Available online at www.sciencedirect.com
SCIENCE DIRECT
International Journal for Parasitology 34 (2004) 1211–1219
www.parasitology-online.com

Nucleic acid binding properties of SmZF1, a zinc finger protein of *Schistosoma mansoni*

C.E. Calzavara-Silva[a], F. Prosdocimi[b], F.G.C. Abath[c], S.D.J. Pena[a], G.R. Franco[a,*]

[a]*Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais-UFMG, Avenida Antônio Carlos 6627, Belo Horizonte, MG 31270-010, Brazil*
[b]*Departamento de Biologia Geral, ICB-UFMG, Minas Gerais, Brazil*
[c]*Centro de Pesquisas Aggeu Magalhães, FIOCRUZ, Recife, Pernambuco, Brazil*

Received 7 May 2004; received in revised form 6 July 2004; accepted 12 July 2004

Abstract

During its life cycle, the flat worm *Schistosoma mansoni* is exposed to diverse environmental conditions and changes its morphological form. Each change calls for distinct patterns of gene expression. In order to understand the regulation of gene expression, it is necessary to identify regulatory elements in the promoter region of genes, and DNA transacting factors that control transcription. Zinc finger protein domains are responsible for transcription regulation of diverse genes in a wide range of organisms and are also involved in the promotion of protein–protein interactions. A transcript homologous to zinc finger gene sequences was isolated from a *S. mansoni* adult worm cDNA library and named SmZF1. It codes for a protein of 164 amino acids presenting three C_2H_2 type zinc finger motifs. The recombinant SmZF1 protein was expressed and used on electrophoretic mobility shift assays to investigate the binding specificity of SmZF1 for DNA and RNA oligonucleotides. Our results demonstrated that SmZF1 binds both ds and ss DNA oligonucleotides, with an apparent preference for the specific D1-3DNA oligonucleotide, and also binds RNA oligonucleotides with lower affinity. Although we found that SmZF1 recognises DNA and RNA oligonucleotides not containing putative target sites, SmZF1 binds preferentially to sequence specific sites. Furthermore, unrelated oligonucleotides are not able to abolish this interaction. In silico studies identified putative SmZF1 binding sites in the complete genome of three model organisms and in partial genome sequences of *S. mansoni.* Six *Drosophila* genes presented these binding sites in their promoter region, indicating that they might be controlled by transcription factors containing zinc fingers motifs. Taken together, these results suggest that SmZF1 acts as a putative transcription factor of *S. mansoni*.

Keywords: *Schistosoma mansoni*; Zinc finger; DNA/RNA binding protein; Electrophoretic mobility shift assays

1. Introduction

Schistosomiasis is a human disease caused by several trematodes of the genus *Schistosoma*. In South America, *Schistosoma mansoni* is the only causative agent of Schistosomiasis. Schistosome worms have a complex life cycle living inside molluscs and vertebrate hosts, as well as short periods as larvae swimming freely in water. The distinct morphogenetic forms of the parasite at different developmental stages are determined, to a large extent, by activation/inactivation of a selected group of genes, resulting in modifications in its morphology, physiology and biochemistry. Indeed, several sex and stage-specific genes with different patterns of expression during the schistosome life cycle have been described (Abath et al., 2000; Bickle and Oldridge, 1999; Fantappié et al., 1999; Freebern et al., 1999; Valle et al., 1999; Rabelo et al., 2000; Hoffmann et al., 2001; de Mendonça et al., 2002).

The knowledge of the control of gene expression in *S. mansoni* is in its infancy. However, some genes coding for proteins that act as transcription factors have been described in the parasite. Among them, is the SmPUR-alpha, a highly

* Corresponding author. Tel.: +55-31-3499-2537; fax: +55-31-3499-2984.
E-mail address: gfranco@icb.ufmg.br (G.R. Franco).

0020-7519/$30.00 © 2004 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2004.07.006

conserved 30 kDa protein that shows strong preference for a sequence containing a stretch of alternating pyrimidines, which binds to the p14 gene upstream region and activates the transcription of a reporter gene in yeast (Fantappié et al., 2000). Nabhan et al. (2002), described a 35 kDa protein, named SmPOH, that functions as a positive modulator of transcription by increasing the stability of the cellular protein c-Jun, thereby making greater amounts of this protein available for transactivation of AP-1-responsive genes. Another protein, named SMYB1, is able to bind to oligonucleotides containing CCAAT motifs, and may cause the activation/deactivation of genes in the parasite (Valadão et al., 2002).

The recognition by proteins of specific DNA and RNA sequences is an important control mechanism for a number of important biological processes, including gene expression, recombination and DNA repair (Choo and Klug, 1997; Alvarez et al., 2003; Gaston and Jayaraman, 2003; Werner, 2003; Wray et al., 2003; Agris, 2004). Many of these proteins present zinc finger motifs in their structure, which are mainly responsible for DNA binding in a specific manner (Bernstein et al., 1994; Clarke and Berg, 1998; Yokono et al., 1998; Leon and Roth, 2000; Iuchi, 2001). The zinc finger is a protein domain having its secondary structure stabilised by a zinc ion bound to cysteines and histidines residues. There are several different types of fingers, grouped by the nature of their zinc-interacting residues. Amongst them, the C_2HC, C_2C_2 and C_2H_2 families are more frequently verified in proteins. The classical C_2H_2 zinc finger motif, focused in this study, consists of the amino acid consensus sequence $CX_{2-4}CX_{12}HX_{2-5}H$. The sequence, number and organisation of the zinc finger motifs are important for the biological function of the protein (Berg, 1990; Clarke and Berg, 1998; Wolfe et al., 1999; Leon and Roth, 2000; Iuchi, 2001). Some C_2H_2 fingers can also bind RNA segments and proteins, either forming homodimers or heterodimers, depending on the number, the distance and the sequence between fingers. In general, it is observed that the greater the number of fingers in one protein, the more different ligands it can bind (Mackay and Crossley, 1998; Bird et al., 2000; Laity et al., 2001; Iuchi, 2001).

