3. RESULTS

3.1 Identification and biochemical characterization of apicoplast replication ori-binding proteins

3.1.1. Isolation of proteins from parasite apicoplast-enriched fraction
3.1.2. Specific DNA-protein complexes are detected at a plDNA ori site
3.1.3. DNA-protein complexes at ori sequence exhibit stage-dependent binding
3.1.4. Competition with DNA fragments covering the entire replication initiation zone in apicoplast genome reveals an additional binding site
3.1.5. Non-ionic interactions play a major role in complex formation and divalent cations are not essential for DNA-binding
3.1.6. The protein(s) interact through the minor groove of the DNA double helix
3.1.7. A 13 bp sequence on the 195 bp ori site is protected in DNase I footprinting
3.1.8. The two DNA-protein complexes exhibit differential affinity for DNA
3.1.9. DNA-protein complex II (CII) has higher temperature-stability than complex I (CI)
3.1.10. ~68-75 kDa proteins are detected by Southwestern blotting of parasite lysate
3.1.11. PfDnaJ is identified as a component of the ori DNA-protein complex
3.1.12. Sequence analysis of PfDnaJ

3.2 Expression and characterization of identified PfDnaJ protein

3.2.1. DnaJ transcript is detected in *P.falciparum*
3.2.2. Amplification, cloning and overexpression of nuclear-encoded gene *PfDnaJ*
3.2.3. Antibodies raised against PfDnaJ recognize nuclear encoded DnaJ in *P. falciparum* lysate

3.2.4. Recombinant PfDnaJ exhibits DNA-binding activity

3.2.5. Antibody-mediated inhibition of binding confirms identity of the protein in complexes obtained with *P. falciparum* lysates

3.2.6. Recombinant PfDnaJ binds non-specifically to DNA

3.2.7. Recombinant PfDnaJ forms dimers in solution

3.2.8. Organellar localization of PfDnaJ analysed by immunofluorescence microscopy

3.3 Transcriptional analysis of *P. falciparum* apicoplast genome

3.3.1. Transcripts of apicoplast genes ORF470/sufB, *clpC* and *tufA* are detected by RT-PCR

3.3.2. Northern blotting of *LSU* and ORF470/sufB transcripts reveals presence a polycistronic transcript

3.3.3. Confirmation of polycistronic nature of transcript

3.3.4. 5'-end mapping of *LSU* transcript using RACE reveals presence of multiple initiation sites
3.1 Identification and biochemical characterization of apicoplast replication ori-binding proteins

In order to determine protein factors that bind to apicoplast DNA replication ori sequences, a panel of plDNA segments from the IR region that exhibited high ori activity was selected (Fig. 1.7). These ori elements have been identified previously in our laboratory (Singh et al., 2003; Singh et al., 2005). The DNA fragments were used in binding reactions with parasite proteins in electrophoretic mobility shift assays (EMSAs).

3.1.1. Isolation of proteins from parasite apicoplast-enriched fraction

To investigate the DNA-protein interaction(s) at apicoplast replication ori, we then standardized a protocol for isolation of proteins from apicoplast-enriched fractions of *P. falciparum* infected erythrocyte culture by differential centrifugation as shown in Fig. 2.2. Presence of the apicoplast proteins in nuclear, cytoplasmic and organellar-enriched protein fractions was confirmed by immuno-precipitation followed by western blotting using anti EF-Tu antibody that specifically recognizes apicoplast-encoded EF-Tu.

Parasites released from the infected erythrocytes by 0.05% saponin lysis were disrupted either by NP-40 lysis buffer or by mechanical disruption (homogenization). In NP-40 lysis, parasites were suspended in a buffer containing 0.14% NP-40. In homogenization, released parasites were suspended in above buffer except NP-40 and homogenized in a homogenizer (15 times with 10 seconds strokes). Disrupted parasites by both methods were centrifuged at 2,000 rpm to pellet the nuclei and processed as described in the material and methods section. The protocol is summarized in Fig. 2.2, A.

Presence of parasite proteins extracted from different fractions by NP-40 lysis and homogenization was carried out by immunoprecipitation followed by western blotting using anti EF-Tu antibody. A band of EF-Tu
Results

protein at ~46 kDa was detected in cytoplasmic and organelar enriched protein fraction isolated from homogenized parasites, whereas no band was detected in nuclear fraction of homogenized cells (Fig. 2.2, B). Hence, the homogenization method was used for preparation of protein extract for further experiments.

3.1.2. Specific DNA-protein complexes are detected at a plDNA ori site

Extracts from organelle-enriched fractions were used to detect DNA-protein complexes in EMSAs. DNA fragments from ori sequences were radiolabelled and used in EMSAs. Protein extract was incubated with labeled probe in presence of poly(dA-dT) as a non-specific competitor, electrophoresed on 4% native polyacrylamide gel and autoradiographed. Two clear shifts were visible in the experimental set in which a 195 bp ori region fragment was used as probe. A complex of high mobility (CI) and a complex of lower mobility (CII) of similar intensity were consistently detected (Fig. 3.1, lane 2). To determine whether CI and CII represented specific DNA-protein interactions, competition experiments were carried out. An excess (~200-fold molar excess) of unlabeled homologous competitor (195 bp fragment) competed for complex formation (Fig. 3.1, lane 3 and 4) while CI and CII remained unaffected in the presence of the similar amount of non-specific competitor pUC18 (Fig. 3.1, lane 5 and 6) demonstrating the specificity of the two complexes. Complexes CI and CII obtained with the 195 bp ori sequence were subjected to further analysis.

