2009

# DNA replication in the human malaria parasite and potential for novel drug development

Li, J

http://hdl.handle.net/10026.1/10331

Royal Society of Chemistry

CHAPTER 13

# *DNA Replication in the Human Malaria Parasite and Potential for Novel Drug Development*

JI-LIANG LI

Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK and School of Biotechnology, Nanfang Medical University, Guangzhou, China

## 13.1 Introduction

Malaria threatens 40% of the world's population and kills 1–3 million humans, mostly children under five, every year.[1] The lack of effective vaccines and the spreading of drug-resistant parasites has led to a resurgence of malaria in many countries.[2,3] It is therefore imperative that our understanding of the fundamental biology and biochemical processes at different stages of the life cycle of the malaria parasite be improved, to facilitate the identification of new targets for the development of novel drugs.

DNA replication represents one such key biochemical process of the parasite. It takes place in at least five distinct points in the parasite life cycle, two of which occur in the human host (*i.e.* exo-erythrocytic and erythrocytic schizogony in liver and red blood cells, respectively); the remainder take place in the mosquito host (*i.e.* gametogenesis in midgut, pre-meiosis after fertilisation and sporogony in oocysts)[4] (Figure 13.1). Thus the selective disruption of DNA synthesis in the parasite might not only inhibit the disease itself but also block parasite transmission by the mosquito vector. During gametogenesis, three successive rounds of genome replication are completed within 10 minutes, raising the DNA content

Molecular Themes in DNA Replication
Edited by Lynne S. Cox

Published by the Royal Society of Chemistry, www.rsc.org

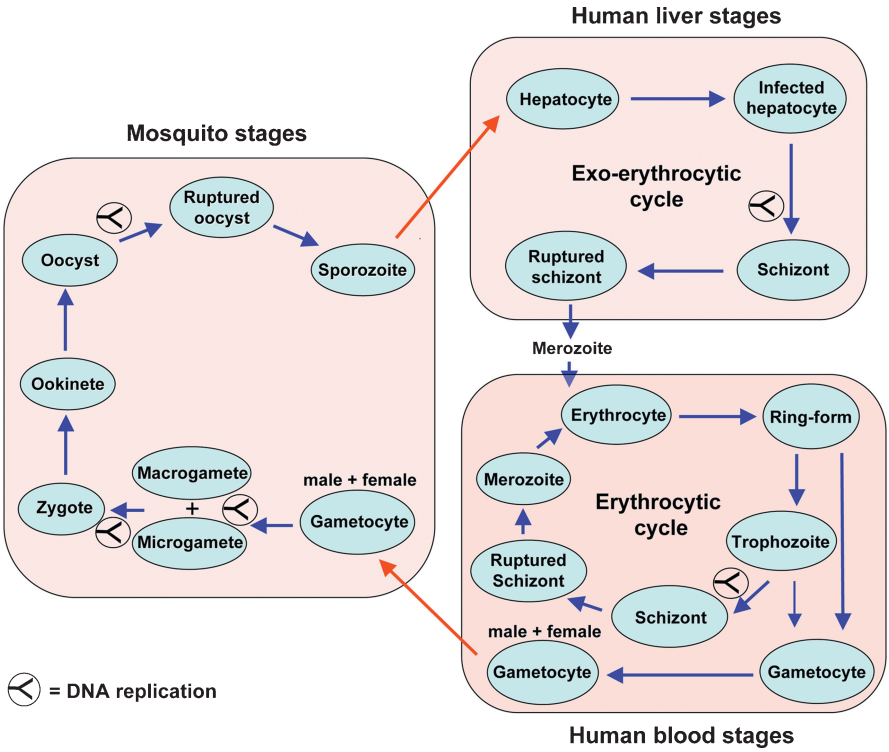



**Figure 13.1** ***Plasmodium falciparum* life cycle.** The life cycle of the malarial parasite involves two hosts: mosquito and human. It may be divided into three stages. One is sexual and takes place inside the mosquito, while the other two are asexual and occur inside the human host: the erythrocytic cycle (in red blood cells) and the exo-erythrocytic cycle (in liver cells). During a blood meal, a malaria-infected female *Anopheles* mosquito inoculates sporozoites into the human host. Sporozoites, upon entering the bloodstream, will reach the liver (hepatocytes) where they mature into schizonts, which rupture and release merozoites. After this initial replication in the liver (exo-erythrocytic schizogony), the parasites undergo asexual multiplication in red blood cells (erythrocytic schizogony). Merozoites infect erythrocytes becoming ring-forms. The ring-forms progress to trophozoites and mature into schizonts. The segmented schizonts rupture releasing merozoites. Some parasites in red blood cells differentiate into sexual erythrocytic stages (gametocytes). The gametocytes, including male (microgametocytes) and female (macrogametocytes), are ingested by an *Anopheles* mosquito during a blood meal. The microgametocytes undergo exflagellation in the midgut generating microgametes. The microgametes fertilise the macrogametes producing zygotes. The zygotes in turn become motile and elongated to form ookinetes that invade the midgut wall of the mosquito where they develop into oocysts. The oocysts grow, rupture and release sporozoites, which migrate to the salivary glands. Inoculation of the sporozoites by the mosquito into a new human host continues to transmit the parasites.

to octoploid values just before exflagellation, suggesting that the genome duplication rate of the parasite is extremely high. Assuming that the rate of replication fork movement in the parasite is similar to that in other eukaryotes, at least 1300 replication origins among the 14 chromosomes of *Plasmodium* would be needed to achieve this rate of replication.[5] It is known that at least two replication origins exist within each inverted repeat region of the 35 kb apicoplast genome of the parasite[6–8] but information concerning replication origins on nuclear chromosomes of the parasite remains absent.

This chapter focuses on the important molecules (Tables 13.1, 13.2 and 13.3) involved in the DNA replication initiation and elongation processes, and their potential as targets for development of novel drugs against *Plasmodium falciparum*—the most malignant pathogen of four malarial parasite species that infect humans.

## 13.2 Replication Initiation Proteins

### 13.2.1 PfORC

The initiation of DNA replication is tightly regulated such that all chromosomes are replicated precisely at specific points in the parasite life cycle (Figure 13.1). In eukaryotes, ORC is a multi-subunit protein complex composed of six polypeptides (ORC1-6) that binds to replication origins and is essential for the initiation of chromosomal DNA replication (see Chapters 1 and 2). It was originally thought that only two or three PfORC homologues exist in *P. falciparum*.[9–11] However, extensive bioinformatics investigations reveal that there may be four putative homologues (*i.e.* PfORC1, 2, 4 and 5) in the parasite genome.

PfORC1 was the first malarial ORC isolated from *P. falciparum* using vectorette technology.[12] PfORC1 is composed of two distinct domains: a variable N-terminal domain (residue positions 1–783) and a highly conserved C-terminal domain (residue positions 784–1189). The N-terminal domain is the largest extension in the ORC1 family and contains several other unique characteristics:

(i) It is rich in serine/threonine and tyrosine (20%), forming a number of potential phosphorylation sites for a range of known protein kinases such as cyclin-dependent kinases, suggesting that PfORC1 may be regulated by phosphorylation.
(ii) It contains a large proportion (30%) of charged amino acid residues (K, R, E and D) that may be involved in protein–protein interactions.
(iii) It has four copies of a heptamer ISSSLT(S)N repeat.
(iv) It includes two putative nuclear localisation signal (NLS) motifs, suggesting that PfORC1 may be a nuclear protein.
(v) It possesses a leucine zipper motif ($LX_6LX_6LX_6L$) that may be involved in DNA-binding activity.

The C-terminal domain of PfORC1 shares 48–61% similarity and 27–40% identity with members of the ORC1 family from other species, and contains all

**Table 13.1** DNA replication initiation proteins in *Plasmodium falciparum*.

| *Replication initiation protein* | *Transcript size (kb)* | *Predicted no. of amino acids* | *Protein molecular weight (kDa)* | *Stage-specific expression* | *Chromosome location* | *Database accession no./Gene ID* | *Reference* |
|---|---|---|---|---|---|---|---|
| PfORC1 | ~5.6 | 1189 | 138.7 | Ase, Gam | 12 | AF373219<br>PFL0150w | 10,12,13 |
| PfORC2 | | 825 | 97.9 | | 7 | MAL7P1.21 | |
| PfORC4 | | 983 | 117.7 | | 13 | PF13_0189 | |
| PfORC5 | | 899 | 103.9 | | 2 | PFB0720c | 9,63 |
| PfCDT1 | | 982 | 114.5 | | 13 | PF13_0237 | |
| PfCDC6 | | 979 | 115 | | 5 | PFE0155w | |
| PfMCM2 | ~3.8 | 971 | 111.4 | Ase, Gam | 14 | AF095948<br>PF14_0177 | [a]<br>11 |
| PfMCM3 | | 962 | 109.7 | | 5 | PFE1345c | 68 |
| PfMCM4 | ~4.0 | 1005 | 115 | Ase, Gam | 13 | AF083323<br>PF13_0095 | 11,13,21 |
| PfMCM5 | ~3.8 | 758 | 85.7 | Ase, Gam | 12 | AF139108<br>PFL0580w | [a]<br>62 |
| PfMCM6 | | 929 | 105.5 | | 13 | PF13_0291 | 11 |
| PfMCM7 | | 821 | 94.1 | | 7 | PF07_0023 | 11 |
| PfMCM8 | | 1024 | 119.2 | | 12 | PFL0560c | 62 |
| PfMCM9 | | 1465 | 171.2 | | 4 | PFD0790c | 68 |

Ase = sexual stages; Gam = gametocyte. [a]Data from unpublished research by J-L. Li.

**Table 13.2** Replication elongation factors in *Plasmodium falciparum*.

| *Elongation factor* | *Transcript size (kb)* | *Predicted no. of amino acids* | *Protein molecular weight (kDa)* | *Stage-specific expression* | *Chromosome location* | *Database accession no./Gene ID* | *Reference* |
|---|---|---|---|---|---|---|---|
| PfPCNA1 | ~1.6; ~2.2 | 275 | 30.6 | Ase, Gam | 13 | P31008<br>PF13_0328 | 46 |
| PfPCNA2 | ~1.8; ~2.5<br>~2.0 | 264 | 30.2 | Ase, Gam | 12 | AF056205<br>AF544241<br>PFL1285c | 47,48 |
| PfRFC1 | ~4.0 | 904 | 104.2 | Ase, Gam | 2 | AF139827<br>PFB0895c | [a] |
| PfRFC2 | ~1.85 | 330 | 38.0 | Ase, Gam | 2 | AF071409<br>PFB0840w | [a]<br>63 |
| PfRFC3 | ~1.8 | 344 | 39.2 | Ase, Gam | 14 | AF069296<br>PF14_0601 | [a] |
| PfRFC4 | ~1.75 | 336 | 37.7 | Ase, Gam | 12 | AF126257<br>PFL2005w | [a] |
| PfRFC5 | | 349 | 40.3 | | 11 | PF11_0117 | 62 |
| PfFen1 | ~2.4<br>~3.2 | 672<br>650 | 76.6 | Gam<br>Ase, Gam | 4 | AF093702<br>PFD0420c<br>MAL4p2.21<br>AF278764 | [a]<br><br><br>51 |
| PfRNase H2 | | 288 | 33.0 | | 6 | PFF1150w<br>MAL6P1.190 | 68 |
| PfRPA1 | ~6.5 | 1145 | 134.1 | Tro, Sch | 4 | AL035475<br>Ma14p2.32<br>PFD0470c | 29 |
| PfRPA1′ | | 484 | 56.1 | | 9 | PFI0235w | 68 |
| PfRPA2 | | 273 | 31.6 | | 11 | PF11_0332 | 62 |
| PfRPA2′ | | 191 | 22.3 | | 11 | PF11_0130 | 62 |
| PfRAP3 | | 135 | 15.4 | | 7 | PF07_0039 | |

**Table 13.2** (*Continued*).

| *Elongation factor* | *Transcript size (kb)* | *Predicted no. of amino acids* | *Protein molecular weight (kDa)* | *Stage-specific expression* | *Chromosome location* | *Database accession no./Gene ID* | *Reference* |
|---|---|---|---|---|---|---|---|
| PfLigI | | 912 | 104 | | 13 | MAL13P1.22 | 52 |
| PfPolα | ~7.0 | 1855 | 205 | Ase, Gam | 4 | L18785<br>PFD0590c | 37 |
| PfPolaB | | 539 | 63.9 | | 14 | PF14_0602 | |
| PfPriL | | 525 | 62.2 | | 9 | PFI0530c | 68 |
| PfPriS | ~2.1 | 452 | 53 | Ase | 14 | X99254<br>PF14_0366 | 39 |
| PfPolδ | ~5.2, ~5.7<br>~4.5 | 1094 | 126.8 | Ase, Gam<br>Tro, Sch | 10 | X62423<br>M64715<br>PF10_0165 | 41,42 |
| PfPolδS | | 498 | 57.6 | | 3 | PFC0340w<br>MAL3P3.4 | 69 |
| PfPolε | | 2907 | 344.5 | | 6 | PFF1470c<br>MAL6P1.125 | 68 |
| PfPolεB | | 624 | 73.5 | Ase | 12 | PFL1655c | 19,70 |

Ase = sexual stages; Gam = gametocyte; Sch = schizont; Tro = trophozoite. [a]Data from unpublished research by J-L. Li.

