PfSR1 controls alternative splicing and steady-state RNA levels in *Plasmodium falciparum* through preferential recognition of specific RNA motifs

Shiri Eshar,1 Lindsey Altenhofen,2,3 Alona Rabner,4 Phil Ross,3 Yair Fastman,1 Yael Mandel-Gutfreund,4 Rotem Karni,5 Manuel Llinás2,3 and Ron Dzikowski1*

1Department of Microbiology and Molecular Genetics, The Kuvin Center for the Study of Infectious and Tropical Diseases and 5Department of Biochemistry and Molecular Biology, The Institute for Medical Research Israel–Canada, The Hebrew University – Hadassah Medical School, Jerusalem, Israel.

2Department of Molecular Biology and Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA.

3Department of Biochemistry and Molecular Biology, Department of Chemistry and Center for Malaria Research, Pennsylvania State University, State College, PA 16802, USA.

4Department of Biology, Israel Institute of Technology-Technion, Haifa, Israel.

Summary

*Plasmodium* species have evolved complex biology to adapt to different hosts and changing environments throughout their life cycle. Remarkably, these adaptations are achieved by a relatively small genome. One way by which the parasite expands its proteome is through alternative splicing (AS). We recently identified PfSR1 as a *bona fide* Ser/Arg-rich (SR) protein that shuttles between the nucleus and cytoplasm and regulates AS in *Plasmodium falciparum*. Here we show that PfSR1 is localized adjacent to the Nuclear Pore Complex (NPC) clusters in the nucleus of early stage parasites. To identify the endogenous RNA targets of PfSR1, we adapted an inducible overexpression system for tagged PfSR1 and performed RNA immunoprecipitation followed by microarray analysis (RIP-chip) to recover and identify the endogenous RNA targets that bind PfSR1. Bioinformatic analysis of these RNAs revealed common sequence motifs potentially recognized by PfSR1. RNA-EMSAs show that PfSR1 preferentially binds RNA molecules containing these motifs. Interestingly, we find that PfSR1 not only regulates AS but also the steady-state levels of mRNAs containing these motifs *in vivo*.

Introduction

Malaria is one of the most devastating infectious diseases, causing over 200 million clinical episodes and hundreds of thousands of deaths every year worldwide (WHO, 2012). *Plasmodium falciparum* is the causative agent of the deadliest form of human malaria. This protozoan parasite has a complex life cycle alternating between anopheline mosquitoes and the human host where it causes malaria upon invasion and replication within the human red blood cells. The complex life cycle of these parasites requires adaptations to different environments in different hosts and cell types, including red blood cells, mosquitoes and the liver, and has been found to be associated with tight regulation of gene expression (Bozdech et al., 2003). Nevertheless, this remarkable developmental complexity is achieved with a rather limited number of genes. One strategy by which eukaryotes can expand their proteome is through alternative splicing (AS) of pre-mRNAs. For *P. falciparum*, a limited number of AS events have previously been reported (Knapp et al., 1991; Muhia et al., 2003; Singh et al., 2004; Fonager et al., 2007; Iriko et al., 2009). Recent studies using whole genome approaches have enabled the detection of hundreds of AS events indicating that in blood stage parasites transcripts from approximately 5% of the genes undergo AS (Otto et al., 2010; Lopez-Barragan et al., 2011; Sorber et al., 2011). These reports suggest that AS plays an important role in post-transcriptional regulation of gene expression in *Plasmodium* parasites.

SR proteins are known to be key regulators of AS in eukaryotes. This family of RNA-binding proteins contain Ser/Arg-rich (SR) domains and a well-conserved RNA recognition motif (RRM) domains and are implicated in multiple roles in constitutive and AS (Black, 2003; Long and Caceres, 2009; Shepard and Hertel, 2009). Although the mechanisms by which SR proteins regulate AS and influence the choice of a particular splice site is not fully understood, AS events in *Plasmodium* parasites are often conserved across species, suggesting that AS is a stable feature in this parasite.
understood, a few models have been proposed suggesting that AS is regulated through the interplay between SR proteins that enhance splicing and others that antagonize it. In general, SR proteins are thought to enhance splicing by binding to specific RNA elements known as Exonic/Intronic Splicing Enhancers, whereas other proteins inhibit splicing by binding to RNA elements called Exonic/Intronic Splicing Silencers (Astr et al., 2004; Wang and Burge, 2008; Long and Caceres, 2009). In addition to their role as splicing factors, SR proteins have been implicated in other aspects of RNA metabolism including transcription as well the regulation of mRNA export, stability and translation (Shepard and Hertel, 2009; Zhong et al., 2009; Ji et al., 2013).

We recently identified a number of putative SR-like proteins in P. falciparum and demonstrated that PISR1 (PF3D7_0517300; PFE0865c) is a bona fide SR protein that functions as an AS factor in P. falciparum and that it is essential for parasite proliferation in human red blood cells (RBCs) (Eshar et al., 2012). Interestingly, PISR1 can rescue the splicing activity of splicing-deficient S100 extracts of mammalian HeLa cells and can function as an AS factor in mammalian minigene systems similar to its human homologue SRSF1 (Eshar et al., 2012). In vitro assays have shown that PISR1 can bind long RNA molecules in vitro (Dixit et al., 2010); however, the RNA molecules that are bound by PISR1 in vivo remain unknown. We hypothesized that by identifying endogenous RNA targets of PISR1, it would be possible to define common sequence motifs bound by PISR1 and discover specific splicing elements in endogenous P. falciparum RNA molecules. Here we adapted an inducible expression system to express epitope tagged PISR1, allowing us to perform RNA immunoprecipitation followed by microarray analysis (RIP-chip) to identify endogenous RNA targets that are bound by PISR1. Comparative analysis of these RNAs revealed common motifs found to be specific binding sites of PISR1. We further demonstrate that PISR1 can regulate AS of RNAs containing these motifs. We also found that, in addition to its role in AS regulation, PISR1 binding of RNA molecules containing these motifs influences their steady-state levels. This suggests that PISR1 influences RNA metabolism in P. falciparum through several different mechanisms.

