# Pfmrk, a MO15-related protein kinase from *Plasmodium falciparum*

## Gene cloning, sequence, stage-specific expression and chromosome localization

Ji-Liang LI[1,2], Kathryn J. H. ROBSON[3], Ji-Li CHEN[1], Geoffrey A. T. TARGETT[1] and David A. BAKER[1]

[1] Department of Medical Parasitology, London School of Hygiene and Tropical Medicine, London UK
[2] Department of Malaria Immunology, The First Medical University, Nanfang Hospital, Guangzhou, Peoples Republic of China
[3] MRC Molecular Haematology Unit, Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford UK

(Received 17 June 1996) – EJB 96 0885/1

Cyclin-dependent kinases (Cdks) play a central role in the regulation of the eukaryotic cell cycle. A novel gene encoding a Cdk-like protein, *Pfmrk*, has been isolated from the human malaria parasite *Plasmodium falciparum*. The gene has no introns and comprises an open reading frame encoding a protein of 324 amino acids with a predicted molecular mass of 38 kDa. Database searches revealed a striking similarity to the Cdk subfamily with the highest similarity to human MO15 (Cdk7). The overall sequence of Pfmrk shares 62% similarity and 46% identity with human MO15, in comparison to the 49–58% similarity and 34–43% identity with other human Cdks. Pfmrk contains two unique inserts: one consisting of 5 amino acids just before the cyclin-binding motif and the other composed of 13 amino acids within the T-loop equivalent region. Southern blots of genomic DNA digests and chromosomal separations showed that *Pfmrk* is a single-copy gene conserved between several parasite strains and is located on chromosome 10. A 2500-nucleotide transcript of this gene is expressed predominantly in the sexual blood stages (gametocytes), suggesting that Pfmrk may be involved in sexual stage development.

*Keywords:* *Plasmodium falciparum*; cyclin-dependent kinase; cell cycle; mRNA expression; chromosome localization.

*Plasmodium falciparum* is the causative agent of the most serious form of human malaria, which remains one of the most prevalent infectious diseases in the tropics and subtropics and affects over 100 million people, resulting in 1–2 million deaths/year. *P. falciparum* has a complicated life cycle, occurring extracellularly in the invertebrate and intracellularly in the vertebrate hosts where it undergoes four cycles of development. Sporozoites, injected by a mosquito bite, rapidly enter liver cells and then develop into exo-erythrocytic forms. After a few days they divide into thousands of merozoites that specifically invade red blood cells. The growth of the asexual erythrocytic stages progresses from ring to trophozoite to schizont, culminating in cell division followed by synchronous rupture of the infected erythrocytes to release free merozoites, causing the clinical symptoms and mortality of malaria. Merozoites invade erythrocytes again, continuing the asexual blood cycle. A small number of ring forms develop into male and female gametocytes, which are nondividing infective sexual stages and are morphologically and biochemically distinct from asexual parasites. After being taken up in a bloodmeal by the mosquito, the parasites emerge from the erythrocyte; the male gametocytes undergo exflagellation (which produces eight gametes from a single male gametocyte). This is followed by fertilization and sporogony (cell proliferation) resulting in the production of infective sporozoites within the mosquito salivary glands. The signals involved in triggering cell growth, proliferation and the molecular events controlling sexual differentiation, such as gametocytogenesis, and fertilization are unknown.

Protein kinases and phosphatases play a central role in regulation of eukaryotic cell division (Norbury and Nurse, 1992; Hunter, 1995). One subfamily of protein kinases is the cyclin-dependent kinases (Cdks) which contain a catalytic domain that requires the association of a regulatory cyclin for full acitivity. They are designated individually by number as Cdk1 to Cdk8 (Solomon, 1994; Morgan, 1995). The first identified Cdk, Cdc2 (Cdk1), is the best understood. In yeast two control points of the cell cycle, the transition through START, the point at which progression to DNA synthesis begins, and the initiation of mitosis, are regulated by Cdk1, the product of the *cdc2* gene in *Schizosaccharomyces pombe* and the *cdc28* gene in *Saccharomyces cerevisiae* (Reed et al., 1985; Simanis and Nurse, 1986). In higher eukaryotes Cdk1 is the active constituent of the maturation promoting factor, a complex required for induction of mitosis (Norbury and Nurse, 1992). Cdk/cyclin complexes have been shown to be required at different points of the eukaryotic cell cycle. Cdk activation is very important in ensuring that certain critical steps are completed before others begin, and a series

*Correspondence to* J.-L. Li, Immunology and Cell Biology Unit, Department of Medical Parasitology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
*Fax:* +44 171 636 8739.

*Abbreviations.* Cdk, cyclin-dependent kinase; Cak, cyclin-dependent kinase-activating kinase; Pfmrk, *Plasmodium falciparum* MO15-related kinase; PFGE, pulsed-field gel electrophoresis; Cki, cyclin-dependent kinase inhibitor.

*Enzymes.* Protein serine/threonine kinase (EC 2.7.1.37); protein phosphatase (EC 3.1.3.16); DNA polymerase (EC 2.7.7.7); restriction endonucleases *Acc*I, *Bam*HI, *Bcl*I, *Eco*RI, *Eco*RV and *Rsa*I (EC 3.1.21.4).

*Note.* The novel nucleotide sequence data reported here have been submitted to the EMBL/GenBank sequence data bank and are available under the accession number 73195.

of checkpoint controls operates to safeguard the fidelity of the system.

The life cycle of *P. falciparum* provides a number of defined points at which Cdks could act. The integral association between the control of the cell cycle and differentiation of the parasite has driven us to investigate the molecular components of these processes. The differences in the biology of the cell cycle between yeast, higher eukaryotes, and the parasite suggest that molecular mechanisms will have evolved to meet the specific requirements of each organism. Clearly, understanding the differences at the molecular level may help the development of novel intervention strategies against malaria. In this paper we report the molecular cloning of a gene encoding a Cdk-like protein from *P. falciparum*, related to the mammalian MO15 (Cdk7).

## MATERIALS AND METHODS

**Parasites.** *P. falciparum* clones 3D7A and T996 (Walliker et al., 1987) were cultivated at 37°C in RPMI-1640 medium (Gibco) supplemented with 37.5 mM Hepes buffer, 10% (by vol.) human serum ($A^+$) and filter-sterilized gas (96% $N_2$, 3% $CO_2$ and 1% $O_2$) in a semi-automated continous flow apparatus. Asexual stages of the parasite were collected at a parasitaemia of 5–10%; gametocytes were harvested after 14–18 days and purified by Percoll gradient centrifugation. The purified gametocytes were mainly of stages III–V.

**Nucleic acid isolation from parasites.** Parasites were recovered from infected erythrocytes by 0.1% saponin treatment, washed with phosphate-buffered saline (NaCl/$P_i$) and then resuspended in DNA lysis buffer (0.1 M NaCl, 10 mM Tris/HCl pH 7.4, 25 mM EDTA, 0.5% SDS, 120 μg/ml proteinase K). The lysate was then placed on a rocking table at room temperature for 2 h and then in a shaking waterbath apparatus at 37°C overnight. The DNA was extracted with phenol/chloroform and precipitated with 0.3 M sodium acetate and 2 vol. ethanol. Total RNA was isolated from asexual and sexual stage parasites by lysis with 4 M guanidium thiocyanate, followed by CsCl centrifugation and sodium acetate/ethanol precipitation (Chirgwin et al., 1979).

