The TATA-box binding protein of *Entamoeba histolytica*: cloning of the gene and location of the protein by immunofluorescence and confocal microscopy

Juan P. Luna-Arias,^{1,3} Rosaura Hernandez-Rivas,¹
Guadalupe de Dios-Bravo,³ Job Garcia¹, Leobardo Mendoza³
and Esther Orozco^{2,3}

Author for correspondence: Esther Orozco. Tel: +52 5 747 7000 ext. 5642. Fax: +52 5 747 7108.
e-mail: esther@mail.cinvestav.mx

^{1,2} Program of Molecular Biomedicine¹, and Department of Experimental Pathology², CINVESTAV, IPN, A.P. 14-740, Mexico 07300 DF, Mexico

³ CICATA, IPN, Legaria 694, Mexico 11590 DF, Mexico

A 309 bp DNA fragment from *Entamoeba histolytica* was amplified by PCR using primers derived from the *Acanthamoeba castellanii* consensus TATA-box binding protein amino acid sequence. The amplified fragment was used to isolate cDNA and genomic DNA clones containing an ORF encoding the complete *E. histolytica* TATA-box binding protein (*Ehtbp*, 702 bp, 234 aa, molecular mass 26 kDa). The EhTBP functional domain showed 55% sequence identity to that of *Homo sapiens*, 54% to *A. castellanii* and 37% to *Plasmodium falciparum* TBPs. In Southern blot experiments we detected a single *Ehtbp* band, which was transcribed as a 1.3 kb mRNA containing a 420 nt 5' untranslated region. However, the probe hybridized with the 0.8 and 1.5 Mb chromosomes, suggesting that this sequence is diploid. *In situ* PCR assays showed two signals in 95% of trophozoites, one located in the nucleus and another in EhkO, the novel DNA-containing organelle recently reported. The recombinant *E. histolytica* TATA-box binding protein was expressed in *Escherichia coli*. Antibodies against it recognized two proteins of 26 and 29 kDa in *E. histolytica* nuclear extracts. Confocal microscopy immunofluorescence analysis located the protein in both the nucleus and EhkO.

Keywords: *Entamoeba histolytica*, TATA-box binding protein, *tbp* gene expression

INTRODUCTION

Entamoeba histolytica is a protozoon responsible for human amoebiasis. Knowing its DNA structure and its mechanisms for regulating gene transcription is a step toward understanding virulence. *E. histolytica* has an A+T rich genome (67 mol%) (Sánchez *et al.*, 1994; Tannich & Horstmann, 1992) and its codon usage is unusual, with 85% A+U in the third codon position (Bruchhaus *et al.*, 1993). Putative TATA boxes with TATTAAA, TATAAA (Sánchez *et al.*, 1994) and TATTTAAA sequences (Bruchhaus *et al.*, 1993; Li *et al.*, 1992) have been located 27 to 50 bp upstream from

the ATG start codon. The ATTCA and ATCA sequences have been described as consensus transcription start sites (Bruchhaus *et al.*, 1993). Remarkably, protein-encoding gene transcription is insensitive *in vitro* to 1 mg α -amanitin ml⁻¹, suggesting that the RNA polymerase differs from those of other eukaryotic organisms (Lioutas & Tannich, 1995), but little is known about the transcription machinery.

Studies of promoter structure and function began recently with the characterization of the *hgl2* and *hgl5* gene promoters, which encode the slightly different 170 kDa heavy subunits of the galactose-inhibitable lectin (Buß *et al.*, 1995; Purdy *et al.*, 1996; Singh *et al.*, 1997). The TATA box, the CCAAT box (*hgl2*) and the GAAC element (*hgl5*) were identified as regulators for full promoter activity *in vivo*. We have studied factors involved in expression of the *EhPgp1* and *EhPgp5* genes, both responsible for the multidrug resistance phenotype

Abbreviations: TAFE, transverse alternating field electrophoresis; TBP, TATA-box binding protein; UTR, untranslated region.

The EMBL accession number for the sequence reported in this paper is Z48307.

in *E. histolytica* (Orozco *et al.*, 1995). The *EhPgp5* gene is induced by the presence of emetine in the medium and has a typical TATA box element that was not found in the promoter of the constitutively expressed *EhPgp1* gene (Gómez *et al.*, 1998; Pérez *et al.*, 1998).