cDNA sequences coding for proteins containing zinc finger motifs have already been identified in *S. mansoni* and may comprise a significant portion (about 5%) of its genome (Verjovski-Almeida et al., 2003). Souza et al. (2001) have cloned and characterised a 2181 bp gene of *S. mansoni*, named *SmZF1*, which has about 47% sequence similarity with some eukaryotic zinc finger proteins. Computational modelling of the protein predicted the presence of three C_2H_2 zinc finger motifs in its structure. The present communication reports the DNA and RNA binding ability of SmZF1, as well as its preferences and specificity of binding. The characterisation of the molecular factors that determine the specificity of binding of putative regulatory factors in *Schistosoma*, including zinc finger proteins, may contribute to a better understanding of the biology of the parasite, as well as to the evaluation of these proteins as targets for drug therapy (Choo and Isalan, 2000; Nagaoka et al., 2001).

2. Materials and methods

2.1. Oligonucleotides

DNA oligonucleotides were synthesised by Alpha DNA (Montreal, Canada), and RNA oligonucleotides by Integrated DNA Technologies (Coralville, USA). The sequences are listed in Table 1. The specific oligonucleotides D1-3DNA and D2 and the partially random oligonucleotide ZNB, containing the probable binding sites for SmZF1, were designed based on the binding properties of the murine zinc finger protein Zif268 (PDB **1MEY**) (Elrod-Erickson et al., 1998). Oligonucleotides ZIF1, ZIF2 and ZIF3 contain the specific binding sites for the three fingers of Zif268, respectively. VGF (a portion of the Vaccinia virus genome, minus strand, positions 183349-183320; gi **6969640**) and M13R oligonucleotides were chosen because they are not specific for recognition of zinc fingers and do not contain any G rich portions, commonly found in binding sites for these protein motifs. All double stranded (ds) DNA oligonucleotides were prepared as follows: To 1 µl (200 pmol) of single stranded (ss) DNA oligonucleotides was added 1 unit of *Escherichia coli* DNA polymerase I Klenow fragment (Stratagene) in 25 µl total reaction buffer (50 mM Tris–HCl, pH 7.5, 7 mM $MgCl_2$, 1 mM DTT and 20 µM of each dNTP). After 16 h of

Table 1
Sequences of DNA and RNA oligonucleotides used as targets to SmZF1 binding

Oligonucleotide	Sequence
D1-3DNA	CGCCAGGGTTTTCCCAGTCACGA-CAG**GCGTGGGGGTGT**CATAGCTGTTTCCTG
D2	CGCCAGGGTTTTCCCAGTCACGACAG-**GAGTGGGAGAGT**CATAGCTGTTTCCTG
ZIF1	CGCCAGGGTTTTCCCAGTCACGA-CAG**GCGTGGGCGTGT**CATAGCTGTTTCCTG
ZIF2	CGCCAGGGTTTTCCCAGTCACGA-CAG**GCGTGGGACCGT**CATAGCTGTTTCCTG
ZIF3	CGCCAGGGTTTTCCCAGTCACGACAG**GC-GAGGGCACGT**CATAGCTGTTTCCTG
ZNB	CGCCAGGGTTTTCCCAGTCACGA-CAG**G*M*G*W*GGG*MVH*GT**CA-TAGCTGTTTCCTG
M13r	GTCATAGCTGTTTCCTG
VGFDNA	TTATAAAAATGCTAAGTATGCGATGTATCT
D1-3RNA	CGCCAGGGUUUUCCCAGUCACGACAG**GG-GAGGGAGU**GUCAUAGCUGUUUCCUG
VGFRNA	UUAUAAAAAUGCUAAGUAUGCGAUGUAUCU

The possible binding sites of SmZF1 are bolded. The degenerated bases are in italics. The oligonucleotides D1-3DNA, D2, ZIF1, ZIF2, ZIF3, ZNB and D1-3RNA are flanked by the sequences of primers M13F-40 and M13r.

incubation at 25 °C, samples were purified by treatment with phenol/chloroform/isoamilic alcohol (25:24:1), ethanol precipitated and resuspended in 20 µl deionised water.

2.2. Cloning of SmZF1 in prokaryote expression vector

SmZF1 open reading frame was PCR amplified using as template a cDNA clone isolated by Souza et al. (2001). Amplifications were performed using primers flanking the start and stop codons and containing *Eco*RI and *Hin*dIII restriction sites, respectively. The PCR reaction mixtures were prepared in a 50 µl final volume containing 25 ng of template DNA, 0.2 pmol μl^{-1} of each primer, 2 mM $MgCl_2$, 200 µM dNTPs and 5 U of *Taq* DNA polymerase (Phoneutria) in the appropriated buffer (50 mM KCl; 10 mM Tris–HCl, pH 8.4; 0.1% Triton X-100; 1.5 mM $MgCl_2$). The fragment obtained after PCR amplification was double digested with *Eco*RI and *Hin*dIII restriction enzymes, purified and cloned into the pMAL-c2G vector (New England Biolabs). Ligation products were used to transform *E. coli* DH5α strain and the rescued plasmid DNA was sequenced using malE primer (New England Biolabs) and the Thermo Sequenase™ Fluorescent Labelled Primer Cycle Sequencing Kit (Amersham Biosciences) on the A.L.F. DNA Sequencer (Amersham Biosciences).