**Oligonucleotide design and synthesis.** The following oligonucleotides were synthesized and used for PCR and sequencing reactions. TK1 (5′ CATAGAGATTTAAGAGCAGCAAAT 3′, nucleotides 533–556) and TK2 (5′ TCCAA/TAACTCCATACATCACT 3′, nucleotides 749–769), used to isolate the *Pfmrk* gene, were based on the consensus sequence of subdomain VI (HRDLAARN) and subdomain IX (SDVWSF/YG) in the catalytic region of protein kinases (Hanks et al., 1988), respectively, and combined with *P. falciparum*-biased codon usage (Saul and Battistutta, 1988; Hyde et al., 1989). PK5 (5′ GTTACTTTATGGTATAGAGCACCAG 3′, nucleotides 689–713) and PK8 (5′ CCAAAATCAGCTAATTTAACTTCTCC 3′, nucleotides 575–690) were used in vectorette PCR and were obtained from Pharmacia. Vectorette I primer, Vectorette I nested primer and Vectorette I sequencing primer were obtained from Cambridge Research Biochemicals.

**Construction and screening of a genomic DNA library.** A λGEM-12 genomic library was constructed using a commercial kit and screened according to the manufacturer's instructions (Promega). Briefly, a *Sau*3A partial digest of 3D7A genomic DNA was end-filled with dATP and dGTP and ligated into the dCTP and dTTP filled-in *Xho*I λGEM-12 arms. After ligation, the DNA was packaged *in vitro* using Packagene extract and plated on *Escherichia coli* strains LE392 or KW251. Final titres of $2\times10^5$ and $4\times10^5$ were obtained for LE392 and KW251, respectively. The library was screened using KW251 according to standard procedures (Sambrook et al., 1989).

**Construction and screening of vectorette libraries.** This method comprises three basic steps: digestion of genomic DNA with an appropriate restriction enzyme, ligation of suitable synthetic oligonucleotide (termed vectorette) onto the digested DNA to construct a vectorette library and PCR using a specific primer and a universal primer directed toward the vectorette. A number of blunt-end vectorette libraries were constructed from 3D7A genomic DNA or a positive phage (gTK3) DNA according to the manufacturer's instructions (Cambridge Research Biochemicals). Briefly, 1 μg DNA was digested with *Hinc*II, *Rsa*I or *Acc*I restriction enzymes and approximately 200 ng of the digested DNA was ligated directly into a blunt-end vectorette. The *Acc*I-digested DNA was end-filled prior to ligation. Based on the known sequence and Southern blot information, a number of specific primers were synthesized and PCR was performed using a specific primer and the universal vectorette I primer. The resulting PCR fragments were sequenced directly or cloned into the pGEM-T vector (Promega) for double-stranded sequencing.

**Southern and northern blotting.** For Southern blotting, approximately 4 μg 3D7A genomic DNA was digested overnight with several restriction enzymes and then fractionated on a 1.0% agarose gel. For northern blotting, approximately 10 μg total RNA extracted from both asexual and sexual blood stages of *P. falciparum* 3D7A were separated on 1.0% agarose gels under denaturing conditions using formamide and formaldehyde in Mops buffer (Robson and Jennings, 1991). DNA gels were then denatured in 0.5 M NaOH, 1.5 M NaCl and neutralized in 1.5 M NaCl, 1 M Tris/HCl pH 8.0. The nucleic acids were transferred to Hybond $N^+$ nylon membrane (Amersham) in 20× NaCl/Cit (NaCl/Cit = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) as described by Southern (1975). The membranes were dried at room temperature and the nucleic acids covalently cross-linked (ultraviolet crosslinker, Stratagene) with ultraviolet light prior to hybridization.

**DNA and RNA hybridization.** Southern, northern and PFGE blots or plaque lift filters were hybridized in a solution containing 50% formamide, 6×NaCl/Cit, 5×Denhardt's solution, 0.5% SDS and 100mg/μl herring sperm DNA at 42°C for 4–6 h. Hybridization with a radiolabeled DNA probe (random hexamer primed DNA labeling kit, Boehringer Mannheim) was performed overnight under the same conditions. Filters were washed at 52–65°C once in NaCl/Cit/0.1% SDS for 30 min and twice in 0.5×NaCl/Cit/0.1% SDS for 40 min. If necessary, the filters were further washed in 0.2×NaCl/Cit/0.1% SDS and autoradiographed at −70°C.

**Polymerase chain reaction.** Reaction mixtures (50 μl total volume) contained 10 mM Tris/HCl pH 8.3, 2.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin, 200 μM each of dNTPs, 0.5 μM of each primer, 100–200 ng DNA as template and 2.5 U *Taq* DNA polymerase (Promega). Samples were overlaid with sterile mineral oil (Sigma) and subjected to 36 cycles of denaturation for 50 s (4 min for the first cycle) at 94°C, annealing for 1 min at 48–68°C and extension for 2 min at 72°C (10 min for the final cycle). PCR products were separated on agarose gels, purified using a Geneclean kit (Bio 101, Inc.) and cloned into a plasmid vector for sequence analysis.

**Sequence analysis.** PCR fragments were cloned into the pGEM-T vector (Promega). DNA sequencing was performed using a Sequenase version 2.0 kit (Amersham) on double-stranded or single-stranded plasmid DNA. Single-stranded DNA was generated by treatment with T7 gene 6 exonuclease (Amersham). Both strands were sequenced using a series of primers spanning the whole sequence. Sequence data derived from PCR were verified using independent fragments. Sequence data were

analyzed using the DNA Inspector IIe programme and the MRC Human Genome Mapping Project Resource Centre computor facilities (UK).

**Pulse-field gel electrophoresis.** Chromosomal DNA-agarose blocks were prepared as described by Wellems et al. (1987). Chromosomes were separated by pulse-field gel electrophoresis (PFGE) using a Bio-Rad CHEF DRII system. 1% agarose (Ultrapure, Gibco) gels were run in 45 mM Tris/borate, 1 mM EDTA pH 8.0 at 80 V for 168 h with a pulse time of 180 s ramped to 900 s. After electrophoresis, the gels were depurinated, denatured, neutralized and blotted onto Hybond $N^+$ nylon membrane. The membrane was fixed with ultraviolet light. Initial and final hybridization of the PFGE blot was performed as described above.

## RESULTS

**Cloning and sequencing strategies.** Protein kinases possess a catalytic domain of approximately 30 kDa within which some sequences are highly conserved among all members of the family (Hanks et al., 1988). Two primers, TK1 and TK2, were synthesized based on the conserved sequences of subdomains VI and IX and the *P. falciparum*-biased codon usage. A band of 236 bp was obtained using PCR from *P. falciparum* genomic DNA (3D7A; data not shown). The sequence data showed that the 236-bp fragment (TK1-2 clone) encoded 79 amino acids, with 59% identity and 74% similarity with Cdc2 of *S. pombe*. The TK1-2 clone possessed an A+T content of 76% and codon usage was typical of *P. falciparum*. Southern blot analysis confirmed that the TK1-2 was derived from *P. falciparum* DNA and revealed the presence of a number of related genes in the genome (data not shown). A λGEM-12 genomic library, therefore, was screened with the TK1-2 probe. After tertiary screening, six positive plaques were purified to homogeneity but only one (gTK3) contained the desired gene, as confirmed by PCR (data not shown). Despite exhaustive attempts, it was not possible to subclone restriction fragments of gTK3 DNA. Consequently, a partial restriction map of the gene was established and vectorette libraries were constructed and screened by PCR (Fig. 1). Two fragments (PK5-*Hin*cII and PK8-*Hin*cII) were amplified from genomic DNA and one fragment (PK5-*Acc*I) was derived from the gTK3 λ-phage DNA. PK8-*Hin*cII contained a putative ATG start codon and PK5-*Acc*I had a TGA stop codon. The sequence of the overlapping fragments showed that the ATG and the TGA were located in the same open reading frame (ORF). In order to confirm the sequence obtained from the overlapping clones, a pair of primers, $PK_S$ and $PK_E$, were used to amplify the complete coding region from both the gTK3 phage and genomic DNA. The two PCR products were the same size and both strands were sequenced directly.

