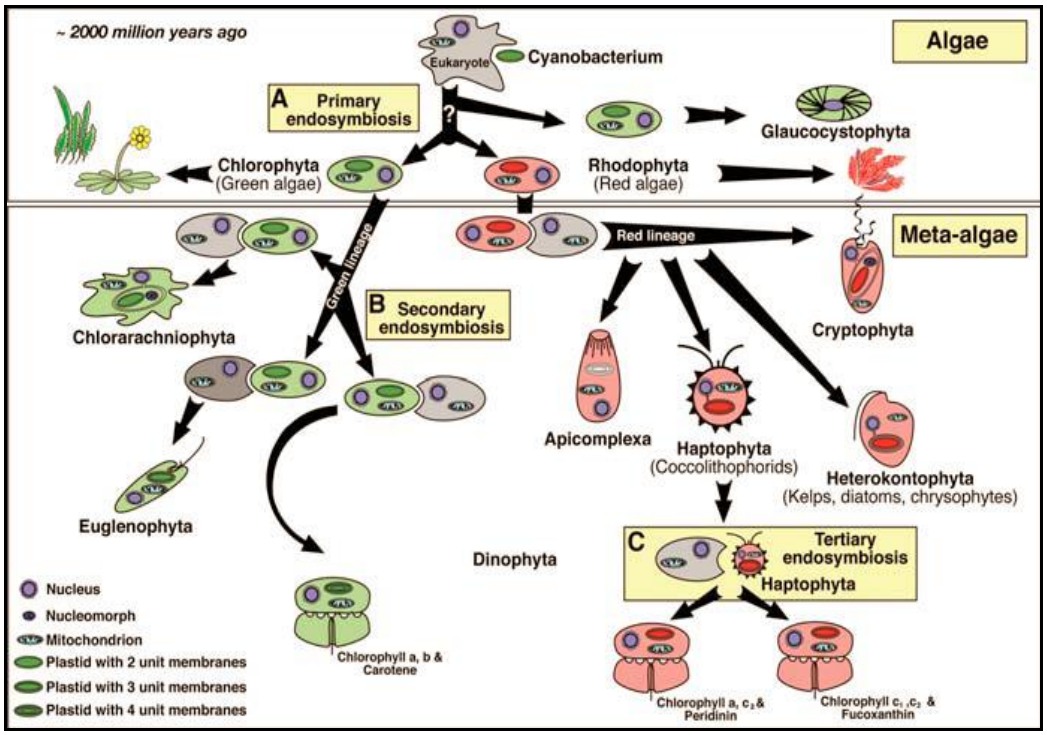

Figure 1.2: Hypothetical evolutionary scheme outlining sequential endocytobioses led from primary algae to meta-algae carrying secondary and even tertiary plastids. (Ref: Wilson, R. J. M., 2005; Parasite plastids: approaching the endgame, Biol. Rev. 80: 129 – 153).

The discovery of Apicoplast is related with the discovery of its circular DNA. In the initial findings, the plastid-like DNA (plDNA), (Figure 1.3), was first noted in *P. knowlesi* as a low-density minor satellite in CsCl gradients, and its role in the cell was tentatively suggested to be mitochondrial. Similar low density satellites were also reported in two rodent malaria parasites, *P. berghei* and *P. chabaudi*. Kilejian was the first person to "see" the plastid – DNA circles. In a pioneering electron microscopic study, she found them in extracts of isolated "mitochondria" of the avian parasite *P. lophurae*. They had a contour length of 10.3 mm (31 kb) and a buoyant density linking them to the previously noted satellites. She understandably assumed them to be of mitochondrial origin (Kilejian, 1975). Some similar studies following this in *P. berghei* and *Toxoplasma gondii* (another Apicomplexan) stated this circular DNA to be of mitochondrial origin but also raised concern on the large cruciform structures seen in the electron micrographs of the DNA (Mcfadden *et. al.*, 1997; Wilson *et. al.*, 1997).