**Table 13.3** Proteins involved in apicoplast replication in *Plasmodium falciparum*.

| *Apicoplast DNA replication factor* | *Transcript size (kb)* | *Predicted no. of amino acids* | *Protein molecular weight (kDa)* | *Stage-specific expression* | *Chromosome location* | *Database accession no./ Gene ID* | *Reference* |
|---|---|---|---|---|---|---|---|
| PfPOM1 (PfPREX) | ~7.0 (Pc) | 2016 | 235.7 | Tro, Sch Gam | 14 | PF14_0112 | 61 |
| PfGyrA | | 1222 | 143 | | 12 | PFL1120c | 56,57 |
| PfGyrB | | 1006 | 116 | | 12 | PFL1915w | 56,57 |

Gam = gametocyte; Sch = schizont; Tro = trophozoite.

conserved sequences of the family including the Walker A (GMPGTGKT, 815–822) and Walker B (DEID, 903–906) motifs of the classical purine nucleotide-binding sites (see Chapter 2), suggestive of ATP binding and hydrolysis activity of PfORC1 (Figure 13.2A). Indeed, ATPase activity of the recombinant C-terminal domain has been recently confirmed *in vitro* and, like other $AAA^+$ ATPases, is dependent on the K821 residue within the Walker A motif.[10] In addition, PfORC1, like most members of the ORC1 family from other species, contains a putative PCNA binding motif, **QKVLFTLF** (913–920), suggestive of an interaction with PfPCNA (see Chapter 3). In fact, PfORC1 has been shown to interact with PfPCNA1 *in vitro* and to co-localise with the majority of the PfPCNA1 replication foci during trophozoite stages.[9] Expression of PfORC1 has been demonstrated in the parasite nucleus during late trophozoite and early schizont stages and in sexual stages,[9,10,12–15] consistent with its potential role in replication initiation in these stages. Recently, it was reported that PfORC1 can specifically interact with telomeres and with various subtelomeric repeats and, thus, act also as a telomere-associated protein in *P. falciparum*.[15] More than 80% of PfORC1 nuclear foci were shown to co-localise with PfSir2, a telomeric protein that was identified as reversibly associating with the promoter regions of silent but not active subtelomeric *var* genes.[16,17] Therefore, PfORC1 may cooperate with PfSiv2 and contribute to telomeric silencing of virulence-factor genes and antigenic variation in *P. falciparum* by a phenomenon called the telomere position effect.[18] Structural modelling of PfORC1 against archeal ORC/cdc6 suggests very strong conservation such that plasmodium PfORC1 may similarly distort the helical axis of DNA (Figure 13.2B), thought to be important for localised melting of the origin during replication initiation (see also Figure 12.1).

PfORC2, like PfORC1, shares 45% similarity and 26% identity with human ORC2 only at the C-terminal region (residue positions 488–825) (J-L. Li unpublished data). PfORC2 is transcribed during late trophozoite and early schizont stages,[14,19] consistent with a role in DNA replication.

PfORC4 contains the Walker A (GMLGCGKT, 160–167) and Walker B motifs (DEND, 270–273) characteristic of other ORC family members and consistent with a proposed action as an $AAA^+$ ATPase (see Chapter 2).

**Figure 13.2** **PfORC1.** (**A**) The ORC1/Cdc6/Cdc18 domain. Sequences were aligned with the CLUSTAL W (1.85) multiple sequence alignment program,[71,72] using data from GenBank/EMBL/DDJB database accession numbers: *P. falciparum* PfORC1, AF373219; human HsORC1, Q13415; and human HsCdc6, U77949. Six conserved blocks that define PfORC1 as a member of the ORC1/Cdc6/Cdc18 family are highlighted with grey and labelled at the bottom of the sequence. A putative PCNA-binding motif (913–920) is highlighted in black and underlined at the bottom of the sequence. (**B**) Predicted structure of PfORC1. Molecular modelling of PfORC1 (green) was conducted using SWISS-MODEL[73,74] and MacPyMOL (http://delsci.com/macpymol/) with archaeal ORC/Cdc6 (PDB accession number 2QBY) (red in figure) as a template.

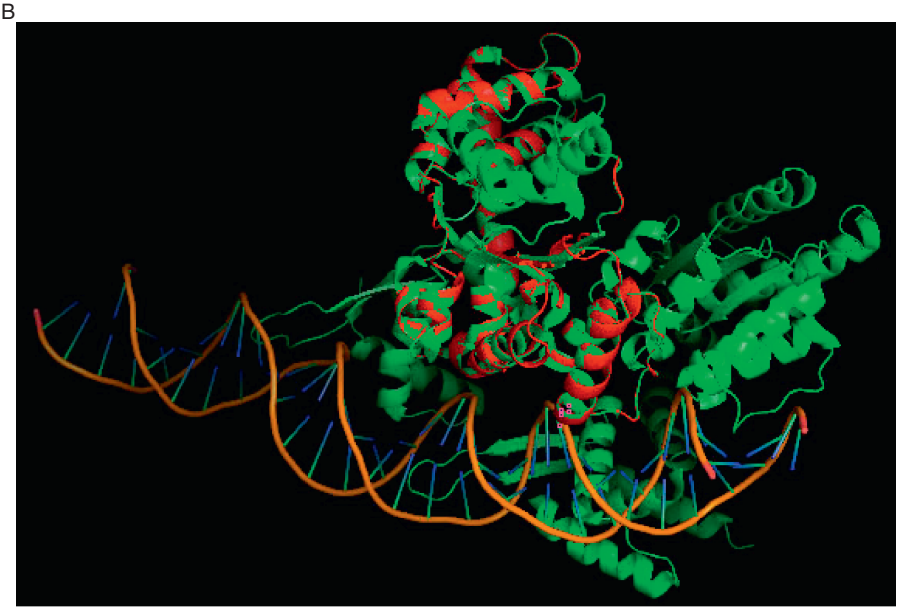
A

```
                  10        20        30        40        50        60        70        80
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  xExxxxxLxxxxxVPExLPxREKExxxIxxFLxxxIxxxxxxxxxLYISGxPGTGKTAxVxxVIQxLQxxxRxxxxxxFxxxx
HsORC1     EEARLRLHVSAVPESLPCREQEFQDIYNFVESKLLDHTGG-CMYISGVPGTGKTATVHEVIRCLQQAAQANDVPPFQYIE
PfORC1     DKAIRMMQLDVVPKYLPCREKEIKEVHGFLESGIKQSGSNQILYISGMPGTGKTATVYSVIQLLQIKSRKKLLPSFNVFE
HsCdc6     YQQAKLVLNTAVPDRLPAREREMDVIRNFLREHICGKKAG-SLYLSGAPGTGKTACLSRILQDLKKELKG-----FKTIM
                                     Block 1 (Walker A)
                  90       100       110       120       130       140       150       160
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  INxMxLxxxxxxYxxxxQQLxxQKxxSxxxxxELLxKxxxxxxxxxxxxxVLIIDELDxLxTxxQxVLYxLFDWPxxxxS
HsORC1     VNGMKLTEPHQVYVQILQKLTGQK-ATANHAAELLAKQFC-TRGSPQETTVLLVDELDLLWTHKQDIMYNLFDWPTHKEA
PfORC1     INGMNVVHPNAAYQVFYKQLFNKKPPNALNSFKIIDRLFNKSQKDNRDVSILIIDEIDYLITKTQKVLFTLFDWPTKINS
HsCdc6     LNCMSLRTAQAVFPAIAQEICQEE-VSRPAGKDMMRKLEKHMTAEKGPMIVLVLDEMDQLDSKGQDVLYTLFEWPWLSNS
                 Block 2                                   Block 3 (Walker B)
                 170       180       190       200       210       220       230       240
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  xLVLIxIANTMDLxDRILxxRxxSRxxxxxxxLxFxPYxxNQIxxILRxRLxxxxxxxxxDNxAIQxxARKVAxVSGDxR
HsORC1     RLVVLAIANTMDLPERIMMNRVSSRLGLT--RMCFQPYTYSQLQQILRSRLKH--LKAFEDD-AIQLVARKVAALSGDAR
PfORC1     KLILIAISNTMDLPDRLIP-RCRSRLAFG--RLVFSPYKGDEIEKIIKERLEN--CKEIIDHTAIQLCARKVANVSGDIR
HsCdc6     HLVLIGIANTLDLTDRILP-RLQAREKCKPQLLNFPPYTRNQIVTILQDRLNQVSRDQVLDNAAVQFCARKVSAVSGDVR
                 Block 4                           Block 5                           Block 6
                 250       260       270       280       290       300       310       320
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  KxLxICRRAxxxxxxxxxxxxxIxxxSxxxxxSxxxxPxxxxIxHxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxI
HsORC1     RCLDICRRAT-----------EICEFSQQKPDS----PGLVTIAHSMEAVDEMFSSSYITA-IKNSSVLEQSFLR---AI
PfORC1     KALQICRKAFENKRGHKIVPRDITEATNQLFDS----PLTNAINYLPWAFKIFLTCLIIELRIINEFVIPYKKVVNRYKI
HsCdc6     KALDVCRRAIEIVESDVKSQTILKPLSECKSPSEPLIPKRVGLIHISQVISEVDGNRMTLSQEGAQDSFPLQQKILVCSL

                 330       340       350       360       370       380       390       400
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  LxxxxxxxxxxxxxxxLxExxxxKLxxxxxxLxRxxxIxxxxxSExxSxxxxLxxxxxxxxxxxxxxxxxxxVxxxxNKxx
HsORC1     LAEFR--------RSGLEEATFQQIYSQHVALCRMEGLPYPTMSETMAVCSHLGSCR-----------LLLVEP--SRND
PfORC1     LIETSGKYIGMCSDNELFKIMLDKLVKMGILLIRPY-IPLENISKNKSKEALLGFNESSKKGNNQKITRAQVSPDIDKES
HsCdc6     MLLIR--------QLKIKEVTLGKLYEAYSKVCRKQQVAAVDQSECLSLSGLLEARG-------------ILGLKRNKET

                 410       420       430       440       450       460       470       480
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  xxxxVxxxVxxxxIxxALxxExxxxxxxxxxxx
HsORC1     LLLRVRLNVSQDDVLYALKDE------------
PfORC1     GDMGIELNVETQLIITALMKDPDCSKKLNFY--
HsCdc6     RLTKVFFKIEEKEIEHALKDKALIGNILATGLP
```

