Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 360 (2007) 163–172
BBRC
www.elsevier.com/locate/ybbrc

Molecular analysis of SmFes, a tyrosine kinase of *Schistosoma mansoni* orthologous to the members of the Fes/Fps/Fer family ☆

Fernanda Ludolf [a,c], Diana Bahia [a,*], Luiza F. Andrade [a,c], Alexandre Cousin [b], Monique Capron [b], Colette Dissous [b], Raymond J. Pierce [b], Guilherme Oliveira [a,c]

[a] *Centro de Pesquisas René Rachou (CPqRR)—FIOCRUZ, Av. Augusto de Lima 1715, Belo Horizonte, MG 30190-002, Brazil*
[b] *Inserm, U 547, Université Lille 2, Institut Pasteur de Lille, IFR 42, 1 rue du Professeur A. Calmette, 59019 Lille, France*
[c] *Programa de Pós-Graduação e Pesquisa, Santa Casa de Belo Horizonte, Av. Francisco Sales 1111 9° andar, Ala D, Belo Horizonte, MG 30150-221, Brazil*

Received 1 June 2007
Available online 14 June 2007

Abstract

A novel protein tyrosine kinase (PTK) was identified in *Schistosoma mansoni* and designated SmFes. SmFes exhibits the characteristic features of Fes/Fps/Fer (fes, feline sarcoma; fps, Fujinami poultry sarcoma; fer, fes related) PTKs, containing three coiled-coil regions, an SH2 (Src-homology-2) and a TK (tyrosine kinase catalytic) domain signature. SmFes is the first gene from the Fes/Fps/Fer family identified in *S. mansoni*, and is a single copy gene. Phylogenetic analyses revealed that SmFes is most closely related to its invertebrate orthologues. The assembly of the SmFes cDNA and genomic sequences indicated the presence of 18 introns in SmFes. Comparison of its genomic structure with those of human Fps/Fes and *Drosophila* Fps indicates that intron positions are conserved within the region encoding the kinase domain. Analysis of partial cDNA clones showed the presence of a 9 bp insertion at the 3′ end of exon 10, producing two different cDNA populations, pointed as an alternative splicing event. In addition, an allele of SmFes containing a 15 bp insertion was observed in the genomic sequence. Quantitative RT-PCR indicated that the overall transcription level of SmFes is rather low in all parasite developmental stages. Moreover, SmFes mRNA levels decrease progressively after cercarial transformation, consistent with a role for the corresponding protein in the early stages of infection.

Keywords: Fes/Fps/Fer; fps proto-oncogene protein; Protein tyrosine kinase; *Schistosoma mansoni*; Signal transduction; Alternative splicing

Schistosoma mansoni is one of three main species of platyhelminth parasite that cause schistosomiasis, a disease that affects over 200 million people, placing over 600 million individuals at risk [1]. The main control strategy involves the use of the only drug available for mass chemotherapy, Praziquantel, an effective, safe and single-dose drug [2]. However, despite advances, especially in genomics [3], the goals of developing new drugs and vaccines have not yet been reached and the eradication of transmission remains a long-term objective.

Schistosoma mansoni has a complex life-cycle including two hosts and six morphologically distinct forms and is characterized by its interaction with the environment, which stimulates physiological, morphological, and biochemical changes. For example, female worm maturation depends on pairing with the male, which induces mitotic activity in female reproductive organs [4,5]. The parasite is highly adapted and can survive in the human host for many decades [6]. Molecular factors produced by the parasitized organism (for a review see [7]) can be sensed and transduced into signals that mediate physiological adaptations of the parasite. In common with other developmental processes, schistosome infection of the host and subsequent

☆ The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint first authors.

* Corresponding author. Address: Rua Botucatu, 862, 6[0] andar, São Paulo, SP 04023-062, Brazil. Fax: +55 11 5571 1095.
E-mail address: dianabahia@hotmail.com (D. Bahia).

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.06.018

maturation require permanent sensing of the environment, communication between cells of the organism and between sexes, all of which involve signal transduction mechanisms. However, although a variety of signaling proteins have been identified in schistosomes [7], including cell surface receptor tyrosine kinases and cytosolic protein kinases [8], many components of the corresponding signaling cascades have yet to be characterized.

Protein tyrosine kinases (PTKs) are key molecules in a large variety of signal transduction pathways, being involved in developmental, differentiation and communication processes of cells [9]. Cellular PTKs are involved in eukaryotic pathways, controlling diverse cellular processes such as adhesion, migration, proliferation, differentiation, cytoskeletal alteration, and survival [10–12]. As well as providing insights into the mechanisms involved in these processes, the study of PTKs may provide new strategies for the development of new drugs, an approach intensively pursued in cancer research [13]. Three cytosolic TKs have so far been described in *S. mansoni*: TK5 is an orthologue of Fyn that is a member of the Src family [14]; TK4 is an orthologue of the Syk protein [15] and TK3 is an orthologue of the Src protein [16].

In the context of the EST program of Schistosome Genome Project [3], we identified a new PTK in *S. mansoni*, SmFes (GenBank Accession No.: AF515706), a putative member of the Fes/Fps/Fer family. Although the cellular functions of the Fes/Fps/Fer kinases are still largely unknown, they have been shown to be implicated in the regulation of cell–cell and cell–matrix interactions and cytoskeletal rearrangement (for review see [17]). In a parallel study [18], we determined that the SmFes protein is localized at the surface of infective larval stages, suggesting that it may play a role in signal transduction pathways involved in larval transformation after penetration into intermediate and definitive hosts.

Here, we report the overall structural conservation of SmFes compared to its orthologues and show that the gene structure is also partially conserved. We further describe the identification of both splicing and allelic isoforms and the level of SmFes mRNA expression throughout the life-cycle and following cercarial transformation.

Materials and methods

Parasites. A Puerto-Rican strain of *S. mansoni* was maintained in *Biomphalaria glabrata* snails and golden hamsters (*Mesocricetus auratus*). Cercariae were released from infected snails and harvested on ice. They were then washed three times by resuspension in 30 ml of Hank's Balanced Salt Solution (Gibco-BRL) in a corex tube (Corning) and centrifugation for 10 min. at 1500*g*. Schistosomula were obtained *in vitro* [19] and were maintained in culture for up to 8 days in 6-well plates (25,000 schistosomula per well) under the conditions previously described [20]. Adult worms were obtained by whole-body perfusion of 6-week infected hamsters [21]. Eggs were obtained from the livers of infected hamsters and hatched under light to obtain miracidia [22]. Primary sporocysts were obtained after overnight axenic culture of miracidia as described [22]. Parasite DNA was extracted from the free-living cercariae using standard methods [23]. Total RNA was extracted from all life-cycle stages using the guanidine thiocyanate/caesium chloride method [24], except in the case of cultured schistosomula from which total RNA was obtained using the RiboPure kit (Ambion).