The circular molecules were first isolated in usable quantity from the simian parasite *P. knowlesi* (Williamson *et. al.*, 1985). They were 11.5 mm in contour length and showed a cruciform structure just like that of *Toxoplasma*. Gardner and co – workers isolated the counterpart from the human parasite *P. falciparum* with an estimated size of 35 kb and substantiated the organellar origin of the molecules by obtaining a sequence corresponding to a fragment of a prokaryotic small-subunit (SSU) rRNA gene (Gardner *et. al.*, 1988; Gardner *et. al.*; 1991b, Gardner *et. al.*, 1993). Various genes and open reading frames similar to the genes of red and blue green algae were found in the 35 kb circular DNA. These genes helped researchers to relate the prokaryotic origin of the organelle and pointed to its plastid like provenance (Gardner *et. al.*, 1991a).

The Plastid DNA

Plastid genomes are similar to those of their bacterial progenitors in that they are super – coiled DNA circles. At 27–35 kb, the apicoplast genome is the smallest known plastid genome, but it appears to have retained a circular, supercoiled architecture. The identity and arrangement of the genes on the 35-kb circle

overwhelmingly favor its evolution from a plastid genome. There are about eight to fifteen copies of 35-kb circles per cell in *P. falciparum.* In other apicomplexans like *Toxoplasma gondii* the DNA molecule number goes upto 25 per cell. The *P. falciparum* plastid DNA circle is packed with genes; usually separated by only a few nucleotides and in some cases even have small overlaps. The sequences of most of the proteins encoded by this malarial circular DNA have diverged considerably from their counterparts in prokaryotes and other plastids. This is associated with considerable bias arising from the comparatively high A+T content of the circle (86%) compared with plastids from other sources. The plastid DNA contains over 60 genes which includes tRNA, rRNA, ribosomal protein, 7 unidentified ORF's and RNA polymerase genes. There are some specific genes encoding elongation factor EF – Tu, Caseinolytic protease ClpC, and SufB proteins. About one – third of the circle includes inverted repeat regions, IR_A and IR_B. The remaining two – third is a single copy region arbitrarily divided into IR_A sector and IR_B sector. A detailed map of the *P. falciparum* DNA is depicted (Figure 1.3).

Inverted Repeat

The inverted repeat (IR) has a total length of approximately 10.5 kb and encodes duplicated genes for the small subunit (*SSU*) (Gardner *et. al.*, 1991b) and large-subunit (*LSU*) *rRNAs* and nine different *tRNAs* (Gardner *et. al.*, 1994b; Preiser *et. al.*, 1995). Both forms of the *rRNA* genes in the IR are flanked by *tRNA* genes for *met* [M], *arg* [R], *val* [V], *arg* [R], *leu* [L], *asp* [N], *ala* [A], *Iso* [I] and *Thr* [T]. The IR extends for 36 nucleotides beyond the *tRNA* genes just downstream of the *LSU rRNA* genes and includes the first three codons of two different open reading frames (*ORF470*/ *SufB* and *rps4*). The disposition of the *rRNA* and *tRNA* genes in the IR differs in several respects from that of other recorded plastids. However, this order is conserved in *T. gondii*, suggesting that it may be an ancestral apicomplexan character derived from numerous rearrangements and deletions in an ancient progenitor of the apicomplexan clade (Wilson *et. al*., 1997).

