

A functionally conserved member of the FTZ-F1 nuclear receptor family from *Schistosoma mansoni*

Ricardo L. de Mendonça^{1,*}, Didier Bouton^{1,†}, Benjamin Bertin¹, Hector Escriva², Christophe Noël¹, Jean-Marc Vanacker², Jocelyne Cornette¹, Vincent Laudet² and Raymond J. Pierce¹

¹INSERM U 547, Institut Pasteur, Lille, France; ²CNRS UMR 49, Ecole Normale Supérieure de Lyon, Lyon, France

The *fushi tarazu* factor 1 (FTZ-F1) nuclear receptor subfamily comprises orphan receptors with crucial roles in development and sexual differentiation in vertebrates and invertebrates. We describe the structure and functional properties of an FTZ-F1 from the platyhelminth parasite of humans, *Schistosoma mansoni*, the first receptor from this family to be characterized in a Lophotrochozoan. It contains a well conserved DNA-binding domain (55–63% identity to other family members) and a poorly conserved ligand-binding domain (20% identity to that of zebrafish FF1a). However, both the ligand domain signature sequence and the activation function 2-activation domain (AF2-AD) are perfectly conserved. Phylogenetic analysis confirmed that SmFTZ-F1 is a member of nuclear receptor subfamily 5, but that it clustered with the *Drosophila* receptor DHR39 and has consequently been named NR5B1. The gene showed a complex structure with 10 exons and an overall size of 18.4 kb. Two major transcripts were detected, involving

alternative promoter usage and splicing of the two 5' exons, but which encoded identical proteins. SmFTZ-F1 mRNA is expressed at all life-cycle stages with the highest amounts in the larval forms (miracidia, sporocysts and cercariae). However, expression of the protein showed a different pattern; low in miracidia and higher in adult male worms. The protein bound the same monomeric response element as mammalian SF-1 (SF-1 response element, SFRE) and competition experiments with mutant SFREs showed that its specificity was identical. Moreover, SmFTZ-F1 trans-activated reporter gene transcription from SFRE similarly to SF-1. This functional conservation argues for a conserved biological role of the FTZ-F1 nuclear receptor family throughout the metazoa.

Keywords: platyhelminth; development; orphan receptor; phylogeny; DNA-binding.

The FTZ-F1 gene subfamily encodes orphan nuclear receptors and appears to be present in all metazoan phyla [1]. The first member of the subfamily, FTZ-F1 α , was isolated from *Drosophila melanogaster* [2,3] and was identified both as a transcriptional regulator and cofactor [4,5] of the homeodomain protein *fushi tarazu* (FTZ), a segmentation gene of the pair-rule class responsible for the formation of alternative segmental units in the *D. melanogaster* embryo [6]. FTZ-F1 α

is expressed in early embryos, concomitant with FTZ expression. A second isoform, FTZ-F1 β , encoded by the same gene [7], is detectable in late-stage embryos through to adults, when FTZ expression is absent, and regulates genes associated with ecdysis and metamorphosis [8]. In the nematode *Caenorhabditis elegans*, *nhr-25*, the homologue of FTZ-F1, is required for epidermal and somatic gonad development and also participates in the regulation of moulting [9,10]. In vertebrates, an FTZ-F1 orthologue was first identified as a steroidogenic factor (Ad4BP/SF-1) present in the adrenal gland and able to bind to proximal promoter regions of cytochrome P450 steroid hydroxylase genes (reviewed in [11]). Further studies performed to identify the tissue expression pattern of SF-1 demonstrated its presence in the steroidogenic compartments of the adrenal gland and gonads [12], at the anterior pituitary gland and at the ventromedial hypothalamic nucleus in the brain. Confirming these histological observations, mice knocked out for the *ftz-fl* gene showed female external genitalia irrespective of genetic sex, consistent with an inability to produce testicular androgens, reduced expression of luteinizing hormone and follicle-stimulating hormone, as well as impaired differentiation of adrenal glands and gonads [13]. Furthermore, the use of pituitary-specific knockout mice [14] has shown that SF-1 is particularly involved in the production of luteinizing hormone and follicle-stimulating hormone. A second subfamily of FTZ-F1, encoded by a separate gene now named NR5A2 [15] is represented by LHR-1 (liver receptor homologue-1) in the mouse [16], FTF (α -fetoprotein transcription factor) in the rat [17], PHR-1 in humans [18]

Correspondence to R. Pierce, INSERM U 547, Institut Pasteur de Lille, 1 rue du Prof. A. Calmette, F-59019 Lille, France.

Fax: + 33 3 20 87 78 88, Tel.: + 33 3 20 87 77 83,

E-mail: Raymond.Pierce@pasteur-lille.fr

Abbreviations: AF2-AD, activation function 2-activation domain; DR, direct repeat of the AGGTCA response element; EMSA, electrophoretic mobility shift assay; FTF, α -fetoprotein transcription factor; FTZ-F1, *fushi tarazu* factor 1; HRE-PAL, palindromic repeat of the AGGTCA element; LRH-1, liver receptor homologue-1; MAPK, mitogen activated protein kinase; SF-1, steroidogenic factor 1; SFRE, SF-1 response element; SmFTZ-F1, *Schistosoma mansoni* FTZ-F1.

Note: Ricardo L. de Mendonça and Didier Bouton contributed equally to the work.

*Present address: Université Libre de Bruxelles, Service de Microbiologie, Hôpital Erasme, Route du Lennik, 1070-Bruxelles, Belgium.

†Present address: Department for Biosciences at Novum, Karolinska Institute, 141 57 Huddinge, Sweden.

(Received 18 March 2002, revised 16 July 2002, accepted 26 September 2002)

and FF1 in the zebrafish [19]. In the mouse LRH-1 is expressed mainly in the liver, whereas the rat orthologue is expressed in gut endodermal cells, including the liver and pancreas and has recently been shown to be required for the regulation of a critical gene in the bile acid biosynthetic pathway [20,21]. The functional importance of the vertebrate NR5A2 gene in the development of digestive organs is also shown by its expression pattern during zebrafish development [19].

As part of a wider investigation of the evolution of the nuclear receptor superfamily in the metazoa we used a PCR-based strategy targeting the conserved DNA-binding C domain to isolate five new nuclear receptors from the platyhelminth human parasite, *Schistosoma mansoni* [22]. We are now studying the properties of these receptors to determine the level of conservation of their function and their role in the complex development of this parasite. One of these, SmRXR, has been the subject of a recent report [23]. In this paper we describe the characterization of a FTZ-F1 homologue from *S. mansoni* designated SmFTZ-F1, the first member of this subfamily to be characterized from a lophotrochozoan. This receptor has since been named NR5B1 under the unified nuclear receptor nomenclature [15]. In view of the key role of FTZ-F1 proteins during the development and sexual differentiation of arthropods and vertebrates, SmFTZ-F1 is likely to be involved as a regulator of these pathways in the schistosome. Despite limited sequence identity, particularly of the E domain, SmFTZ-F1 indeed showed functional conservation. Notably, in comparison to the human SF-1 protein, SmFTZ-F1 shared comparable functional features of DNA binding and transcriptional activation in transfected cell lines.

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 centrifuged for 10 min at 1500 g. Schistosomula were obtained *in vitro* [24] and were maintained in culture for up to 8 d under the conditions described previously [25]. Adult worms were obtained by whole-body perfusion of 6-week-old infected hamsters [26]. Eggs were obtained from the livers of infected hamsters and hatched out under light to obtain miracidia [27]. Primary sporocysts were obtained after overnight axenic culture of miracidia as described [27]. Parasite DNA was extracted from the free-living cercariae using standard methods [28]. Total RNA was extracted from all life-cycle stages using the guanidine thiocyanate/caesium chloride method [29] and poly A⁺ RNA was purified on oligo-dT cellulose [30].

Library screening

About 1×10^6 recombinant phage from an adult worm cDNA library constructed in lambda ZAP II (Stratagene), a kind gift of R. Harrop and A. Wilson (University of York, UK), were screened with a 128-bp PCR-generated fragment

corresponding to the C-domain of *S. mansoni* FTZ-F1 [22]. Hybridization was carried out by standard methods [28]. Inserts were sequenced using an Applied Biosystems 377 automated sequencer and methods and reagents of the supplier. In order to extend the cDNA sequence in both directions, 5' and 3'-RACE was carried out using the SMART RACE kit (Clontech) according to the manufacturer's instructions.