Cloning of the SmFes cDNA. A cDNA clone containing a 1891 bp insert encoding part of a tyrosine kinase catalytic domain was originally identified by random sequencing of clones from an *S. mansoni* egg stage library (G.O., unpublished data). The full-length cDNA sequence of SmFes was obtained by RACE-PCR (5′ and 3′ rapid amplification of cDNA ends). Complementary DNA was synthesized from adult worm RNA using the SMART RACE cDNA amplification kit (Clontech), according to the manufacturer's instructions. 5′-RACE was carried out using SmFesRev followed by SmFesRevnested and 3′-RACE with SmFesFwd followed by SmFesFwdnested (oligonucleotide primers used in the study are listed in Table 1) along with adaptor primers to amplify cDNA fragments of approximately 2200 and 1800 bp, respectively. Sequencing reactions were performed using the dye terminator cycle sequencing kit and analyzed on an ABI Prism 377 DNA sequencer (Perkin-Elmer Biosystems). Sequence analysis was performed using DNAStar (DNAstar Inc.). The sequence of SmFes was deposited with the GenBank Accession No. AF515706.

Computational sequence analysis and determination of intron–exon boundaries. Analyses of the SmFes amino acid sequence domains were made with RPS-BLAST [25]. The probability of SmFes to fold into a coiled-coil conformation was tested using the program Coils version 2.1 by comparison against a database of well-known parallel two-stranded coiled-coils using the MTIDK matrix [26]. The subdomains of the catalytic domain described by Hanks et al. [27] were manually inspected in SmFes after alignment with a set of Fps/Fes/Fer protein members using ClustalW [28] and displayed graphically with the use of the BioEdit v7.0.1 package (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The hydrophobic profile of SmFes was determined as described by Kyte and Doolittle [29] with a window size of 19 residues.

The SmFes cDNA sequence was compared with the genomic sequences available at the Wellcome Trust Sanger Institute *S. mansoni* Blast server (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_mansoni). Intron–exon junctions were detected by eye (5′-GT and 3′-AG) on sequence alignments obtained using Megalign (DNAStar Inc.) and contigs were assembled using Seqman (DNAStar Inc.). Some introns were confirmed by PCR followed by restriction digestion (not shown).

Table 1
Oligonucleotide primers used in the study

Primer	Sequence (5′–3′)
For cloning SmFes cDNA	
SmFesRev	CACGAAGACGACATTGATATGTACGTG
SmFesRevnested	CATAGATTGTCGTACACTTTGCCAAAGACG
SmFesFwd	CAGGAGCATGTCTTGTGACACCAATTTC
SmFesFwdnested	CTTGTCGTGTTGATATGACTGCATCGG
For generating probes for Southern blots	
SmFesTK4	TCATATTCTACTGATCCTCC
SmFesPK1	GACAGTTGGATTATGTAAAGG
SmGOTK1	ATGCCAAAAGACAGATAAAACCATACG
SmGOTK2	ATAACCCTTTACATAATCCAACTGTCAATG
E	AATCGTAAAATTATTGACAG
P2rev	GATTAGCAAACACAGCACGCCATTGAGAT
Characterization of the SmFes variant form	
SmFesTK2	AATAAGAATAGCCTAATAAGCCTTAATGGA
SmFesTK6	CCTTTACATAATCCAACTGTC
SmFesTK7	CAAGTTCTATACCAAATACACG
SmFesTK1	ATGCCAAAAGACAGATAAAACCGTACC
SmFesP4	CGAAATTTTGACATAGATTGTCGTACAC
Real-time quantitative PCR	
SmRTF	TCAATACAAAACGATATACTGATTCTGTGA
SmRTR	TGGACAGTTTTTTCAGCTTTGACT

Data mining and phylogenetic analysis. For phylogenetic analysis, a set of Fps/Fes/Fer protein members were selected from GenBank and aligned with ClustalW [28]. The sequences from each organism were selected based on the presence of the domains characteristic of the Fps/Fes/Fer family. The sequences used and their GenBank accession numbers were: FBS of Fujinami sarcoma virus (NP_955606); Fert2 of *Mus musculus*, (AAH58100); Fps/Fes of *Mus musculus* (AAN33122); Fer of *Mus musculus* (AAB18988); PTK of *Ephydatia fluvialis* (BAA81721); Fer of *Canis familiaris* (AAF00543); Fer of *Homo sapiens* (NP_005237); Fps/Fes of *H. sapiens* (P07332); Fps85D(dFer) of *Drosophila melanogaster* (P18106); Fer-frk-1 of *Caenorhabditis elegans* (NP_501818); Fes/Fps *Felis catus* (TVCTFF) and PTK of *Sycon raphanus* (CAA7660). Phylogenetic analysis was conducted by PHYLIP 3.6 (http://evolution.genetics.washington.edu/phylip.html). The complete sequences or the SH2-PK domains were analyzed separately using parsimony and bootstrap with 1000 replicates. The tree was constructed using TreeView 1.6.6 [30].

Southern blot. DNA from adult worms was obtained using the Wizard genomic DNA purification kit (Promega). For the Southern blot analyses, genomic *S. mansoni* DNA (10 μg) was digested with HindIII and EcoRI restriction enzymes and separated on 1% agarose gels. The genomic DNA was transferred onto a nylon membrane (Roche). Probes were obtained by incorporation of digoxigenin-labeled nucleotide during PCR amplification of the cDNA of *S. mansoni*, using as primers (Table 1) SmFesTK4 (position 2065–2084) and SmFesPK1 (reverse 2398–2418) (TK4PK1 probe), SmGOTK1 (reverse 2785–2811) and SmGOTK2 (forward 2393–2422) (TK1TK2 probe), or E (forward 991–1010) and P2rev (reverse 1503–1531) (P2revE probe). Blots were hybridized at 37 °C in the presence of 50% formamide, then washed twice in both SSC 2× 0.1% SDS for 15 min at 68 °C and in SSC 0.5× 0.1% SDS for 15 min at 37 °C. 1.5 μl of the anti-digoxigenin antibody (750 U/ml, Roche) was added to a 2.5% blocking solution (Roche) for 30 min. CSPD ready to use (Roche) was added for detection and the membrane exposed to Kodak X-Omat film.

Characterization of the SmFes variant form. In order to detect the presence of a 9 bp insert in transcripts, RNA was extracted from individual male or female worms, pooled adult males and females, eggs and cercariae, using the Rneasy Kit (Qiagen). RNA was treated with DNase I (deoxyribonuclease) (Invitrogen) to eliminate genomic DNA contamination and cDNA synthesis was done with Superscript II reverse-transcriptase (Invitrogen), according to the manufacturer's instructions. Standard PCR amplification proceeded for 35 cycles for 1 min at 95 °C, 1 min at 52 °C and 1 min at 72 °C, using TaqGold (Applied Biosystems) in an ABI 9700 thermocycler. For the reaction three sense oligonucleotide primers were used: SmFesTK2, (forward 2392–2422), SmFesTK6 (forward 2398–2418), SmFesTK7 (forward 2273–2294). All reactions were performed with one anti-sense primer SmFesTK1 (reverse 2798–2826), in individual reactions. Restriction digestion was carried out with 8 μl of the PCR product with 1 U of SfcI (New England Biolabs) at 25 °C for 16 h.