Detection of DNA-protein complex formation using protein extracts prepared from the nuclear, cytoplasmic and organelle-enriched components of P.falciparum was carried out by incubating the extracts with the 195 bp probe. While no shift was observed with nuclear proteins, a shift of much lower mobility than CI and CII was obtained with the cytoplasmic fraction indicating that protein components of CI and CII are specific to the organellar fraction (Fig. 3.2, A).
**Results**

**Fig. 3.1.** DNA-protein complexes at a pIDNA ori site. Two complexes (CI and CII) were detected in EMSAs using organellar extract (lane 2). Competition with unlabeled 195 bp DNA (lanes 3 and 4) and non-specific competitor pUC18 (lanes 5 and 6). Lane 2 is normal binding. FP, free probe.

**Fig. 3.2.** (A) Complexes obtained with nuclear (N), cytoplasmic (C) and organellar protein extracts (O). CI and CII are indicated by arrows. (B) Complex formation with parasite lysates from different erythrocytic stages. R, early-middle rings; T, early-middle trophozoites; LT, late trophozoites; T-S, late trophozoites-early schizonts; S, middle-late schizonts.
3.1.3. DNA-protein complexes at ori sequence exhibit stage-dependent binding

We investigated complex formation using equal quantities of organellar extracts prepared from synchronised cells at the early/middle ring, early/middle trophozoite, late trophozoite, late trophozoite-early schizont and middle/late schizont stages (Fig. 3.2, B). While both CI and CII were detected with extracts from all stages, the intensity of the shifts was stage-dependent. While no clear difference in intensity was visible for the lower complex (CI), the upper complex (CII) was maximally detected in the late trophozoite-early schizont stage with much lower intensities in the ring and schizont stages. The higher intensity of CII observed at the stage of the parasitic intra-erythrocytic cycle at which plDNA replication takes place is indicative of a possible role of its constituent protein(s) in the replication process.

3.1.4. Competition with DNA fragments covering the entire replication initiation zone in apicoplast genome reveals an additional binding site

The IR region of plDNA serves as a zone of initiation with the utilization of multiple ori elements with differential efficiencies (Fig. 1.7). To investigate whether the protein(s) that complexes with the 195 bp ori region of the IR also bound at ori sites within the IR, restriction fragments covering the 1820 bp and 1988 bp IR segments (Fig. 3.3, C) were used as competitors. While none of the other fragments of the 1988 bp segment could compete with the 195 bp probe for complex formation (Fig. 3.3, B), a 200-fold excess of the 113 bp fragment of the 1820 bp segment successfully competed out both CI and CII (Fig. 3.3, A) indicating that the organellar protein(s) recognized at least one more binding site within the IR.

3.1.5. Non-ionic interactions play a major role in complex formation and divalent cations are not essential for DNA-binding
**Results**

**Fig. 3.3.** Competition with IR fragments of plDNA. (A) and (B) show competition with fragments of the 1820 bp and 1988 bp IR regions. Lane 2 in (A) and lane 1 in (B) depicts normal binding. FP denotes free probe. (C) Position of restriction fragments of the plDNA IR used as competitors in EMSAs.

**Fig. 3.4.** Binding characteristics of the protein-ori DNA complexes. Complex formation in presence of increasing concentrations of NaCl (A); and EDTA (B) indicates that binding is essentially non-ionic and does not require divalent cations.
Experiments were carried out to determine the interactions involved in complex formation at the 195 bp ori sequence. When the binding reactions were carried out in the presence of increasing concentrations of NaCl, the intensity of CI and CII decreased gradually as a function of salt concentration but was not completely abolished even at 2 M NaCl (Fig. 3.4, A). This suggested that ionic interactions might be only marginally involved in complex formation.

When binding was carried out in the presence of increasing concentrations of EDTA (Fig. 3.4, B), the complex remained unaffected till 50 mM EDTA with a decrease in intensity only at 100 mM EDTA, demonstrating that divalent cations do not play a major role in complex formation.

3.1.6. The protein(s) interact through the minor groove of the DNA double helix

To determine whether the DNA-binding protein(s) interact through the major or the minor groove of the DNA molecule, binding reactions were carried out in the presence of actinomycin D, distamycin A or methyl green. While methyl green interacts specifically with the major groove of DNA, actinomycin D and distamycin A are minor groove-binding drugs (Copenhaver, et al., 1993; Kim, et al., 1993). Incubation of the 195 bp radiolabeled probe with the drugs alone did not affect mobility of the probe. The minor-groove binding drugs, actinomycin D and distamycin A, inhibited both complexes CI and CII in a dose-dependent manner in EMSAs (Fig. 3.5, A and B). However, the complexes remained intact upon pre-incubation of the DNA probe with methyl green even when 1 mM of the drug was used in the binding reaction (Fig. 3.5, C). These results demonstrate that the protein components of the complexes approach the DNA through the minor groove of the double helix.
**Fig. 3.5.** Inhibition of binding with increasing concentrations of minor-groove binding drugs actinomycin D (A), distamycin A (B) and major-groove binding compound methyl green (C). Lane 2 is the 195 bp probe with inhibitor alone, and lane 3 is normal binding in the absence of inhibitor. FP; Free probe.
3.1.7. A 13 bp sequence on the 195 bp ori site is protected in DNase I footprinting

Binding site of the protein(s) on the 195 bp DNA probe with the organellar lysate was analyzed by DNase I footprinting analysis (Fig. 3.6). Free radiolabeled DNA probe as well as the DNA-protein complex was incubated with DNase I for 2 minutes and separated on the 6% urea-polyacrylamide gel along with G and G+A sequencing ladder and subjected to autoradiography. DNase I cleaves the phosphodiester bond of the DNA; as a result the region bound to protein remains protected to enzymatic cleavage. The nucleotide sequence of the protected region was determined with the help of the G and G+A sequencing ladder.