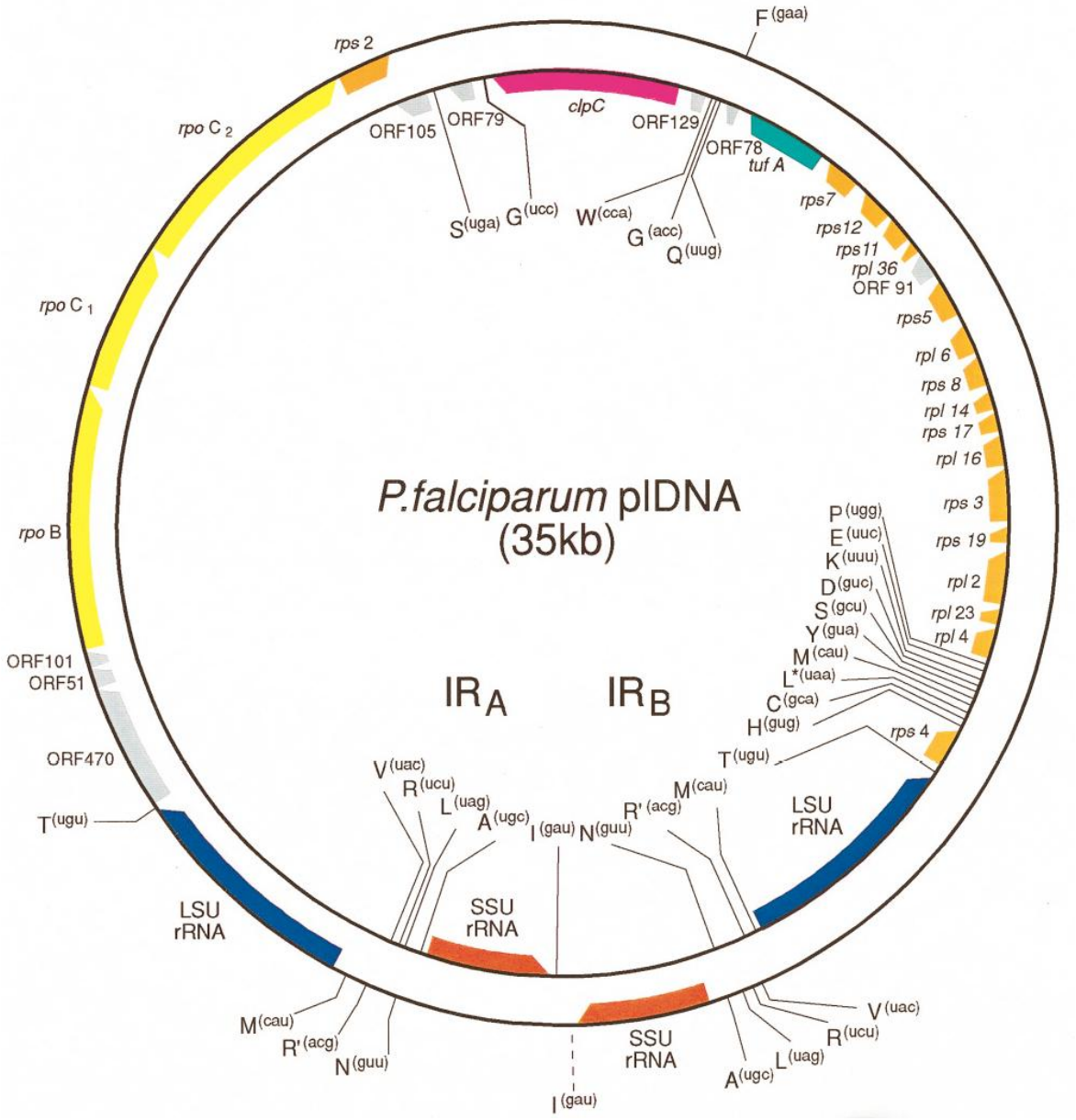

Figure 1.3: Gene map of the 35 kb circular DNA of *Plasmodium falciparum*. The two halves (A and B) of the inverted repeat (IR) are indicated. *tRNA* genes are specified by the anticodon, as well as the single letter amino acid code. ORFs specify the number of amino acid residues in the open reading frame. Genes on the outer strand are transcribed clockwise, those on the inner strand anti-clockwise. (Ref: Wilson, R. J. M., *et. al*., 1996; Complete gene map of the Plastid – like DNA of the malaria parasite *Plasmodium falciparum*, J. Mol. Biol. 261: 155 – 172).

Single-Copy Region (IR_A Sector)

Immediately downstream of the inverted repeat on the IR_A sector and encoded on the same DNA strand as the *LSU rRNA* and *tRNA* lay *ORF470*/ *SufB* which is understood to be a homologue gene encoding SufB protein essential for the growth of cyanobacteria *Synechocystis* species. It is proposed that the gene might be functioning as an aid in iron metabolism in the apicoplast. Downstream to *SufB/ORF470* are two unassigned ORFs (*ORF101* and *ORF51*). Immediately 3' to the above three ORFs on the same DNA strand are present three larger genes whose recognition was one of the first clues to the plastid ancestry of the circle. Designated as *rpoB*, *rpoC*1, and *rpoC*2, they encode subunits β, β', and β", respectively, of an RNA polymerase similar to that found in cyanobacteria and chloroplasts and typically not in mitochondria (Gardner *et. al.*, 1991a; Gardner *et. al.*, 1994a).