For sequencing of the genomic fragments containing the 15 bp insertion, DNA was amplified with primers SmFesTK4 (forward 2065–2084) and SmFesP4 (reverse 2207–2234) and directly sequenced using the DYEnamic ET Dye terminator sequence kit (GE Healthsciences) in a Megabace 500 automated sequencer (GE Healthsciences).

Real-time quantitative PCR. Analysis of the SmFes transcript levels at different developmental stages was performed by quantitative RT-PCR. Total RNA from the different life-cycle stages was reverse-transcribed using the Thermoscript™ RT-PCR System (Invitrogen). cDNAs were used as templates for PCR amplification using the SYBR® Green PCR Master Mix in an ABI PRISM 7000 sequence detection system (Applied Biosystems). Primers specific for *S. mansoni* tubulin (M80214, positions 851–873 and 925–904) and SmFes (AF515706, SmRTF, positions 356–389, and SmRTR, 486–509, Table 1) were designed by the Primer Express Program (Applied Biosystems). Amplification reactions were carried out in triplicate. PCR cycling conditions were of 40 cycles at 95 °C for 15 s and at 60 °C for 1 min. In order to determine the efficiency of the PCRs with each primer pair, C_t values were obtained for cercarial cDNA in amounts ranging from 40 pg to 100 ng. The standard curves obtained (not shown) showed high linearity (Pearson correlation coefficient $r > 0.99$). The real-time PCR efficiency (E) of one cycle in the exponential phase was calculated according to the equation: $E = 10^{[-1/\text{slope}]}$ [31]. The investigated transcripts showed very high and comparable efficiency rates; SmFes, 1.94; *S. mansoni* α-tubulin, 2.00. For graphical representation of quantitative PCR data, ΔC_t values were obtained by deducting the raw cycle threshold (C_t values) obtained for α-tubulin mRNA, the internal standard, from the C_t values obtained for SmFes in miracidia, cercariae, schistosomula, and adult male and female worms. The efficiency rates of the PCRs allow the ratios of expression to be calculated using the $2^{-\Delta\Delta C_t}$ ratio [32,33] compared to adult male worms.

Results and discussion

Sequence analysis and domain characterization of SmFes

The complete SmFes cDNA obtained by 5′ and 3′ RACE PCR (4911 bp) contained an open-reading frame (ORF) coding for a protein of 1259 aa (Fig. 1A) with a predicted molecular weight of 143.2 kilodaltons (kDa) and a theoretical isoelectric point of 6.93. The hydrophobic profile of SmFes indicated the absence of a transmembrane region in the protein. Structural analysis by RPS-BLAST indicated that SmFes exhibited characteristic features of the Fes/Fps/Fer protein tyrosine kinase family: three coiled-coil regions, one SH2 (gn1|CDD|16538) and one PTK catalytic domain (gn1|CDD|5392) signatures (Fig. 1A).

Interestingly, no FCH amino terminal domain (Fes/CIP4 homology domain), usually found in proteins of the Fes/Fps/Fer family [17], was identifiable in SmFes using RPS-BLAST and this result was confirmed by InterProScan analysis. However, a closer examination of the N-terminal region of SmFes and particularly the first coiled-coil domain (Fig. 1A) showed some similarity to the FCH domains of Fes and Fer proteins of other species (Fig. 1B). Moreover, this alignment is supported by the presence of two conserved intron positions immediately downstream of the exon encoding this region (see Expression of SmFes during the *S. mansoni* life-cycle). It is therefore possible that SmFes contains a degenerate FCH domain. This domain is found in proteins that are involved in the regulation of cytoskeletal rearrangement, vesicular transport and endocytosis, and microfilament association [17]. The FCH domain of Fes has been reported to bind to microtubules and tubulin [34] and in this way to be involved in microtubule nucleation. A Fes mutant, lacking the FCH domain, still co-localized with microtubules, but it could not bind to soluble tubulin [35]. However, not all metazoan Fes family members possess this domain. In *C. elegans,* 42 genes of the 52 cytoplasmic PTKs, belong to the Fer subfamily [36]. However, none of these predicted PTKs contain the FCH or coiled-coil domains that characterize the Fes/Fps/Fer family. This absence and that of a strongly conserved FCH domain from SmFes probably represent a loss of function during evolution since the Fes family member from the marine sponge *S. raphanus* contains a conserved FCH domain ([37]; Fig. 1B). SmFes does contain two further coiled-coil domains (Fig. 1A). These mediate the formation of oligomers, homopentamers