The TATA-box binding protein (TBP) is associated with several transcription factors to make up the multiprotein complexes SL1 (Comai *et al.*, 1992), TFIID (Dynlacht *et al.*, 1991) and TFIIB (Kassavetis *et al.*, 1992; Taggart *et al.*, 1992), which are used by a specific RNA polymerase (White & Jackson, 1992). TBP is the first molecule in the assembly of TFIID, indicating its central role in transcription (Chen *et al.*, 1994). Its functional domain, a 180 aa sequence located at the C-terminus, is highly conserved in most *tbp* genes (Horikoshi *et al.*, 1989; Nikolov *et al.*, 1992; Struhl, 1994). However, the *Plasmodium falciparum* *tbp* sequence diverges significantly from other *tbp* genes. Due to the unusual A + T richness of the *P. falciparum* and *E. histolytica* genomes (Pollack *et al.*, 1982; Bruchhaus *et al.*, 1993; McAndrew *et al.*, 1993), the transcription

factors in these parasites may differ from those reported in humans, making them a good target for drug design. In this paper we report the cloning, molecular characterization and expression of the *E. histolytica* *tbp* gene (*Ehtbp*). The localization of the protein by immunofluorescence and confocal microscopy was also performed.

METHODS

***E. histolytica* cultures.** Trophozoites of *E. histolytica* clone A, strain HM1:IMSS were cultured in TYI-S-33 medium and harvested during the exponential growth phase (Diamond *et al.*, 1978).

Cloning and sequencing of the *Ehtbp* gene. Two oligonucleotide primers (sense: 5' GCTTTACATGCTAGAAAT-GCT 3' and antisense: 5' TGGAAATAATTCTGGTTCATA 3') corresponding to the nucleotide sequence in conserved regions in the C-terminal domain of *A. castellanii* TBP (*Actbp*, bp 354–374 and 651–671) (Wong *et al.*, 1992a) were synthesized (model 381A, Applied Biosystems) using *E. histolytica* codon preferences (Tannich & Hortsman, 1992).

Fig. 1. (a) Comparison of the *E. histolytica*, *A. castellanii*, *P. falciparum* and *H. sapiens* TBP amino acid sequences. Direct repeats are boxed; the basic domain is double underlined; +, identical amino acids; -, gaps introduced during the comparison; +, basic residues in conserved position within the basic domain. Numbers correspond to the residue number for each TBP. (b) Diagram of the complete *Ehtbp* ORF with the elements found in other TBPs: DR1, direct repeat 1 (residues 57–116); DR2, direct repeat 2 (residues 147–207); B domain, basic domain (residues 109–145); σ region, σ -factor homology region (residues 170–201).

Table 1. Percentage amino acid identities in the C-terminal domain of EhtBP and other TBPs

	C terminus*	DR1†	DR2‡	B domain§	σ region¶
<i>A. castellanii</i>	54	65	62	47	59
<i>P. falciparum</i>	37	43	41	40	50
<i>H. sapiens</i>	55	63	64	44	59

* C terminus is the functional domain for each TBP (residues 48–234 in EhTBP).

† DR1 and DR2 correspond to direct repeats 1 (residues 57–116 in EhTBP) and 2 (residues 147–207 in EhTBP), respectively.

‡ B domain is the basic domain (residues 109–145 in EhTBP).

§ σ factor region (residues 170–201 in EhTBP).

For PCR, 100 ng *E. histolytica* genomic DNA was used as the template. Denaturation was carried out at 94 °C for 1 min, annealing at 45 °C for 1 min and extension at 72 °C for 2 min, for a total of 35 cycles. The 309 bp PCR-amplified fragment was sequenced with Sequenase version 2.0 (US Biochemicals). After labelling with [α -³²P]dATP and a Random Primed DNA Labelling Kit (Boehringer), the fragment was used to screen *E. histolytica* cDNA and genomic DNA libraries constructed in λ ZAPII (Descoteaux *et al.*, 1992; Sánchez *et al.*, 1994). Both strands of cDNA and DNA clones containing the full length *Ehtbp* gene were sequenced. The predicted amino acid sequence of EhTBP was compared with other TBP sequences in GenBank using the algorithm of Altschul *et al.* (1990). For hydropathy analyses of EhTBP, the algorithm of Kyte & Doolittle (1982) was used.