3.1.8. The two DNA-protein complexes exhibit differential affinity for DNA

The affinity of the protein(s) for the ori DNA was determined by calculation of the $K_d$ value of the complexes CI and CII. The apparent $K_d$ value for specific DNA-protein interaction was estimated by a method based on EMSA (Liao, et al., 1992; Habib and Hasnain, 1996). A constant amount (3 μg) of P.falciparum organellar lysate was incubated with increasing concentrations of the radiolabeled 195 bp probe in separate binding reactions (Fig. 3.7, A). The extent of complex formation was quantitated by scintillation counting of excised labeled bands of CI, CII and the free probe as well as by densitometry analysis of bands. A plot bound/free versus the bound probe yielded a straight line with the inverse of the slope equal to the dissociation constant. Separate determinations gave a $K_d$ value of $1.4 \pm 0.98 \times 10^{-10}$ for the lower complex CI and a $K_d$ value of $4.4 \pm 1.5 \times 10^{-11}$ for the upper complex CII (Fig. 3.7, B). Although protein(s) of both complexes exhibited high affinity for DNA, the protein(s) of the upper complex had a ~3-fold higher affinity for the DNA probe.
Fig. 3.6. DNase I footprinting. The probe was incubated with increasing concentrations of organellar lysate (2 μg to 10 μg) in the binding buffer used for EMSAs (lanes 4-8). Lane 3 is DNase I-treated probe in the absence of protein. Nucleotide sequence of the protected region is shown.
3.1.9. DNA-protein Complex II (CII) has higher temperature-stability than complex I (CI)

The stability of the complexes at different temperatures was determined by incubating separate binding reactions at temperatures ranging from 4°C to 85°C for 10 min (Fig. 3.8). The lower complex (CI) remained stable till 37°C followed by a reduction in intensity at 45°C and complete disappearance of the complex at 65°C. In contrast, the intensity of the upper complex (CII) increased with temperature till 45°C where an intense CII band was visible. CII was also visible at 65°C and was completely abolished only at 85°C. These results indicate that the upper complex is not only more heat stable than the lower complex but is preferentially formed at higher temperatures.

3.1.10. ~68 to 75 kDa proteins are detected by Southwestern blotting of parasite lysate

Southwestern analysis of the organellar protein using the radiolabeled 195 bp fragment as a probe was carried out to determine the molecular mass of the protein complexing with the ori. For this, proteins from organellar- enriched fraction extracted at late trophozoite-early schizont stage were fractionated on a 10% SDS-polyacrylamide gel. The proteins were transferred onto nitrocellulose membrane, partially refolded with blocking buffer and hybridized with the labeled 195 bp fragment. Two bands of ~68 kDa and at ~75 kDa were visible in autoradiographs (Fig. 3.9, A). No such bands were observed when the membrane was hybridized with radiolabeled pUC18 fragment as control (Fig. 3.9, B). This suggested that the size of the protein that bound the 195 bp ori in EMSAs was in the range of ~68 to 75 kDa.

3.1.11. PfDnaJ is identified as a component of the ori DNA-protein complex

In order to identify proteins that comprise the two DNA-protein complexes observed in EMSAs with the 195 bp probe, the complexes CI and CII were electro-eluted from the EMSA gels. DNA and proteins of both
**Results**

**Fig. 3.7.** (A) EMSA gel for $K_d$ value determination. CII has greater affinity for DNA (B) $K_d$ for CI and CII calculated from binding reactions with 3 μg organellar lysate and increasing concentrations of probe. Values from two separate determinations are shown.

**Fig. 3.8.** Effect of temperature on complex formation. CII is more temperature stable than CI.

**Fig. 3.9.** Southwestern analysis of organellar lysate probed with the 195 bp fragment (A); pUC18 (B). ~68 - ~75 kDa protein bands are visible when probed with 195 bp probe.
complexes were separated by SDS-PAGE and visualized by silver staining. A prominent band of ~69 kDa was seen in lanes containing eluted products from both CI and CII (Fig. 3.10, A). An additional band of ~37 kDa was visible in the CI lane while two additional bands of ~37 kDa and ~42 kDa were seen in the CII lane. Autoradiograph of the gel indicating position of the labeled DNA probe ensured that the DNA migrated differently from the protein bands. All the identified protein bands were cut from the gel, digested with trypsin (In-gel digestion) and subjected to peptide mass fingerprinting by MALDI-TOF. The peptide mass fingerprint generated from the ~69 kDa band (Fig. 3.10, B) from both CI and CII was similar and identified PfDnaJ (gene ID: PFD0462w in PlasmoDB 5.2) as a protein component of both complexes with a strong probability value. The experiment was repeated three times and similar peptide mass fingerprints were generated each time. The fact that the same protein (PfDnaJ) was identified as a component of both CI and CII that are well separated in EMSAs suggested that it was not a non-DNA binding background protein that would be invisible in EMSAs but likely to show up in silver stains of the separated components of the eluted complexes. The peptide mass fingerprints generated from the other proteins did not give a hit with strong probability when searched against the P. falciparum database. PfDnaJ was thus identified as a candidate protein capable of interacting with the 195 bp ori probe. The candidature of PfDnaJ as a DNA-binding protein in the apicoplast was strengthened further by Southwestern analysis of P. falciparum organelar lysate using the 195 bp ori region as probe (Fig. 3.9) that also identified protein of ~68 kDa as capable of interacting with the 195 bp ori.