Subdivision of the *E. coli rpoC* equivalent into two genes, *rpoC*1 and *rpoC*2, is another plastid-like feature of the diminutive malarial circular genome. Phylogenetic analysis of the *rpoB* gene showed that it has diverged considerably from other plastid sequences. The plastid sequence of the protist *Euglena gracilis* is the closest. To date, only a small fragment of the malarial *rpoC*1 gene has been analyzed phylogenetically, and again the *Euglena* plastid sequence showed closest homology. However, the level of conservation of the predicted peptides encoded by *rpoB* and *rpoC* is much lower. For example, *rpoB* of *P. falciparum* encodes a peptide that has only 29% identity with the β subunit predicted from the plastid sequence of spinach. Despite this, all the known functional domains of the protein are readily recognized in the predicted malarial peptide, which resembles the plastid type much more than it resembles the bacterial type (Wilson *et. al.*, 1997).

As in the red alga *P. purpurea* but unlike in higher plants, the malarial *rpoC*1 gene does not have an intron. An intergenic region of 11 nucleotides separates *rpoC*1 from *rpoC*2, which codes for the largest subunit of the chloroplast-encoded RNA polymerase. A structural feature formed by about 600 amino acids encoded in the

central region of *rpoC*2 sequences from higher plants is absent from *E. coli* and the plastid sequences of both *E. gracilis* and *P. falciparum*. The red alga *P. purpurea* has a sequence corresponding to about half of it. As is typical of many plastid genomes, immediately downstream of *rpoC*2 lies the ribosomal protein gene *rps*2, whose sequence in the malarial case is not highly conserved. Downstream of *rps*2, lies coding switches to the complementary strand, with the direction of transcription pointing away from IR_B (Wilson *et. al*., 1997).

Single-Copy Region (IR_B Sector)

An ORF lying a few nucleotides 3' of the *tRNA*[T] gene marking the end of IR_B was identified by a homology search as the ribosomal protein gene *rps4*, although the level of predicted amino acid identity to its closest known counterpart, from *Marchantia polymorpha*, was only 27%. The malarial *rps4* sequence shares the same first three codons as *ORF470/ SufB* at the other end of the rDNA palindrome. Immediately downstream of *rps4* lies a cluster of 10 *tRNA* genes (*his* [H], *cys* [C], *leu* [L], *met* [M], *tyr* [Y], *ser* [S], *asp* [D], *lys* [K], *glu* [E] and *pro* [P]). The leucine gene holds the only intron so far recognized on the circle, located, as with other plastid homologs, within the anticodon. Downstream of the tRNA genes, a series of ORFs encode ribosomal proteins, arranged much as in other plastid genomes. The first ORF in the series has only a low level (19%) of predicted amino acid identity to *rpl4*, but the likelihood that this gene has been correctly identified is supported by the subsequent ordered series of ORFs corresponding to ribosomal protein genes like those encoded by the *S10*, *spc*, *alpha*, and *str* operons of *E. coli* and chloroplast genomes. Following *rpl4* on the malarial plDNA, are identified *rpl23*, *rpl2*, *rps19*, *rps3*, *rpl16*, and *rps17*, corresponding to the *S10* operon. None of the predicted peptides is highly conserved, the best being *rpl16* with 33% identity to its closest homolog, *Marchantia polymorpha* (Wilson *et. al*., 1997).

A frameshift and a single-base overlap with *rps17* lead into sequences corresponding to the *spc* operon. After *rpl14*, are identified *rps8*, *rpl6*, *rps5*, and the small but highly conserved ribosomal protein gene *rpl36* (*secX*). Between the last two genes lies an unidentified sequence (*ORF91*). Downstream of this *spc*-like operon, another

frameshift leads to *rps11*, a member of the *alpha* operon of *E. coli*. By contrast with the latter, *rpoA* has been deleted from the malarial circle, possibly transposed to the nucleus. A few nucleotides after *rps11* on the 35-kb circle lies a final pair of ribosomal protein genes, namely, *rps12* and *rps7*, corresponding to components of the *str* operon of *E. coli*. As in other algal genomes, these precede a *tufA* gene, which encodes the elongation factor Tu (EF-Tu), a G-protein crucial in the elongation step of protein synthesis. Three of the conserved functional domains in the predicted peptide lie in the N-terminal half of the protein and correspond to the GTP binding site. As in other plastid *tufA* genes, the malarial gene has an insertion which encodes a 10-amino-acid extended loop that may be involved in defining the tertiary structure of the GTP-binding domain. It is notable that *tufA* occurs on the plastid genome of many algae but not on that of higher plants (Wilson *et. al*., 1997).