Southern and Northern blot analysis. Total DNA from *E. histolytica* was isolated by discontinuous caesium chloride gradients (Sánchez *et al.*, 1994) and digested with various endonucleases. The digested DNA (5 μ g) was electrophoresed through 0.8% (w/v) agarose gels and transferred to nylon membranes (DuPont NEN). On the other hand, *E. histolytica* DNA was separated by transverse alternating field electrophoresis (TAFE) in 1% (w/v) agarose gels and transferred to nylon membranes (Orozco *et al.*, 1993). Hybridization was performed at 42 °C and filters washed at 65 °C (Orozco *et al.*, 1993). For Northern blot experiments, total RNA from *E. histolytica* was isolated using 4 M guanidine isothiocyanate and caesium chloride centrifugation as described (Ausubel *et al.*, 1994; Sánchez *et al.*, 1994). Poly(A)⁺ RNA was purified through poly(U)-Sephadex columns (Gibco-BRL) according to the manufacturer's instructions. Poly(A)⁺ RNA (3–6 μ g) was electrophoresed through formaldehyde 1% (w/v) agarose gels, transferred to nylon membranes and hybridized with *Ehtbp* DNA using the conditions described above.

In situ PCR. This was performed using a GeneAmp *In situ* PCR System 1000 (Perkin Elmer), following the manufacturer's instructions. As primers we used the TBP1 (5' CCGGG-ATCCATGTCAACACCTGGAGA 3') and TBP2 (5' CCGG-AATTCTTATTGATTTGATATATCTTC 3') oligonucleotides.

Determination of the *Ehtbp* transcription start site. Primer extension experiments were performed as described by Gómez *et al.* (1998) and Pérez *et al.* (1998) with some modifications. Total RNA (50 μ g) was used to extend the pextbp (5' ACTCAGTTGAAGTTGACATA 3') or ptbpn-01 (5' TGT-CTTTTCATTTAAGGCTCA 3') primers. 100 U SuperScript II (Gibco BRL) were used per reaction; the conditions used were 94 °C for 1 min, 42 °C for 10 min and 37 °C for 10 min.

Expression and purification of EhTBP. The full length *Ehtbp* was cloned in-frame into the pRSETA expression vector (Invitrogen) after PCR amplification of the *E. histolytica* genomic clone with the oligonucleotides TBP1 and TBP2. A recombinant plasmid was isolated and used to transform *E. coli* BL21(DE3) allowing expression of the recombinant EhTBP as a fusion polypeptide at the N-terminus, containing 4 aa (MRGS) followed by a 6 H tag and the sequence GMASMTGGQMQMRDLYDDDDKDRWGS, in which there is an enterokinase recognition sequence. Expression of recombinant EhTBP was induced by 1 mM IPTG for 1.5 h. Cells were harvested by centrifugation and the recombinant EhTBP was partially purified under denaturing conditions through a Ni²⁺-NTA affinity column (QIAGEN). Further purification was obtained by 12% SDS-PAGE and electroelution of the protein (Ausubel *et al.*, 1994).

Production, purification of anti-EhTBP rabbit polyclonal antibodies and Western blot analysis. Recombinant EhTBP

(100 μ g) was electroeluted from a 12% SDS-polyacrylamide gel and subcutaneously injected with 500 μ l Freund's complete adjuvant into a New Zealand rabbit (2–4 months old). Before the serum was obtained, three additional immunizations were done at 2–3 week intervals using Freund's incomplete adjuvant. Anti-EhTBP-specific antibodies were affinity purified using the recombinant EhTBP immobilized on a nitrocellulose membrane (Ausubel *et al.*, 1994). Western blot assays were done using the purified recombinant TBP and nuclear extracts (Gómez *et al.*, 1998; Pérez *et al.*, 1998) and the ECL detection system (Amersham).

Confocal and immunofluorescence microscopy. Trophozoites were grown overnight on sterile coverslips, fixed with 3.8% (v/v) paraformaldehyde, permeabilized with 0.1% (v/v) Triton X-100 and blocked with 1% (w/v) BSA. The fixed trophozoites were incubated with a 1:10 dilution of rabbit anti-EhTBP polyclonal antibody and then with a 1:1000 dilution of goat anti-rabbit polyclonal antibody labelled with Cy2 Dye (Amersham). Finally, cells were counterstained with 0.5 mg propidium iodide ml⁻¹ in PBS for 5 min and examined through a Nikon microscope attached to a laser confocal scanning system MRC 1024 (Bio-Rad).

