

1998-09-15

A putative protein serine/threonine phosphatase from *Plasmodium falciparum* contains a large N-terminal extension and five unique inserts in the catalytic domain

Li, Ji-Liang

<http://hdl.handle.net/10026.1/10327>

10.1016/s0166-6851(98)00106-6

Molecular and Biochemical Parasitology

Elsevier BV

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Short communication

A putative protein serine/threonine phosphatase from *Plasmodium falciparum* contains a large N-terminal extension and five unique inserts in the catalytic domain¹

Ji-Liang Li^{a,b,*}, David A. Baker^a

^a Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

^b Institute of Tropical Medicine, The First Medical University, Nanfang Hospital, Guangzhou GD510515, People's Republic of China

Received 12 June 1998; received in revised form 10 July 1998; accepted 15 July 1998

Keywords: *Plasmodium falciparum*; Malaria; Protein serine/threonine phosphatase; Chromosome; mRNA; Gametocyte

Plasmodium falciparum possesses a complex life cycle involving two different hosts and interactions with multiple cell types. During the various stages of the life cycle the parasite undergoes cell growth and division, development and differentia-

tion. Sexual stage development, gametocytogenesis in the host and gametogenesis in the mosquito, is accompanied by biochemical and morphological changes in the parasite. The molecular and cellular mechanisms involved in regulation of proliferation, development and differentiation of *P. falciparum* are unclear. However, inhibitors of protein kinases and phosphatases can interfere with parasite growth [1], suggesting a requirement for phosphorylation–dephosphorylation in control of the parasite life cycle. Reversible phosphorylation has been shown to play an important role in invasion of erythrocytes by merozoites [2] and in intraerythrocytic growth and development [3]. Recently, a number of protein serine/threonine kinase genes have been isolated from *P. falciparum* [4–7].

Abbreviations: PPP, PP-1/PP-2A/PP-2B protein serine/threonine phosphatase family; PfPP- α , *Plasmodium falciparum* protein serine/threonine phosphatase α ; PfPP- β , *Plasmodium falciparum* protein serine/threonine phosphatase β ; PFGE, pulse-field gel electrophoresis.

* Corresponding author. Present address: Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK. Tel.: +44 1865 275722; fax: +44 1865 275721; e-mail: jlli@bioch.ox.ac.uk

¹ Note: Nucleotide sequence data reported in this paper are available in the GenBank™, EMBL and DDJB databases under the accession number U88869.

To understand the cell cycle events and signal transduction pathways regulated by reversible phosphorylation in *P. falciparum*, apart from the protein kinases, the relevant protein phosphatases involved have to be identified and characterized. In a previous paper, we have described a PP-2A-like protein serine/threonine phosphatase from *P. falciparum* [8]. Here, we report the molecular characterization of another novel PP-1/PP-2A/PP-2B (PPP)-related protein serine/threonine phosphatase from *P. falciparum*. This phosphatase, termed *P. falciparum* protein serine/threonine phosphatase α (PfPP- α), not only contains five unusual inserts in the catalytic domain but also has the largest N-terminal extension amongst known protein phosphatases of the PPP family.

On the basis of two highly conserved motifs (GDYVDRG and DLLWSDP) in the catalytic domain of the PPP family [9], two degenerate oligonucleotide primers, termed P1 (5' GGIGAT-TAT/CGTIGAT/CA/CGIGG 3', nucleotides 2043–2062) and P3 (5' GGATCIIICCAIAA/GIAA/GA/GTC 3', nucleotides 2415–2434), were constructed and then used in PCR to amplify the desired gene from 3D7A genomic DNA (data not shown). The sequence of one PCR clone (P1–P3) revealed a 392 bp insert encoding 131 amino acids of a protein serine/threonine phosphatase-like protein. The A + T content and codon usage of the nucleotide sequence are typical of the coding region of *P. falciparum* genes. Southern blotting confirmed that the P1–P3 clone is derived from parasite DNA (data not shown). Therefore, the P1–P3 clone was employed to screen a λ GEM-12 genomic library [6]. Restriction digestion of one positive plaque (λ P1) DNA with a variety of enzymes gave rise to a fragment of \approx 1.5 kb, which was shown by hybridization to contain the target gene. The fragment was cloned into the pUC19 vector (clone P2-2). Sequence data showed that the P2-2 clone contained only a partial sequence of the desired gene. Accordingly, attempts were made to clone other fragments including a 3.0 kb *Bam*HI fragment and a 4.5 kb *Eco*RI fragment. This proved unsuccessful. Consequently, vectorette libraries were constructed and screened by PCR. Based on the southern blotting data and the sequence obtained from the P1–P3