Results

Inducible expression of PISR1-GFP maintains its cellular localization

We previously showed that stable overexpression of PISR1 influences parasite growth and gene expression (Eshar et al., 2012). We were therefore interested in creating a regulatable ectopic expression system for PISR1 to circumvent these issues. To do this, we adapted a recently developed methodology to manipulate protein levels that uses protein fusions to a de-stabilization domain in which the protein fails to fold properly in the absence of a stabilizing ligand (Muralidharan et al., 2011). Episomally expressed PISR1 was fused to the dihydrofolate reductase (DHFR) degradation domain (DDD), which targets the fusion protein for proteasome degradation. However, the DDD protein is stabilized by supplementation with the folate analog trimethoprim (TMP) in the culture media (Iwamoto et al., 2010). The DDD domain was also fused to GFP and an HA (haemoglutinin) epitope tag (Fig. 1A), thus enabling cellular localization and affinity isolation of the protein from transfected parasites (Muralidharan et al., 2011). The PISR1-GFP-DDD expression vector was transfected into dhfr-resistant NF54 transgenic line (Salazar et al., 2012) followed by selection for parasites that stably carry these episomes using blastcidin. These parasites were then grown on increasing concentration of TMP to induce dose-dependent overexpression of the stabilized episomal PISR1-GFP-DDD (Fig. 1B). Encouraged by the ability to induce functional ectopic expression using TMP, we measured whether PISR1-GFP-DDD maintained the cellular dynamics of the endogenous PISR1. We previously showed that PISR1 shuttles between the nucleus and the cytoplasm during intraerythrocytic developmental cycle (IDC) (Eshar et al., 2012), which is similar to what has been reported for several SR proteins in higher eukaryotes, including SRSF1, the mammalian homologue to PISR1 (Caceres et al., 1997). We performed immuno-fluorescence assays (IFAs) as well as in vivo imaging and showed that in early ring stages PISR1-GFP-DDD localizes to the nucleus in foci that are mainly located at the nuclear periphery. This pattern changes later in the IDC at the trophozoite and schizont stages, when PISR1-GFP-DDD can be detected in both the nucleus and in the cytoplasm (Fig. 1C and Fig. S1). Overall, this approach enabled expression of tagged PISR1 in an inducible manner and demonstrated that the GFP-DDD tagged PISR1 maintains its native cellular localization.

PISR1 binds a subset of endogenous Plasmodium RNA targets

PISR1 contains two RRM domains and was previously shown to bind RNA molecules in vitro (Dixit et al., 2010). To identify the endogenous P. falciparum RNAs bound by PISR1 during IDC, we used the DDD stabilized PISR1 parasite line described above to induce ectopic expression of the tagged PISR1 and performed RIP-chip (Fig. 2A). PISR1 protein was stabilized by adding 5 μM TMP to the culture media for 24 h before performing immunoprecipitation (IP) in unsynchronized cultures containing a similar
were significantly enriched (i.e. at least 3 standard deviations above average), in the IPs of the transgenic line versus the controls (Table S1). Some of these transcripts were also validated by reverse transcription polymerase chain reaction (RT-PCR) (Fig. 2D and Fig. S2), and their predicted function and exon numbers are presented in Fig. 2E, Fig. S3 and Table S1. These results suggest that PfSR1 specifically binds a subset of transcripts in blood stage parasites.

RNAs that are bound by PfSR1 share common binding motifs

In order determine if the RNAs that were enriched in the PfSR1-GFP-DDD IP share common binding motifs, the enriched probe sequences and their flanking regions were analyzed using the motif prediction algorithms DRIMust (Leibovich et al., 2013), and an in-house modified DRIMust algorithm. Our analysis identified two RNA motifs, which we named SR1 Binding Motif 1 and 2 (SBM-1 and SBM-2) (Fig. 3A and B respectively). Interestingly, SBM-1 shows high similarity to the purine-rich motif that has been identified as the binding site for human SRSF1 (Sanford et al., 2009; Janes et al., 2011; Clery et al., 2013). To demonstrate that these motifs directly bind PfSR1, we performed electrophoretic mobility shift assays using recombinant PfSR1 (PfSR1-His) and radiolabeled RNA probes (RNA-EMSA) containing either SBM-1 or SBM-2. Incubation of the purified recombinant PfSR1-His with the radiolabeled SBM-1 probe results in a clear shift in gel migration indicative of the formation of a protein-RNA complex (Fig. 3C). To examine if PfSR1 preferentially binds to SBM-1, we performed competition assays using either the unlabeled SBM-1 probe or an unrelated RNA sequence as alternative competitors. We found that the unlabeled SBM-1 probe interfered with the formation of the PfSR1-His–SBM-1 complex in a dose-dependent manner; however, the formation of this complex was less affected by competition using unrelated RNA probes (Fig. 3C and Fig. S4A). Similarly, PfSR1-His preferentially binds the SBM-2 motif compared with unrelated RNA molecules that did not contain the motif (Fig. 3D and Fig. S4B). These data are compatible with the known promiscuous in vitro binding of RNAs by SR proteins (Liu et al., 1998) (compared, e.g., with binding of transcription factors to DNA), which is reflected in the higher binding affinity to RNA molecules containing their preferred cis-regulatory binding elements.