RESULTS AND DISCUSSION

Cloning the *Ehtbp* gene

A 309 bp fragment of *E. histolytica* DNA was amplified by PCR using primers from bp 354–374 and 651–671 of the *A. castellanii* *tbp* gene (not shown). This fragment was used to isolate two identical genomic DNA and cDNA clones, indicating the absence of introns in the *Ehtbp* gene, as also reported for other *E. histolytica* genes (Descoteaux *et al.*, 1992; Sánchez *et al.*, 1994). The genomic DNA clone contained a 2190 bp insert with a

Fig. 2. Southern and Northern blot analysis of the *E. histolytica* *tbp* gene. (a) Southern blot of *E. histolytica* genomic DNA digested with *EcoRI* and *AccI* enzymes. (b) Southern blot of high molecular mass DNA separated by TAFE. (c) Northern blot of poly(A)⁺ RNA.

Fig. 3. Confocal microscopy of trophozoites. (a) Exponentially growing cells were used to amplify the full length *Ehtbp* gene by *in situ* PCR with the oligonucleotides TBP1 and TBP2, and Cy5-dCTP. Cells were observed in the blue channel. (b) Cells counterstained with propidium iodide and observed in the red channel. (c) Transmission microscopy of the cells observed in (a) and (b). (d). Cells treated as in (a), but no AmpliTaq was used in the *in situ* PCR reaction. N, nucleus; EhkO, EhkO organelle.

702 bp ORF encoding a predicted 234 aa (26 kDa) polypeptide with a pI of 9.27. The predicted amino acid sequence of this ORF showed 55% identity to the *H. sapiens* TBP, 54% to the *A. castellanii* TBP and 37% to the TBP of *P. falciparum* (Fig. 1a). *Ehtbp* contained an ATCA consensus sequence 8 bp upstream of the ATG start codon, which is similar to the sequence described as a transcription start site (Bruchhaus *et al.*, 1993). In addition, downstream of the stop codon we located three TAATT sequences at 48, 75 and 180 nt, proposed as the polyadenylation signal for several genes (Bruchhaus *et al.*, 1993), and at 76 nt an AATTAA sequence proposed as the polyadenylation signal for the *Ehtatub1* gene (Sánchez *et al.*, 1994). As reported for other genes (Tannich & Horstmann, 1992), the *Ehtbp* gene has 68% A+U base preference in the coding regions.

Characterization of the EhTBP sequence

The EhTBP C-terminal domain (residues 48–229) comprises 182 residues and its alignment with other TBP functional domains showed 55% and 54% identity with the *H. sapiens* and *A. castellanii* TBPs, respectively, and only 37% identity with the TBP from *P. falciparum* (Fig. 1a; Table 1), which is a protozoan with a similar A+T content in its genome.

The EhTBP C-terminal domain (Fig. 1b) contains all the structural elements identified in other TBPs. It has two

Fig. 4. Determination of the transcription start site for the *Ehtbp* mRNA. Primer extension experiments were performed with the oligonucleotide ptbpn-01 and total *E. histolytica* RNA (50 µg). The arrow shows the primer extension product that mapped to the ACGC sequence (coding strand). X, primer extension lane.

Fig. 5. Purification and immunodetection of the recombinant EhTBP. (a) 12 % SDS-PAGE. Lanes: 1, molecular mass markers; 2 and 3, *E. coli* BL21(DE3) cells transformed with pRSETA (lane 2) or pRSETA/Ehtbp plasmids (lane 3), induced with 1 mM IPTG as described in Methods; 4, recombinant EhTBP purified by Ni²⁺-NTA affinity chromatography. (b). Western blot analysis of molecular mass markers (lane 1), 100 µg *E. histolytica* nuclear extracts (lane 2) and 500 ng purified recombinant EhTBP (lane 3).

direct repeats (DR1 and DR2, residues 57–116 and 147–207, respectively; Fig. 1b) described in other organisms as responsible for protein–DNA binding (Nikolov *et al.*, 1992). The EhTBP repeats show 41–65 % identity to those of *A. castellanii*, *P. falciparum* and *H. sapiens* TBPs (Table 1). A region with homology to the sigma factor (σ factor 2.3 and 2.4 domains) (Horikoshi *et al.*, 1989), located at residues 170–201, is contained in DR2 (Fig. 1b). A highly basic region (residues 109–145) connects the two direct repeats in EhTBP (Fig. 1b). This core has eight basic residues, seven of which are conserved in all TBPs studied (Fig. 1a). We found 40–47 % identity with the *A. castellanii*, *P. falciparum* and *H. sapiens* TBPs in this region of EhTBP (Table 1). The *Arabidopsis thaliana* TBP2 basic domain forms an α -helix that is involved in TFIIA–TBP interactions (Nikolov *et al.*, 1992).