3.1.12. Sequence analysis of PfDnaJ

The annotated PfDnaJ (PFD0462w) in PlasmoDB 5.0 is encoded by Chr.4 of P. falciparum and has a GO assignment for the apicoplast. Proteins targeted to the apicoplast carry a bipartite targeting sequence comprising an N-terminal signal sequence followed by a transit peptide (Foth et al., 2003). Although prediction for a signal sequence using
**Fig. 3.10.** Identification of a protein bound at the 195 bp ori. (A) Silver stained SDS-PAGE of DNA-protein complexes eluted from EMSAs. (B) MALDI-TOF analysis of excised ~69 kDa band after digestion with trypsin.
Results

SignalP 3.0 for the presence of a signal sequence is not very strong (score Mean S=0.503), algorithm for prediction of a transit peptide sequence for apicoplast targeting (PATS) give a very high score for the presence of a transit peptide within the first 60 amino acids. The protein is also part of the list of nuclear genes with apicoplast signal sequences predicted to be involved in replication and transcrip­tion in the *P. falciparum* apicoplast (Malaria Parasite: Metabolic Pathways site http://huji.ac.il/malaria). Among the 35 *P. falciparum* proteins predicted to have homology with the J- domain of DnaJ, PFD0462w is the only one that contains the conserved J- domain together with the Zn-finger and C-terminal domains (Fig. 3.11 and 3.12). This PfDnaJ homolog carries a ~30 kDa C-terminal extension containing low complexity regions that does not exhibit significant homology to any known protein (Fig. 3.11).

3.2. Expression and characterization of identified PfDnaJ protein:

3.2.1. DnaJ transcript is detected in *P. falciparum*

The presence of the DnaJ transcript in the parasite was confirmed by RT-PCR using gene-specific primers for amplification. A band of 1076 bp was seen in the reaction containing reverse transcriptase, while no such band was amplified in the control reaction lacking reverse transcriptase thus confirming that DnaJ gene located at Chr4 is transcriptionally active (Fig. 3.13).

3.2.2. Amplification, cloning and over-expression of nuclear-encoded gene *PfDnaJ*

In order to confirm whether PfDnaJ identified by MALDI-TOF analysis actually exhibited DNA-binding activity, *P. falciparum* DnaJ encoded by Chr4 was PCR-amplified cloned and over-expressed in *E. coli*. PfDnaJ protein contains a N-terminal presequence (SignalP score=0.503) and a transit peptide sequence from amino acid 1-60 (i.e. from nucleotide 1 to 180). A region from amino acid 61-405 (nucleotide 181 to 1256) shows homology with DnaJ protein and the remaining ~30 kDa C-terminal extension (amino acid 406-672; nucleotide 1257 to 2019) contains low
Fig. 3.11. Sequence alignment of DnaJ protein from different organisms with DnaJ region (aa 1-405) of *P. falciparum* DnaJ protein (PFD0462w) using clustalW programme. Conserved regions present in DnaJ protein are underlined. (*Rickettsia typhi*, AAU03659.1); *Esherichia coli*, NC000913.2; *Arabidopsis thalina*, AY085227.1; *Trypanosoma cruzi*, XM_805422.1; *Theileria parva*, XM_759572.1).
**Fig. 3.12.** Schematic representation of conserved domains and motifs in PfDnaJ (PFD0462w). The first 60 amino acids comprise predicted apicoplast signal and transit sequences.

**Fig. 3.13.** PfDnaJ is transcriptionally active. Lane M: 1 Kb DNA ladder. The expected band of 1076 bp was seen in +RT reaction (lane 2). This band was absent in -RT control (lane 3).
complexity regions that do not exhibit significant homology to any known protein. We PCR-amplified cloned and expressed two regions of PfDnaJ as shown in Fig. 3.12.

a) Region covering nucleotide 181 to 1256 (amino acid 61-414) termed DnaJ_{int}.

b) Region covering nucleotide 181 to 2019 (amino acid 61-672) termed DnaJ_{m}.

Both DnaJ_{int} and DnaJ_{m} regions were PCR amplified by SacI-tagged upstream primer and HindIII-tagged downstream primer using *P. falciparum* genomic DNA as template.

The expected PCR products of 1076 bp for DnaJ_{int} and 1948 bp for DnaJ_{m} were obtained (Fig. 3.14, A and Fig. 3.15, A; respectively). Both PCR fragments were cloned into pQE-30 *E.coli* expression vector to give the constructs pQE30-DnaJ_{int} and pQE30-DnaJ_{m} (Fig. 3.14, C and Fig. 3.15, C; respectively).

Digestion of the constructs with SacI and HindIII enzymes release inserts of expected sizes (Fig. 3.14, B; lane 1 and Fig. 3.15, B; lane 2). Orientation of cloned DnaJ gene was confirmed by restriction enzyme analysis as well as by DNA sequencing. A schematic representation of the location of restriction sites used for characterization of sequence cloned in pQE-30 vector and the PCR-amplified fragment is depicted in (Fig. 3.14, C and Fig. 3.15, C).

Recombinant N-terminal His-tagged DnaJ_{int} and DnaJ_{m} proteins were expressed in *E.coli* (TG-1 strain) co-transformed with RIG plasmid (kind gift from Prof. W.G.J. Hol) by inducing cultures at 0.5 OD_{600} with 0.5 mM IPTG. Total cell protein was prepared and analyzed by SDS-PAGE. On induction with IPTG, both proteins (DnaJ_{int} and DnaJ_{m}) were expressed in the soluble form and were purified by Ni-NTA affinity chromatography (Fig. 3.16, A and Fig. 3.17, A). Western blotting of the different fractions of
Results

cell lysate (such as uninduced, induced, insoluble and soluble) was carried out with anti-His antibody (sigma). While an intact ~39 kDa band was obtained for PfDnaJ\textsubscript{int} (Fig. 3.16, B), purified PfDnaJ\textsubscript{m} had a minor proportion of the mature ~69 kDa product with additional degraded bands at ~50 kDa, ~39 kDa and ~28 kDa in coomassie stained gel (Fig. 3.17, A). These ~50 kDa and ~39 kDa degradation products, but not the ~28 kDa band, were recognized by anti-His antibody in western blot (Fig. 3.17, B and C) indicating that the ~50 kDa and ~39 kDa bands were N-terminal segments of PfDnaJ. 28 kDa and 39 kDa bands may be C-terminal and N-terminal degradation products of full length PfDnaJ, respectively.