**Nucleotide sequence analysis of the *Pfmrk* gene.** The sequence obtained from the overlapping clones of the λGEM-12 genomic library and vectorette PCR-amplified genomic DNA fragments consists of 1147 bp (Figs 1 and 2). The ORF of the *Pfmrk* gene starts with a putative ATG codon at nucleotide 134 and ends with a TGA stop codon at nucleotide 1108. It appears that the entire coding region of *Pfmrk* has been obtained for the following reasons. First, there are two in-frame stop codons located at nucleotide positions −24 and −45, respectively, upstream of the ATG initiation codon and three in-frame stop codons at positions +19, +28 and +34, respectively, downstream of the TGA stop codon. Second, around the putative start and stop codon and two unique insert sites of the deduced amino acids (see below), there are no intron-exon consensus boundary

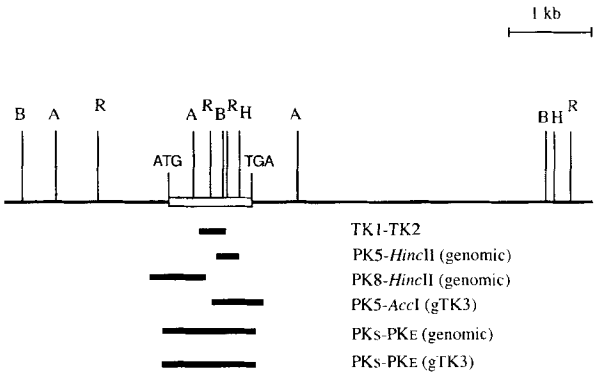


**Fig. 1. A schematic representation of the partial restriction map of the *Pfmrk* gene and the overlapping PCR fragments used to determine the nucleotide sequence of the *Pfmrk* gene.** A, *Acc*I; B, *Bcl*I; H, *Hin*cII; and R, *Rsa*I. The open box indicates the coding region of *Pfmrk*.

```
   1 ATATATCCAC AAAGAAATAT GTAAATGTAT TATATATGTG TTGCAATCCT TTTGGAAGAA
  60 AGATTAATCT GAAAGATTGC ATTTTCTTTT TAATTTCTTT TTATATATTT AATTAAAAAA
 121 TATTTATGTT CTCATGGAAA ATAATTCAAC AGAAAGATAT ATATTCAAGC CTAATTTTTT
                 M  E  N  N  S  T  E  R  Y  I  F  K  P  N  F  L   16
 181 AGGTGAGGGT TCCTATGGTA AAGTTTATAA GGCATATGAT ACAATTTTAA AAAAAGAAGT
      G  E  G  S  Y  G  K  V  Y  K  A  Y  D  T  I  L  K  K  E  V   36
 241 TGCAATAAAG AAGATGAAAT TAAATGAGAT AAGTAATTAT ATTGATGATT GTGGTATAAA
      A  I  K  K  M  K  L  N  E  I  S  N  Y  I  D  D  C  G  I  N   56
 301 TTTTGTTTTG TTAAGAGAAA TTAAAATAAT GAAGGAAATT AAACATAAAA ATATTATGAG
      F  V  L  L  R  E  I  K  I  M  K  E  I  K  H  K  N  I  M  S   76
 361 TGCTTTAGAT TTGTATTGTG AGAAAGATTA TATAAATTTA GTAATGGAAA TAATGGACTA
      A  L  D  L  Y  C  E  K  D  Y  I  N  L  V  M  E  I  M  D  Y   96
 421 TGATTTATCT AAGATAATAA ATCGAAAAAT TTTTCTAACA GATAGTCAAA AAAAGTGTAT
      D  L  S  K  I  I  N  R  K  I  F  L  T  D  S  Q  K  K  C  I  116
 481 ACTTTTACAA ATTTTAAATG GTCTAAATGT ATTACATAAA TATTATTTTA TGCATCGAGA
      L  L  Q  I  L  N  G  L  N  V  L  H  K  Y  Y  F  M  H  R  D  136
 541 TTTATCACCA GCAAATATAT TTATAAATAA AAAAGGAGAA GTTAAATTAG CTGATTTTGG
      L  S  P  A  N  I  F  I  N  K  K  G  E  V  K  L  A  D  F  G  156
 601 TTTATGTACA AAATATGGTT ATGATATGTA TTCAGATAAA CTATTTAGAG ATAAATATAA
      L  C  T  K  Y  G  Y  D  M  Y  S  D  K  L  F  R  D  K  Y  K  176
 661 AAAAAATTTA AATCTTACAA GTAAAGTTGT TACTTTATGG TATAGAGCAC CAGAATTATT
      K  N  L  N  L  T  S  K  V  V  T  L  W  Y  R  A  P  E  L  L  196
 721 ATTGGGAAGT AATAAATATA ATTCATCTAT TGATATGTGG AGTTTTGGTT GTATTTTTGC
      L  G  S  N  K  Y  N  S  S  I  D  M  W  S  F  G  C  I  F  A  216
 781 TGAACTATTA CTACAAAAAG CTCTCTTCCC AGGAGAAAAT GAAATTGATC AATTAGGAAA
      E  L  L  L  Q  K  A  L  F  P  G  E  N  E  I  D  Q  L  G  K  236
 841 AATATTTTTC CTTTTAGGTA CTCCTAACGA AAATAATTGG CCAGAAGCTC TTTGTCTTCC
      I  F  F  L  L  G  T  P  N  E  N  N  W  P  E  A  L  C  L  P  256
 901 CTTATATACA GAATTTACAA AAGCTACAAA AAAAGATTTT AAAACATATT TTAAAATAGA
      L  Y  T  E  F  T  K  A  T  K  K  D  F  K  T  Y  F  K  I  D  276
 961 TGATGATGAT TGTATTGATT TGTTAACGTC ATTTTTAAAA TTAAATGCTC ATGAACGTAT
      D  D  D  C  I  D  L  L  T  S  F  L  K  L  N  A  H  E  R  I  296
1021 CAGTGCAGAA GACGCTATGA AACACAGATA TTTTTTTAAT GATCCTTTGC CATGTGATAT
      S  A  E  D  A  M  K  H  R  Y  F  F  N  D  P  L  P  C  D  I  316
1081 ATCACAATTA CCTTTCAATG ATTTGTGAAT AAACATTGTT TTCGTATAAT ACAATTAAAA
      S  Q  L  P  F  N  D  L  *                                      324
1141 GTAGATC
```

**Fig. 2. Nucleotide and the deduced amino acid sequences of the *Pfmrk* gene (one-letter code).** The nucleotides and the amino acids are numbered on the left and right respectively. The amino acid sequences used to synthesis PCR primers for identification of this gene are underlined. The in-frame start and stop codons are bold-face type. A potential polyadenylation signal is underlined. The asterisk indicates the termination codon.

sequences. Furthermore, PCR amplification of the proposed coding region of *Pfmrk* from genomic DNA and cDNA generated fragments of the same length (data not shown). Third, the coding region (975 bp) has an A+T content of 75.6%, while the 5′ upstream (133 bp) and 3′ downstream (39 bp) non-coding regions possess A+T contents of 80.5% and 79.5%, respectively, characteristically higher than that of the coding region. A final line of evidence is that the length of the ORF and the amino acid composition of the predicted protein are completely compa-

