Chapter 4

RESULTS
RESULTS

Helicases are nucleic acid dependent ATPases known to play important roles in almost all aspects of nucleic acid metabolism. DEAD-box family is a small subset of otherwise large helicase group and are known to play unique and indispensable roles in every facet of RNA metabolism (Tanner and Linder, 2001). *Plasmodium falciparum*, a parasitic protozoan has 22 DEAD-box genes, whose role in parasite’s development and nucleic acid metabolism are not known (Tuteja and Pradhan, 2006). Moreover, a previous pioneering study determined that RNA helicase like proteins might be involved in the antimalarial action of the drug choloquine (Thelu et al., 1994). This previous study along with the lack of knowledge about the roles of these DEAD-box proteins prompted us to isolate and characterise a bonafide member of the DEAD-box helicase family from *Plasmodium falciparum*. The member with PlasmoDB number PF14_0563 is homologue of Dbp5/Rat8, which is involved in the export of mRNA from the nucleus to the cytoplasm. In Yeast Dbp5p is required for providing an overall directionality to the process of mRNA export and is indispensable for the organism as knockouts of this gene are lethal (Tseng et al., 1998; Tran et al, 2007).

4.1 Molecular cloning and sequencing of the gene encoding helicase from *Plasmodium falciparum*

Using the bioinformatics tools previous studies in the lab were able to assign PF14_0563 as the bonafide homologue of yeast and mammalian Dbp5 (Tuteja and Pradhan, 2006). This DEAD-box helicase shows high degree of homology with other Dbp5 homologues and also has the characteristic signature motifs of Dbp5. Helicase gene ‘PF14_0563’ has all the 9 conserved motifs of the DEAD-box helicases. The complete ORF has 2226bp and is located on the 14th chromosome of *Plasmodium falciparum* and is not interrupted by introns. N-terminal of Dbp5 is only required for the interaction with the components of the NPC and doesn’t play any active role in the activity of the protein. Therefore an N-terminal truncation
RESULTS

was made for the purpose of cloning and expression of PF14_0563. The primers 20F1 (forward primer) and 20R1 (reverse primer) were designed for the purpose of amplification of PF14_0563 from the genomic DNA template. The primers 20F1 and 20R1 contain HindIII and XhoI sites respectively for the purpose of manipulation of the DNA. PCR amplification of PF14_0563 was done from genomic DNA and the PCR products were then analysed on an agarose gel. Upon PCR amplification of PF14_0563, a single band of ~1.73kb was observed (Figure 1A). The PCR product was then cloned into the pGEMT-Easy vector and the positive clones were then analysed by colony PCR and restriction digestion. The DNA clone were then sequenced using the dideoxy sequencing reaction. The ~1.73kb DNA band was excised using HindIII and XhoI enzymes and gel purified for subcloning into the expression vector pET 28b. The sequence analysis showed the complete clone with the methionine at the start and termination codon at the end. The sequence was finally submitted to GenBank and the accession numbers is EF612437 (Figure 2).

Figure 1: PCR amplification and pGEMT-Easy cloning of PfD66

(A) Lane 1, 1kb ladder; Lane 2 and 3, PCR product of PfD66 using genomic DNA as the template (B) Lane 1 and 2, clone 1 and clone 2 genomic DNA (C) Lane 1, 1kb ladder; Lane 2 and 3, HindIII and XhoI digested clone 1 and clone 2 of PfD66-pGEMT-easy.
RESULTS

The 1731 base pair genomic sequence codes for a protein of the size of ~66kDa [hence the protein was named *Plasmodium falciparum* Dbp5 66kDa in size (PfD66)]. PlasmoDB defines PF14_0563 as a putative DEAD-box protein of unknown functionality. The PfD66p is an acidic protein having an isoelectronic point of 4.87. Moreover, the sequence analysis shows the presence of all the 9 conserved motifs characteristic of the DEAD-box protein family (Figure 3). Figure shows the genomic DNA sequence of PfD66 and the schematic diagram showing various motifs of PfD66 (Figure 3).

```
ATGCCAAGTGTGATCTTTAAAAAAAAATCCAAGTGATTTAATAACAAAAAGAGAGAAAATG
AAGATATATATATCTACATCGGATACACAAAAATGTTAGTAATAATAATAACAAAAAGT
GAAGTGATACACTAGCACAACATACACAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
TTGTTTATGTTACACTACTCAGAGAAAAATGTTAGTAATAATAATAATATAAAA
TAATGTTGAGCAGAAAAAAGTTTATCGGAGAGAGACACTAATAATAATAAATAGAGAAG
```

Figure 2: Nucleotide Sequence of amplified PF14_0563/PfD66 from genomic DNA as submitted to Genbank (Accession No. EF612437)
RESULTS

Figure 3: Schematic diagram showing the various conserved motifs of PfD66. Open boxes represent the conserved helicase motifs and the conserved amino acid sequence of each motif is written by a single letter code inside the box. The numbers above the motifs are the number of amino acids separating the various motifs.

4.2 Amino acid alignment of PfD66 with the homologues in other organisms.

PfD66 is unannotated in PlasmoDB and moreover, very little is known about the function of this protein in *Plasmodium falciparum*. In order to determine its role in the parasite PSI-BLAST was carried out to determine the homologues of the protein. PfD66 shows high degree of homology ranging between 91-95% to its other related homologues of *Plasmodium* species like *Plasmodium berghei* (PB000472.01.0), *Plasmodium vivax* (PVX_117510), *Plasmodium yoelii* (PY06529), *Plasmodium chaubaudi* (PC000054.02.0) and *Plasmodium knowlesi* (PKH_124710). Dbp5 homologues of both *Plasmodium falciparum* and *Plasmodium vivax* have an unusually large N-terminal as opposed to the homologues of higher eukaryotes, (Figure 4). Moreover, PfD66 homologues are present in other apicomplexan species like *Toxoplasma gondii* (XP_002369936.1), *Cryptosporidium parvum* (XP_627359), *Theileria parva* (XP_765582), *Cryptosporidium muris* (XP_002142519), *Babesia bovis* (XP_001609208) and *Theileria annulata* (XP_954649). The domain organisation of PfD66 with the orthologous proteins from other organisms was done and there was a consistent conservation in the C terminal domain of all the homologues (Figure 4).
RESULTS

Orthologues of Dbp5p are also annotated as “DEAD-box RNA helicase” in other species like *Plasmodium vivax* (PVX_117510), *Plasmodium yoelii* (PY06529), *Plasmodium chabaudi* (PC000054.02.0) and *Plasmodium berghei* (PB000472.01.0). BLAST (Altschul et al., 1997) search using yeast Dbp5 as the query, revealed the presence of very closely related proteins in other apicomplexan species like *Cryptosporidium parvum* (XP_627359), *Babesia bovis* (XP_001609208), *Toxoplasma gondii* (EEB02796), *Theileria parva* (XP_765582), *Theileria annulata* (XP_954649) and *Cryptosporidium hominis* (XP_666618). Multiple sequence alignment of these genes was done using the CLUSTAL W program with the help of Mac vector software version 7.2.2 (Figure 5). DEAD-box proteins differ from each other due to the individual variations in the N and C terminal sequences as well as the distance between the conserved motifs. The amino acid alignment data clearly indicates that PfD66 has evolved from its human orthologue through a different lineage (Figure 6).
RESULTS

Figure 5 Multiple amino acid sequence alignment of PD66
Figure 6 Phylogenetic tree of PfD66. Dendrogram showing the phylogenetic relationship between the deduced PfD66 and other homologues. The names are mentioned to the right side. The phylogenetic tree was constructed using the CLUSTAL W program based on the amino acid sequence similarity of PfD66. A value of 0.1 corresponds to a difference of 10% between the two sequences.

4.3 Homology Modelling of PfD66 Based on Crystal Structure of other DEAD-Box Proteins.

The complete sequence of PfD66 was submitted to the Swiss-Model program (http://swissmodel.expasy.org) for the purpose of molecular modelling. PfD66 primary sequence shows a 36.67% identity with the human DEAD-box RNA-helicase DDX19 chain B. The structure modelling of the PfD66 was therefore done using the known crystal structure of the DEAD-box RNA-helicase DDX19 as the template (3ewsB) (Figure 7A). The ribbon model of PfD66 created by the Swiss-Model software was visualized using the chimera molecular modelling program (UCSF) (Figure 7B). It is interesting to note that even though
RESULTS

the sequence identity of PfD66 and the template (human DEAD-box RNA-helicase DDX19) is only ~37% but the overall structure of PfD66 is highly similar and superimposable to that of human DEAD-box RNA-helicase DDX19 (Figure 7C).