The EhTBP N-terminal domain is 47 aa long and rich in hydrophobic residues (Fig. 1a). In contrast, the human TBP N-terminus is 158 aa long, but is also rich in hydrophobic residues, with a segment of 38 consecutive glutamines (Q-run) (Peterson *et al.*, 1990) that was not

Fig. 6. Confocal immunofluorescence microscopy of EhTBP. *E. histolytica* trophozoites were exponentially grown and treated with anti-EhTBP rabbit polyclonal antibodies and counterstained with propidium iodide. (a) Control cells treated only with the goat anti-rabbit polyclonal antibodies coupled to Cy2 Dye and observed in the green channel. (b) Cells stained with propidium iodide and observed in the red channel. (c) Cells treated with anti-EhTBP rabbit polyclonal antibodies and goat anti-rabbit polyclonal antibodies coupled to Cy2 Dye and observed in the green channel. (d) Cells treated as in (b) and (c) observed simultaneously in both channels. N, nucleus; EhkO, EhkO organelle.

found in the EhTBP N-terminal domain. One of the challenges in the study of human parasites is to find differences between host and parasite molecules that can be used as drug targets, and to design vaccine targets. Differences in the TBP N-terminal domains could be useful to identify *E. histolytica*-specific transcription factors or other molecules involved in gene-expression regulation, like the DNA-dependent protein kinase that hyperphosphorylates the N-terminal domain of *Xenopus* TBP (Labhart, 1996). This phosphorylation was seen to be fully dependent on the presence of the C-terminal core domain.

Copy number of the *Ehtbp* gene

To assess the copy number of the *Ehtbp* gene, we carried out Southern blot experiments using both DNA digested with several restriction enzymes and high molecular mass DNA separated by TAFE (Orozco *et al.*, 1993). We used the full-length gene as a probe. Single bands of 7 and 20 kb were observed in DNA digested with *EcoRI* and *AccI*, respectively, enzymes that cut outside the ORF (Fig. 2a). Analysis with DNA digested with several other enzymes also revealed a single band (data not shown), suggesting that there is a single copy of the *Ehtbp* gene in the genome. The PCR fragments, and the cDNA and genomic DNA clones contained the same sequence, also suggesting the presence of a unique gene. However, in TAFE-separated DNA, *Ehtbp* hybridized with 0.8 and 1.5 Mb chromosomes (Fig. 2b). These two bands, reproducibly obtained in TAFE, could correspond to two copies of the same gene located on different chromosomes, suggesting diploidy of *Ehtbp* or cross-hybridization with other sequences in the genome.

In situ PCR

To investigate the cellular location of the *Ehtbp* gene we carried out *in situ* PCR experiments. The TBP1 (sense) and TBP2 (antisense) oligonucleotides used for the *in situ* PCR amplified the 702 bp full length *tbp* gene. In 100% of the cells the Cy5 fluorochrome detected the *tbp* gene in the nuclei (Fig. 3). In contrast, 95% of the trophozoites also presented small fluorescent spots outside the nucleus, probably corresponding to the EhkO organelle. This strongly suggests that there is more than one copy of the *tbp* gene in trophozoites. The two copies of the *tbp* gene must show similar sequence, because a single restriction band was found in Southern blot experiments (Fig. 2a) using the full length gene as a probe, total DNA and enzymes cutting outside the gene. Thus, the two DNA bands detected in TAFE experiments using whole trophozoites may correspond to similar *tbp* genes located on different DNA molecules, one in the nucleus and another in EhkO. However, some reports indicate that there is a TBP-related factor in *Drosophila* that is capable of replacing TBP for *in vitro* transcription (Buratowski, 1997; Hansen *et al.*, 1997). We have not yet investigated the presence of a TBP-related factor in *E. histolytica*, whose gene could cross-hybridize with *Ehtbp*.