2.3. Expression and purification of recombinant proteins

Positive *E. coli* clones harbouring the pMAL-c2G and pMAL-SmZF1 constructs were selected and the expression of the proteins Maltose Binding Protein (MBP) and MBP-SmZF1 was induced by 0.6 mM IPTG during 4 h at 37 °C. Cells were harvested by centrifugation at $4000 \times g$ for 20 min and the supernatants were discarded. The cell pellets were three times resuspended in 50 ml column buffer (20 mM Tris–HCl, pH 7.5, 200 mM NaCl, 1 mM EDTA and 1 mM DTT), frozen in dry ice–methanol bath and allowed to thaw at 37 °C. The cell suspensions were centrifuged at $9000 \times g$ for 30 min, and the supernatant containing crude extracts was loaded onto a 10% SDS-PAGE to confirm the presence of the fusion proteins. *E. coli* lysates containing the induced MBP and MBP-SmZF1 fusion protein were diluted 1:5 in column buffer and loaded into an amylose resin column (New England Biolabs). Purification procedures were performed according to the manufacturer's protocol. Protein samples were analysed on 10% SDS-PAGE and visualised by Coomassie blue staining.

2.4. Eletrophoretic mobility shift assay (EMSA)

In the EMSA experiments, different concentrations of the SmZF1 protein were incubated with several ds and ss specific and non-specific oligonucleotides. The reactions were carried out for 15 min at 4 °C, in a final volume of 20 µl. The binding solution consisted of 4 mM Tris–HCl, pH 8.0, 40 mM NaCl and 4 mM $MgCl_2$ in 5% glycerol. After incubation, samples were fractionated in 4% non-denaturing polyacrylamide gel in TBE buffer (89 mM Tris–borate, pH 8.0, 2 mM EDTA), at constant 25 mA at 4 °C, to separate the bound complex from the free oligonucleotides. The resulting gels were stained with VISTRA Green DNA specific dye (Amersham Biosciences), according to the manufacturer's protocol. The radioactive EMSA (R-EMSA) were performed using labelled ds and ss oligonucleotide D1-3DNA and the ss oligonucleotide D1-3RNA. The oligonucleotides were end labelled with γ^{32}-P(dNTP) using T4 polynucleotide kinase (Promega) and incubated with the purified protein SmZF1 (2 µg). The reactions were carried out for 15 min at 4 °C, in a final volume of 20 µl. After incubation, reactions were fractionated in 4% non-denaturing polyacrylamide gels, as described above. Gels were dried under vacuum and exposed to X-ray film with an intensifying screen for about 6 h. For competition experiments, SmZF1 protein was pre-incubated for 15 min at 4 °C with 5, 50, 200 or 500 times molar excess of the appropriate unlabeled specific and non specific competitors (ss and ds) before the addition of the labelled oligonucleotides.

2.5. In silico analyses

SmZF1 binding sites were searched in the complete genome of *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster*, downloaded from Entrez genomes at NCBI (http://www.ncbi.nlm.nih.gov) and also searched in partial genome sequences of *S. mansoni*, obtained from TIGR (http://www.tigr.org) and SANGER Institute (http://www.sanger.ac.uk). The search for SmZF1 binding sites on each organism sequences was performed using the BLAST package algorithm seedtop (ftp://ftp.ncbi.nlm.nih.gov/blast). To find genes putatively controlled by model organisms counterparts of SmZF1, we developed an algorithm named TGFinder (http://dawkins.icb.ufmg.br/). This software looks for a given transcription factor binding sites on sequences in annotated genomes and searches for the genes present at a chosen distance (500 nt) from these binding sites. TGFinder was used only in the genome of model organisms, since there is no GenBank annotated genome file for *S. mansoni*.

3. Results

3.1. SmZF1 is a typical zinc finger, binding both ss and ds DNA oligonucleotides

Because SmZF1 is highly homologous to members of the family of C_2H_2 zinc finger proteins, which includes mainly DNA binding proteins and putative transcription factors, we investigated its ability to bind to DNA segments.

Studies based on the three dimensional structure of the murine protein Zif268 revealed an important aspect of the DNA binding preferences of the mouse protein: the specificity of some residues (located at positions

−1, 2, 3 and 6 of the α-helix, that is part of the finger motif) for particular bases, mainly guanines and cytosines, in the DNA binding site (Elrod-Erickson et al., 1998). Taking advantage of the Zif268 studies, we first designed a partially random oligonucleotide to test the binding activity of SmZF1 in DNA molecules. Looking at the amino acid residues positioned in those locations in the three finger motifs of SmZF1, it was possible to predict the majority of the nucleotides in the binding site of the oligonucleotide. However, for some positions, degenerations were introduced, giving rise to a total oligonucleotide collection composed of 72 different possible sequence combinations, the ZNB oligonucleotide.

We used ss and ds ZNB oligonucleotides (the mixture of the 72 possible oligonucleotide combinations) as probes in EMSA experiments. Fig. 1, representative of three different experiments, shows that SmZF1 was able to interact with both ss and ds ZNB oligonucleotides, leading to the formation of a single complex. Varying the concentrations of the protein (0–2.5 μg), it can be seen that the SmZF1 affinity is apparently the same for ss and ds oligonucleotides. It is important to note that the binding of the fusion protein to the oligonucleotides is done by the SmZF1 component, since MBP was not able to shift the migration of the ZNB oligo, even when an excess of this protein (4 μg) was used.