3.2.3. Antibodies raised against PfDnaJ recognize nuclear encoded DnaJ in \textit{P. falciparum} lysate

Polyclonal antibodies against the recombinant PfDnaJ\textsubscript{int} protein were raised in rabbit as well as in mice to detect DnaJ protein in \textit{P. falciparum} lysates by western blot analysis. The recombinant PfDnaJ\textsubscript{int} protein used for antibody generation was purified by gel elution, concentrated and checked on SDS-polyacrylamide gel. High titre anti-PfDnaJ\textsubscript{int} serum were raised in rabbit as well as in mice after giving one booster dose in Freund’s complete adjuvant followed by one booster injection of purified recombinant PfDnaJ\textsubscript{int} in incomplete adjuvant. The titre of the raised antisera was analysed by ELISA. This immune serum was used to purify the antibody using membrane bound antigen as described in the methods section. The specificity of the raised antibodies was determined by western blot analysis using purified recombinant PfDnaJ\textsubscript{int} and PfDnaJ\textsubscript{m} protein as well as proteins isolated from uninfected red blood cells, \textit{E.coli} cells and \textit{P.falciparum} parasite cultures. Recombinant PfDnaJ\textsubscript{int} and PfDnaJ\textsubscript{m} gave bands of expected sizes (for PfDnaJ\textsubscript{int} at ~39 kDa and for PfDnaJ\textsubscript{m} at ~69 kDa, ~50 kDa, ~39 kDa) when probed with anti-PfDnaJ\textsubscript{int} sera raised in rabbit (Fig. 3.18, A; lane 1 and 2 respectively). This band pattern is similar when recombinant PfDnaJ\textsubscript{int} and PfDnaJ\textsubscript{m} were probed with anti-His antibody in western blotting. Lysates prepared from \textit{E.coli} cells
**Fig 3.14.** PCR amplification and characterization of dnaJ<sub>int</sub> fragment cloned in pQE30 vector. (A) PCR amplification of dnaJ<sub>int</sub> fragment. M; 1 kb DNA marker. (B) Digestion of pQE30-DnaJ<sub>int</sub> with *SacI* and *HindIII* releases 1076 bp fragment (lane 1); Lane 3 shows *XbaI* digested pQE30-DnaJ<sub>int</sub> with expected fragments of 2.6 kb and 1.9 kb. (C) Position of *XbaI* used for the characterization of DnaJ<sub>int</sub> cloned in pQE30 vector.
Fig 3.15. PCR amplification and characterization of dnaJm at cloned in pQE30 vector. (A) PCR amplification of DnaJm fragment. M; 1 kb DNA marker, (B) Digestion of pQE30-DnaJm with SacI and HindIII releases 1076 bp fragment (lane 1); Lane 3 shows EcoRI digested pQE30-DnaJm with expected fragments of 3.4 and 1.9 kb, (C) Position of EcoRI used for the characterization of DnaJm cloned in pQE30 vector.
Fig. 3.16. Expression and purification of His-tagged PfDnaJ_{int}.

(A) Coomassie-stained gel of purified PfDnaJ_{int}. M; protein molecular weight marker. (B) Western blot probed with anti-His antibody. Lysates prepared from *E. coli* TG-1 cells cotransformed with pQE30-DnaJ_{int} and RIG plasmid. Lane 1, uninduced sample; lane 2, total cell protein after induction with IPTG; lane 3, insoluble fraction; lane 4, soluble fraction; lane 5, protein purified through Ni-NTA affinity column.

Fig. 3.17. Expression and purification of His-tagged PfDnaJ_{m}.

(A) Coomassie-stained gel of purified PfDnaJ_{m}. M; protein molecular weight marker. (B) Western blot probed with anti-His antibody. Lysates prepared from TG-1 cells co-transformed with pQE30-DnaJ_{m} and RIG plasmid. Lane 1, uninduced sample; lane 2, insoluble fraction after induction with IPTG; lane 3, soluble fraction. (C) Western blot of DnaJ_{m} purified through Ni-NTA affinity column. Apart from the ~69 kDa band prominent degraded products of ~50 and ~39 kDa bands are visible.

containing pQE-30 vector alone did not gave any band when probed with anti-PfDnaJ_{int} serum (Fig. 3.18, A; lane 3).
The anti-PfDnaJ\textsubscript{int} purified Ab was used as a tool to detect presence of nuclear encoded DnaJ protein in \textit{P. falciparum}. Western blot analysis of total parasite lysate prepared from synchronized cells at different intra-erythrocytic stages was carried out. A prominent band of \textasciitilde 69 kDa representing processed mature PfDnaJ (expected size 68.6 kDa) (Fig. 3.18, B; panel i) was specifically detected in the late ring to schizont stages. A faint band at \textasciitilde 76 kDa that is likely to be the unprocessed form of PfDnaJ (expected size 75.9 kDa) was also observed. Additional bands of \textasciitilde 62 kDa, 54 kDa and 39 kDa that may represent degradation products of PfDnaJ were recognized at different stages. However, the possibility of these bands being other J-domain proteins recognized by anti-PfDnaJ\textsubscript{int} antibodies cannot be ruled out. No signal was obtained with pre-immune serum or from RBC lysate (Fig. 3.18, B; panel i, Fig. 3.18, A; lane 4, respectively) confirming that the detected protein bands were specific to the parasite. Although the combined signal from all bands was maximal at the middle trophozoite to schizont stages, the 69 kDa mature form of PfDnaJ was seen maximally at the early trophozoite stage when protein levels were normalized with a housekeeping protein tubulin (Fig. 3.18, B; panel ii and ii). These results demonstrate that the \textit{dnaJ} homolog on Chr4 of \textit{P. falciparum} is translated in the parasite with the intact 69 kDa product seen prominently from the early trophozoite stage onwards.

