Cellular and Molecular Life Sciences

MULTI-AUTHOR REVIEW

The apicomplexan plastid and its evolution

Shigeharu Sato

Abstract Protistan species belonging to the phylum Apicomplexa have a non-photosynthetic secondary plastid—the apicoplast. Although its tiny genome and even the entire nuclear genome has been sequenced for several organisms bearing the organelle, the reason for its existence remains largely obscure. Some of the functions of the apicoplast, including housekeeping ones, are significantly different from those of other plastids, possibly due to the organelle's unique symbiotic origin.

Keywords Apicomplexa · Evolution · Housekeeping functions · Plastid · Symbiosis

Abbreviations
NEP   Nuclear-encoded T7-like RNA polymerase
NTD   Nucleotide-binding domain
PBGS Porphobilinogen synthase
SPP   Stromal-processing peptidase
TPT   Triosephospate/phosphate transporter

Introduction

Apicomplexa comprise a eukaryotic phylum that, along with two others, Ciliophora and Dinoflagellata, form the superphylum Alveolata under the protistan kingdom Chromalveolata [1, 2]. To date, about 6,000 apicomplexan species have been named, but sequencing data of environmental samples suggest there may be millions more

S. Sato (✉)
Division of Parasitology, MRC National Institute for Medical Research, Mill Hill, London NW7 1AA, UK
e-mail: ssato@nimr.mrc.ac.uk

species belonging to this phylum [3]. Most named apicomplexans are obligate parasites, and some of them cause important human or animal diseases such as malaria (caused by *Plasmodium* spp.), toxoplasmosis (*Toxoplasma gondii*), coccidiosis in poultry (*Eimeria* spp.), babesiosis (*Babesia* spp.), theileriosis (*Theileria* spp.), and cryptosporidiosis (*Cryptosporidium* spp.). However, many apicomplexans are not pathogenic to their host.

Although apicomplexans apparently lack photosynthesis, they have a secondary plastid—the apicoplast (Fig. 1). Because of their clinical, veterinary, or economical importance, disease-related apicomplexans have been extensively researched, and in several instances, both the apicoplast and the nuclear genomes have been sequenced. Metabolisms involving the apicoplast have attracted attention as potential targets for disease-controlling drugs, since they might directly contribute to the survival of the apicomplexan cell. By contrast, housekeeping functions of the organelle have attracted less attention, though they are often unique or significantly different from those of other organisms. In this article, housekeeping functions unique to the apicoplast are mainly discussed.

The apicoplast—the plastid of apicomplexan cells

Discovery of the plastid in apicomplexan cells

An extremely A + T rich extrachromosomal DNA was identified in *Plasmodium* spp. [4]. The DNA was initially believed to be mitochondrial [5–8], but later was shown to be related to the plastid DNA of algae and plants [9]. Other apicomplexan genera such as *Toxoplasma*, *Eimeria*, and *Theileria* have a similar plastid-like DNA [9, 10] sharing highly conserved coding sequences [11, 12]. However,
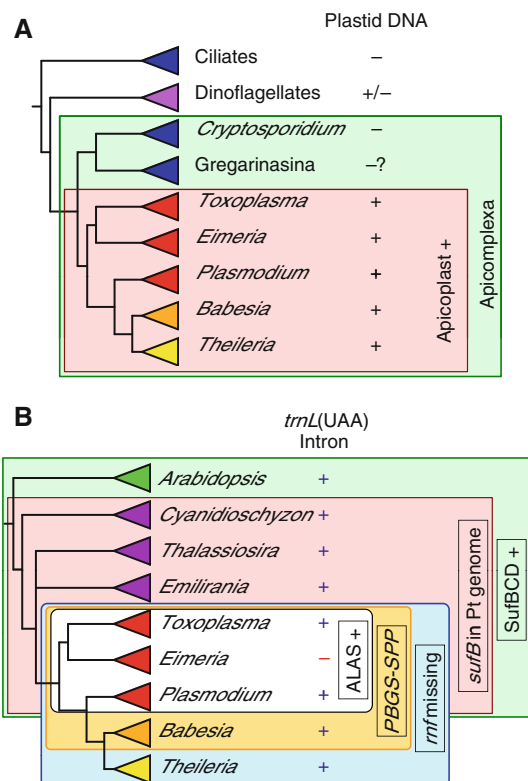
Fig. 1 Apicomplexans and the plastid. **a** Phylogeny of alveolates and distribution of the plastid. The phylogenetic tree was drawn based on the nuclear-encoded 18S rRNA sequences of representative species available in the databases and suggests only topological relationships between taxa. Distribution of the plastid in most *Gregarinasina* species has not yet been studied (see text). **b** Phylogeny of the plastids and their variety. Like red plastids (*purple*), apicoplasts (*red, orange, yellow*) have the genome encoding *sufB*, while organisms with green plastids (*green*) have the gene encoded in the nucleus. Unlike other plastids, the apicoplast genomes lack 5S rRNA gene (*rnf*). Like other plastid-bearing organisms, *Toxoplasma*, *Eimeria*, and *Plasmodium* have the Suf system incorporating the SufBCD complex, while *Babesia* and *Theileria* lack genes specifying the components of the complex. *Toxoplasma*, *Eimeria*, and *Plasmodium* have a unique hybrid-type heme pathway that involves mitochondrial ALA synthase (ALAS). Although *Babesia* lacks the heme pathway, it has the *PBGS* gene forming a tight gene cluster with the *SPP* gene in the nuclear genome like *Toxoplasma*, *Eimeria*, and *Plasmodium*. Unlike other organisms, *Eimeria* lacks the intron in the *trnL*(UAA) gene in the plastid genome

intensive studies to detect these genes in *Cryptosporidium parvum* [13] and *Gregarina niphandrodes* [14] have failed, indicating that some apicomplexans species lack the plastid-like DNA.

McFadden et al. [15] reported that rRNA encoded by the plastid-like DNA is accumulated in a small ovoid organelle located anterior to the nucleus of *T. gondii*. Independently, Köhler et al. [10] showed by in situ hybridization that the plastid-like DNA is strictly localized in an organelle, and created a new term "apicoplast" that has become popular and is widely used today by abbreviating the phrase

"apicomplexan plastid". Chaubey et al. [16] reported that the EF-Tu protein expressed from the gene on the plastid-like DNA localizes in the apicoplast as distinct from the mitochondrion in *P. falciparum*. It is generally believed that every apicomplexan species with the plastid-like DNA has an apicoplast, though this has yet to be confirmed experimentally in genera other than *Plasmodium* and *Toxoplasma*. Perhaps some species have an apicosome—an apicoplast without organellar DNA, though such has yet to be reported.

Morphology and potential association with other organelles

The number of membranes surrounding the apicoplast is generally recognized as four [15], and no internal thylacoid-like membranous structure has ever been observed. However, Hopkins et al. [17] proposed that the *P. falciparum* apicoplast probably had only three membranes. They also reported that the organelle harbors two unique membrane complexes. The inner membrane complex was predicted to be a rolled-up myelin-like invagination of the innermost membrane. On the other hand, an outer membrane complex, which lies between the outermost and the middle membrane, remains of uncertain origin. Both the outer and inner membrane complexes increase in size and in complexity during the parasite's development from merozoite to trophozoite.

Köhler [18] reported that the apicoplast membrane of *T. gondii* was not consistent throughout the organelle and at least one extensive sector appeared to be bordered solely by two membranes. All apicoplasts examined, including those in organellar division, have a single pocket-like invagination of about 50–200 nm in width, often located in close proximity to a voluminous evagination of the outer membrane of the parasite's nuclear envelope.

Tomova et al. [19] isolated *Sarcocystis* sp. from roe deer and analyzed the ultrastructure of its apicoplast using a combination of high-pressure freezing, freeze-substitution, and electron tomography. This apicoplast had four continuous membranes: two inner membranes of circular profile with a constant distance between them and two outer membranes of irregular shape. The outermost membrane displayed protuberances into the parasite cytoplasm and was associated with the endoplasmic reticulum (ER) at 'contact sites'. Similar contact sites were observed in *T. gondii*, but no fusion point was observed [20]. Thus it is uncertain if the apicoplast is connected to the ER like the secondary plastids of cryptophytes [21].

Hopkins et al. [17] mentioned that the intracellular position of the apicoplast in *P. falciparum* varied considerably depending on the developmental stage of the parasite, but there was a close association between the

apicoplast and mitochondrion in all cellular stages, including merozoites. Kobayashi et al. [22] reported the apicoplast and the mitochondrion were inseparable from one another using two different techniques—Percoll density gradient centrifugation and fluorescence-activated organelle sorting. This agrees with the suggestion now prevalent that the apicoplast in *Plasmodium* is physically bound to the mitochondrion via the cytoskeleton.