To determine whether PfSR1 preferentially binds one of the two motifs, SBM-2 was radiolabeled and competed for PfSR1-His binding using either unlabeled SBM-1 or SBM-2 as competitors. We found that when the unlabeled SBM-1 was used as a competitor no complex was formed between SBM-2 and PfSR1-His, indicating that SBM-1 binds PfSR1-His with higher affinity than SBM-2 (Fig. 3E and Fig.
Fig. 2. PISR1 binds endogenous RNA in vivo. RNA immunoprecipitation followed by microarray analyses (RIP-chip) reveals RNA molecules that are bound by PISR1 in vivo.

A. Schematic of the RIP-chip experimental procedure. PISR1-GFP expression was induced using 5 μM TMP in NF54 parasites carrying the PISR1-GFP-DDD vector. Complexes of RNA and tagged PISR1 were precipitated using anti-GFP antibodies and protein G beads. RNA was extracted and analyzed by microarray compared with RIP of untransfected NF54 parasites as a negative control.

B. Western blot analysis of the IP samples using anti-GFP antibodies. Lane 1: IP on NF54 parasites used as a negative control; lane 2: IP on parasites that carried the PISR1-GFP-DDD plasmid and grown on regular media; lane 3: IP on parasites that carried the PISR1-GFP-DDD plasmid and grown on 5 μM TMP; Lane 4: IP using anti-myc antibodies as a negative control on parasites that carried the PISR1-GFP-DDD plasmid and grown on 5 μM TMP.

C. Microarray analysis of RNAs that were enriched by PISR1-GFP-IP. Histogram representing the frequencies in the differences between the log2 ratios of probes’ signal in the RIP-chip of PISR1-GFP-DDD compared with their signal in the RIP-chip of untransfected NF54 parasites that were used as a negative control. Probes that were enriched by at least 3 STDV above average are marked by black box and presented in Table S1.

D. Validation of the RNA-IP of representative genes (PF3D7_1412600, PF3D7_0826100 and PF3D7_0909700 respectively), using RT-PCR on the PISR1-GFP-IP and control NF54-IP samples.

E. Pie chart of putative gene annotation of the RNA molecules that were specifically enriched in the PISR1-GFP-IP versus control IP. The number of genes in each group are presented. Gene annotations were derived from PlasmoDB using annotated GO process.
We also performed mutational analyses to determine if the PfSR1 binding affinity was altered in the presence of RNAs with mutated SBM RNA motifs. RNA EMSAs with several competitor ligands containing mutated binding motifs showed that replacement of four nucleotides with uridines at different locations in the SBM-1 motif resulted in decreased PfSR1-His binding (Fig. 4A). However, mutants in which cytosine was introduced were still able to compete for PfSR1-His binding similar to what was shown for SRSF1 (Liu et al., 1998; 2000). Similarly, cytosine replacement throughout the SBM-2 motif had little effect on its ability to bind while replacement of adenine or uridine bases at different location indicated that the GUU sequence at the core of SBM-2 is important for PfSR1-His binding (Fig. S5). Interestingly, we also found that the flanking regions of SBM-1 also influence its ability to bind PfSR1-His (Fig. 4B). Altogether, these results indicate that PfSR1 preferentially binds RNA molecules containing the specific SBM sequence motifs.

PfSR1 regulates AS as well as steady-state mRNA levels of endogenous RNA targets containing the SBM motifs in vivo

Having demonstrating that PfSR1 binds to RNA molecules that contain specific motifs in vivo, we further investigated...
its function on the metabolism of these RNAs. We have previously shown that PfSR1 regulates AS of a few *P. falciparum* genes (Eshar et al., 2012) and had hypothesized that the AS pattern of transcripts containing these motifs could be altered when PfSR1 is overexpressed. For example, ectopic expression of PfSR1 changes the splicing pattern of the multi-exon gene PF3D7_0509700, which contains the SBM-1 motif and several repeats of SBM-2 (Fig. 5A). To link the presence of these motifs to alterations in AS, we re-used the RNA samples from the PfSR1-GFP-DDD immunoprecipitation described above and performed RT-PCR with primers specific to the first and fourth exons of PF3D7_0509700. We were able to detect PF3D7_0509700 only from the PfSR1-GFP-DDD-specific IP samples and not from IPs of the un-transfected NF54 parasites (Fig. 5B), indicating that PfSR1-GFP-DDD specifically interacts with this RNA in vivo. In addition, we show that ectopic expression of PfSR1-GFP-DDD affects the splicing pattern of PF3D7_0509700, favoring the production of a 485 bp splice product (Fig 5C and D). These results suggest that PfSR1 binds the PF3D7_0509700 RNA and influences its splicing pattern.

Our RIP-chip analysis allowed us to define the first set of endogenous *P. falciparum* RNAs that interact with PfSR1. Interestingly, some of these transcripts have no introns, suggesting that PfSR1 may be involved in other aspects of RNA metabolism, such as RNA stabilization, in addition to its role as an AS factor. To test this, we first performed a genome-wide search for SBM-1 and 2 motifs and found that these motifs are particularly enriched within exons (96.6%) and in many genes that have only one or two exons. In addition, although longer transcripts showed a higher motif occurrence, normalized motif hits per transcript length shows a negative correlation with...
transcript length. This uneven distribution throughout the genome indicates that these motifs are enriched in a specific subset of genes. (Fig. 6A and Table S2).