PCR and P2-2 pUC19 clones, a partial restriction map of the gene was established (Fig. 1A). Two specific primers, PP1 (to obtain further sequence in the 3' direction) and PP10 (to obtain further sequence in the 5' direction), were used in PCR to walk outwards in both directions. The sequence data thus acquired permitted construction of the PP14 primer and subsequent complete sequencing across the gene (Fig. 1A).

The sequence derived from the overlapping genomic clone and vectorette PCR fragments consists of 3128 bp and contains one open reading frame with a putative ATG start codon at nucleotide 174 and a TAA stop codon at nucleotide 2799. The ORF encodes a protein of 875 amino acids with a predicted molecular mass of \approx 99 kDa. The predicted PfPP- α protein comprises two distinct domains. The N-terminal segment consists of 516 amino acids (see below), whereas the C-terminal segment has 359 amino acids and contains a region of high similarity to the catalytic subunits of the other protein phosphatases in the PPP family. The catalytic domain of PfPP- α contains the majority (40 out of 45) of the conserved residues in the PPP family [10]. Some of these residues have been known to be involved in metal cation binding, substrate recognition, catalysis and toxin binding by genetic mutagenesis [11] and more recently by crystallographic studies of rabbit and human PP-1 [12]. These include (relative to the residue numbers of rabbit muscle PP-1 α) [12]: (a) Asp95, Arg96, Asn124, His125, Asp208 and Tyr272, forming part of the substrate binding and catalytic sites; (b) Asp64, His66, Asp92, Asn124, His173, and His248, which coordinate Mn²⁺ and Fe²⁺ binding in PP-1; (c) Arg246, being essential for the function of phosphatase enzymes; (d) Tyr272, involved in binding to toxins such as tautomycin, calyculin A and okadaic acid; and (e) the motif Ile121-Arg122-Gly123-Asp124-His125-Glu126, which is a signature of the catalytic domain of serine/threonine specific protein phosphatases in the PPP family. However, five well-conserved residues (Gly80, Pro196, Gly199, Arg211, and Pro270) in the PPP family [10] are replaced in PfPP- α by Lys80, Ser196, Pro199, Thr211, and Thr270, respectively. Database searches revealed further that the catalytic domain

A.