The apicoplast genome

The complete gene content and map of the apicoplast genome were first determined for *P. falciparum* [9], and then for two coccidian species, *T. gondii* (Kissinger et al., unpublished: database sequence with accession no. U87145) and *E. tenella* [23], followed by two piroplasmids, *Theileria parva* [24] and *Babesia bovis* [25]. The gene content of the apicoplast genomes is highly conserved apart from a few lineage specific genes; the genome of each species commonly encodes SSU and LSU rRNAs (*rrs* and *rrl*), three subunits of the bacteria-type RNA polymerase (*rpoB*, *rpoC1*, *rpoC2*), 16 ribosomal proteins, an EF-Tu, a ClpC-like protein and 24 tRNA species, the minimum sufficient for translation without importing a tRNA from the cytosol. On the other hand, genes for DNA metabolic enzymes, the α subunit of the bacteria-type RNA polymerase (*rpoA*) and some ribosomal proteins are missing from the apicoplast genome. These and other genes specifying most if not all proteins involved in the organellar functions are encoded by the nuclear genome.

The apicoplast genomes of *Plasmodium* and coccidian genera contain an inverted repeat (IR) each half of which consists of *rrs*, *rrl*, and nine tRNA genes. Unlike those of other plastid DNAs, the *rrs* and *rrl* in each half of the IR are arranged head-to-head, and all protein-coding genes are arranged in two large clusters following the *rrl* genes. By contrast, there is no IR in the apicoplast genome of piroplasmids, despite the fact that it contains a very similar set of genes to *Plasmodium* and coccidians. All genes in the piroplasmid apicoplast genome, except for the duplicated *clpC* and several repetitive hypothetical ORFs, are single-copy and arranged on the same DNA strand, potentially forming one big gene cluster.

Likewise, topology of the apicoplast DNA varies depending on species. Essentially all molecules of apicoplast DNA in *P. falciparum* are circular and about 35 kb in size; only a minor population (about 3%) are linear [26]. On the other hand, the circular form represents only 9% of the total mass of the apicoplast DNA in *T. gondii*; the remaining >90% existing as linear concatemers of the 35-kb units up to dodecamer [27]. The apicoplast DNAs of *E. tenella* [28] and *Neospora caninum* [29] also have been reported to comprise linear molecules of several different sizes. The topology of the apicoplast DNA for other apicomplexan species is unknown.

Nuclear-encoded apicoplast proteins

Like those of other secondary plastids [30–33], apicoplast proteins encoded by the nuclear genome generally have a bipartite organellar targeting sequence at the N terminus [34, 35]. The transit peptide following the N terminal signal sequence of the bipartite apicoplast targeting sequence of *P. falciparum* is rich in asparagine and lysine residues rather than the serine/threonine found in plants [34]. The transit peptide often contains putative Hsp70-binding sites, a feature that was included with earlier judging criteria to create the prediction algorithm PlasmoAP [36] which predicts that more than 500 proteins encoded by the nuclear genome of *P. falciparum* have an apicoplast targeting sequence at the N terminus [36, 37]. About 150 of them show significant sequence similarity to other proteins of known function or structure in the sequence databases: the most prominent include enzymes involved in de novo biosynthesis of isoprenoid, fatty acid and heme, as well as housekeeping proteins such as DNA polymerase, DNA gyrase subunits, ribosomal proteins, molecular chaperones, and components of a Suf type Fe–S cluster assembly system [37, 38]. There remain over 350 hypothetical "stromal" proteins with obscure functions. In addition to these 500+ proteins, a nuclear-encoded protein that lacks a typical bipartite organellar targeting sequence has been identified to localize to the apicoplast of *P. falciparum*. This exceptional protein, PfoTPT, is one of the two triosephosphate/phosphate transporters (TPT) that mediate moving of phosphorylated C3, C5, and C6 compounds such as phospho*enol*pyruvate across the apicoplast membranes [39]. It has been proposed that PfoTPT is inserted into the ER membrane by the first of its ten transmembrane domains that functions as the internal signal peptide, before being transferred to the outermost membrane of the apicoplast [39]. The apicoplast-related genes in the nuclear genomes of other *Plasmodium* spp. including that of oTPT are remarkably similar to those of *P. falciparum* [40, 41] with the exon/intron structures almost perfectly conserved.

The nuclear genomes of three other plastid-bearing apicomplexan species, *T. parva* [24], *T. annulata* [42], and *B. bovis* [25] have also been sequenced. Apart from unequally expanded gene families, both *Theileria* nuclear genomes contain an almost identical set of genes [42]. While localization of nuclear-encoded apicoplast proteins has not been shown experimentally, manual inspection predicts that 345 of the 4,035 proteins annotated in the nuclear genome of *T. parva* are targeted to the apicoplast; of these, 69 have functions predictable from orthologs in

other organisms [24]. Some apicoplast proteins identified in *Plasmodium* spp., such as the enzymes for synthesis of fatty acid and heme and other housekeeping proteins such as SufC involved in plastidic Fe–S cluster assembly system, are missing from *Theileria* spp. These omissions indicate that the function of the apicoplast has been greatly streamlined in these organisms. The percentage of apicoplast-targeted proteins that have been identified in the nuclear genome of *B. bovis* (47 proteins of total 3,671 = 1.3%) [25] is remarkably lower than in *P. falciparum* (8.8–10.3%) [37, 43], *T. parva* (8.6%) [24] or *A. thaliana* (7.9%) [44]. This suggests that the *B. bovis* genome encodes more apicoplast-targeted proteins than those detected by the prediction algorithms [25]. Indeed, an almost complete set of orthologs of putative apicoplast-targeted proteins of *T. parva* seem to be present in the nuclear genome, though most of them have not been predicted as apicoplast targeted [25]. Sequencing projects for two coccidian species *T. gondii* and *E. tenella* are still going on and putative nuclear-encoded 'apicoplast-targeted proteins' are listed in the partial sequence data [45]. Apicoplast localization has been confirmed for only a few of these proteins to date.

Non-housekeeping functions of the apicoplast

Apicoplasts are presumed to contribute to the metabolism of the cells that maintain them. Catalogues of nuclear-encoded enzymes predict that the apicoplast of *P. falciparum* is involved in isoprenoid biosynthesis via the DOXP/MEP pathway, fatty acid synthesis with the type II fatty acid synthase, and heme biosynthesis in collaboration with the mitochondrion [37, 43, 46, 47]. Similar predictions have been made for coccidian parasites such as *T. gondii* [45]. By contrast, the apicoplasts of piroplasmids such as *T. parva* [24] and *B. bovis* [25] are unlikely to contribute to fatty acid or heme biosynthesis. Studies using specific inhibitors have been inclined to suggest the importance of some apicoplast metabolisms, but they are not conclusive by themselves as predicted apicoplast metabolisms might be redundant. For example, triclosan, the specific inhibitor of enoyl-ACP-reductase in the type II fatty acid synthesis system, strongly suppressed the growth of *P. falciparum* [48, 49], though the drug also strongly inhibited the growth of *T. parva* that lacks enoyl-ACP-reductase [50]. This suggests there is a yet-to-be-identified target of triclosan in the organism, and perhaps in other apicomplexans as well. Another example is fosmidomycin, the specific inhibitor of DOXP reductoisomerase involved in apicoplast isoprenoid biosynthesis. Fosmidomycin strongly inhibits the growth of *P. falciparum* [51] but has little inhibitory effect on *T. gondii* [52] and *T. parva* [50].

Perhaps subtle differences in the structure of the target enzyme DOXP reductoisomerase or the accessibility of the inhibitor to the enzyme are responsible for this difference [52]. Alternatively, the importance of the isoprenoid metabolism may depend on species. Recently, Brooks et al. [53] made a *T. gondii* strain in which the apicoplast membrane-localized phosphate translocator (TgAPT) gene was knocked out conditionally. Analysis of the phenotype suggested that isoprenoid biosynthesis in the apicoplast is essential for parasites survival. A similar genetic approach should be applied to establish the importance of individual metabolisms hypothetically attributed to the apicoplast of each species.

Housekeeping functions of the apicoplast

Maintenance of the apicoplast DNA

The blood-stage *P. falciparum* contains not more than three copies of organellar genome prior to replication; this increases by more than ten times to distribute apicoplast DNA to each daughter cell in schizogony [26]. Two different mechanisms are involved in the replication of the DNA. One uses twin D-loops in the IR region, forming a θ-form intermediate and is highly sensitive to the DNA topoisomerase inhibitor ciprofloxacin [26, 54, 55]. The other, less drug sensitive mechanism probably involves rolling circles likely to initiate outside the IR, but has not been characterized.

