Primary structure and sexual stage-specific expression of a LAMMER protein kinase of *Plasmodium falciparum*

Ji-Liang Li\^a,b,*, Geoffrey A.T. Targett\^a, David A. Baker\^a

\^aDepartment of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
\^bInstitute of Tropical Medicine, The First Medical University, Nanfang Hospital, Guangzhou GD510515, China

Received 9 November 2000; received in revised form 2 January 2001; accepted 2 January 2001

Abstract

We have isolated a LAMMER-like gene from *Plasmodium falciparum* by vectorette technique. The gene consists of 3316 bp encoding a protein 881 amino acids with a predicted molecular mass of approximately 106.7 kDa. The encoded protein, termed PfLAMMER, is composed of two distinct domains. The N-terminal domain is not related to any previously described protein kinases and has several interesting features including multiple consensus phosphorylation sites for a range of protein kinases, a number of RS/SR dipeptides, a large proportion of charged amino acids, two putative nuclear localisation signals and 14 copies of a tetramer DKYD repeats. The C-terminal domain is characteristic of a kinase in the LAMMER family with the highest homology to the *Arabidopsis thaliana* AFC3 kinase. Genomic restriction analysis showed that PfLAMMER is encoded by a single copy gene in the parasite genome. A single transcript of approximately 3800 nucleotides is expressed specifically in the sexual stage, indicating that PfLAMMER may be important in regulating the processes of sexual differentiation of the parasite. © 2001 Australian Society for Parasitology Inc. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Malaria; Protein kinase; LAMMER; Gametocyte; *Plasmodium falciparum*

The resistance of *Plasmodium falciparum* to drugs and the resistance of mosquitoes to insecticides have resulted in a resurgence of malaria in the tropics and subtropics. Therefore, development of effective vaccines and new anti-malarial drugs has become a top priority. A better understanding of the various cellular processes at different stages of the parasite will help to identify new targets for vaccine and drug development. Gametocytogenesis within the human erythrocytes and gametogenesis in the mosquito midgut are two key processes in the parasite life cycle. However, the molecular mechanisms involved in control of the sexual development are unclear (Lobo and Kumar, 1998). We are interested in signal transduction pathways involved in the sexual differentiation of *P. falciparum*. The first step toward this goal has been to identify the components, mainly protein kinases and phosphatases, of the signal transduction pathways. Recently, we have reported several sexual stage-specific genes encoding either protein serine/threonine phosphatases (Li and Baker, 1997, 1998) or protein kinases (Li et al., 1996, 2000). In this paper, we describe the molecular cloning and characterisation of a novel gene encoding a *P. falciparum* LAMMER protein kinase (PfLAMMER). PfLAMMER possesses a large N-terminal extension and is expressed specifically in the sexual stage of the parasite life cycle, indicating that it may be important in regulating the processes of sexual stage development.

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). Initially, although our interests lay in the *P. falciparum* protein kinases in general, cyclin-dependent kinases were targeted. Therefore, two primers, cdcI (5'GAAA/AA/GITIGGIGAA/GGGIAC-3', 1798±1817) and cdcII (5'TCCAA/TAACTCCATACATCACT-3', 2395±2415), were constructed based on the conserved sequences of subdomains I (EKIGEGT) and IX (SDVWSF/YG) and used in PCR. A band of 617 bp was amplified from *P. falciparum* 3D7A genomic DNA (data not shown) and inserted into the pGEM-T vector (Promega). The sequence data revealed that the PCR clone (cdcI-II) encodes 205 amino acids of a kinase-like protein. The A + T content and codon usage are typical of the coding region of *P. falciparum* genes (Weber, 1988). Southern blot analysis

* Corresponding author. Present address: Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9D5, UK. Tel.: + 44-1865-222419; fax: + 44-1865-222431.
E-mail address: lij@icrf.icnet.uk (J.-L. Li).

---

Nucleotide sequence data reported in this paper are available in the GenBank\textsuperscript{™}, EMBL and DDBJ databases under the accession number AF104915.

© 2001 Australian Society for Parasitology Inc. Published by Elsevier Science Ltd. All rights reserved.