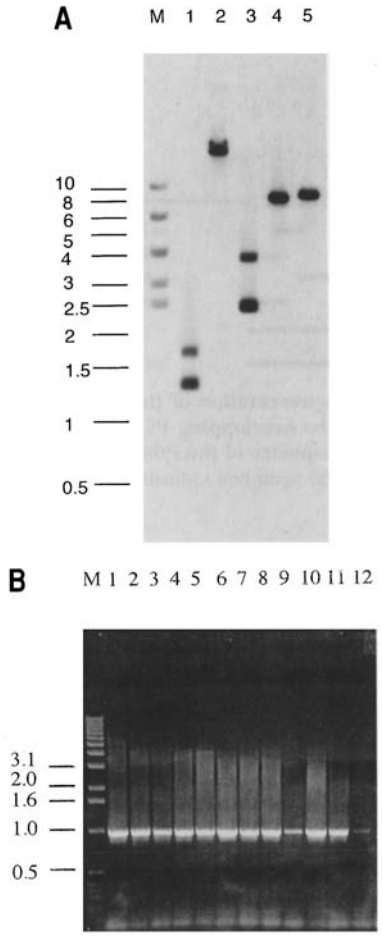


**Fig. 3. Southern blot and PCR analysis of the *Pfmrk* gene.** (A) Southern blot analysis. 4 µg genomic DNA from 3D7A *P. falcipaurm* was digested with restriction enzymes, electrophoresed on a 1.0% agarose gel, transferred onto nylon membrane, and probed with a PCR fragment corresponding to the complete coding region of *Pfmrk*. Lanes 1–5 correspond to digests with *Acc*I, *Bam*HI, *Bcl*I, *Eco*RI and *Eco*RV. The sizes of the DNA markers are given in kilobase pairs on the left side. (B) Agarose gel analysis of PCR products corresponding to the complete coding region of *Pfmrk* amplified from genomic DNA of *P. falciparum*. Lanes 1–12 contain DNA from NF54, K1, T996, T994R, FC27, Honduras, H1, 7G8, V1S, T994, 3D7A and an infected patient blood sample from The Gambia, respectively. The DNA markers are a 1-kb DNA ladder (BRL).

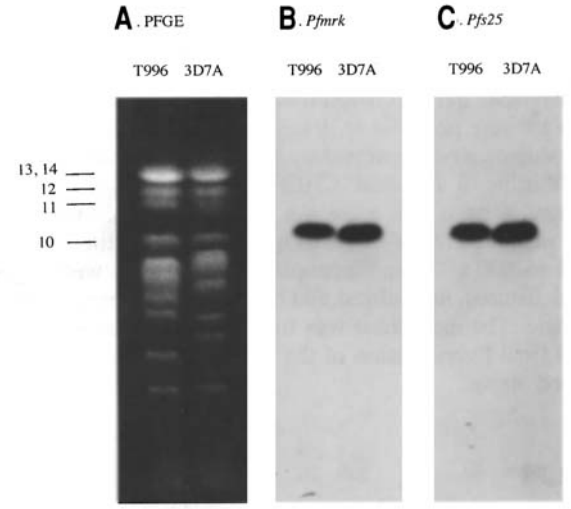


**Fig. 4. Chromosome localization of the *Pfmrk* gene.** Parasite chromosomes from *P. falciparum* 3D7A and T996 were separated by pulsed-field gel electrophoresis, stained with ethidium bromide, blotted onto nylon membrane and probed with a PCR fragment corresponding to the coding region of *Pfmrk* or *Pfs25*. According to the yeast chromosome markers and hyridization of several *P. falciparum* chromosome marker genes, chromosome 10–14 were identified on the PFGE blot (A). Chromosome 10 was hybridized with *Pfmrk* (B) and *Pfs25* (C), respectively.

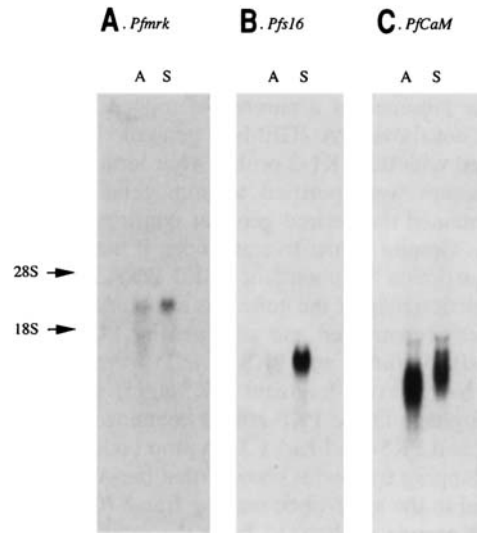


**Fig. 5. Northern blot analysis of the *Pfmrk* gene.** 10 µg total RNA extracted from asexual and sexual erythrocytic stages of *P. falciparum* 3D7A were fractionated in a denaturing formaldehyde gel, blotted onto nylon membrane and hybridized to radioactive probes. A, asexual erythrocytic stages; S, sexual blood stages. The positions of *P. falciparum* rRNA subunits (18S and 28S) are indicated. (A), (B) and (C) are autoradiographs of the blotted nylon membrane probed with the *Pfmrk* gene, the *Pfs16* sexual stage-specific gene and the calmodulin housekeeping gene (*PfCaM*) respectively. *Pfs16* hybridized with a band of approximately 1400 nucleotides in the sexual blood stages and *PfCaM* hybridized with two bands of 1300 and 1000 from the parasite.

rable to the homologues of other species (see below). Sequence and codon usage in the coding region is typical of *P. falciparum* genes. The sense strand has a higher content of A (40.1%) than T (35.5%); 71% (230/324) of the codons is *P. falciparum*-preferred and the frequency distributions are markedly biased, 87% of the codons possess an A or T at the third position, G is predominant in the first position (52.7%) and T is predominant in the last position (42.9%). No *P. falciparum* or protozoan consensus translation initiation sequences (Saul and Battistutta, 1990; Yamauchi, 1991) occur in the *Pfmrk* gene.

**Properties of the predicted Pfmrk protein.** The ORF encodes a protein of 324 amino acids with a predicted molecular mass of 38 kDa. Pfmrk contains all 11 conserved subdomains of the protein kinase family and has almost all of the characteristic features of a kinase (Knighton et al., 1991; DeBondt et al., 1993). These include (relative to the residue numbers for the catalytic subunit of bovine cAPK-$\alpha$, refer to Hanks et al., 1988): (a) the glycine loop Gly50-Xaa-Gly52-Xaa-Xaa-Gly55, forming part of the ATP-binding site; (b) the catalytic loop Arg165-Asp166-Xaa-Xaa-Xaa-Xaa-Asn171, involved in catalysis and in guiding the peptide substrate into the proper orientation so that catalysis can occur; (c) the triad composed of the side chain of Lys72, Asp184 and Glu91, which is close to the $\gamma$-phosphate of ATP and plays a key role in recognition of the phosphate of