Electron microscopic analysis of apicoplast DNA prepared from *T. gondii* tachyzoites found replicating molecules in lariat rather than θ-form replication intermediates [27]. Compared to blood-stage *P. falciparum,* *T. gondii* tachyzoites contain much more apicoplast DNA: at least 25 units of the apicoplast genome per cell [56]. In addition, organellar division occurs much more frequently in *T. gondii* tachyzoites than in blood-stage *P. falciparum*. Such factors might account for the different mode of DNA replication. As mentioned earlier, the apicoplast DNAs of *T. parva* and *B. bovis* lack the IR, all genes including those of rRNAs and tRNAs being arranged on the same DNA strand. These features suggest that piroplasmid apicoplast DNAs might replicate by a rolling circle mechanism like *T. gondii*, though experimental evidence is unavailable.

The DNA polymerase responsible for DNA replication in the apicoplast has yet to be determined. Nonetheless, one nuclear-encoded enzyme predicted to be involved has been characterized and named the plastidic DNA replication/repair enzyme complex (Prex). An apicoplast targeting sequence precedes a multifunctional polypeptide that comprises a primase–helicase domain resembling the Twinkle helicase of mammalian mitochondria, and an

exonuclease–polymerase domain like DNA polymerase I of *Aquifex aeolicus* [57]. An ortholog occurs in the nuclear genome of every plastid-bearing apicomplexan species examined, but not in the genomes of either plastid-lacking *Cryptosporidium* spp. or non-apicomplexan species. This implies that Prex is important for the maintenance of apicoplast DNA.

Like the bacteria from which it evolved, the plastid of some algae such as *Cyanidioschyzon merolae* have an HU protein that assembles organellar DNA into the structure known as the nucleoid [58]. An HU ortholog with an apicoplast targeting signal is encoded in the nuclear genome of each plastid-bearing apicomplexan species examined to date. In the case of *P. falciparum,* HU localizes in the apicoplast and binds the apicoplast DNA in a sequence-independent manner [59, 60]. PfHU is distributed throughout the interior of the apicoplast of *P. falciparum* [60], agreeing with the distribution of the apicoplast DNA suggested by electron microscopy [17] and visualization with DNA-binding fluorescent dyes [61]. Likewise, Matsuzaki et al. [56] and Köhler [18] mentioned that apicoplast DNA permeates the entire apicoplast of *T. gondii*. Together these observations imply that the apicoplast in general has a rather-spread nucleoid whose size is equal or very close to the size of the organelle itself, though the organelle may have a compact nucleoid like those of other plastids, at certain developmental stages of the cell.

Transcription and splicing

All apicoplast genomes so far analyzed encode genes specifying the β, β' and β'' subunits of the bacteria-type RNA polymerase (*rpoB*, *rpoC1*, *rpoC2*). By analogy, these subunits should form a complex with a dimer of the α subunit [62]. Unlike other plastid genomes, the apicoplast genome lacks the gene for the α subunit (*rpoA*); instead the nuclear genome of each plastid-bearing apicomplexan encodes two different *rpoA* genes. The α subunit of other organisms is comprised of two conserved domains: the N terminal domain required for assembly of the RNA polymerase complex and basal transcription [63, 64] and the C terminal domain that interacts with transcription activators recognizing the upstream promoter element [65]. Both α subunit proteins encoded by the apicomplexan nuclear genome have an predicted apicoplast targeting sequence and a well conserved N terminal domain, but neither contains sequence corresponding the C terminal domain conserved among other α subunit proteins. Furthermore, the gene specifying the σ subunit, which binds to the promoter element and promotes transcription of the gene by the bacteria-type RNA polymerase [66, 67], seems to be missing from the genomes of apicomplexan species, as are the conserved -35 and -10 elements from the 5′ sequence upstream of the apicoplast genes. These peculiarities suggest that transcription in the apicoplast is regulated differently from bacteria or other plastids.

Transcription in the mitochondrion is governed by a nuclear-encoded RNA polymerase distantly related to the RNA polymerase of T7 bacteriophage [68]. A similar nuclear-encoded T7-like RNA polymerase (NEP) is present in plastids and involved in transcription of genes in the organelle [69]. The apicomplexan nuclear genome encodes one gene for a T7-like RNA polymerase which has been annotated as a mitochondrion-specific RNA polymerase [70]. In chloroplasts, the plastid-encoded bacteria-type RNA polymerase is believed to transcribe photosynthesis-related genes whereas NEP is used for all others [69, 71]. As the apicoplast genome lacks photosynthetic genes, importing NEP for transcription could be more beneficial than keeping genes of bacteria-type RNA polymerase subunits. Consequently, the nuclear-encoded T7-like RNA polymerase might be targeted to the apicoplast and function like NEP, though no experimental evidence is available.

Like those of cyanobacteria and plastids that have a group-I type self-splicing intron [72], the *trnL*(UAA) gene in the apicoplast genome has an intervening sequence at the corresponding position. By analogy, the insert has been predicted as a self-splicing intron, though there is no experimental proof. Unlike other apicomplexans, the apicoplast *trnL*(UAA) gene of *E. tenella* lacks the insertion, and it seems likely that it was lost specifically in the *Eimeria* lineage after it evolved away from the one leading to *Toxoplasma*. A gene for a potential reverse transcriptase associated with retrotransposons and retroviruses (XP_001238615) is annotated in the nuclear genome of *E. tenella* [73], whereas retroelements have not been annotated in either *Plasmodium* or *Toxoplasma* [74]. Whilst there is no evidence that the reverse transcriptase is expressed in *E. tenella* and targeted to the apicoplast, the absence of the potential group-I intron from the *trnL*(UAA) gene in this particular coccidian might have something to do with the presence of the nuclear-encoded reverse transcriptase.

Translation

The 5S rRNA is a component of the ribosome in almost all living organisms [75]. In general, plastid genomes contain the gene for 5S rRNA (*rrf*) immediately downstream *rrl*. The exceptions so far known are dinoflagellates with the mini-circle-type plastid genome [76], *Chromera velia* [77] and apicomplexans. Mammalian mitochondrial genomes lack *rrf* and a structural model suggested the mitochondrial rRNA lack a 5S rRNA-binding domain [78]. However, it has been reported that the organelle imports nuclear-encoded 5S

rRNA from the cytosol [79, 80]. Whether the imported 5S rRNA is essential for the function of the ribosomes in mammalian mitochondria is still unknown [81]. Perhaps the apicoplast imports the 5S rRNA from the cytosol like mammalian mitochondria. Otherwise, there might be an *rrf* whose sequence is too divergent from others to be recognized. Or, the apicoplast may be a very exceptional plastid whose ribosomes are independent of 5S rRNA.

Rpl11 (L11) is a ribosomal protein that together with other ribosomal proteins—S4, L6 and L14 and the stalk proteins L10 and L7/L12—forms the "factor binding site" in prokaryote type ribosomes, including those in plastids [82]. The apicoplast genome encodes S4, L6 and L14 proteins and the nuclear genomes of *P. falciparum*, *T. parva* and *B. bovis* contain genes for L10 and L7/L12, which are probably targeted to the apicoplast [24, 25, 43]. One *rpl11* gene has been annotated in the apicoplast genomes of *T. gondii* and *E. tenella*, and the product of the *ORF129* gene in the *P. falciparum* apicoplast genome seems to be related to coccidian L11 [11]. These data imply that ribosomes in the apicoplast of these species probably have the factor-binding site like other prokaryote ribosomes. However, no gene in the *T. parva* or the *B. bovis* apicoplast genome seems to specify a product related to L11. This might be because piroplasmid *rpl11* genes are too divergent to be identified by a simple sequence similarity search, or their ribosomes might not require the L11 protein.

It is intriguing to see in frame UGA/UAA codons in some "genes" in the apicoplast genome of coccidians. Strictly, there is no direct evidence that these genes actually express proteins. However, the deduced amino acid sequences resemble those of *Plasmodium* and piroplasmids, especially when each UGA is regarded as a tryptophan codon, like those in bacteria with extremely A + T rich genomes [83]. This implies that the coccidian genes are not pseudogenes but encode expressed proteins. As in bacteria, translation termination in plastids relies on two different types of peptide releasing factors, RF1 and RF2. Both RFs bind specific nonsense codons—RF1 to UAA and UAG, whereas RF2 binds to UAA and UGA— and promote the release of the ribosome from mRNA [84, 85]. As the apicoplast genome lacks a gene for either RF, correct termination of translation in the organelle depends on imported RFs. Unlike those of *Plasmodium* and piroplasmids, the apicoplast genome of *Eimeria* and *Toxoplasma* probably contains no gene with UGA as the termination codon. Perhaps the coccidian apicoplast is free from a functional RF2 on which the apicoplast genomes of *Plasmodium* and piroplasmids apparently depend.