Downstream of *tufA* lie another four tRNA genes: *phe* (F), *gln* (Q), *gly* (G), and *trp* (W). *tRNA*[F] is on the complementary strand and in this respect is distinct from almost all the other genes on the IR_B arm. Another short tentative ORF, *ORF129*, then leads to the final large ORF on the IR_B single copy region, provisionally identified as *clpC*, a member of the *hsp100* gene family; these genes encode ubiquitous heat shock or stress proteins that act as molecular chaperones with diverse functions. Once again, a corresponding gene is present on the plastid genome of the red alga *Porphyra purpurea* but not on the plastids of higher plants. The gene is also absent from the only other fully sequenced plastid genome available from a protist, that of *Euglena gracilis*. The malarial *clpC* gene is unusual in that only the second of the two ATP binding domains is conserved. In *Plasmodium falciparum* the N-terminal half of plastid *clpC* is atypical and thus its function is doubtful, but the protein is found better conserved in *T. gondii* (Wilson, 2002). Following the *clpC*-like gene, are present two tRNA genes encoding alternative codons for *gly* (G) and *ser* (S) which are separated by a short region of some 240 nt that contains an unassigned ORF (*ORF79*). Downstream of *tRNA*[S], the 3' end of another short potential ORF (*ORF105*) overlaps the *rps2* gene on the opposite strand by a few nucleotides, marking the transcription crossover point with the IR_A sector (Wilson *et. al.*, 1997; Wilson *et. al*., 2002).

Table 1.1: Gene content of the 35-kb circular DNA of *P. falciparum*

Class	Genes
rRNA	16S, 23S
tRNA[a,b]	A_{UGC} C_{GCA} D_{GUC} E_{UUC} F_{GAA} G_{ACC} G_{UCC} H_{GUG} I_{GAU} K_{UUU}
	L_{UAG} L_{UAA}[b] M_{CAU} M_{CAU} N_{GUU} P_{UGG} Q_{UUG} R_{UCU} R_{ACG} S_{GCU} S_{UGA} T_{UGU} V_{UAC} W_{CCA} Y_{GUA}
Ribosomal proteins	*rps* 2, 3, 4, 5, 7, 8, 11, 12, 17, 19
	rpl 2, 4, 6, 14, 16, 23, 36
RNA polymerase	*rpoB, rpoC*$_1$, *rpoC*$_2$
Other proteins	*clpC, tufA,* ORF470/ *SufB*
Unassigned ORFs	51, 78, 79, 91, 101, 105, 129

a Single-letter amino acid code and anticodon; b Intron. (Ref: Wilson, R. J. M., *et. al*., 1996; Complete gene map of the Plastid – like DNA of the malaria parasite *Plasmodium falciparum*, J. Mol. Biol. 261: 155 – 172).

Apicoplast: Division and Functions

Most of our understanding of plastid division in the Apicomplexa to date comes from morphological and ultrastructural observations. Since the plastid in Apicomplexan parasites is not pigmented, GFP has been used in transgenic parasites to mark the apicoplast division (Waller *et. al.*, 2000; He *et. al.*, 2001). The general morphology of apicoplast division in *Plasmodium falciparum* was first outlined by Waller *et. al.* (2000) in striking fluorescence images depicting the asexual parasite cell cycle within erythrocytes (Figure 1.4).

In ring stage parasites (the earliest stage after initial infection of red blood cells), the single apicoplast is crescent shaped. It subsequently rounds up into a sphere in early trophozoites and grows in size. As schizont formation begins (the stage at which the parasite begins to segment into multiple daughter parasites), the apicoplast starts forming a reticulate, branched structure, which divides (apparently simultaneously) into as many plastids as there are daughter cells in the maturing schizont. After red blood cell rupture, every merozoite (the stage that travels through the plasma from the spent blood cell and invades a new host erythrocyte) contains one slightly elongated plastid (Waller *et. al.*, 2000).