Transcription of the *Ehtbp* gene

Northern blot experiments revealed a 1.3 kb transcript in poly(A)⁺ RNA (Fig. 2c). This mRNA is larger than the *Ehtbp* ORF, which is 702 bp long. These results agree with other reports indicating that several *tbp* genes are transcribed as larger mRNAs than their ORFs. The ORFs of the *Drosophila* (Hoey *et al.*, 1990), human (Peterson *et al.*, 1990) and *Plasmodium* (McAndrew *et al.*, 1993) *tbp* genes are 1059, 1017 and 684 bp long, respectively, while their transcripts are 1.6, 1.8–2 and 1.8 kb, respectively. As in other organisms, *Ehtbp* mRNA contains untranslated regions (UTRs), probably important for its stability or localization. Human and *Drosophila* *tbp* genes have long 5' UTRs (241 and 173 nt, respectively) (Peterson *et al.*, 1990; Hoey *et al.*, 1990), while *A. castellanii* *tbp* contains a 67 nt 5' UTR (Wong *et al.*, 1992b).

To confirm that the *tbp* gene has a long 5' UTR and that it was transcribed in a single transcript, we carried out primer extension experiments using the pexttbp oligonucleotide located 100 bases downstream of the ATG start codon (not shown) or the ptbpn-01 oligonucleotide located at 150 bases upstream of the ATG. In all cases, primer extension results showed a single transcription start site for the *tbp* gene at the ACGC sequence located 420 bases upstream of the ATG codon (Fig. 4). The *Ehtbp* mRNA 3' UTR may be 180 bases long to give the 1.3 kb transcript. Most *E. histolytica* genes have short 5' UTRs (Bruchhaus *et al.*, 1993), although the *EhPAK* and *EhMCM3* mRNAs have 265 and 126 nt 5' UTRs (Gangopadhyay *et al.*, 1997). The *Ehtbp* 5' UTR is the longest reported for amoebic genes. In these three cases, the transcription start site sequences (CAATT, TTAA and ACGC for *EhPAK*, *EhMCM3* and *Ehtbp* mRNA, respectively) differed from the ATTCA or ATCA sequences reported as the consensus start sites (Bruchhaus *et al.*, 1993). A putative TATA box (ATTAATT) was located at 28 bp upstream of the transcription start site.

Expression of EhTBP

We cloned the *Ehtbp* ORF in-frame into the pRSETA expression vector and its expression was IPTG induced. Transformed bacteria produced a 30 kDa polypeptide (Fig. 5a), 4 kDa larger than the molecular mass deduced from the *Ehtbp* ORF. This was due to the small fusion polypeptide encoded by the expression vector (described in Methods). Rabbit polyclonal antibodies against the recombinant EhTBP were generated and they detected two 26 and 29 kDa polypeptides in nuclear extracts (Fig. 5b). The 29 kDa protein could be the phosphorylated form of the 26 kDa TBP, as has been reported for other TBPs (White *et al.*, 1995; Ghavidel & Schultz, 1997; Chibazakura *et al.*, 1997), although protein degradation can not be disregarded.

Intracellular location of EhTBP

Confocal microscopy and immunofluorescence analysis revealed that EhTBP is located in the nucleus and in

subcellular structures corresponding to EhkO, the novel organelle recently reported (Orozco *et al.*, 1997). In agreement with our *in situ* PCR experiments (Fig. 3), EhTBP colocalized with propidium-iodide-stained material in EhkO and in the nucleus, confirming that this novel organelle contains genetic material, which may be transcriptionally active. As reported, the diameter of EhkO varied from 0.5 to 5 μ m (Orozco *et al.*, 1997). We are currently studying EhkO function and the nature of its genetic material.

In summary, *Ehtbp* encodes a 234 aa protein with a molecular mass of 26 kDa. It contains a 182 aa conserved C-terminal domain with all the structural elements that have been found in other TBPs (two direct repeats, a basic domain and a σ -factor homology domain). As in other organisms, the 47 aa N-terminal domain is rich in hydrophobic residues and is not conserved, differing totally from human TBP. Antibodies against the recombinant protein detected two polypeptides of 26 and 29 kDa in nuclear extracts and immunostained both the nucleus and EhkO. *In situ* PCR also revealed that these organelles contained at least one copy of *Ehtbp*.