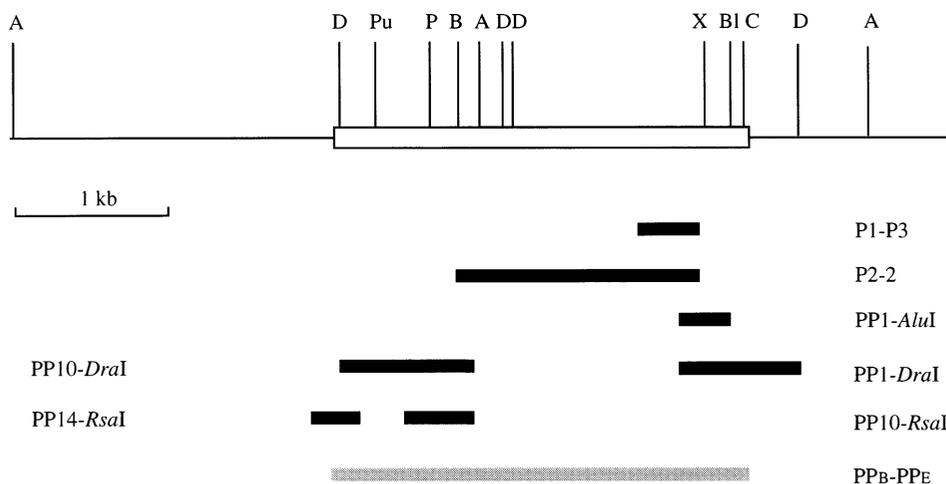


Fig. 1. (A) A schematic representation of a partial restriction map of the *Pfpp- α* gene and the overlapping clones of the genomic library and PCR fragments used to determine the nucleotide sequence of *Pfpp- α* . A, *Acc*I; B, *Bam*HI; BI, *Bgl*II; C, *Cla*I; D, *Dra*I; P, *Pst*I; Pu, *Pvu*II, and X, *Xba*I. The open box indicates the coding region of the *Pfpp- α* gene. The PPB-PPE fragment is derived from genomic DNA. (B) Alignment of the predicted amino acid sequences of PfPP- α with human PP-1 α and PP-2A α . The GenBank™/EMBL database accession numbers are as follows: *P. falciparum* PfPP- α , U88869; human HuPP-1 α , X70848; human HuPP-2A α , X12646. 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. Inserts I-V are labelled on the top of sequences. The positions of secondary structural elements [12] are shown above the sequences. (C) A schematic representation of the structural comparison of PfPP- α with protein phosphatases in the PPP family. The black boxes denote the conserved catalytic domain of protein phosphatases in the PPP family and the open boxes represent the N- or C-terminal extension of protein phosphatases. The amino acid numbers of each phosphatase are shown to the right.

of PfPP- α is more closely related to PP-1 than to PP-2A (Fig. 1B). It shares 48–54% similarity and 40–46% identity with PP-1 from the different species, in comparison to the 42–46% similarity and 33–38% identity with PP-2A and to the 36–42% similarity and 31–35% identity with PP-2B of other organisms. Sequence comparison of PfPP- α with PP-1 showed that PfPP- α contains a number of residues conserved in the PP-1 subfamily (see Fig. 1B). However, it is different from PP-1 in several respects. There are changes in more than 40 residues which are well conserved in all members of the PP-1 group [10], including some within the Gly274-Glu275-Phe276-Asp277 motif in the β 12- β 13 loop of PP-1, which is thought to be involved in the binding of okadaic

acid [11]. In addition, PfPP- α contains five unique inserts in the catalytic domain: insert I, consisting of eight amino acids, located within the α A helix based on the tertiary structure of PP-1 [12]; insert II, containing 14 amino acid residues, upstream of the β 3 strand; insert III, comprising six amino acids, between the α E and α F helices; insert IV, composed of six amino acid residues, positioned upstream of the α H helix; and insert V, possessing eight amino acids, following the β 8 strand. To investigate the potential function of these inserts, the three-dimensional structure of the catalytic domain of PfPP- α was modelled using the computer programme Swiss-Pdb Viewer (Version 2.11) on the basis of the crystal structure of the catalytic domain of rabbit muscle PP-1 α [12] (data

C.