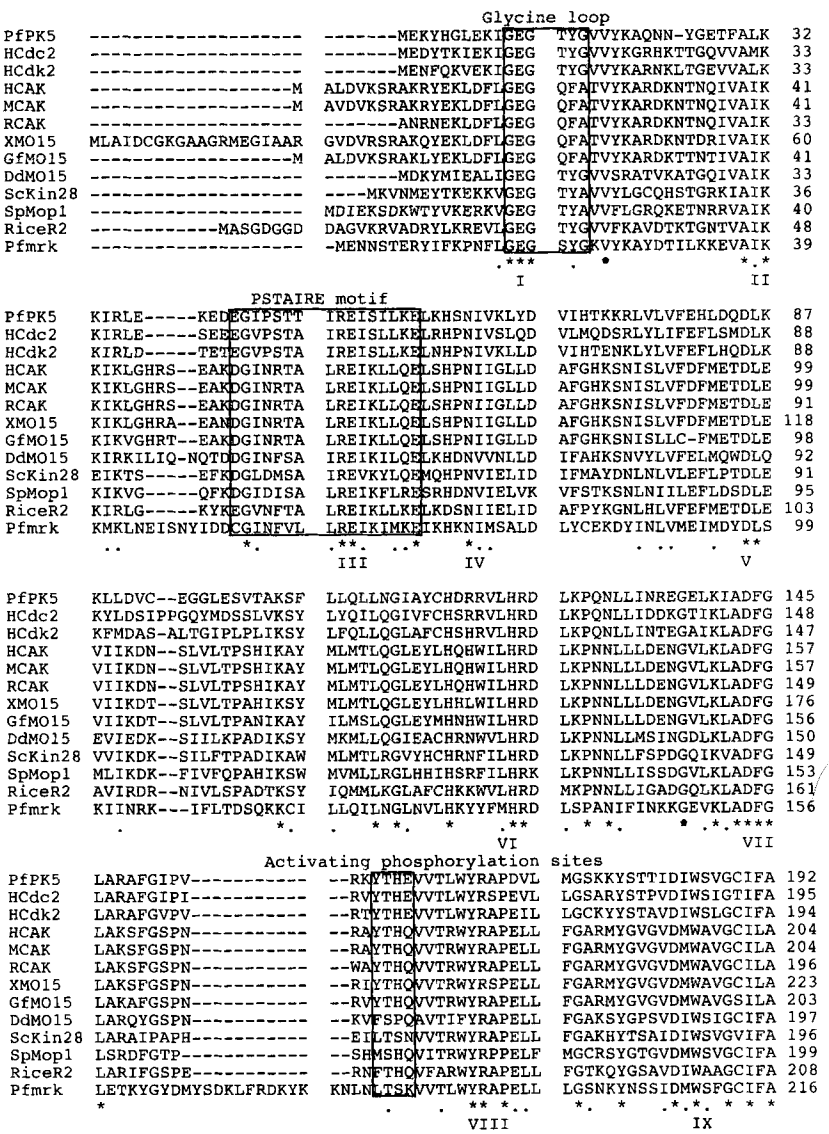


**Fig. 6. Alignment of the predicted amino acid sequences of Pfmrk with PfPK5, human Cdc2 and Cdk2 and the MO15 homologues of different species.** The EMBL/GenBank Database accession numbers are as follows: *P. falciparum* PfPK5, X61921; human HCdc2, X05360; human HCdk2, X61622; HCAK (HCdk7), X79193; mouse MCAK, X74145; rat RCAK, S51085; *Xenopus levis* XMO15, X53962; goldfish GfMO15, D38631; slime mold *D. discoideum* DdMO15, S79590; *S.cerevisiae* ScKin28, X04423; *S. pombe* SpMop1, L47353; rice R2, X58194. The sequences were aligned with the CLUSTAL W (1.60) multiple sequence alignment programme. Identical amino acids are indicated by an asterisk and conservative changes by a dot. The glycine loop, PSTAIRE motif, activating phosphorylation sites and GDSEIDQ motif are boxed. The conserved kinase domains referred to in the text are marked I–XI.

ATP; (d) Asp184, Asn171 and Asp166, which are also identified as a sequence motif implicated in ATP binding; (e) Asp220, Glu208 and Arg280, involved in the stabilization of protein kinases; (f) Ala206 and Glu208, which are diagnostic of the catalytic domain of protein kinases. Pfmrk also contains almost invariant amino acids corresponding to Phe185, Gly186, Trp222 and Gly225 whose functions have not been defined. The sequences (Asp-Leu-Ser-Pro-Ala-Asn) in subdomain VI and (Val-Thr-Leu-Trp-Tyr-Arg-Ala-Pro-Glu) in subdomain VIII indicate that Pfmrk is a serine/threonine kinase rather than a tyrosine kinase.

***Pfmrk* is a single-copy gene conserved in different strains.** In order to determine the gene copy number of *Pfmrk* in *P. falciparum*, 3D7A genomic DNA was digested with different restriction enzymes and analysed by Southern blotting, using the PCR fragment corresponding to the complete coding region as a probe. The results (Fig. 3A) are consistent with the restriction map (Fig. 1), indicating that Pfmrk is encoded by a single-copy gene in the parasite genome. PCR analysis of different parasite strains suggests that this gene is highly conserved in *P. falciparum* (Fig. 3B).

**Chromosome localization.** The chromosomes of clones 3D7A and T996 of *P. falciparum* were separated by pulsed-field gel electrophoresis (Fig. 4A), blotted onto nylon membrane and hybridized with the *Pfmrk* gene probe. A single band was observed corresponding to chromosome 10 (Fig. 4B). This was confirmed by re-probing the same blot with the *Pfs25* gene (Fig. 4C), which has been assigned to chromosome 10. The single band obtained on the PFGE blot also supports the conclusion that *Pfmrk* is a single-copy gene per haploid genome.

**mRNA size and expression.** Northern blots of total RNA prepared from cultures enriched in stages III–V gametocytes and

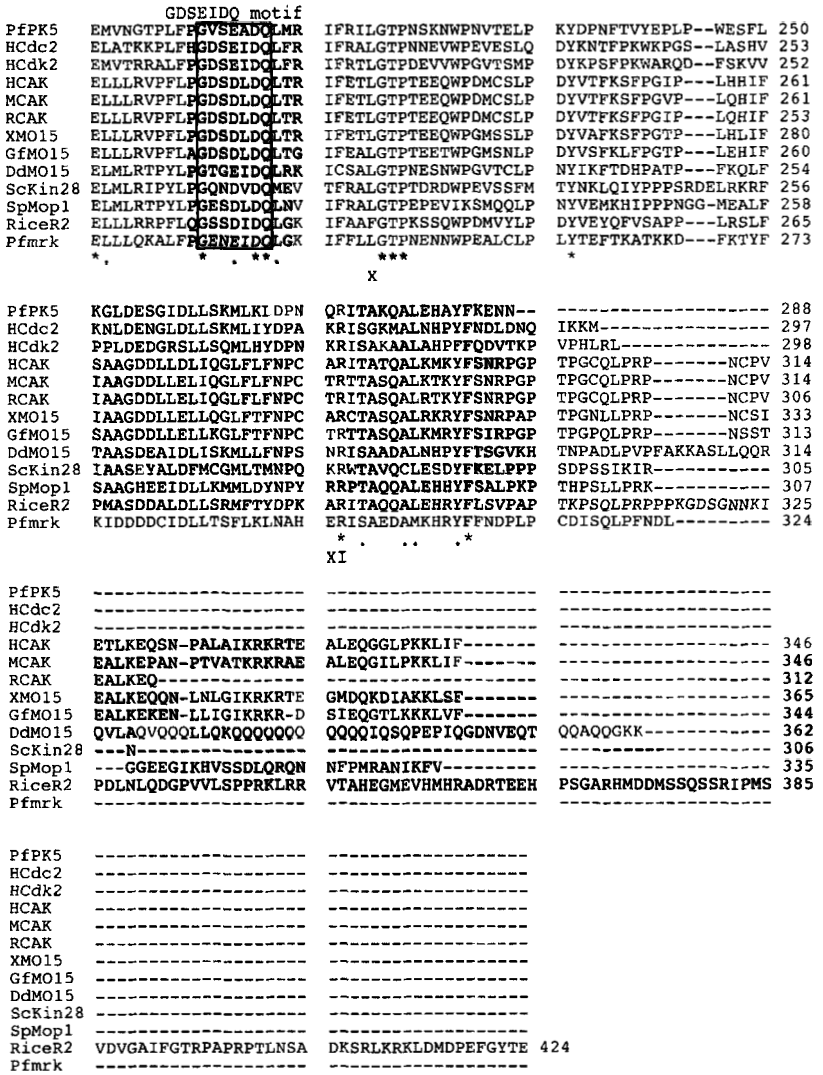


**Fig. 6.** Continuation

**Table 1. Comparison of Pfmrk with human Cdks, PfPK5 and other MO15 homologues.**

| Cdk | Identity | Similarity | MO15/Cdk7 | Identity | Similarity |
|---|---|---|---|---|---|
| | % | | | % | |
| HCdc2 | 40.6 | 57.7 | HCAK | 46.2 | 61.8 |
| HCdk2 | 41.4 | 58.0 | MCAK | 43.4 | 58.9 |
| HCdk3 | 41.0 | 48.7 | RCAK | 44.5 | 60.5 |
| HCdk4 | 38.6 | 54.6 | XMO15 | 41.7 | 57.6 |
| HCdk5 | 42.7 | 57.6 | DdMO15 | 42.5 | 59.4 |
| HCdk6 | 37.6 | 57.8 | GfMO15 | 43.9 | 58.7 |
| HCdk7 | 46.2 | 61.8 | RiceR2 | 44.3 | 62.6 |
| HCdk8 | 34.4 | 51.9 | ScKin28 | 35.9 | 57.1 |
| PfPK5 | 41.8 | 56.3 | SpMop1 | 36.5 | 58.2 |

from mixed asexual erythrocytic stages were probed with a fragment corresponding to the complete coding region (Fig. 5A). A band of approximately 2500 nucleotides in size was detected, migrating between the 28S and 18S ribosomal RNA species, in both stages. The result suggests that the mature *Pfmrk* transcript contains approximately 1300 nucleotides of 5′ and/or 3′ untranslated sequences. The mRNA transcript was much more abundant in the sexual stage than in the asexual erythrocytic stages. Probes corresponding to *Pfs16*, a sexual stage-specific gene (Bruce et al., 1990), and calmodulin, a house-keeping gene (Robson and Jennings, 1991), were used as controls (Fig. 5, B and C).