PII: S0020-7519(01)00126-6

0200-7519/01/$20.00 © 2001 Australian Society for Parasitology Inc. Published by Elsevier Science Ltd. All rights reserved.
confirmed that the cdcI-II clone is derived from parasite DNA (data not shown). To isolate the full-length gene, two specific primers, PSK1 (5'-GTTCGAACAATTCTTACACGGTAG-3', 1863–1887) (to obtain further sequence in the 5’ direction) and PSK2 (5'-CTCCAGAAGTTATATTAAATTTGGG-3', 2357–2381) (to obtain further sequence in the 3’ direction), were constructed on the basis of known sequences and used in PCR to screen vectorette libraries (Li et al., 1996). Four fragments (PSK1-Rsal, PSK1-Sau3AI, PSK2-Rsal and PSK2-DraI) (Fig. 1) were obtained and sequenced. As expected, PSK1-Sau3AI contains all sequence of PSK1-Rsal while PSK2-DraI covers a full-length of PSK2-Rsal. The sequence data of PSK1-Sau3AI permitted construction of the PSK5 primer (5'-CATACTGTGTGCTATTTGCTCTC-3', 1048–1072) and subsequent screening of vectorette libraries. Based on the sequence of PSK5-TaqI, the PSK27 primer (5'-CATCTGTGTGAATTTGTA-3', 518–543) was designed to produce the PSK27-DraI fragment. Database searches revealed that in the P. falciparum tag database there was a PNAQ25TR fragment containing all sequence of PSK27-DraI. Nucleotide sequence analysis revealed a putative start codon at nucleotide 2776. The open reading frame starts with an ATG codon at nucleotide 133 and terminates with a TAA termination codon. A diffuse G+T-rich sequence occurs at nucleotide +24, downstream of the TAA termination codon. The sequence and codon usage in the coding region are typical for a P. falciparum gene (Weber, 1988). The A + T contents of both flanking regions (91.9 and 84.4%, respectively) are characteristically higher than that of the coding region (75.5%). Two putative polyadenylation signals (AATAAA) are found at nucleotides +242 and +321, downstream of the TAA termination codon. A diffuse G+T-rich sequence occurs at nucleotide +24, downstream of the first polyadenylation signal. The existence of introns in the coding sequence was excluded by reverse transcription-PCRs (Li and Baker, 1997; Li et al., 2000; Li and Cox, 2000) using a number of primer pairs that cover the whole coding region.

The open reading frame encodes a protein of 881 amino acids with a predicted Mr of approximately 106.7 kDa. Database searches revealed that the amino acid sequence of PfLAMMER shares 56–60% similarity and 37–41% identity with kinases in the LAMMER family across the catalytic domain. PfLAMMER has the highest homology to the Arabidopsis thaliana AFC3 kinase (41% identity, 60% similarity) (Bender and Fink, 1994). Fig. 2 shows a sequence alignment of LAMMER kinases from diverse species. PfLAMMER is composed of two distinct domains. The N-terminal domain, consisting of 547 amino acids with a predicted Mr of approximately 106.7 kDa, is not related to any previously described protein domain and the PCR product was sequenced in both strands (see Fig. 1). The sequence derived from overlapping fragments consists of 3316 bp and contains an open reading frame starting with an ATG codon at nucleotide 133 and terminating with a TAA codon at nucleotide 2776. The sequence and codon usage in the coding region are typical for a P. falciparum gene (Weber, 1988). The A + T contents of both flanking regions (91.9 and 84.4%, respectively) are characteristically higher than that of the coding region (75.5%). Two putative polyadenylation signals (AATAAA) are found at nucleotides +242 and +321, downstream of the TAA termination codon. A diffuse G+T-rich sequence occurs at nucleotide +24, downstream of the first polyadenylation signal. The existence of introns in the coding sequence was excluded by reverse transcription-PCRs (Li and Baker, 1997; Li et al., 2000; Li and Cox, 2000) using a number of primer pairs that cover the whole coding region.

The open reading frame encodes a protein of 881 amino acids with a predicted Mr of approximately 106.7 kDa. Database searches revealed that the amino acid sequence of PfLAMMER shares 56–60% similarity and 37–41% identity with kinases in the LAMMER family across the catalytic domain. PfLAMMER has the highest homology to the Arabidopsis thaliana AFC3 kinase (41% identity, 60% similarity) (Bender and Fink, 1994). Fig. 2 shows a sequence alignment of LAMMER kinases from diverse species. PfLAMMER is composed of two distinct domains. The N-terminal domain, consisting of 547 amino acid residues, is not related to any previously described protein domain and the PCR product was sequenced in both strands (see Fig. 1). The sequence derived from overlapping fragments consists of 3316 bp and contains an open reading frame starting with an ATG codon at nucleotide 133 and terminating with a TAA codon at nucleotide 2776. The sequence and codon usage in the coding region are typical for a P. falciparum gene (Weber, 1988). The A + T contents of both flanking regions (91.9 and 84.4%, respectively) are characteristically higher than that of the coding region (75.5%). Two putative polyadenylation signals (AATAAA) are found at nucleotides +242 and +321, downstream of the TAA termination codon. A diffuse G+T-rich sequence occurs at nucleotide +24, downstream of the first polyadenylation signal. The existence of introns in the coding sequence was excluded by reverse transcription-PCRs (Li and Baker, 1997; Li et al., 2000; Li and Cox, 2000) using a number of primer pairs that cover the whole coding region.

