Parasitology Research (2019) 118:3019–3031
https://doi.org/10.1007/s00436-019-06438-z

PROTOZOOLOGY - ORIGINAL PAPER

An in vitro characterisation of the *Trichomonas vaginalis* TATA box-binding proteins (TBPs)

Olivia Parra-Marín[1] · Lluvia Rosas-Hernández[1] · Karla López-Pacheco[1] · Bernardo Franco[1,2] · Alexandra Ibáñez-Escribano[1,3] · Roberto Hernández[1] · Imelda López-Villaseñor[1]

Received: 8 May 2019 / Accepted: 22 August 2019 / Published online: 31 August 2019

Abstract

The protozoan parasite *Trichomonas vaginalis* is a common human pathogen from one of the earliest-diverging eukaryotic lineages. At the transcriptional level, the highly conserved Inr element of RNA pol II-transcribed genes surrounds the transcription start site and is recognised by IBP39, a protein exclusive of *T. vaginalis*. Typical TATA boxes have not been identified in this organism but, in contrast, BLAST analyses of the *T. vaginalis* genome identified two genes encoding putative TATA-binding proteins (herein referred to as TvTBP1 and TvTBP2). The goal of this work was to characterise these two proteins at the molecular level. Our results show that both TvTBPs theoretically adopt the saddle-shaped structure distinctive to TBPs and both Tv*tbp* genes are expressed in *T. vaginalis*. TvTBP1 did not complement a *Saccharomyces cerevisiae* mutant lacking TBP; however, TvTBP1 and TvTBP2 proteins bound *T. vaginalis* DNA promoter sequences in EMSA assays. We propose that TvTBP1 may be part of the preinitiation transcription complex in *T. vaginalis* since TvTBP1 recombinant protein was able to bind IBP39 in vitro. This work represents the first approach towards the characterisation of general transcription factors in this early divergent organism.

Keywords *Trichomonas vaginalis* · Transcription · TBP · IBP39

Introduction

An important factor for eukaryotic and archaebacterial transcription is the TATA box-binding protein (TBP) (White and Jackson 1992; Hernandez 1993; Vannini 2013). Classically, in TATA box-containing promoters, TBP binds to the TATA box element (TATAWAWN) in the minor groove and drastically distorts the structure of the DNA (Kim and Burley 1994; Juo et al. 1996). This interaction results in the recruitment of the pre-initiation complex (PIC) to the promoter, which positions RNA polymerase II (RNA pol II) over the gene transcription start site (Hahn 2004). However, several *cis* elements have been described at the promoter sites, rendering PIC formation highly complex and variable (Danino et al. 2015).

TBP was first identified as the TATA box-binding protein subunit of Transcription Factor IID (TFIID), which suggested an exclusive function in the transcription of RNA pol II promoters containing a TATA box (Hernandez 1993; Thomas and Chiang 2006). Nevertheless, TBP is involved in initiation complexes formed on TATA-containing and TATA-less promoters of genes transcribed by RNA polymerases I, II and III (Hernandez 1993); therefore, TBP has been considered a key component of the transcription initiation complexes in the three RNA polymerases: the SL1 (Selectivity Factor 1) complex for RNA pol I, TFIID for RNA pol II and TFIIIB for RNA pol III. TBP is recruited to promoters through different strategies that involve either direct binding to the DNA, in the

Olivia Parra-Marín and Lluvia Rosas-Hernández contributed equally to this work.

Handling Editor: Julia Walochnik

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00436-019-06438-z) contains supplementary material, which is available to authorized users.

✉ Imelda López-Villaseñor
imelda@biomedicas.unam.mx

1 Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Ciudad Universitaria, Mexico City, Mexico

2 Present address: Departamento de Biología, División de Ciencias Naturales y Exactas, Universidad de Guanajuato, Guanajuato, Gto., Mexico

3 Present address: Departamento de Microbiología y Parasitología, Facultad de Farmacia, Universidad Complutense de Madrid, Madrid, Spain

case of TATA box-containing promoters, or protein-protein interactions, in the case of TATA-less promoters (Rigby 1993; Hahn 2004; Vannini 2013). In addition, all three polymerases require structurally and functionally related TFIIB-like factors for transcription initiation: Rrn7 (or TAF1B in human) for RNA pol I, TFIIB for RNA pol II and TFIIB-related factor 1 (BRF1) for RNA pol III. TBP is associated with these TFIIB-like factors in the three transcription initiation complexes, which are required for polymerase recruitment to the promoter (Vannini and Cramer 2012). Many metazoan eukaryotes contain multiple TBP paralogs (TBP-related factors, TRFs) that have specific and distinct functions in vivo and may or may not bind to DNA (Davidson 2003; Akhtar and Veenstra 2011). A phylogenetic analysis of TBPs from eukaryotes (Castañon-Sanchez et al. 2010) showed that these proteins can be grouped into three branches: (1) TBPs from most eukaryotes, (2) TBP-related proteins and (3) two proteins homologous to TBPs from *T. vaginalis*, which have not been previously characterised although they are referred to as TBPs in several papers (Castañon-Sanchez et al. 2010; Smith and Johnson 2011; Koster et al. 2015).