One example of a single intron gene that interacts with PISR1 is the well-known gametocyte-associated gene PF3D7_0509700 (Pf11-1) (Fig. 6B), which is comprised of a very long exon of approximately 27 kb and contains hundreds of SBM motif repeats. Using RIP followed by RT-PCR, we confirmed our previous RIP-chip analysis and demonstrate that the RNA of PF3D7_0509700 interacts with PISR1-GFP-DDD in vivo (Fig. 6C). This gene was suggested to encode a gametocyte specific protein with a role in gametogenesis (Scherf et al., 1992). Indeed, RT-PCR of blood stage NF54 parasites failed to detect expression of this gene, similar to what has been previously observed (data not shown). However, in parasites in which PISR1 was ectopically expressed (Eshar et al., 2012), we were able to detect its steady state transcript (Fig. 6D and E), whereas transcription of most gametocyte marker genes was not elevated (Figs S6 and S7). To support our observation that PISR1 is involved in regulating steady-state transcript levels of genes containing

Fig. 5. PISR1 binds in vivo to endogenous RNAs containing SBM motifs and regulate AS. 
A. Schematic gene structure of PF3D7_0509700 including the location of SBM-1 (red) and SBM-2 (grey). Exons are shown in boxes and introns in lines. P1 and P2 represent the position of the primers that were used to detect PF3D7_0509700 isoforms.
B. PISR1 binds the PF3D7_0509700 transcript in vivo. RT-PCR on RNA extracted from the IP samples of tightly synchronized late stage (40 hpi) of PISR1-GFP-DDD and NF54 parasites using PF3D7_0509700 specific primers. The 500 base pair size is arrowed.
C. PISR1 influences the choice for alternative splice sites of PF3D7_0509700 in vivo. RT-PCR using P1 & 2 primers on RNA extracted from NF54 parasites (lane 1) and parasites that express mock-GFP growing on TMP and blasticidin (lane 2) or parasites expressing PISR1-GFP-DDD growing on TMP and blasticidin (lane 3). Schematic of the amplified isoforms is shown on the left.
D. Quantification of changes in the relative abundance of each isoform following inducible ectopic expression of PISR1. The relative copy number of each of the three isoforms were measured by RT-qPCR and calculated relative to the housekeeping gene arginyl-tRNA synthetase (PFL0900c). The ratio of the RCN of each isoform out of the sum of the RCN of all 3 isoforms in each sample is presented as relative abundance. All values presented are the average of triplicates. Error bars represent standard deviation. Significant differences between spliced isoform in parasite overexpressing PISR1 and those transfected with mock plasmid were tested by unpaired *t*-test (*P < 0.05; **P < 0.01).
SBM-1 and SBM-2, we screened for additional examples of genes containing (i) a low number of exons and (ii) numerous SBM recognition motifs, which have elevated steady-state transcript levels when PfSR1 is overexpressed. In addition to Pf11-1, we found that PfSR1 is involved in controlling the RNA level of several other genes that contain numerous SBM-1 and SBM-2 repeats (Fig. S7). These results suggest that PfSR1 may have a secondary role in controlling the steady state levels of a subset of RNAs in *P. falciparum*.

**PfSR1 is localized adjacent to the nuclear pore complexes**

While shuttling into the nuclei of mammalian cells, human SRSF1 was shown to be located in nuclear speckles and proposed to play a role in regulating the assembly of pre-mRNA processing factors in these speckles (Caceres et al., 1997; Tripathi et al., 2012). In addition, recent studies have coupled pre-mRNA processing by SR proteins with transcription (Ji et al., 2013). In *P. falciparum*, the regions of the nuclear periphery adjacent to the clusters of nuclear pore complexes (NPCs) are composed of loose genetic matter that enable transcription (Weiner et al., 2011; Dahan-Pasternak et al., 2013), supporting the gene gating hypothesis (Blobel, 1985). In view of the recent evidence coupling transcription and splicing in other eukaryotes, we tested whether PfSR1 colocalizes with the NPC clusters. By co-transfecting PfSR1-myc (Eshar et al., 2012) and PfSec13-GFP, an NPC marker (Dahan-Pasternak et al., 2013), into NF54 parasites, we could localize both proteins by IFA.
The nuclear localization of PfSR1 is a characteristic of eukaryotes. SR proteins are known to regulate splicing and AS is well conserved between these evolutionarily distant organisms. In recent years, with the advent of RNAseq and ChIPseq analyses, there is a growing appreciation of the complexity and importance of AS events in metazoan species. However, many components of the splicing machinery found in mammalian cells also to recognize and bind to the mammalian RNA sequences found on the minigenes (Eshar et al., 2012). Although it was previously shown that PfSR1 can bind RNA molecules in vitro (Dixit et al., 2010), to begin understanding how PfSR1 regulates AS in P. falciparum, it was necessary to determine which RNA sequences are specifically recognized in vivo and identify its cis-acting RNA elements.