3.2. SmZF1 binds DNA in a specific manner

Using the ZNB oligonucleotide, it was shown that SmZF1 is able to bind to DNA sequences designed for recognition by zinc finger motifs. However, it was not possible to precise which components of the oligonucleotide collection were responsible for the binding activity. Thus, reporting again to the binding preferences of Zif268, we designed two new oligonucleotides containing the putative binding sites for each of the three SmZF1 fingers: D1-3DNA oligonucleotide, as preferential target for fingers 1 and 3, and D2 oligonucleotide, for finger 2. These two oligonucleotides were used in EMSA, together with ZNB, to verify the binding preferences of the parasite protein. Additionally a non-specific oligonucleotide (VGFDNA) was used as a control. This oligonucleotide does not contain any putative zinc fingers binding site. Fig. 2 shows that the protein binds to the three specific oligonucleotides, as well. However, 0.5 μg of SmZF1 is sufficient to delay the electrophoretic mobility of oligonucleotides D1-3DNA and ZNB, but not oligonucleotide D2, indicating that there is an apparent ordered preference for the D1-3DNA oligonucleotide, followed by ZNB and by D2. It can also be seen that an excess of SmZF1 (2 and 4 μg) forces it to bind the non-specific oligonucleotide VGFDNA.

In order to investigate whether the DNA binding of SmZF1 is sequence-specific, we performed R-EMSA using the labelled oligonucleotide D1-3DNA in competition assays with unlabelled D1-3DNA, D2, ZNB, ZIF1, ZIF2, ZIF3 and the non-specific oligonucleotide M13r. The M13r oligonucleotide was chosen as non-specific competitor because all other oligonucleotides used in this experiment contained a flanking region composed of M13 sequences. At the molar ratio 50:1, only D1-3DNA, D2 and ZNB unlabeled oligonucleotides were able to compete for SmZF1 binding (see Fig. 3). Zif268 specific oligonucleotides (ZIF1, ZIF2 and ZIF3) competed for SmZF1 binding only when present in the ratio 500:1 (unlabelled:labelled). The non-specific oligonucleotide M13r was unable to compete with the oligonucleotide D1-3DNA for SmZF1 binding, even when 500-fold molar excess was used.

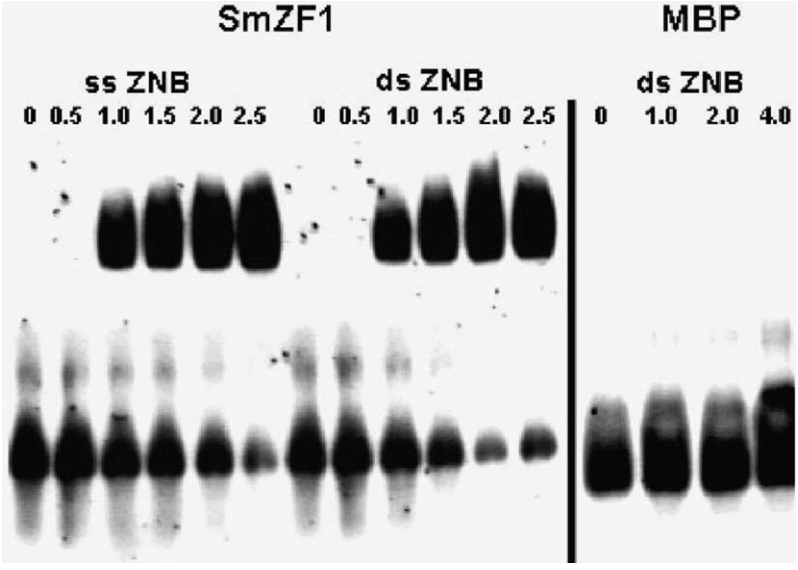

Fig. 1. Electrophoresis mobility shift assay shows that SmZF1 binds both ss and ds DNA oligonucleotides. Proteins SmZF1 (0–2.5 μg) and MBP (0.5–4 μg) were incubated with, respectively, 20 and 25 ng of the ss and ds ZNB oligonucleotide and submitted to electrophoresis in 4% non-denaturing polyacrylamide gel stained with VISTRA Green. The binding of the protein to the oligonucleotide can be visualised by the presence of an upper spot in the gel.

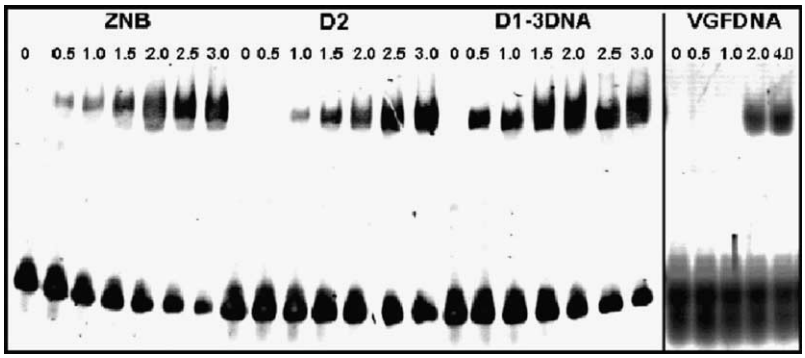

Fig. 2. SmZF1 recognises DNA oligonucleotides in a sequence-specific manner. 0.5–4.0 μg of SmZF1 was incubated with 20 ng of the ss oligonucleotides ZNB, D2 and D1-3DNA, and with 25 ng of the ss oligonucleotide VGFDNA. The samples were submitted to gel shift assay in 4% non-denaturing polyacrylamide gel stained with VISTRA Green.