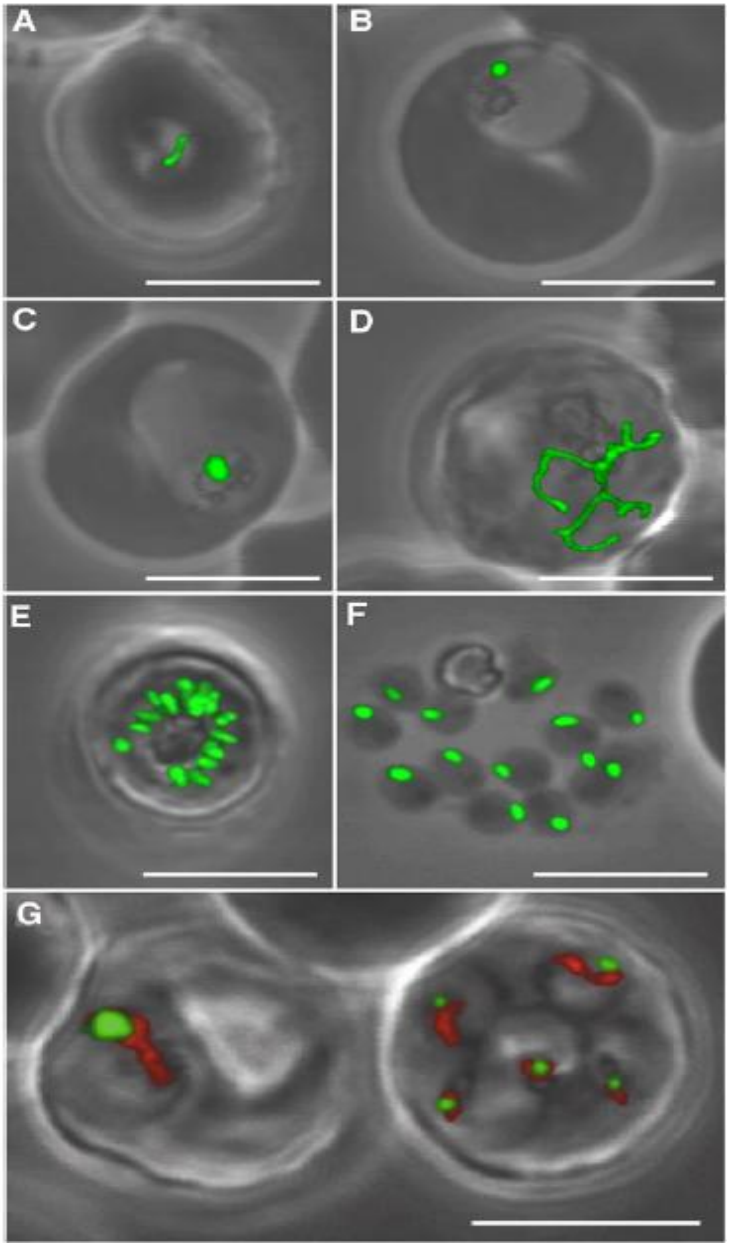

Figure 1.4: Apicoplast morphology throughout the asexual life cycle of blood stage malaria parasites expressing apicoplast-targeted GFP. The images show red blood cells infected by *Plasmodium falciparum* in various life stages. The panels represent parasites (A) just after red blood cell invasion (ring form), (B, C) during the major growth phase (trophozoites), (D, E) during cell division (schizont stage), and (F) as free daughter cells that are ready to infect new red blood cells (merozoites). (G) Costaining with Mitotracker Red shows that the apicoplast (green) is clearly distinct from but located in close proximity to the mitochondrion (red). The erythrocyte on the right contains multiple infections. Scale bars, 5µm. (Ref: 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)

In *P. falciparum* the number of daughter cells produced per schizont is variable and must obviously match the number of replicated nuclei. In addition, the division of the apicoplast and its genome must yield as many daughter plastids as there are merozoites being formed. This suggests that nuclear division and plastid division and/or segregation are intimately linked in these parasites (Waller *et. al.*, 2000; Foth *et. al.*, 2003).

The function of the apicoplast has been debated since its discovery. The parasite is absolutely dependent on this curious organelle, which has led to speculation that the apicoplast is a potential 'Achilles'Heel' of the malaria parasite. Parasites die after treatment with drugs that interrupt apicoplast genome replication, transcription or translation (Ralph, *et. al.*, 2001; Fichera, *et. al.*, 1997). Moreover, mutant parasites lacking an apicoplast are not viable (He, *et. al.*, 2001). Immediately after apicoplast perturbation (either pharmacological or genetic), parasites continue to grow normally in the host cell. However, the parasites subsequently arrest and die after infecting a new host cell. Presumably, whatever the apicoplast provides for the parasite is crucial for a viable infection process. This could be a component of the parasitophorous vacuole, which surrounds parasites in the host cell, or perhaps a resource that is usually replenished at the time of host-cell invasion (Roos *et. al.*, 1999; Sullivan *et. al.*, 2000; McKean *et. al.*, 2002; Ralph *et. al.*, 2004).