Figure 7 Computer based structural modelling of PfD66. (A) Ribbon model of the crystal structure of the human DEAD-box RNA-helicase DDX19 in complex with ADP (shown in salmon colour) shown using the UCSF chimera program. (B) Three dimensional structure of PfD66 (shown in yellow colour) obtained using the DDX19 (3ewsB) as the template. (C) Superimposition of the PfD66 structure on the parent template of the human DEAD-box protein DDX19 (3ewsB).
4.4 Expression and purification of PfD66

For the purpose of expression and further biochemical characterization of PfD66, 1731 bp of the complete open reading frame was subcloned into the expression vector pET28b from Novagen (Madison, WI, USA). PfD66 already cloned into pGEMT-Easy vector was digested with HindIII and XhoI enzymes and the insert was stitched into the pET28b multiple cloning site using the HindIII and XhoI sites to generate the recombinant construct pET28B-PfD66. The recombinant construct so generated was confirmed for the correct orientation and size by digesting with the HindIII and XhoI (Figure 8B, lane 3). Moreover, the correct reading frame and orientation was further confirmed by sequencing. The recombinant construct pET28B-PfD66 was transformed into E. coli strain Rosetta (DE3) pLysS and the expression of recombinant protein was induced using 1mM IPTG. The expressed protein was purified using Ni-NTA (Qiagen, GmbH, Germany) affinity chromatography. The recombinant Histidine tagged protein was eluted with 150mM imidazole in protein buffer having 20mM Tris-HCl, pH 8.0, 250mM NaCl, 10% glycerol and protease inhibitor cocktail from Sigma (St. Louis, MO, USA). The purity of the Ni-NTA purified recombinant PfD66 was checked using a 10% SDS-PAGE gel. SDS-PAGE analysis showed a ~66 kDa protein band on the SDS gel having 95% purity (Figure 8C). An identical gel was loaded and proteins were electrophoretically transferred onto the nitrocellulose membrane. The membrane was incubated with anti-His antibody and developed using standard protocols (Figure 8E). The purity of the protein was further checked with silver staining method (Figure 8D) (Sambrook et al., 1989). SDS-PAGE along with silver staining containing two different concentration of the purified protein showed that PfD66 is homogenous and contain no contaminating impurities. The purified protein was used for biochemical characterization and production of polyclonal antibodies in mice.
Figure 8 Cloning, expression, purification, silver staining and western blot analysis of PfD66 protein.

(A) Vector map of pET28B, the gene was cloned into HindIII and XhoI site. (B) Restriction analysis of the construct pET 28b-PfD66 (Lane M- 1kb ladder, Lane 2 - single cut pET 28b vector, Lane 3 – Hind III and Xho I digested recombinant plasmid. (C) The purified protein PfD66 expressed in E. coli and visualized by coomassie blue staining (Lane M contains the protein molecular weight marker and Lane 1 and 2 contain 0.5µg of the protein each (D) Silver stained protein PfD66. (E) western blot of the same samples.
RESULTS

4.5 DNA helicase activity of PfD66

The expressed and purified PfD66p was used for determining its biochemical properties. Various activities which are characteristic of DEAD-box proteins like helicase and ATPase were performed with the recombinant purified protein as described in the following sections.

DNA helicase activity of PfD66 was measured by assaying the displacement of $^{32}$P-labeled DNA from a partial duplex DNA substrate. The DNA substrate used for studying helicase assay is a $^{32}$P-labelled 47-mer DNA oligodeoxynucleotide (5'- (T)$_{15}$GTTTTCAGTCAGCAG(T)$_{15}$-3') annealed to M13mp19 phage ssDNA to create a partial duplex. At both the 5' and 3'- ends, this oligodeoxynucleotide contains 15bp of non-complementary region for the helicase to load itself, as most of the helicase require single stranded regions to load themselves. Ten nanograms of this oligodeoxynucleotide was labelled at 5'-end with T4 polynucleotide kinase (5U) in 50mM Tris-HCl, pH 8.0, 10mM MgCl$_2$, 5mM DTT, 0.1mM EDTA, 0.1mM spermidine and 1.85 MBq of [$\gamma$-32P]ATP (specific activity 222 TBq/mmol). The labelled oligodeoxynucleotide was then annealed with 2.5µg of single-stranded circular M13mp19 (+) DNA in 20mM Tris-HCl (pH 7.5), 10mM MgCl$_2$, 100mM NaCl and 1mM DTT by heating at 95 ºC for 1 min, transferring immediately to 65 ºC for 2 min and then cooling slowly to room temperature. Non-hybridized oligodeoxynucleotide was removed by gel filtration through a 1 ml Sepharose 4B column (Pharmacia, Sweden) with 10mM Tris-HCl (pH 7.5), 1mM EDTA and 100mM NaCl. The substrate and products were separated by electrophoresis on a nondenaturing 12% PAGE, dried, and the gel was exposed to hyper film for autoradiography. Figure 9 (lane 1-6) shows the substrate purified using the sepharose 4B column, whereas the lane 7 shows the boiled sample of fraction 2.
RESULTS

Figure 9 Helicase assay substrate. Lanes 1-6 are the purified fractions of the substrate on a sepharose 4B column. Lane 7 is the boiled sample of the Fraction 2.

4.6 Kinetics of unwinding

The biochemical properties of the Protein PfD66 were studied after expression and purification of the truncated protein. For that different kind of helicase and ATPase activities were carried out with purified recombinant PfD66 protein. The DNA unwinding activity of PfD66 was characterized by assaying the displacement of $^{32}$P-labeled DNA from a partial duplex DNA substrate. The assay was carried out using ~1 ng (40 pM or 0.40 pmole/10 μl) of the substrate and 20 ng of PfD66 enzyme (20 nM or 0.20 pmole/10 μl). The substrate described in section 4.5 was used for assaying the unwinding activity of PfD66. The substrate used for most of the characterization contains hanging tails of 15 bp on both the 5' and 3' ends. The structure of the substrate is shown at left side of the gel (Figure 10). The reaction tubes having heat denatured/boiled substrate and without enzyme (control) were used to determine the activity as controls (Figure 10, Lane 1 and lane2, respectively). Since, PfD66 did not show activity if Mg$^{2+}$(Lane 3) or ATP (lane 4) was omitted from the reaction. It was concluded that PfD66 like the other DEAD-box proteins contained the characteristic ATP and Mg-dependent DNA unwinding activity (Figure 10, Lane 5).
RESULTS

Figure 10 Helicase activity of purified PfD66. Lane 1, heat denatured substrate Lane 2, control reaction without enzyme, Lane 3, reaction without Mg\(^{2+}\), Lane 4, reaction without ATP and Lane 5, reaction with enzyme in the presence of ATP and Mg\(^{2+}\).

4.7 Kinetics of Unwinding in a Time and Concentration Dependent Fashion

The strand displacement assay measures the unwinding of \(^{32}\)P-labelled DNA fragment from a partially duplex nucleic acid substrate. The kinetics of the DNA unwinding reaction of PfD66 using 20ng of the purified enzyme and optimal assay conditions and \(~\)1ng (40pM) of the substrate in buffer having 2 mM ATP, 0.5 mM MgCl\(_2\) and 75 mM KCl showed a linear rate up to 60 min (Figure 11A, lanes 3-7). The structure of the substrate is shown on the left side of the autoradiogram of the gel (Figure 11). On further incubation a deviation in the linearity of the activity was observed. Whereas the titration of unwinding activity with increasing amount of purified PfD66 showed linearity up to 20ng of protein (Figure 11B, lanes 3-7). Further increase in the enzyme activity did not occur on concomitant increase in the concentration of the enzyme (Figure 11B, lanes 8-10). To find out the optimum concentration of ATP, Mg\(^{2+}\) and KCl the DNA unwinding activity of PDH47 was performed at increasing concentration of these factors.
RESULTS

(A) Time curve

(B) Concentration curve

Figure 11: DNA unwinding activity of PfD66. (A) Time dependent curve of PfD66, the time of reaction is mentioned on the top of each lane of autoradiogram. (B) Concentration dependent curve of PfD66, the various concentration of the pure PfD66 are mentioned on the top of the autoradiogram.