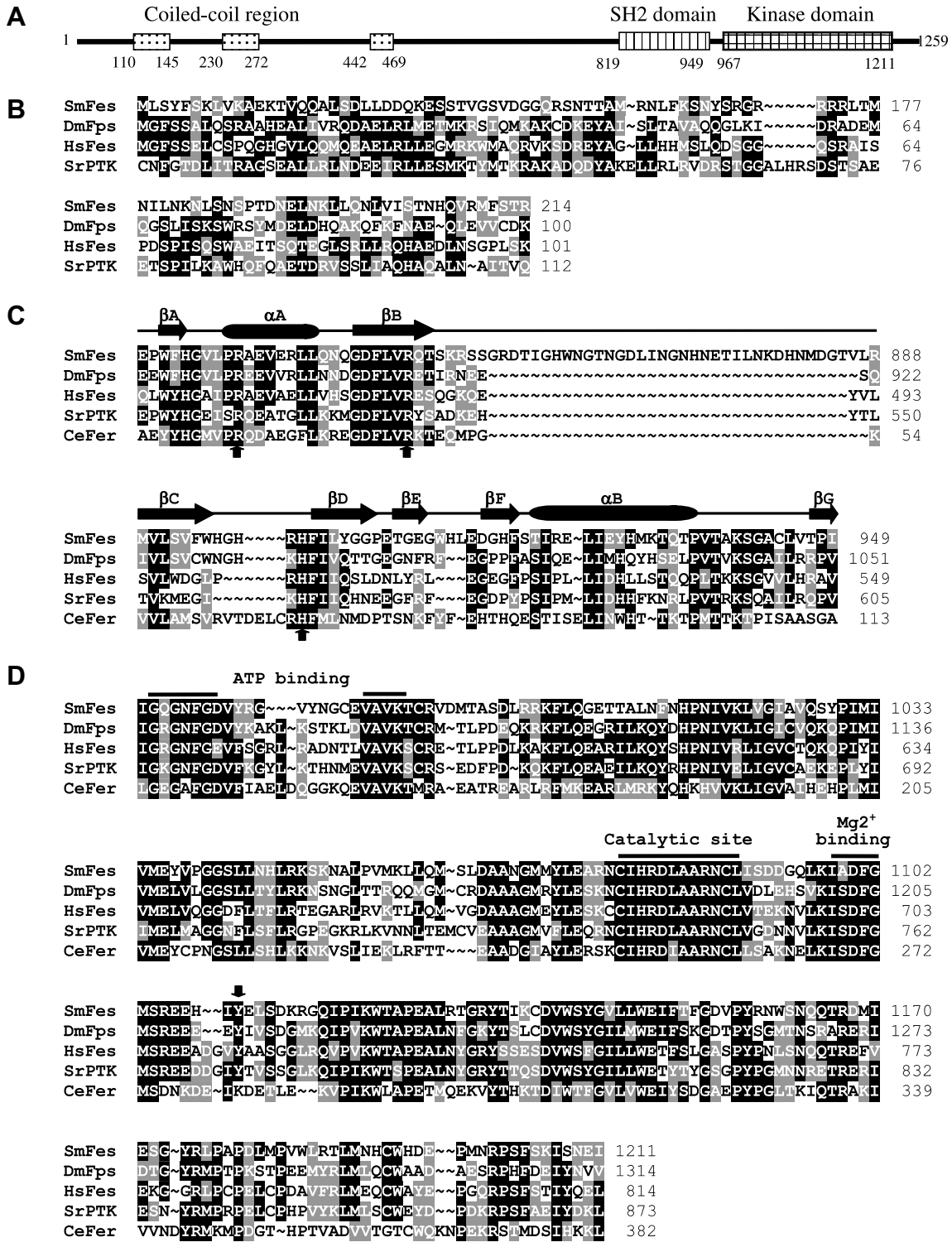

Fig. 1. Domain structure and conservation of SmFes. (A) Schematic diagram of the domain structure of the SmFes protein. The numbers indicate the amino acid sequence limits of each domain. (B) Alignment of the putative FCH domain of SmFes with those of *D. melanogaster* Fps85D (DmFps; Accession No.: P18106) human Fps/Fes (HsFes; P07332), and *S. raphanus* PTK (SrPTK; Y17051). (C) Alignment of the SH2 domains of SmFes, HsFes, DmFps, SrPTK and *C. elegans* Fer-frk-1 (CeFer; NP_501818). The conserved structural motifs (α-helices and β-strands, according to [39]) are shown above the alignment. Arrows indicate the conserved residues involved in interactions with phosphotyrosines. (D) Alignment of the kinase domains of SmFes, HsFes, DmFps, SrPTK, and CeFer. Conserved motifs involved in ATP and Mg^{2+} binding and catalysis are overlined. The arrow indicates the conserved autophosphorylated tyrosine residue. Alignments were carried out using Clustal W and the graphical representation prepared using BioEdit v7.0.1 package (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

or even higher order structures in the case of Fps/Fes, which are thought to facilitate *trans*-autophosphorylation [17]. Moreover, together with the SH2 domain, the coiled-coil motifs regulate the activity of Fes within cells [38].

BLAST-p of SmFes against the nr database indicated 40–50% identity of the SH2 and kinase domains with those of the Fes/Fps/Fer family. The SH2 domain recognizes phosphotyrosine residues in ligand peptides via the phosphotyrosine (pTyr) pocket and a specificity determining region which recognizes downstream residues. The overall architecture of the SH2 domain of SmFes (two α-helices and seven β-strands ordered βαββββββαβ; [39]) is conserved (Fig. 1C) as are the three basic residues, R^{828}, R^{844}, and H^{900} involved in interactions with phosphotyrosines. However, compared to the overall architecture of SH2 domains [39] and other members of the family (Fig. 1C), a large insertion is present, the functional significance of which is unclear.

The SmFes catalytic domain conserves all the characteristics of a protein kinase (Fig. 1D) including the ATP and Mg^{2+} binding sites and the catalytic site [27]. The eleven subdomains of the kinase domain, previously characterized [27], were also observed in SmFes at the expected positions, with conserved subdomains VI and VIII specifying the tyrosine kinase nature of SmFes (not shown). The autophosphorylation tyrosine residue, related to position 713 in the human Fps/Fes protein [40], is located at position 1110 in SmFes (Fig. 1D). Tyrosine autophosphorylation sites are able to regulate the kinase activity of the domain (Tyr^{713} is in the activation loop of human Fps/Fes) and also recruit proteins containing SH2 domains. Autophosphorylation of these residues seems to be an intermolecular event, as is the case for growth factor RTKs for example [40].

Taken together, these results indicate that SmFes is probably functional, containing most of the structural motifs and the TK signature. Nevertheless, the poorly conserved FCH domain and the presence of an insertion in the SH2 domain, suggest that SmFes may have different protein partners and/or substrates compared to other members of the Fes/Fps/Fer family.

Phylogenetic analysis

Phylogenetic reconstruction was carried out to determine from which branch of the Fes/Fps/Fer family SmFes is derived. Phylogenetic analysis was performed with the SH2 and PK domains, or the complete amino acid sequences of Fes/Fps/Fer of other species. The resulting dendrogram of the SH2-PK sequence clustered SmFes with other invertebrate orthologues with a significant bootstrap value (Fig. 2). The same observation was made when the complete protein sequences were used (not shown). The exception was the *C. elegans* Fer-frk-1 protein that was grouped in the invertebrate branch in the SH2-PK tree, while in the complete protein tree it was grouped with

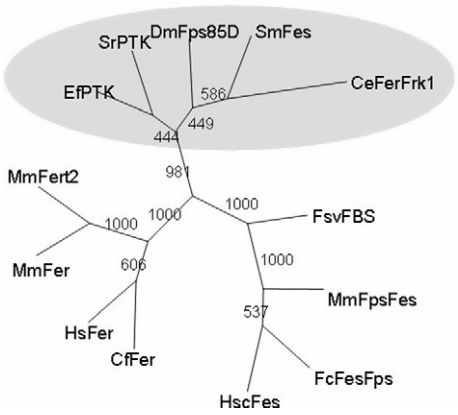

Fig. 2. Unrooted phylogenetic tree of SmFes with other Fps/Fes/Fer member proteins: Sm, *Schistosoma mansoni*; Hs, *Homo sapiens*; Cf, *Canis familiaris*; Mm, *Mus musculus*; Ce, *Caenorhabditis elegans*; Fs.virus, *Fujinami sarcoma virus*; Ef, *Ephydatia fluvialis*; Dm, *Drosophila melanogaster*; Fc, *Felis catus*; Sr, *Sycon raphanus*. (A) Alignment of the SH2 and PK domains. Bootstrap values are shown at branch points. Area shaded grey indicates invertebrate proteins.

the Fer subfamily branch. This anomaly is likely due to the fact that this protein, along with the other members of this family in *C. elegans*, comprises only the SH2 and kinase domains and lacks the coiled-coil domains. CeFer-frk-1 was chosen for this analysis among the 42 members of this family present in the nematode [36] since it gave the highest Blast score when compared to SmFes. The overall analysis shows that, despite the low conservation of the FCH domain, SmFes is clearly a member of the Fes/Fps/Fer family. The analysis of orthologues from other lophotrochozoan species will be necessary to determine whether or not the divergence of the N-terminal part of the protein is a schistosome-specific event.