**Comparison of the amino acid sequence of Pfmrk with MO15 and other Cdks.** Database (Trembl and Nbrf) analysis revealed that Pfmrk has a striking similarity to members of the Cdk subfamily, with the highest similarity to human MO15 protein kinase and a rice Cdc2-like kinase (Fig. 6 and Table 1). The amino acid sequence of Pfmrk shares 62% similarity and 46% identity with the human MO15 (Cdk7) kinase, in comparison to the 49−58% similarity and 34−43% identity with other human Cdks (Cdk1−6 and Cdk8), and 56% similarity and 42% identity with the PfPK5, another Cdk-related protein kinase of *P. falciparum* (Ross-MacDonald et al., 1994). Sequence analysis showed that Pfmrk contains all four major regions that are specific features of the Cdk subfamily in addition to the catalytic domains found in all serine/threonine kinases. These include an ATP-binding domain, cyclin-binding domain, activating phosphorylation sites and the GDSEIDQ motif which is thought to be involved in the control of phosphorylation of the activating

sites and the binding of cyclins. However, sequence comparison of Pfmrk with human Cdks in these domains revealed that Pfmrk is much more similar to MO15 than to the other Cdks. Pfmrk contains residues equivalent to Ser14 and Tyr15 in the ATP-binding domain, more like the rice MO15 homologue and the yeast kinases Kin28 and Mop1. The potential activating phosphorylation sites of Pfmrk are quite different to human MO15, but similar to those of Kin28. The fact that Pfmrk only has Glu208, Ile209, Asp210, Phe240 and Lys242 but lacks Asp206, Phe213, Lys237, Ser239, Pro241 and Trp243 residues (relative to the human Cdk2), which is thought to be involved in binding of inhibitory CksHs1 subunit by human Cdk2 (Bourne et al., 1996), implies further that Pfmrk is not closely related to Cdc2 and Cdk2. Despite the interesting similarities to known Cdks, Pfmrk has two unique inserts: a small insert, consisting of five amino acids compared to human Cdc2, Cdk2 and MO15 homologues of rice and yeast, that appears just before the cyclin-binding motif, and a larger insert, composed of 13 amino acids, located within the T-loop domain correponding to residues 152–170 of human Cdk2.

## DISCUSSION

The number of identified Cdks in mammalian cells and in budding yeast is rapidly expanding (Meyerson et al., 1992; Morgan, 1995). In *P. falciparum*, only two genes related to that encoding Cdk, *PfPK5* and *Pfcrk1*, have been isolated and neither of these has a defined function (Ross-MacDonald et al., 1995; Doerig et al., 1995). *PfPK5* is highly similar to that encoding Cdc2 from human and yeast but, unlike the human gene, was unable to complement yeast *cdc2* in a functional assay. However, this feature is common to all protozoan *cdc2* homologues described to date (Mottram et al., 1993). *Pfcrk1* encodes a protein related to $p58^{GTA}$, a negative regulator of cell growth (Bunnell et al., 1990). In this study, we have identified a novel *cdk*-like gene from *P. falciparum*, which is conserved in different parasite strains and located on chromosome 10. The gene has no introns since no obvious intron-exon consensus boundaries were observed in the coding region and nucleotide sequences within the ORF were not A+T-rich (Weber, 1988). This was confirmed by PCR amplification of the coding region from cDNA. Database searches revealed that the encoded protein is more similar to MO15 than to other Cdks. Consequently, we designated it as Pfmrk, *P. falciparum* MO15-related kinase. The gene for MO15 has been cloned from a variety of species including *Xenopus* (Shuttleworth et al., 1990), human (Darbon et al., 1994; Levedakou et al., 1994; Tassan et al., 1994; Wu et al., 1994), mouse (Stepanova et al., 1994), rat, goldfish (Onoe et al., 1993), fission and budding yeast (Simon et al., 1986; Damagnez et al., 1995; Buck et al., 1995), slime mold *Dictyostelium discoideum* (Michaelis et al., 1995) and possibly rice (Hata, 1991).

A comparison between conserved domains of Pfmrk and human Cdks may not help in defining a function for Pfmrk but it may give some clues as to its likely control mechanisms. Cdk activity is tightly controlled by four highly conserved biochemical mechanisms. The cyclin subunit is a primary regulator of Cdk acitivity. Cyclin binds to one side of the catalytic cleft, interacting with the N-terminal and C-terminal lobes of Cdk to form a large, continuous protein-protein interface (DeBondt et al., 1993, Jeffrey et al., 1995). The $\alpha$1 helix composed of 16 amino acids, called the PSTAIRE motif of Cdk, is central to the interface. In Pfmrk, there is a small insert consisting of five amino acids (compared to Cdc2 and Cdk2) just before the PSTAIRE motif (see Fig. 6). In the human Cdk2, the residues of Glu42, Gly43, Val44, Ile49, Ile52, Leu54, Lys56 and Glu57 are involved in cyclin A binding (Jeffrey et al., 1995). However, Pfmrk, more like mammalian MO15 (see Fig. 6), contains only Gly43, Ile52, Lys56 and Glu57, implying that the potential cyclin of Pfmrk, if it exists, is different from cyclin A. In fact, the partner of Cdk7 is cyclin H (Fisher and Morgan, 1994; Makela et al., 1994). In additon to cyclin binding, complete Cdk activation depends upon phosphorylation at Thr160 (human Cdk2) or its equivalent residue in the T-loop. The kinase responsible for the phosphorylation is the Cdk-activating kinase, Cak (Solomon, 1994). It has been demonstrated that MO15/Cdk7 can activate Cdk1, Cdk2, Cdk3, Cdk4 and Cdk6 *in vitro* and/or *in vivo* by phosphorylating this threonine residue (Fesquet et al., 1993; Solomon et al., 1993; Poon et al., 1993, 1994; Fisher and Morgan, 1994; Labbe et al., 1994; Kato et al., 1994; Matsuoka et al., 1994; Aprelikova et al., 1995; Desai et al., 1995). In human Cdk2, Thr160 lies in the T-loop that blocks the catalytic cleft between the N- and C-terminal lobes (DeBondt et al., 1993; Jeffrey et al., 1995). Interestingly, in Pfmrk there is a hydrophilic insert composed of 13 amino acids following the fifth residue in the equivalent T-loop position (see Fig. 6). It is not clear how this insert would influence the phosphorylation of Thr182 (equivalent to Thr160 in Cdk2 and Thr170 in Cdk7) by Cak. Understanding the precise structure and function of the insert will be dependent upon crystallographic analysis of Pfmrk. Phosphorylation of Thr14 and Tyr15 in the glycine loop is particularly important in the control of Cdk activity. Like yeast Kin28, Mop1, and the rice and slime mold MO15 homologues, Pfmrk contains these residues (Thr14 replaced by Ser14) but the corresponding sites are not conserved in the MO15 of goldfish, *Xenopus*, human, mouse and rat, suggesting that these sites may not be important for the regulation of Cdk7 activity. The fourth major mechanism for Cdk regulation involves a diverse family of proteins, termed the Cki (cyclin-dependent kinase inhibitors), which bind and inactivate Cdk-cyclin complexes. Recently, crystallographic studies showed that CksHs1, a human Cki, bound via all four $\beta$ strands to the Cdk2 C-terminal lobe (Bourne et al., 1996). The GDSEID motif at the beginning of the $\alpha$5 helix and the (S/T)FPXW motif in the L14 segment of Cdk are implicated in the Cks binding. Pfmrk possesses the GDSEID motif but not the (S/T)FPXW motif, implying that the inhibitor of Pfmrk, if presented in the parasite, would be different to human Cki. Indeed, a few proteins have been observed to associate with the Cdk7-cyclin H complex (Tassan et al., 1994; Yee et al., 1995; Adamczewski et al., 1996).