Genomic DNA clones containing part of the *Smftz-f1* gene were obtained by screening a *S. mansoni* lambdaEMBL3 library grown at high density using duplicate plaque lifts on Hybond N+ filters with the 2775 bp cDNA insert as a probe labelled by random priming (see Results). In order to obtain the 5' end of the gene we then screened the *S. mansoni* BAC library [31] on high density nylon filters, again using the cDNA insert as a probe. Growth of BAC clones and BAC DNA preparations were as described previously [31]. In order to sequence both lambda and BAC clones, a strategy of gene walking was used, with oligonucleotides initially based on the cDNA sequence, and subsequently on the genomic sequence obtained.

Sequence analysis and phylogenetic tree construction

Alignment of the SmFTZ-F1 E domain with homologues was carried out after prediction of its secondary structure using the PROTEIN SEQUENCE ANALYSIS system programs (Biomolecular Engineering Research Center, Boston University, USA). The prediction is based on technical notes described in [32–34]. For phylogenetic analyses, sequence alignments of SmFTZ-F1 C and E domains with homologues were carried out using the MUST programme which allows alignment by eye [35]. The mouse GCNFI receptor (accession no. NP_034394) was used as an outgroup and for artificial rooting of the phylogenetic tree constructed. Phylogenetic analyses were carried out by distance analysis using NEIGHBOR from the PHYLIP [36] package and by Maximum Likelihood (ML) with TREE-PUZZLE 5.0 [37]. Maximum likelihood analyses were performed using the JTT amino acid substitution model and a rate heterogeneity model with gamma distributed rates over eight categories plus one invariable (JTT + I + Γ). The α parameter and the amino acid frequencies were estimated from the data. The confidence of the nodes was estimated by 1000 bootstrap replicates (PRODIST) and 10 000 quartet puzzling steps (TREE-PUZZLE). The bootstrap replicates of PRODIST were generated using SEQBOOT and compiled in a consensus tree with CONSENSE. In addition we have performed a second ML analysis using the programme MRBAYES [38] with the JTT model and four categories plus one invariable (JTT + I + Γ) in order to confirm the ML tree topology obtained with TREE-PUZZLE.

Northern Blot

Electrophoresis of total RNA from larvae and adult worms (20 μ g per lane) was carried out alongside RNA size markers (Invitrogen life technologies) in a 1.0% (w/v) agarose/3% (v/v) formaldehyde gel [39] that was then blotted onto a Hybond N+ nylon membrane (Amersham). Hybridization with a cDNA probe was carried out as described [28] and blots were exposed overnight to X-Omat AR film (Kodak).

RT-PCR

Reverse transcription of 5 µg of total RNA from each life-cycle stage was carried out using 40 pmoles of random hexamers (Promega) and the Superscript™ kit (Invitrogen life technologies). The resulting cDNA was then amplified in a 50 µL total volume with 10 mM Tris/HCl (pH 9.0), 50 mM KCl, 0.1% (v/v) Triton X-100, 1.5 mM MgCl₂, 0.2 mM dNTPs, 2.5 U of *Taq* DNA polymerase (Promega) and 30–40 pmoles of forward (SmFTZ-F1: CAA CCA gTT gCT ggA ACT AgT ATT C; Sm28GST: ggC gAg CAT ATC AAg gTT ATC) and reverse (SmFTZ-F1: CAC AgC TgC TCg TCA TCT gAA ACC; Sm28GST: CCC AAg AgC TTT CCT gT) primers. After 3 min at 95 °C, 25 cycles of 95 °C for 15 s, 60 °C for 30 s and 72 °C for 1 min were carried out. Analysis of the products was carried out on 1.2% (w/v) agarose gels in Tris/acetate/EDTA buffer. Quantification was carried out by removing aliquots of the polymerase reaction every four cycles starting at the eighth cycle, dot-blotting the samples on to a charged nylon membrane and hybridizing exactly as previously described [40] with ³²P-end-labeled oligonucleotide probes (SmFTZ-F1: CTT CAT CCT CCg gAA CTC CTC AgC g and Sm28GST: CCT CgT TTT CAC CCA TC). The quantity of product for SmFTZ-F1 after 24 amplification cycles was compared to the *S. mansoni* 28 kDa glutathione S-transferase (Sm28GST) product obtained after 16 cycles. Dot blots were scanned using a PhosphorImager (Molecular Dynamics) and the results expressed as the relative intensity of the mean integrated signal (three determinations) for SmFTZ-F1 compared to Sm28GST.

Antibodies

An ovalbumin-coupled peptide (supplied by Synt:em, France) covering the residues 427–441 (AVA-SETAAPEGVSSDD) of SmFTZ-F1 was used to immunize New Zealand Rabbits (IFFA-Credo, France) as described [41]. Sera of immunized rabbits were collected and tested for the presence of specific anti-(SmFTZ-F1) Igs two months after the initial injection using ELISA [42] with uncoupled peptide adsorbed onto Maxisorp plates (Nunc). For Western blotting the purified IgG fraction was used [43]. Rabbit antisera to recombinant Sm28GST [44] and adult worm soluble protein extract were prepared as described [41].

Western blot

Parasites from each life-cycle stage were suspended in 10 mM potassium phosphate buffer, pH 7, containing 154 mM KCl, 1 mM EDTA and 0.1 mM phenylmethylsulfonyl fluoride and sonicated three times for 10 s (maximum power, Microson XL, Misonix). The protein content of the supernatant obtained after centrifugation at 20 000 g for 30 min was measured using the BCA assay kit (Pierce). One µg of protein from each stage was separated on a 10% (v/v) SDS–polyacrylamide gel and blotted on to a nitrocellulose membrane [45]. Blots were developed with the primary antiserum diluted 1 : 500 and the peroxidase-coupled anti-rabbit IgG (Sanofi-Pasteur) at 1 : 7000. Detection was carried out by chemiluminescence using the Renaissance kit (NEN).

Electrophoretic mobility shift assay

Full length *SmFTZ-F1* was cloned into the *Hind*III/*Sma*I restriction sites of pTL1 (a modified version of pSG5; Stratagene), for *in vitro* translation and transient transfection assays. The ORF of human SF-1 cloned into the pJ3Ω vector was a kind gift from P. de Santa Barbara, CNRS UPR 1142, Montpellier, France.

Recombinant SmFTZ-F1 and SF-1 proteins were produced *in vitro* using the rabbit reticulocyte TNT kit (Promega). Electrophoretic mobility shift assays (EMSAs) were performed using 40 × 10³ c.p.m. of ³²P-end-labeled double strand oligonucleotide probe and 2 µL of *in vitro* synthesized proteins. Binding reactions were performed according to [46]. Reaction products were run on a 5% (v/v) native polyacrylamide gel in Tris/borate/EDTA. For supershift experiments, *in vitro*-produced SmFTZ-F1 protein was incubated with polyclonal anti-SmFTZ-F1 Ig for 30 min on ice before adding the end-labelled probe to the binding reaction.

Transient transfection assays

Concatemers of 3× synthetic SF-1 response element were cloned into the pGL-2 luciferase reporter plasmid (Promega). Cell lines were maintained in Dulbecco's modified medium supplemented with 10% (v/v) fetal bovine serum. Cells were transfected by 1 µg total DNA per assay using 4 µL of Ex Gen500 (Euromedex, France) under the conditions recommended by the supplier. The pTL1 plasmid was used as carrier when necessary. Cells were lysed 48 h after transfection and assayed for luciferase activity. For detection of recombinant SmFTZ-F1 expressed in transfected cells, these were cultured on round cover slips for 48 h, washed twice in NaCl/P_i, fixed in NaCl/P_i containing 4% (v/v) paraformaldehyde for 20 min at 4 °C, washed twice in NaCl/P_i, permeabilized in NaCl/P_i containing 0.15% (v/v) Triton X-100 for 2 min, washed twice in NaCl/P_i, incubated for 20 min at 4 °C in NaCl/P_i containing 1% (v/v) ethanolamine and rinsed twice more in NaCl/P_i containing 0.5% (w/v) BSA. Cells were then incubated in the presence of primary antibody diluted 1 : 100 in NaCl/P_i/BSA (0.5%, w/v) for 1 h at 37 °C. After four washes in NaCl/P_i/BSA the cells were incubated in fluorescein isothiocyanate (FITC)-labelled anti-rabbit Ig (Dako) diluted 1 : 100 in NaCl/P_i/BSA for 30 min at 37 °C. They were then washed four times in NaCl/P_i/BSA (0.5%), twice in NaCl/P_i, mounted on slides in fluoprep (BioMérieux) and observed under a fluorescence microscope (Leica) equipped with a Leica WILD camera.