Unlike UGA, UAA is the codon preferentially chosen for the translation terminator by many genes in the apicoplast genome. The apicoplast genomes of *Eimeria* and *Toxoplasma* encode the β'' subunit of prokaryote-type RNA polymerase by the *rpoC2* gene that contains one in frame UAA codon. By contrast, the *rpoC2* in the *P. falciparum* apicoplast genome requires a frame shift to produce the right translation product [9]. The coccidian UAA codons and the *P. falciparum* frame shift occur in the same region of poorly conserved sequence. These data could suggest that the multi-subunit RNA polymerase of the apicoplast is unique, the β'' subunit being split into two halves. Indeed, the apicoplast genomes of *T. parva* and *B. bovis* also have split *rpoC2* genes and the N and C terminal halves have been ascribed to separate genes [24, 25]. Besides *rpoC2*, there is another in frame UAA codon in the *rps8* gene of the *T. gondii* apicoplast genome. Because the amino acid sequence specified by this gene is conserved between different apicomplexan species, the gene probably express the polypeptide, though to what amino acid residue the UAA is translated cannot be predicted from data available.

Import and maturation of organellar proteins

In plants, nuclear-encoded plastid proteins have a transit peptide at the N terminus that is cleaved from the matured form of the polypeptide by a stromal-processing peptidase (SPP) [86]. Like other secondary plastids, apicoplasts import nuclear-encoded proteins that have a signal peptide preceding the transit peptide [34]. The signal peptide is removed by the signal peptidase as the proteins are co-translationally released into the ER lumen and the resulting intermediate form with the exposed transit peptide is targeted to the stroma of the apicoplast [87]. Each membrane-crossing step is probably mediated by specific translocons such as ERDA/Der1 [88, 89]. Once the protein reaches the apicoplast stroma, the transit peptide is cleaved off by SPP as in primary plastids and the resulting polypeptide is matured with the help of organellar chaperones.

The gene for a putative SPP is found in the nuclear genome of each plastid-bearing apicomplexan species but is absent from *Cryptosporidium*. Although the SPP gene in *Theileria* encodes a protein with an apparent apicoplast targeting sequence [24, 42], those in *P. falciparum* and *T. gondii* appear to lack the bipartite organellar targeting sequence. Instead, the SPP gene of both species is preceded by the gene for apicoplast-targeted porphobilinogen synthase (PBGS) [46, 90], and exons encoding the organellar targeting sequence of PBGS are utilized to synthesize the SPP mRNA by alternative splicing [91, 92]. This unique micro-synteny occurs even in the nuclear genome of *Babesia* spp., though it is absent from *T. parva* and *T. annulata* [24, 42]. One explanation is that the unique gene expression system for SPP/PBGS is an ancient feature—probably established in a common ancestor of

Apicomplexa—and has been kept in each plastid-bearing lineage thereafter, except for *Theileria*.

Protein maturation in organelles requires the type I chaperonin system that originated from the bacterial GroESL system [93]. Like proteobacteria, the mitochondrion has a system comprised of one species each of Cpn60 and Cpn10. By contrast, the plastid, which shares the same origin as cyanobacteria, contains a chaperonin system generally made of two different Cpn60 subunits and Cpn20 [94, 95]. In addition, plant plastids have a Cpn10 that is supposed to localize in the thylakoid lumen [96]. The nuclear genome of each plastid-bearing apicomplexan species encodes two Cpn60, one Cpn10 and one Cpn20 [97]. It has been shown in *P. falciparum* that one of the two nuclear-encoded Cpn60 proteins is targeted specifically to the apicoplast along with Cpn20, while the other Cpn60 and Cpn10 are localized exclusively to the mitochondrion [97–99]. The type I chaperonin system that involves Cpn20 and only one species of Cpn60 is unique to the apicoplast. The lack of Cpn10 is probably because the organelle has no thylakoid. The nuclear genome of the plastid-less *Cryptosporidium* spp. encodes orthologs of mitochondrial Cpn60/Cpn10 but lacks genes encoding orthologs of plastidic Cpn60/Cpn20, supporting the strict organelle-specific localization of the two type I chaperonin systems of plastid-bearing apicomplexan species.

All apicoplast genomes so far analyzed contain the *clpC* gene that specifies a member of the HSP100/Clp chaperone family. These are ATP-dependent protein unfoldases belonging to the AAA^+ family, and ClpC belongs to a subgroup with two nucleotide binding domains (NTD) along with ClpA and ClpB [100]. The phylogenetically related Clp proteins of cyanobacteria and plastids are essential for normal growth [101–103]. In photobionts, ClpC forms a complex with ClpP peptidase (and related ClpR protein) in an ATP-dependent manner [104, 105], and the ClpCP(R) complex proteolyses mistargeted substrates. In addition, it has been presumed that ClpC bound to Tic110 maintains the solubility of proteins imported to plant plastids prior to transfer to other chaperones [106] and provides the driving force for complete translocation into the stroma [107]. A similar intrinsic chaperone activity preventing aggregation of unfolded polypeptides and resolubilizing and refolding aggregated proteins into their native structures, has been reported for cyanobacterial ClpC [108].

The nuclear genome of apicomplexan species with the apicoplast encodes genes for ClpP and ClpR with the apicoplast targeting sequence. Like those of other plastids, these proteins are supposed to form a ClpCPR protease complex in the apicoplast. However, ClpC specified by the gene in the apicoplast genome has only one NTD in the molecule, unlike other plastidic ClpCs. Nevertheless, the apicomplexan nuclear genome encodes several HSP100/Clp family proteins with two NTDs, and some of them have a putative apicoplast targeting sequence. Perhaps one of these nuclear-encoded Clp proteins form the ClpCPR protease complex in the apicoplast whereas others including the apicoplast-encoded ClpC with only one NTD are required for different functions. It is known that ClpB, a paralog to ClpC, and Hsp70 (DnaK) comprise a bi-chaperone system that is important for disaggregation and refolding of intracellular protein aggregates [109]. Like cyanobacteria, plastids of other organisms in general have both ClpB and Hsp70 whereas the apicoplast seems to lack Hsp70. Perhaps, the apicoplast-encoded ClpC is an unusual ClpB that has evolved to compensate the lack of Hsp70. Alternatively, the protein might directly acts as a substitute for the missing Hsp70.

The *sufB* gene, formerly named *ycf24*, is found in the organellar genomes of *Plasmodium* and coccidian species whereas it is missing in piroplasmids. In *E. coli*, *sufB* is tightly linked with other genes—*sufA*, *sufC*, *sufD*, *sufE* and *sufS*—forming an operon whose transcription is induced during exposure to hydrogen peroxide [110] and iron starvation [111]. SufS is a cysteine desulphurase that is involved in Fe–S cluster assembly and SufE enhances the function of SufS [112, 113]. SufB forms a ternary complex with SufC and SufD [114] and stimulates the function of SufE on SufS [115, 116]. Plastids have inherited Suf proteins form their bacterial ancestor [117], though the genes have mostly transferred to the nuclear genome. No Suf genes remain in the plastid genome of Viridiplantae, but a *sufB*–*sufC* cluster is generally found on the genome of red plastids. The apicoplast genome with *sufB* appears to be intermediate as it lacks *sufC*; the gene is encoded by the nuclear genome together with other Suf genes.

Like *Plasmodium* and coccidians, piroplasmids have several apicoplast proteins requiring Fe–S clusters. But they lack *sufB*, *sufC*, *sufD* and *sufE* genes, while *sufS* gene is encoded in the nuclear genome. One explanation is that the piroplasmid apicoplast acquires Fe–S clusters from outside the organelle importing them by an unknown mechanism and that SufS is required only for other metabolisms such as tRNA modification. Alternatively, the organelle might have a simplified Suf system that would resemble the mitochondrial Isc system where an Hsp70 takes the place of the SufBCD complex [118]. Perhaps the ClpC protein encoded by the duplicated apicoplast genes participates in the system, acting as the functional substitute for the Hsp70.