ACKNOWLEDGEMENTS

Dr E. Orozco is an International Research Scholar from the Howard Hughes Medical Institute (USA). This work was also supported by CONACyT (Mexico).

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D., Seidman, J. G., Smith, J. A. & Struhl, K. (1994). *Current Protocols in Molecular Biology*. New York: Wiley.
- Bruchhaus, I., Leippe, M., Lioutas, C. & Tannich, E. (1993). Unusual gene organization in the protozoan parasite *Entamoeba histolytica*. *DNA Cell Biol* **12**, 925–933.
- BuB, H., Lioutas, C., Dobinsky, S., Nickel, R. & Tannich, E. (1995). Analysis of the 170 kDa lectin gene promoter of *Entamoeba histolytica*. *Mol Biochem Parasitol* **72**, 1–10.
- Buratowski, S. (1997). Multiple TATA-binding factors come into style. *Cell* **91**, 13–15.
- Chen, J.-L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. & Tjian, R. (1994). Assembly of recombinant TFIID reveals differential coactivators requirements for distinct transcriptional activators. *Cell* **79**, 93–105.
- Chibazakura, T., Watanabe, F., Kitajima, S., Tsukada, K., Yasukochi, Y. & Teraoka, H. (1997). Phosphorylation of human general transcription factors, TATA-binding protein and transcription factor IIB by DNA-dependent protein kinase synergistic stimulation of RNA polymerase II basal transcription *in vitro*. *Eur J Biochem* **247**, 1166–1173.
- Comai, L., Tanese, N. & Tjian, R. (1992). The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor SL1. *Cell* **68**, 965–976.
- Descoteaux, S., Ayala, P., Orozco, E. & Samuelson, J. (1992). Primary sequences of two P-glycoprotein genes of *Entamoeba histolytica*. *Mol Biochem Parasitol* **54**, 201–212.
- Diamond, L. S., Harlow, D. R. & Cunnick, C. C. (1978). A new medium for the axenic cultivation of *Entamoeba histolytica* and another *Entamoeba*. *Trans R Soc Trop Med Hyg* **72**, 431–432.
- Dynlacht, B. D., Hoey, T. & Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**, 563–576.
- Gangopadhyay, S. S., Ray, S. S., Sinha, P. & Lohia, A. (1997). Unusual genome organisation in *Entamoeba histolytica* leads to two overlapping transcripts. *Mol Biochem Parasitol* **89**, 73–83.
- Ghavidel, A. & Schultz, M. C. (1997). Casein kinase II regulation of yeast TFIIB is mediated by the TATA-binding protein. *Genes Dev* **11**, 2780–2789.
- Gómez, C., Pérez, D. G., López-Bayghen, E. & Orozco, E. (1998). Transcriptional analysis of the *EhPgp1* promoter of *Entamoeba histolytica* multidrug-resistant mutant. *J Biol Chem* **273**, 7277–7284.
- Hansen, S. K., Takada, S., Jacobson, R. H., Lis, J. T. & Tjian, R. (1997). Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* **91**, 71–83.
- Hoey, T., Dynlacht, B. D., Peterson, M. G., Pugh, B. F. & Tjian, R. (1990). Isolation and characterization of the *Drosophila* gene encoding the TATA box binding protein, TFIID. *Cell* **61**, 1179–1186.
- Horikoshi, M., Wang, C. K., Fujii, H., Cromlish, J. A., Weil, P. A. & Roeder, R. G. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature* **341**, 299–303.
- Kassavetis, G. A., Joazeiro, C. A. P., Pisano, M., Geiduschek, E. P., Colbert, T., Hahn, S. & Blanco, J. A. (1992). The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIB. *Cell* **71**, 1055–1064.
- Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105–132.
- Labhart, P. (1996). Phosphorylation of the N-terminal domain of *Xenopus* TATA-box binding protein by DNA-dependent protein kinase depends on the C-terminal core domain. *FEBS Lett* **386**, 110–114.
- Li, E., Kunz-Jenkins, C. & Stanley, S. L., Jr (1992). Isolation and characterization of genomic clones encoding a serine-rich *Entamoeba histolytica* protein. *Mol Biochem Parasitol* **50**, 355–358.
- Lioutas, C. & Tannich, E. (1995). Transcription of protein-coding genes in *Entamoeba histolytica* is insensitive to high concentrations of α -amanitin. *Mol Biochem Parasitol* **73**, 259–261.
- McAndrew, M. B., Read, M., Sims, P. F. G. & Hyde, J. E. (1993). Characterization of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T-rich human malaria parasite *Plasmodium falciparum*. *Gene* **124**, 165–171.
- Nikolov, D. B., Hu, S.-H., Lin, J., Gasch, A., Hoffman, A., Horikoshi, M., Chua, N.-H., Roeder, R. G. & Burley, S. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* **360**, 40–46.
- Orozco, E., Báez-Camargo, M., Gamboa, L., Flores, E., Valdés, J. & Hernández, F. (1993). Molecular karyotype of related clones of *Entamoeba histolytica*. *Mol Biochem Parasitol* **59**, 29–40.
- Orozco, E., Pérez, D. G., Gómez, M. C. & Ayala, P. (1995). Multidrug resistance in *Entamoeba histolytica*. *Parasitol Today* **11**, 473–475.
- Orozco, E., Gharaibeh, R., Riveron, A. M., Delgadillo, D. M., Mercado, M., Sanchez, T., Gomez-Conde, E., Vargas, M. A. &