4.8 Unwinding in the Presence of Different Concentration of ATP and MgCl₂

In order to find the optimum amount of ATP and Mg²⁺ required for DNA unwinding, helicase assay was carried out in the presence of different concentration of ATP and MgCl₂. Unwinding assay was carried out in optimal assay conditions as determined through previous experiments using ~1 ng (40 pM) of the substrate and 20 ng of PfD66 enzyme (40 nM). It was observed that the unwinding activity of PfD66 is ATP dependent, as no activity was observed in the absence of ATP (Figure 12A, lane 2). At 2 mM conc. of ATP PfD66 exhibited maximum unwinding activity (Figure 12A, lane 6). At 8 mM conc. of ATP (Figure 12A, lane 9) the activity of PfD66 was inhibited. In case of Mg²⁺ (Figure 12B, lanes 3-9), 1 mM concentration (Figure 12B, lane 5) showed highest unwinding activity but 8 mM Mg²⁺ concentration (Figure 12B, lane 9) showed complete inhibition.
RESULTS

Figure 12 Concentration curve of ATP and Magnesium for DNA unwinding by PfD66
The helicase reaction was performed using pure PfD66 and 1 ng of $^{32}$P labelled substrate using varying concentrations of ATP (A) and Magnesium (B). The concentration used are mentioned on the top of the autoradiogram, C is control without enzyme and B is boiled substrate.

4.9 Unwinding in the presence of different kind of divalent cations and different kind of NTP's/dNTP's

To find the optimum assay condition for different kind of ions and different dNTP/NTP requirement DNA unwinding activity was carried out using ~1 ng (40 pM or 0.40 fmole/10 μl) of the substrate and 20 ng of PfD66 enzyme (40 nM). It was observed that for the purpose of unwinding PfD66 utilises Mg$^{2+}$ for activity followed by Mn$^{2+}$ and Zn$^{2+}$. Whereas other divalent cations like Co$^{2+}$, Ni$^{2+}$, Ca$^{2+}$ and Ag$^{2+}$ were unable to support any significant unwinding activity (Figure 13A, lane 5, 6, 8 and 9 respectively). For unwinding activity, ATP and dATP were the best utilized cofactor by the PfD66 (Figure 13B, lanes 3 and 4 respectively) while other NTPs/dNTPs such as dGTP, GTP, CTP, dCTP, TTP, and UTP were effective to some extent (Figure 13B, lanes 5-10 respectively).
RESULTS

Figure 13 Effect of various cations and NTP/dNTPs on the unwinding activity of PfD66. The helicase reaction was performed using pure PfD66, 1 ng of $^{32}$P labelled substrate using various types of cations (A) and NTP/dNTPs (B). The ions and NTPs/dNTPs used are mentioned on the top of the autoradiograms, C is control without enzyme and B is boiled substrate.

4.10 Determination of direction of unwinding of PfD66

Helicases are motor proteins that require a sense of direction while translocating over bipolar DNA duplex. The direction of unwinding by helicase is defined by the strand to which the helicase binds and preferentially unwinds in a polar fashion by moving on the bound strand. Most of the SF2 family members have 3′-5′ directionality with a few exceptions like the XPD helicases. In order to determine the direction of translocation of PfD66, two direction specific substrates were prepared, one specific for the 3′-5′ direction and the other one specific for 5′-3′ direction. The preparation of the 3′-5′ and 5′-3′ direction substrates was done using a 32mer oligodeoxynucleotide having the sequence 5′-TTCGAGCTCGGTACCCGGGGATCCTCTAGAGT-3′. For constructing 5′ to 3′
unwinding substrate, the oligodeoxynucleotide 32-mer was first annealed to M13mp19 ssDNA and then labelled at 3'-OH end in 40mM Tris–HCl (pH 7.5), 10mM MgCl₂, 50mM NaCl and 1mM DTT with 50μCurie [α-³²P]dCTP and 5 units of DNA polymerase I (large fragment) at 23°C for 20 min. The incubation was continued for an additional 20 min at 23°C after increasing the dCTP to 50mM using unlabelled dCTP. This was digested with SmaI and purified by gel filtration through 1 ml sepharose 4B column. The substrate consisting of long linear M13ssDNA with short duplex ends for 3' to 5' unwinding was prepared by first 5'-end labeling of 32-mer oligodeoxynucleotide and then annealing with M13mp19 ssDNA as described above. The annealed substrate was digested with SmaI and purified by gel filtration through 1 ml of Sepharose 4B. Both the substrates are shown in Figure 14A and B and the structure is shown in upper panels of Figure 15A and B. The unwinding activity of PfD66 was determined using ~1 ng of these direction-specific substrates and 4 different concentrations of pure PfD66 enzyme. The results of this DNA unwinding assay using these two substrates indicated that the enzyme PfD66 is bipolar in nature. The results further confirmed that the helicase activity of PfD66 with both of these substrates was concentration dependent (Figure 15A, lanes 3-6 and Figure 15B, lanes 3-6). It was interesting to note that DNA unwinding activity was more with 3' to 5' direction specific substrate (Figure 15A, lanes 3-6) as compared to the 5' to 3' direction specific substrate (Figure 15B, lanes 3-6).

![Figure 14](Image)

Figure 14 Autoradiogram showing (A) substrate for 3'-5' directionality and (B) substrate for 5'-3' directionality.
Figure 15 Direction of unwinding of PfD66. Bipolar DNA unwinding activities of PfD66. The construction of the 3'-5' (A) directionality and 5'-3' (B) directionality substrate are mentioned at the top of the panels. In the autoradiogram the enzyme concentration is mentioned at the top of the autoradiogram where C is the substrate without enzyme and B is heat denatured substrate.
RESULTS

4.11 ATPase activity

Helicase use the energy of ATP hydrolysis to unwind duplex molecules or for translocation over the nucleic acid strand. Most of the DEAD-box proteins show nucleic acid dependent ATPase activity. Therefore the ability of purified PfD66 protein to hydrolyze and release free radioactive phosphate (Pi) from [γ-32P] ATP was tested in presence and absence of Mg2+ and ssDNA. The reaction products were separated by ascending chromatography on polyethyleneimine-cellulose thin layer plates followed by autoradiography. No ATPase activity was observed in the absence of protein (Figure 16). Absence of ssDNA or MgCl2 from reaction also did not support the activity (Figure 16, lane 2 and lane 3). PfD66 exhibited MgCl2 and ssDNA dependent ATPase activity (Figure 16, lane 4).

Figure 16 ATPase activity of purified PfD66. Lane 1, reaction without enzyme, Lane 2, reaction without Mg2+, Lane 3, reaction without single stranded mp13 DNA (ssDNA) and Lane 4, reaction with enzyme in the presence of ssDNA and Mg2+.
RESULTS

4.12 Time dependence of ATPase activity

Time dependence of ATPase activity was studied by studying the release of radioactive phosphate (Pi) from [$\gamma^{32}$ P] ATP by PfD66 at different time points respectively. The percent release of radioactive phosphate (Pi) from [$\gamma^{32}$ P] ATP by PfD66 enzyme was calculated and plotted as graphs. The kinetics of the ssDNA-dependent ATPase reaction was studied under standard assay condition using 30 ng of purified PfD66 enzyme and 1665 Bq [$\gamma^{32}$P]ATP as a substrate in the presence of 100 ng of M13 ssDNA. The percent release of radioactive phosphate (Pi) from [$\gamma^{32}$P] ATP showed linearity up to 120 minutes (Figure 17, lane 7) and deviation from linearity was observed upon longer incubation.