Using inducible expression of tagged PfSR1 and RIP-chip analysis, we identified common motifs (SBM-1 & 2) found in a subset of RNA molecules that are bound by PfSR1 in vivo. We found that PfSR1 preferentially binds these motifs and regulates expression of genes containing these motifs through AS. Interestingly, SBM-1, which binds PfSR1 with greater affinity, shows high similarity to a GAAGAA motif that was predicted in silico to be one of the common ESEs found in human genes (Fairbrother et al., 2002). More recently, whole genome CLIP-seq analysis on human embryonic kidney cells identified a similar purine rich octamer motif containing a core of GAAGAA as the binding site of SRSF1 in vivo (Sanford et al., 2009; Janes et al., 2011). This motif is also similar to several SRSF1 recognition sites identified by in vitro binding SELEX experiments and mutation analyses of splicing enhancers (Caputi et al., 1994; Ramchatesingh et al., 1995; Tacke and Manley, 1995). Interestingly, nucleotide replacements in PfSR1-binding motifs with stretches of cytosine did not interfere with their ability to bind PfSR1 (Fig. 4) similar to what was shown for SRSF1 (Liu et al., 1998; 2000). These sequence similarities between SRSF1 ESE and PfSR1 binding sites, as well as the structural similarity in the RRM1 domains of these two SR proteins (Eshar et al., 2012), likely explain the ability of PfSR1 to function as an SR protein in mammalian cells and minigenes. It will be interesting in the future to create an experimental system using transgenic parasites that carry Plasmodium-based minigenes and determine whether these binding sites are functional ESEs.

Discussion

Alternative splicing of pre-mRNA is an important post-transcriptional mechanism by which eukaryotes regulate gene expression. In recent years, it has become clear that the extent of AS events in *P. falciparum* during its IDC is much more abundant than previously reported (Otto et al., 2010; Lopez-Barragan et al., 2011; Sorber et al., 2011). Nonetheless, our knowledge of AS events in other stages of the parasite’s life cycle, such as in the mosquito and liver stages, remains unknown. Several putative proteins have been identified as SR-like encoded in *P. falciparum* genome; however, to our knowledge, thus far, only PfSR1 was shown to be a bona fide SR protein that functions as an AS factor in vitro and in vivo (Eshar et al., 2012). PfSR1’s ability to influence the choice for AS sites in mammalian cells and on mammalian-based minigenes (Eshar et al., 2012) indicated that the mechanism by which it regulates AS is well conserved between these evolutionarily distant eukaryotes. SR proteins are known to regulate splicing and affect the choice for AS sites by binding to specific cis-RNA elements (ESEs and ISEs) and by interacting with other components of the splicing machinery that enhance splicing and inclusion of the adjacent exon in the spliced mRNA (Long and Caceres, 2009). The similar AS function of PISR1 and SRSF1 in mammalian-based experimental systems indicated that PISR1 was able to interact with other components of the splicing machinery found in mammalian cells and also to recognize and bind to the mammalian RNA sequences found on the minigenes (Eshar et al., 2012). Although it was previously shown that PISR1 can bind RNA molecules in vitro (Dixit et al., 2010), to begin understanding how PISR1 regulates AS in *P. falciparum*, it was necessary to determine which RNA sequences are specifically recognized in vivo and identify its cis-acting RNA elements.

Using inducible expression of tagged PISR1 and RIP-chip analysis, we identified common motifs (SBM-1 & 2) found in a subset of RNA molecules that are bound by PISR1 in vivo. We found that PISR1 preferentially binds these motifs and regulates expression of genes containing these motifs through AS. Interestingly, SBM-1, which binds PISR1 with greater affinity, shows high similarity to a GAAGAA motif that was predicted in silico to be one of the common ESEs found in human genes (Fairbrother et al., 2002). More recently, whole genome CLIP-seq analysis on human embryonic kidney cells identified a similar purine rich octamer motif containing a core of GAAGAA as the binding site of SRSF1 in vivo (Sanford et al., 2009; Janes et al., 2011). This motif is also similar to several SRSF1 recognition sites identified by in vitro binding SELEX experiments and mutation analyses of splicing enhancers (Caputi et al., 1994; Ramchatesingh et al., 1995; Tacke and Manley, 1995). Interestingly, nucleotide replacements in PISR1-binding motifs with stretches of cytosine did not interfere with their ability to bind PISR1 (Fig. 4) similar to what was shown for SRSF1 (Liu et al., 1998; 2000). These sequence similarities between SRSF1 ESE and PISR1 binding sites, as well as the structural similarity in the RRM1 domains of these two SR proteins (Eshar et al., 2012), likely explain the ability of PISR1 to function as an SR protein in mammalian cells and minigenes. It will be interesting in the future to create an experimental system using transgenic parasites that carry Plasmodium-based minigenes and determine whether these binding sites are functional ESEs. SR proteins are found in plants (Lopato et al., 1996; Lorkovic et al., 2000; Yoshimura et al., 2002), metazoan (Zahler et al., 1992) and in some unicellular...
eukaryotes that undergo AS (Gross et al., 1998). Other eukaryotes, such as *Saccharomyces cerevisiae* that use AS to control transcript levels rather than to expand proteome diversity (Kawashima et al., 2014), do not have ‘classic’ SR proteins. The presence of SR proteins in Apicomplexan parasites and the conservation of the GAAGAA motif as the binding site of PfSR1 in the overall AT rich genome of *P. falciparum* indicate that the regulatory roles of PfSR1 in RNA metabolism confer advantageous functions in *Plasmodium* biology that have been maintained throughout its evolution.