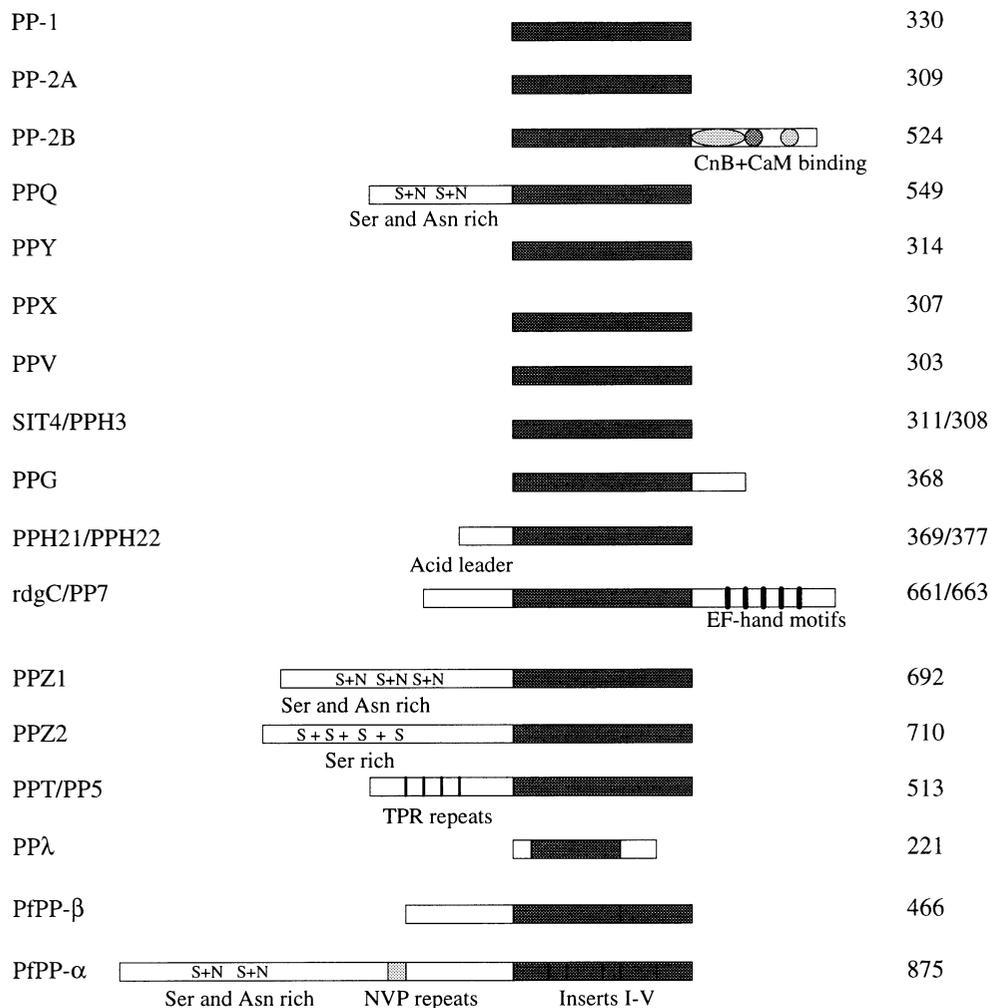


Fig. 1. (Continued)

The amino-terminal region of PfPP- α is not related to any previously described protein serine/threonine phosphatases and has several interesting features. Firstly, there are six copies of a degenerate trimer N(I)VP(Q) repeat at residue positions 419–436, which is reminiscent of the NVDP repeats found in the *P. falciparum* circumsporozoite protein. Secondly, in contrast to the catalytic domain, this region is relatively rich in serine

(8.5%), asparagine (9.5%) and threonine (6.2%). The N-terminal segment represents about 59.0% of PfPP- α in size, however, 71.0% of serine (44/62) and asparagine (49/69), and 72.7% of threonine (32/44) residues are located within the N-terminal portion. It is noteworthy that the N-terminal domains of PPQ (comprising 236 residues), PPZ1 (containing 360 amino acids) and PPZ2 (consisting of 395 residues) are also rich in serine and/or