B



The central region (residues 431–660) of PfORC4 displays up to 31% identity and 50% similarity with other members of ORC4 family. Oligonucleotide array analysis showed that PfORC4 is expressed not only in trophozoite and early schizont stages,[19] but also in sexual and sporozoite stages.[14]

PfORC5, like PfORC1, comprises two distinct domains: a variable N-terminal domain and a highly conserved C-terminal domain. The N-terminal extension contains a putative NLS motif (residue positions 56–72) and a D/N/K-rich repeat (residue positions 81–170). The C-terminal domain of PfORC5 exhibits 25% identity and 46% similarity (between the 572–815 region) with *Drosophila* ORC5. PfORC5, like *Saccharomyces cerevisiae* ORC5 (ScORC5), contains the ATP-binding motif (GLPGMGKT, 303–310) but lacks the Walker B motif designated as a nucleotide hydrolysis domain (see Chapter 2). Between the ATP-binding motif and the C-terminal conserved region, however, five sequence insertions ranging from seven to 56 amino acid residues are found, the largest one having the D/N/K-rich repeat. Surprisingly, PfORC5 has been demonstrated to have ATPase activity, *i.e.* not only binding but also hydrolysis of ATP.[9] Chimeric ORC5 (composed of the N-terminus of ScORC5 and the C-terminus of PfORC5), but not full-length PfORC5, was shown to be able to complement an ScORC5-deficient yeast strain, confirming that PfORC5 is a true homologue of ORC5.[9] PfORC5 is predominantly expressed in trophozoite and early schizont stages, and in sexual stages.[9,14] It was reported that PfORC5 co-immunoprecipitates with PfPCNA1 (see Section 13.3.4) from mixed trophozoite extract, and co-localises with PfPCNA1 replication foci *in vivo* during early-to-mid replicating trophozoite stages, with PfORC5 and PfPCNA1 starting to dissociate from each other during further growth progression and finally separating completely during late schizont stages.[9] PfORC5 was also shown to co-localise with PfORC1 foci from mid-trophozoite to mid-schizont stages until PfORC1 was completely diminished in late schizont stages.[9]

### 13.2.2 PfCDC6

CDC6 is an ATP-binding protein that plays a crucial role in the assembly of the pre-replication complex (pre-RC) (see Chapters 1 and 2). CDC6 requires ORC to associate with chromatin and is in turn required for MCM2-7 chromatin association. The CDC6 homologue was thought to be absent in *P. falciparum*;[10,11] however, our extensive bioinformatics studies reveal that a gene located on chromosome 5 does encode a PfCDC6-like protein in the parasite. Like PfORC1 (Figure 13.2A), there are six blocks in the central region of PfCDC6 showing strong similarity with members of the ORC1/CDC6/Cdc18 family including the Walker A (GPSGQGKT) and B (DELD) motifs,[12] suggestive of an ATPase activity of PfCDC6. PfCDC6 contains the conserved region corresponding to domains I and II of archaeal Cdc6 and can presumably form a two-lobed, cashew-shaped molecule as does archaeal Cdc6.[20]

Interestingly, between the first two conserved blocks, there is an amino acid insertion of more than 50 residues, which may present a target for development of novel anti-malarial drugs. PfCDC6 seemed to be predominantly expressed in gametocyte but weakly in late trophozoite stages.[14,19]

### 13.2.3 PfCDT1

CDT1 protein, like CDC6, is required to load the MCM helicase at the replication origin to form the licensed pre-RC (see Chapter 1). Our bioinformatics analysis reveals that there is a putative PfCDT1 protein in the parasite genome. PfCDT1 shares up to 21% identity and 43% similarity with other members of the Cdt1 family between residues 544–700, although two short helices (H1 and H2) and loop L1 in mouse Cdt1 that contact with the N-terminal part of the geminin dimer are absent in PfCDT1. Perhaps this is less surprising in light of the absence of any identifiable geminin-like protein in the parasite genome (as is also the case in yeast; geminin may be a feature of metazoan replication control). Near the C-terminus is a leucine zipper motif ($LX_6LX_6LX_6L$, positions 921–950), suggesting that PfCDT1 may possess DNA-binding activity. It will be interesting to investigate whether PfCDT1 is required in replication initiation as would be predicted from its similarity to Cdt1 from other species. Consistent with a role in DNA replication, PfCDT1 is transcribed in both asexual (trophozoite and schizont) and sexual (gametocyte) stages.[14]

### 13.2.4 PfMCM

In eukaryotes, the MCM complex is composed of six conserved proteins (MCM2-7) and is recruited to replication origins by Cdc6 and Cdt1 to function as the replicative DNA helicase (see Chapter 3). Although PfMCM4 was the first malarial MCM member isolated from the parasite,[21] all six MCMs (PfMCM2-7) have now been identified in the parasite, the smallest being PfMCM5 (758 amino acid residues, predicted molecular mass 85.7 kDa) and the largest being PfMCM4 (1005 residues with predicted molecular mass 115 kDa) (Table 13.1). Although it has been reported that none of the PfMCM genes contain introns,[11] we have found that four out of six (*i.e.* PfMCM2, 3, 5 and 7) contain between one and three introns (Li *et al.,* unpublished data).

Comparisons of the PfMCM sequences reveal that the central region of approximately 200 amino acids is conserved in all six members, including the Walker A and B motifs and the MCM signature motif (IDEFDKM). In addition, a zinc finger motif ($CX_2CX_{18}CX_{2-4}C$), characteristic of DNA-binding domains, is present in the N-terminal region of four PfMCMs (PfMCM2, 4, 6 and 7), as would be expected for these four members, although there is a slight variation for the motif in PfMCM2 ($CX_2CX_{21}CX_2C$) and in PfMCM4 ($CX_{13}CX_{18}CX_2C$). Moreover, all PfMCMs contain various unique sequence insertions ranging from five to 84 residues compared with members of each

MCM family in other species. For example, PfMCM4 has five unique amino acid insertions with sizes from five to 84 residues located in different places in the sequence[21] (see Figure 13.3).

Co-immunoprecipitation demonstrated that PfMCM2, 6, and 7 are present in a protein complex; among them, PfMCM6 is the only protein that is tightly associated with chromatin, suggesting that PfMCM6 may directly interact with chromatin, while other PfMCMs probably associate *via* protein-protein interactions.[11] PfMCM2-6 all appear to be transcribed during trophozoite, schizont and gametocyte stages[14,19,21] while some (*i.e.* PfMCM3, 4, 5 and 7) are also expressed in sporozoites.[14] At the protein level, PfMCM2 has been demonstrated to peak in late schizont/segmented schizont stages, PfMCM6 in late trophozoite and schizont/segmented schizont stages, and PfMCM7 in schizont/segmented schizont stages.[11] Immunofluorescent staining revealed the PfMCM4 protein in trophozoite, schizont and gametocyte stages.[13] In other eukaryotic systems, the MCM complex is regulated at replication initiation by phosphorylation by cyclin-dependent kinases (*e.g.* Clb5/6-CDK1 in yeast, together with Cdc7-Dbf4). Moreover, association of DNA pol α-primase at the pre-initiation complex is usually mediated by Cdc45.[22–24] However, despite extensive analysis, we have been unable to identify Cdc45 and Dbf4-Cdc7 (DDK) homologues in the parasite genome.

In this context, it is interesting to note that structural changes in yeast MCM5 can remove the need for Dbf4-kinase activation of the MCM complex at initiation.[25] It will therefore be interesting to see whether the PfMCM complex can similarly be activated without the need for Cdc7-Dbf4-dependent phosphorylation.

### 13.2.5 PfMCM8 and PfMCM9

MCM8 and MCM9 proteins are conserved in a diverse array of eukaryotes but are lacking in most fungi, *Caenorhabditis elegans* and *Giardia lamblia,* and thus were originally thought to be vertebrate-specific proteins[26] (see also Chapter 3). However, we have identified MCM8 and MCM9 homologues in the malarial parasite, and found that they are very similar to even their human orthologues (Li *et al.,* unpublished data): PfMCM8 shares 34% identity and 55% similarity with human hMCM8 in two separate regions and contains all MCM2-7 family