RESULTS

Characterization of a schistosome FTZ-F1 homologue

An adult worm cDNA library was screened with a PCR-generated probe similar to the C domain of members of the FTZ-F1 subfamily [22]. The screening yielded a single clone (2775 bp) encompassing a deduced amino acid sequence of 731 residues and an apparent mass of 78 kDa (GenBank accession number AF158103). Sequence analysis showed that schistosome protein had all the modular domains characteristic of the nuclear hormone receptor superfamily.

Moreover, there were stop codons in all three potential reading frames upstream from the predicted methionine + 1 and no other potential translation initiation codons between this methionine and the first stop codon of the 5' UTR. A second ATG codon is present in the same reading frame just upstream of the C domain (see below). We thus concluded that this clone contained the complete primary sequence. This was confirmed by performing both 5' and 3' RACE on single-stranded cDNA that also allowed us to extend the 5' and 3' UTRs to produce a 3.8 kb sequence. Both the 5' and 3' extensions were confirmed by repeating RACE PCR with primers closer to the new ends. At the 5' end, two alternative sequences were detected, apparently in roughly equal amounts, that corresponded to the alternative splicing of the two 5' exons of the gene (see below). The 3' end of the sequence is not supported by the presence of a classical consensus polyadenylation signal upstream of the detected poly A tail, although the sequence GATAAA is present at -20 to -15 and might constitute such a signal.

Compared to the FTZ-F1 receptors isolated so far, SmFTZ-F1 is among the largest. For example, human, zebrafish, shrimp and *C. elegans* orthologs have 461, 516, 545 and 568 residues, respectively. Only the *Drosophila* homologues FTZ-F1 α and β FTZ-F1/DHR39 [47], which have 1198 and 808 amino acid residues, respectively, are larger. Homology searches of the amino acid sequence clearly place the schistosome protein into the FTZ-F1 group (NR5) of the nuclear receptor superfamily. Figure 1A shows a schematic organization of the SmFTZ-F1 protein. The putative start and end points of each domain are indicated. Sequence alignment of the C domain of SmFTZ-F1 with those of *HsSF-1*, *DmFTZ-F1 α , *DmDHR39*, *DrFF1a*, *MeFTZ-F1* and *CeNHR25* showed between 55% and 63% sequence identity (Fig. 1A,B). Although the identity scores are lower than those observed for other FTZ-F1 proteins, SmFTZ-F1 has a FTZ-F1 box, a specific feature of this group [49] (Fig. 1B). The same analysis performed with the putative ligand-binding domain (E domain) showed lower identity scores, ranging from 14% for *CeFTZ-F1* to 32% for *DmDHR39*. The overall structural features of nuclear receptor ligand-binding domains are retained, however. This is particularly the case for the ligand-binding domain-specific signature, a motif which is common to several members of the nuclear hormone receptor superfamily [23,48], and the activation function 2-activation domain (AF2-AD, Fig. 1C), a core domain that interacts with transcriptional cofactors in a ligand- [50,51] or phosphorylation- [52] dependent manner. In addition, the region described as a dimerization interface mapped at helix 10 (identity box, I-box) in a variety of receptors [53,54], but which has been shown to be involved in coactivator recruitment in the zebrafish FTZ-F1 homologue (*DrFF1A*) [55], is well conserved.*

Organization of the *Smftz-f1* gene and alternative promoter usage

The *Smftz-f1* gene was characterized from a λ EMBL-3 genomic clone and three BAC clones, and completely sequenced (GenBank accession numbers AY028787, AY028788). The overall gene organization is shown in Fig. 2A and comprises 10 exons. The alternative 5' end sequences of the cDNA mentioned above are generated by

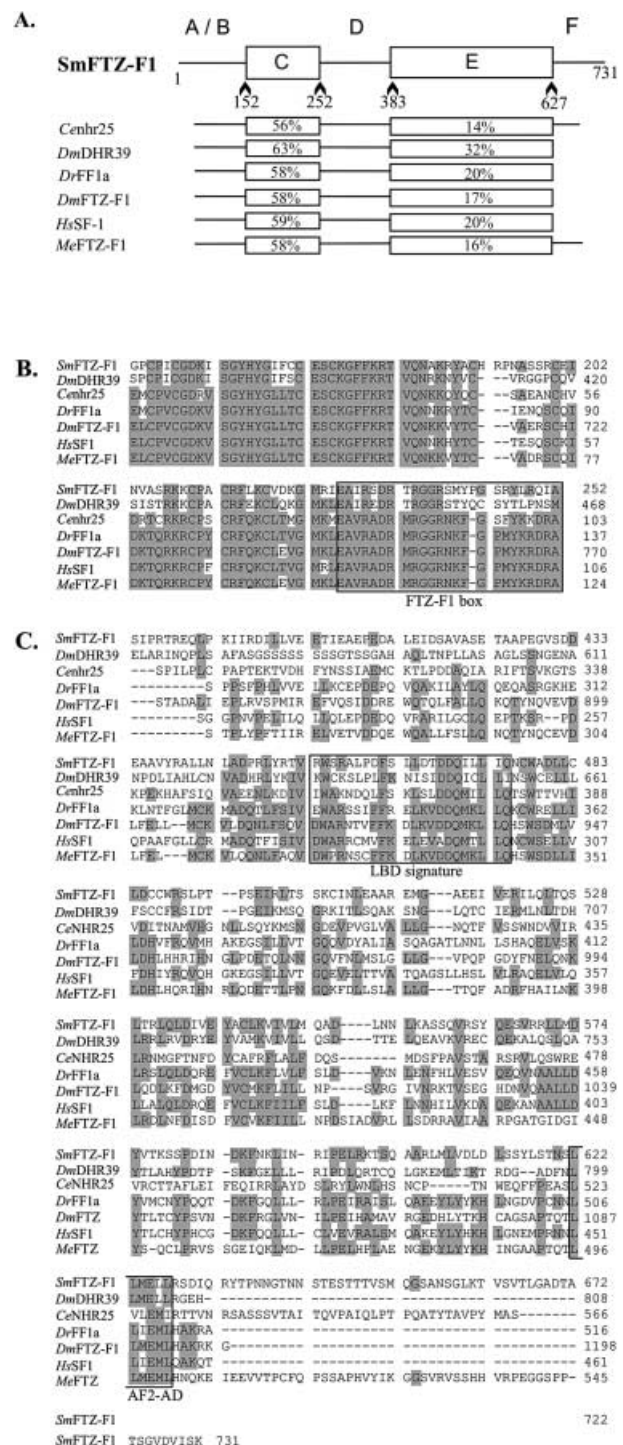


Fig. 1. Alignment of SmFTZ-F1 C and E domains to members of the FTZ-F1 nuclear receptor family. (A) Domain structure of SmFTZ-F1 and levels of identity of the peptide sequences of the C and E domains to those of *C. elegans* nhr25 (Cenhr25, accession no. AF179215), *D. melanogaster* DHR39 (*DmDHR39*, accession no. Q05192) *Danio rerio* FF1a (*DrFF1a*, accession no. AF014926), *D. melanogaster* FTZ-F1 α (*DmFTZ-F1*, accession no. M63711), human SF-1 (*HsSF-1*, accession no. XM_044809) and *Metapenaeus ensis* FTZ-F1 (*MeFTZ-F1*, accession no. AF159132). (B) Alignment of the C domains and the FTZ-F1 box (boxed). Shaded residues are conserved in a majority of the sequences. (C) Alignment of the E domains showing the ligand-binding domain signature region [48] (boxed) and AF2-AD domain (boxed).

the use of alternative promoters and the splicing of exon 1 within exon 2. The two forms thus start either with the exon 1 sequence, or with exon 2. Exon–intron junctions always have GT at the 5' end of the intron and AG at the 3' end. The complete gene contains more introns than any other members of this gene family, apart from *NHR-25* from *C. elegans*, which also has 10 exons [9], and measures approximately 18.4 kb in length. Southern blotting (not shown) indicates that there is only one copy of the gene and this is supported by the presence of only three positive

clones in the *S. mansoni* BAC library which has an approximate eightfold overall genome coverage [31]. Within the coding region the intron–exon structure of *Smftz-f1* is conserved in relation to members of the NR5A2 family (including human *FTF/LRH* (accession no. NT_021968) and zebrafish *ff1* [19]) and the NR5A1 family such as the mouse ELP (SF-1) gene [56], with three of six intron positions conserved. When compared to the *D. melanogaster fitz-f1* gene (accession no. AE003519) only two intron–exon junctions are conserved. One of these positions is also conserved in all the vertebrate genes as well as *Smftz-f1* and is at the end of the C domain just upstream of the Ftz box. Interestingly, none of these intron positions are conserved in the gene encoding *D. melanogaster* DHR39 (accession no. AE003669).