Expression of SmFes during the S. mansoni life-cycle

We investigated the relative expression levels of SmFes during the parasite life-cycle using quantitative RT-PCR (Fig. 3). This showed that SmFes transcript levels were relatively low at all stages. For example, in adult male worms there was approximately 620-fold less SmFes mRNA than α-tubulin mRNA. The comparison of SmFes mRNA levels at the different life-cycle stages showed that male worms had the highest levels (Fig. 3A) and that expression was markedly reduced in both female worms and cercariae (about 4-fold less). Interestingly, SmFes mRNA levels were not increased in miracidia. Miracidia are the free-living ciliated larvae which invade the fresh-water snail intermediate host and we have previously observed that a number of *S. mansoni* transcripts show markedly increased levels of expression at this stage. This is notably the case for the transcripts encoding other schistosome proteins such as SmPKC1 [41] and enzymes involved in gluconeogenesis [42] and is thought to reflect the stocking of messengers ready for translation immediately after host invasion [41]. The fact that SmFes mRNA levels are highest in the

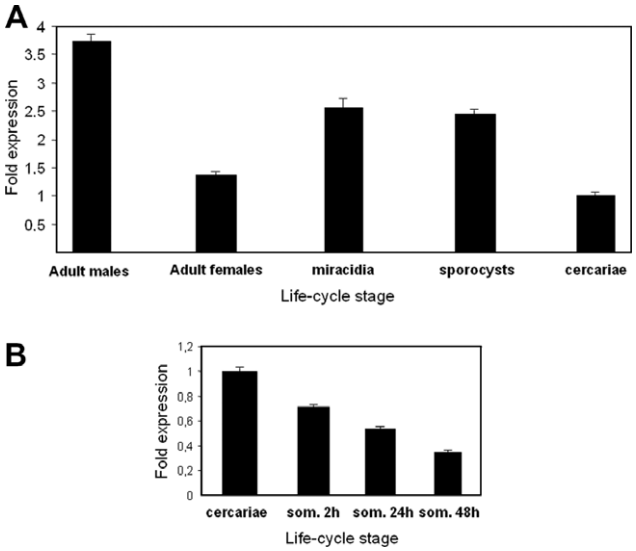

Fig. 3. (A) Transcript levels of SmFes at the different development stages of *S. mansoni* determined by quantitative RT-PCR. The relative amounts of SmFes transcripts in adult male worms, adult female worms, miracidia, sporocysts, and cercariae are shown. The $2^{-\Delta\Delta C_t}$ values were calculated relative to the expression in adult male worms (see real-time quantitative PCR). (B) Transcript levels in cercariae compared to those in newly transformed schistosomula (som. 2 h) and schistosomula maintained in culture for 24 h (som. 24 h) or 48 h (som. 48 h). Results shown are from a representative experiment, one of three carried out.

differentiated male worms may indicate a specific role for SmFes in these worms.

In a parallel study [18], we showed that the SmFes protein is particularly expressed at the terebratorium of miracidia and the tegument of cercaria and skin-stage schistosomula. Moreover, no SmFes protein was found in schistosomula maintained in culture for 7 days. We therefore carried out a comparison of the expression of SmFes mRNA in cercariae, newly transformed schistosomula and larvae incubated in culture (Fig. 3B). The results of three separate experiments show that SmFes mRNA levels decrease steadily after transformation of cercariae into schistosomula and subsequent culture, consistent with the corresponding decline in the amounts of protein. After 48 h mRNA levels were reduced by more than 60%. This finding is consistent with a putative role for SmFes in signaling during the process of host invasion.

The genomic structure of SmFes: comparison with the human Fps/Fes and Drosophila Fes genes

The *SmFes* gene structure was determined by screening the individual and assembled genomic sequences at the Wellcome Trust Sanger Institute *S. mansoni* Blast server. The results of the survey are shown in detail on Table 1 in Supplementary material online. The gene comprises 19 exons and 18 introns of which 7 are of indeterminate size due to gaps in the contig assembly which we were not able to fill by chromosome walking with individual shotgun sequences. The overall gene size is >52,990 bp with intron sizes ranging from 1057 to >7912 bp. The sizes of some of the introns was checked by PCR (not shown) and notably intron 18 was found to be about 1100 bp in size (>755 bp from the genomic assembly). There are therefore no very small introns which are often present at the 5′ end of schistosome genes such as the three ~30 bp introns present in the *SmPKC1* gene [41]. Moreover, it is unlikely that very large introns are present such as the >40 kbp introns in two schistosome *Hox* genes [43].

The genome survey indicated that only one copy of the SmFes gene was present and this view was supported by Southern blotting. We used PCR-generated cDNA fragments corresponding to exon 10 to probe Southern blots with genomic DNA digested with HindIII or EcoR1 (Fig. 1 in Supplementary material online).

The comparison of the organization of the *SmFes* gene with those of the Human *Fps/Fes* (*HsFps/Fes*) and *Drosophila Fps* (*DmFps*) genes showed that some intron positions were conserved, particularly at the 5′ end of the genes and in the regions encoding the SH2 and kinase domains. The *HsFps/Fes* gene (on genomic contig NT_010274.16), like *SmFes*, contains 19 exons including one non-coding exon, but is more compact; the whole gene spans only 11,297 bp. This contrasts with the human *Fer* gene (on genomic contig NT_034772) which contains 20 exons spanning 439,845 bp (not shown). The direct comparison of the *HsFps/Fes* and *SmFes* gene structures (Table 2) and the location of the intron positions on the peptide

Table 2
Comparison of the genomic organization of the SmFes and *HsFps/Fes* genes

HsFes[a]				SmFes[a]			
Exons	Size[b]	Introns	Size	Exons	Size	Introns	Size
1	66	1	493	1	22	1	2963
2	222	2	154	2	685	2	1057
3	**174**	3	1376	**3**	**174**	3	>5298
4	**97**	4	130	**4**	**97**	4	6807
5	184	5	1936	5	199	5	>4613
6	138	6	74	6	108	6	>387
7	120	7	204	7	129	7	1792
8	123	8	130	8	294	8	3830
9	187	9	123	9	331	9	>2882
10	**84**	10	498	10	421[c]	10	>7912
11	210	11	363	11	110	11	>1638
12	123	12	382	**12**	**65**	12	3482
13	54	13	596	13	273	13	2229
14	119	14	275	14	136	14	1613
15	95	15	97	**15**	**211**	15	>1395
16	**124**	16	240	**16**	**125**	16	2317
17	158	17	125	17	91	17	2020
18	123	18	1358	**18**	**184**	18	>755
19	360			19	1280		

The SmFes gene structure is detailed in Table 1 in Supplementary material online. Numbers in bold and underlined are conserved exons between the human and schistosome genes. Numbers in bold are introns conserved between the SmFes gene and the *D. melanogaster* Fps gene (not shown).

[a] *HsFes* gene size = 11,297 bp, SmFes > 52,990 bp.

[b] Size given in base pairs.