Our experimental system, *T. vaginalis*, is the aetiologic agent of the most common sexually transmitted infection worldwide (World Health Organization 2011). *T. vaginalis* belongs to the phylum Parabasalia within the kingdom Excavata, argued to be among the earliest-diverging eukaryotic lineages (Gunderson et al. 1995). At the genomic level, protein-coding genes are composed of bipartite promoters, with a conserved core promoter region and gene-specific distal regulatory elements (Vanacova et al. 2003; Smith and Johnson 2011). RNA pol II core promoters in *T. vaginalis* possess a highly conserved initiator element (Inr) (Quon et al. 1994; Liston and Johnson 1999) that is recognised by the amino terminal domain of the Inr-binding protein IBP39 (Liston et al. 2001; Lau et al. 2003; Schumacher et al. 2003), whereas the carboxy terminal domain recognises the RNA polymerase II CTD (Schumacher et al. 2003; Lau et al. 2006). Additional core elements have been described (Ong et al. 2004; Smith et al. 2011), but a classic TATA box has not been identified in *T. vaginalis* promoters (Liston and Johnson 1999; Smith and Johnson 2011). The only exception is the finding of a canonical TATA-box element positionally conserved in the *T. vaginalis* U6 snRNA gene. However, the function of this putative TATA-box has not been determined (Simoes-Barbosa et al. 2008). Two genes encoding putative TBPs are annotated in the *T. vaginalis* genome (Carlton et al. 2007), but the function of these proteins (herein referred to as TvTBP1 and TvTBP2) has not been experimentally addressed. The aim of the present work was to characterise these two proteins at the functional and biochemical levels. We show here that (i) both Tv*tbp* genes are expressed in *T. vaginalis*, (ii) TvTBP1 was not able to complement *Saccharomyces cerevisiae* cells lacking TBP, (iii) TvTBP1 recombinant protein interacted with IBP39 in pull-down assays and (iv) TvTBP1 and TvTBP2 bound *T. vaginalis* DNA sequences in electrophoretic mobility shift assays (EMSAs).

Materials and methods

In silico analysis of TvTBPs

A BLASTP search of the *T. vaginalis* genome with the human TBP (NP_003185.1) or *S. cerevisiae* TBP (NP_011075.3) identified two hypothetical protein sequences: XP_001329427.1 and XP_001300812.1 **(TrichDB:** TVAG_291560 and TVAG_285070, respectively). A percent identity matrix of various TBPs (ESM 1 Table) was created with Clustal Omega using C-terminal amino acid sequences (Sievers et al. 2014). TvTBP1 (region 48-235), TvTBP2 (region 45-232) and ScTBP (region 61-240) amino acid sequences were modelled using Modeller in the Bioinformatics Toolkit platform (Webb and Sali 2016; Zimmermann et al. 2018) using *S. cerevisiae* TBP (4boa crystal, chain A) as the template. The quality of the models was analysed with the Verify 3D program (Bowie et al. 1991; Lüthy et al. 1992) and visualised with the Chimera program (Pettersen et al. 2004). Multiple sequence alignment was performed using the Clustal Omega (version 1.2.4) program in the UniProt web server.

T. vaginalis cell culture

The *T. vaginalis* strain CNCD147 (Alvarez-Sánchez et al. 2000) was used throughout the experimental work and cultured as previously described (Diamond 1957; Clark and Diamond 2002) using trypticase-yeast extract-maltose (TYM) medium supplemented with 10% heat-inactivated horse serum at 37 °C.

Expression of the TvTBPs

Non-quantitative reverse transcription polymerase chain reactions were performed with mRNA poly(A+) to amplify cDNA fragments, ~ 708 and ~ 702 bp long, which correspond to the expected sizes for Tv*tbp1* and Tv*tbp2* transcripts, respectively. Briefly, the cells were grown to a density of approximately 1.5×10^6 cells/ml. Total RNA was then extracted using the TRIzol reagent (Invitrogen). The RNA poly(A+) fraction was isolated from total RNA using a commercial oligodeoxyribonucleotide (oligo)-dT kit (illustra™ QuickPrep™ Micro mRNA Purification Kit, GE Healthcare, UK). For cDNA synthesis, each reaction was set up using ~ 500 ng of RNA poly(A+), 200 ng of the AP primer (Invitrogen) and 0.5 mM each dNTP. The samples were incubated at 65 °C for 5 min, followed by 1 min on ice. One hundred units of SuperScript™ III reverse transcriptase (Invitrogen), 1× first strand buffer and 5 mM dithiothreitol (DTT) were added to each reaction. The reactions were incubated at 42 °C for 15 min, followed by 50 °C for 30 min, and stopped by 15 min of incubation at 70 °C. Subsequently, 1 unit of RNase H (Invitrogen) was added, and the samples were incubated at 37 °C for 35 min. Three cDNA synthesis reactions were performed: (1) the complete reaction, (2) the reaction in the absence of reverse transcriptase and (3) the reaction without RNA poly(A+). The above three samples were used for each PCR. We used a forward specific oligonucleotide for each Tv*tbp* gene (ESM 2 Table), and the universal AUAP-B primer (Invitrogen). The amplicons were cloned into the pGEM®-T Easy vector (Promega) and sequenced. To confirm the results, a subsequent nested PCR amplification was performed with specific primers.