**Figure 13.3** **Sequence comparisons between *P. falciparum* and human MCM4.** The sequences were aligned with the CLUSTAL W (1.85)[71,72] using data derived from GenBank/EMBL/DDJB database accession numbers: *P. falciparum* PfMCM4, AF083323; human MCM4 (aka CDC21), P33991; and mouse MCM4, P49717. Inserts A–E are labelled and underlined at the bottom of sequence. The zinc-finger motif ($CX_nCX_{18}CX_2C$) is highlighted in grey and the Walker A and B motifs are highlighted and indicated at the bottom of sequence. Note that murine MCM4 is included to demonstrate the very high degree of conservation of MCM4 in mammals.

```
                    10        20        30        40        50        60        70        80
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  MxSPxxxxxRRxSxxxxxxxxxSxxxxxSxSxxxxxxxxxxxxSTxxxxxxxTxxxxxLxxxxxxNxxxSxxxxMxxxxx
MCM4_MOUSE MSSPASTPSRRSSRRGRVTPTQSLRSEESRSSPNRRRRGEDS-STGELLPMPTSPGADLQSPPAQNALFSSPPQMHSLAI
MCM4_HUMAN MSSPASTPSRRGSRRGRATPAQTPRSEDARSSPSQRRRGEDSTSTGELQPMPTSPGVDLQSPAAQDVLFSSPPQMHSSAI
PfMCM4     MGTPRRRLGQQNNNNNSPFALSSSNIFGSNNEIFGSNFMHTPMSSRRTKNSKSFLNSMLNESRYLNQSNAGSQFIKYGHT

                    90       100       110       120       130       140       150       160
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  PLxxxxxxxxxxxxxxxxxExxxxxxxxGxxxxxxxxxxxxxKxLxxxxxxDxxxxxxxxxxDxxxxEQSLxQxLVIxxx
MCM4_MOUSE PLDFDVSSPLTYGTPSSRVEGTPRSGVRGTPVRQRPDLGSARKGLQVDLQSDGAAAE-----DIVPSEQSLGQKLVIWGT
MCM4_HUMAN PLDFDVSSPLTYGTPSSRVEGTPRSGVRGTPVRQRPDLGSAQKGLQVDLQSDGAAAE-----DIVASEQSLGQKLVIWGT
PfMCM4     PLAIRRIKCARADIGDVGREAFMEDEESGRLPHFIDSNLEQIKELFNQFFDEFNITNYSDVLDFTDEDRSISEYILLHRD
                                                                          Insert A
                    170       180       190       200       210       220       230       240
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  DVxVxxxxxxFQxxLxxFxDxxxKxExxxxNxxxDxxExxxxxxxGxxNIxxxxxxxVNxxHIxxFNKxLYRxLIxYPxE
MCM4_MOUSE DVNVATCKENFQRFLQCFTDPLAKEE---ENVGIDITQPLYMQQLGEINITGEPFLNVNCEHIKSFSKNLYRQLISYPQE
MCM4_HUMAN DVNVAACKENFQRFLQRFIDPLAKEE---ENVGIDITEPLYMQRLGEINVIGEPFLNVNCEHIKSFDKNLYRQLISYPQE
PfMCM4     NLKVYLAYYGWK--MIKFIETGRQNECRLNNTNYEDDDENNENSEGIRNLEHIKSFEIDLTHIFFFNKKLYKLIIEYPSD

                    250       260       270       280       290       300       310       320
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  xIxxxDxxVNxxFxxxxxxxxxxxxxxxxxDRYPxxSxxExxxQVRxFNxxxxxSxRxLxPxxIDxLIxIxGMVIRxSxL
MCM4_MOUSE VIPTFDMAVNEIFF---------------DRYP-DSILEHQIQVRPFNALKTKSMRNLNPEDIDQLITISGMVIRTSQL
MCM4_HUMAN VIPTFDMAVNEIFF---------------DRYP-DSILEHQIQVRPFNALKTKNMRNLNPEDIDQLITISGMVIRTSQL
PfMCM4     CISEIDKIISTKYNSLLALVLEGDTRSSSSDKYPLSSTKQDYCRVRFFNKKHKDTPRKLGPNQIETLVCVKGVIIRCSNI
                          Insert B
                    330       340       350       360       370       380       390       400
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  IPEMxxAxFQCxxxxxxxxxxxQxCxxxxxxxIxRGxIxEPxxCxxCHxxHSxxLxHNxxxFSxKQMIKLxExxExMxxG
MCM4_MOUSE IPEMQEAFFQC-----------QVCAHTTRVEIDRGRIAEPCSCVHCHTTHSMALIHNRSFFSDKQMIKLQESPEDMPAG
MCM4_HUMAN IPEMQEAFFQC-----------QVCAHTTRVEMDRGRIAEPSVCGRCHTTHSMALIHNRSLFSDKQMIKLQESPEDMPAG
PfMCM4     IPEMTMAAFKCTSKKRIGVNNYEKCNEEVYEHVIQGEVQEPVTCSNCNNKNTFELWHNNCCFSSKQLIKLSEVTEHLKQG
                Insert C
                    410       420       430       440       450       460       470       480
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  QTPxTIxLFAHNDLVDxxQPGDxVxVTGIxRAxPIRVNPRxxxxxSVYKTHIDVIHxRKxDxKRLxxxDExxxxxxxxx
MCM4_MOUSE QTPHTIVLFAHNDLVDKVQPGDRVNVTGIYRAVPIRVNPRVSNVKSVYKTHIDVIHYRKTDAKRLHGLDE---------
MCM4_HUMAN QTPHTVILFAHNDLVDKVQPGDRVNVTGIYRAVPIRVNPRVSNVKSVYKTHIDVIHYRKTDAKRLHGLDE---------
PfMCM4     ETPQSISIYAYDDLIDYTKPGDTVELTGILKASPVRLNPRSRCYNSVHRTYINVIHIKKENKQKMKLTEQNDTANIILKR

                    490       500       510       520       530       540       550       560
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  xxxxxxxxxxxxxxExxxxLFSxKxVExLKELSRxPDIYERLxxALAPSIYxxxDIKKGILxQLFGGTRxxxSxxxxxKF
MCM4_MOUSE --------------EAEQKLFSEKRVKLLKELSRKPDIYERLASALAPSIYEHEDIKKGILLQLFGGTRKDFSHTGRGKF
MCM4_HUMAN --------------EAEQKLFSEKRVELLKELSRKPDIYERLASALAPSIYEHEDIKKGILLQLFGGTRKDFSHTGRGKF
PfMCM4     NEDGTVEENFEKLNEQGNLLFTTEVIQKMEQLSKDPNIYQRLVDSIAPSIYGRGDIKKGLLCQLFGGSK--ITDKYNNKY
           Insert D
                    570       580       590       600       610       620       630       640
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  RAEINILLxGDPxTSKSQLLxYVYxLxPRGxYTSGKGSSAVGLTAYVxKDxETRQxVLQTGALVLSDxGICCIDEFDKMN
MCM4_MOUSE RAEINILLCGDPGTSKSQLLQYVYNLVPRGQYTSGKGSSAVGLTAYVMKDPETRQLVLQTGALVLSDNGICCIDEFDKMN
MCM4_HUMAN RAEINILLCGDPGTSKSQLLQYVYNLVPRGQYTSGKGSSAVGLTAYVMKDPETRQLVLQTGALVLSDNGICCIDEFDKMN
PfMCM4     RSEIHILLRGDPSTAKSQLLHYVHKLSPRGIYTSGKGSSSVGLTAFISKDSETKEYILESGAVVLSDKGICCIDEFDKMD
           Walker A                                                   Walker B
                    650       660       670       680       690       700       710       720
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  ESxRSVLHEVMEQQTLSIAKAGIIxxLNARTSVLAAANPIxSQWNxxKxxIENIxLPxTLxSRFDLIFLMLDxxDExxDR
MCM4_MOUSE ESTRSVLHEVMEQQTLSIAKAGIICQLNARTSVLAAANPIESQWNPKKTTIENIQLPHTLLSRFDLIFLMLDPQDEAYDR
MCM4_HUMAN ESTRSVLHEVMEQQTLSIAKAGIICQLNARTSVLAAANPIESQWNPKKTTIENIQLPHTLLSRFDLIFLLLDPQDEAYDR
PfMCM4     DSARAILHEVMEQQTVTIAKAGIVATLNARTSILASANPINSRYDKNKAVVENINLPPSLFSRFDLIYLVIDQANEDEDR

                    730       740       750       760       770       780       790       800
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  RLAxxLxxxYxxSxEQxEEExxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
MCM4_MOUSE RLAHHLVSLYYQSEEQVEEE-----------------------------------------------------------
MCM4_HUMAN RLAHHLVALYYQSEEQAEEE-----------------------------------------------------------
PfMCM4     KLATVLCKNFSYNPEEEEDEDQEDQEEDEPNYITQQRARKSKGTSRKNERENYYNDGDNDDDDDISNYLNDSNDAQNKRG
                                                                    Insert E
                    810       820       830       840       850       860       870       880
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  xxxxxxxxxxxxxxxxxxxxxxxxxLDxxxLxxYIAYxxxTxxPxLSxEAxQxLIExYVxMRxxxxSRxxxSAxPRQLEx
MCM4_MOUSE -----------------------FLDMAVLKDYIAYAHSTIMPRLSEEASQALIEAYVNMRKIGSSRGMVSAYPRQLES
MCM4_HUMAN -----------------------LLDMAVLKDYIAYAHSTIMPRLSEEASQALIEAYVDMRKIGSSRGMVSAYPRQLES
PfMCM4     SWANVNISYDEYNNSSNKKTSKNYLIDSNTLALYIAYCRITCNPIISLESKKIIIEEYIKMRCKEGTKS-PTASPRQLEG

                    890       900       910       920       930       940       950       960
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  LIRLAEAxAKVRxxxxVxxxDxxEAxRLxxxAxxQSxxDPxTGxVDxxxLxxGxSAxxRKRxExLxExLxxxILxKxxTx
MCM4_MOUSE LIRLAEAHAKVRFSNKVEAIDVEEAKRLHREALKQSATDPRTGIVDISILTTGMSATSRKRKEELAEALRKLILSKGKTP
MCM4_HUMAN LIRLAEAHAKVRLSNKVEAIDVEEAKRLHREALKQSATDPRTGIVDISILTTGMSATSRKRKEELAEALKKLILSKGKTP
PfMCM4     LVRLSQSLAKMKLKRVVSPEEANEAVRLMNIATFQSLIDPLSGRIDFDQVNLGQTSQHKKKSDLIKDIIMNALVLKNMTK

                    970       980       990      1000      1010      1020      1030      1040
           ----:----|----:----|----:----|----:----|----:----|----:----|----:----|----:----|
Consensus  xxxxxxxxxExIxxxxQxxxAIxKxxFEEAxxxLxxxxxLTxxxxxVxxxx
MCM4_MOUSE ALKYQQLFEDIRG--QSDTAITKDMFEEALRALADDDFLTVTGKTVRLL-
MCM4_HUMAN ALKYQQLFEDIRG--QSDIAITKDMFEEALRALADDDFLTVTGKTVRLL-
PfMCM4     DELLTHCHETIMNDPQHTTSMDRKSFEEAFYDLEKSQEITRLCSGLYKKK
```

domains including the Walker A (GDPGLGKS) and Walker B (CIDELDKI) motifs as well as the conserved zinc finger motif ($CX_2CX_{18}CX_4C$). However, compared with hMCM8, PfMCM8 has three sequence insertions of six, 26 and 186 amino acid residues, respectively. In addition, near the N-terminus, there are six copies of a hexameric GNKN(Y)G(E)K repeat. PfMCM8 appears to be transcribed in trophozoite and schizont stages, but at slightly higher levels in the sporozoite.[14]

PfMCM9 is the largest of the PfMCM family members, comprising 1465 amino acid residues with a predicted molecular mass of 171.2 kDa. The *PfMCM9* gene is transcribed in late trophozoite and early schizont stages, and in sexual stages.[14,19] In five separate regions between residues 412–1400, PfMCM9 shares 37% identity and 57% similarity with human and mouse MCM9, including the Walker A (GDPGTGKS) and B (CIDEFCLM) motifs as well as the full MCM2-7 family domain. However, there are a number of unique amino acid insertions in PfMCM9. Compared with other members in the MCM9 family, PfMCM9 lacks the C-terminal region but has the large N-terminal extension. In addition, there are two types of amino acid sequence repeats, *i.e.* 13 copies of NN(Y)D(G)DNK at positions 149–226 and 12 copies of INV(G)D(N)D(N) at positions 721–780. How these repeats arose and whether they serve a function in parasite replication remain open questions.

## 13.3 Replication Elongation Proteins

### 13.3.1 PfRPA

Replication protein A (RPA) functions as a heterotrimeric complex that stabilises single-stranded DNA (ssDNA), and coordinates the sequential assembly and disassembly of DNA processing proteins on ssDNA during replication, and also during repair and transcription.[27,28] The largest subunit, RPA1 (hRPA70 in humans), contains the primary ssDNA-binding activity, while the two smaller subunits, RPA2 (hRPA32) and RPA3 (hRPA14), stabilise the complex and mediate interactions with replication and repair machinery,[28] (see Chapter 6). In *P. falciparum,* we have identified five putative subunits, *i.e.* two RPA1-type proteins (PfRPA1 and PfRPA1′), two RPA2-like molecules (PfRPA2 and PfRPA2′), and one RPA3 protein (PfRPA3) (Li *et al.,* unpublished data).

PfRPA1 was the first malarial RPA to be experimentally identified, through affinity purification of a 55 kDa factor possessing the major ssDNA-binding activity in asexual stage extracts.[29] Mass spectral analysis of 11 tryptic peptides demonstrated that the 55 kDa protein is the C-terminal fragment of PfRPA1.