The open reading frame encodes a protein of 881 amino acids with a predicted Mr of approximately 106.7 kDa. Database searches revealed that the amino acid sequence of PfLAMMER shares 56–60% similarity and 37–41% identity with kinases in the LAMMER family across the catalytic domain. PfLAMMER has the highest homology to the Arabidopsis thaliana AFC3 kinase (41% identity, 60% similarity) (Bender and Fink, 1994). Fig. 2 shows a sequence alignment of LAMMER kinases from diverse species. PfLAMMER is composed of two distinct domains. The N-terminal domain, consisting of 547 amino acid residues, is not related to any previously described protein domain and the PCR product was sequenced in both strands (see Fig. 1). The sequence derived from overlapping fragments consists of 3316 bp and contains an open reading frame starting with an ATG codon at nucleotide 133 and terminating with a TAA codon at nucleotide 2776. The sequence and codon usage in the coding region are typical for a P. falciparum gene (Weber, 1988). The A + T contents of both flanking regions (91.9 and 84.4%, respectively) are characteristically higher than that of the coding region (75.5%). Two putative polyadenylation signals (AATAAA) are found at nucleotides +242 and +321, downstream of the TAA termination codon. A diffuse G+T-rich sequence occurs at nucleotide +24, downstream of the first polyadenylation signal. The existence of introns in the coding sequence was excluded by reverse transcription-PCRs (Li and Baker, 1997; Li et al., 2000; Li and Cox, 2000) using a number of primer pairs that cover the whole coding region.

The open reading frame encodes a protein of 881 amino acids with a predicted Mr of approximately 106.7 kDa. Database searches revealed that the amino acid sequence of PfLAMMER shares 56–60% similarity and 37–41% identity with kinases in the LAMMER family across the catalytic domain. PfLAMMER has the highest homology to the Arabidopsis thaliana AFC3 kinase (41% identity, 60% similarity) (Bender and Fink, 1994). Fig. 2 shows a sequence alignment of LAMMER kinases from diverse species. PfLAMMER is composed of two distinct domains. The N-terminal domain, consisting of 547 amino acid residues, is not related to any previously described protein domain and the PCR product was sequenced in both strands (see Fig. 1). The sequence derived from overlapping fragments consists of 3316 bp and contains an open reading frame starting with an ATG codon at nucleotide 133 and terminating with a TAA codon at nucleotide 2776. The sequence and codon usage in the coding region are typical for a P. falciparum gene (Weber, 1988). The A + T contents of both flanking regions (91.9 and 84.4%, respectively) are characteristically higher than that of the coding region (75.5%). Two putative polyadenylation signals (AATAAA) are found at nucleotides +242 and +321, downstream of the TAA termination codon. A diffuse G+T-rich sequence occurs at nucleotide +24, downstream of the first polyadenylation signal. The existence of introns in the coding sequence was excluded by reverse transcription-PCRs (Li and Baker, 1997; Li et al., 2000; Li and Cox, 2000) using a number of primer pairs that cover the whole coding region.

The open reading frame encodes a protein of 881 amino acids with a predicted Mr of approximately 106.7 kDa. Database searches revealed that the amino acid sequence of PfLAMMER shares 56–60% similarity and 37–41% identity with kinases in the LAMMER family across the catalytic domain. PfLAMMER has the highest homology to the Arabidopsis thaliana AFC3 kinase (41% identity, 60% similarity) (Bender and Fink, 1994). Fig. 2 shows a sequence alignment of LAMMER kinases from diverse species. PfLAMMER is composed of two distinct domains. The N-terminal domain, consisting of 547 amino acid residues, is not related to any previously described protein domain and the PCR product was sequenced in both strands (see Fig. 1). The sequence derived from overlapping fragments consists of 3316 bp and contains an open reading frame starting with an ATG codon at nucleotide 133 and terminating with a TAA codon at nucleotide 2776. The sequence and codon usage in the coding region are typical for a P. falciparum gene (Weber, 1988). The A + T contents of both flanking regions (91.9 and 84.4%, respectively) are characteristically higher than that of the coding region (75.5%). Two putative polyadenylation signals (AATAAA) are found at nucleotides +242 and +321, downstream of the TAA termination codon. A diffuse G+T-rich sequence occurs at nucleotide +24, downstream of the first polyadenylation signal. The existence of introns in the coding sequence was excluded by reverse transcription-PCRs (Li and Baker, 1997; Li et al., 2000; Li and Cox, 2000) using a number of primer pairs that cover the whole coding region.
kinases and has several interesting features. Firstly, it is rich in serine/threonine [14% (74/547)] and tyrosine [14% (73/547)], forming a number of potential phosphorylation sites for a range of known protein kinases (Kennelly and Krebs, 1991). Thus, there is a strong possibility for regulation of the PILAMMER activity by reversible phosphorylation of the N-terminal segment. Secondly, the N-terminal domain contains many serine and arginine residues [15% (92/616)].