Recombinant proteins

The complete TvTBP1, TvTBP2, TvCRK5 (Amador et al. 2017) and IBP39 (Liston et al. 2001) ORFs were amplified by PCR from *T. vaginalis* DNA with specific oligonucleotides (ESM 2 Table). TvTBPs and TvCRK5 were cloned into a 6xHis-tag pQE30 expression vector (QIAGEN). Expression of the recombinant proteins in *Escherichia coli* (M15 strain) was induced by 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37 °C. Proteins were purified under native conditions through affinity chromatography (QIAexpress® Ni-NTA Fast Start Kit, QIAGEN) following the manufacturer's instructions. IBP39 was cloned into pGEX4T1, transformed into *E. coli* (BL21 strain) and induced by 1 mM IPTG for 2.5 h at 37 °C. IBP39 was purified using Glutathione Sepharose™ beads (GE Healthcare). It is important to mention that TvTBP2 was always expressed in a much lower amount than TvTBP1 despite the use of different clones and was more difficult to purify.

Preparation of TvTBP polyclonal antiserum

An anti-TvTBP1 rabbit polyclonal antibody was produced by immunisation of a New Zealand eight-week-old rabbit with the purified 6xHis-TvTBP1 recombinant protein. Pre-immune serum was obtained 2 weeks prior to starting the immunisation scheme. Three immunisations of ~ 250 μg each of the recombinant protein were injected subcutaneously with 2-month intervals, the first with complete Freund's adjuvant and the subsequent ones with incomplete adjuvant. The collected blood was allowed to clot for 3 h, then was incubated overnight at 4 °C and centrifuged at 4000 rpm for 20 min at 4 °C. The serum aliquots were kept at − 70 °C until used. The immunisation procedure was carried out in strict accordance with the recommendations in the Official Mexican Norm: *Technical specifications for the production, care and use of laboratory animals* (NOM-062-ZOO-1999) and the protocol was approved by the Institutional Commission for the Care and Use of Laboratory Animals (CICUAL) of the Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México (Folio ID 6299). The rabbit in this study was housed in research animal facilities staffed with trained husbandry, technical and veterinary personnel.

Preparation of nuclear extracts

Nuclear extract preparations were obtained by a modification of previously described protocols (Liston and Johnson 1999; Ruvalcaba-Salazar et al. 2005). Briefly, approximately 1×10^6 *T. vaginalis* mid-log cells were collected by centrifugation, the pellet was washed twice with cold 1× phosphate-buffered saline (PBS) and the cells were resuspended in five volumes of resuspension buffer (10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM tosyl-L-lysyl-chloromethane hydrochloride (TLCK), 1 mM leupeptin, 1× complete Protease Inhibitor Cocktail (05892791001, Roche)). The sample was transferred to a glass Dounce homogenizer (Wheaton) and incubated for 30 min on ice. Cells were broken by 20 strokes with an A pestle (Tight). The cytoplasmatic fraction was collected by centrifugation at 14,000 rpm for 10 min at 4 °C, and the pellet was resuspended in extraction buffer (20 mM HEPES-KOH pH 7.9, 400 mM NaCl, 1 mM EDTA pH 8, 1 mM EGTA pH 8, 1 mM DTT, 1 mM PMSF, 1 mM TLCK, 1 mM leupeptin). The sample was vigorously mixed for 30 min in a cold room, and the extract was centrifuged at 14,000 rpm for 30 min at 4 °C. The supernatant, containing nuclear proteins, was collected and used immediately.

Western blot

Purified recombinant proteins 6xHis-TvTBP1 and 6xHis-TvTBP2 and the *T. vaginalis*-enriched nuclear extract preparations were used to test for dimerisation in western blot assays. SDS-PAGE (with or without β-mercaptoethanol (BME) in the sample buffer) was performed. The polyclonal TvTBP1 antiserum was diluted 1:10,000 to test the recombinant proteins and 1:500 to test the enriched nuclear extracts from *T. vaginalis*.