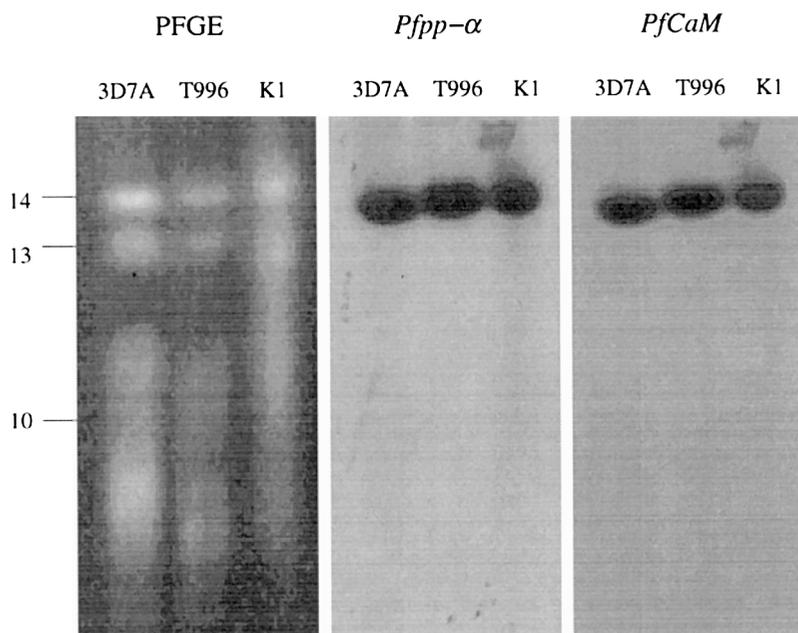
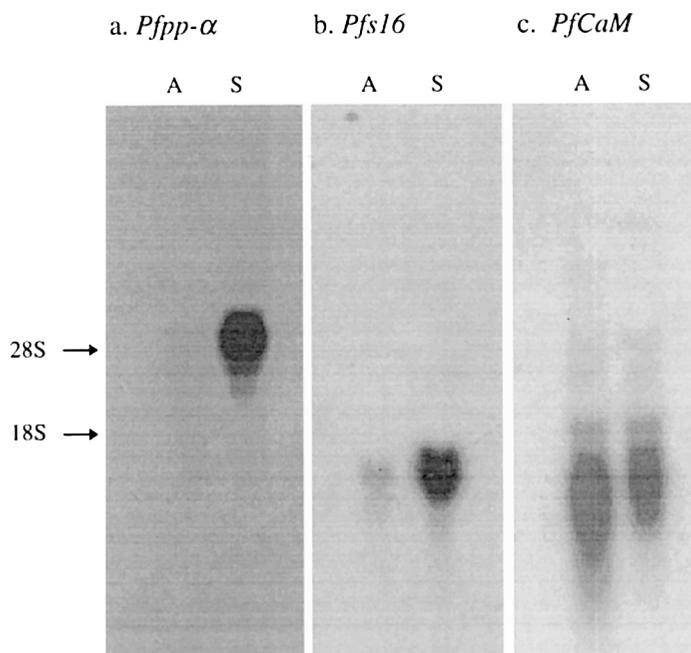
A.**B.**

Fig. 2.

asparagine [13–15]. Thirdly, the N-terminal domain contains multiple potential phosphorylation sites for a range of known protein kinases [16]. For instance, if the sequence R/K-X-X-S/T is considered to be the minimal recognition motif for cAMP-dependent protein kinase, more than ten serine/threonine residues may act as potential phosphorylation sites. Similarly, if the motif S/T-X-K/R is a minimal requirement for protein kinase C-dependent phosphorylation, at least eight serine/threonine residues could be considered as potential phosphorylation sites. Thus, there is a strong possibility of regulation of PfPP- α by reversible phosphorylation. It is well known that protein serine/threonine phosphatase activity can also be regulated through the interaction of the catalytic subunits with regulatory proteins [17]. Therefore, it is also tempting to speculate that the N-terminal extension of PfPP- α may regulate the specificity of the C-terminal catalytic domain by binding one or more accessory subunits that target the enzyme to its correct location. Finally, the N-terminal segment is the largest extension amongst the known protein phosphatases in the PPP family yet described. The phosphatases carrying an N-terminal extension include *Drosophila* rdgC/human PP7 [18,19], *Saccharomyces cerevisiae* PPH21, PPH22 [20], CNA1/CMP1 [21], CNA2/CMP2 [22], PPQ [13], PPZ1 [14], PPZ2 [15], PPT (PP5 in human) [23] and *P. falciparum* PfPP- β [8]. PPZ2 possesses the largest N-terminal extension (≈ 395 residues) amongst the above mentioned phosphatases. However, the N-terminal region (516 amino acids) of PfPP- α is substantially longer than that of PPZ2.

In order to reveal the precise relationship of PfPP- α with other protein phosphatases identified so far, comparison of the overall sequence of PfPP- α with other relevant phosphatases in the PPP family was carried out (Fig. 1C). The results clearly show that PfPP- α is a member of the PPP family but cannot be assigned to any particular subgroup with the PPP family. Therefore, PfPP- α seems to comprise a distinct subgroup in the PPP family.