389
indicating that PfLAMMER is an SR-like protein (Nayler et al., 1997). Close inspection of the sequence revealed 11 RS/SR dipeptides and two RSRS motifs in this region, implying that the N-terminal segment may interact with other SR proteins (Wu and Maniatis, 1993; Kohtz et al., 1994). Thirdly, a large proportion of charged amino acids [37% (203/547) K, R, E and D] exists in the N-terminal segment, suggesting further that this region may be involved in the protein-protein interactions. Fourthly, the N-terminal segment possesses at least two putative nuclear localisation signals (Dingwall and Laskey, 1991), implying that PfLAMMER may be a nuclear protein. Fifthly, there are 14 copies of a tetramer DKYD repeats at residue positions 307 to 482. Finally, in contrast to those of the known kinases in the LAMMER family, the N-terminal segment is the largest extension of LAMMERs yet described, placing PfLAMMER in a distinct category within the LAMMER family. The kinase catalytic domain of PfLAMMER is composed of 334 residues and contains all 11 conserved subdomains of the protein kinase family (Hanks et al., 1988). It has almost all of the characteristic features of a kinase including 15 invariant residues, the ATP-binding motif and the catalytic motif (Hanks et al., 1988; Knighton et al., 1991; DeBondt et al., 1993). The conserved sequences (HTDLKPENIL) in subdomain VI and (EHLAMMESII) in subdomain X are characteristic for LAMMER kinases (Yun et al., 1994). However, a highly conserved R residue in the LAMMER motif is replaced by an S residue in PfLAMMER. A similar change was also found in mammalian CLK3 kinases (Hanes et al., 1994; Becker et al., 1996; Nayler et al., 1997). By analogy with the crystallised structure of two protein kinases (Knighton et al., 1991; DeBondt et al., 1993), the LAMMER motif lies in an $\alpha$-helix below the substrate-binding cleft, potentially allowing it to contact with substrates. Therefore, the R to S change in the LAMMER motif might affect substrate specificity of PfLAMMER. The sequences (DLKPEN) in subdomain VI indicate that PfLAMMER is a serine/threonine kinase (Hanks et al., 1988); however, the sequences (TRQYRAPE) in subdomain VIII suggest that PfLAMMER could also be a tyrosine kinase (Hanks et al., 1988; Howell et al., 1991). Indeed, many kinases in the LAMMER family have been demonstrated to exhibit a dual-specificity, phosphorylating not only serine and threonine but also tyrosine (Ben-David et al., 1991; Howell et al., 1991; Yun et al., 1994; Duncan et al., 1995; Lee et al., 1996; Sessa et al., 1996). Therefore, PfLAMMER may be a parasite dual-specificity protein kinase. It is worth noting that, up to date, no tyrosine kinase has been reported in P. falciparum.

To investigate the structural organisation of the Pflammer gene in the P. falciparum genome, clone 3D7A genomic DNA was digested with a number of restriction enzymes and analysed by Southern blotting. Hybridisation of the cdcI-II (see Fig. 1) probe revealed a single band in digests with BamHI, BclI, EcoRI, EcoRV, HincII and RsaI, and two bands in digest with AccI, consistent with the restriction map (Fig. 1a), suggesting strongly that Pflammer is encoded by a single copy gene in the parasite genome. However, hybridisation of the same blot with the PSK1-Sau3AI fragment detected two bands (one predominant band and the other faint) in digests with AccI, BclI, EcoRI, EcoRV and HincII, and three bands in digest with RsaI, contradictory to