Yeast complementation analyses

Open reading frames from the TvTBP1 and ScTBP (*SPT15* yeast gene) were amplified by PCR from *T. vaginalis* and

S. cerevisiae genomic DNA. The PCR products were cloned into the pYES2 vector (Invitrogen), which contains the *GAL1* promoter, which is inducible by galactose, and a *URA3* selection marker. Additionally, site-directed mutagenesis was performed on TvTBP1-pYES2 to generate TvTBP1mut containing the amino acids involved in DNA binding in *S. cerevisiae* (Table 1). All oligonucleotides used are described in Table ESM 2. Plasmid constructs were transformed using the lithium acetate method in the *S. cerevisiae spt15*-knockout heterozygous diploid strain BY4743 (*MATa/α his3Δ1/his3Δ1*, *leu2Δ0/leu2Δ0*, *LYS2/lys2Δ0*, *met15Δ0/MET15*, *ura3Δ0/ura3Δ0*, *SPT15/spt15Δ::KANMX*) (Thermo Fisher Scientific cat. YSC1021-674345). After transformation, cells were plated, selected on synthetic dropout medium lacking uracil and containing 2% galactose and 1% raffinose and grown at 30 °C. The clones were then transferred to a pre-sporulation medium (0.8% yeast extract, 0.3% peptone, 10% anhydrous dextrose, 2% agar) and finally to a sporulation medium (1% potassium acetate, 0.02% raffinose, 5 mg/l histidine, 25 mg/l leucine, 2% agar). Complementation was determined by tetrad dissection and spore viability on YPGal medium (1% yeast extract, 2% peptone, 2% galactose). An empty pYES2 vector was used as a negative control. Additionally, to confirm that the germinated cells contain the allele interrupted by a *KANMX* module (complemented cells), the colonies were patched on both YPGal and G418-containing YPGal plates.

The expression of the recombinant TvTBP1 and TvTBP1mut in the transformed yeast cultures was demonstrated by western blot using the anti-TvTBP1 antibody (Dunn and Wobbe 1993). Yeast *spt15Δ* mutant diploid cells were transformed with the pYES2 constructs and grown in induction medium (synthetic dropout medium lacking uracil and containing 2% galactose, 1% raffinose). Cultures were grown to mid-log and cells were collected by centrifugation. The pellet was washed twice with ice-cold water. Cells were resuspended in three volumes of glass bead disruption buffer (20 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM EDTA, 5% glycerol, 1 mM DTT, 300 mM ammonium sulfate, 1× complete Protease Inhibitor Cocktail (05892791001, Roche), 1 mM PMSF) and 4 volumes of acid-washed glass beads was added. Cells were lysed by six cycles of vortex at maximum speed for 1 min and incubated on ice for 1 min. Protein lysates were centrifuged at 12,000×*g* for 60 min at 4 °C and the supernatant was collected. The protein extracts were loaded in a 12% SDS-PAGE and analysed by western blot with anti-TvTBP1 antiserum (1:1000). Expression of TvTBP2 in the transformed yeast cultures could not be corroborated with the anti-TvTBP1 antiserum which mainly recognises TvTBP1 (see Fig. 3b).

Pull-down assays

Transformed *E. coli* cells with the recombinant glutathione S-transferase GST-IBP39 fusion protein, or GST alone as a control, were induced as described in the 'Recombinant proteins' section. Fifty millilitre of induced *E. coli* cultures expressing GST-IBP39, GST or 6xHis-TvTBP1 was lysed in 2 ml of resuspension buffer (16 mM Tris-HCl pH 8, 0.8 mM EDTA pH 8, 1 mg/ml lysozyme, 100 U Benzonase™ nuclease), and incubated 30 min on ice. The lysates were centrifuged at 10,000×*g* for 15 min at 4 °C. The GST-IBP39 or GST supernatants were incubated with Glutathione Sepharose™ beads (GE Healthcare) (135 μl of 75% slurry previously equilibrated with 1× PBS) for 30 min at room temperature with gentle shaking, followed by 3 washes (500 μl each) with 1× PBS. The beads were then incubated with the 6xHis-TvTBP1 lysates (2 ml) at room temperature for 30 min and then washed four times in 1× PBS before elution with

Table 1 Summary of the predicted relevant amino acid residues involved in TBP binding to DNA

ScTBP	N69	V71	F99	F116	V122	T124	N159	V161	F190	V203	L205	F207	L214	T215
HsTBP	•	•	•	•	•	•	•	•	•	•	•	•	•	•
EhTBP	•	•	•	•	•	•	•	•	•	T	•	•	•	•
EhTRF1	•	•	•	•	•	•	•	•	•	•	•	•	F	•
TvTBP1	•^{56}	•^{58}	•^{86}	•^{103}	N^{109}	V^{111}	•^{147}	•^{149}	Y^{180}	T^{193}	M^{195}	Y^{197}	I^{204}	•^{205}
TvTBP2	•^{53}	I^{55}	I^{84}	•^{101}	N^{107}	V^{109}	S^{144}	•^{146}	•^{177}	T^{190}	•^{192}	•^{194}	•^{201}	•^{202}
TvTBP1mut	•	•	•	•	•	•	•	•	•	•	•	•	I	•
TcTBP	G	I	T	R	S	I	S	T	•	S	T	•	I	L
TbTRF4	G	I	C	R	G	I	S	M	•	C	T	•	V	L
LmTBP	A	Q	I	H	S	I	S	A	•	S	S	•	•	M
GlTBP	G	N	Y	S	T	R	S	T	M	C	S	•	I	V

