Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular & Biochemical Parasitology 141 (2005) 179–189
MOLECULAR & BIOCHEMICAL PARASITOLOGY

A role for poly(dA)poly(dT) tracts in directing activity of the *Plasmodium falciparum* calmodulin gene promoter

Hannah E.J. Polson [1], Michael J. Blackman *

Division of Parasitology, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

Received 20 August 2004; received in revised form 8 February 2005; accepted 13 February 2005
Available online 23 March 2005

Abstract

Expression of the *Plasmodium falciparum* calmodulin gene (*pfcam*) is developmentally regulated throughout the blood-stage cycle. The promoter lies within ∼1 kb of intergenic sequence that separates the *pfcam* open reading frame (ORF) from an upstream inverted ORF encoding a product homologous to the co-chaperone STI1. Using the oligo-capping method, which selectively reverse-transcribes cDNA from only full-length, capped transcript, we have mapped multiple transcription-initiation sites for both genes. Transcription of the *pfSTI1* gene initiates over a 150 bp region centred ∼350 bp upstream of the ORF. The *pfcam* transcription start sites cluster into four ∼30 bp regions lying within 180 bp upstream of the *pfcam* ORF, generating transcripts with 5′ untranslated regions (UTR) of 3–173 nucleotides in length. Remarkably, splicing was found to be related to UTR length, with apparent preferential splicing of longer transcripts. Activity of the *pfcam* promoter diminished in a linear fashion to undetectable levels upon step-wise removal of sequence between 625 and 230 bp upstream of the start ATG. Electromobility-shift assays demonstrated nuclear factor binding to eight oligonucleotide probes spanning 657 bp of the *pfcam* ORF proximal upstream sequence. The degree of binding correlated with the density of poly(dA)poly(dT) tracts within the probes, and in all cases could be inhibited by excess synthetic poly(dA)poly(dT), but not by poly(dAdT)poly(dAdT). The multiple transcription-initiation sites of both *pfSTI1* and *pfcam* genes lie just downstream of 25 bp-long poly(dA)poly(dT) tracts, and the intergenic region contains over 20 poly(dA)poly(dT) tracts of 4 bp or more. Our results suggest that the basal *pfcam* promoter is situated between approximately −300 and −230 bp upstream of the *pfcam* ORF and that the *P. falciparum* transcription-initiation complex has a low degree of sequence-specificity for the sites of initiation but preferentially acts downstream of long poly(dA)poly(dT) tracts.

Keywords: *Plasmodium*; Promoter; Calmodulin; Transcription; Poly(dA)poly(dT); Transfection

1. Introduction

Gene transcription in *Plasmodium falciparum* is developmentally regulated as the parasite progresses through its multi-stage life cycle. Each major morphological stage is accompanied by distinct patterns of mRNA expression and protein synthesis, and within the asexual erythrocytic cycle numerous genes have been shown to be subject to temporal variation in expression [1–5]. Several housekeeping genes that are often constitutively expressed in eukaryotes also show developmentally regulated expression profiles; these include genes encoding ribosomal RNAs [6], components of the cytoskeleton [7,8] and calmodulin [9]. Large-scale microarray studies of gene expression in the erythrocytic cycle have shown widespread regulation of steady-state mRNA levels [10], and direct evidence of transcriptional control of gene expression has been obtained by nuclear run-on analysis, e.g. [4,11]. All of the above data point to the existence of highly dynamic mechanisms controlling differential gene expression in all life stages of *P. falciparum*. The development of transfection technology in *Plasmodium* [12,13] has allowed

Abbreviations: CAT, chloramphenicol acetyl transferase; CNE, crude nuclear extracts; EMSA, electromobility-shift assays; TAP, tobacco acid pyrophosphatase; USS, upstream sequence

* Corresponding author. Tel.: +44 208 816 2127; fax: +44 208 816 2730.
E-mail address: mblackm@nimr.mrc.ac.uk (M.J. Blackman).
[1] Present address: Parasites Molecular Immunology Unit, Metchnikoff Building, No. 67, Institut Pasteur, 25 Rue Du Dr Roux, 75015 Paris, France.

0166-6851/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2005.02.008

direct experimental approaches to be used to demonstrate that *Plasmodium* gene promoter structure follows the common eukaryotic bipartite model [14], comprising a basal promoter region capable of driving low-level gene expression, and upstream *cis*-acting elements involved in gene-specific regulation events [15–19]. Several different *cis*-acting elements have been identified, some of which have been shown to mediate nuclear factor-binding [5,15,18,20]. A TATA-box-binding protein has been described [21] and the largest subunit of RNA polymerase II has been identified and investigated at the primary sequence level [22]. However, progress in identifying DNA sequence elements that could be involved in directing transcription-initiation events has been slow, probably for two main reasons. Firstly, *P. falciparum* intergenic sequences are often over 90% A + T-rich and are generally composed of highly repetitive sequence or long homopolymeric adenosine and thymidine (poly(dA)poly(dT)) tracts with few or no distinguishing elements identifiable within the putative promoter regions [5,16,23,24]. Secondly, transcription initiation appears generally to occur at multiple sites and often over a large area of sequence [25].