To detect further alternative transcripts of the *Smftz-f1* gene, we performed 5' and 3' RACE PCR with primers located in exons 5 and 6, as well as RT-PCR with primers in exons 1, 2 and 10. The RACE PCR confirmed the presence of the two major splicing isoforms mentioned above but failed to detect any alternative splicing of exons in the coding region. Notably, no isoforms were detected that would lead to the alternative usage of two ATG initiation codons within exon 5 (Fig. 2B). This contrasts with the mouse ELP gene encoding SF-1 among other isoforms [56] in which alternative splicing determines the usage of two ATG initiation codons within the third exon. Furthermore, no splicing isoforms were detected that would alter the coding sequence, encoding for example proteins truncated after the C domain, as in the case of the short variant of *Xenopus laevis* FF1a [57], or which lack the C domain entirely, as with the *C. elegans* *nhr25β* isoform [9]. This was confirmed by PCR on single-stranded cDNA using primers located in exons 1 or 2 and 10. Both the variants generated by the alternative usage of exons 1 and 2 had identical exon compositions, confirmed by sequencing the single PCR products obtained in each case (not shown). Thus, unlike the other members of the FTZ-F1 receptor family, the *Smftz-f1* gene does not give rise to major splicing isoforms encoding different proteins. The significance of the two variants that differ only in the 5' noncoding region remains to be determined.

The promoter region upstream of exon 1 shows some conserved features and similarities to SF-1 promoters in vertebrates (Fig. 2C). A TATA element is present, but this is at –74 and therefore may not be functional, although a TATA element is present at a similar distance from the transcription initiation site in the *S. mansoni* α -tubulin gene [58]. The transcription initiation site itself conforms to the mammalian consensus (TTA₊₁TATA compared to PyPyA₊₁NT/APyPy [59]). Two elements shown to be essential for the expression of the mammalian *fitz-f1* gene [60], a CCAAT box and an E-box, are also present in the promoter. An inverted CCAAT box is at –182 which overlaps with an E-box at –183. A second E-box is present at –10. The second promoter region upstream of exon 2 also contains a TATA element well upstream (–80) of the transcription initiation site, which in this case does not conform to the consensus sequence. There is no proximal CCAAT element, but the E-box at –10 in promoter 1 is at –242 in promoter 2. A striking feature of the latter region is the presence of three tandem and one inverted degenerate repeats of the nuclear receptor consensus response element, AGGTCA. These are, respectively, AGGCTA, AGGTCT

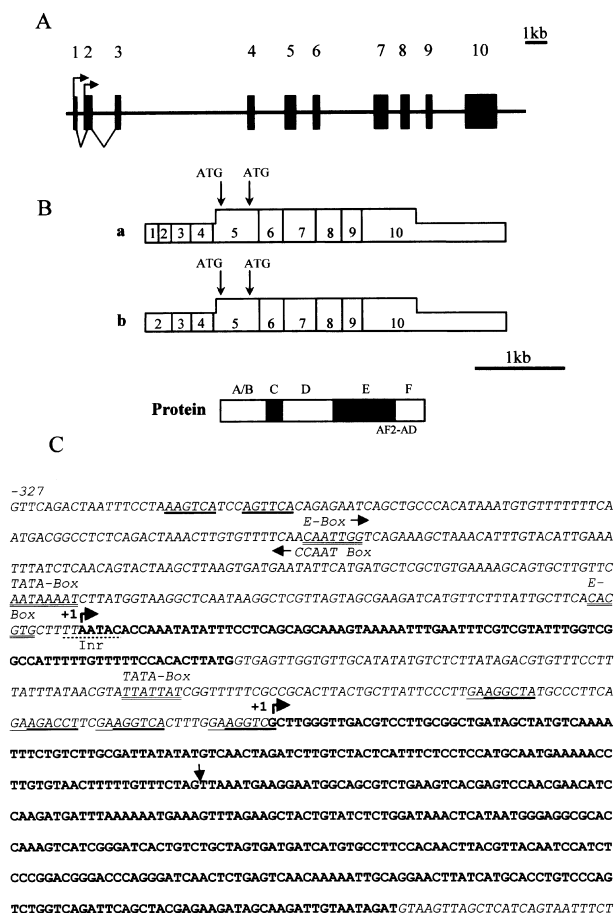


Fig. 2. Structure and alternative promoter usage of the *Smftz-f1* gene. (A) Structure of the *Smftz-f1* gene. Exons are shown by boxes and introns by intervening lines. Alternative promoter usage and splicing at the 5' end are shown below the gene structure. Exon and intron sizes are, respectively, 167, 436, 150, 235, 413, 268, 509, 303, 190 and 1377 bp and 143, 1049, 5824, 1224, 696, 2643, 557, 794 and 1531 bp. (B) Diagram of the transcripts encoded by the *Smftz-f1* gene. Numbers inside boxes represent exon numbers, and translated and untranslated regions are indicated by wide and narrow rectangular boxes, respectively. The position of the two ATG codons in exon 5 are indicated by arrows. The domain structure of the corresponding protein is aligned with the exons making up the transcripts. (C) Nucleic acid sequence of the *Smftz-f1* gene promoter region. Exons 1 and 2 are shown in bold and 5' and intron sequences in italics. The two transcriptional start sites are indicated with bent arrows. The splice site for exon 1 within exon 2 is shown by a vertical arrow. Putative TATA elements, E boxes and a CCAAT box are double underlined. The conserved transcription initiation sequence (Inr) for promoter 1 is underlined in dots. Nuclear receptor response elements are underlined in bold.

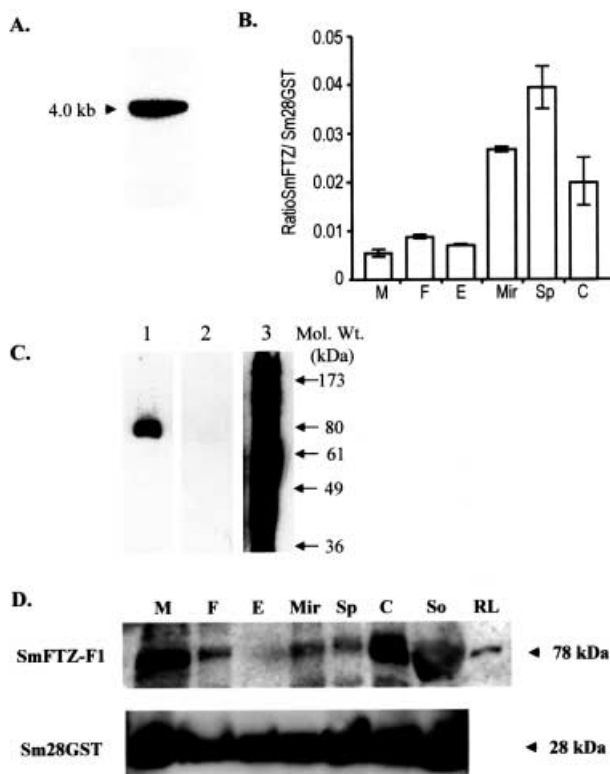


Fig. 4. Expression of SmFTZ-F1 during the schistosome life cycle. SmFTZ-F1 mRNA and protein are differentially expressed at different life-cycle stages. (A) Northern blot of adult worm RNA showing a unique band at 4 kb (B) Semi-quantitative RT-PCR of SmFTZ-F1 mRNA relative to Sm28GST mRNA in adult male worms (M), adult female worms (F), eggs (E), miracidia (Mir), sporocysts (Sp) and cercariae (C). (C) Western blot of protein extract of adult worms probed with (1) antiserum to SmFTZ-F1 peptide, (2) preimmune serum from the rabbit immunized with SmFTZ-F1 peptide, (3) anti-serum to protein extract of adult worms. (D) Western blot of protein extracts of schistosome life cycle stages (as above with the addition of schistosomula, So, and *in vitro* translated SmFTZ-F1, RL) with antisera to SmFTZ-F1 peptide and Sm28GST (separate gels with the same extracts).

detected by RT-PCR (Fig. 4D). Interestingly, cercariae and schistosomula show high levels of the protein, suggesting that its synthesis may be up-regulated immediately prior to parasite invasion of the definitive host.