PfRPA1 consists of 1145 amino acids with a predicted mass of 134.1 kDa, but it shares 30–39.2% identity with other eukaryotic RPA1 family members only within the C-terminal 466 amino acids. PfRPA1 possesses three ssDNA-binding domains:[30] DBD-A (positions 685–778), DBD-B (817–919) and DBD-C (958–1134); and a consensus zinc finger motif ($CX_2C_{26}CX_2C$, 986-1019) within DBD-C. However, in the N-terminal region of PfRPA1, no

significant homology to any known protein was identified; this non-conserved region may provide a suitable target for development of anti-malarial agents. PfRPA1 appears to be expressed in late trophozoite and schizont stages[29] as well as in sexual (gametocyte) and sporozoite stages.[14]

The second malarial RPA1, which we term PfRPA1′, consists of 484 amino acids with a calculated molecular mass of 56.1 kDa, and shares 26% identity and 50% similarity with PfRPA1. Like PfRPA1, PfRPA1′ contains the DNA binding domains DBD-A (10–109), DBD-B (140–247), DBD-C (275–451), and the zinc finger motif ($CX_2C_{28}CX_2C$, 315–350) characteristic of RPA1, but lacks the N-terminal protein interaction domain. PfRPA1′ seems to be expressed in trophozoite and schizont stages, and in sexual (gametocyte) and in sporozoite stages.[14,19]

Two putative malarial RPA2 subunits, PfRPA2 and PfRPA2′, consist of 273 and 191 amino acids with calculated masses of 31.6 and 22.3 kDa, respectively, and share 39% identity and 57% similarity each other in the central region. Importantly, PfRPA2 and PfRPA2′ each contain the consensus DBD-D DNA binding domain, together with several potential interface sites with RPA1 DBD-C and RPA3, and possible phosphorylation sites in their N-terminal regions. PfRPA2 seems to be weakly expressed in sporozoite and in early schizont and gametocyte stages, while PfRPA2′ is expressed in gametocyte and early trophozoite stages.[14]

The smallest malarial subunit, PfRPA3, consists of 135 amino acids with a predicted molecular mass of 15.4 kDa, and contains several putative interface sites with RPA2 DBD-D and RPA1 DBD-C. PfRPA3 appears to be expressed very weakly in gametocytes.[14]

## 13.3.2 DNA Polymerases

DNA polymerases α, δ and ε are the three established polymerases responsible for the bulk of DNA replication (see Chapter 4) that are sensitive to aphidicolin but insensitive to dideoxynucleotide analogues. There have been several reports on stage-dependent, aphidicolin-sensitive and aphidicolin-resistant DNA polymerase activities in the asexual stage extracts of *P. falciparum*.[31–33] DNA polymerase activity was undetectable in ring-form extracts, but increased in trophozoites and peaked in schizonts. Seven DNA polymerase activities have been identified and partially purified from *P. berghei*, five of which are aphidicolin-sensitive, while two are resistant.[34]

### *13.3.2.1 PfPolα*

The pol α-primase complex is responsible for making a chimeric RNA–DNA primer of ~40 nucleotides for initiating DNA synthesis, and is composed of a large catalytic subunit (Polα/p180), an intermediate B subunit (p70) and two small subunits (PriL/p55 and PriS/p48).[35] The largest subunit possesses the polymerase catalytic activity, while the two small subunits, PriS and PriL,

together function as the core primase for creating an RNA primer of 6–15 nucleotides.[36] The malarial Polα catalytic subunit (PfPolα), B subunit (PfPolαB) and two small primase subunits (PfPriS and PfPriL) have been found in the parasite genome.

PfPolα shares 14–17% overall identity only with other members of the Polα family but contains all seven motifs I–VII conserved in DNA polymerases, together with four of five Polα-specific motifs (A–E). However, PfPolα has a number of unique features. Firstly, motif A, present in the human counterpart, is absent in PfPolα. Secondly, in motif D, a highly conserved glycine (residue G493 in ScPolα), implicated in the interaction with DNA primase in yeast, is replaced by Leu in *P. falciparum*. Thirdly, there are at least six sequence insertions with sizes ranged from 10 to 91 amino acids interspersed between the polymerase-specific motifs, two of which contain repeated sequences. Insertion 2 has seven degenerate copies of repeat QQSVVS, while insertion 4 has four copies of the degenerate repeat KNIHSD. Oddly, the QQSVVS motif is shared with an RNA-directed RNA polymerase from bovine viral diarrhoea virus (Swissprot accession P19711). Fourthly, there are four Asn-rich tracts, three in the N-terminal region and one in the C-terminal region. Finally, PfPolα has a novel C-terminal extension of 98 residues compared with the human counterpart. PfPolα is transcribed in asexual and sexual stages.[37]

PfPolαB shares 24% identity and 48% similarity with its human counterpart in the region of 148–514. However, PfPolαB contains two unique sequence insertions with nine and 12 residues, respectively, compared with the human orthologue. PfPolαB seems to be expressed in asexual and sexual (gametocyte) stages, and the sporozoite stage.[14,19]

Primase in the parasite is encoded by two genes, *PfPriL* and *PfPriS*. PfPriL shares 33% identity and 53% similarity with its human orthologue, including a putative iron–sulfur cluster coordinated by four conserved cysteine residues (C362, 438, 456 and 496) that contributes to enzymatic activity of the core primase.[38]

PfPriS, encoded by a gene containing 15 introns, displays 36% identity and 53% similarity with its human counterpart. It contains all five domains (I–V) conserved in eukaryotic primase small subunits, including three catalytic aspartates (D130, 132 and 347) and the consensus sequence SGXRG (residues 181–185) involved in nucleotide binding. Additional conserved regions, termed Ia, VI and VII, are found in mammalian and yeast homologues.[39] However, PfPriS has three unique sequence insertions, two flanking domain Ia and one before domain V. Primase activity has been confirmed for recombinant PfPriS expressed in insect cells.[39] Both PfPriL and PfPriS appear to be expressed in trophozoites, schizonts, gametocytes and sporozoites.[14]

#### *13.3.2.2 PfPolδ*

DNA polymerase δ functions as a heterotetramer in fission yeast and mammals (see Chapter 4), and consists of a tightly associated dimer of a large catalytic

subunit (p125) and a small subunit (p50), which is associated with two other subunits (p68 and p12). However, in budding yeast, only three subunits have been identified.[40] In *P. falciparum*, two subunits, the large catalytic subunit (PfPolδ) and the small regulatory subunit (PfPolδS) have been identified.

PfPolδ shares 46% identity and 67% similarity with *Schizosaccharomyces pombe* Polδ between amino acids 97–1013.[41,42] PfPolδ contains all seven major motifs (I–VII) used to identify DNA polymerases in a correct spatial order, five additional regions (δ1-5) conserved in members of the Polδ family, and two putative zinc finger motifs ($CX_2CX_9CX_2C$, 1003-1019; $CX_2CX_9CX_4C$, 1049-1067). In addition, PfPolδ has the 3′–5′ exonuclease domain (residues 126–471) responsible for conferring proofreading activity, and retains several conserved residues (D308, I309, E310, Y390, D396 and D509) that presumably form the exonuclease catalytic site. PfPolδ is transcribed in asexual (trophozoite and schizont) and sexual (gametocyte) stages[41,42] as well as in sporozoites.[14] Western blotting analysis revealed that PfPolδ protein is expressed in trophozoite and schizont stages, but not in the ring stage.[43]

PfPolδS is slightly less conserved, showing up to 33% identity and 56% similarity with other members of the Polδ2 family. PfPolδS seems to be transcribed in asexual (trophozoite, schizont), sexual (gametocyte) and sporozoite stages.[14]

### *13.3.2.3 PfPolε*

Pol ε comprises four different subunits in vertebrates (Polε A–D) and budding yeast (Pol2/256kDa, Dpb2/79kDa, Dpb3/23kDa, and Dpb4/22kDa).[40] The DNA polymerase and 3′–5′ exonuclease of polymerase ε reside in the largest subunit. In *P. falciparum,* genes encoding both the catalytic subunit (PfPolε) and the second largest subunit (PfPolεB) have been found in the parasite genome. PfPolε displays significant homology to its human counterpart in four different regions:

- 30% identity and 54% similarity in the region of 27–175;
- 37% identity and 57% similarity in the region of 259–1406;
- 28% identity and 49% similarity in the region of 1729–1841;
- 25% identity and 46% similarity in the region of 2714–2899.

PfPolε contains all seven ordered motifs (I–VII) conserved in DNA polymerases in the N-terminal region, the exonuclease domain in the region of 337–635 including several conserved residues (D344, I345, E346, Y433, D439 and D533) that form the exonuclease catalytic site, plus two putative zinc finger motifs ($CX_2CX_{26}CX_2C$, 2782-2815; $CX_2CX_{11}CXC$, 2845-2863) conserved in the Polε family at the C-terminal side. In addition, a putative PCNA-binding motif, **QKKITSFF** (1312-1319), is found in the middle part of PfPolε, suggestive of an interaction with PfPCNA. It is noted that PfPolε also possesses numerous sequence insertions compared with its human counterpart; such additional sequences may prove useful in directed drug design.

PfPolεB shares 23% identity and 44% similarity with human PolεB in the region of 168–575. PfPolεB appears to be expressed in trophozoite and schizont stages.[14,19]

### 13.3.3 PfRFC

RFC consists of five non-identical subunits (RFC1–5) and is a DNA-dependent ATPase that functions as the clamp loader for PCNA (see Chapters 2 and 3) to confer processivity on DNA Pol δ and ε. We have identified all five subunits in the parasite and our Northern blot hybridisation shows that PfRFC1-5 are all expressed in asexual and sexual stages (J-L. Li and A. Goldsmith, unpublished data).

Each PfRFC subunit contains three domains (domains I–III) and the primary sequence of these domains is homologous among the five subunits; only PfRFC1, the large subunit, has additional N- and C-terminal regions. There is an overall similarity between the parasite PfRFC and corresponding orthologues in most other eukaryotes. All RFC boxes (I–VIII) occur in PfRFC, ranged from the N-terminus to the C-terminus.

- Box I is present only in PfRFC1 and consists of about 95 amino acids from positions 185–280. This region shows homology to the DNA ligase homology domain and BRCT, the BRCA1 C-terminus domain.
- Box II, at the N-terminus of PfRFC2-5, shows the WV(I,L)E-KYR(S)PXXL(I) consensus sequence.
- Box III, contains the phosphate-binding motif (P-loop) with the consensus sequence GXXGXGKT(S). It is notable that a classic GKT(S) sequence in the consensus P-loop is replaced by GKK in both human hRFC38 and yeast ScRFC5; however, PfRFC5 retains the motif (GKS) in the P-loop.
- Box IVcontains the L(I)EL(F)NASD sequence although a motif variation (LELQCFE) occurs in PfRFC5.
- Box V has the DEA(V)D motif, which for PfRFC5 is replaced by KDAE, suggesting that PfRFC5 might not be able to hydrolyse ATP.
- Box VI is different in the small and the large PfRFC subunits, as would be expected.
- Box VIa, only present in PfRFC1, has the consensus sequence of GMSsGDKGGstaI while box VIb, present in the PfRFC2-5 subunits, shows xM(L)T(S)xxAQxxL(M)RRI(t)M(I/L)E.
- Box VII, the arginine finger motif (SRC), is conserved in all small subunits (PfRFC2-5) but only the cysteine is present in PfRFC1, as would be expected for members of the RFC1 family.
- Box VIII has the consensus sequence GDL(M/I)RxxA(M/I)L(V)xxLQ, which for PfRFC5 is replaced by THGRNLRKVI.

These sequence characteristics suggest that all PfRFC subunits may bind ATP (*via* Walker A boxes), but that only PfRFC2-4 may effectively hydrolyse the bound ATP by virtue of their Walker B boxes.

### 13.3.4 PfPCNA

Proliferating cell nuclear antigen (PCNA) functions as a DNA sliding clamp that tethers the replicative DNA polymerases to the template, and also interacts with and regulates the activities of numerous proteins involved in DNA and chromatin processing during DNA synthesis (see Chapter 3). In yeast and mammals, only one form of PCNA has been found; this associates into a homotrimer with pseudo-sixfold symmetry.[44,45] By contrast, two different PCNAs have been described in the malarial parasite[46–48] (note that three different forms with subtly different interaction profiles have been characterised in Crenarchaea—see Chapter 12).

PfPCNA1 shares 29% identity and 53% similarity with PfPCNA2. PfPCNA1 and 2 are not only expressed in asexual trophozoite and schizont stages,[43,47,48] but also in the sexual gametocyte stage,[47] consistent with a role in rapidly proliferating cells. Both PfPCNA1 and PfPCNA2 contain all conserved motifs of members of the PCNA family, including the potential DNA-binding and protein interaction domains. PfPCNA2 may form a homotrimeric structure and PfPCNA1 may multimerise, but whether it forms dimers or trimers is unclear.[47,48]

Compared with human hPCNA, both PfPCNAs have a unique sequence insertion (PfPCNA1 with nine amino acid residues and PfPCNA2 with seven residues) just before the I2 motif. In addition, PfPCNA2 contains a second insertion with four amino acid residues located just before the βH1 region. We have conducted structural modelling of PfPCNA2 based on the published crystal structure for human PCNA,[44] and as shown in Figure 13.4, the additional regions present in PfPCNA2 but absent in hPCNA (white arrows in Figure 13.4) may be solvent-exposed and thus accessible to drugs.