[c] The genomic sequence contains a 15bp insertion.

sequence (not shown) demonstrates that several intron positions are conserved, leading to conserved exons (Table 2). Notably, the positions of exons 3 and 4 in each gene are perfectly conserved and have identical sizes. These exons are immediately downstream of exon 2 which in each case encodes the (putative in the case of SmFes) FCH domain and support the peptide sequence similarity detected in this region. The central exons are not conserved, either in terms of size or of position. Exon 10 in *HsFps/Fes* and exon 12 in *SmFes* both encode the start of the SH2 domain and are conserved. Similarly, exons 16 in each gene are also conserved and encode part of the kinase domain. Comparison with the *DmFps* gene (on genomic contig NT_033777.2) showed two further conserved exons (Table 2) in the region encoding the kinase domain (exons 15 and 18 in *SmFes*, exons 10 and 13 in *DmFps*, not shown). Moreover, exons 3 and 4 of *DmFps* are also conserved, further supporting the identity of this region in *SmFes*. The conservation of intron positions in schistosome genes is variable. In general, as in the *Hox* genes [43], intron positions are identical to those of orthologous genes within and around highly conserved functional domains (the homeodomain for example), but not elsewhere in the gene. This is the case for *SmFes*, but, in contrast, in *SmPKC1* [41] all the intron positions were conserved compared to the human *PKCβ* gene.

Whilst, we have detected a degree of polymorphism in the *SmFes* gene (see Polymorphism of SmFes), it differs from its orthologues in the apparent absence of major alternatively spliced isoforms. No potentially alternatively major spliced forms (in which exons were deleted) were detected, either during the 5′ and 3′ RACE experiments, or during PCR amplification of internal fragments. This contrasts with the *DmFps* gene for which four different transcripts are identified in GenBank [CG8874-RA (NM_079564), CG8874-RB (NM_169274), CG8874-RC (NM_169275), and CG8874-RD (NM_169276)] generated by alternative splicing or alternative promoter usage (CG8874-RD). Similarly, the human *Fer* gene, which has an identical genomic structure to *HsFps/Fes*, apart from the presence of an additional 5′ non-coding exon, generates a truncated isoform in testis, FerT, the transcription of which is driven from a testis-specific promoter [17].

Polymorphism of SmFes

Sequencing of multiple SmFes cDNA clones permitted the identification of clones containing a 9 bp insert just before the SH2 domain (nt position 2556) and coding for the peptide sequence VSE (Fig. 4A). Comparison of the cDNA sequence with the corresponding genomic region suggested the existence of an alternative splice site at the 5′end of the intron 10, resulting in the cDNA populations that are observed (Fig. 4B). The verification of alternatively spliced messages was performed by RFLP (restriction-fragment length polymorphism) as the presence of the 9 bp insertion created a restriction site for SfcI. Three different PCR amplifications of the insertion region were conducted with cDNA obtained from adult worms. All PCR fragments were partially digested, indicating that the two cDNA populations were present (Fig. 4C). The same observation was made for cDNA obtained from pools of male or female worms, cercariae, eggs and from individual male or female worms, indicating that alternative splicing of SmFes is present throughout the life-cycle of the parasite and could coexist in a single individual (Fig. 4D). Control DNA was fully digested (not shown). These alternative transcripts could determine tissue specific expression, cellular localization or mechanisms of protein regulation. The function of the alternatively spliced SmFes is unknown. It is possible, however, that the three residues could alter the protein structure and modulate the function of SmFes.

Comparison of the genomic and cDNA sequences indicated the possibility of a 15 bp insertion at position 2170 of the cDNA (within exon 10) coding for the peptide LQHQQ. None of the 11 cDNAs sequenced had the 15 bp insertion (Fig. 4E). However, all three genomic sequences from the Sanger database had the insertion, as did some of the new genomic sequences generated by us. Taken together, the results indicate that these two types of sequence may be alleles of the same *SmFes* gene. Two different populations, containing or not the 15 bp insertion, were identified among several genomic sequences of *S. mansoni*, and can be classified as alleles since no splicing elements were observed and since *SmFes* is a single copy gene. Allele-specific polymorphisms were also identified at the *S. mansoni* Fyn-like protein kinase [14].

Single nucleotide polymorphisms (SNPs) at position 2166 (C-A) and 2169 (G-A) were also observed. The SNPs always appeared linked to one of the two genomic alleles described above (Fig. 4E). The first SNP results in a substitution of a histidine for glutamine in the allele containing the 15 bp insertion, while the second SNP is silent and maintains the glutamine residue (Fig. 4E). Analysis of genomic and cDNA clones sequences indicated the possible existence of single nucleotide polymorphisms (SNP) at position 2416 (G/A) of the cDNA, just before the PK domain. The polymorphism results in a substitution of valine for isoleucine.

Conclusion

SmFes shows a high degree of structural conservation, particularly in the kinase catalytic domain, but also original structural aspects compared to other members of the Fps/Fes/Fer family. The low level of conservation of the putative FCH domain indicates that it may not interact with microtubules in the same way as does mammalian Fps/Fes [34] and the large insertion in the SH2 domain may influence the binding specificity of the domain and the type of protein partner recognized. Moreover, we are currently constructing a 3D model of the SH2 domain to determine the potential effect on the domain structure both of the large insertion and of the VSE insertion which just