Previous work by others has shown that transcription of the *P. falciparum* calmodulin gene (*pfcam*) is temporally regulated over the course of the blood-stage cycle. Rojas and Wasserman [9] showed that calmodulin steady-state mRNA levels are dramatically up-regulated from 28 h post-invasion, peak at 40 h, then decrease over the last hours of the asexual erythrocytic cycle. Low-level protein expression was observed in the early ring stage parasites and increased exponentially in abundance from 30 to 48 h. The *pfcam* upstream sequence (USS) has been partially characterised; a major transcriptional start site was mapped to −62 bp relative to the start ATG [26], and the core promoter region was delineated by Crabb and Cowman [17] to lie within 625 bp of the proximal 5′ flanking sequence. The *pfcam* promoter is one of very few to be used to date in transgene expression studies in *P. falciparum*, so its further characterisation will have important practical uses as well as increasing our broader understanding of transcriptional regulation in *P. falciparum*. Here, we show that initiation of transcription of the *pfcam* gene occurs at multiple sites, that correct splicing of pre-mRNA is dependent upon the site of transcription initiation, and that nuclear factor-binding and promoter activity may involve long poly(dA)poly(dT) tracts within the basal promoter region and the USS.

2. Materials and methods

2.1. Parasite culture and transfection

P. falciparum clone 3D7 parasites were cultured as described by Trager and Jensen [27]. Parasites were synchronised by rounds of Percoll-enrichment of late-stage schizonts, addition to fresh red blood cells (RBC) to allow invasion to occur for 4–6 h, then sorbitol treatment to lyse residual schizonts [28]. Transfection was performed as described by Wu et al. [12] using 100 μg of MAXIprep (QIAGEN)-purified DNA unless stated otherwise. Preliminary time-course experiments indicated that optimal levels of luciferase and chloramphenicol acetyl transferase (CAT) reporter activity were obtained in cultures harvested at 22 h and 46 h post-transfection respectively, and so these time-points were used routinely henceforth. Harvested parasite cultures were resuspended in 100 μL of 0.15% (w/v) saponin, incubated at room temperature (RT) for 2 min and centrifuged at 14,000 × *g* for 10 min at RT. The supernatant was removed and the parasite pellets snap frozen on dry ice. Samples were stored at −70 °C and assayed for reporter activity within 1 week.

2.2. Full-length cDNA production and sequencing

Total RNA was purified from Percoll-enriched schizonts as previously described [29]. Full-length cDNA was selectively amplified using the 'oligo-capping' or RNA ligase-mediated rapid amplification of 5′ cDNA ends (RLM–RACE) method, originally described by Sugano and co-workers [25,30–32], using an RLM–RACE kit (Ambion). Total RNA was treated with calf intestinal phosphatase (CIP), then with tobacco acid pyrophosphatase (TAP) before ligation to a synthetic RNA adaptor. Control RNA samples were not treated with TAP but were otherwise treated identically to experimental samples. Random primers were annealed to the modified RNA, and cDNA produced using reverse transcriptase. The 5′ ends of gene-specific clones were then amplified by nested PCR using two RNA adapter-specific forward primers: (outer, 5′-GCT GAT GGC GAT GAA TGA ACA CTG-3′ or inner, 5′-CGC GGA TCC GAC ACT CGT TTG CTG GCT TTG ATG-3′) and one or more of the following different sets of gene-specific reverse primers: (a) 5′-CTC AAT CCA TGT TCT GCA CAC CC-3′; (b) 5′-CCT AAT ATA CCC TTT AGG CCA ATC C-3′; (c) 5′-GAA AAG GAA CAA TTA CAT ATT ATA AAC AAT G-3′; (d) 5′-CAT ATA TAC ATA AGA AAT AAA AAG ATT CAA TAG G-3′; (e) 5′-AGG TTA GAA ATT CGG GAA AAT CGA TCG-3′; (f) 5′-ACC ATC TCT ATC AAA AAC TCC GGA AGG-3′. Products were cloned into *pMOS*Blue (Amersham Pharmacia Biotech) and analysed by automated sequencing using an ABI PRISM 377 DNA sequencer. Sequencing reactions were performed using the T7-promoter primer (5′-CTA ATA CGA CTC ACT ATA GGG A-3′) and U20 primer (5′-GGG TTT TCC CAG TCA CGA CG-3′), which hybridise to sequences flanking the cloning site in *pMOS*Blue. Sequence data were analysed using AutoAssembler™ sequence analysis software (Applied Biosciences).

2.3. Deletion mutagenesis and preparation of constructs for promoter analysis

Construct pHEP.luc was a kind gift of Dr. Pradeep Patnaik (present address: William Paterson University, Wayne, NJ, USA). Briefly, it contains the firefly luciferase coding