### 3.2.4. Recombinant PfDnaJ exhibits DNA-binding activity

To confirm whether the protein identified by us actually exhibited DNA-binding activity, purified recombinant PfDnaJ\textsubscript{m} and PfDnaJ\textsubscript{int} proteins were incubated with the 195 bp probe in binding reactions for EMSAs. Both proteins exhibited DNA-binding activity indicating that the DNA-binding domain lay within the 39 kDa PfDnaJ\textsubscript{int} region. However, a single band of lower mobility was detected with recombinant PfDnaJ\textsubscript{int} (Fig. 3.19, A; lane 1) while two bands of DNA-protein complexes were observed with
**Fig 3.18.** (A) Western blots of recombinant PfDnaJ_{int} (lane 1) and DnaJ_{m} (lane 2); *E.coli* lysate (lane 3) and RBC (lane 4) probed with anti-PfDnaJ_{int} antibodies. (B) Western blot of total parasite lysate with rabbit pre-immune serum (panel i; lane 1) and parasite proteins isolated from different intraerythrocytic stages of the parasite and probed with anti-PfDnaJ_{int} antibodies (panel i; lanes 2-5). The blot was stripped and reprobed with anti-tubulin antibody to normalize the signals (panel ii). (C) PfDnaJ_{m}/tubulin ratio at different erythrocytic stages. The protein is expressed from early trophozoite onwards with maximal expression at the early trophozoite stage.
recombinant PfDnaJm in EMSA (Fig. 3.19, B; lane 5 and 6). The position of shifts obtained with the *P.falciparum* organelle-enriched fraction and recombinant PfDnaJm was comparable (Fig. 3.19, B; lane 3 and 6).

3.2.5. Antibody-mediated inhibition of binding confirms identity of the protein in complexes obtained with *P.falciparum* lysates

In order to confirm the identity of the protein in *P.falciparum* organellar lysates, antibodies raised against PfDnaJint were used in supershift/inhibition assays in EMSAs. While the rabbit immune serum gave a supershift, anti-PfDnaJint antibodies purified against the recombinant protein inhibited DNA binding of PfDnaJint in a dose-dependent manner when incubated with the protein before addition of probe (Fig. 3.19, A; lane 2-4). Purified anti-PfDnaJint, added to the binding reaction after DNA probe-PfDnaJint complexation also inhibited DNA-binding. Both immune serum and purified anti-PfDnaJint antibodies also caused supershifting and inhibition of binding, respectively, with *P.falciparum* organellar lysate (Fig. 3.19, B; lane 2 and 4). When rabbit pre-immune serum was added to the binding reaction with *P.falciparum* organellar lysate before and after the addition of probe, no effect in DNA-protein complex formation was observed in EMSAs (Fig. 3.19, C). These results confirm that the protein from organellar lysates that binds the 195 bp probe is PfDnaJ.

3.2.6. Recombinant PfDnaJ binds non-specifically to DNA

In order to determine whether recombinant PfDnaJ specifically recognized replication ori, we carried out competition with unlabeled DNA fragments representing plDNA ori (113 bp fragment and the 195 bp self competitor), a non-ori sequence (from *clpC*) and a non-*Plasmodium* sequence (pUC18). The 195 bp DNA probe-PfDnaJint complex was competed out with 200-fold molar excess of all DNA fragments demonstrating the non-specific nature of DNA recognition by recombinant PfDnaJ (Fig. 3.19, D). This contrasts with our observation with *P.falciparum* organelle-enriched fractions where the two complexes could
**Fig. 3.19.** Properties of recombinant PfDnaJ. (A) Inhibition of PfDnaJ\textsubscript{int} binding with increasing concentrations of purified anti-PfDnaJ\textsubscript{int} antibody. (B) Pre-immune serum added to the binding before reaction (lane 3) or after addition of the organelar-enriched parasite lysate, does not cause super-shifting or inhibition of binding. (C) DNA-binding of PfDnaJ\textsubscript{m} (lanes 5 and 6) and antibody-mediated inhibition (lane 2), super shifting (lane 4) of complexes obtained with organelar lysate. Ab\textsubscript{p}, purified antibody; Ab\textsubscript{is}, immune serum; OL, organelar lysate. (D) Competition of PfDnaJ\textsubscript{int} binding with plDNA ori and non-ori fragments and pUC18. Lane 4 is normal binding in the absence of competitor.
not be competed out with several ori fragments (Fig. 3.3) leading us to speculate that as yet unidentified protein(s) bound at or near plDNA ori may help recruit DnaJ to specific ori sites.

3.2.7. Recombinant PfDnaJ forms dimers in solution

The oligomeric nature of the PfDnaJ was determined by chemical cross-linking as well as by size exclusion chromatography. In chemical cross-linking, 100 µg purified recombinant PfDnaJ<sub>int</sub> was treated with glutaraldehyde as described in methods section; treated and untreated samples were analyzed on 8% SDS-polyacrylamide gel. A band at ~68 kDa position was visualized in glutaraldehyde-treated sample indicating that PfDnaJ form dimers in solution (Fig. 3.20, A; lane 3).