SmFTZ-F1 has similar functional properties to human SF-1

To determine the DNA binding specificity of SmFTZ-F1, EMSAs were performed with the *in vitro* synthesized SmFTZ-F1 protein and double stranded oligonucleotide probes corresponding to the response element for SF-1, SFRE (TCTAGGTCA). SmFTZ-F1 binds to SFRE as observed in Fig. 5, lane 1. The identity of the protein present in the complex was confirmed by a supershift with specific anti-(SmFTZ-F1) Ig (Fig. 5, lane 4). No such shift was obtained when preimmune serum was added to the protein-DNA complex (Fig. 5, lane 5). The specificity of binding was investigated by competition experiments with unlabelled oligonucleotide competitors (Fig. 5, lanes 2, 3 and

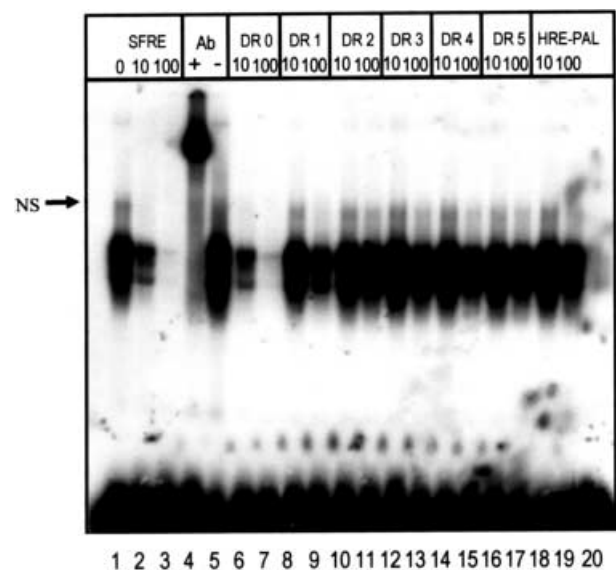


Fig. 5. SmFTZ-F1 binds to the monomeric SF-1 response element (SFRE). EMSA of binding of *in vitro* translated SmFTZ-F1 to a [³²P]end-labelled double-stranded oligonucleotide containing the SF-1 response element (SFRE; TCTAGGTCA, lane 1). Competition assays were carried out with 10- or 100-fold molar excess of SFRE (lanes 2 and 3) or of direct repeats of the AGGTCA core sequence separated by 0 (DR0) to 5 (DR5) nucleotides (lanes 6–17) or of the same sequence in a palindromic repeat (HRE-PAL; lanes 18 and 19). Lanes 4 and 5 show binding to SFRE in the presence of antibody against SmFTZ-F1 and preimmune serum, respectively. A band corresponding to non-specific binding is indicated by the arrow. Lane 20 shows the absence of binding when the empty pTL1 vector was transcribed and translated *in vitro* and the products used in EMSA.

6–19). A 10-fold molar excess of cold SFRE or DR-0 led to a reduction in the signal (Fig. 5, lanes 2 and 6) and a 100-fold excess of the same competitor completely abolished the binding of the labelled probe. This is expected since the DR0 element (AGGTCAAGGTCA) contains a consensus SFRE. Again, as expected, no significant reduction of binding was observed when unlabelled DR-1 to DR-5 (Fig. 5, lanes 8–17) or unrelated HRE-PAL response elements (Fig. 5, lanes 18 and 19) were used as competitors. Finally, no retarded bands were observed when the empty pTL1 vector was used in the assay (Fig. 5, lane 20).

These results led us to investigate the sequence requirements for SmFTZ-F1 binding to SFRE and to compare this to those observed for human SF-1. To do this, EMSAs were performed on *in vitro* synthesized SmFTZ-F1 or SF-1 proteins bound to the wild type radiolabelled SFRE probe. Competition experiments were carried out by adding a 100-fold molar excess of unlabelled, point-mutated SFREs. The ability of each modified SFRE to compete was evaluated by scoring the signal intensity of each shifted band. These scores are reported in Table 1 ranging from – (no competition) to + + + (abolition of the signal). The results summarized in Table 1 clearly show that both SmFTZ-F1 and human SF-1 have the same sequence requirements, substitutions in the second and third nucleotide positions in the nonamer element having the most dramatic effect on the binding of both receptors. We next tested the capacity of SmFTZ-F1 to transactivate transcription from the SFRE

Table 1. Specificity of binding of SmFTZ-F1 and SF-1 to mutated response elements. Binding was assessed in competition EMSA experiments. Scores range from – (no competition) to +++ (abolition of the signal).

Response element	SmFTZ-F1	SF-1
TCA AGGTCA	+++	+++
<u>A</u> CA AGGTCA	++	+++
<u>C</u> CA AGGTCA	++	+++
<u>G</u> CA AGGTCA	++	+++
<u>T</u> GA AGGTCA	+	+
<u>T</u> TA AGGTCA	+	+
<u>T</u> AA AGGTCA	–	–
<u>T</u> CT AGGTCA	+	+
<u>T</u> CG AGGTCA	+	++
<u>T</u> CC AGGTCA	–	–
TCA <u>G</u> GGTCA	++	+
TCA AGGT <u>C</u>	++	++
TCA AGGT <u>C</u> G	+++	+++
TCA AGGT <u>C</u> T	++	+++
TCA AGGT <u>G</u> A	++	++
TCA AGGT <u>T</u> A	+++	+++
TCA AGGT <u>A</u> A	++	+

element in a mammalian cell line. Initially, a control construct constitutively expressing SmFTZ-F1 (SV-40 promoter) was transfected in CV-1 cells. The protein was expressed in the nucleus, as expected (not shown). To investigate the transcriptional properties of SmFTZ-F1, transient cotransfection assays of CV-1 cells were performed with reporter constructs under the control of SFRE sites (3× SFRE). As observed in Fig. 6, SmFTZ-F1 activates transcription through SFRE seven- to eightfold compared to the vector alone. Human SF-1 gave similar results under the same conditions (not shown) confirming our previous DNA-binding data.

DISCUSSION

The *S. mansoni* FTZ-F1 nuclear receptor described here diverges markedly from most arthropod and vertebrate members of this subfamily in terms of its size, peptide sequence and the absence of alternatively spliced isoforms. However, it does conserve the basic functional characteristics of the subfamily. It binds to a monomeric response element with the same specificity as mammalian SF-1. Moreover, it can transactivate transcription of a reporter gene in mammalian cell lines. This demonstrates that it can interact with mammalian coactivators of transcription and that similar cofactors probably exist in schistosomes.

This functional conservation is probably due to the fact that whilst identity scores, particularly those observed for the E domain (Fig. 1) were relatively weak, SmFTZ-F1 presents the modular structure characteristic of the nuclear receptor superfamily and all signatures present in this group. The specific signature of the FTZ-F1 subfamily, the FTZ-F1 box [49], is also observed, although this is truncated in SmFTZ-F1. The ligand-binding domain shows a particularly low level of conservation (14–20%), but the ligand-binding domain-specific signature, located between helix three and five [48], is perfectly conserved in SmFTZ-F1. This signature is a common feature throughout

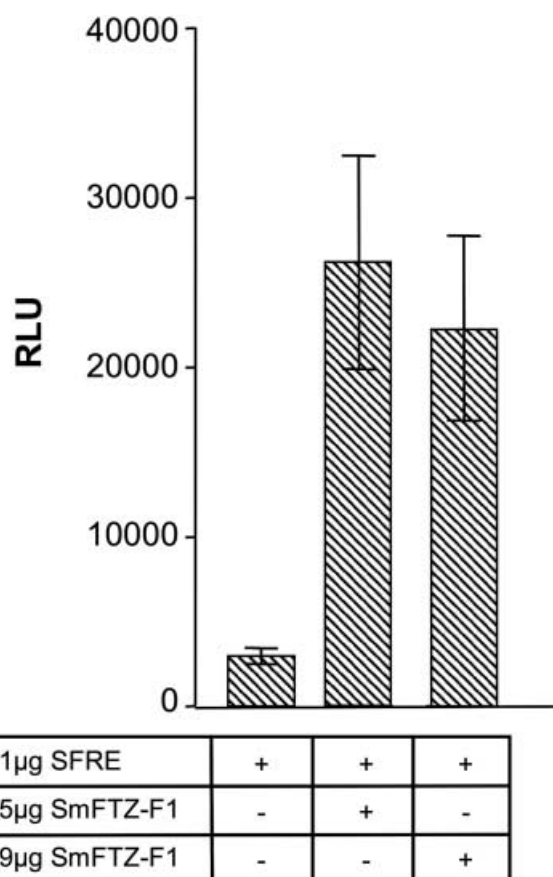


Fig. 6. SmFTZ-F1 transactivates transcription of a reporter gene under the control of the SFRE. CV-1 cells cotransfected with plasmids containing the luciferase gene downstream of a promoter containing the SFRE and expressing SmFTZ-F1 under the control of the SV40 promoter express the reporter gene seven- to eightfold more than cells transfected with the reporter plasmid alone. Results are expressed in Relative Luminescence Units (RLU) and represent the means of three separate experiments performed in triplicate.

the nuclear receptor superfamily and constitutes part of the coactivation-binding surface in a ligand-induced conformational state [50,62]. As a result, there has been considerable speculation that members of the FTZ-F1 family, and in particular SF-1, may be activated by a ligand. 25-Hydroxycholesterol was found to activate transcription by this receptor [63] but the relevance of this observation was refuted by the demonstration that this occurred only with very high concentrations of the ligand and that 25-hydroxycholesterol failed to increase transcription from a variety of SF-1-dependent promoters [64]. The overall low level of sequence identity of the SmFTZ-F1 E domain, in keeping with members of the family from other species, further suggests that even if a ligand exists, it is different from that bound by mammalian SF-1. Moreover, the binding of transcriptional coactivators in a ligand-independent manner is also possible. Mouse SF-1 is activated by phosphorylation at Ser203 by the mitogen activated protein kinase (MAPK) signaling pathway [51]. However, this precise regulation mechanism may not exist in the case of SmFTZ-F1 since the MAPK consensus phosphorylation site (PXnS/TP) present as PYASP in mouse and human SF-1 and as PYTSSP in

human FTF is absent from the schistosome sequence, although potential phosphorylation sites for other kinases are present.