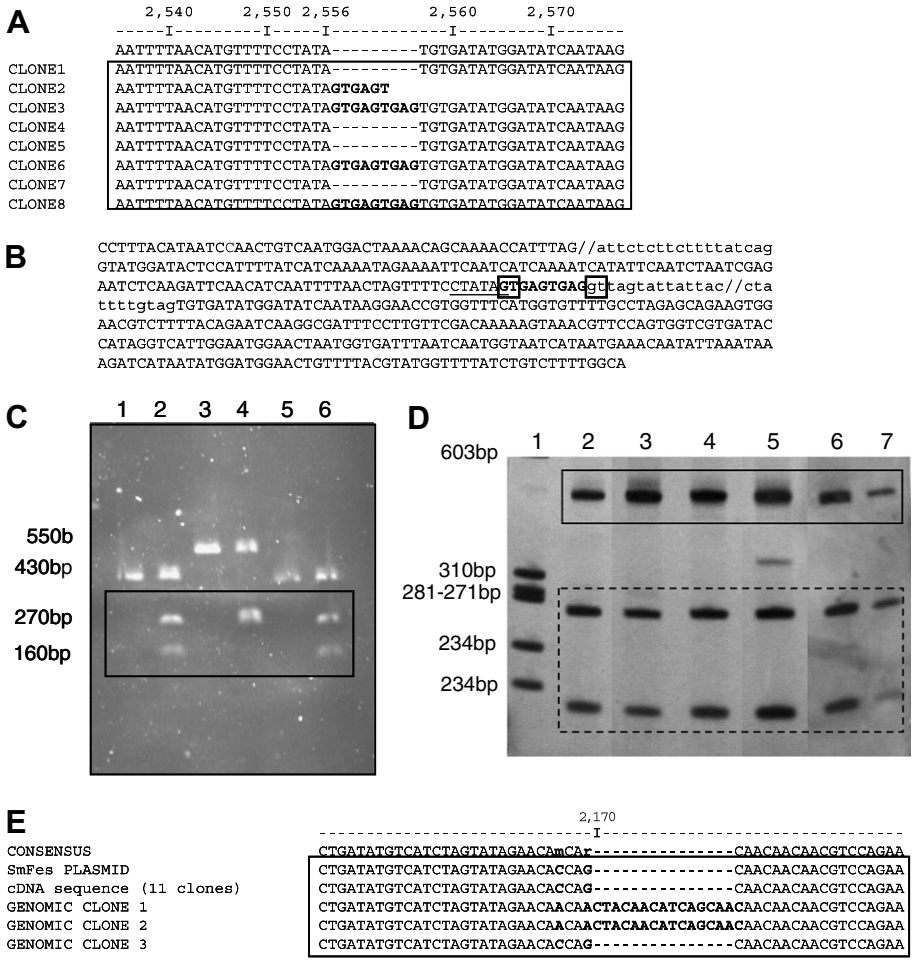

Fig. 4. (A) Alternative splicing of SmFes. (A) Sequences of 8 cDNA clones indicating the 9 bp insertion at position 2556. (B) Genomic sequence of the region where the 9 bp insertion occurs. The coding sequence is in capital letters, introns in lower case letters. The squares indicate the alternative splice donor sites. The SfcI restriction site is underlined. (C) Restriction pattern of PCR products obtained from the amplification of adult worm cDNA. Lanes 1–2, 3–4, and 5–6 were obtained by amplification with SmFesTK1/SmFesTK2, SmFesTK1/SmFesTK7, and SmFesTK1/SmFesTK6 primer pairs, respectively. Lanes 2, 4, and 6 were digested with SfcI. The undigested cDNA is a band of 428 bp for lane 1, 547 bp for lane 3 and 423 bp for lane 5. After digestion, the fragment is cleaved into 159 and 269 bp for lane 2, 278 and 269 bp for lane 4, 154 and 269 bp for lane 6, shown in the dotted box. Molecular weight markers are shown on the left. (D) Restriction pattern of PCR products obtained from cDNA amplified with the primers, SmFesTK1 and SmFesTK2, and digested with the enzyme SfcI. (1) Molecular weight markers, (2) adult worm pool, (3) male pool, (4) female pool, (5) cercariae pool, (6) individual female worm, (7) individual male worm. (E) *SmFes* gene allele containing a 15 bp insertion. The 15 bp are inserted after the base 2170 of the cDNA in exon 10. 11 cDNA clones showed no insertion. Several genomic clones, either with or without the 15 bp insertion were observed.

precedes the domain. We have shown that the SmFes protein is expressed in locations in infective larvae (cercarial tegument, miracidial terebretorium) consistent with a role in host invasion [18]. Indeed, one of the defined roles of mammalian Fes/Fps/Fer is to regulate cell–cell and cell–matrix interactions [17] and we can speculate that in schistosomes SmFes could similarly regulate cytoskeletal rearrangements associated with host penetration. Future work will be aimed at identifying protein partners of SmFes, interacting either with the putative FCH domain or the SH2 domain as well as protein substrates of SmFes in order to determine its biological role. A number of orthologues of partners of Fps/Fes/Fer [17] have been found in *S. mansoni* by *in silico* screening of the EST and genomic databases (F.L. and D.B., not shown) including cortactin, PI3K, and β-catenin, as well as other PKs such as SRC and JAK. We will extend and complete this survey and select potential partners in order to test the conservation or otherwise of the protein interaction networks involving SmFes.

Acknowledgments

The work was supported by the Inserm-Fiocruz collaborative programme, the Institut National de la Santé et de la Recherche Médicale (U547), the Institut Pasteur de Lille, the Centre National de la Recherche Scientifique and the Microbiology program of the Ministère de l'Education Nationale, de la Recherche et de la Technologie (MENRT). The financial support from FIOCRUZ/PDTIS (GO), PAPES (GO 400315/2006-8) and FAPEMIG (DB CBB-174/02) is also acknowledged. F.L. is supported by

an NIH-Fogarty grant (TW007012-01). G.O. is a CNPq fellow.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2007.06.018.