![Fig. 3. Southern blot analysis of the Pflammer gene. Four micrograms of genomic DNA from P. falciparum clone 3D7A were digested with restriction enzymes, electrophoresed on a 1% agarose gel, transferred onto a nylon membrane, and probed with the cdcI-II fragment (a) and PSK1-Sau3AI fragment (b and c), respectively. The blot was washed either in a normal condition (at 56°C) (a and b) (Li et al., 1996) or in a higher condition (at 65°C) (c). Lanes 1–7 correspond to digests with AccI, BamHI, BclI, EcoRI, EcoRV, HincII and RsaI. The sizes of 1–Kb DNA markers (M) are given in kilobase pairs to the left.](image-url)
the membrane probed with the (18 and 28S) are indicated by arrows. a, b, c and d are autoradiographs of the Pfpkac P. results suggest that only one LAMMER kinase exists in diseased to radiolabeled probes. The positions of P. falciparum rRNA subunits (18 and 28S) are indicated by arrows. a, b, c and d are autoradiographs of the membrane probed with the Pflammer gene (cdc1-II and PSK1-Sau3AI), the Pfpkac gene (Li and Cox, 2000) and the Pfpp-α gene (Li and Baker, 1998), and exposed for 44, 1 and 24 h, respectively. *Pflammer* detected a transcript of approximately 3800 nucleotides in the sexual stage. Pfpkac hybridised with a band of approximately 1800 nucleotides in the asexual stage and Pfpp-α hybridised with a band of approximately 3900 nucleotides in the sexual stage.

In order to obtain some information on how *Pflammer* mRNA levels are regulated during parasite development and differentiation, a northern blot containing equal quantities of total RNA prepared from cultures enriched in stage III to stage V gametocytes and from mixed asexual erythrocytic stages was probed with the cdc1-II clone. A single transcript of approximately 3800 nucleotides in size was detected only in the lane containing the sexual stage RNA, migrating just behind of the 28S rRNA bands (Fig. 4a). This was also confirmed by hybridising the same blot with the PSK1-Sau3AI fragment (Fig. 4b). The result is in sharp contrast to the lammer genes of humans, mice and insects, for which multiple transcripts were detected due to the alternative and incomplete splicings (Howell et al., 1991; Johnson and Smith, 1991; Hanes et al., 1994; Yun et al., 1994; Duncan et al., 1995; Nayler et al., 1997). As internal controls for hybridisation to the asexual and sexual stage mRNA, the same blot was hybridised with Pfpkac, an asexual stage-specific gene (Li and Cox, 2000), and Pfpp-α, a sexual stage-specific gene (Li and Baker, 1998) (Fig. 4c,d). The result suggests that PfLAMMER is involved in sexual stage-specific events. In rodent, it has been shown that mCLK2, mCLK3, mCLK4 and rCLK3 are predominantly expressed in testis, indicating that these kinases regulate a testicular function (Becker et al., 1996; Nayler et al., 1997). In *Arabidopsis*, the AFC1 kinase gene is able to complement yeast signal transduction mutants via activation of the transcription factor STE12 (Bender and Fink, 1994). In *Drosophila*, DOA is critical to the development of the fly embryo and affects eye differentiation and sex determination by phosphorylation of the SR proteins and consequently regulation of alternative pre-mRNA splicing (Yun et al., 1994; Du et al., 1998). In mammalian cells, over expression of the CLK1 kinase in PC12 cells seems to initiate their differentiation into neural derivatives possibly through activation of a mitogen-activated protein kinase cascade (Myers et al., 1994). The mouse CLKs have been demonstrated to bind to and phosphorylate SR-rich mRNA splicing factors (Colwill et al., 1996; Nayler et al., 1997). Overexpression of CLK1 in COS cells leads to the subcellular redistribution of the SR-proteins, and to alternations in mRNA splicing in vivo (Colwill et al., 1996; Duncan et al., 1997). Recently, it has been shown that both mCLK1 and hCLK2 can phosphorylate and activate the tyrosine phosphatases PTP-1B and YPTP1 in vitro (Moeslein et al., 1999). By analogy, therefore, we speculate that PfLAMMER may play an important role in signal transduction pathway(s) involved in the sexual differentiation of the parasite perhaps by phosphorylation of the sexual stage-specific mRNA splicing factors and/or protein phosphatases.

Acknowledgements

This work was supported in part by funds from the Wellcome Trust. J.-L. Li was supported by the Sino-British Friendship Scholarship Scheme (SBFSS), the Great Britain-China Educational Trust and Henry Lester Trust.

References


Bender, J., Fink, G.R., 1994. AFC1, a LAMMER kinase from *Arabidopsis*.