![Graph showing time dependence of ATPase activity of PfD66.](image)

**Figure 17 Time dependence of ATPase activity of PfD66.** The standard reaction contained 100ng ssDNA, purified protein and 0.5 mM MgCl$_2$. The enzyme activity data was plotted in form of a graph. The time of incubation is mentioned at the top of the autoradiogram and C is the control reaction without enzyme.
4.13 ATPase Activity in the Presence of Different Kind of Nucleic Acid Species

The ATPase activity of purified PfD66 was tested by its ability to release radioactive Pi from $[^\gamma-32P]ATP$ in the presence of different kind of nucleic acid to find its preference for the same. This activity was done to ascertain that the ATPase activity of PfD66 is stimulated by Poly (A) rich RNA species or not. The reaction products were separated by ascending chromatography on polyethyleneimine-cellulose thin layer plates followed by autoradiography (Figure 18A). For quantititative data the %Pi release was calculated and plotted as a histogram for the various nucleic acid species (Figure 18B). The effect of different polynucleotides along with the standard mp13ss DNA on the ATPase activity of PfD66 (30 ng) was checked. The amount of these species used was 100 ng in each case, which allowed maximum stimulation of ATPase activity. The data showed that the PfD66 exhibited less stimulation of ATPase activity in the presence of same amount (100 ng) of M13 ss DNA (lane 8 of Figure 18A and B) as compared to Plasmodium falciparum (Pf) total RNA (lane 7 of Figure 18A and B). The ATPase activity was RNA stimulated as out of total RNA, mRNA is the natural substrate for the enzyme. In the presence of partial duplex DNA helicase substrate the PfD66 showed more activity (Figure 18A and B, lane 8). Out of the four poly (N) RNA oligos, namely poly (G) (lane 3 of Figure 18A and B), poly (T) (lane 4 of Figure 18A and B), Poly (C) (lane 5 of Figure 18A and B) and poly (A) oligo (lane 6 of Figure 18A and B), poly (A) oligo fraction showed the highest overall ATPase activity. All these results suggest that it is the poly (A) fraction of the mRNA which stimulates the enzyme. Moreover the stimulation of the ATPase activity by poly (G), poly (T) and poly (C) RNA oligo was roughly the same as the M13 ss DNA (lane 8 of Figure 18A and B). All these results suggest that PfD66 activity is probably stimulated by the poly (A) rich mRNA fraction of the RNA in vivo.
Figure 18 Stimulation of the ATPase activity of PfD66 in the presence of various nucleic acid species. The % Pi release was calculated and plotted as histogram below the autoradiogram. The numbers below the histogram correspond to the lanes of the autoradiogram. Various DNA/ RNA substrates were used to stimulate the ATPase activity are mentioned at the top of the autoradiogram.
4.14 Preparation of the truncated derivatives of PfD66

Q motif is the most recently discovered motifs of the DEAD-box family of helicases (Tuteja and Pradhan, 2006). In order to check the roles of the Q motif and the helicase N-terminal in the ATPase, DNA unwinding and RNA binding activities, two truncated derivatives of PfD66 were made: PfDT1 (~44kDa) and PfDT2 (~30kDa). Two serial truncations of PfD66 were made, where in the first truncation (PfDT1) Q motif was completely deleted and in the second truncation (PfDT2) helicase N-terminal comprising of the Q, PTRELS, TPGK and DEAD was completely deleted (Figure 19A). PfDT1 was amplified using the forward primer T1F (5'CAAgCTTgCTTATgCATTgCCTATA3') and the reverse primer T1R (5'CCTCGAGTTAATTTTTCAATTTGGTCA3') whereas, PfDT2 was amplified using the forward primer T2F (5'CAAGCTTATGTCTTCACAAGTAgAAACT3') and the reverse primer T2R (5'CCTCGAGTTAATTTTTCAATTTGGTCA3') and the same conditions as described previously for purifying the protein PfD66 (under section 3.2.14; Materials and Methods). The resulting clones were verified by sequencing and the fragments were sub-cloned in the protein expression vector pET-28b obtained from Novagen (Madison, WI, USA). The truncated proteins were purified using the same protocol as previously described for purifying the protein PfD66 under section 3.2.14 (Materials and Methods) and using a 10% SDS-PAGE gel (Figure 19B). Same concentrations (20 ng) of PfD66, PfD66T1 and PfD66T2 were checked for DNA helicase and ssDNA-dependent ATPase activities. The results clearly showed that only PfD66 showed the DNA helicase and ATPase activities, whereas the truncated proteins were ineffective in showing any activity (Figure 19C, D).
Figure 19 Characterization of truncated derivatives of PfD66 (A) Schematic drawing showing the various conserved motifs of PfD66, PfDT1 and PfDT2. Open boxes represent the conserved motifs and the name of the motif is written inside the box. (B) The purified proteins expressed in E. coli and visualized by coomassie blue staining. Lane M contains the protein molecular weight marker and Lane 1- PfDT1 and Lane 2 contain PfDT2. (C) DNA helicase activity of PfD66 and its truncated derivatives: PfDT1 and PfDT2. Lane 1 is boiled substrate and Lane 2 is control without enzyme. (D) ssDNA dependent ATPase activity of PfD66 and its truncated derivatives. Lane 1 is control without enzyme and position of ATP and Pi are marked on the left hand side of the autoradiogram.
RESULTS

4.15 RNA binding activity of PfD66 and its truncation derivatives.

In order to check the efficiency of PfD66 and its truncated derivatives to bind to RNA, the RNA binding propensity analysis was performed. The bioinformatics analysis of PfD66 sequence using the program RNA interface residue prediction was done for the protein 3D structure obtained from Swiss-Model program ([http://swissmodel.expasy.org](http://swissmodel.expasy.org)). The software available on ([http://yayoi.kansai.jaea.go.jp/qbglkygl/index.php](http://yayoi.kansai.jaea.go.jp/qbglkygl/index.php)) was used to map the residues having high propensity to bind RNA. The interface residue prediction program labels the amino acid having high probability of being present at interface as red and the buried amino acids (not considered as an interface residue) in deep blue (Figure 20A). Using the RNA binding propensity parameters, the 3-D structure of PfD66 was modelled in such a way that individual amino acids were coloured according to color coding scheme showing their ability to bind RNA (Figure 20B). In order to find the RNA binding efficiency of PfD66, PfDT1 and PfDT2, the RNA binding assay was performed using the methods described previously (under section 3.2.29 in Materials and Methods). The experiment was repeated three times and the result was found to be reproducible and a representative autoradiogram of the result is shown in Figure 20D. The result of the experiment showed that PfD66 shows maximum efficiency in binding RNA (Figure 20D, Lane 3) than PfDT1 (Figure 20D, Lane 4) and PfDT2 (Figure 20D, Lane 5) and BSA was used as a control, which showed no binding to the RNA (Figure 20D, Lane 2). The results were expressed as percentage RNA binding with PfD66 having 100% RNA binding efficiency. The histogram shows that PfDT1 and PfDT2 bind RNA with 58% and 17% RNA binding respectively as compared to the full length PfD66 (Figure 20E). An identical blot of PfD66, PfDT1 and PfDT2 was probed with anti-his antibody, which confirmed equal loading of protein (Figure 20C, Lane 1-3).
Figure 20 RNA binding activity of PfD66 and its truncated derivatives (A) Color coded scale of RNA binding. (B) Model of PfD66 showing the residue responsible for RNA binding according to the color coded scheme shown in (A). (C) Western blot probed with anti-his antibody. (D) RNA binding activity. (E) A graphical representation of the data of panel (D).
RESULTS

4.15 RNA helicase activity of PfD66

After we established that PfD66 contains RNA dependent ATPase activity, we determined the RNA helicase activity of PfD66. RNA helicase assay was done under optimized conditions using PfD66 and partially duplex substrate. The RNA helicase substrate was prepared by using the RNA oligonucleotides synthesized from Primm srl (Milan, Italy). The sequence of the oligonucleotides is as follows: 39-mer 5'-GGGAGAAUAUCACUCGUGAGGCUAUCCGUAAAGCACGC-3' and 13-mer 5'-AUAGCCUCAACCG-3'. 10 ng of the 13-mer oligonucleotide was labeled at 5'-end with T4 polynucleotide kinase (5U) and the labeled oligonucleotide was then annealed with the 39-mer oligonucleotide and purified as described for DNA helicase substrate. The RNA helicase assay was performed in the same way using this substrate. The RNA helicase activity of PfD66 was checked by using this RNA-RNA substrate with two different concentrations of PfH45 (50 nM and 100 nM). The results clearly show that PfD66 contains concentration dependent RNA unwinding activity (Figure 21, lanes 2 and 3 respectively).