To determine the copy number of the *Pfpp*- α gene in the parasite genome, southern blot analyses were performed using genomic DNA (3D7A) digested with a number of restriction enzymes. Hybridization of the blots were performed with the P1–P3 probe and the P2-2 fragment, respectively. The results (data not shown) are consistent with the restriction map (see Fig. 1A), indicating strongly that *Pfpp*- α is a single-copy gene per haploid genome. To determine the chromosomal location of the *Pfpp*- α gene, hybridization of the P2-2 clone was carried out using a pulse-field gel electrophoresis (PFGE) blot containing 3D7A, T996 and K1 chromosomes. A single band was detected corresponding to chromosome 14 (Fig. 2A). This was repeated using the same probe with an independent PFGE blot (data not shown). The result was further confirmed by re-probing the same blots with the calmodulin gene (*PfCaM*) which is located on chromosome 14. The single band result on the PFGE blots also supports the conclusion that *Pfpp*- α is a single-copy gene in the parasite genome.

Analysis of *Pfpp*- α mRNA expression in the erythrocytic life cycle stages of *P. falciparum* may give some clues as to its possible function in the

Fig. 2. (A). Chromosome localization of *Pfpp*- α . Parasite chromosomes from *P. falciparum* 3D7A, T996 and K1 were separated by pulse-field gel electrophoresis, stained with ethidium bromide, blotted onto a nylon membrane and hybridized with radiolabeled probes. According to the yeast chromosome markers and hybridization of several *P. falciparum* chromosome marker genes, the positions of chromosome 10, 13, and 14 were identified on the ethidium bromide-stained gel. *Pfpp*- α and *PfCaM* both hybridized to chromosome 14. (B) Northern blot analysis of the *Pfpp*- α gene. Ten μg of total RNA extracted from the asexual erythrocytic stages (A) and sexual erythrocytic stages (mainly stage III–V) (S) of *P. falciparum* (clone 3D7A) were fractionated in a denaturing formaldehyde gel, blotted onto a nylon membrane and hybridized to radiolabeled probes. The positions of *P. falciparum* rRNA subunits (18S and 28S) are indicated by arrows. (a), (b) and (c) are autoradiographs of the membrane probed with the *Pfpp*- α gene, the *Pfs16* sexual stage-specific gene and the calmodulin housekeeping gene (*PfCaM*), and exposed for 24 h, 50 min and 96 h, respectively. *Pfpp*- α detected a transcript of ≈ 3900 nucleotides in the sexual erythrocytic stages, *Pfs16* hybridized with a band of ≈ 1400 nucleotides in the sexual erythrocytic stages and *PfCaM* hybridized with two bands of 1300, and 1000 nucleotides from the parasite.

parasite development and differentiation. Northern blot analysis was performed by using the P2-2 probe. A transcript of ≈ 3900 nucleotides in size was only detected in the lane containing sexual stage RNA, migrating just behind the 28S ribosomal RNA band (Fig. 2B). This suggests that PfPP- α is involved in sexual stage-specific events. It is known that in vertebrates protein dephosphorylation plays a key role in a number of testicular functions including germ cell meiosis [24] and sperm motility [25]. Testis-specific forms of PP-1 and PP-2B have been identified in mammalian systems [25]. Recently, a male specific protein phosphatase, PPY, has been localized in somatic cells of the *Drosophila* testis [26]. In the plant *Arabidopsis*, it has been also shown that a PP-1 homologue is predominantly expressed in the tapetum, the developing and mature pollen and in the ovaries [27]. Interestingly, PfKIN, an SNF1-type protein kinase, and PfMRP, a MAP kinase-related protein, have been reported to be specifically expressed in the sexual stages of the life cycle [5,7]. These findings imply that reversible protein phosphorylation events may be particularly important in regulation of sexual stage development in *P. falciparum*. Various inhibitors and activators of signal transduction pathways have been shown to influence events such as gametocytogenesis and exflagellation [28,29]. DNA replication, cell division (exflagellation), and fertilization occur sequentially in the sexual stage. Interestingly, Pfcrk-1 and Pfmrk, two *P. falciparum* Cdk-related protein kinases, have been recently shown to be expressed predominantly in gametocytes [4,6]. Taken together it is concluded that PfPP- α may play an important role in the control of cell cycle and signal transduction involved in the development of sexual stage and/or in fertilization.