The SmFTZ-F1 AF2-AD domain, one of the two activation function domains present in nuclear receptors and located at helix 12 of the ligand-binding domain, is also conserved. This domain in mammalian SF-1 is required, but not sufficient, for potentiation via coactivators [65] which also requires the phosphorylation of the AF-1 function in the D domain [52]. This, together with the conserved ligand-binding domain-specific signature, indicates that schistosome receptor may interact with transcriptional cofactors which are common among metazoans.

The phylogenetic tree derived from the alignment of the conserved C and E domains of SmFTZ-F1 with FTZ-F1 family members clearly places the schistosome receptor within this family. However, the clustering of SmFTZ-F1 with DHR39 was unexpected. DHR39 is expressed as an 'early late' transcript in third instar larvae under the control of ecdysone [66], and binds to the same response element as FTZ-F1 α [67]. SmFTZ-F1 shares some of the peptide sequence characteristics of DHR39, including an identical AF2 domain and a truncated FTZ-box. However, the *Smftz-fl* gene shares three intron-exon boundaries within the coding region of the gene with mammalian SF-1 and FTF, including a widely conserved intron position at the C-terminal end of the C domain, whereas the *DHR39* gene does not. This may indicate that the schistosome receptor is not a true orthologue of DHR39.

The overall gene structure of *Smftz-fl* is complex, with 10 exons, and the presence of four noncoding exons in the 5' region of the gene is particularly surprising. The significance of the alternative transcripts, initiating either from exon 1, spliced into exon 2, or from the start of exon 2 is unknown. Their existence implies the presence of alternative promoters and the examination of the sequences upstream of the respective transcription initiation sites reveals the presence of elements common notably to the promoters of mammalian SF-1. The presence of overlapping E and CCAAT boxes in promoter 1 (upstream of exon 1) in the *Smftz-fl* gene resembles the close juxtaposition of these elements in the rat *ftz-fl* gene. In the latter case, a cooperative interaction has been demonstrated between the proteins binding these elements [60], and the situation may be similar in the case of the schistosome gene. The putative promoter 2 region strikingly contains four repeats of nuclear receptor response elements. The mammalian *ftz-fl* gene contains one such monomeric element [61] that has been shown to be involved in an autoregulatory loop by which SF-1 regulates the transcription of its own gene. Overall, the presence of these conserved sequence elements may indicate a degree of similarity between the mechanisms of the control of *ftz-fl* gene expression between platyhelminths and mammals.

The alternative transcripts both correspond to full-length mRNAs containing all the coding exons. This is unlike the situation for the mouse ELP/SF-1 gene, in which alternative splicing of two noncoding 5' exons signals the production of distinct coding isoforms [56]. In particular, the splicing of exon 1 within exon 3 leads to the use of an alternative ATG start codon and the production of the SF-1 mRNA, in contrast to the ELP1 isoform that uses an upstream ATG within the same exon. In the *Smftz-fl* gene two possible translation initiating ATGs are present in exon 5, but we

have found no splicing isoforms that lead to the use of one or other of these start codons. Indeed, it is striking that no alternatively spliced transcripts of *Smftz-fl* were found that would encode protein isoforms, despite an extensive search by RT-PCR. The presence of alternatively spliced variants that give rise to distinct protein isoforms is a feature of the FTZ-F1 family in all species so far investigated. It is thus surprising that no such variants were detectable for the *Smftz-fl* gene. Moreover, the functional significance of the alternative promoter usage that we detected is enigmatic. One hypothesis would be that the corresponding mRNAs would interact differently with the translational machinery or have different stabilities, possibly accounting for the differences we detected between the relative amounts of mRNA and the corresponding protein at different life-cycle stages.

Analysis of the expression of *Smftz-fl* by RT-PCR showed that *Smftz-fl* mRNA was detected in all life-cycle stages, with higher levels in larval intermediates miracidia, sporocysts and cercariae (about five times higher than in male and female adult worms). This was previously observed for another schistosome nuclear receptor, SmRXR [23], and probably reflects the high level of protein synthesis characteristic of these stages. Interestingly, the levels of detected protein vary considerably throughout the life cycle, in a manner different from the mRNA levels. The highest levels of SmFTZ-F1 protein were detected in male adult worms and cercariae. In contrast, very low levels were detected in all other intermediates, including miracidia. Our protein preparations were tested using antisera specific for Sm28GST, which has a highly reproducible pattern of expression throughout the life cycle (Fig. 4D), indicating that variations observed for SmFTZ-F1 were not due to differences in protein concentration or degradation. Thus, these results suggest that SmFTZ-F1 is necessary for transcriptional regulation of a number of genes throughout the life cycle, but it is more abundant in mature male worms and cercariae, indicating that SmFTZ-F1 fulfils an important role during the invasion of, and adaptation to, the definitive host. They also suggest that *Smftz-fl* expression could be controlled, at least in part, post-transcriptionally.

FTZ-F1 and its mammalian homologue SF-1, bind as monomers to DNA consensus sequences called SFREs (TCAAGGTCA) [2]. This is in contrast to most other members of the nuclear receptor superfamily, which bind to repeated consensus elements either as homodimers, or with heterodimer partners for transcriptional activation [68]. However, in *Drosophila*, FTZ-F1 α interacts with the homeobox protein *ftz* facilitating its binding to DNA and allowing interactions with weak affinity sites [4,5]. This type of interaction has also been shown for the rat orphan receptor NOR-1 and the Six3 homeodomain protein [69], and may represent a widespread function of orphan receptors. As is the case for other members of the group, we demonstrated that SmFTZ-F1 binds to SFRE. We thus investigated the specificity of binding by competition experiments with unlabelled oligonucleotides. Of the dimeric response elements tested, only DR0, which encompasses an SFRE site, specifically interacted with SmFTZ-F1. Comparison of the binding profiles of SF-1 and SmFTZ-F1 in competition experiments with mutated SFREs showed that these receptors have the same specificity. Various

authors have shown that the members of the FTZ-F1 family bind to essentially the same response element and can compete with each other in *in vitro* assays, as in the case of the *Drosophila* receptors FTZ-F1 α and DHR39 [70].

In keeping with the results of the gel shift experiments, SmFTZ-F1 was able to transactivate transcription of a reporter gene from the SFRE at a similar level to SF-1 in CV-1 cells. This indicates that at least some of the mammalian coactivators are capable of interacting with SmFTZ-F1, most probably through its conserved AF2 domain. This in turn implies that similar cofactors are likely to be present in the schistosome.

In all the metazoan species so far studied, apart from *C. elegans*, two distinct genes encode FTZ-F1 family members that have distinct expression profiles and biological roles. In *C. elegans* only one FTZ-F1 orthologue is present, but the gene encodes two protein isoforms. However, one of these (the nhr25 β isoform) lacks a C domain and may act as an inhibitor. The nhr25 α isoform is crucial in embryo morphogenesis and gonad development. In *S. mansoni* only one *ftz-f1* gene family member has been found so far and this strikingly encodes only one protein. It is therefore possible that one receptor fulfils multiple functions that are shared between different receptors in other metazoans. The presence of SmFTZ-F1 in the parenchyma of adult male worms and in all the life-cycle stages of the parasite argues for such multiple roles in development, and we will next attempt to determine whether this receptor also shares mechanisms for the control of its activity with its vertebrate or ecdysozoan orthologues.