Figure 21 RNA helicase activity of PfD66. The structure of the substrate is shown on the left side of the autoradiogram. Asterisk denotes the 32P-labelled end. The concentrations of PfD66 used are indicated at the top of each lane. Control (lane 1) and boiled (lane 4) are reaction without enzyme and heat denatured substrate respectively.
4.16 Effect of DNA-Interacting Ligands on DNA Unwinding Activity of PfD66

Antihelicase activity of a variety of DNA intercalating agents was checked using the standard helicase assay. Three classes of nucleic acid interacting agents used in the study were: (i) DNA-intercalating agents, such as daunorubicin, ethidium bromide, ellipticine, nogalamycin, cyclophosphamide, mitoxantrone, and actinomycin D, and (ii) minor groove binders, such as distamycin and netropsin, and (iii) non-intercalating topoisomerase inhibitors, such as camptothecin, VP-16, novobiocin and DNA polymerase alpha inhibitor like aphidicolin. The effect of these ligands on the DNA unwinding activity of PfD66 (20 ng) was examined by including 50 μM of each ligand separately in the standard helicase assay using 1 ng of the hanging tails substrate (Figure 22). The results indicated that compounds like actinomycin D, camptothecin, DAPI, ethidium bromide, nalidixic acid, netropsin, daunorubicin and nogalamycin inhibited the activity of PfD66 effectively (Figure 22, lane 4, 5, 7, 8, 10, 11, 15 and 16 respectively) at 50μM concentration. Whereas the other compounds like aphidicolin, Cyclophosphamide, mitoxantrone, novobiocin, VP-16, ellipticine and geniticin (Figure 22, lane 3, 6, 9, 12, 13, 14 and 17 respectively) were unable to inhibit the DNA-unwinding activity of PfD66 at 50μM concentration. The results indicated that compounds like actinomycin D, camptothecin, nalidixic acid, ethidium bromide, DAPI, netropsin, Daunorubicin and nogalamycin inhibited the activity of PfD66 effectively (Figure 22, lane 4, 5, 7, 8, 10, 11, 15 and 16 respectively) at 50μM concentration. Whereas the other compounds like aphidicolin, Cyclophosphamide, mitoxantrone, novobiocin, VP-16, ellipticine and geniticin (Figure 22, lane 3, 6, 9, 12, 13, 14, 17 respectively) were unable to inhibit the DNA-unwinding activity of PfD66 at 50μM concentration.
RESULTS

Figure 22 The effect of different DNA intercalating compounds on the DNA unwinding activity of PfD66. Effect of DNA intercalating ligands on DNA unwinding activity of PfD66. The reaction was performed by purified enzyme, 40pM of the substrate and 50μM of the ligand. The compound added is mentioned at the top of the histogram and C is control, No inhibitor is control having enzyme without inhibitor and Boiled is having the heat denatured substrate. Asterisks denote the 32P-labelled end.

The compounds, which inhibited the helicase activity, were investigated further for the kinetics of inhibition. For this, each inhibitor was included in the helicase reaction at the final concentration ranging from 0.25 to 50μM. The concentrating curve of these inhibitors is shown in figure 23 A-D. The apparent IC50 was calculated for each of the inhibitors from the data. The most effective inhibitors were netropsin and ethidium bromide, with apparent IC50 values of 0.5μM and 1μM respectively (Figure 23A, B) followed by DAPI and Nogalamycin with apparent IC50 values of 1.9 and 5μM respectively (Figure 23 C, D). The structure of some of these compounds which inhibit the helicase activity is shown in Figure 26.
Figure 23 A-D Kinetics of inhibition of DNA unwinding activity of PfD66. The standard helicase reaction was performed with 20Nμm of PfD66 and about 1ng of the substrate. The compounds used are written above the autoradiogram. The DNA helicase reaction was performed in the presence of increasing concentration of various compounds (written on the top of the autoradiogram). The concentration curve for each of these compounds is shown on the right side of each autoradiogram. In each panel, C and B are reactions without enzyme and heat denatured substrate respectively, whereas no inhibitor (0μM) is the lane with enzyme and without any compound but in the presence of 1μl of DMSO.
RESULTS

4.17 Effect of DNA-intercalating compounds on ATPase activity of PfD66

For studying the effect of DNA intercalating compounds described above on the ssDNA dependent ATPase activity, different compounds were added to the reaction mixture prior to the addition of enzyme. The effect of the compounds on the ssDNA dependent ATPase activity of PfD66 was studied by including 50μM of the compound in each case in a standard ATPase assay. The results showed that Nogalamycin, DAPI, Ethidium bromide and Netropsin (Figure 24, lane 9, 10, 13 and 18) showed significant inhibition of the ssDNA dependent activity of PfD66. Moreover, on the other hand most of the compounds like Camptothecin, Aphidicolin, cyclohexamide, VP-16, Actinomycin D, Distamycin, Ellipticine, Mitoxantrone, nalidixic acid, Cyclophosphamide, Daunorubicin and Genticin (Figure 24, lane 3, 4, 5, 6, 7, 8, 11, 12, 14, 15, 16 and 17) were not able to inhibit the ssDNA dependent ATPase activity of PfD66 significantly at 50 μM concentration.

Figure 24 The effect of different DNA intercalating compounds on the ATPase activity of PfD66. Effect of DNA intercalating ligands on DNA unwinding activity of PfD66. The reaction was performed by purified enzyme, 40pM of the substrate and 50μM of the ligand. The compound added is mentioned at the top of the histogram and C is control without enzyme and lane enzyme is control having enzyme without inhibitor. No inhibitor is control having enzyme without inhibitor and boiled lane is having the heat denatured substrate.
RESULTS

The compounds, which inhibited the ssDNA-dependent ATPase activity, were investigated further for the kinetics of inhibition. Each inhibitor was included in the reaction at the final concentration ranging from 0.5 to 10μM and the results of the concentration curve of the inhibitors are shown in Figure 25 A-D. The results show that the most effective inhibitor is Netropsin with an apparent IC$_{50}$ value of 1μM (Figure 25 A), followed by Ethidium bromide, Dapi and Nogalamycin with the IC$_{50}$ values of 1.5μM, 2μM and 3.2μM respectively (Figure 25 B-D). It was concluded from the inhibitor studies that Netropsin was the most potent inhibitor of PfD66 having the lowest Ki values for helicase and ATPase activities.

Figure 25 Kinetics of inhibition of DNA-dependent ATPase activity of PfD66 by different inhibitors. Kinetics of the inhibition of the ATPase activity of PfD66 by Netropsin (A), Ethidium bromide (B), DAPI (C) and Nogalamycin (D). The ATPase reaction was performed in the presence of increasing concentration of the inhibitor using 20nM of the pure enzyme. Various concentrations of each ligand used are mentioned on the top of each lane. Lane C in all the essay is the control without protein. The quantitative curve is shown on the right hand side of each radiogram of the TLC.
RESULTS

Figure 26 Structures of some of the DNA interacting ligands /compounds. Name of each compound is written below the structure.

4.18 Localization of PfD66 by immunoflourescence assay

Localization studies were done in order to know the biological role of the PfD66 helicase. Antibodies against the protein were raised by injecting BALB/c mice with 20μg of the purified recombinant protein. IgG fraction of the polyclonal antibody was purified using Protein A sepharose column. The purified antibodies were then used to localize this protein in
RESULTS

the various intra-erythrocytic stages of *Plasmodium falciparum*. The synchronized parasite culture was used for making thin smears on glass slides which after processing were probed with IgG against PfD66 protein followed by secondary FITC antibody and then counterstained with DAPI. PfD66 is expressed in all the three developmental stages of the parasite (the ring, the trophozoite and the schizont). The results suggest that protein is mainly distributed in the cytoplasm and small amount of the protein in the nuclear fraction is also visible (Figure 27).