Acknowledgements

We are very grateful to Professor Geoffrey A.T. Targett for his tremendous support and encouragement on this project, to Dr Adam A. Witney for his generous assistance with the computer modelling, to Dr Belinda S. Hall, and Dr Martin R. Goodier for their critical comments of the

manuscript, and to Olalekan O. Daramola for his help with parasite cultivation. We are greatly indebted to Dr Jane M.R. Carlton, University of Edinburgh, for her suggestion on the pulse-field gel electrophoresis. This work was supported in part by a programme grant (to G.A.T. Targett) and a project grant from the Wellcome Trust. J.-L. Li was also supported by the Sino-British Friendship Scholarship Scheme (SBFSS) and the Great Britain–China Educational Trust.

References

- [1] Ward GE, Fujioka H, Aikawa M, Miller LH. Staurosporine inhibits invasion of erythrocytes by malarial merozoites. *Exp Parasitol* 1994;79:480–7.
- [2] Rangachari K, Dluzewski A, Wilson RJM, Gratzer WB. Control of malarial invasion by phosphorylation of the host cell membrane cytoskeleton. *Nature* 1986;324:364–5.
- [3] Yuthavong Y, Limpaboon T. The relationship of phosphorylation of membrane proteins with the osmotic fragility and filterability of *Plasmodium berghei*-infected mouse erythrocytes. *Biochim Biophys Acta* 1987;929:278–87.
- [4] Doerig C, Doerig C, Horrocks P, et al. Pfcrk-1, a developmentally regulated cdc2-related protein kinase of *Plasmodium falciparum*. *Mol Biochem Parasitol* 1995;70:167–74.
- [5] Bracchi V, Langsley G, Thelu J, Eling W, Ambroise-Thomas P. PfKIN, an SNF1 type protein kinase of *Plasmodium falciparum* predominantly expressed in gametocytes. *Mol Biochem Parasitol* 1996;76:299–303.
- [6] Li JL, Robson KJH, Chen JL, Targett GAT, Baker DA. Pfmrk, a MO15-related protein kinase from *Plasmodium falciparum*: gene cloning, sequence, stage-specific expression and chromosome localization. *Eur J Biochem* 1996;241:805–13.
- [7] Lin DT, Goldman ND, Syin C. Stage-specific expression of a *Plasmodium falciparum* protein related to the eukaryotic mitogen-activated protein kinases. *Mol Biochem Parasitol* 1996;78:67–77.
- [8] Li JL, Baker DA. Protein phosphatase β , a putative type-2A protein phosphatase from the human malaria parasite *Plasmodium falciparum*. *Eur J Biochem* 1997;249:98–106.
- [9] Cohen PTW, Brewis ND, Hughes V, Mann DJ. Protein serine/threonine phosphatases: an expanding family. *FEBS Lett* 1990;268:355–9.
- [10] Barton GJ, Cohen PTW, Barford D. Conservation analysis and structure prediction of the protein serine/threonine phosphatases. *Eur J Biochem* 1994;220:225–37.
- [11] Zhang Z, Zhao S, Long F, et al. A mutant of protein phosphatase-1 that exhibits altered toxin sensitivity. *J Biol Chem* 1994;269:16997–7000.