In addition to Cdk activation, Cak may have other functions. MO15 and cyclin H are components of the transcription factor TFIIH and phosphorylate the C-terminal repeat domain (CTD) of the largest subunit of RNA polymerase II in yeast (Cismowski et al., 1995; Buck et al., 1995; Damagnez et al., 1995) and mammals (Feaver et al., 1994; Serizawa et al., 1995). Cak may be involved in the control of transcription initiation and nucleotide excision repair (Roy et al., 1994; Shiekhattar et al., 1995). It has been described that MO15 is associated with several other components of TFIIH (Shiekhattar et al., 1995; Yee et al., 1995). However, so far there is no information on the constitution of the transcription factors in malaria parasites. In this case, Pfmrk may be useful in searching the constituent molecules of the equivalent transcription factors such as TFIIH in *P. falciparum*.

Expression of *Pfmrk* mRNA, though detectable in the asexual erythrocytic stages, occurs predominantly in the sexual stage as showed by northern blotting. Whether Pfmrk is involved in the molecular events controlling sexual development such as gametocyte maturation, exflagellation and fertilization remains to be investigated.

We are very grateful to Drs J. M. Kelly, B. S. Hall for useful comments and discussions, to Drs M. R. Goodier, P. M. Kaye for critical

reading of the manuscript and to Mr O. Daramola for help with parasite cultivation. This work was supported in part by the Wellcome Trust. J.-L. Li was supported by the Sino-British Friendship Scholarship Scheme (SBFSS), the Great Britain-China Educational Trust and the Henry Lester Trust Limited.

## REFERENCES

Adamczewski, J. P., Rossignol, M., Tassan, J.-P., Nigg, E. A., Moncollin, V. & Egly, J.-M. (1996) MAT1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH, *EMBO J. 15*, 1877–1884.

Aprelikova, O., Xiong, Y. & Liu, E. T. (1995) Both p16 and p21 families of cyclin-dependent kinase (CDK) inhibitors block the phosphorylation of cyclin-dependent kinases by the CDK-activating kinase, *J. Biol. Chem. 270*, 18195–18197.

Bourne, Y., Watson, M. H., Hickey, M. J., Homes, W., Rocque, W., Reed, S. I. & Tainer, J. A. (1996) Crystal structure and mutational analysis of the human CDK2 kinase complex with cell cycle-regulatory protein CksHs1, *Cell 84*, 863–874.

Bruce, M. C., Baker, D. A., Alano, P., Rogers, N. C., Graves, P. M., Targett, G. A. T. & Carter, R. (1990) Sequence coding for a sexual stage specific protein of *Plasmodium falciparum*, *Nucleic Acids Res. 18*, 3637.

Buck, V., Russell, P. & Millar, J. B. (1995) Identification of a cdk-activating kinase in fission yeast, *EMBO J. 14*, 6173–6183.

Bunnell, B. A., Heath, L. S., Adams, D. E., Lahti, J. M. & Kidd, V. J. (1990) Increased expression of a 58-kDa protein kinase leads to changes in the CHO cell cycle, *Proc. Natl Acad. Sci. USA 87*, 7467–7471.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease, *Biochemistry 18*, 5294–6299.

Cismowski, M. J., Laff, G. M., Solomon, M. J. & Reed, R. I. (1995) KIN28 encodes a C-terminal domain kinase that controls mRNA transcription in *Saccharomyces cerevisiae* but lacks cyclin-dependent kinase-activating kinase (CAK) activity, *Mol. Cell. Biol. 15*, 2983–2992.

Damagnez, V., Makela, T. P. & Cottarel, G. (1995) *Schizosaccharomyces pombe* Mop1-Mcs2 is related to mammalian CAK, *EMBO J. 14*, 6164–6172.

Darbon, J.-M., Devault, A., Taviault, S., Fesquet, D., Martinez, A.-M., Galas, S., Cavadore, J.-C., Doree, M. & Blanchard, J.-M. (1994) Cloning, expression and subcellular localization of the human homolog of the $p40^{MO15}$ catalytic subunit of cdk activating kinase, *Oncogene 9*, 3127–3138.

DeBondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O. & Kim, S.-H. (1993) Crystal structure of cyclin-dependent kinases 2, *Nature 363*, 595–602.

Desai, D., Wessling, H. C., Fisher, R. P. & Morgan, D. O. (1995) The effect of phosphorylation by CAK on cyclin binding by CDC2 and CDK2, *Mol. Cell. Biol. 15*, 345–350.

Doerig, C., Doerig, C., Horrocks, P., Coyle, J., Carlton, J., Sultan, A., Arnot, D. & Carter, R. (1995) Pfcrk-1, a developmentally regulated cdc2-related protein kinase of *Plasmodium falciparum*, *Mol. Biochem. Parasitol. 70*, 167–174.

Feaver, W. J., Svejstrup, J. Q., Henry, N. L. & Kornberg, R. D. (1994) Relationship of cdk-activating kinase and RNA-polymerase II CTD kinase TFIIH/TFIIK, *Cell 79*, 1103–1109.

Fesquet, D., Labbe, J.-C., Derancourt, J., Capony, J.-P., Galas, S., Girard, F., Lorca, T., Shuttleworth, J., Doree, M., & Cavadore, J. C. (1993) The MO15 gene encodes the catalytic subunit of a protein kinase that activates cdc2 and other cyclin-dependent kinases (CDKs) through phosphorylation of Thr161 and its homologues, *EMBO J. 12*, 3111–3121.

Fisher, R. P. & Morgan, D. O. (1994) A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase, *Cell 78*, 713–724.

Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains, *Science 241*, 42–52.

Hata, S. (1991) cDNA cloning of novel cdc2+/CDC28-related protein kinase from rice, *FEBS Lett. 279*, 149–152.

Hunter, T. (1995) Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling, *Cell 80*, 225–236.

Hyde, J. E., Kelly, S. L., Holloway, S. P., Snewin, V. A. & Sims, P. F. G. (1989) A general approach to isolating *Plasmodium falciparum* genes using non-redundant oligonucleotides inferred from protein sequences of other organisms, *Mol. Biochem. Parasitol. 32*, 247–262.

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. & Pavletich, N. P. (1995) Mechanism of CDK activation revealed by the structure of a cyclin A-CDK2 complex, *Nature 376*, 313–320.

Kato, J.-Y., Matsuoka, M., Strom, D. K. & Sherr, C. J. (1994) Regulation of cyclin D-dependent kinase 4 (cdk4) by cdk4-acitvating kinase, *Mol. Cell. Biol. 14*, 2713–2721.

Knighton, D. R., Zheng, J., TenEyck, L. F., Ashford, V. A., Xuong, N.-H., Taylor, S. S. & Sowadski, J. M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase, *Science 253*, 407–414.

Labbe, J.-C., Martinez, A.-M., Fesquet, D., Capony, J.-P., Darbon, J.-M., Derancourt, J., Devault, A., Morin, N., Cavadore, J.-C. & Doree, M. (1994) $p40^{MO15}$ associates with a p36 subunit and requires both nuclear translocation and Thr176 phosphorylation to generate cdk-activating kinase activity in *Xenopus* oocytes, *EMBO J. 13*, 5155–5164.