References

[1] D. Engels, L. Chitsulo, A. Montresor, L. Savioli, The global epidemiological situation of schistosomiasis and new approaches to control and research, Acta Trop. 82 (2002) 139–146.
[2] D. Cioli, L. Pica-Mattoccia, Praziquantel, Parasitol. Res. 90 (Suppl. 1) (2003) S3–S9.
[3] G. Oliveira, N.B. Rodrigues, A.J. Romanha, D. Bahia, Genome and genomics of schistosomes, Can. J. Zool. 82 (2004) 375–390.
[4] W. Kunz, Schistosome male–female interaction: induction of germ-cell differentiation, Trends Parasitol. 17 (2001) 227–231.
[5] P.T. Loverde, L. Chen, Schistosome female reproductive development, Parasitol. Today 7 (1991) 303–308.
[6] A.R. Harris, R.J. Russell, A.D. Charters, A review of schistosomiasis in immigrants in Western Australia, demonstrating the unusual longevity of *Schistosoma mansoni*, Trans. R. Soc. Trop. Med. Hyg. 78 (1984) 385–388.
[7] D. Bahia, L.F. Andrade, F. Ludolf, R.A. Mortara, G. Oliveira, Protein tyrosine kinases in *Schistosoma mansoni*, Mem. Inst. Oswaldo Cruz 101 (2006) 137–143.
[8] C. Dissous, N. Khayath, J. Vicogne, M. Capron, Growth factor receptors in helminth parasites: signalling and host–parasite relationships, FEBS Lett. 580 (2006) 2968–2975.
[9] K. Neet, T. Hunter, Vertebrate non-receptor protein-tyrosine kinase families, Genes Cells 1 (1996) 147–169.
[10] H. Mano, The Tec family protein-tyrosine kinases: a subset of kinases for a subset of signalings, Int. J. Haematol. 69 (1999) 6–12.
[11] T.E. Smithgall, J.A. Rogers, K.L. Peters, J. Li, S.D. Briggs, J.M. Lionberger, H. Cheng, A. Shibata, B. Scholtz, S. Schreiner, N. Dunham, The c-Fes family of protein-tyrosine kinases, Crit. Rev. Oncog. 9 (1998) 43–62.
[12] S.M. Thomas, J.S. Brugge, Cellular functions regulated by Src family kinases, Ann. Rev. Cell Dev. Biol. 13 (1997) 513–609.
[13] P. Traxler, Tyrosine kinases as targets in cancer therapy—successes and failures, Expert Opin. Ther. Targets 7 (2003) 215–234.
[14] K. Kapp, P. Schussler, W. Kunz, C.G. Grevelding, Identification, isolation and characterization of a Fyn-like tyrosine kinase from *Schistosoma mansoni*, Parasitology 122 (2001) 317–327.
[15] J. Knobloch, R. Winnen, M. Quack, W. Kunz, C.G. Grevelding, A novel Syk-family tyrosine kinase from *Schistosoma mansoni* which is preferentially transcribed in reproductive organs, Gene 294 (2002) 87–97.
[16] K. Kapp, J. Knobloch, P. Schussler, S. Sroka, R. Lammers, W. Kunz, C.G. Grevelding, The *Schistosoma mansoni* Src kinase TK3 is expressed in the gonads and likely involved in cytoskeletal organization, Mol. Biochem. Parasitol. 138 (2004) 171–182.
[17] P. Greer, Closing in on the biological functions of Fps/Fes and Fer, Nat. Rev. Mol. Cell Biol. 3 (2002) 278–289.
[18] D. Bahia, R.A. Mortara, J.R. Kusel, L.F. Andrade, F. Ludolf, P.R. Kuser, L. Avelar, J. Trolet, C. Dissous, R.J. Pierce, G. Oliveira, *Schistosoma mansoni*: expression of Fes-like tyrosine kinase SmFes in the tegument and terebratorium suggests its involvement in host penetration, Exp. Parasitol. 116 (2007) 225–232.
[19] F.J. Ramalho-Pinto, G. Gazzinelli, R.E. Howells, T.A. Mota-Santos, E.A. Figueiredo, J. Pellegrino, *Schistosoma mansoni*: defined system for stepwise transformation of cercaria to schistosomule in vitro, Exp. Parasitol. 36 (1974) 360–372.
[20] R. Harrop, R.A. Wilson, Protein synthesis and release by cultured schistosomula of *Schistosoma mansoni*, Parasitology 107 (Pt. 3) (1993) 265–274.
[21] S.R. Smithers, R.J. Terry, The infection of laboratory hosts with cercariae of *Schistosoma mansoni* and the recovery of the adult worms, Parasitology 55 (1965) 695–700.
[22] T.P. Yoshino, J.R. Laursen, Production of *Schistosoma mansoni* daughter sporocysts from mother sporocysts maintained in synxenic culture with *Biomphalaria glabrata* embryonic (Bge) cells, J. Parasitol. 81 (1995) 714–722.
[23] J. Sambrook, D.W. Russel, Molecular Cloning: A Laboratory Manual, third ed., Cold Spring Harbour Laboratory Press, New York, 2001.
[24] J.M. Chirgwin, A.E. Przybyla, R.J. MacDonald, W.J. Rutter, Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease, Biochemistry 18 (1979) 5294–5299.
[25] A. Marchler-Bauer, J.B. Anderson, C. DeWeese-Scott, N.D. Fedorova, L.Y. Geer, S. He, D.I. Hurwitz, J.D. Jackson, A.R. Jacobs, C.J. Lanczycki, C.A. Liebert, C. Liu, C. Madej, G.H. Marchler, R. Mazumder, A.N. Nikolskaya, A.R. Panchenko, B.S. Rao, B.A. Shoemaker, V. Simonyan, J.S. Song, P.A. Thiessen, S. Vasudevan, Y. Wang, R.A. Yamashita, J.J. Yin, S.H. Bryant, CDD: a curated Entrez database of conserved domain alignments, Nucleic Acids Res. 31 (2003) 383–387.
[26] A. Lupas, M. Van Dyke, J. Stock, Predicting coiled coils from protein sequences, Science 252 (1991) 1162–1164.
[27] S.K. Hanks, A.M. Quinn, T. Hunter, The protein kinase family: conserved features and deduced phylogeny of the catalytic domains, Science 241 (1988) 42–52.
[28] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[29] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[30] R.D. Page, TreeView: an application to display phylogenetic trees on personal computers, Comput. Appl. Biosci. 12 (1996) 357–358.
[31] R. Rasmussen, Quantification on the Light Cycler, in: S. Meuer, C. Wittwer, K. Nakagawara (Eds.), Rapid Cycle Real-time PCR, Methods and Applications, Springer Press, Heidelberg, 2001, pp. 21–34.
[32] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (2001) e45.
[33] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method, Methods 25 (2001) 402–408.
[34] S. Takahashi, R. Inatome, A. Hotta, Q. Qin, R. Hackenmiller, M.C. Simon, H. Yamamura, S. Yanagi, Role for Fes/Fps tyrosine kinase in microtubule nucleation through is Fes/CIP4 homology domain, J. Biol. Chem. 278 (2003) 49129–49133.
[35] C.E. Laurent, F.J. Delfino, H.Y. Cheng, T.E. Smithgall, The human c-Fes tyrosine kinase binds tubulin and microtubules through separate domains and promotes microtubule assembly, Mol. Cell. Biol. 24 (2004) 9351–9358.
[36] G.D. Plowman, S. Sudarsanam, J. Bingham, D. Whyte, T. Hunter, The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms, Proc. Natl. Acad. Sci. USA 96 (1999) 13603–13610.
[37] H. Cetkovic, I.M. Muller, W.E. Muller, V. Gamulin, Characterization and phylogenetic analysis of a cDNA encoding the Fes/FER related, non-receptor protein–tyrosine kinase in the marine sponge *Sycon raphanus*, Gene 216 (1998) 77–84.
[38] Y. Takashima, F.J. Delfino, J.R. Engen, G. Superti-Furga, T.E. Smithgall, Regulation of c-Fes tyrosine kinase activity by coiled-coil and SH2 domains: analysis with *Saccharomyces cerevisiae*, Biochemistry 42 (2003) 3567–3574.
[39] K. Machida, B.J. Mayer, The SH2 domain: versatile signaling module and pharmaceutical target, Biochem. Biophys. Acta 1747 (2005) 1–25.

[40] J.A. Rogers, R.D. Read, J. Li, K.L. Peters, T.E. Smithgall, Autophosphorylation of the Fes tyrosine kinase. Evidence for an intermolecular mechanism involving two kinase domain tyrosine residues, J. Biol. Chem. 271 (1996) 17519–17525.

[41] D. Bahia, L. Avelar, R.A. Mortara, N. Khayath, Y. Yan, C. Noel, M. Capron, C. Dissous, R.J. Pierce, G. Oliveira, SmPKC1, a new protein kinase C identified in the platyhelminth parasite *Schistosoma mansoni*, Biochem. Biophys. Res. Commun. 345 (2006) 1138–1148.

[42] N. Khayath, G. Mithieux, C. Zitoun, C. Coustau, J. Vicogne, A.G. Tielens, C. Dissous, Glyceroneogenesis: an unexpected metabolic pathway for glutamine in *Schistosoma mansoni* sporocysts, Mol. Biochem. Parasitol. 147 (2006) 145–153.

[43] R.J. Pierce, W. Wu, H. Hirai, A. Ivens, L.D. Murphy, C. Noel, D.A. Johnston, F. Artiguenave, M. Adams, J. Cornette, E. Viscogliosi, M. Capron, G. Balavoine, Evidence for a dispersed Hox gene cluster in the platyhelminth parasite *Schistosoma mansoni*, Mol. Biol. Evol. 22 (2005) 2491–2503.