- [12] Goldberg J, Huang H, Kwon Y, Greengard P, Nairn AC, Kuriyan J. Three-dimensional structure of the catalytic subunit of serine/threonine phosphatase-1. *Nature* 1995;376:745–53.
- [13] Chen MX, Chen YH, Cohen PTW. PPQ, a novel protein phosphatase containing a Ser + Asn-rich amino terminal domain, is involved in the regulation of protein synthesis. *Eur J Biochem* 1993;218:689–99.
- [14] Posas F, Casamayor A, Morral N, Arino J. Molecular cloning and analysis of a yeast protein phosphatase with an unusual amino-terminal region. *J Biol Chem* 1992;267:11734–40.
- [15] Hughes V, Muller A, Stark MJR, Cohen PTW. Both isoforms of protein phosphatase Z are essential for the maintenance of cell size and integrity in *Saccharomyces cerevisiae* in response to osmotic stress. *Eur J Biochem* 1993;216:269–79.
- [16] Kennelly PJ, Krebs EG. Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. *J Biol Chem* 1991;266:15555–8.
- [17] Wera S, Hemmings BA. Serine/threonine protein phosphatases. *Biochem J* 1995;311:17–29.
- [18] Steele FR, Washburn T, Rieger R, O'Tousa JE. *Drosophila* retinal degeneration C (*rdgC*) encodes a novel serine/threonine protein phosphatase. *Cell* 1992;69:669–76.
- [19] Huang X, Honkanen RE. Molecular cloning, expression, and characterization of a novel human serine/threonine protein phosphatase, PP7, that is homologous to *Drosophila* retinal degeneration C gene product (*rdgC*). *J Biol Chem* 1998;273:1462–8.
- [20] Sneddon AA, Cohen PT, Stark MJ. *Saccharomyces cerevisiae* protein-phosphatase 2A performs an essential cellular function and is encoded by two genes. *EMBO J* 1990;9:4339–46.
- [21] Liu Y, Ishii S, Tokai M, et al. The *Saccharomyces cerevisiae* genes (CMP1 and CMP2) encoding calmodulin-binding proteins homologous to the catalytic subunit of mammalian protein phosphatase 2B. *Mol Gen Genet* 1991;227:52–9.
- [22] Cyert MS, Kunisawa R, Kaim D, Thorner J. Yeast has homologs (CNA1 and CNA2 gene products) of mammalian calcineurin, a calmodulin-regulated phosphoprotein phosphatase. *Proc Natl Acad Sci USA* 1991;88:7376–80.
- [23] Chen MX, McPartlin AE, Brown L, Chen YH, Barker HM, Cohen PTW. A novel human protein serine/threonine phosphatase, which possesses four tetratricopeptide repeat motifs and localizes to the nucleus. *EMBO J* 1994;13:4278–90.
- [24] Alphey L, Jimenez J, White-Cooper H, Dawson I, Nurse P, Glover DM. *Twine*, a *cdc25* homologue that functions in the male and female germline of *Drosophila*. *Cell* 1992;69:977–88.
- [25] Muramatsu T, Giri PR, Higuchi S, Kincaid RL. Molecular cloning of a calmodulin-dependent phosphatase from murine testis: identification of a developmentally expressed nonneural isoenzyme. *Proc Natl Acad Sci USA* 1992;89:529–33.
- [26] Armstrong CG, Mann DJ, Berndt N, Cohen PTW. *Drosophila* PPY, a novel male specific protein serine/threonine phosphatase localised in somatic cells of the testis. *J Cell Sci* 1995;108:3367–75.
- [27] Arundhati A, Feiler H, Traas J, Zhang H, Lunness PA, Doonan JH. A novel Arabidopsis type I protein phosphatase is highly expressed in male and female tissues and functionally complements a conditional cell cycle mutant of *Aspergillus*. *Plant J* 1995;7:823–34.
- [28] Kawamoto F, Alejo-Blanco R, Fleck SL, Kawamoto Y, Sinden RE. Possible role of Ca^{2+} and cGMP as mediators of the exflagellation of *Plasmodium berghei* and *Plasmodium falciparum*. *Mol Biochem Parasitol* 1990;42:101–8.
- [29] Ogwan'g R, Mwangi J, Gachihi G, Nwachukwu A, Roberts CR, Martin SK. Use of pharmacological agents to implicate a role for phosphoinositide hydrolysis products in malaria gamete formation. *Biochem Pharmacol* 1993;46:1601–6.