![Figure 27 Localisation of PfD66 in various erythrocytic stages of P. falciparum by immunofluorescence microscopy](image)

The four time points are listed to the left side of the figure and the panel type is mentioned at the top of the figure.
RESULTS

4.19 Western blot analysis of PfD66

Expression of PfD66 at the proteomic level during the intraerythrocytic developmental stages was done by the western blot analysis. A mixed stage culture having the parasite from all the three stages was harvested for the purpose of western blotting. Total protein (100 μg) from the lysate representing the various stages of the parasite was separated by SDS PAGE and the proteins were transferred on to a nitrocellulose membrane. The blot was probed with 1:1000 dilution of purified PfD66 antisera. After washing the blot was incubated with the appropriate secondary antibody coupled to alkaline phosphate from Sigma (St. Louis, MO, USA) and developed using BCIP and NBT obtained from Sigma according to manufacturer’s instructions. The western blotting results with the protein extracts from the mixed intraerythrocytic developmental stages of Plasmodium falciparum shows that full length PfD66 (84.3 kDa) is expressed in the parasite. (Figure 28A, Lane 1)

Figure 28 Western blotting of PfD66 (A) Western blot showing the expression of PfD66 in the Mixed stage parasite culture lysate. Lane M, Protein molecular weight marker, Lane 1, culture lysate of the mixed stage parasite.
RESULTS

4.20 Components of the mRNA export pathway of *Plasmodium falciparum*

After characterizing PfD66, which is a homologue of DBP5 and is involved in mRNA transport, we studied the components of mRNA transport pathway of *Plasmodium falciparum* using bioinformatics approach. With the completion of the genome sequencing in 2002, there has been a surge in the efforts for genome annotation. *P. falciparum* contains about 5300 putative protein coding genes, of which only 2060 have been annotated by sequence similarity and manual curation (Gardner et al., 2002; Ginsburg, 2009). *In silico* methods for the purpose of gene annotation hold much promise as the classical tools of genetics and biochemistry have been slow to yield results. The A-T richness of the *P. falciparum* further hampers the gene annotation. Little is known about mRNA export in general and even less is known about mRNA export in *P. falciparum*. Most of the information about mRNA export has been deciphered using yeast and other higher eukaryotes as the model systems. Using the tools of bioinformatics we have identified the components of the mRNA export pathway in *P. falciparum*. These components were identified mainly by the analysis of the genome using BLAST and profile-Hidden Markov Model searches. Most of the components of the mRNA export pathway are conserved in *P. falciparum* but there are a few peculiarities which are described in respective sections.

4.21 SR proteins

SR proteins are serine-arginine rich abundantly found pre mRNA splicing factors, which are involved in multiple steps of RNA metabolism (Bourgeois et al., 2004). The shuttling SR proteins apart from carrying out splicing are able to couple splicing and export by acting as export adapters for the mRNA export factor (Huang et al., 2004). Moreover, SR proteins like Npl3 has been shown to provide directionality to the process of mRNA export by undergoing a cycle of phosphorylation and dephosphorylation events as described previously. SR
RESULTS

proteins like Gbp2, Hrb1 and Npl3, are recruited to mRNA by the TREX complex (Hurt et al., 2004; Huang and Steitz, 2005). Components of the TREX (THO) complex are responsible for recruiting Grb2 and Hrb1 but not Npl3, which is recruited independently by the interaction with components of the cap binding complex (Hacker and Krebber, 2004; Hurt et al., 2004). BLAST search using the *Saccharomyces cerevisiae* proteins Npl3 and Gbp2 as query in the *P. falciparum* database “PlasmoDB” (http://www.plasmodb.org/) revealed two proteins with PlasmoDB number PF10_0217 and PF10_0066 respectively. These proteins PF10_0217 and PF10_0066 are annotated in the PlasmoDB as pre-mRNA splicing factor and RNA binding protein respectively. Moreover, using bioinformatics approach we were not able to detect the homologue of the SR protein Hrp1 in *P. falciparum*. PF10_0217 and PF10_0066 contain 538 and 246 amino acids and show considerable similarity to the yeast proteins Npl3 (4e-11) and Gbp2 (2e-17) respectively (Figure 29). PfNpl3 (PF10_0217) homologue contains two RRMs in the N terminal region of the protein, whereas the corresponding yeast homologue Npl3 has the RRMs located in the middle region of the protein (Figure 29A). PfGbp2 (PF10_0066) homologue surprisingly has only two RRMs, whereas the corresponding yeast and human homologues Gbp2 and heterogeneous nuclear ribonucleoprotein have three RRMs each and are quite longer than their counterparts in apicomplexans (Figure 29B). The difference in the homologous proteins of *P. falciparum* and other higher eukaryotes is due to the divergence during the course of evolution. The structure modelling of the RRMs of the PfGbp2 was done using the RNA binding protein Fir protein as the template. The results of this modelling show that there is significant structure conservation in the RRMs of the two proteins as the two structures are completely superimposable (Figure 30).
Figure 29 Domain organizations of Npl3 and Gbp2

(A) Schematic diagram showing the domain organization in the various homologues of Npl3 in 1: Plasmodium falciparum, 2: Plasmodium vivax, 3: Toxoplasma gondii, 4: Saccharomyces cerevisiae and 5: Homo sapiens.

(B) Schematic diagram showing the domain organization in the various homologues of Gbp2 in 1: Plasmodium falciparum, 2: Plasmodium vivax, 3: Toxoplasma gondii, 4: Saccharomyces cerevisiae and 5: Homo sapiens.
Figure 30 Computer based structure modeling of Gbp2
Superimposition of the computer based structure modeling of the RRM domain of *Plasmodium falciparum* Gbp2 (shown in pink) on the RRM domain of Fir (2qfj) as a parent template (shown in yellow).

4.22 TREX complex

TREX complex is a highly conserved protein complex involved in the export of mRNA in eukaryotes. In yeast, the TREX complex is made up of a THO subcomplex consisting of Tho2, Hpr1, Mft1, Thp2, and the export factors Sub2 and Yra1 along with Tex1, a protein of unknown function (Chavez et al., 2000; Reed and Cheng, 2005). TREX complex is itself loaded cotranscriptionally to the elongating transcript and is further responsible for recruiting many proteins involved in the process of splicing and mRNA export. In *P. falciparum*, although some of the components of the TREX complex like Tho2, UAP56 (PfU52) and REF are present, but using the bioinformatics approach we were unable to detect the rest of the components of the TREX complex. It has been reported previously that all the components of the TREX complex are not thoroughly conserved across the various species like homologues of Mft1 and Thp2 have not been found in *Drosophila* and humans (). Moreover, the homologues of yeast Hpr1 and Tex1 are present in *Drosophila* and humans but using the
bioinformatics approach we were unable to detect the orthologous proteins in *P. falciparum*. It might be possible that some other proteins are playing the role of TREX homologues in *P. falciparum* and therefore the TREX complex of this organism is slightly different from the other higher eukaryotes. Using the yeast Tho2 protein sequence as the query, the homologue with the PlasmoDB number PFL2390c (PfTho2) was identified in the *Plasmodium* genome. PfTho2 is unusually long and contains 2932 amino acids and shows ~27% identity and ~49% homology to the yeast Tho2, which is only 1597 amino acid long. Tho2 in various species lacks a conserved domain and therefore no conserved domain was observed in PfTho2 using the InterProSan tool (http://www.ebi.ac.uk). Moreover, PfTho2 is unannotated and has been described as a conserved hypothetical protein in PlasmoDB. The corresponding homologues in *P. vivax* (PVX_101385) and *P. yoelii* (PY01809) are also unannotated hypothetical proteins.

Apart from Tho2, UAP56 homologue (PfU52) and Ref/Aly (described separately) are also present in the *P. falciparum*’s TREX complex (Shankar et al., 2008). UAP56 is a member of DEAD box family of RNA helicase, involved in the ATP dependent assembly of spliceosome. UAP56 is an essential protein, that has been implicated in the export of mRNA and it has been shown recently that *P. falciparum* homologue is an RNA dependent ATPase and it also has a role in the splicing processes (Jensen et al., 2001; Macmorris et al., 2003; Shankar et al., 2008). It was first identified as an interacting partner of the U2AF$^{65}$ in yeast two hybrid screens for proteins interacting with U2AF$^{65}$ (Fleckner et al, 1997). In yeast it was observed that there is rapid accumulation of poly(A) RNA upon shifting of temperature sensitive Sub2/UAP56 mutants to non permissive temperatures (Jensen et al, 2001). In Drosophila also double-stranded RNA (ds RNA) mediated depletion of HEL/UAP56 leads to growth inhibition and robust accumulation of poly(A) RNAs in the nucleus (Gatfield et al,
RESULTS

The structure modeling of the PfU52 was done using the human UAP56 as the template and it was observed that although the overall structure is conserved but insertions in the protein tend to loop out (Figure 31).