It has become widely accepted that, in addition to their role in catalyzing spliceosome assembly and regulating AS, SR proteins also participate in a wider range of functions in mRNA metabolism starting at the level of transcription, as well as functions downstream of splicing such as mRNA export, stability and translation (Long and Caceres, 2009; Shepard and Hertel, 2009; Zhong et al., 2009). Indeed, recent whole genome CLIP-seq experiments have demonstrated that SRSF1 binds *in vivo* to numerous GAAGAA binding sites that are found in functionally diverse intronless genes that are not processed by pre-mRNA splicing (Sanford et al., 2009). Similarly, our current findings show that in *P. falciparum*, PfSR1 is specifically bound to several RNA molecules containing a single exon that are not spliced. Furthermore, our genome wide *in silico* search identified the SBM motifs in numerous intronless genes. This may imply that PISR1 contributes to the regulation of RNA metabolism and gene expression not only through splicing and AS. We demonstrated that PfSR1 is involved in the regulation of steady state levels of a long exonic transcript of PF3D7_1038400, *(pf11-1)* which contains many repeats of the SBM binding motifs. Pf11-1 was implicated in gametogenesis (Scherf et al., 1992). In agreement with these previous observations, we did not detect the transcript of *pf11-1* in blood stages of wild-type NF54 parasites; however, in parasites ectopically expressing PfSR1, this transcript could be detected by RT-PCR. Our hypothesis is that PfISR1 may regulate the steady-state levels of these long transcripts by affecting their RNA stability. This activity has been demonstrated in higher eukaryotes where SR proteins are known to regulate polyadenylation of pre-mRNA (Lou et al., 1998) and thus contribute to stabilization and trafficking of the mature mRNA and requires further investigation in *Plasmodium*.

Another possible hypothesis is that PfISR1 could influence the expression of single-exon genes by promoting transcription. There is mounting evidence that a subset of SR proteins colocalize with active genes and play a direct role in transcriptional elongation (Zhong et al., 2009; Ji et al., 2013). In addition, certain SR proteins bind directly to the histone H3 tail (Loomis et al., 2009) and some associate with the phosphorylated C terminal domain of RNA polymerase II (Misteli and Spector, 1999). A model has been proposed by which SR proteins are recruited to the site of active transcription to facilitate transcriptional elongation, thus linking transcription to the initiation of RNA splicing (Zhong et al., 2009). Recently, SRSF2 was shown to play a direct role in transcriptional activation through binding to promoter associated nascent RNA (Ji et al., 2013). Although the role of PfSR1 in transcriptional regulation in *P. falciparum* is unknown, its nuclear localization indicates that such a role may be relevant. In the current study, we show that, at early stages of the IDC, when PISR1 is nuclear, it is found mainly at distinct foci adjacent to the cluster of the NPC. We previously showed that in *P. falciparum*, the region at the nuclear periphery adjacent to the NPC is composed mainly of loose euchromatic genetic matter that facilitates transcription (Weiner et al., 2011; Dahan-Pasternak et al., 2013). In mammalian cells, SR proteins, as well as other components of the splicing machinery, are concentrated in nuclear speckles of two kinds: interchromatin granules clusters that serve as storage for splicing factors, and perichromatin fibrils, which were suggested to be sites of actively transcribing genes and co-transcriptional splicing (Spector, 1996). The sub-nuclear localization of PfISR1 in the euchromatic region adjacent to the nuclear periphery may also indicate that the act of transcription is coupled with pre-mRNA splicing in *P. falciparum* as well. Furthermore, several SR proteins that shuttle between the nucleus and the cytoplasm were implicated in facilitating export of intronless as well as spliced mRNAs (Long and Caceres, 2009; Shepard and Hertel, 2009; Zhong et al., 2009). Similar to other SR proteins, PISR1 also shuttles between the nucleus and the cytoplasm during the IDC (Eshar et al., 2012). Therefore, PISR1’s location adjacent to the NPC in regions shown to be typically euchromatic in *P. falciparum* raises the possibility that efficient coupling between a gene’s transcription, pre-mRNA splicing and mRNA export have evolved in this parasite. This subnuclear localization is in agreement with the ‘gene gating’ hypothesis suggested years ago, which states that the non-random distribution of the NPC at the nuclear periphery and their specific interaction with euchromatic regions in the genome creates a gating organelle that allows efficient transport of mRNAs out of the nucleus (Blobel, 1985).

In the current study, we have applied RIP-chip under native conditions to gain the first insight into PISR1 binding sites *in vivo*. Interestingly, there is very little overlap between the transcripts that were shown to be bound PISR1 in our study and the genes reported recently to undergo AS in blood stages parasites (Otto et al., 2010; Lopez-Barragan et al., 2011; Sorber et al., 2011). One
possible explanation is that PISR1 has a more transient binding pattern when regulating AS and a more stable association when it is involved in other mechanisms of RNA metabolism. Therefore, it is possible that by performing RIP under native conditions and stringent analysis of transcript enrichment on the array, we gained insight on a limited subset of transcripts that are stably bound to PISR1. Using cross-linking methods such as CLIP-seq in future analyses will allow us to gain further insight into the genome-wide landscape of PISR1 in Plasmodium. Furthermore, there are additional predicted SR proteins in Plasmodium and each could be regulating a different subset of RNAs. A recent genome-wide study on the binding landscape of SRSF1 and SRSF2 revealed that splicing regulation in the mammalian genome is dependent on the orchestrated interaction of one SR protein with other SR proteins (Pandit et al., 2013). In summary, we have identified an important role for PISR1 in the regulation of RNA metabolism in the blood stages parasites. This work emphasizes the need for future investigation of the role of SR proteins throughout the entire P. falciparum cell cycle.