3.3. RNA binding activity and nucleic acid binding preferences of SmZF1

Because some zinc finger proteins can bind RNA molecules in a specific manner, we tested the binding ability of SmZF1 for RNA oligonucleotides. Fig. 4 clearly shows that SmZF1 binds preferentially to the specific D1-3RNA oligonucleotide, but barely binds to the non-specific oligonucleotide VGFRNA. This was verified only when an excess of protein was used (2 and 4 μg), suggesting that its binding to RNA oligonucleotides is sequence-specific. Further, we compared the DNA/RNA binding specificity of SmZF1, using the SmZF1 specific ss oligonucleotides D1-3DNA, D1-3RNA and the non-specific ss oligonucleotides VGFDNA, VGFRNA and M13r to compete with labelled D1-3DNA and D1-3RNA oligonucleotides for protein binding. Fig. 5 shows that when unlabelled and labelled oligonucleotides were used in the ratio 50:1 only the D1-3DNA was able to self-compete for SmZF1 binding. It is important to note that D1-3DNA, when used in 50× excess, was also able to cross-compete with D1-3RNA. On the other hand, D1-3RNA was able to displace the interaction between the protein and D1-3DNA just when high concentrations of D1-3RNA (500× excess) was used, indicating that SmZF1 binds to RNA segments with lower affinity. Furthermore, it can also be verified that the unlabelled oligonucleotides VGFDNA, VGFRNA and M13r, even when present in 500 fold molar excess, were unable to cross-compete with DNA and RNA labelled oligonucleotides for SmZF1 binding. The MBP protein was not able to shift the electrophoretic mobility of either D1-3DNA or D1-3RNA labelled oligonucleotides.

3.4. SmZF1 binding sites are present on promoter regions of genes from model organisms

Considering that SmZF1 binds specifically to the oligonucleotides D1-3 and D2, we used the seedtop software to search for these oligonucleotides on genome sequences of *S. mansoni* and three model organisms. The sequence pattern of D1-3DNA was more frequently found on all searched genomes, when compared with D2. Both binding

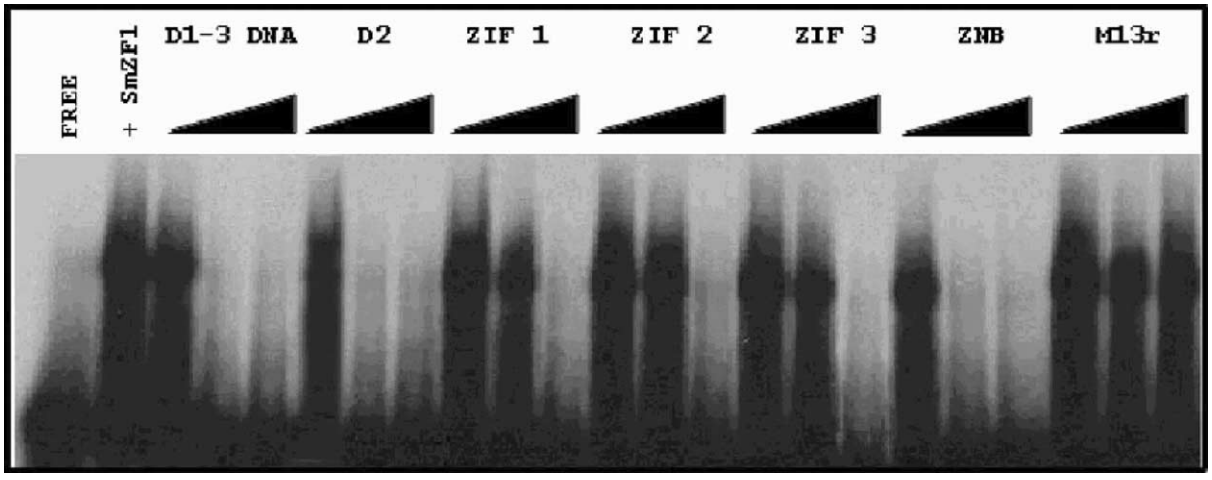

Fig. 3. SmZF1 binds preferentially to specifically designed oligonucleotides. SmZF1 (2 μg) was pre-incubated with ds SmZF1 specific (D1-3DNA, D2), Zif268 specific (ZIF1, ZIF2 and ZIF3), the random ZNB and the non-specific M13r unlabeled oligonucleotides (5×, 50× and 500× concentration of labelled oligonucleotide, shown by incline) and incubated with the ds oligonucleotide D1-3DNA labelled with γ-ATP. After incubation, the samples were submitted to electrophoresis in 4% non-denaturing polyacrylamide gel.

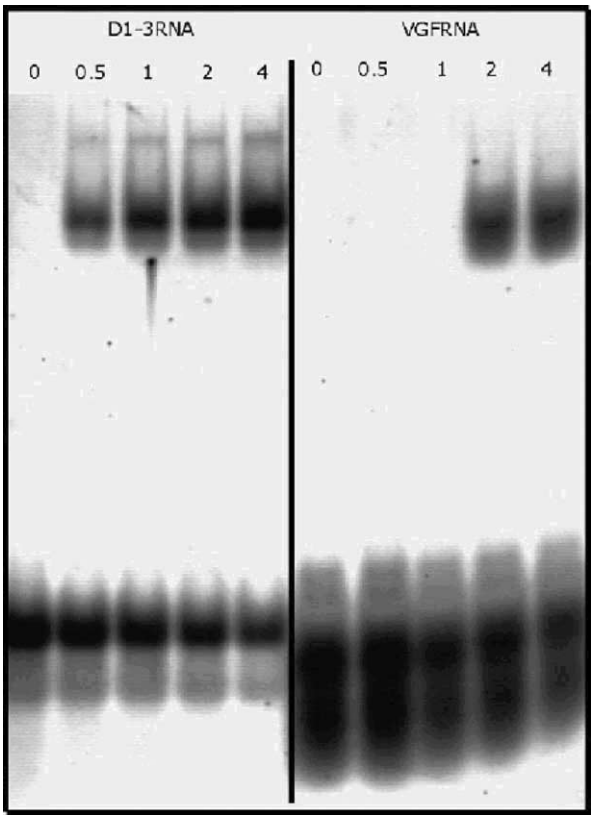

Fig. 4. SmZF1 recognises RNA oligonucleotides in a sequence-specific manner. 0.5–4.0 μg of SmZF1 was incubated with 20 ng of the indicated RNA oligonucleotides and submitted to gel shift assay in 4% non-denaturing polyacrylamide gel stained with VISTRA Green.

sites were more present in the *D. melanogaster* genome, although one hit was found in the *S. mansoni* genome. To identify the SmZF1 target genes we used the software TGFinder that was able to find six genes putatively controlled by a possible SmZF1 counterpart in the *D. melanogaster* genome (see Table 2).