Apicoplast: A Drug Target

A number of pathways that are reported in the parasite are not found in its vertebrate host, and help provide insights into apicoplast function. The apicoplast has been convincingly shown to be the site of *de novo* fatty acid biosynthesis (Waller *et. al.*, 1998; Waller *et. al.*, 2000; Jelenska *et. al.*, 2001; McLeod *et. al.*, 2001; Surolia and Surolia, 2001) and isoprenoid biosynthesis (Jomaa *et. al.*, 1999; Wiesner *et. al.*, 2000). It also appears to contain a ferredoxin-based redox system (Vollmer *et. al.*, 2001) and part of a heme synthesis pathway (Sato and Wilson, 2002; van Dooren *et. al.*, 2002). An integrated *in silico* metabolism for apicoplast isopentenyl diphosphate, fatty-acid and haem biosynthesis is shown in Figure 1.5.

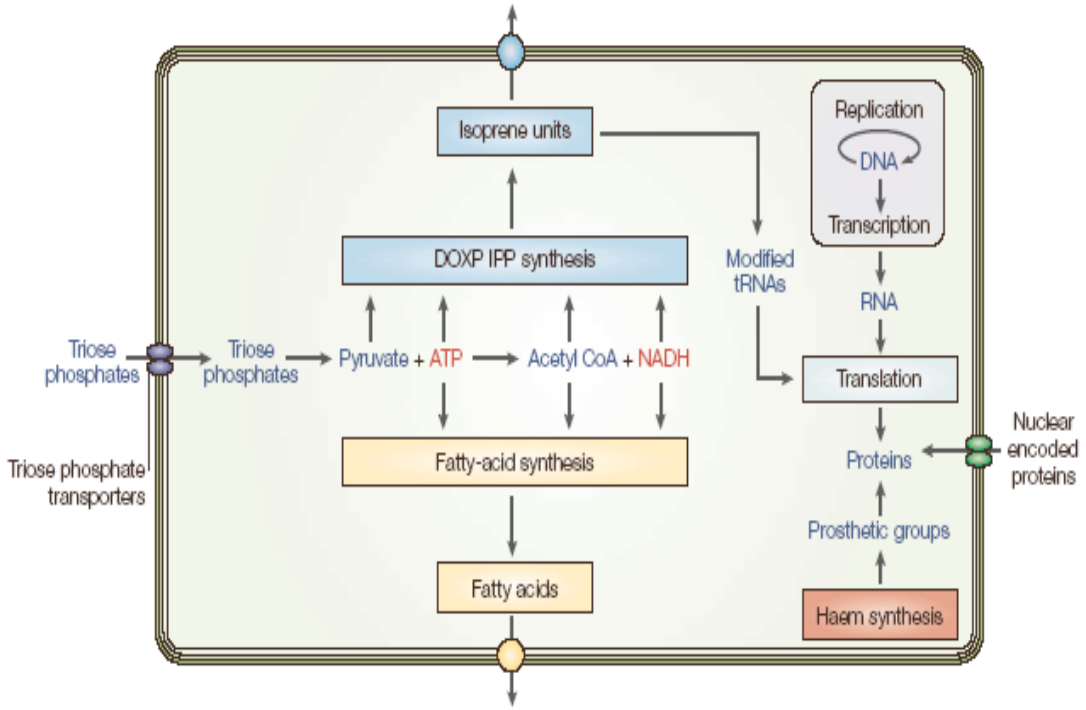

Figure 1.5: Overview of apicoplast metabolism and pathways**.** The apicoplast apparently imports trioses that are converted to either fatty acids or isopentenyl diphosphate (isoprenoid precursors) by the DOXP (1-deoxy-D-xylulose-5-phosphate) IPP (isoprenoid precursor) synthesis pathway. These acyl products are likely to be exported for use elsewhere in the parasite cell, perhaps even in formation of the parasitophorous vacuole within the host. Numerous nuclear encoded proteins are imported to join the handful of endogenously produced proteins for these activities. (Ref: Ralph, S. A., *et. al.*, 2004; Metabolic Maps and Functions of the Plasmodium falciparum apicoplast, Nat. Rev. 2: 203 – 216).