Dots indicate amino acid identity as compared with ScTBP. Shadowed boxes show amino acid conservation, with darker colours representing higher conservation. The numbers in the first row correspond to the ScTBP sequence. TvTBPs amino acid numbers are denoted in superscript. TvTBP1mut was generated through site-directed mutagenesis using TvTBP1 as the template. The following changes were made: N109V, V111T, Y180F, T193V, M195L, Y197F

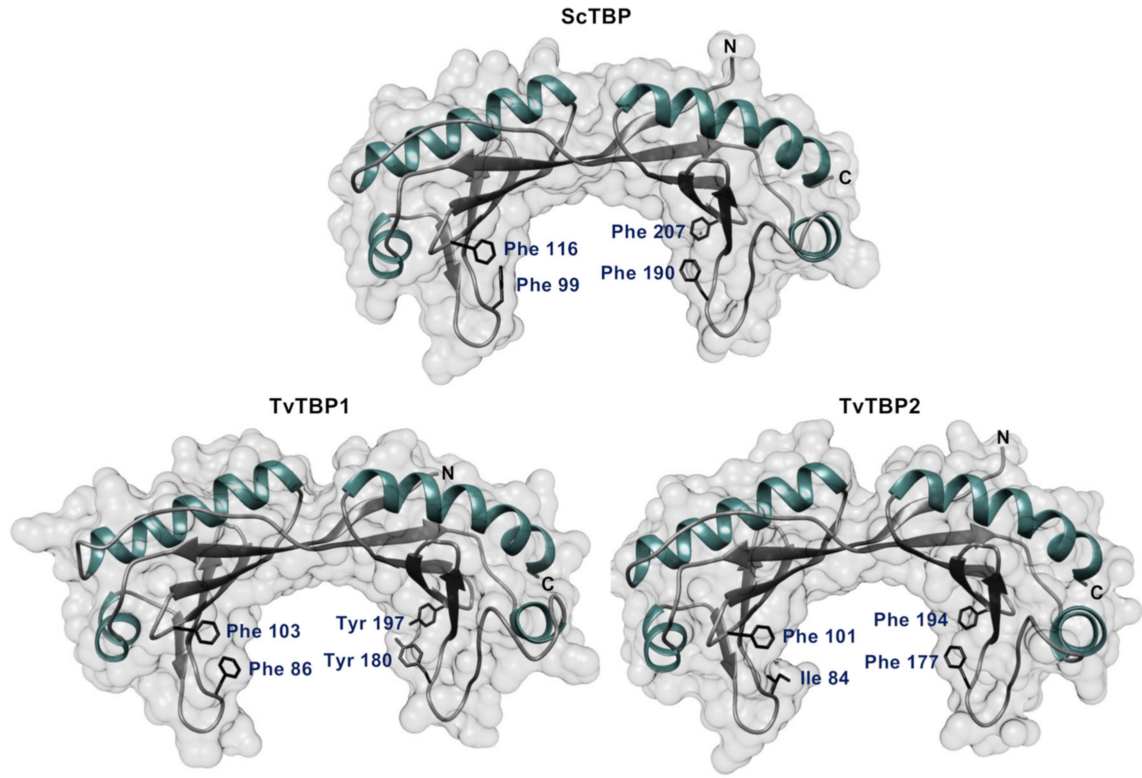


Fig. 1 Tertiary structure models for *T. vaginalis* TBPs. Theoretical structures of TvTBP1 and TvTBP2 adopt the typical saddle-shaped structure shown by a canonical TBP from *Saccharomyces cerevisiae* (ScTBP). Alpha helices are shown in green, and beta strands are shown in dark grey. Important amino acids for DNA binding are indicated and numbered according to the sequences shown in Fig. 2. Models were obtained with the Modeller web server and visualised with Chimera as described in the 'Materials and methods' section

10 mM reduced glutathione. The eluted proteins were separated by 12% SDS-PAGE and analysed by western blotting with HisProbe™-HRP (Thermo Scientific) to identify the interactions.