4. Discussion

Considering that SmZF1 is highly homologous to the C_2H_2 zinc finger family, we verified its ability to bind to ss and ds DNA using band shift assays. Based on these experiments, we concluded that SmZF1 binds to ss and ds DNA oligonucleotides apparently with the same affinity. This fact is very interesting since the putative binding sites for each zinc finger motif of SmZF1 are composed by nucleotides present in both strands of a ds DNA molecule. Moreover, we observed that there is an apparent binding preference of SmZF1 for the D1-3DNA oligonucleotide, followed by ZNB and by D2. Such preference for the D1-3DNA oligonucleotide, as compared with D2, could be explained by the fact that fingers 1 and 3 bind with high avidity to the D1-3DNA oligonucleotide, whereas just finger 2 binds with high avidity to oligonucleotide D2. The DNA binding specificity of the *S. mansoni* protein was demonstrated by R-EMSA experiments, confirming the high specificity of the D1-3DNA-SmZF1 binding. However,

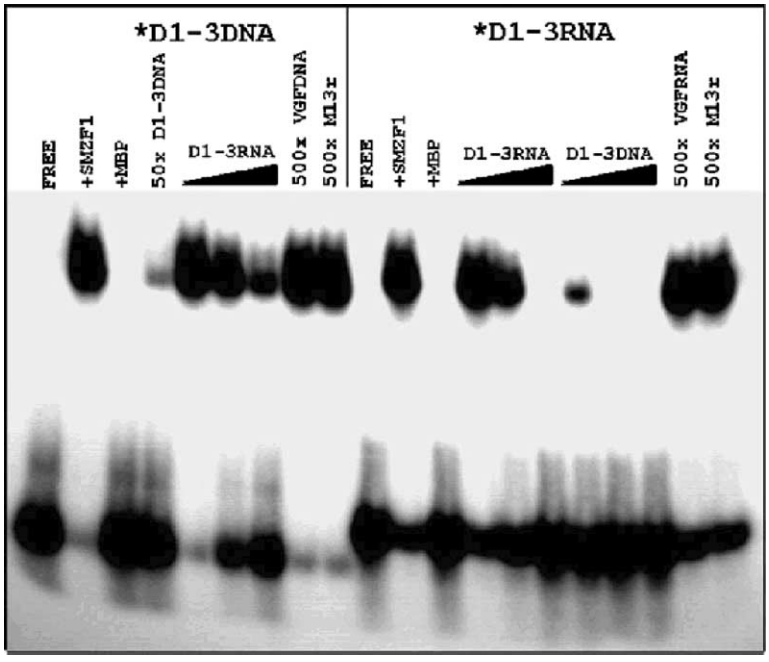

Fig. 5. SmZF1 binds preferentially to specific DNA oligonucleotides when compared to RNA oligonucleotides. SmZF1 (2 μg) was pre-incubated with either specific ss (D1-3DNA, D1-3RNA) or non-specific ss (M13r, VGFDNA and VGFRNA) unlabeled oligonucleotides (50×, 200× and 500× concentration of the labelled oligonucleotide) and incubated with D1-3DNA and D1-3RNA labelled with γ-ATP. After incubation, the samples were submitted to electrophoresis in 4% non-denaturing polyacrylamide gel.

Table 2
In silico analyses to verify the number of SmZF1 binding sites and target genes putatively regulated by these elements present in the genome of model organisms and *Schistosoma mansoni*

Organism	Sequence pattern	Number of patterns	Number of target genes	Description of the putative target gene[a]
Saccharomyces cerevisiae	D1-3DNA	0	0	No genes found
	D2	0	0	No genes found
Caenorhabditis elegans	D1-3DNA	6	0	No genes found
	D2	5	0	No genes found
Drosophila melanogaster	D1-3DNA	49	5	CG16857 gene-involved in cell adhesion, contains immunoglobulin-like domains CG14812 gene-Unknown CG6394 gene-GalNAc-Transferase 2 CG6318 gene-Rad51D CG5024 gene-involved in calcium ion binding, contains an EF-hand family domain
	D2	16	1	CG7344 gene-Unknown
Schistosoma mansoni	D1-3DNA	1	–	–
	D2	0	–	

[a] Information about the genes putatively controlled by the transcription factor was retrieved from FlyBase (http://flybase.bio.indiana.edu/).

SmZF1 was able to bind to a set of oligonucleotides (ZIF1, ZIF2, ZIF3) that does not contain the SmZF1 proper binding site. The binding to these oligonucleotides can be explained by the fact that, although the oligonucleotides are non-specific to SmZF1, they have a guanine rich region that is a putative target to most zinc finger proteins. These inconsistencies may result from context-dependent effects caused by adjacent fingers and the effect of 'inside environments' such as size and type of residue side chains (Iuchi, 2001). It is also observed that when an excess (>2.0 μg) of SmZF1 is incubated with a non-specific oligonucleotide (VGFDNA), the protein shifts the migration of the oligonucleotide, indicating that although SmZF1 shows affinity for its specific targets, it presents binding promiscuity when great amounts of the protein are used. Indeed, some promiscuity in nucleic acid binding proteins has been observed. An example is the protein SmYB1, a member of the Y box family in *S. mansoni*, binds to CCAAT motifs in a specific manner, but also recognises DNA promiscuously (Franco et al., 1997; Valadão et al., 2002).