ACKNOWLEDGEMENTS

The work was supported by the Institut National de la Santé et de la Recherche Médicale (U 547), the Centre National de la Recherche Scientifique (UMR 49), the Institut Pasteur de Lille, the Ecole Normale Supérieure de Lyon and the Microbiology programme of the Ministère de l'Éducation Nationale de la Recherche et de la Technologie. DB benefited from grants by the Institut Pasteur de Lille, the Région Nord-Pas de Calais and the Fondation pour la Recherche Médicale. RM was supported by the Fondation des Treilles and by the Institut Pasteur de Lille. HE was supported by the European Molecular Biology Organization. We are grateful for the support of Volvic S.A. in the maintenance of the parasite life cycle.

REFERENCES

- Laudet, V. (1997) Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J Mol Endocrinol.* **19**, 207–226.
- Ueda, H., Sonoda, S., Brown, J.L., Scott, M.P. & Wu, C. (1990) A sequence-specific DNA-binding protein that activates fushi tarazu segmentation gene expression. *Genes Dev.* **4**, 624–635.
- Lavorgna, G., Ueda, H., Clos, J. & Wu, C. (1991) FTZ-F1, a steroid hormone receptor-like protein implicated in the activation of fushi tarazu. *Science* **252**, 848–851.
- Guichet, A., Copeland, J.W., Erdelyi, M., Hlousek, D., Zavorszky, P., Ho, J., Brown, S., Percival-Smith, A., Krause, H.M. & Ephrussi, A. (1997) The nuclear receptor homologue Ftz-F1 and the homeodomain protein Ftz are mutually dependent cofactors. *Nature* **385**, 548–552.
- Yu, Y., Li, W., Su, K., Yussa, M., Han, W., Perrimon, N. & Pick, L. (1997) The nuclear hormone receptor Ftz-F1 is a cofactor for the *Drosophila* homeodomain protein Ftz. *Nature* **385**, 552–555.
- Lawrence, P.A., Johnston, P., MacDonald, P. & Struhl, G. (1987) Borders of parasegments in *Drosophila* embryos are delimited by the fushi tarazu and even-skipped genes. *Nature* **328**, 440–442.
- Lavorgna, G., Karim, F.D., Thummel, C.S. & Wu, C. (1993) Potential role for a FTZ-F1 steroid receptor superfamily member in the control of *Drosophila* metamorphosis. *Proc. Natl Acad. Sci. USA* **90**, 3004–3008.
- Yamada, M., Murata, T., Hirose, S., Lavorgna, G., Suzuki, E. & Ueda, H. (2000) Temporally restricted expression of transcription factor β FTZ-F1: significance for embryogenesis, molting and metamorphosis in *Drosophila melanogaster*. *Development* **127**, 5083–5092.
- Gissendanner, C.R. & Sluder, A.E. (2000) nhr-25, the *Caenorhabditis elegans* ortholog of *ftz-f1*, is required for epidermal and somatic gonad development. *Dev Biol.* **221**, 259–272.
- Asahina, M., Ishihara, T., Jindra, M., Kohara, Y., Katsura, I. & Hirose, S. (2000) The conserved nuclear receptor Ftz-F1 is required for embryogenesis, molting and reproduction in *Caenorhabditis elegans*. *Genes Cells* **5**, 711–723.
- Parker, K.L. & Schimmer, B.P. (1997) Steroidogenic Factor 1: a key determinant of endocrine development and function. *Endocrine Rev.* **18**, 361–377.
- Ikeda, Y., Lala, D.S., Luo, X., Kim, E., Moisan, M.P. & Parker, K.L. (1993) Characterization of the mouse FTZ-F1 gene, which encodes a key regulator of steroid hydroxylase gene expression. *Mol Endocrinol.* **7**, 852–860.
- Luo, X., Ikeda, Y. & Parker, K.L. (1994) A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* **77**, 481–490.
- Zhao, L., Bakke, M., Krimkevich, Y., Cushman, L.J., Parlow, A.F., Camper, S.A. & Parker, K.L. (2001) Steroidogenic factor 1 (SF1) is essential for pituitary gonadotrope function. *Development* **128**, 147–154.
- Laudet, V., Auwerckx, J., Gustafsson, J.A. & Wahli, W. (1999) A unified nomenclature system for the nuclear receptor superfamily. *Cell* **97**, 161–163.
- Tugwood, J.D., Isserman, I. & Green, S. (1991) Mouse liver receptor homologous protein (LRH-1) mRNA, *Genbank. Nucleotide sequence database.*
- Galarneau, L., Pare, J.F., Allard, D., Hamel, D., Levesque, L., Tugwood, J.D., Green, S. & Belanger, L. (1996) The α 1-fetoprotein locus is activated by a nuclear receptor of the *Drosophila* FTZ-F1 family. *Mol Cell Biol.* **16**, 3853–3865.
- Becker-Andre, M., Andre, E. & DeLamarter, J.F. (1993) Identification of nuclear receptor mRNAs by RT-PCR amplification of conserved zinc-finger motif sequences. *Biochem. Biophys. Res. Commun.* **194**, 1371–1379.
- Lin, W., Wang, H.W., Sum, C., Liu, D., Hew, C.L. & Chung, B. (2000) Zebrafish *ftz-f1* gene has two promoters, is alternatively spliced, and is expressed in digestive organs. *Biochem. J.* **348**, 439–446.
- Del Castillo-Olivares, A. & Gil, G. (2000) α 1-fetoprotein transcription factor is required for the expression of sterol 12 α -hydroxylase, the specific enzyme for cholic acid synthesis. *J. Biol. Chem.* **275**, 17793–17799.
- Repa, J.J. & Mangelsdorf, D.J. (1999) Nuclear receptor regulation of cholesterol and bile acid metabolism. *Curr. Opin. Biotechnol.* **10**, 557–563.
- Escriva, H., Safi, R., Langlois, M.C., Saumitou-Laprade, P., Stehelin, D., Capron, A., Pierce, R.J. & Laudet, V. (1997) Ligand binding was acquired during evolution of nuclear receptors. *Proc. Natl Acad. Sci. USA* **94**, 6803–6808.
- De Mendonça, R.L., Escriva, H., Bouton, D., Zelus, D., Vanacker, J.M., Bonnelye, E., Cornette, J., Pierce, R.J. & Laudet, V. (2000) Structural and functional divergence of a nuclear receptor of the RXR family from the trematode parasite *Schistosoma mansoni*. *Eur J Biochem.* **267**, 3208–3219.