Conclusion and remarks

The lineages of *Plasmodium* and coccidians such as *Toxoplasma* and *Eimeria* are estimated to have diverged from

each other about 350–824 million years ago, predating the divergence of *Plasmodium* and piroplasmids [74, 119, 120]. The remarkable syntheny of the entire apicoplast genome between *Plasmodium* and coccidians is unlikely to have arisen secondarily by convergence. Probably the differently organized piroplasmid genomes were generated by later rearrangements. Despite these changes, the gene content of the piroplasmid apicoplast genomes is almost the same as *Plasmodium* and coccidians, suggesting the present gene content is the minimum acceptable for these secondary plastids.

All apicoplast genomes so far analyzed contain *clpC* in addition to genes involved in transcription/translation. This fact could imply that the apicoplast genome exists in order to express ClpC. Perhaps the apicoplast-encoded ClpC is the functional substitute for the Hsp70 missing from the stroma of the organelle. The apicoplast genome of piroplasmids contains duplicated *clpC* genes but lacks *sufB* present in the *Plasmodium* and the coccidian apicoplast genomes. The mitochondrial Fe–S cluster assembly system (the Isc system) involves an Hsp70-family protein in place of the SufBCD complex of the plastidic Suf system [114, 118]. Again, one of the duplicated ClpCs might function as substitute for Hsp70, compensating the absence of the SufBCD complex in the Suf type Fe–S cluster assembly in the piroplasmid apicoplast.

The most distant ancestor of the apicoplast must have been a photosynthetic primary plastid. However, the immediate origin of the apicoplast acquired by the ancestor of Apicomplexa and other related protists is unknown. The plastid could have been primary or secondary, but it must have come from an alga with a plastid whose organellar genome encoded SufB. Recently, the plastid genome was reported for a previously undescribed alga CCMP3155, a photosynthetic alveolate that is phylogenetically close to the Apicomplexa [77]. This plastid genome contains all the genes in the apicoplast genome, and the order of genes in the apicoplast genome can be reconstructed from the CCMP3155 plastid genome with only a small number of hypothetical rearrangements. These data imply the Apicomplexa and CCMP3155 are close related, though not necessarily in direct line.

When the plastid-less ancestor of Apicomplexa acquired its symbiotic alga as secondary plastid, some metabolic pathways in the organism must have been duplicated. Competition between duplicated pathways would have ensued to reduce the redundancy. Heme synthesis was one such pathway. In general, plastid-bearing organisms synthesize heme exclusively in the plastid using a characteristic metabolic pathway also used for chlorophyll synthesis. By contrast, plastid-less organisms have a distinct heme pathway that involves the mitochondrion and the cytosol. Probably because the plastid pathway is

sufficient to supply all required heme, organisms with a photosynthetic secondary plastid, such as diatoms, maintain the complete plastid-localized heme pathway, losing the non-plastid pathway (Fig. 2). Like other plastid-bearing organisms, apicomplexans such as *P. falciparum* have an algal-type PBGS that localizes in the apicoplast [46, 90]. However, they lack algal enzymes required for synthesizing substrate for PBGS in the plastid [121] and have ALA synthase, which is found exclusively in the plastid-less organisms, in the mitochondrion instead [46]. This fact might suggest that chlorophyll synthesis in the ancestral apicoplast in the ancient Apicomplexa in which the selection took place had already become dispensable. In other words, the alga from which the apicoplast originated could have been non-photosynthetic, contrary to the commonly held belief. Even so, the apicomplexan ancestor must have benefited from the acquired plastid as it would be a useful source of various metabolites even when non-photosynthetic [122]. Why then has the plastidic PBGS been maintained instead of the non-plastidic one? One possible explanation is that the algal PBGS gene could not be lost because it was inserted by chance within the SPP gene to make the unique PBGS/SPP gene cluster.
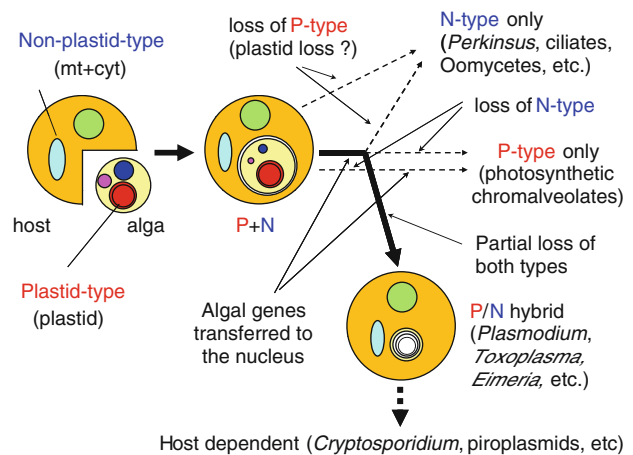
Fig. 2 Evolution of heme biosynthesis in organisms with secondary plastids (hypothesis). The plastid-less ancestor had a non-plastid (*N*) type heme pathway whereas the algal endosymbiont, which donated the secondary plastid, had a distinct pathway in the plastid (*P*). Algal genes were transferred to the host nuclear genome one by one. The ancestral organism with a secondary plastid initially had both P- and N-type pathways, but subsequently, one of the two was selected and the other wiped out as redundant. Because chlorophyll synthesis depends on the P-type heme pathway, photosynthetic organisms chose the P-type, throwing the N-type away. By contrast, organisms that were not dependent on photosynthesis kept the N-type; this was often accompanied by loss of the plastid. The apicomplexan ancestor was non-photosynthetic and would lose the P-type. However, the gene complex of PBGS and SPP in the nuclear genome made the loss incomplete, giving rise to a unique P/N hybrid heme pathway (*thick line*). Thereafter some apicomplexans lost the entire pathway

The proteome of the apicoplast of piroplasmids has shrunk considerably compared to those of *Plasmodium* and coccidians. This is probably because piroplasmids have acquired alternative sources of some metabolites that used to be supplied from the apicoplast. Apicoplast DNA is apparently missing from *C. parvum* [13] and the fact that the nuclear genome of *C. parvum* lacks genes that are required for the biogenesis of the apicoplast such as Cpn60 and SPP [123] suggests that the species completely lacks the apicoplast as an organelle. Gregarinasina, one of the major apicomplexan subclasses that contains at least 1,600 species [2, 124], has been shown to be closely related to the genus *Cryptosporidium*. Some if not all Gregarinasina can lack the apicoplast like *Cryptosporidium*. Indeed, it was reported that *G. niphandrodes* lacks apicoplast DNA [14].

To investigate the past and future of the apicoplast, the genome of many more apicomplexans should be studied, especially those placed phylogenetically between *Plasmodium* and *T. parva/B. bovis* as well as those belonging to Gregarinasina. The analysis of non-apicomplexans closely related to Apicomplexa also is essential. To elucidate the functions of present day apicoplasts, it is critical to establish an isolation method, as biochemical and proteomic analysis of the purified organelle is definitely required.

References

1. Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52:399–451

2. Levine ND (1988) Protozoan phylum Apicomplexa. CRC Press, Boca Raton

3. Adl SM, Leander BS, Simpson AG, Archibald JM, Anderson OR, Bass D, Bowser SS, Brugerolle G, Farmer MA, Karpov S, Kolisko M, Lane CE, Lodge DJ, Mann DG, Meisterfeld R, Mendoza L, Moestrup O, Mozley-Standridge SE, Smirnov AV, Spiegel F (2007) Diversity, nomenclature, and taxonomy of protists. Syst Biol 56:684–689

4. Gutteridge WE, Trigg PI, Williamson DH (1971) Properties of DNA from some malarial parasites. Parasitology 62:209–219

5. Kilejian A (1975) Circular mitochondrial DNA from the avian malarial parasite *Plasmodium lophurae*. Biochim Biophys Acta 390:276–284

6. Dore E, Frontali C, Forte T, Fratarcangeli S (1983) Further studies and electron microscopic characterization of *Plasmodium berghei* DNA. Mol Biochem Parasitol 8:339–352

7. Williamson DH, Wilson RJM, Bates PA, McCready S, Perler F, Qiang BU (1985) Nuclear and mitochondrial DNA of the primate malarial parasite *Plasmodium knowlesi*. Mol Biochem Parasitol 14:199–209

8. Gardner MJ, Bates PA, Ling IT, Moore DJ, McCready S, Gunasekera MB, Wilson RJM, Williamson DH (1988) Mitochondrial DNA of the human malarial parasite *Plasmodium falciparum*. Mol Biochem Parasitol 31:11–17

9. Wilson RJM, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH (1996) Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J Mol Biol 261:155–172

10. Köhler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJM, Palmer JD, Roos DS (1997) A plastid of probable green algal origin in apicomplexan parasites. Science 275:1485–1489