In addition to DNA, many C_2H_2 zinc finger proteins bind to RNA. The mode of binding is not clear, but the ability of different species and structures of RNA, such as ds RNA, ss RNA, and even DNA-RNA heteroduplexes to recognise the protein indicates that the mode of C_2H_2 zinc finger-RNA binding is not as simple as that of DNA binding. We showed that SmZF1 is a zinc finger protein able to bind to ss RNA molecules. However, the affinity for RNA is lower when compared with DNA. Nakagama et al. (1995) described the binding properties of WT1, a transcription factor coded by the Wilm's tumour gene, containing four C_2H_2 zinc fingers domains. This protein binds to DNA in a very specific manner, using all the four fingers for the recognition of the molecule. WT1 is also able to interact weakly with RNA molecules through just one of its fingers. The ability of SmZF1 to bind to RNA segments, as well as its preference for binding to DNA, either ds or ss, might be explained, as observed in WT1, by the number and differences in amino acid composition among the three fingers. The high avidity of SmZF1 for DNA could be due to binding of two of its three fingers, probably fingers 2 and 3, which are close together, but the third one (finger 1) could improve the binding. In contrast, the binding of SmZF1 to RNA might be performed by just the finger 1.

Two other possible models for recognition of nucleic acids by SmZF1 can be suggested: fingers 1 and 3 may confer the binding affinity of SmZF1 for DNA and finger 2 may be responsible to increase this affinity. On the other hand, this finger could confer binding specificity, as suggested for SP1, a zinc finger protein containing three Kruppel-like fingers (Yokono et al., 1998). Another possibility to explain the action of SmZF1 is that all its three fingers are dedicated to DNA and RNA. The closest fingers 2 and 3 are involved in sequence-specific DNA and/or RNA binding and the N-terminal finger 1 could mediate protein interaction and be involved in SmZF1 dimerisation. Corroborating this hypothesis, the murine protein Ikaros contains six zinc fingers and is able to interact with DNA and proteins. The four N-terminal fingers bind to the DNA and the two C-terminal fingers bind each other to form homodimers. Binding of the N-terminal fingers to the target sequence GGGAA is stronger in the presence of fingers 5 and 6, suggesting that the homodimerisation strengthens the DNA binding (Sun et al., 1996).

Once we predicted the best theoretical DNA binding sites for SmZF1 and demonstrated the interaction between the oligonucleotides D1-3DNA and D2 with SmZF1 by EMSA, we decided to search for genes putatively controlled by

possible SmZF1 counterparts on complete genome sequences of model organisms using the TGFinder software. This software was not used on *S. mansoni* partial genome sequences because they are not annotated. Using the seedtop algorithm just one single pattern (D1-3DNA) was observed in the parasite genome, probably because there is still a limited number of *S. mansoni* sequences in public databases. On the other side, six genes containing binding sites in their promoter regions for possible SmZF1 orthologs were found only in the *D. melanogaster* genome. Amongst them, two are unknown and information regarding the transcription control of the other four identified genes was not found in the current literature.

In this work, we showed that the SmZF1 is a zinc finger protein presenting structural characteristics allowing it to specifically bind to DNA, RNA and proteins. The DNA and RNA binding affinity and specificity of SmZF1 suggest that it can act as a *S. mansoni* transcription factor. The property of SmZF1 to form homo and heterodimers still remains to be tested. Future studies will be directed to further investigate the binding properties, by working with SmZF1 mutated at different binding domains, and to demonstrate the biological activity of the protein.

Acknowledgements

The authors thank Patrícia Nogueira da Gama e Silva and Cláudio Antônio Bonjardim (Laboratório de Vírus-Departamento de Microbiologia, ICB, UFMG) for directions on EMSA and R-EMSA assays. Kátia Barroso and Marina Mourão for carrying out automated DNA sequencing and Neuza Antunes Rodrigues for technical support. This work received financial support from the following sources: FAPEMIG (CBB2475/98) and CNPq (Proc. No. 470425/2001-6).

References

Abath, F.G.C., Xavier, E.M., Allen, R., Gomes, Y.M., Lucena-Silva, N., Baliza, M., Simpson, A.J.G., 2000. Characterisation of Sm13, a tegumental antigen of Schistosoma mansoni. Parasitol. Res. 86, 745–752.

Agris, P.F., 2004. Decoding the genome: a modified view. Nucleic Acids Res. 32, 223–238.

Alvarez, M., Rhodes, S.J., Bidwell, J.P., 2003. Context-dependent transcription: all politics is local. Gene 313, 43–57.

Berg, J.M., 1990. Zinc fingers and other metal-binding domains. Elements for interactions between macromolecules. J. Biol. Chem. 265, 6513–6516.

Bernstein, B.E., Hoffmanm, R.C., Klevitm, R.E., 1994. Sequence-specific DNA recognition by Cys2, His2 zinc fingers. Ann. NY Acad. Sci. 726, 92–102.

Bickle, Q.D., Oldridge, J., 1999. Characterization of a stage-specific Mr16000 schistosomular surface glycoprotein antigen of *Schistosoma mansoni*. Mol. Biochem. Parasitol. 100, 85–94.

Bird, A., Evans-Galea, M.V., Blankman, E., Zhao, H., Luo, H., Winge, D.R., Eide, D.J., 2000. Mapping the DNA binding domain of the ZAP1 zinc responsive transcriptional activator. J. Biol. Chem. 275, 16160–16166.

Choo, Y., Isalan, M., 2000. Advances in zinc finger engineering. Curr. Opin. Struct. Biol. 10, 411–416.

Choo, Y., Klug, A., 1997. Physical basis of a protein-DNA recognition code. Curr. Opin. Struct. Biol. 7, 117–125.