### 13.3.5 PfFen1

Fen1, a structure-specific 5′ endo/exonuclease, cleaves a 5′-unannealed flap and degrades nucleotides from a nick or a gap on dsDNA. It plays a key role in the removal of RNA primers during Okazaki fragment maturation in lagging strand DNA synthesis (see Chapter 5), long-patch base excision repair and maintaining genome stability.

The malarial Fen1, PfFen1, shares up to 54% identity and 72% similarity with other members of the Fen1 family and consists of five regions. The N and I regions, like other members in the Fen1 family, constitute the nuclease core domain (residues 1–345) which contains the helical clamp region (residues 89–136), the helix-three turn-helix motif (residues 234–266), and seven critical acidic residues that cluster to form two active sites. By analogy to human FEN1, four of the acidic amino acids in PfFen1 (D34, D88, E160 and E162) are predicted to form the first metal ion-binding site, and three others (D181, D183 and D246) are likely to be involved in forming the second metal ion-binding site. The C region, located internally, contains the conserved PCNA-binding motif (**QRRLDNFF**, 350–357) that in fact forms a part of the βA–αA–βB

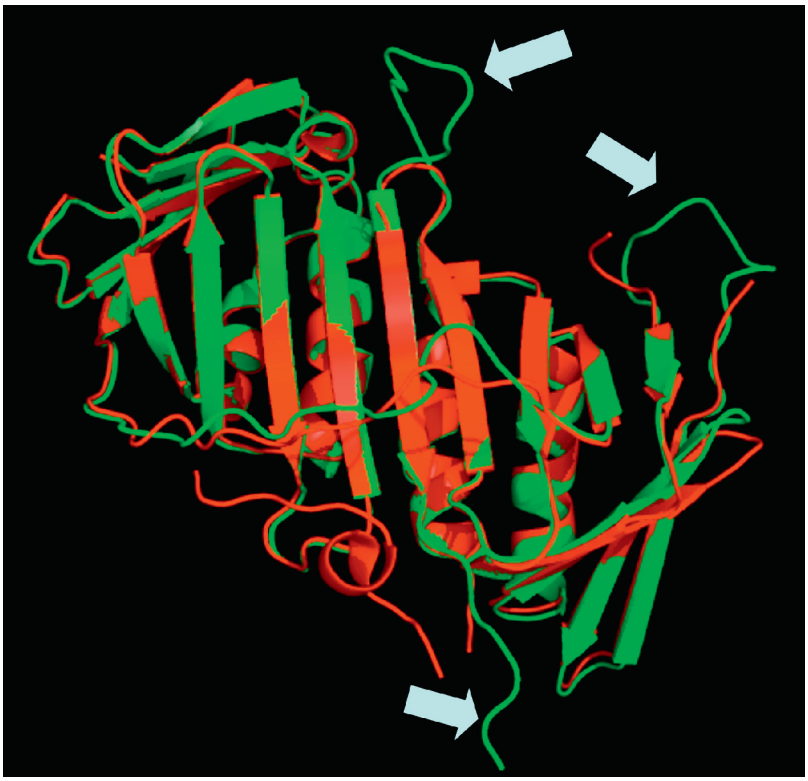


**Figure 13.4** **Predicted structural model of PfPCNA2 monomer.** Homology modelling of PfPCNA2 (green) was conducted using SWISS-MODEL[73,74] using the coordinates obtained for yeast PCNA[44] (PDB accession number 1PLR) as a template, and superimposed on human PCNA (red in figure) using MacPyMOL (http://delsci.com/macpymol/). The additional regions present in PfPCNA2 but absent in hPCNA are highlighted by white arrows.

structure, the main interacting interface with PCNA.[49,50] PfFen1 can generate a nicked DNA substrate that can be ligated by PfLigI.[51,52] PfFen1 seems to be transcribed in both asexual and sexual stages (J-L. Li, unpublished data). Western blot analysis of parasite lysates from the erythrocytic stages revealed that PfFen1 protein is present in all asexual stages.[51]

The R region, unique to PfFen1, has the DdeKXX hexamer repeated 12 times (residues 421–492). The E region (residues 493–672), the unique large-extended C-terminal segment, is present in a natively disordered state with no discernible secondary structure but possesses a putative NLS motif (residues 645–667). Although such unique extensions may appear to be prime targets for designing agents to inhibit malarial DNA replication without impacting on the function of essential host enzymes, functional analysis is critical in determining target validity. For example, the R and E regions of PfFen1 negatively influence

the enzyme's activities, since a PfFen1 C-terminal truncation consisting of about 400 amino acids and lacking the R and E regions possesses endonuclease and exonuclease activities 300-fold and 30-fold higher than full-length PfFen1, respectively.[51] Thus, these domains serve a regulatory role and their disruption may promote rather than block parasite DNA replication.

### 13.3.6 PfRNase H

RNase H specifically degrades the RNA moiety in RNA–DNA hybrids and plays an important role in DNA replication by removal of the RNA primer of Okazaki fragments (see Chapter 5). There are two major types of RNase H: RNase H1 is the smaller enzyme with major activity in prokaryotes, while RNase H2 is the RNase H large subunit which is most abundant in eukaryotes and archaea.[53]

Bioinformatics investigations failed to detect any RNase H1 candidate in the parasite genome (Li *et al.,* unpublished data), suggesting that RNase H1 might not be conserved in *P. falciparum*. However, a malarial orthologue of RNase H2, PfRNase H2, has been found in the parasite genome. PfRNase H2 is highly similar to the human RNase H large subunit, sharing 51% identity and 69% similarity; it retains all the conserved domains, including the active site triad residues (D22, D130 and D158) and the nearby serine (S168) and glutamate (E23) discovered in the archaeal orthologue.[53] An effort to define the functional capacity of PfRNase H2 is urgently needed.

### 13.3.7 PfLigase I

DNA ligase I constitutes the primary ligase utilised in DNA replication and plays an essential role in the joining of Okazaki fragments and nick sealing. The malarial DNA ligase I, PfLigI, shares 30% overall identity with human DNA ligase I, with even greater homology (60% identity) in the C-terminal region. PfLigI retains all of the conserved domains common to the ATP-dependent DNA ligase family, including important residues from motifs I–V that constitute the nucleotide binding pocket and motif VI which, together with motifs I–V, comprises the catalytic core of DNA ligase molecules.

In addition, PfLigI contains a bipartite nuclear localisation motif (residues 100–117), suggesting that PfLigI may be a nuclear protein. However, a unique apicoplast signal sequence, consisting of a signal peptide and a transit peptide, occurs in the N-terminus of PfLigI instead of the PCNA binding domain that normally exists in mammalian counterparts, suggesting that PfLigI might also be transported into apicoplasts (see Section 13.4.1). In addition, in the conserved region there is a sequence insertion consisting of about 100 amino acids compared with mammalian counterparts. PfLigI protein appears to be expressed in the asexual schizont stage.[52] Recombinant PfLigI expressed in HEK293 cells catalysed phosphodiester bond formation on a singly nicked

DNA substrate in an ATP-dependent manner, and joined RNA-DNA substrates only when the RNA sequence was upstream of the nick.[52]

## 13.4 Potential Targets for Novel Drug Development

*P. falciparum* contains most of the essential DNA replication components that are conserved in eukaryotes, suggestive of a similar core replication machinery of the parasite to that of the human host. However, the parasite seems to lack a number of core replication proteins including ORC3, ORC6, Cdc45, Dbf4-Cdc7, geminin and DNA2 helicase, suggestive of some differences in replication and regulatory pathways between the parasite and the human host, thereby presenting excellent targets for development of novel anti-malarial drugs.

### 13.4.1 Targeting Unique Replication Pathways: Apicoplast DNA Replication

The apicoplast, a chloroplast-like plastid present in *P. falciparum* and absent in humans, is no longer photosynthetic, but like chloroplasts, it contains its own genome and this is essential for parasite survival.[54] Apicoplast DNA replication is dependent on nuclear-encoded apicoplast-targeted proteins (Table 13.3). DNA gyrase, a typical type II topisomerase that can introduce negative supercoils in DNA, is required for parasite apicoplast DNA replication.[54,55] The malarial gyrase A and gyrase B, PfGyrA and PfGyrB, are encoded in the nucleus and then transported into the apicoplast. PfGyrB contains strong intrinsic ATPase activity and, together with the N-terminal domain of PfGyrA, can efficiently cleave supercoiled DNA.[56,57] The fluoroquinolone antibiotic, ciprofloxacin, that targets bacterial DNA gyrase, has been demonstrated to inhibit apicoplast DNA replication and growth of the parasite *in vitro*.[8,58,59] Novobiocin, a specific inhibitor of bacterial GyrB, has been shown to inhibit the ATPase activity of PfGyrB and caused parasite death in culture.[57,60] Thus, PfGyrA and PfGyrB have become attractive targets for use of existing antimicrobials and for development of novel anti-malarial therapeutics.

PfPOM1, known as *P. falciparum* plastidic DNA replication/repair enzyme complex (PfPREX), contains multiple domains with DNA primase, DNA helicase, DNA polymerase and 3′–5′ exonuclease activities. The N-terminal sequence has been demonstrated to direct PfPOM1 to the apicoplast, suggesting that PfPOM1 may have an important role in apicoplast DNA replication.[61] As the polymerase and primase subunits of PfPOM1 have no direct orthologues in human host cells, PfPOM1 presents a perfect target for new drug development.

### 13.4.2 Targeting Unique Protein Sequences

The sequences of malarial replication proteins exhibit high homology to their human counterparts, but also contain a number of unique features such as large

N- or C-terminal extensions, numerous sequence insertions within highly conserved regions, and various amino acid sequence repeats. These unique properties of the replication proteins offer potentially ideal targets for new drug therapy. In particular, the extra regions in PfPCNA, and the 12 repeats in the R region of PfFen1 pose very exciting potential targets, since high resolution crystallographic data exist on human Fen1–PCNA complexes[49] (see also Figure 3.8), which can be used to model similarities and differences in the protein folds and interactions. Moreover, the genes encoding these malarial proteins have been cloned and recombinant protein expressed and assayed *in vitro*.[46–48,51] Relatively simple *in vitro* assays for Fen1 nuclease activity[51] and PCNA binding already exist,[48,50] both of which should be adaptable to high throughout methodologies, to allow rapid screening of compounds that specifically and selectively inhibit malarial but not human Fen1. Given more investment in the field of malarial research, the identification and characterisation of *Plasmodium* replication components should provide a veritable gold mine of 'druggable' targets.

## 13.5 Concluding Remarks

With the completion of the *P. falciparum* genome sequencing project,[62,63] our knowledge of the DNA replication machinery in the human malarial parasite has developed rapidly in the past decade. The parasite contains most DNA replication components conserved in eukaryotes. Nearly every replication protein described in *P. falciparum* has a close orthologue in other malarial species such as *P. vivax, P. knowlesi, P. yoelii, P. berghei,* and *P. chabaudi*.[64–66] However, some differences between *P. falciparum* and other eukaryotes (particularly humans) have been exploited, particularly at the replication initiation phase. Based on protein sequence comparisons, many proteins that are absolutely essential for eukaryotic DNA replication cannot be found in the parasite genome. In addition, a large proportion of replication proteins of the parasite display unique and interesting features such as terminal sequence extensions, amino acid sequence insertions and sequence repeats. These attractive characteristics present ideal potential targets for development of novel anti-malarial drugs.