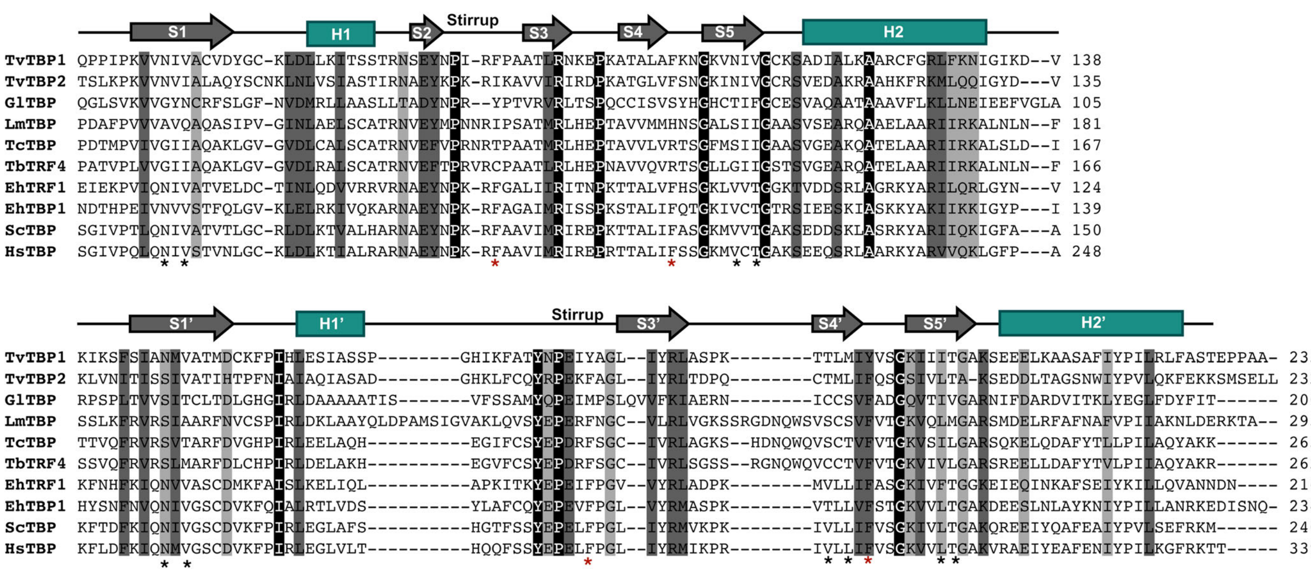


Fig. 2 Alignment of TBP amino acid sequences. Carboxi-terminal sequences of TBPs from *T. vaginalis* (TvTBP1 (A2DQT7) and TvTBP2 (A2G4N7)), *Giardia lamblia* (GlTBP (Q86GG9)), *Leishmania major* (LmTBP (Q4QD71)), *Trypanosoma cruzi* (TcTBP (Q4DK05)), *Trypanosoma brucei* (TbTRF4 (Q387R9)), *Entamoeba histolytica* (EhTBP1 (P52653) and EhTRF1 (A7UFC2)), *Saccharomyces cerevisiae* (ScTBP (P13393)) and *Homo sapiens* (HsTBP (P20226)) were aligned using the Clustal Omega program. Accession numbers are from the UniProt database. The secondary structure was determined from the TvTBP1 amino acid sequence (H, α helix; S, β strand). Phenylalanine residues (red) and other residues (black) involved in DNA binding are marked with asterisks