Clarke, N.D., Berg, J.M., 1998. Zinc finger in *Caenorhabditis elegans*: finding families and probing pathways. Science 282, 2018–2022.

de Mendonça, R.L., Bouton, D., Bertin, B., Escriva, H., Noel, C., Vanacker, J., Cornette, J., Laudet, V., Pierce, R.J., 2002. A functionally conserved member of the FTZ-F1 nuclear receptor family from *Schistosoma mansoni*. Eur. J. Biochem. 269, 5700–5711.

Elrod-Erickson, M., Benson, T.E., Pabo, C.O., 1998. High-resolutions structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. Curr. Biol. 6, 451–464.

Fantappié, M.R., Galina, A., Mendonça, R.L., Furtado, D.R., Secor, E.W., Colley, D.G., Corrêa-Oliveira, R., Freeman Jr.., J., Tempone, A.J., Camargo, L.L., Rumjanek, F.D., 1999. Molecular characterization of a NADH ubiquinone oxidoreductase subunit 5 from *Schistosoma mansoni* and inhibition of mitochondrial respiratory chain function by testosterone. Mol. Cel. Biochem. 202, 149–158.

Fantappié, M.R., Osman, A., Niles, E.G., LoVerde, P.T., 2000. Identification and functional characterization of a member of the PUR-alpha family of *Schistosoma mansoni*. Mol. Biochem. Parasitol. 110, 373–390.

Franco, G.R., Garratt, R.C., Tanaka, M., Simpson, A.J., Pena, S.D., 1997. Characterization of a *Schistosoma mansoni* gene encoding a homologue of the Y-box binding protein. Gene 198, 5–16.

Freebern, W.J., Osman, A., Niles, E.G., Christen, L., LoVerde, P.T., 1999. Identification of a cDNA encoding a retinoid X receptor homologue from *Schistosoma mansoni*. Evidence for a role in female-specific gene expression. J. Biol. Chem. 274, 4577–4585.

Gaston, K., Jayaraman, P.S., 2003. Transcriptional repression in eukaryotes: repressors and repression mechanisms. Cell. Mol. Life Sci. 60, 721–741.

Hoffmann, K.F., Davis, E.M., Fischer, E.R., Wynn, T.A., 2001. The guanine protein coupled receptor rhodopsin is developmentally regulated in the free-living stages on *Schistosoma mansoni*. Mol. Biochem. Parasitol. 112, 113–123.

Iuchi, S., 2001. Three classes of C_2H_2 zinc finger proteins. Cell. Mol. Life Sci. 58, 625–635.

Laity, J.H., Lee, B.M., Wright, P.E., 2001. Zinc finger proteins: new insights into structural and functional diversity. Curr. Opin. Struct. Biol. 11, 39–46.

Leon, O., Roth, M., 2000. Zinc fingers: DNA binding and protein–protein interactions. Biol. Res. 33, 21–30.

Mackay, J.P., Crossley, M., 1998. Zinc fingers are sticking together. Trends Biol. Sci. 23, 1–4.

Nabhan, J.F., Hamdan, F.F., Ribeiro, P., 2002. A *Schistosoma mansoni* Pad1 homologue stabilizes c-Jun. Mol. Biochem. Parasitol. 121, 163–172.

Nagaoka, M., Shiraishi, Y., Sugiura, Y., 2001. Selected base sequence outside the target binding site of zinc finger protein Sp1. Nucleic Acids Res. 29, 4920–4929.

Nakagama, H., Heinrich, G., Pelletier, J., Housman, D.E., 1995. Sequence and structural requirements for high-affinity DNA binding by the WT1 gene product. Mol. Cell. Biol. 15, 1489–1498.

Rabelo, E.M., Coelho, P.M., Pena, S.D., 2000. Identification of stage-specific genes in *Schistosoma mansoni*. Ann. NY Acad. Sci. 916, 654–657.

Souza, P.R.E., Valadão, A.F., Calzavara-Silva, C.E., Franco, G.R., de Morais Jr., M.A., Abath, F.G., 2001. Cloning and characterization of SmZF1, a gene encoding a *Schistosoma mansoni* zinc finger protein. Mem. Inst. Oswaldo Cruz. 96, 123–130.

Sun, L., Liu, A., Georgopoulos, K., 1996. Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. Eur. Mol. Biol. Org. J. 15, 5358–5369.

Valadão, A.F., Fantappié, M.R., LoVerde, P.T., Pena, S.D.J., Rumjanek, F.D., Franco, G.R., 2002. Y-box binding protein from *Schistosoma mansoni*: interaction with DNA and RNA. Mol. Biochem. Parasitol. 125, 47–57.

Valle, C., Festucci, A., Calogero, A., Macrý, P., Mecozzi, B., Liberti, P., Cioli, D., 1999. Stage-specific expression of a *Schistosoma mansoni* polypeptide similar to the vertebrate regulatory protein stathmin. J. Biol. Chem. 274, 33869–33874.

Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E., Ojopi, E.P., Paquola, A.C., Piazza, J.P., Nishiyama Jr., M.Y., Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F., Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L., Leite, R.A., Malaquias, L.C., Marques, R.C., Miyasato, P.A., Nascimento, A.L., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A., Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T., Rodrigues, V., Madeira, A.M., Wilson, R.A., Menck, C.F., Setubal, J.C., Leite, L.C., Dias-Neto, E., 2003. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. Nat. Genet. 35, 148–157.

Werner, T., 2003. Promoters can contribute to the elucidation of protein function. Trends Biotechnol. 21, 9–13.

Wolfe, S.A., Nekludova, L., Pabo, C.O., 1999. DNA recognition by Cys2His2 zinc finger proteins. Ann. Rev. Biophys. Biomol. Struct. 3, 183–212.

Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V., Romano, L.A., 2003. The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. 20, 1377–1419.

Yokono, M., Saegusa, N., Matsushita, K., Sugiura, Y., 1998. Unique DNA binding mode of the N-terminal zinc finger of transcription factor Sp1. Biochemistry 37, 6824–6832.