24. Ramalho-Pinto, F.J., Gazzinelli, G., Howells, R.E., Mota-Santos, T.A., Figueiredo, E.A. & Pellegrino, J. (1974) *Schistosoma mansoni*: defined system for stepwise transformation of cercariae to schistosomula *in vitro*. *Exp. Parasitol.* **36**, 360–372.
25. Harrop, R.A. & Wilson, R.A. (1993) Protein synthesis and release by cultured schistosomula of *Schistosoma mansoni*. *Parasitology* **107**, 265–274.
26. Smithers, S.R. & Terry, R.J. (1965) The infection of laboratory hosts with cercariae of *Schistosoma mansoni* and the recovery of the adult worms. *Parasitology* **55**, 695–700.
27. Yoshino, T.P. & Laursen, J.R. (1995) Production of *Schistosoma mansoni* daughter sporocysts from mother sporocysts maintained in syntenic culture with *Biomphalaria glabrata* embryonic (Bge) cells. *J. Parasitol.* **81**, 714–722.
28. Sambrook, J. & Russel, D.W. (2001) *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, USA.
29. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. & Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonucleases. *Biochemistry* **18**, 5294–5299.
30. Aviv, H. & Leder, P. (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl Acad. Sci. USA* **69**, 1408–1412.
31. Le Paslier, M.C., Pierce, R.J., Merlin, F., Hirai, H., Wu, W., Williams, D.L., Johnston, D., LoVerde, P.T. & Le Paslier, D. (2000) Construction and characterization of a *Schistosoma mansoni* bacterial artificial chromosome library. *Genomics* **65**, 87–94.
32. Stultz, C.M., White, J.V. & Smith, T.F. (1993) Structural analysis based on space-state modeling. *Protein Sci.* **2**, 305–314.
33. Stultz, C.M., White, J.V. & Smith, T.F. (1994) Protein classification by stochastic modeling and optimal filtering of amino acid sequences. *Math Biosci.* **119**, 35–75.
34. Stultz, C.M., Numbudripad, R., Lathrop, R.H. & White, J.V. (1997) Predicting protein structure with probabilistic models. In *Protein Structural Biology in Bio-Medical Research* (Allewell, N. & Woodward, C., eds). JAI Press, Greenwich, CT, USA.
35. Philippe, H. (1993) MUST, a computer package of Management Utilities for Sequences and Trees. *Nucleic Acids Res.* **21**, 5264–5272.
36. Felsenstein, J. (1995) Phylogeny Inference Package 3.57c edit. Seattle.
37. Strimmer, K. & Von Hessler, A. (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol Biol Evol.* **13**, 954–959.
38. Huelsenbeck, J.P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.
39. Lehrach, H., Diamond, D., Wozney, J.M. & Boedtker, H. (1977) RNA molecular weight determinations by gel electrophoresis under denaturing conditions: a critical reexamination. *Biochemistry* **16**, 4743–4747.
40. Pereira, C., Fallon, P.G., Cornette, J., Capron, A., Doenhoff, M.J. & Pierce, R.J. (1998) Alterations in cytochrome-c oxidase expression between praziquantel-resistant and susceptible strains of *Schistosoma mansoni*. *Parasitology* **117**, 63–73.
41. Vaitukatis, J., Robbins, J.B., Nieschalaf, E. & Ross, G.T. (1971) A method for producing specific antisera with small doses of immunogen. *J. Clin. Endocrinol.* **17**, 988–991.
42. Hancock, C.D. & Evans, G.I. (1992) Production and characterization of antibodies against synthetic peptides. In *Methods in Molecular Biology: Immunochemical Protocols* (Manson, M., ed.), pp. 33–41. Humana Press, Totowa, NJ, USA.
43. Steinbuch, M. & Audran, R. (1969) Isolation of IgG immunoglobulin from human plasma using caprylic acid. *Rev. Fr. Etud. Clin. Biol.* **14**, 1054–1058.
44. Taylor, J.B., Vidal, A., Torpier, G., Meyer, D.J., Roitsch, C., Balloul, J.M., Southan, C., Sondermeyer, P., Pemble, S. & Lecocq, J.P. (1988) The glutathione transferase activity and tissue distribution of a cloned Mr28K protective antigen of *Schistosoma mansoni*. *EMBO J.* **7**, 465–472.
45. Towbin, H., Staehelin, T. & Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels onto nitrocellulose sheets. *Proc. Natl Acad. Sci. USA* **76**, 4350–4354.
46. Vanacker, J.M., Laudet, V., Adelmant, G., Stehelin, D. & Rommelaere, J. (1993) Interconnection between thyroid hormone receptor signalling pathways and parvovirus cytotoxic functions. *J. Virol.* **67**, 7668–7672.
47. Ohno, C.K. & Petkovich, M. (1992) FTZ-F1 β , a novel member of the *Drosophila* nuclear receptor family. *Mech Dev.* **40**, 13–24.
48. Wurtz, J.M., Bourguet, W., Renaud, J.P., Vivat, V., Chambon, P., Moras, D. & Gronemeyer, H. (1996) A canonical structure for the ligand-binding domains of nuclear receptors. *Nat. Struct. Biol.* **3**, 87–94.
49. Ueda, H., Sun, G.C., Murata, T. & Hirose, S. (1992) A novel DNA-binding motif abuts the zinc finger domain of insect nuclear hormone receptor FTZ-F1 and mouse embryonal long terminal repeat-binding protein. *Mol Cell Biol.* **12**, 5667–5672.
50. Kraichely, D.M., Collins, J.J.I., DeLisle, R.K. & MacDonald, P.N. (1999) The autonomous transactivation domain in helix H3 of the vitamin D receptor is required for transactivation and coactivator interaction. *J. Biol. Chem.* **274**, 14352–14358.
51. Durand, B., Saunders, M., Gaudon, C., Roy, B., Losson, R. & Chambon, P. (1994) Activation function 2 (AF-2) of retinoic acid receptor and 9-*cis* retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of the response element on AF-2 activity. *EMBO J.* **13**, 5370–5382.
52. Hammer, G.D., Krylova, I., Zhang, Y., Darimont, B., Simpson, K., Weigel, N.L. & Ingraham, H.A. (1999) Phosphorylation of the nuclear receptor SF-1 modulates cofactor recruitment: integration of hormone signaling in reproduction and stress. *Mol Cell.* **3**, 521–526.
53. Liu, D., Le Drean, Y., Ekker, M., Xiong, F. & Hew, C.L. (1997) Teleost FTZ-F1 homolog and its splicing variant determine the expression of the salmon gonadotropin II β subunit gene. *Mol Endocrinol.* **11**, 877–890.
54. Perlmann, T., Umehono, K., Rangarajan, P., Forman, B. & Evans, R. (1996) Two distinct dimerization interfaces differentially modulate target gene specificity of nuclear hormone receptors. *Mol Endocrinol.* **10**, 958–966.
55. Liu, D., Chandry, M., Lee, S.K., Le Drean, Y., Ando, H., Xiong, F., Woon Lee, J. & Hew, C.L. (2000) A zebrafish ftz-F1 (Fushi tarazu factor 1) homologue requires multiple subdomains in the D and E regions for its transcriptional activity. *J Biol Chem.* **275**, 16758–16766.
56. Ninomiya, Y., Okada, M., Kotomura, N., Suzuki, K., Tsukiyama, T. & Niwa, O. (1995) Genomic organization and isoforms of the mouse ELP gene. *J. Biochem.* **118**, 380–389.
57. Ellinger-Ziegelbauer, H., Glaser, B. & Dreyer, C. (1995) A naturally occurring short variant of the FTZ-F1-related nuclear orphan receptor xFF1rA and interactions between domains of xFF1rA. *Mol Endocrinol.* **9**, 872–886.
58. Duvaux-Miret, O., Baratte, B., Dissous, C. & Capron, A. (1991) Molecular cloning and sequencing of the α -tubulin gene from *Schistosoma mansoni*. *Mol Biochem. Parasitol.* **49**, 337–340.
59. Jahavery, R., Khachi, A., Lo, K., Zenie-Gregory, B. & Smale, S. (1994) DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol.* **14**, 116–127.
60. Daggett, M.A., Rice, D.A. & Heckert, L.L. (2000) Expression of steroidogenic factor 1 in the testis requires an E box and CCAAT box in its promoter proximal region. *Biol. Reprod.* **62**, 670–679.
61. Nomura, M., Nawata, H. & Morohashi, K. (1996) Auto-regulatory loop in the regulation of the mammalian *ftz-fl* gene. *J Biol Chem.* **271**, 8243–8249.

62. Brzozowski, A., Pike, A., Dauter, Z., Hubbard, R., Bonn, T., Engstrom, O., Öhman, L., Greene, G., Gustafsson, J.A. & Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389**, 753–758.
63. Lala, D.S., Syka, P.M., Lazarchik, S.B., Mangelsdorf, D.J., Parker, K.L. & Heyman, R.A. (1997) Activation of the orphan nuclear receptor steroidogenic factor 1 by oxysterols. *Proc. Natl. Acad. Sci. USA* **94**, 4895–4900.
64. Mellon, S.H. & Bair, S.R. (1998) 25-Hydroxycholesterol is not a ligand for the orphan nuclear receptor steroidogenic factor-1 (SF-1). *Endocrinology* **139**, 3026–3029.
65. Crawford, P.A., Polish, J.A., Ganpule, G. & Sadovsky, Y. (1997) The activation function-2 hexamer of steroidogenic factor-1 is required, but not sufficient for potentiation by SRC-1. *Mol Endocrinol.* **11**, 1626–1635.
66. Huet, F., Ruiz, C. & Richards, G. (1995) Sequential gene activation by ecdysone in *Drosophila melanogaster*: the hierarchical equivalence of early and early late genes. *Development* **121**, 1195–1204.
67. Ayer, S., Walker, N., Mosammaparast, M., Nelson, J.P., Shilo, B.Z. & Benyajati, C. (1993) Activation and repression of *Drosophila* alcohol dehydrogenase distal transcription by two steroid hormone receptor superfamily members binding to a common response element. *Nucleic Acids Res.* **21**, 1619–1627.
68. Gronemeyer, H. & Laudet, V. (1995) Transcription factors 3: nuclear receptors. In *Protein Profile*, pp. 1173–1308. Academic Press, London.
69. Ohkura, N., Ohkubo, T., Maruyama, K., Tsukada, T. & Yamaguchi, K. (2001) The orphan nuclear receptor NOR-1 interacts with the homeobox containing protein Six3. *Dev Neurosci.* **23**, 17–24.
70. Ohno, C.K., Ueda, H. & Petkovich, M. (1994) The *Drosophila* nuclear receptors FTZ-F1 β and FTZ-F1 α compete as monomers for binding to a site in the fushi tarazu gene, *Mol Cell Biol.* **14**, 3166–3175.