Levedakou, E. N., He, M., Baptist, E. W., Craven, R. J., Cance, W. G., Welcsh, P. L., Simmons, A., Naylor, S. L., Leach, R. J., Lewis, T. B., Bowcock, A. & Liu, E. T. (1994) Two novel human serine/threonine kinases with homologies to the cell cycle regulating *Xenopus* MO15 and NIMA kinases: cloning and characterization of their expression pattern, *Oncogene 9*, 1977–1988.

Makela, T. P., Tassan, J. P., Nigg, E. A., Frutiger, S., Hughes, G. J. & Weinberg, R. A. (1994) A cyclin associated with the CDK-activating kinase MO15, *Nature 371*, 254–257.

Matsuoka, M., Kato, J., Fisher, R. P., Morgan, D. O. & Sherr, C. J. (1994) Activation of cyclin-dependent kinase 4 (cdk4) by mouse MO15-associated kinase, *Mol. Cell. Biol. 14*, 7265–7275.

Meyerson, M., Enders, G. H., Wu, C. L., Su, L. K., Gorka, C., Nelson, C., Harlow, E. & Tsai, L. H. (1992) A family of human cdc2-related protein kinases, *EMBO J. 11*, 2909–2917.

Michaelis, C., Luo, Q. & Week, G. (1995) A *Dictyostelium discoideum* gene, which is highly related to mo15 from *Xenopus*, is expressed during growth but not during development, *Biochem. Cell Biol. 73*, 51–58.

Morgan, D. O. (1995) Principles of CDK regulation, *Nature 374*, 131–134.

Mottram, J. C., Kinnaird, J. H., Shiels, G. R., Tait, A. & Barry, J. D. (1993) A novel CDC2-related protein kinase from *Leishmania mexicana*, LmmCRK1, is post-translationally regulated during the life cycle, *J. Biol. Chem. 268*, 21044–21052.

Norbury, C. & Nurse, P. (1992) Animal cell cycles and their control, *Annu. Rev. Biochem. 61*, 441–470.

Onoe, S., Yamashita, M., Kajiura, H., Katsu, Y., Jianquao, J. & Nagahama, Y.(1993) A fish homolog of the cdc2-related protein $p40^{MO15}$: its cDNA cloning and expression in oocytes, *Biomed. Res. 14*, 441–444.

Poon, R. Y. C., Yamashita, K., Adamczewski, J. P., Hunt, T. & Shuttleworth, J. (1993) The cdc2-related protein $p40^{MO15}$ is the catalytic subunit of a protein kinase that can activate $p33^{cdk2}$ and $p34^{cdc2}$, *EMBO J. 12*, 3123–3132.

Poon, R. Y. C., Yamashita, K., Howell, M., Ershler, M. A., Belyavsky, A. & Hunt, T. (1994) Cell cycle regulation of the $p34^{cdc2}/p33^{cdk2}$-activating kinase $p40^{MO15}$, *J. Cell Sci. 107*, 2789–2799.

Reed, S. I., Hadwiger, J. A. & Lorincz, A. T. (1985) Protein kinase activity associated with the product of the yeast cell division cycle gene *cdc28*, *Proc. Natl Acad. Sci. USA 82*, 4055–4059.

Robson, K. J. H. & Jennings, M. W. (1991) The structure of the calmodulin gene of *Plasmodium falciparum*, *Mol. Biochem. Parasitol. 46*, 19–34.

Ross-MacDonald, P. B., Graeser, R., Kappes, B., Franklin, R. & Williamson, D. H. (1994) Isolation and expression of a gene specifying a cdc2-like protein kinase from the human malaria parasite *Plasmodium falciparum*, *Eur. J. Biochem. 220*, 693–701.

Roy, R., Adamczewski, J. P., Seroz, T., Vermuelen, W., Tassan, J. P., Schaeffer, L., Nigg, E. A., Hoeijmakers, J. H. J. & Egly, J.-M. (1994)

The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor, *Cell 79*, 1093–1101.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.

Saul, A. & Battistutta, D. (1988) Codon usage in *Plasmodium falciparum*, *Mol. Biochem. Parasitol. 27*, 35–42.

Saul, A. & Battistutta, D. (1990) Analysis of the sequences flanking the translational start sites of *Plasmodium falciparum*, *Mol. Biochem. Parasitol. 42*, 55–62.

Serizawa, H., Makela, T. P., Conaway, J. W., Conaway, R. C., Weinberg, R. A. & Young, R. A. (1995) Association of Cdk-activating kinase subunits with transcription factor TFIIH, *Nature 374*, 280–282.

Shiekhattar, R., Mermelstein, F., Fisher, R. P., Drapkin, R., Dynlacht, B., Wessling, H. C., Morgan, D. O. & Reinberg, D. (1995) Cdk-activating kinase (CAK) complex is a component of human transcription factor IIH, *Nature 374*, 283–287.

Shuttleworth, J., Godfrey, R. & Colman, A. (1990) $p40^{MO15}$, a cdc2-related protein kinase involved in negative regulation of meiotic maturation in *Xenopus* oocytes, *EMBO J. 9*, 3233–3240.

Simanis, V. & Nurse, P. (1986) The cell cycle control gene *cdc2+* for fission yeast encodes a protein kinase potentially regulated by phosphorylation, *Cell 45*, 261–268.

Simon, M., Seraphin, B. & Faye, G. (1986) *KIN28*, a yeast split gene coding for a putative protein kinase homologous to CDC28, *EMBO J. 5*, 2697–2701.

Solomon, M. J. (1994) The function(s) of CAK, the $p34^{cdc2}$-activating kinase, *Trends Biochem. Sci. 19*, 496–500.

Solomon, M. J., Harper, J. W. & Shuttleworth J. (1993) CAK, the $p34^{cdc2}$ activating kinase, contains a protein identical or closely related to $p40^{MO15}$, *EMBO J. 12*, 3133–3142.

Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J. Mol. Biol. 98*, 503–517.

Stepanova, L. Y., Ershler, M. A. & Belyavsky, A. V. (1994) Sequence of the cDNA encoding murine CRK4 protein kinase, *Gene (Amst.) 149*, 321–324.

Tassan, J. P., Schultz, S. J., Bartek, J. & Nigg, E. A. (1994) Cell cycle analysis of the activity, subcellular localization and subunit composition of human CAK (CDK-activating kinase), *J. Cell Biol. 127*, 467–478.

Walliker, D., Quakyi, I. A., Wellems, T. E., McCutchan, T. F., Szarfman, A., London, W. T., Cocoran, L. M., Burkot, T. R. & Carter, R. (1987) Genetic analysis of the human malaria parasite *Plasmodium falciparum*, *Science 236*, 1661–1666.

Weber, J. L. (1988) Molecular biology of malaria parasites, *Exp. Parasitol. 66*, 143–170.

Wellems, T. E., Walliker, D., Smith, C. L., do Rosario, V. E., Maloy, W. L., Howard, R. J., Carter, R. & McCutchan, T. F. (1987) A histidine-rich protein gene marks a linkage group favored strongly in a genetic cross of *Plasmodium falciparum*, *Cell 49*, 633–642.

Wu, L., Yee, A., Liu, L., Carbonaro-Hall, D., Venkatesan, N., Tolo, V. T. & Hall, F. L. (1994) Molecular cloning of the human CAK1 gene encoding a cyclin-dependent kinase-activating kinase, *Oncogene 9*, 2089–2096.

Yamauchi, K. (1991) The sequence flanking translational initiation site in protozoa, *Nucleic Acids Res. 19*, 2715–2720.

Yee, A., Nichols, M. A., Wu, L., Hall, F. L., Kobayashi, R. & Xiong, Y. (1995) Molecular cloning of CDK7-associated human MAT1, a cyclin-dependent kinase-activating kinase (CAK) assembly factor, *Cancer Res. 55*, 6058–6062.