11. Denny P, Preiser P, Williamson D, Wilson I (1998) Evidence for a single origin of the 35 kb plastid DNA in apicomplexans. Protist 149:51–59

12. Lang-Unnasch N, Reith ME, Munholland J, Barta JR (1998) Plastids are widespread and ancient in parasites of the phylum Apicomplexa. Int J Parasitol 28:1743–1754

13. Zhu G, Marchewka MJ, Keithly JS (2000) *Cryptosporidium parvum* appears to lack a plastid genome. Microbiology 146(Pt 2):315–321

14. Toso MA, Omoto CK (2007) *Gregarina niphandrodes* may lack both a plastid genome and organelle. J Eukaryot Microbiol 54:66–72

15. McFadden GI, Reith ME, Munholland J, Lang-Unnasch N (1996) Plastid in human parasites. Nature 381:482

16. Chaubey S, Kumar A, Singh D, Habib S (2005) The apicoplast of *Plasmodium falciparum* is translationally active. Mol Microbiol 56:81–89

17. Hopkins J, Fowler R, Krishna S, Wilson I, Mitchell G, Bannister L (1999) The plastid in *Plasmodium falciparum* asexual blood stages: a three-dimensional ultrastructural analysis. Protist 150:283–295

18. Köhler S (2005) Multi-membrane-bound structures of Apicomplexa: I. The architecture of the *Toxoplasma gondii* apicoplast. Parasitol Res 96:258–272

19. Tomova C, Geerts WJ, Muller-Reichert T, Entzeroth R, Humbel BM (2006) New comprehension of the apicoplast of *Sarcocystis* by transmission electron tomography. Biol Cell 98:535–545

20. Tomova C, Humbel BM, Geerts WJ, Entzeroth R, Holthuis JC, Verkleij AJ (2009) Membrane contact sites between apicoplast and ER in *Toxoplasma gondii* revealed by electron tomography. Traffic 10:1471–1480

21. Gould SB, Waller RF, McFadden GI (2008) Plastid evolution. Annu Rev Plant Biol 59:491–517

22. Kobayashi T, Sato S, Takamiya S, Komaki-Yasuda K, Yano K, Hirata A, Onitsuka I, Hata M, Mi-ichi F, Tanaka T, Hase T, Miyajima A, Kawazu S, Watanabe Y, Kita K (2007) Mitochondria and apicoplast of *Plasmodium falciparum*: behaviour on subcellular fractionation and the implication. Mitochondrion 7:125–132

23. Cai X, Fuller AL, McDougald LR, Zhu G (2003) Apicoplast genome of the coccidian *Eimeria tenella*. Gene 321:39–46

24. Gardner MJ, Bishop R, Shah T, de Villiers EP, Carlton JM, Hall N, Ren Q, Paulsen IT, Pain A, Berriman M, Wilson RJM, Sato S, Ralph SA, Mann DJ, Xiong Z, Shallom SJ, Weidman J, Jiang L, Lynn J, Weaver B, Shoaibi A, Domingo AR, Wasawo D, Crabtree J, Wortman JR, Haas B, Angiuoli SV, Creasy TH, Lu

C, Suh B, Silva JC, Utterback TR, Feldblyum TV, Pertea M, Allen J, Nierman WC, Taracha EL, Salzberg SL, White OR, Fitzhugh HA, Morzaria S, Venter JC, Fraser CM, Nene V (2005) Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. Science 309:134–137

25. Brayton KA, Lau AO, Herndon DR, Hannick L, Kappmeyer LS, Berens SJ, Bidwell SL, Brown WC, Crabtree J, Fadrosh D, Feldblum T, Forberger HA, Haas BJ, Howell JM, Khouri H, Koo H, Mann DJ, Norimine J, Paulsen IT, Radune D, Ren Q, Smith RK Jr, Suarez CE, White O, Wortman JR, Knowles DP Jr, McElwain TF, Nene VM (2007) Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. PLoS Pathog 3:1401–1413

26. Williamson DH, Preiser PR, Moore PW, McCready S, Strath M, Wilson RJM (2002) The plastid DNA of the malaria parasite *Plasmodium falciparum* is replicated by two mechanisms. Mol Microbiol 45:533–542

27. Williamson DH, Denny PW, Moore PW, Sato S, McCready S, Wilson RJM (2001) The in vivo conformation of the plastid DNA of *Toxoplasma gondii*: implications for replication. J Mol Biol 306:159–168

28. Dunn PP, Stephens PJ, Shirley MW (1998) *Eimeria tenella*: two species of extrachromosomal DNA revealed by pulsed-field gel electrophoresis. Parasitol Res 84:272–275

29. Gleeson MT, Johnson AM (1999) Physical characterisation of the plastid DNA in *Neospora caninum*. Int J Parasitol 29:1563–1573

30. Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG (2009) Protein targeting into secondary plastids. J Eukaryot Microbiol 56:9–15

31. Apt KE, Clendennen SK, Powers DA, Grossman AR (1995) The gene family encoding the fucoxanthin chlorophyll proteins from the brown alga *Macrocystis pyrifera*. Mol Gen Genet 246:455–464

32. Lang M, Apt KE, Kroth PG (1998) Protein transport into "complex" diatom plastids utilizes two different targeting signals. J Biol Chem 273:30973–30978

33. Schwartzbach SD, Osafune T, Loffelhardt W (1998) Protein import into cyanelles and complex chloroplasts. Plant Mol Biol 38:247–263

34. Waller RF, Keeling PJ, Donald RG, Striepen B, Handman E, Lang-Unnasch N, Cowman AF, Besra GS, Roos DS, McFadden GI (1998) Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. Proc Natl Acad Sci USA 95:12352–12357

35. Waller RF, Reed MB, Cowman AF, McFadden GI (2000) Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. EMBO J 19:1794–1802

36. Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, Cowman AF, McFadden GI (2003) Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. Science 299:705–708

37. Ralph SA, van Dooren GG, Waller RF, Crawford MJ, Fraunholz MJ, Foth BJ, Tonkin CJ, Roos DS, McFadden GI (2004) Tropical infectious diseases: metabolic maps and functions of the *Plasmodium falciparum* apicoplast. Nat Rev Microbiol 2:203–216

38. Sato S, Wilson RJM (2005) The plastid of *Plasmodium* spp.: a target for inhibitors. Curr Top Microbiol Immunol 295:251–273

39. Mullin KA, Lim L, Ralph SA, Spurck TP, Handman E, McFadden GI (2006) Membrane transporters in the relict plastid of malaria parasites. Proc Natl Acad Sci USA 103:9572–9577

40. Pain A, Bohme U, Berry AE, Mungall K, Finn RD, Jackson AP, Mourier T, Mistry J, Pasini EM, Aslett MA, Balasubrammaniam S, Borgwardt K, Brooks K, Carret C, Carver TJ, Cherevach I, Chillingworth T, Clark TG, Galinski MR, Hall N, Harper D, Harris D, Hauser H, Ivens A, Janssen CS, Keane T, Larke N, Lapp S, Marti M, Moule S, Meyer IM, Ormond D, Peters N, Sanders M, Sanders S, Sargeant TJ, Simmonds M, Smith F, Squares R, Thurston S, Tivey AR, Walker D, White B, Zuiderwijk E, Churcher C, Quail MA, Cowman AF, Turner CM, Rajandream MA, Kocken CH, Thomas AW, Newbold CI, Barrell BG, Berriman M (2008) The genome of the simian and human malaria parasite *Plasmodium knowlesi*. Nature 455:799–803

41. Carlton JM, Angiuoli SV, Suh BB, Kooij TW, Pertea M, Silva JC, Ermolaeva MD, Allen JE, Selengut JD, Koo HL, Peterson JD, Pop M, Kosack DS, Shumway MF, Bidwell SL, Shallom SJ, van Aken SE, Riedmuller SB, Feldblyum TV, Cho JK, Quackenbush J, Sedegah M, Shoaibi A, Cummings LM, Florens L, Yates JR, Raine JD, Sinden RE, Harris MA, Cunningham DA, Preiser PR, Bergman LW, Vaidya AB, van Lin LH, Janse CJ, Waters AP, Smith HO, White OR, Salzberg SL, Venter JC, Fraser CM, Hoffman SL, Gardner MJ, Carucci DJ (2002) Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. Nature 419:512–519