Unfortunately, the majority of DNA replication components of the parasite are only at this stage 'predicted proteins', being revealed from the parasite genome using data-mining bioinformatics approaches. It has been shown that approximately 24% of the parasite genes in current databases were predicted incorrectly, including one or more additional introns found in some so-called 'intron-free' genes, different sizes and locations of introns in numerous predicted 'intron-containing' genes, and alternative splicings for many genes available in current databases.[67] Therefore it is essential to verify amino acid sequences and real identities of the replication proteins experimentally in the near future.

The challenges now are to express replication proteins in heterologous systems and to purify sufficient amounts of the recombinant proteins for more detailed biochemical and three-dimensional structural studies. In addition, to reconstitute a functional malarial replication system *in vitro* will greatly facilitate our understanding of the replication mechanisms in detail, and permit large-scale screening of natural product and chemical compound libraries for development of new anti-malarial drugs.

## Acknowledgements

I thank Dr. Lynne S. Cox very much for her critical reading and comments on the chapter and her tremendous help in preparing Figures 13.2B and 13.4.

## References

1. R. W. Snow, C. A. Guerra, A. M. Noor, H. Y. Myint and S. I. Hay, The global distribution of clinical episodes of *Plasmodium falciparum* malaria, *Nature*, 2005, **434**, 214–217.
2. J. K. Baird, Effectiveness of antimalarial drugs, *N. Engl. J. Med.*, 2005, **352**, 1565–1577.
3. S. M. Todryk and A. V. Hill, Malaria vaccines: the stage we are at, *Nat. Rev. Microbiol.*, 2007, **5**, 487–489.
4. J. H. White and B. J. Kilbey, DNA replication in the malaria parasite, *Parasitol. Today*, 1996, **12**, 151–155.
5. C. J. Janse, P. F. van der Klooster, H. J. van der Kaay, M. van der Ploeg and J. P. Overdulve, DNA synthesis in *Plasmodium berghei* during asexual and sexual development, *Mol. Biochem. Parasitol.*, 1986, **20**, 173–182.
6. D. Singh, S. Chaubey and S. Habib, Replication of the *Plasmodium falciparum* apicoplast DNA initiates within the inverted repeat region, *Mol. Biochem. Parasitol.*, 2003, **126**, 9–14.
7. D. Singh, A. Kumar, E. V. Raghu Ram and S. Habib, Multiple replication origins within the inverted repeat region of the *Plasmodium falciparum* apicoplast genome are differentially activated, *Mol. Biochem. Parasitol.*, 2005, **139**, 99–106.
8. D. H. Williamson, P. R. Preiser, P. W. Moore, S. McCready, M. Strath and R. J. Wilson, The plastid DNA of the malaria parasite *Plasmodium falciparum* is replicated by two mechanisms, *Mol. Microbiol.*, 2002, **45**, 533–542.
9. A. Gupta, P. Mehra and S. K. Dhar, Plasmodium falciparum origin recognition complex subunit 5: functional characterization and role in DNA replication foci formation, *Mol. Microbiol.*, 2008, **69**, 646–665.
10. P. Mehra, A. K. Biswas, A. Gupta, S. Gourinath, C. E. Chitnis and S. K. Dhar, Expression and characterization of human malaria parasite *Plasmodium falciparum* origin recognition complex subunit 1, *Biochem. Biophys. Res. Commun.*, 2005, **337**, 955–966.

11. S. Patterson, C. Robert, C. Whittle, R. Chakrabarti, C. Doerig and D. Chakrabarti, Pre-replication complex organization in the atypical DNA replication cycle of *Plasmodium falciparum*: characterization of the mini-chromosome maintenance (MCM) complex formation, *Mol. Biochem. Parasitol.*, 2006, **145**, 50–59.
12. J. L. Li and L. S. Cox, Characterisation of a sexual stage-specific gene encoding ORC1 homologue in the human malaria parasite *Plasmodium falciparum*, *Parasitol. Int.*, 2003, **52**, 41–52.
13. A. Gupta, P. Mehra, R. Nitharwal, A. Sharma, A. K. Biswas and S. K. Dhar, Analogous expression pattern of *Plasmodium falciparum* replication initiation proteins PfMCM4 and PfORC1 during the asexual and sexual stages of intraerythrocytic developmental cycle, *FEMS Microbiol. Lett.*, 2006, **261**, 12–18.
14. K. G. Le Roch, Y. Zhou, P. L. Blair, M. Grainger, J. K. Moch, J. D. Haynes, P. De La Vega, A. A. Holder, S. Batalov, D. J. Carucci and E. A. Winzeler, Discovery of gene function by expression profiling of the malaria parasite life cycle, *Science*, 2003, **301**, 1503–1508.
15. L. Mancio-Silva, A. P. Rojas-Meza, M. Vargas, A. Scherf and R. Hernandez-Rivas, Differential association of Orc1 and Sir2 proteins to telomeric domains in *Plasmodium falciparum*, *J. Cell Sci.*, 2008, **121**, 2046–2053.
16. M. T. Duraisingh, T. S. Voss, A. J. Marty, M. F. Duffy, R. T. Good, J. K. Thompson, L. H. Freitas-Junior, A. Scherf, B. S. Crabb and A. F. Cowman, Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*, *Cell*, 2005, **121**, 13–24.
17. L. H. Freitas-Junior, R. Hernandez-Rivas, S. A. Ralph, D. Montiel-Condado, O. K. Ruvalcaba-Salazar, A. P. Rojas-Meza, L. Mancio-Silva, R. J. Leal-Silvestre, A. M. Gontijo, S. Shorte and A. Scherf, Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites, *Cell*, 2005, **121**, 25–36.
18. J. A. Baur, Y. Zou, J. W. Shay and W. E. Wright, Telomere position effect in human cells, *Science*, 2001, **292**, 2075–2077.
19. Z. Bozdech, M. Llinas, B. L. Pulliam, E. D. Wong, J. Zhu and J. L. DeRisi, The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*, *PLoS Biol.*, 2003, **1**, E5.
20. J. Liu, C. L. Smith, D. DeRyckere, K. DeAngelis, G. S. Martin and J. M. Berger, Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control, *Mol. Cell*, 2000, **6**, 637–648.
21. J. L. Li and L. S. Cox, Identification of an MCM4 homologue expressed specifically in the sexual stage of *Plasmodium falciparum*, *Int. J. Parasitol.*, 2001, **31**, 1246–1252.
22. A. Gambus, R. C. Jones, A. Sanchez-Diaz, M. Kanemaki, F. van Deursen, R. D. Edmondson and K. Labib, GINS maintains association of Cdc45

with MCM in replisome progression complexes at eukaryotic DNA replication forks, *Nat. Cell Biol.*, 2006, **8**, 358–366.
23. S. E. Moyer, P. W. Lewis and M. R. Botchan, Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase, *Proc Natl Acad Sci U S A*, 2006, **103**, 10236–10241.
24. M. Pacek, A. V. Tutter, Y. Kubota, H. Takisawa and J. C. Walter, Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication, *Mol. Cell*, 2006, **21**, 581–587.
25. M. L. Hoang, R. P. Leon, L. Pessoa-Brandao, S. Hunt, M. K. Raghuraman, W. L. Fangman, B. J. Brewer and R. A. Sclafani, Structural changes in Mcm5 protein bypass Cdc7-Dbf4 function and reduce replication origin efficiency in *Saccharomyces cerevisiae*, *Mol. Cell. Biol.*, 2007, **27**, 7594–7602.
26. M. Lutzmann, D. Maiorano and M. Mechali, Identification of full genes and proteins of MCM9, a novel, vertebrate-specific member of the MCM2-8 protein family, *Gene*, 2005, **362**, 51–56.
27. E. Fanning, V. Klimovich and A. R. Nager, A dynamic model for replication protein A (RPA) function in DNA processing pathways, *Nucleic Acids Res.*, 2006, **34**, 4126–4137.
28. Y. Zou, Y. Liu, X. Wu and S. M. Shell, Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses, *J. Cell. Physiol.*, 2006, **208**, 267–273.
29. T. S. Voss, T. Mini, P. Jenoe and H. P. Beck, *Plasmodium falciparum* possesses a cell cycle-regulated short type replication protein A large subunit encoded by an unusual transcript, *J. Biol. Chem.*, 2002, **277**, 17493–17501.
30. E. Bochkareva, S. Korolev, S. P. Lees-Miller and A. Bochkarev, Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA, *EMBO J.*, 2002, **21**, 1855–1863.
31. L. Abu-Elheiga, D. T. Spira and U. Bachrach, *Plasmodium falciparum*: properties of an alpha-like DNA polymerase, the key enzyme in DNA synthesis, *Exp. Parasitol.*, 1990, **71**, 21–26.
32. P. Chavalitshewinkoon, E. de Vries, J. G. Stam, F. F. Franssen, P. C. van der Vliet and J. P. Overdulve, Purification and characterization of DNA polymerases from *Plasmodium falciparum*, *Mol. Biochem. Parasitol.*, 1993, **61**, 243–253.
33. I. Choi and R. B. Mikkelsen, Cell cycle-dependent biosynthesis of *Plasmodium falciparum* DNA polymerase-alpha, *Exp. Parasitol.*, 1991, **73**, 93–100.
34. E. de Vries, J. G. Stam, F. F. Franssen, P. C. van der Vliet and J. P. Overdulve, Purification and characterization of DNA polymerases from Plasmodium berghei, *Mol. Biochem. Parasitol.*, 1991, **45**, 223–232.
35. D. N. Frick and C. C. Richardson, DNA primases, *Annu. Rev. Biochem.*, 2001, **70**, 39–80.
36. S. H. Lao-Sirieix, L. Pellegrini and S. D. Bell, The promiscuous primase, *Trends Genet.*, 2005, **21**, 568–572.

37. J. H. White, B. J. Kilbey, E. de Vries, M. Goman, P. Alano, S. Cheesman, S. McAleese and R. G. Ridley, The gene encoding DNA polymerase alpha from *Plasmodium falciparum*, *Nucleic Acids Res.*, 1993, **21**, 3643–3646.
38. B. E. Weiner, H. Huang, B. M. Dattilo, M. J. Nilges, E. Fanning and W. J. Chazin, An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase, *J. Biol. Chem.*, 2007, **282**, 33444–33451.
39. S. Prasartkaew, N. M. Zijlstra, P. Wilairat, J. P. Overdulve and E. de Vries, Molecular cloning of a *Plasmodium falciparum* gene interrupted by 15 introns encoding a functional primase 53 kDa subunit as demonstrated by expression in a baculovirus system, *Nucleic Acids Res.*, 1996, **24**, 3934–3941.
40. A. Johnson and M. O'Donnell, Cellular DNA replicases: components and dynamics at the replication fork, *Annu. Rev. Biochem.*, 2005, **74**, 283–315.
41. B. A. Fox and D. J. Bzik, The primary structure of *Plasmodium falciparum* DNA polymerase delta is similar to drug sensitive delta-like viral DNA polymerases, *Mol. Biochem. Parasitol.*, 1991, **49**, 289–296.
42. R. G. Ridley, J. H. White, S. M. McAleese, M. Goman, P. Alano, E. de Vries and B. J. Kilbey, DNA polymerase delta: gene sequences from *Plasmodium falciparum* indicate that this enzyme is more highly conserved than DNA polymerase alpha, *Nucleic Acids Res.*, 1991, **19**, 6731–6736.
43. P. Horrocks, M. Jackson, S. Cheesman, J. H. White and B. J. Kilbey, Stage specific expression of proliferating cell nuclear antigen and DNA polymerase delta from *Plasmodium falciparum*, *Mol. Biochem. Parasitol.*, 1996, **79**, 177–182.
44. T. S. Krishna, X. P. Kong, S. Gary, P. M. Burgers and J. Kuriyan, Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA, *Cell*, 1994, **79**, 1233–1243.
45. J. M. Gulbis, Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan, Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA, *Cell*, 1996, **87**, 297–306.
46. B. J. Kilbey, I. Fraser, S. McAleese, M. Goman and R. G. Ridley, Molecular characterisation and stage-specific expression of proliferating cell nuclear antigen (PCNA) from the malarial parasite, *Plasmodium falciparum*, *Nucleic Acids Res.*, 1993, **21**, 239–243.
47. J. L. Li, A. V. Warren and L. S. Cox, Identification of a second proliferating cell nuclear antigen in the human malarial pathogen *Plasmodium falciparum*, *Int. J. Parasitol.*, 2002, **32**, 1683–1692.
48. S. Patterson, C. Whittle, C. Robert and D. Chakrabarti, Molecular characterization and expression of an alternate proliferating cell nuclear antigen homologue, PfPCNA2, in *Plasmodium falciparum*, *Biochem. Biophys. Res. Commun.*, 2002, **298**, 371–376.
49. S. Sakurai, K. Kitano, H. Yamaguchi, K. Hamada, K. Okada, K. Fukuda, M. Uchida, E. Ohtsuka, H. Morioka and T. Hakoshima, Structural basis for recruitment of human flap endonuclease 1 to PCNA, *EMBO J.*, 2005, **24**, 683–693.
50. E. Warbrick, D. P. Lane, D. M. Glover and L. S. Cox, Homologous regions of Fen1 and p21Cip1 compete for binding to the same site on