**Experimental procedure**

**Cell cultures and transfection**

All parasites used were derivatives of the NF54 parasite line. Parasites that were used to establish the inducible ectopic expression system contained genomic insertion of hDHFR cassette (Salazar et al., 2012). Parasites were cultivated at 5% hematocrit in RPMI 1640 medium, 0.5% Albumax II (Invitrogen), 0.25% sodium bicarbonate and 0.1 mg ml\(^{-1}\) gentamicin and incubated at 37°C in an atmosphere of 5% oxygen, 5% carbon dioxide and 90% nitrogen. Parasites were transfected as previously described (Deitsch et al., 2001). Briefly, 0.2 cm electroporation cuvettes were loaded with 0.175 ml of erythrocytes and 

\[\sim 100 \, \mu g \, \text{of plasmid DNA in incomplete cytomix solution.}\]

Stable transfectants were initially selected on 2 \(\mu g \, ml^{-1}\) blasticidin-S-hydrochloride (Invitrogen). For inducible ectopic expression of PISR1 non-synchronized parasite containing all stages (but no gametocytes) were split and exposed to increasing concentrations of TMP. For IPs non-synchronized cultures containing equal ratio of late and early stages were used. Both PISR1-GFP and hDHFR-containing NF54 controls were put on TMP 24 h prior to IP.

**DNA constructs**

*pfsr1* was cloned into the expression vector pHBIRH (Epp et al., 2008) and fused to a regulatable fluorescent affinity (RFA) as previously described (Muralidharan et al., 2011) to generate pHBISR1-GFP-DDD. The RFA tag was amplified using the primers GFP-DDDF-SpeI and GFP-DDDR-SacI (see Supplementary materials) and cloned using *SacI* and *SpeI* restriction sites. *pfsr1* was amplified using the primers PISR1F – PISR1R (Eshar et al., 2012) and cloned into the pHBIRH plasmid using *SalI* and *SpeI* restriction sites. PF3D7_0600400 was amplified using the primers PF3D7_0600400F and PF3D7_0600400R (see Supplementary materials) and cloned into pH-SEC13GFP-IDH (Dahan-Pasternak et al., 2013) using *PstI* and *XmaI*.

**Recombinant protein purification and Western blot analysis**

Expression and purification of recombinant PISR1-His and Western blot analyses was performed as previously described (Eshar et al., 2012). Immuno-detection was carried out using mouse-anti-GFP antibody (Roche Applied Science), followed by rabbit anti-mouse antibody conjugated to Horseradish Peroxidase (HRP) (Jackson, ImmunoResearch Laboratories) and developed in EZ/ECL solution (Israel Biological Industries).

**Fluorescence microscopy**

For live imaging parasites were washed once in PBS, stained for 20 min with Hoechst stain (Sigma), washed again in PBS and taken immediately to microscopic detection. Images were taken using Apochromat oil immersion objective using \(\times 100\) magnification (Olympus) on an Olympus IX71S8F microscope equipped with Exi Blue™ Fast camera (Qimaging). Z stack images were processed by 3D deconvolution using cellSens Dimension module with the advanced maximum likelihood estimation algorithm (ADVMLE).

**IFA**

Immuno-fluorescence assay was performed as described (Eshar et al., 2012) Primary antibodies were diluted in blocking solution as follows: mouse-anti-GFP (Roche Applied Science) 1:1000, rabbit-anti-myc 1:300 (CellScience), incubated for 1.5 h and washed three times in PBST. Secondary antibodies (anti-mouse-Alexa488 and anti-rabbit-AlexaCy3, Invitrogen) were diluted 1:250 and incubated for 1 h. After final washes, slides were sealed with ProLong Gold Reagent (Invitrogen). Imaging was performed as described above.

**RNA immuno-precipitation**

RNA immunoprecipitation was done as described before (Mair et al., 2006). Briefly, RBCs from 200 ml cultures at 5% parasitemia were lysed in saponin. Parasites were
washed twice in cold PBS and re-suspend in NET2 buffer (50 mM Tris-HCl, pH = 7.4, 150 mM NaCl, 1% NP-40, protease inhibitor, RNasin, 2 mM DTT). Lysates were incubated for 5 min on ice and centrifuged (14000 r.p.m., 5 min in 4°C). Supernatant was collected and incubated with 1.2 μg of anti-GFP (Roche) or anti-myc (Santa-Cruz) antibodies for 1 h. Protein-G sepharose beads (Santa-Cruz Biotechnology) were washed in NET-2 buffer and added to the lysates. Bound RNA was centrifuged, washed in NET-2 buffer and collected in TRIZOL LS Reagent® and purified on PureLink column (Invitrogen) according to manufacturer’s protocol. Half of each sample was collected in sample buffer and analyzed by western blot for the presence of the GFP-tagged PISR1.