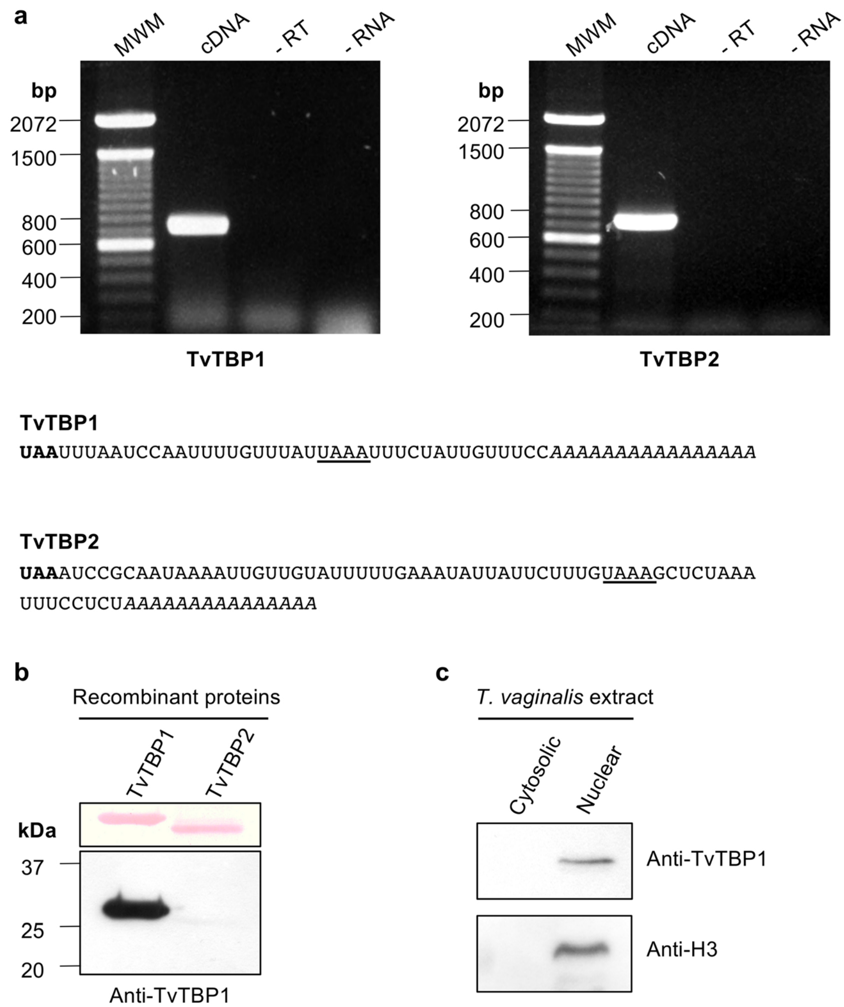


Fig. 3 **a** Both Tv*tbp* genes are expressed in *T. vaginalis*. Agarose gels stained with ethidium bromide show RT-PCRs amplified with specific primers for Tv*tbp1* and Tv*tbp2*; the 3′-UTR cDNA sequences are shown underneath. MWM, molecular weight markers; cDNA, the complete reaction; −RT, reaction without reverse transcriptase; −RNA, reaction without RNA. The UAA stop codons are shown in bold. The *T. vaginalis* UAAA polyadenylation signal (Fuentes et al. 2012) is underlined, and the poly(A) tail is denoted in italics. **b** The TvTBP1 recombinant protein is preferentially recognised by the anti-TvTBP1 antiserum as compared with the TvTBP2 recombinant protein. **c** The TvTBP1 protein is detected in the *T. vaginalis* nuclear protein fraction. An anti-histone H3 antibody was used as a nuclear marker

Electrophoretic mobility shift assays

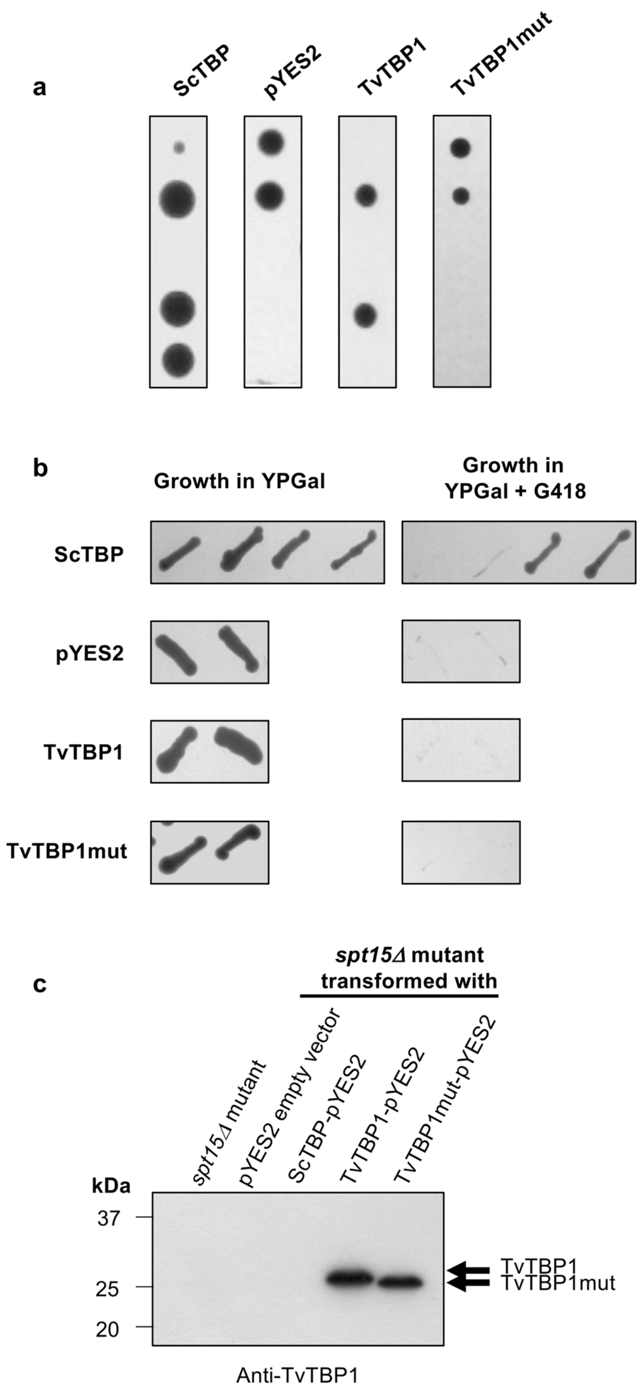
EMSAs were performed using purified recombinant 6xHis-TvTBP1, 6xHis-TvTBP2 and 6xHis-TvCRK5 (as negative control) proteins. The DNA probes were obtained by PCR amplification using specific primers and *T. vaginalis* genomic DNA as the template (ESM2 Table). The probes were purified by Sephadex G50 and quantified by a NanoDrop™ instrument. Briefly, 400–500 ng of purified protein was used for each assay. The recombinant proteins were incubated with each probe (200 ng) for 30 min at room temperature in binding buffer (10 mM HEPES pH 7.9, 4 mM $MgCl_2$, 5 mM ammonium sulfate, 8% (v/v) glycerol, 2% (w/v) PEG, 55 mM KCl, 5 mM β-mercaptoethanol, 0.2 mM EDTA and 0.2 mM PMSF) (Glossop 2004). Then, 100 ng of bovine serum albumin (BSA) was added to the reaction before loading. Reactions were loaded in a non-denaturing polyacrylamide gel (4.5% acrylamide-bisacrylamide 29:1, 0.5% TBE and 2.5% (v/v) glycerol) and electrophoresed in 0.5× TBE as the running buffer. The gels had previously been pre-run at 100 V for 30 min before loading. Gels were stained with ethidium bromide and visualised by UV. For clarity, the images in the results section are shown inverted.