42. Pain A, Renauld H, Berriman M, Murphy L, Yeats CA, Weir W, Kerhornou A, Aslett M, Bishop R, Bouchier C, Cochet M, Coulson RM, Cronin A, de Villiers EP, Fraser A, Fosker N, Gardner M, Goble A, Griffiths-Jones S, Harris DE, Katzer F, Larke N, Lord A, Maser P, McKellar S, Mooney P, Morton F, Nene V, O'Neil S, Price C, Quail MA, Rabbinowitsch E, Rawlings ND, Rutter S, Saunders D, Seeger K, Shah T, Squares R, Squares S, Tivey A, Walker AR, Woodward J, Dobbelaere DA, Langsley G, Rajandream MA, McKeever D, Shiels B, Tait A, Barrell B, Hall N (2005) Genome of the host-cell transforming parasite *Theileria annulata* compared with *T. parva*. Science 309:131–133

43. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419:498–511

44. Richly E, Leister D (2004) An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. Gene 329:11–16

45. Fleige T, Limenitakis J, Soldati-Favre D (2010) Apicoplast: keep it or leave it. Microbes Infect 12:253–262

46. Sato S, Clough B, Coates L, Wilson RJM (2004) Enzymes for heme biosynthesis are found in both the mitochondrion and plastid of the malaria parasite *Plasmodium falciparum*. Protist 155:117–125

47. Roos DS, Crawford MJ, Donald RG, Fraunholz M, Harb OS, He CY, Kissinger JC, Shaw MK, Striepen B (2002) Mining the *Plasmodium* genome database to define organellar function: what does the apicoplast do? Philos Trans R Soc Lond B Biol Sci 357:35–46

48. McLeod R, Muench S, Rafferty J, Kyle D, Mui E, Kirisits M, Mack D, Roberts C, Samuel B, Lyons R, Dorris M, Milhous W, Rice D (2001) Triclosan inhibits the growth of *Plasmodium falciparum* and *Toxoplasma gondii* by inhibition of apicomplexan Fab I. Int J Parasitol 31:109–113

49. Surolia N, Surolia A (2001) Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of *Plasmodium falciparum*. Nat Med 7:167–173

50. Lizundia R, Werling D, Langsley G, Ralph SA (2009) *Theileria* apicoplast as a target for chemotherapy. Antimicrob Agents Chemother 53:1213–1217

51. Jomaa H, Wiesner J, Sanderbrand S, Altincicek B, Weidemeyer C, Hintz M, Turbachova I, Eberl M, Zeidler J, Lichtenthaler HK, Soldati D, Beck E (1999) Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. Science 285:1573–1576

52. Clastre M, Goubard A, Prel A, Mincheva Z, Viaud-Massuart MC, Bout D, Rideau M, Velge-Roussel F, Laurent F (2007) The methylerythritol phosphate pathway for isoprenoid biosynthesis in coccidia: presence and sensitivity to fosmidomycin. Exp Parasitol 116:375–384

53. Brooks CF, Johnsen H, van Dooren GG, Muthalagi M, Lin SS, Bohne W, Fischer K, Striepen B (2010) The *Toxoplasma* apicoplast phosphate translocator links cytosolic and apicoplast metabolism and is essential for parasite survival. Cell Host Microbe 7:62–73

54. Singh D, Chaubey S, Habib S (2003) Replication of the *Plasmodium falciparum* apicoplast DNA initiates within the inverted repeat region. Mol Biochem Parasitol 126:9–14

55. Singh D, Kumar A, Raghu Ram EV, Habib S (2005) Multiple replication origins within the inverted repeat region of the *Plasmodium falciparum* apicoplast genome are differentially activated. Mol Biochem Parasitol 139:99–106

56. Matsuzaki M, Kikuchi T, Kita K, Kojima S, Kuroiwa T (2001) Large amount of apicoplast nucleoid DNA and its segregation in *Toxoplasma gondii*. Protoplasma 218:180–191

57. Seow F, Sato S, Janssen CS, Riehle MO, Mukhopadhyay A, Phillips RS, Wilson RJM, Barrett MP (2005) The plastidic DNA replication enzyme complex of *Plasmodium falciparum*. Mol Biochem Parasitol 141:145–153

58. Sato N, Terasawa K, Miyajima K, Kabeya Y (2003) Organization, developmental dynamics, and evolution of plastid nucleoids. Int Rev Cytol 232:217–262

59. Ram EV, Naik R, Ganguli M, Habib S (2008) DNA organization by the apicoplast-targeted bacterial histone-like protein of *Plasmodium falciparum*. Nucleic Acids Res 36:5061–5073

60. Sasaki N, Hirai M, Maeda K, Yui R, Itoh K, Namiki S, Morita T, Hata M, Murakami-Murofushi K, Matsuoka H, Kita K, Sato S (2009) The *Plasmodium* HU homolog, which binds the plastid DNA sequence-independent manner, is essential for the parasite's survival. FEBS Lett 583:1446–1450

61. Sano-Maeda K, Sato S, Ueda T, Yui R, Itho K, Hata M, Nakano A, Kita K, Murakmi-Murofushi K, Sasai N (2009) Visualization of mitochondrial and apicoplast nucleoids in the human malaria parasite *Plasmodium falciparum* by SYBR Green I and Pico-Green staining. Cytologia 74:449–455

62. Zhang G, Darst SA (1998) Structure of the *Escherichia coli* RNA polymerase α subunit amino-terminal domain. Science 281:262–266

63. Igarashi K, Ishihama A (1991) Bipartite functional map of the *E. coli* RNA polymerase α subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. Cell 65:1015–1022

64. Blatter E, Ross W, Tang H, Gourse R, Ebright R (1994) Domain organization of RNA polymerase α subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. Cell 78:889–896

65. Ishihama A (1993) Protein–protein communication within the transcription apparatus. J Bacteriol 175:2483–2489

66. Burgess RR, Anthony L (2001) How sigma docks to RNA polymerase and what sigma does. Curr Opin Microbiol 4:126–131

67. Lysenko EA (2007) Plant sigma factors and their role in plastid transcription. Plant Cell Rep 26:845–859

68. Masters BS, Stohl LL, Clayton DA (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. Cell 51:89–99

69. Hedtke B, Borner T, Weihe A (2000) One RNA polymerase serving two genomes. EMBO Rep 1:435–440

70. Li J, Maga JA, Cermakian N, Cedergren R, Feagin JE (2001) Identification and characterization of a *Plasmodium falciparum* RNA polymerase gene with similarity to mitochondrial RNA polymerase. Mol Biochem Parasitol 113:261–269

71. Krause K (2008) From chloroplasts to "cryptic" plastids: evolution of plastid genomes in parasitic plants. Curr Genet 54:111–121

72. Kuhsel MG, Strickland R, Palmer JD (1990) An ancient group I intron shared by eubacteria and chloroplasts. Science 250:1570–1573

73. Ling KH, Rajandream MA, Rivailler P, Ivens A, Yap SJ, Madeira AM, Mungall K, Billington K, Yee WY, Bankier AT, Carroll F, Durham AM, Peters N, Loo SS, Isa MN, Novaes J, Quail M, Rosli R, Nor Shamsudin M, Sobreira TJ, Tivey AR, Wai SF, White S, Wu X, Kerhornou A, Blake D, Mohamed R, Shirley M, Gruber A, Berriman M, Tomley F, Dear PH, Wan KL (2007) Sequencing and analysis of chromosome 1 of *Eimeria tenella* reveals a unique segmental organization. Genome Res 17:311–319

74. Templeton TJ, Enomoto S, Chen WJ, Huang CG, Lancto CA, Abrahamsen MS, Zhu G (2010) A genome-sequence survey for *Ascogregarina taiwanensis* supports evolutionary affiliation but metabolic diversity between a *Gregarine* and *Cryptosporidium*. Mol Biol Evol 27:235–248

75. Bogdanov AA, Dontsova OA, Dokudovskaya SS, Lavrik IN (1995) Structure and function of 5S rRNA in the ribosome. Biochem Cell Biol 73:869–876

76. Barbrook AC, Santucci N, Plenderleith LJ, Hiller RG, Howe CJ (2006) Comparative analysis of dinoflagellate chloroplast genomes reveals rRNA and tRNA genes. BMC Genomics 7:297

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

78. Mears JA, Sharma MR, Gutell RR, McCook AS, Richardson PE, Caulfield TR, Agrawal RK, Harvey SC (2006) A structural model for the large subunit of the mammalian mitochondrial ribosome. J Mol Biol 358:193–212

79. Yoshionari S, Koike T, Yokogawa T, Nishikawa K, Ueda T, Miura K, Watanabe K (1994) Existence of nuclear-encoded 5S rRNA in bovine mitochondria. FEBS Lett 338:137–142

80. Magalhaes PJ, Andreu AL, Schon EA (1998) Evidence for the presence of 5S rRNA in mammalian mitochondria. Mol Biol Cell 9:2375–2382