The dimeric nature of PfDnaJ was again confirmed by the size exclusion chromatography using Superdex 200HR 10/300 column. Gel filtration of purified recombinant PfDnaJ<sub>int</sub> on a Superdex 200HR column, calibrated with the various molecular weight standards, showed two peaks at retention volumes of 14.4 ml and 15.5 ml (Fig. 3.20, B). From standard molecular weight calibration, it was deduced that the molecular weight corresponding to the elution volume 14.4 ml is ~65 kDa and for 15.5 ml is ~40 kDa. As the molecular weight of PfDnaJ is 39 kDa, the peak centered at 14.4 ml represents the dimer of the protein in solution. The possibility of additional multimeric forms can not be excluded as broad peaks were seen at lower elution volumes in gel filtration and some cross-linked protein were visible in the well in the glutaraldehyde cross-linked experiment (Fig. 3.20, A).

3.2.8. Organellar localization of PfDnaJ is analysed by immunofluorescence microscopy

The possibility of PfDnaJ being a mitochondrial protein that bound plDNA ori in EMSAs using organellar extracts was addressed by immunofluorescence confocal microscopy using PfDnaJ<sub>int</sub> antibody and
Fig. 3.20. PfDnaJ is capable of forming dimers in solution. (A) Products obtained after chemical crosslinking of PfDnaJint. Lane 2 is PfDnaJint in the absence of glutaraldehyde, Lane 3 shows a crosslinked product in the absence of glutaraldehyde. (B) Size-exclusion chromatography of PfDnaJint, solid arrow heads indicate molecular weight of the standard proteins; arrows denote the peaks at which protein is eluted. 1, 2, 3 possibly represent oligomeric, dimeric and multimeric forms of the protein, respectively.
Mito-Tracker Red. Distinct, non-overlapping signals were obtained for the mitochondrial marker (red) and PfDnaJ (green) indicating that PfDnaJ was not mitochondrial (Fig. 3.21, panel A and B). The PfDnaJ signal was specific to *P. falciparum*-infected cells and localized in extended organellar structures in the middle to late trophozoite stage indicating that apicoplast may be the possible site of PfDnaJ localization.

To confirm the localization of PfDnaJ in apicoplast, the ACP\textunderscore \text{leader}GFP transfected cell line of *P. falciparum* was used. Examination of parasites under immunofluorescence microscope revealed that PfDnaJ (red) signal overlapped with GFP (green) signal but not with DAPI (blue) (Fig. 3.21, panel C and D). Taken together these results confirmed that PfDnaJ localizes to the apicoplast, not to nucleus or the mitochondria.
Fig. 3.21. Immunolocalization of PfDnaJ.
Panels A and B are images obtained after confocal microscopy and show MitoTracker Red signal (1), PfDnaJ fluorescence (2) and their overlap (3) in the trophozoite. Image 4 in each panel is the corresponding phase-contrast scan while image 5 is the overall merged image. Panel A depicts images from a late trophozoite while panel B depicts images from a cell containing two middle trophozoites in a single red cell. Dotted line in panel A, image 4 demarcates parasite boundary. Panels C and D are images obtained from a fluorescence microscope. PfDnaJ (red, image 3) and ACP-GFP (green, image 2) colocalize to the apicoplast. PfDnaJ was probed with purified rabbit anti-PfDnaJ antibody (1:50 dilution), ACP-GFP was probed with monoclonal anti-GFP antibody (1:500 dilution). In both C and D panels, nuclei were counterstained with DAPI (image 1).
3.3 Transcriptional analysis of the *P.falciparum* apicoplast genome:

The apicoplast genome contains several ORFs that have been maintained in different genera despite significant sequence diversity. Studies have shown that tRNAs encoded by plDNA are transcribed (Prieser *et al.*, 1995). Transcripts have also been found for genes like *clpC*, *tufA*, *rpoB*, *rpoC1* and *rpoC2* (Wilson *et al.*, 1996).

This study initiated identification of regulatory sequence(s) and/or promoter regions of apicoplast genes. There are various softwares, based on different algorithms, which predict the consensus regulatory sequence(s) present upstream of a gene. Since the apicoplast genome is highly A+T-rich (~86.9%), presently available bioinformatic tools do not provide significant information on promoter region identification. The start sites of transcription are also not clearly defined for the apicoplast. Therefore, we first determined the nature of apicoplast transcripts (i.e. whether monocistronic or polycistronic) followed by determination of the position of the transcription initiation site(s).

3.3.1. Transcripts of apicoplast genes, ORF470/*sufB*, *clpC* and *tufA*, are detected by RT-PCR

The transcriptional status of ORF470/*sufB*, *clpC* and *tufA* genes was confirmed by RT-PCR. Total *P.falciparum* RNA was isolated at the late trophozoite-early schizont stage from parasite culture, quantitated at 260 nm and treated with DNase I to remove any DNA contamination. DNA-free RNA was used to synthesize first strand cDNA with gene specific downstream primers of the respective genes. cDNA of ORF470/*sufB*, *clpC* and *tufA* were PCR amplified using gene-specific upstream and downstream primers. For all the three transcripts, control reaction in which reverse transcriptase was not added was also carried out. PCR products along with DNA marker were electrophoresed on 1% agarose gel. Presence of product in ORF470/*sufB*, *clpC* and *tufA* RT-PCR reactions at expected position (690 bp, 545 bp and 528 bp, respectively) indicated that these genes were transcriptionally active (Fig. 3.22 A, B, C). These PCR
products were specific as no product was amplified in the control reactions lacking reverse transcriptase. The transcription of ORF470/sufB, clpC and tufA was thus confirmed.