These pathways include a number of enzymes and are considered as the putative drug targets (Soldati *et. al.*, 1999; Gornicki, 2003). Some of these proteins are synthesized within the organelle but remaining proteins are targeted by the nucleus to the apicoplast. This transport of proteins takes place by a two – fold mechanism and includes a secretory pathway (Ralph *et. al*., 2004).

The apicoplast targeted proteins are fused at the N – terminal with a bipartite leader sequence. The N-terminus starts with a typical hydrophobic signal peptide that can usually be recognized by a neural network (http://www.cbs.dtu.dk/services/SignalP/), while the remainder of the N-terminal extension represents a plastid transit peptide. Deletion of just the transit peptide causes proteins, that now only contained an N-terminal signal peptide, to be secreted from the cell into the parasitophorous vacuole (DeRocher *et. al.*, 2000; Waller *et. al.*, 2000), while removal of the signal peptide alone lead to accumulation of the protein in the cytosol (Waller *et. al.*, 2000). The signal peptides are believed to mediate traffic across the outermost membrane of the apicoplast, while the transit peptides mediate traffic across the inner two membranes. The traffic through the innermost membrane is hypothesized to take place via transit peptide receptors/ TOC-TIC complex (Foth *et. al.*, 2003).

Analysis of the genes encoded by the apicoplast genomes of *P. falciparum* offers some glimpse of functions inside the apicoplast. The compact apicoplast genome consists predominantly of genes involved in protein expression within the organelle (Wilson *et. al.*, 1996). Examples include ribosomal RNAs and proteins, transfer RNAs, an RNA polymerase and the translation elongation factor TufA. The only apicoplast-encoded genes with (potentially) different function are *clpC* and *sufB* and seven unassigned ORFs. These genes are some ways involved in above discussed pathways, protein transport mechanisms and many housekeeping processes, such as DNA replication, transcription, translation and post-translational modification of apicoplast-encoded proteins. They are thus considered potential and excellent drug targets (Mcfadden *et. al.*, 1997; Foth *et. al.*, 2003).

Aims and Objectives

There are only a few reports on the partial genes of *Plasmodium vivax* apicoplast genome. Especially in the Indian context as yet there are no reports giving any information about the *P. vivax* apicoplast genome. *P. vivax* is responsible for causing over 60 – 70 %, of malaria cases annually throughout India. The organelle Apicoplast and its genome are considered as putative drug targets. A study detailing genes of this apicoplast genome from Indian *P. vivax* may thus give vital lead towards the use of established or novel therapeutic agents for the treatment of malaria.

This work aims at:

- Amplification and Sequencing of the genes from *Plasmodium vivax* Apicoplast genome.
- Study of the differences of the corresponding genes of apicoplast genome from other parasites such as *Plasmodium falciparum.*
- Localization of Apicoplast in *P. vivax* using antibodies raised in rodent models.
- Evaluation of peptides based on different apicoplast genes, to investigate antibody profiles present in naturally infected individuals.
- Evaluation of data using bio-informatics tools to investigate protein structure and function.
- Study of the evolutionary position of sequenced *P. vivax* apicoplast genes using phylogenetic analysis tools.

Organization

This work aims to identify, characterize and analyze various major genes from the Apicoplast genome of *P. vivax* from Indian isolates. The thesis is organized according to aims set as above. Chapter 1 gives introduction to the topic based upon the literature available till date, Chapter 2 gives general information related with various materials and methods used to achieve the aims, Chapter 3 gives the details about the reactions used for amplification of various genes from *P. vivax* apicoplast genome,

and also details the analysis of the obtained sequences for these genes. The protein expression studies of EF – Tu protein and peptide scanning is detailed in Chapter 4. Chapter 5 details the localization studies carried out using the designed peptides and Chapter 6 gives the work done for Homology Modelling of EF – Tu and RpoC proteins. The phylogenetic analysis to understand the evolutionary position of our identified genes from *P. vivax* are studied in Chapter 7 and Chapter 8 lists the conclusions drawn out of complete work.