81. Smirnov AV, Entelis NS, Krasheninnikov IA, Martin R, Tarassov IA (2008) Specific features of 5S rRNA structure—its interactions with macromolecules and possible functions. Biochemistry (Mosc) 73:1418–1437

82. Brodersen DE, Nissen P (2005) The social life of ribosomal proteins. FEBS J 272:2098–2108

83. Ohama T, Inagaki Y, Bessho Y, Osawa S (2008) Evolving genetic code. Proc Jpn Acad Ser B Phys Biol Sci 84:58–74

84. Nakamura Y, Ito K (1998) How protein reads the stop codon and terminates translation. Genes Cells 3:265–278

85. Scolnick E, Tompkins R, Caskey T, Nirenberg M (1968) Release factors differing in specificity for terminator codons. Proc Natl Acad Sci USA 61:768–774

86. Richter S, Lamppa GK (1998) A chloroplast processing enzyme functions as the general stromal processing peptidase. Proc Natl Acad Sci USA 95:7463–7468

87. Tonkin CJ, Struck NS, Mullin KA, Stimmler LM, McFadden GI (2006) Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites. Mol Microbiol 61:614–630

88. Kalanon M, Tonkin CJ, McFadden GI (2009) Characterization of two putative protein translocation components in the apicoplast of *Plasmodium falciparum*. Eukaryot Cell 8:1146–1154

89. Spork S, Hiss JA, Mandel K, Sommer M, Kooij TWA, Chu T, Schneider G, Maier UG, Przyborski JM (2009) An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. EUKARYOTIC CELL 8:1134–1145

90. Sato S, Wilson RJM (2002) The genome of *Plasmodium falciparum* encodes an active δ-aminolevulinic acid dehydratase. Curr Genet 40:391–398

91. van Dooren GG, Su V, D'Ombrain MC, McFadden GI (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

92. Shanmugam D, Wu B, Ramirez U, Jaffe EK, Roos DS (2010) Plastid-associated porphobilinogen synthase from *Toxoplasma gondii*: kinetic and structural properties validate therapeutic potential. J Biol Chem 285:22122–22131

93. Hemmingsen SM, Woolford C, van der Vies SM, Tilly K, Dennis DT, Georgopoulos CP, Hendrix RW, Ellis RJ (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. Nature 333:330–334

94. Bertsch U, Soll J, Seetharam R, Viitanen PV (1992) Identification, characterization, and DNA sequence of a functional "double" groES-like chaperonin from chloroplasts of higher plants. Proc Natl Acad Sci USA 89:8696–8700

95. Nishio K, Hirohashi T, Nakai M (1999) Chloroplast chaperonins: evidence for heterogeneous assembly of α and β Cpn60 polypeptides into a chaperonin oligomer. Biochem Biophys Res Commun 266:584–587

96. Schlicher T, Soll J (1996) Molecular chaperones are present in the thylakoid lumen of pea chloroplasts. FEBS Lett 379:302–304

97. Sato S, Wilson RJM (2005) Organelle-specific cochaperonins in apicomplexan parasites. Mol Biochem Parasitol 141:133–143

98. Sato S, Rangachari K, Wilson RJM (2003) Targeting GFP to the malarial mitochondrion. Mol Biochem Parasitol 130:155–158

99. Sato S, Wilson RJM (2004) The use of DsRED in single- and dual-color fluorescence labeling of mitochondrial and plastid organelles in *Plasmodium falciparum*. Mol Biochem Parasitol 134:175–179

100. Neuwald AF, Aravind L, Spouge JL, Koonin EV (1999) AAA$^+$: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res 9:27–43

101. Clarke AK, Eriksson MJ (1996) The cyanobacterium *Synechococcus* sp. PCC 7942 possesses a close homologue to the chloroplast ClpC protein of higher plants. Plant Mol Biol 31:721–730

102. Adam Z, Clarke AK (2002) Cutting edge of chloroplast proteolysis. Trends Plant Sci 7:451–456

103. Sjogren LL, MacDonald TM, Sutinen S, Clarke AK (2004) Inactivation of the *clpC1* gene encoding a chloroplast Hsp100 molecular chaperone causes growth retardation, leaf chlorosis, lower photosynthetic activity, and a specific reduction in photosystem content. Plant Physiol 136:4114–4126

104. Halperin T, Ostersetzer O, Adam Z (2001) ATP-dependent association between subunits of Clp protease in pea chloroplasts. Planta 213:614–619

105. Sokolenko A, Lerbs-Mache S, Altschmied L, Herrmann RG (1998) Clp protease complexes and their diversity in chloroplasts. Planta 207:286–295

106. Nielsen E, Akita M, Davila-Aponte J, Keegstra K (1997) Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. EMBO J 16:935–946

107. Jackson-Constan D, Akita M, Keegstra K (2001) Molecular chaperones involved in chloroplast protein import. Biochim Biophys Acta 1541:102–113

108. Andersson FI, Blakytny R, Kirstein J, Turgay K, Bukau B, Mogk A, Clarke AK (2006) Cyanobacterial ClpC/HSP100 protein displays intrinsic chaperone activity. J Biol Chem 281:5468–5475

109. Haslberger T, Bukau B, Mogk A (2010) Towards a unifying mechanism for ClpB/Hsp104-mediated protein disaggregation and prion propagation. Biochem Cell Biol 88:63–75

110. Zheng M, Wang X, Templeton LJ, Smulski DR, LaRossa RA, Storz G (2001) DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. J Bacteriol 183:4562–4570

111. Outten FW, Djaman O, Storz G (2004) A suf operon requirement for Fe–S cluster assembly during iron starvation in *Escherichia coli*. Mol Microbiol 52:861–872

112. Loiseau L, Ollagnier-de-Choudens S, Nachin L, Fontecave M, Barras F (2003) Biogenesis of Fe–S cluster by the bacterial Suf system: SufS and SufE form a new type of cysteine desulfurase. J Biol Chem 278:38352–38359

113. Ollagnier-de-Choudens S, Lascoux D, Loiseau L, Barras F, Forest E, Fontecave M (2003) Mechanistic studies of the SufS–SufE cysteine desulfurase: evidence for sulfur transfer from SufS to SufE. FEBS Lett 555:263–267

114. Wollers S, Layer G, Garcia-Serres R, Signor L, Clemancey M, Latour JM, Fontecave M, Ollagnier de Choudens S (2010) Iron–sulfur (Fe–S) cluster assembly: the SufBCD complex is a new type of Fe–S scaffold with a flavin redox cofactor. J Biol Chem 285:23331–23341

115. Outten FW, Wood MJ, Munoz FM, Storz G (2003) The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe–S cluster assembly in *Escherichia coli*. J Biol Chem 278:45713–45719

116. Wada K, Sumi N, Nagai R, Iwasaki K, Sato T, Suzuki K, Hasegawa Y, Kitaoka S, Minami Y, Outten FW, Takahashi Y, Fukuyama K (2009) Molecular dynamism of Fe–S cluster biosynthesis implicated by the structure of the SufC$_2$–SufD$_2$ complex. J Mol Biol 387:245–258

117. Balk J, Lobreaux S (2005) Biogenesis of iron–sulfur proteins in plants. Trends Plant Sci 10:324–331

118. Lill R, Mühlenhoff U (2006) Iron–sulfur protein biogenesis in eukaryotes: components and mechanisms. Annu Rev Cell Dev Biol 22:457–486

119. Escalante AA, Ayala FJ (1995) Evolutionary origin of *Plasmodium* and other Apicomplexa based on rRNA genes. Proc Natl Acad Sci USA 92:5793–5797

120. Kuo CH, Wares JP, Kissinger JC (2008) The apicomplexan whole-genome phylogeny: an analysis of incongruence among gene trees. Mol Biol Evol 25:2689–2698

121. Surolia N, Padmanaban G (1992) De novo biosynthesis of heme offers a new chemotherapeutic target in the human malarial parasite. Biochem Biophys Res Com 187:744–750

122. Neuhaus HE, Emes MJ (2000) Nonphotosynthetic metabolism in plastids. Annu Rev Plant Physiol Plant Mol Biol 51:111–140

123. Rider SD Jr, Zhu G (2010) *Cryptosporidium*: genomic and biochemical features. Exp Parasitol 124:2–9

124. Clopton RE (2002) Phylum Apicomplexa Levine, 1970: order Eugregarinorida Léger, 1900. In: Lee JJ, Leedale G, Patterson D, Bradbury PC (eds) Illustrated guide to the protozoa. Society of Protozoologists, Lawrence, pp 205–288