Figure 31 Computer based structure modeling of PfU52
Superimposition of the computer based structure modeling of the PfU52 of *Plasmodium falciparum* (shown in blue) on the Human Uap56 (1xti), which was used as a parent template (shown in yellow). The portion which has no overlap is shown in pink.

4.23 REF (RNA and export factor binding proteins) family

ALY/REF (Yra1p in yeast) is an evolutionary conserved family of hnRNP-like proteins, called REF (RNA and export factor binding proteins), which plays the role of an adaptor protein between RNA export factor and mRNPs. Yra1p was originally identified as a yeast nuclear protein which exhibits RNA annealing activity (Portman et al., 1997). ALY/REF is also a key component of the EJC (exon junction complex) as it remains peripherally associated with the EJC core (Bono et al., 2006; Reichert et al., 2002). At least one member of the REF family is present in higher eukaryotes. Multiple members of the REF family have been reported only in *Mus musculus*, *Xenopus laevis*, *Caenorhabditis elegans*, *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* (Stutz et al., 2000). In *Saccharomyces cerevisiae* the two members of the REF family are known by the names of
RESULTS

Yra1 and Yra2. Yra2p when overexpressed is able to complement the deletion of YRA1 in vivo, suggesting that these proteins may have redundant functions (Zenklusen et al., 2001). In *P. falciparum* PSI-BLAST (Altschul et al., 1997) search using the full length protein of yeast Yra1 as a query revealed an ORF with PlasmoDB number PFF0760w, showing ~28% identity and ~55% homology with the yeast protein. Interestingly, PFF0760w shows even more similarity to its human homologue ALY, showing ~34% identity and ~62% homology. This protein (PFF0760w) has been annotated in PlasmoDB as a putative RNA and export factor binding (REF) protein, whereas the corresponding homologues in *P. vivax* (PVX_113920) and *P. knowlesi* (PKH_113330) are annotated as RNA binding proteins. Orthologues of *P. falciparum* REFp (PlasmoDB number PFF0760w) are annotated as hypothetical proteins in *P. yoelii* (PY05533) and *P. berghei* (PB102089.00.0). Although only one member of the REF family is present in the *P. falciparum* and *P. vivax* but our bioinformatics analysis reveals that two members of the REF family are present in both *P. yoelii* (PY07541 and PY05533) and *P. berghei* (PB102089.00.0 and PB406000.00.0). Surprisingly, the proteins of the REF family from *Plasmodium* species range between 100-160 amino acids and are smaller than their eukaryotic counterparts, which contain about 200-300 amino acids (Figure 32 A). Moreover, REF family proteins are characterized by the presence of a central RNA binding domain (RBD), which is flanked on both the sides by two conserved N and C terminal domains (Figure 26 B) (Stutz et al., 2000; Zenklusen et al., 2001; Rodrigues et al., 2001). The conserved N and C terminals of the REF proteins are separated by variable insertions of positively charged amino acids (Stutz et al., 2000). In *Plasmodium* species, the RBD of the REF family is located more towards the C terminal end with a conserved N terminal domain. The C terminal domain of REF in apicomplexans is inconspicuous and is almost fused to the RBD (Figure 32 B).
Figure 32 Domain organisation of *Plasmodium falciparum* REF

(A) Schematic diagram showing the domain organization in the REF family of 1: *Plasmodium falciparum*, 2: *Plasmodium vivax*, 3: *Toxoplasma gondii*, 4: *Saccharomyces cerevisiae* and 5: *Homo sapiens*. (B) Comparison of the domain organization in the Ref proteins of higher eukaryotes and apicomplexans. The conserved N and C terminal domains flanking the RRM motif are shown as REF-N and REF-C respectively. N-v and C-v represents the N and C terminal variable regions (The figure has been prepared with the help of information given in Stutz et al., 2000).

4.24 Tap-p15 pathway

Tap/Mex67 is the principal export factor involved in eukaryotes which mediates the export of bulk of mRNA from the nucleus (Hurt et al., 2000). In yeast there is a single Tap/NXF homologue known as Mex67, but in higher eukaryotes there are multiple members of the NXF family formed by separate gene duplication events (Herold et al., 2000). Tap/Mex67 is a modular protein having three distinct domains: a leucine rich repeat domain (LRR), a NTF2
like middle domain and a C terminal UBA (ubiquitin associated fold) (Senay et al., 2003; Suyama et al., 2000). Mex67p heterodimerizes with Mtr2p with the help of it’s NTF2 like fold to form heterodimer complex (Suyama et al., 2000). In *P. falciparum* database PF14_0305 is annotated as the homologue of the nuclear mRNA export factor TAP. This homologue (PF14_0305) shows ~21% identity and significant similarity (1.4e-05) with the yeast export factor Mex67. The corresponding homologues of TAP in *P. vivax* (PVX_084925) and *P. berghei* (PB300366.00.0) are annotated as conserved hypothetical proteins. PF14_0305 and PVX_084925 have a very long N terminal region with insertions of arginine and lysine and are about twice the size of the other eukaryotic counterparts. The modular architecture of PF14_0305 (PfTAP homologue) is not clear as the NTF2 and UBA domains are not apparent in the bioinformatics analysis using the InterProScan sequence search software (Hunter et al., 2009). The absence of the modular architecture in PF14_0305 might be due to the sequence divergence during the course of evolution. It is important to note that the Nxt1/ P15 homologue is also absent in *P. falciparum* and other apicomplexans as revealed by the PSI-BLAST search. Nxt1 homologues are apparently present in fungi and other higher eukaryotes but they are absent in protozoans (Mans et al., 2004).

### 4.25 Remodeling ATPase Dbp5 and its substrate

Dbp5 is a yeast RNA helicase involved in remodelling RNPs extruding out of the NPC by removing Mex67 and other hnRNP like proteins from the exported mRNP (Lund and Guthrie 2005; Tran et al, 2007). In Dbp5 mutant, there is an increased association of Mex67 with the mRNAs, suggesting a direct role of this protein in RNP remodelling at the nuclear rim (Lund and Guthrie, 2005). Most of the physiological substrates of DBP5, like Mex67 and other mRNP bound proteins tend to accompany the mRNP into the cytoplasm through the
RESULTS

nucleoplasm but are not found in association with the polysomes (Anderson et al., 1993; Green et al., 2002; Windgassen et al., 2004). All of these results suggest that these accompanying proteins are removed upon entry into the cytoplasm from the NPC, just prior to translation. It has been shown that the bound proteins are displaced by the ADP bound form of Dbp5 (Tran et al., 2007). Recently, it has been shown that Nab2p is one the physiological substrate for Dbp5p. Moreover, Nab2p bound RNPs are targeted by Dbp5p at the cytoplasmic face of the nuclear pore, where the RNP bound proteins like Nab2 are removed from the mRNP by Dbp5p (Tran et al., 2007).

In *Plasmodium falciparum* genome the gene with PlasmoDB number PF14_0563 is the homologue of Dbp5/Rat8 and is annotated in PlasmoDB as a putative “DEAD-box RNA helicase” (Tuteja et al., 2006). We have therefore characterized the Dbp5 homologue from *P. falciparum* and the results are presented in this study.