Microarray analysis and motif search

RNA from three independent IPs at concentration of 7–10 ng μl⁻¹ in total volume of 40 μl was prepared for hybridization by indirect amino-allyl cDNA labeling as previously described (Painter et al., 2013) Samples were hybridized against a common 3D7 reference pool on custom Agilent 8 × 15K microarrays spotted with 60mer oligonucleotide probes (Kafaschak et al., 2012; GEO Platform GPL15130). The arrays were scanned using an Agilent G2505B Microarray Scanner (Agilent Technologies, Santa Clara, CA, USA), and images were analyzed using Agilent Feature Extractor Software Version 9.5. Validation of enriched probes was done by RT-PCR with specific primers (Table S1). The difference between the log2(Cy5/Cy3) ratio of PISR1-IP vs. ctrl-IP was calculated. Probes that were significantly enriched (3 standard deviations above average) were chosen for further analysis of specific RNA binding motifs. RNA binding motifs were searched in silico using two different softwares: DRIMust (Leibovich et al., 2013) and in-house modified version of DRIMust, which included secondary structure information in addition to the sequence data. In our search, we analyzed the sequence of the enriched microarray probe (60 bp) and the flanking genomic sequences of 60 bp upstream and downstream the probe. For both algorithms, the enriched probes (top of the list) were defined as the test sequences while the background included the sequences at the bottom of the list. DRIMust parameters were set to: sequence length: 5–10 bp and the statistical significance threshold: P-value < 0.001. The exact P-values found were ≤ 1.3E−15. and ≤ 6.18E−5 for the SBM-1 and SBM-2 motifs respectively.

Genome-wide motif search

A genome-wide motif search was performed using FIMO (v4.9.0) (Grant et al., 2011), and the respective position weight matrices for SBM1 and SBM2 obtained via DRIMust using a P-value thresholds of 1e−2 to 1e−4. In order to determine which motif hits overlapped with CDCs and intergenic regions, BEDTools (v2.20.1) (Quinlan and Hall, 2010) was used to calculate the intersection. Once the degree of overlap was calculated between motif hits and different genomic features, custom written python scripts (provided) utilizing the pandas (v0.14.0) module were used to calculate how many motif hits per gene were found, how well they correlated with transcript length (spearman coefficient) and which genes were found to be enriched with motifs SBM1 and SBM2 by looking three standard deviations above the mean number of normalized motif hits. All Plasmodium data sets and annotations were downloaded directly from v11.0 of PlasmoDB (http://plasmodb.org/common/downloads/release-11.0).

RNA EMSA

RNA ligands were commercially synthesized (BIONEER) and radiolabeled by end labeling of single-stranded oligonucleotides incubate with [γ−32P] dATP and T4 polynucleotide kinase (M0201S; BioLabs) according to the supplier’s instructions. RNA ligands sequences were identical to the endogenous sequences of PF3DT_0305500 (previously PFC0245c, SBM-1) and PF3DT_0903700 (previously PFI0180w, SBM-2). Probe sequences were as follow:

**SBM-1:** 5′-UAGAUGUCGAUGAAGAUGAACAAUGU-3′

**SBM-2:** 5′-CUGCGUAUGUAACCAUC-3′

Non-specific:

5′-GUAAUACAAUAAACAGUAGGGAGGG-3′

For gel shift assays a standard 20 μl binding reaction mixture contained: Tris-Cl 10 mM, MgCl₂ 1 mM, RNase inhibitor 10U, BSA 0.01 mg ml⁻¹, 5% glycerol, KCl 10 mM, DTT 1 mM, HEPES 14 mM and 25 fmol labeled RNA ligands. Competitive unlabeled ligands were added in increasing concentrations as indicated. Purified PISR1 was added last to ensure that the probe and its competitors had the same probability of binding. The reaction mixtures were incubated for 30 min at room temperature; then samples were loaded on a 0.8% agarose gel in TAE buffer [6.7 mM Tris-acetate, 3.3 mM sodium acetate, 1 mM EDTA (pH = 7.5)]. Electrophoresis was conducted at 100 V in a cold room for 1 h. Protein–RNA complexes were visualized and quantified by a Bio Imaging Analyzer (BAS1000; Fuji).

RNA extraction, cDNA synthesis, RT-PCR and real-time RT-qPCR

RNA extraction and cDNA synthesis was performed as described (Dzikowski et al., 2006). Briefly, RNA was extracted from synchronized parasite cultures at 36 hours

© 2015 John Wiley & Sons Ltd, Molecular Microbiology, 96, 1283–1297
post invasion. RNA was extracted with the TRIZOL LS Reagent® (Life Technologies) as described and purified on PureLink column (Invitrogen) according to manufacturer's protocol. Isolated RNA was then treated with recombinant Deoxynucleobase II® (Takara) to degrade contaminating gDNA. cDNA synthesis was performed from 750 ng total RNA with Primerscript RT reagent® (Takara) with oligo dT (Takara) as described by the manufacturer. RT-PCR was performed for 40 cycles with optimized temperature for each primers pair used. Transcript copy numbers were determined by RT-qPCR using the formula 2^−ΔΔCt as described in the Applied Biosystems User Bulletin 2 using NF54 gDNA as the calibrator. Specifically, relative copy number was calculated as 2 exponential negative [(Ct target gene in gDNA − Ct reference gene in cDNA) − (Ct target gene in gDNA − Ct target gene in gDNA)]. The relative copy number was calculated relative to the housekeeping gene arginyl-tRNA synthetase (PF3D7_1218600). Primer sequences for PF3D7_1038400 and PF3D7_0509700 are found in Table S2. All RT-qPCR assays were performed in triplicates.

Acknowledgements

This work was supported in part by a grant from the Israel Science Foundation (141/13) and the Minerva Stiftung, ARCHES award for Research Cooperation and High Excellence in Science. ML received support from the Centre for Quantitative Biology (P50GM071508). The plasmids used for cloning of the protein destabilization domain were kindly donated by Professor Dan Goldberg.

References


**Supporting information**

Additional supporting information may be found in the online version of this article at the publisher’s web-site.