3.3.2. Northern blotting of LSU and ORF470/sufB transcripts reveals presence of a polycistronic transcript

The size of LSU and ORF470/sufB transcripts was determined by northern blotting. Total \textit{P.falciparum} RNA was isolated at the late trophozoite stage, electrophorsed on 1.2\% denaturing agarose gel, transferred to nylon membrane and hybridized with labeled probe covering entire LSU region. Autoradiograph showed bands of sizes 4.2 kb, 3.5 kb and 2.9 kb (Fig. 3.23, A). The predicted size of LSU gene is 2695 bp, so the presence of different large- and small-size fragments relative to the predicted size of LSU indicated there might be a large precursor transcript formed which was processed during intra-erythrocytic stages of the parasite. Similarly, northern blotting of ORF470/sufB showed bands of 4.2 kb, 3.5 kb and 1.9 kb sizes suggesting that ORF470/sufB is also a part of a large transcript since the size of the ORF470/sufB gene is 1.4 kb (Fig. 3.23, B). In the apicoplast genome, the LSU gene is upstream and the \textit{rpoB} gene is downstream of the ORF470/sufB thus ORF470/sufB might be polycistronic with LSU or \textit{rpoB}. These possibilities were confirmed by performing RT-PCR using primer pairs spanning different regions of the LSU, ORF470/sufB and \textit{rpoB} genes.

3.3.3. Confirmation of polycistronic nature of transcript

The transcriptional status (monocistronic and/or polycistronic nature) of the ORF470/sufB transcript was determined by RT-PCR. For this, three sets of RT-PCR reactions were carried out using gene-specific primers spanning different regions of LSU, ORF470/sufB and \textit{rpoB} genes.
Results

Fig. 3.22. Transcriptional analysis of apicoplast genes by RT-PCR. Products were generated by RT-PCR from parasite RNA using gene specific primers for ORF470 (A); clpC (B); and tufA (C). Size of PCR product for corresponding genes is shown; no such amplification was seen in control -RT reaction. M; DNA molecular weight marker.

Fig. 3.24. Northern blot analysis of Large subunit rRNA (LSU) (A); and ORF470/sufB transcripts (B).
A diagrammatic representation of the experimental plan is given in Fig. 3.24. The primer combinations used were $\text{GIN}(\text{U})-\text{GIN}(\text{D})$, $\text{LSU}(\text{U})-\text{GIN}(\text{D})$ and $\text{GIN}(\text{U})-\text{rpoB}(\text{D})$ ($\text{GIN}$ denotes the internal region of ORF470/sufB gene). The following possibilities were considered:

a). If ORF470/sufB is polycistronic with $\text{rpoB}$, a product of 1.58 kb spanning the region between $\text{rpoB}$ and an internal region of ORF470/sufB was expected.

b). If ORF470/sufB is polycistronic with LSU then the 1.5 kb product covering the region between the LSU and ORF470/sufB gene would be obtained.

c). If ORF470/sufB is monocistronic i.e. not the part of either LSU or $\text{rpoB}$ transcript, no product would be amplified with $\text{GIN}(\text{U})-\text{rpoB}(\text{D})$ as well as with $\text{LSU}(\text{U})-\text{GIN}(\text{D})$ primer pairs.

After PCR amplification with different primer combinations, no product was seen in $\text{GIN}(\text{U})-\text{rpoB}(\text{D})$ RT-PCR reaction while a band of 1.5 kb was amplified in RT-PCR with primer pair $\text{LSU}(\text{U})-\text{GIN}(\text{D})$. This PCR product was specific since no band was present in RT reactions lacking reverse transcriptase (Fig.3.24). These results indicated that LSU and ORF470 genes are transcribed together i.e. ORF470/sufB is transcribed polycistronically with the LSU, not with the $\text{rpoB}$ gene.

3.3.4. 5'-end mapping of LSU transcript using RACE reveals presence of multiple initiation sites

Transcriptional initiation site of the LSU transcript was analysed by RACE (rapid amplification of cDNA ends, Frohman et al, 1988). RACE is a PCR based technique to determine the 5'-and 3'-end of the transcript. To determine the 5'-end of the LSU transcript, cDNA was synthesized using LSU downstream primer. Following reverse transcription, the first strand cDNA was PCR amplified with upstream universal primer and LSU downstream primer. RACE PCR products were resolved on 1.2% agarose gel. DNA fragments of 450 bp, 380 bp, 340 bp and 250 bp were eluted.
Fig. 3.24. Determination of nature of suf B transcript. (A) Diagram showing primers (arrows) used in RT-PCR and the expected products if these genes are transcribed polycistronically. Table shows possibility of products in RT-PCR using different primer pairs. PCR of cDNA with suf B primers amplified a 690 bp fragment (panel B, lane 2), primer pair LSU-sufB amplified a 1542 bp fragment (panel C, lane 2) whereas no product was amplified with sufB-rpoB primers (Panel B, lane 4) These PCR products are specific as no product was amplified in –RT reaction (panel B and C, lane 3). M; DNA molecular weight marker.
from the gel (Fig 3.25, A), cloned into T/A cloning vector (pGEM-T easy vector, Promega)) and positive clones containing insert were sequenced. Sequencing of these clones showed that the LSU transcript contains five putative start sites (Fig. 3.25, B).

In order to identify the 5'-end of the other apicoplast transcripts such as ORF470/sufB, tufA and clpC, 5'-RACE for these genes was also attempted. However, nonspecific RACE amplification products were obtained even when downstream primers were designed from different regions of these genes. This may be due to high A-T richness of the apicoplast genome due to which primers annealed nonspecifically.
Fig. 3.25. 5'-end analysis of LSU transcript by 5'-RACE. Amplified RACE products of LSU transcript were resolved on 1.2% agarose gel (panel A, lane 2), eluted, cloned into pGEM-T easy vector and sequenced. M; 100 bp DNA marker (panel B) Nucleotide sequence from 5'-region of LSU gene and arrows indicate start sites identified by RACE. The underlined nucleotides indicate reverse primer sequence used in RACE reaction.