Using the protein sequences of Nab2 (yeast) and the human homologue (Zinc finger CCCH domain-containing protein 14/NY-REN-37) as query. Nab2 homologue was detected in PlasmoDB. The gene with PlasmoDB number PFF1110c is *Plasmodium falciparum* Nab2 homologue and it shows ~46% homology with the yeast protein and ~63% homology (4.5e-17) with the human homologue NY-REN-37. It is important to note that this gene (PFF1110c) is unannotated and has been described as conserved hypothetical protein in the database. The orthologous proteins of *P. vivax* (PVX_114270), *P. yoelli* (PY03499) and *P. berghie* (PB000664.00.0) are also described as hypothetical proteins. Furthermore the results of the InterProScan (http://www.ebi.ac.uk) of the corresponding (PFF1110c, PfNab2) protein revealed the presence of a Zinc finger CCCH type motif and a conserved panther domain (PTHR14738) (Figure 33). The results of our analysis show that the Zinc finger CCCH type motif is only present in the orthologous proteins of *P. vivax, Homo sapiens, Caenorhabditis*
RESULTS

elegans and absent in the Saccharomyces cerevisiae homologue Nab2. Moreover P. falciparum and P. vivax homologues of Nab2 also contain a PWI domain at their N terminal (Figure 27A). PWI domain is a novel nucleic acid binding motif and the proteins bearing this motif are involved in processes like transcription, 5'-end capping, splicing, 3'-end processing (cleavage and polyadenylation), surveillance, turnover, transport, and translation (Szymczyna et al., 2003). Since PWI domain is absent in the human as well as the yeast orthologues, we propose here that PfNab2 might be playing roles other than ones played by Nab2 protein of yeast and it might be a key protein involved in the RNA metabolism. The structure modelling of the CCCH type zinc finger domain of P. falciparum was done using the zinc-finger domain of the template KIAA1064 protein (2cqA) using the software Swiss Model (http://swissmodel.expasy.org) (Figure 27 B).
Figure 33 Domain organization and structure modelling of Pf Nab2

(A) Schematic diagram showing the domain organization in the various homologues of Nab2 in 1: *Plasmodium falciparum*, 2: *Plasmodium vivax*, 3: *Toxoplasma gondii*, 4: *Saccharomyces cerevisiae* and 5: *Homo sapiens*. (B) Superimposition of the computer based structure modeling of the Zinc finger CCCH type motif of *Plasmodium falciparum* Nab2 (shown in gold) on the Zinc-finger domain in KIAA1064 protein (2cqe), which was used as a parent template (shown in spring green).
4.23 Hypothetical role of the components of mRNA export in *P. Falciparum*

A myriad of protein are associated with mRNA and play important role in the process of mRNA export. In yeast these mainly include mRNA binding proteins (Npl3p, Nab2p, Hrp1, Yra1p, and Mex67p), NPC or NPC-associated proteins (Mtr2p, Gle1p, Rip1p, Gle2p, Nup116p, Nup84 and Nup159 NPC sub-complexes) and Dbp5p, a DEAD box ATPase/RNA helicase. For most of these proteins *Plasmodium falciparum* orthologues were identified, which strongly suggests that the basic mechanism for mRNA export is highly conserved across various species. These components were identified mainly by the analysis of the genome using BLAST and profile-Hidden Markov Model searches. Most of the components of the mRNA export pathway are conserved in *Plasmodium falciparum* but there are a few peculiarities. In *Plasmodium falciparum* it was observed that except for Tho2 and Sub2/UAP56, other members of the THO complex are absent (Table 1). Moreover, homologues of Mtr2, Hrbl and Gle1 are too absent from the complex (Table 1). We hypothesize the following model for the mRNA export machinery in *Plasmodium falciparum*. Components of the PfTHO complex like PfTho2 (PFL2390c) and PfU52 (UAP56 of *Plasmodium falciparum*) bind to the mRNA cotranscriptionally. PfTho2 and PfU52 help in recruiting the RNA export factor Ref/ PFF0760w and other RNA associated proteins to the mRNA for the purpose of RNA processing and export. Ref/ PFF0760w in turn recruits the mRNA export factor Tap/ PF14_0305 to the RNA, which facilitates the passage of mRNA into the cytoplasm through the NPC (nuclear pore complex). On the cytoplasmic face of the NPC remodelling factor PfD66 (Dbp5) strips the proteins like PfNab2 off the mRNA, allowing the latter to be transferred to the translational machinery. This removal of proteins is a key step for imparting an overall directionality to the process of mRNA export (Tran et al., 2007;
RESULTS

Stewart, 2007). The removed proteins are then shuttled back into the nucleoplasm to continue a new cycle of mRNA export.

<table>
<thead>
<tr>
<th>Serial No.</th>
<th>Yeast name/ Accession No.</th>
<th>Human name/ Accession No.</th>
<th>P. falciparum name/PlasmoDB ID</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Tho2 or Rlr1/AAA93160</td>
<td>Tho2 or CXorf3/ NP_065182</td>
<td>Hypothetical protein/ PFL2390c</td>
<td>DNA recombination, mRNA processing, mRNA export, RNA elongation from RNA polymerase and component of the THO complex</td>
</tr>
<tr>
<td>2.</td>
<td>Hpr1 or Trf1/CAA88220</td>
<td>hHpr1 or Thoc1/ NP_005122</td>
<td>-</td>
<td>DNA recombination, mRNA processing, mRNA export, RNA elongation from RNA polymerase and component of the THO complex</td>
</tr>
<tr>
<td>3.</td>
<td>Mfl1/CAA86259</td>
<td>-</td>
<td>-</td>
<td>DNA recombination, mRNA export, RNA elongation from RNA polymerase</td>
</tr>
<tr>
<td>4.</td>
<td>Thp2/AAB66025</td>
<td>-</td>
<td>-</td>
<td>DNA recombination, mRNA export, RNA elongation from RNA polymerase</td>
</tr>
<tr>
<td>5.</td>
<td>Sub2/CAA98650</td>
<td>Bat1, Uap56/ BAF31287</td>
<td>Helicase, putative/ PFB0445c</td>
<td>mRNA export, RNA splicing and RNA 3’ end processing</td>
</tr>
<tr>
<td>6.</td>
<td>Yra1/AAC09951</td>
<td>Aly/Q86V81</td>
<td>Ref/ PFF0760w</td>
<td>RNA binding, mRNA export, Transcription export complex</td>
</tr>
<tr>
<td>7.</td>
<td>Mex67/CAA97875</td>
<td>Tap or Nxfl/ NP_006353</td>
<td>Tap/ PF14_0305</td>
<td>mRNA export, Large ribosomal subunit export</td>
</tr>
</tbody>
</table>

Table 1

List of components of the mRNA export in Plasmodium falciparum
### RESULTS

<table>
<thead>
<tr>
<th></th>
<th>Protein/ Description</th>
<th>Function</th>
<th>mRNA Process</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Mtr2/CAA82029 Nxt1 or P15</td>
<td>mRNA export, Large ribosomal subunit export</td>
<td>-</td>
<td>mRNA export, Large ribosomal subunit export</td>
</tr>
<tr>
<td>9</td>
<td>Npl3 or Nab1/CAA50291 Splicing factor, arginine/serine-rich 9/</td>
<td>pre-mRNA splicing factor/</td>
<td>mRNA export</td>
<td>mRNA export and nuclear mRNA splicing</td>
</tr>
<tr>
<td></td>
<td>NP_037380</td>
<td>NP_003760</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>GBP2/AAT92826 Heterogeneous nuclear Ribonucleoprotein in/ NP_112480</td>
<td>RNA binding protein/</td>
<td>Telomere maintenance mRNA export,</td>
<td>mRNA export and nuclear mRNA splicing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAI16413</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Hrb1 or Tom34/CAA95863 Polyadenylate-binding protein 4/ CAI16413</td>
<td>-</td>
<td>poly(A)+ mRNA export from nucleus</td>
<td>mRNA export and nuclear mRNA splicing</td>
</tr>
<tr>
<td>12</td>
<td>Nup42 or Rip1/CAA88706 Nucleoporin 98kDa/ EAX02562</td>
<td>Hypothetical protein/</td>
<td>protein export and import, export of all classes of RNA</td>
<td>mRNA export, rRNA export and nuclear pore organization</td>
</tr>
<tr>
<td>13</td>
<td>Nup 159/ AAC41652 Nucleoporin 214kDa/ EAW87960</td>
<td>Hypothetical protein/</td>
<td>mRNA export, rRNA export and nuclear pore organization</td>
<td>mRNA export, rRNA export and nuclear pore organization</td>
</tr>
<tr>
<td>14</td>
<td>Dbp5 or Rat8/CAA99237 DDX19A/ NP_060802</td>
<td>DEAD box Helicase/</td>
<td>mRNA export from the nucleus, translation termination</td>
<td>mRNA export, rRNA export and mRNA stability</td>
</tr>
<tr>
<td>15</td>
<td>Nab2/CAA96630 NY-REN-37/ NP_079100</td>
<td>Hypothetical protein/</td>
<td>mRNA polyadenylation, mRNA export and mRNA stability</td>
<td>mRNA export, Translation initiation and termination</td>
</tr>
<tr>
<td>16</td>
<td>Gle1 or Brr3/CAA98785 Gle1/ AAC25561</td>
<td>-</td>
<td>mRNA export, Translation initiation and termination</td>
<td>mRNA export, Translation initiation and termination</td>
</tr>
</tbody>
</table>