PCNA: a potential mechanism to co-ordinate DNA replication and repair, *Oncogene*, 1997, **14**, 2313–2321.
51. L. J. Casta, J. S. Buguliskis, Y. Matsumoto and T. F. Taraschi, Expression and biochemical characterization of the *Plasmodium falciparum* DNA repair enzyme, flap endonuclease-1 (PfFEN-1), *Mol. Biochem. Parasitol.*, 2008, **157**, 1–12.
52. J. S. Buguliskis, L. J. Casta, C. E. Butz, Y. Matsumoto and T. F. Taraschi, Expression and biochemical characterization of *Plasmodium falciparum* DNA ligase I, *Mol. Biochem. Parasitol.*, 2007, **155**, 128–137.
53. L. Lai, H. Yokota, L. W. Hung, R. Kim and S. H. Kim, Crystal structure of archaeal RNase HII: a homologue of human major RNase H, *Structure*, 2000, **8**, 897–904.
54. M. E. Fichera and D. S. Roos, A plastid organelle as a drug target in apicomplexan parasites, *Nature*, 1997, **390**, 407–409.
55. V. Weissig, T. S. Vetro-Widenhouse and T. C. Rowe, Topoisomerase II inhibitors induce cleavage of nuclear and 35-kb plastid DNAs in the malarial parasite *Plasmodium falciparum*, *DNA Cell Biol.*, 1997, **16**, 1483–1492.
56. M. A. Dar, A. Sharma, N. Mondal and S. K. Dhar, Molecular cloning of apicoplast-targeted *Plasmodium falciparum* DNA gyrase genes: unique intrinsic ATPase activity and ATP-independent dimerization of PfGyrB subunit, *Eukaryot. Cell*, 2007, **6**, 398–412.
57. E. V. Raghu Ram, A. Kumar, S. Biswas, A. Kumar, S. Chaubey, M. I. Siddiqi and S. Habib, Nuclear gyrB encodes a functional subunit of the *Plasmodium falciparum* gyrase that is involved in apicoplast DNA replication, *Mol. Biochem. Parasitol.*, 2007, **154**, 30–39.
58. E. L. Dahl and P. J. Rosenthal, Multiple antibiotics exert delayed effects against the *Plasmodium falciparum* apicoplast, *Antimicrob. Agents Chemother.*, 2007, **51**, 3485–3490.
59. C. D. Goodman, V. Su and G. I. McFadden, The effects of anti-bacterials on the malaria parasite *Plasmodium falciparum*, *Mol. Biochem. Parasitol.*, 2007, **152**, 181–191.
60. A. A. Divo, A. C. Sartorelli, C. L. Patton and F. J. Bia, Activity of fluoroquinolone antibiotics against *Plasmodium falciparum in vitro*, *Antimicrob. Agents Chemother.*, 1988, **32**, 1182–1186.
61. F. Seow, S. Sato, C. S. Janssen, M. O. Riehle, A. Mukhopadhyay, R. S. Phillips, R. J. Wilson and M. P. Barrett, The plastidic DNA replication enzyme complex of *Plasmodium falciparum*, *Mol. Biochem. Parasitol.*, 2005, **141**, 145–153.
62. M. J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M. S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C.

Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser and B. Barrell, Genome sequence of the human malaria parasite *Plasmodium falciparum*, *Nature*, 2002, **419**, 498–511.

63. M. J. Gardner, H. Tettelin, D. J. Carucci, L. M. Cummings, L. Aravind, E. V. Koonin, S. Shallom, T. Mason, K. Yu, C. Fujii, J. Pederson, K. Shen, J. Jing, C. Aston, Z. Lai, D. C. Schwartz, M. Pertea, S. Salzberg, L. Zhou, G. G. Sutton, R. Clayton, O. White, H. O. Smith, C. M. Fraser, M. D. Adams, J. C. Venter and S. L. Hoffman, Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*, *Science*, 1998, **282**, 1126–1132.
64. J. M. Carlton, J. H. Adams, J. C. Silva, S. L. Bidwell, H. Lorenzi, E. Caler, J. Crabtree, S. V. Angiuoli, E. F. Merino, P. Amedeo, Q. Cheng, R. M. Coulson, B. S. Crabb, H. A. Del Portillo, K. Essien, T. V. Feldblyum, C. Fernandez-Becerra, P. R. Gilson, A. H. Gueye, X. Guo, S. Kang'a, T. W. Kooij, M. Korsinczky, E. V. Meyer, V. Nene, I. Paulsen, O. White, S. A. Ralph, Q. Ren, T. J. Sargeant, S. L. Salzberg, C. J. Stoeckert, S. A. Sullivan, M. M. Yamamoto, S. L. Hoffman, J. R. Wortman, M. J. Gardner, M. R. Galinski, J. W. Barnwell and C. M. Fraser-Liggett, Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*, *Nature*, 2008, **455**, 757–763.
65. J. M. Carlton, S. V. Angiuoli, B. B. Suh, T. W. Kooij, M. Pertea, J. C. Silva, M. D. Ermolaeva, J. E. Allen, J. D. Selengut, H. L. Koo, J. D. Peterson, M. Pop, D. S. Kosack, M. F. Shumway, S. L. Bidwell, S. J. Shallom, S. E. van Aken, S. B. Riedmuller, T. V. Feldblyum, J. K. Cho, J. Quackenbush, M. Sedegah, A. Shoaibi, L. M. Cummings, L. Florens, J. R. Yates, J. D. Raine, R. E. Sinden, M. A. Harris, D. A. Cunningham, P. R. Preiser, L. W. Bergman, A. B. Vaidya, L. H. van Lin, C. J. Janse, A. P. Waters, H. O. Smith, O. R. White, S. L. Salzberg, J. C. Venter, C. M. Fraser, S. L. Hoffman, M. J. Gardner and D. J. Carucci, Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii* yoelii, *Nature*, 2002, **419**, 512–519.
66. A. Pain, U. Bohme, A. E. Berry, K. Mungall, R. D. Finn, A. P. Jackson, T. Mourier, J. Mistry, E. M. Pasini, M. A. Aslett, S. Balasubrammaniam, K. Borgwardt, K. Brooks, C. Carret, T. J. Carver, I. Cherevach, T. Chillingworth, T. G. Clark, M. R. Galinski, N. Hall, D. Harper, D. Harris, H. Hauser, A. Ivens, C. S. Janssen, T. Keane, N. Larke, S. Lapp, M. Marti, S. Moule, I. M. Meyer, D. Ormond, N. Peters, M. Sanders, S. Sanders, T. J. Sargeant, M. Simmonds, F. Smith, R. Squares, S. Thurston, A. R. Tivey, D. Walker, B. White, E. Zuiderwijk, C. Churcher, M. A. Quail, A. F. Cowman, C. M. Turner, M. A. Rajandream, C. H. Kocken, A. W. Thomas, C. I. Newbold, B. G. Barrell and M. Berriman, The genome of the simian and human malaria parasite *Plasmodium knowlesi*, *Nature*, 2008, **455**, 799–803.
67. F. Lu, H. Jiang, J. Ding, J. Mu, J. G. Valenzuela, J. M. Ribeiro and X. Z. Su, cDNA sequences reveal considerable gene prediction inaccuracy in the *Plasmodium falciparum* genome, *BMC Genomics*, 2007, **8**, 255.

68. N. Hall, A. Pain, M. Berriman, C. Churcher, B. Harris, D. Harris, K. Mungall, S. Bowman, R. Atkin, S. Baker, A. Barron, K. Brooks, C. O. Buckee, C. Burrows, I. Cherevach, C. Chillingworth, T. Chillingworth, Z. Christodoulou, L. Clark, R. Clark, C. Corton, A. Cronin, R. Davies, P. Davis, P. Dear, F. Dearden, J. Doggett, T. Feltwell, A. Goble, I. Goodhead, R. Gwilliam, N. Hamlin, Z. Hance, D. Harper, H. Hauser, T. Hornsby, S. Holroyd, P. Horrocks, S. Humphray, K. Jagels, K. D. James, D. Johnson, A. Kerhornou, A. Knights, B. Konfortov, S. Kyes, N. Larke, D. Lawson, N. Lennard, A. Line, M. Maddison, J. McLean, P. Mooney, S. Moule, L. Murphy, K. Oliver, D. Ormond, C. Price, M. A. Quail, E. Rabbinowitsch, M. A. Rajandream, S. Rutter, K. M. Rutherford, M. Sanders, M. Simmonds, K. Seeger, S. Sharp, R. Smith, R. Squares, S. Squares, K. Stevens, K. Taylor, A. Tivey, L. Unwin, S. Whitehead, J. Woodward, J. E. Sulston, A. Craig, C. Newbold and B. G. Barrell, Sequence of *Plasmodium falciparum* chromosomes 1, 3–9 and 13, *Nature*, 2002, **419**, 527–531.
69. S. Bowman, D. Lawson, D. Basham, D. Brown, T. Chillingworth, C. M. Churcher, A. Craig, R. M. Davies, K. Devlin, T. Feltwell, S. Gentles, R. Gwilliam, N. Hamlin, D. Harris, S. Holroyd, T. Hornsby, P. Horrocks, K. Jagels, B. Jassal, S. Kyes, J. McLean, S. Moule, K. Mungall, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, S. Rutter, J. Skelton, R. Squares, S. Squares, J. E. Sulston, S. Whitehead, J. R. Woodward, C. Newbold and B. G. Barrell, The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*, *Nature*, 1999, **400**, 532–538.
70. Z. Bozdech, J. Zhu, M. P. Joachimiak, F. E. Cohen, B. Pulliam and J. L. DeRisi, Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray, *Genome Biol.*, 2003, **4**, R9.
71. R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. Gibson, D. Higgins and J. Thompson, Multiple sequence alignment with the Clustal series of programs, *Nucleic Acids Res.*, 2003, **31**, 3497–34500.
72. M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins, Clustal W and Clustal X version 2.0, *Bioinformatics*, 2007, **23**, 2947–2948.
73. K. Arnold, L. Bordoli, J. Kopp and T. Schwede, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics*, 2006, **22**, 195–201.
74. F. Kiefer, K. Arnold, M. Kunzli, L. Bordoli and T. Schwede, The SWISS-MODEL repository and associated resources, *Nucleic Acids Res.*, 2009, **37**, D387–D392.