- Lopez-Revilla, R. (1997).** A novel cytoplasmic structure containing DNA networks in *Entamoeba histolytica* trophozoites. *Mol Gen Genet* **254**, 250–257.
- Pérez, D. G., Gómez, C., López-Bayghen, E., Tannich, E. & Orozco, E. (1998).** Transcriptional analysis of the *EhPgp5* promoter of *Entamoeba histolytica* multidrug-resistant mutant. *J Biol Chem* **273**, 7277–7284.
- Peterson, M. G., Tanese, N., Pugh, B. F. & Tjian, R. (1990).** Functional domains and upstream activation properties of cloned human TATA binding factor. *Science* **248**, 1625–1630.
- Pollack, Y., Katzen, A. L. & Golenser, J. (1982).** The genome of *Plasmodium falciparum*. I: DNA base composition. *Nucleic Acids Res* **10**, 539–546.
- Purdy, J. E., Lana, T. P., Mann, B. J. & Petri, W. A., Jr (1996).** Upstream regulatory elements controlling expression of the *Entamoeba histolytica* lectin. *Mol Biochem Parasitol* **78**, 91–103.
- Sánchez, M. A., Peattie, D. A., Wirth, D. & Orozco, E. (1994).** Cloning, genomic organization and transcription of the *Entamoeba histolytica* α -tubulin-encoding gene. *Gene* **146**, 239–244.
- Singh, U., Rogers, J. B., Mann, B. J. & Petri, W. A., Jr (1997).** Transcription initiation is controlled by three core promoter elements in the *hgl5* gene of the protozoan parasite *Entamoeba histolytica*. *Proc Natl Acad Sci USA* **94**, 8812–8817.
- Struhl, K. (1994).** Duality of TBP, the universal transcription factor. *Science* **263**, 1103–1104.
- Taggart, A. K. P., Fisher, T. S. & Pugh, B. F. (1992).** The TATA-binding protein and associated factors are components of Pol III transcription factor TFIIB. *Cell* **71**, 1015–1028.
- Tannich, E. & Horstmann, R. D. (1992).** Codon usage in pathogenic *Entamoeba histolytica*. *J Mol Evol* **34**, 272–273.
- White, R. J. & Jackson, S. P. (1992).** The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. *Trends Genet* **8**, 284–288.
- White, R. J., Gottlieb, T. M., Downes, C. S. & Jackson, S. P. (1995).** Mitotic regulation of a TATA-binding-protein containing complex. *Mol Cell Biol* **15**, 1983–1992.
- Wong, J.-M., Liu, F. & Bateman, E. (1992a).** Cloning and expression of the *Acanthamoeba castellanii* gene encoding transcription factor TFIID. *Gene* **117**, 91–97.
- Wong, J.-M., Liu, F. & Bateman, E. (1992b).** Isolation of genomic DNA encoding transcription factor TFIID from *Acanthamoeba castellanii*: characterization of the promoter. *Nucleic Acids Res* **20**, 4817–4824.

Received 12 May 1998; revised 20 August 1998; accepted 28 August 1998.