Fig. 4 Yeast complementation assay for a *S. cerevisiae* TBP mutant. **a** Tetrad dissections show that only two spores grew in YPGal medium when the yeast had been transformed with plasmids coding for TvTBP1 and TvTBP1mut, while all four spores grew when the yeast had been complemented with the *S. cerevisiae* TBP (ScTBP); pYES2 empty vector was used as a negative control. A single dissection is shown as an example for each assay. **b** All viable yeast haploids grew in YPGal medium. When G418 antibiotic was added to this medium, only two spores from the ScTBP control grew, as expected for the segregation of the *KANMX* allele. No growth was observed in the TvTBP1 or TvTBP1mut transformants in the presence of G418 antibiotic. **c** Yeast *spt15Δ* mutant strain was transformed with the indicated pYES2 constructs and propagated in induction medium. Protein extracts were obtained and electrophoresed in a 12% SDS-PAGE for western blotting with anti-TvTBP1 antiserum. The expression of both TvTBP1 and TvTBP1mut proteins was clearly identified in the transformed cultures. The yeast ScTBP is not detected by the anti-TvTBP1 antiserum

Results and discussion

Two genes in *T. vaginalis* encode protein homologues of TBP, as deduced by an in silico analysis

Two genes are annotated in the *T. vaginalis* genome as putative transcription initiation factor TFIID genes: TVAG_291560 and TVAG_285070 (Carlton et al. 2007); their deduced protein sequences have similar characteristics to canonical TBPs in that they have the specific carboxy terminal domain of the TBP superfamily (Pfam: PF00352) (Finn et al. 2014), and the models of their tertiary structures clearly adopt the distinctive saddle-shaped structure (Fig. 1). These two genes are therefore herein referred to as Tv*tbp1* and Tv*tbp2*, respectively.



We compared the deduced C-terminal amino acid sequences of the TvTBPs with those of TBPs from *Giardia lamblia*, *Entamoeba histolytica*, trypanosomatid parasites, yeast and human (Fig. 2). TvTBP1 and TvTBP2 share 45.16% identity between themselves and 22.35–46.67% identity with the *G. lamblia*, *Leishmania major*, *Trypanosoma cruzi*, *Trypanosoma brucei*, *E. histolytica*, *Homo sapiens* and *S. cerevisiae* TBPs (see table in ESM 1). TvTBPs have the highest identity with ScTBP. In addition, we carried out a detailed analysis among the TvTBPs and the abovementioned proteins in order to compare the amino acid residues involved in the recognition of and interactions with the TATA box (Juo et al. 1996; Millán-Pacheco et al. 2009) (Table 1). TvTBP1 has the two asparagine residues (N56, N147), located in the central region, that are important for TATA box recognition. Four phenylalanine residues are also relevant for DNA interaction, and in TvTBP1, two of these amino acids are tyrosine (Y180, Y197). TvTBP2 possesses only one of the two conserved asparagine residues (N53); the second one is replaced by serine (S144), as in trypanosomatid and *G. lamblia* TBPs. Three of the four phenylalanine residues (F101, F177, F194) in TvTBP2 are conserved, the other one corresponds to isoleucine (I84), as in the *Leishmania* proteins. Both TvTBPs possess additional amino acid substitutions (see Table 1).

The alignment of the TvTBPs with the TBP superfamily, their putative saddle-shaped structure and the overall amino acid conservation show that these proteins have characteristics similar to those of a canonical TBP, as has already been mentioned by other authors (Castañon-Sanchez et al. 2010; Smith and Johnson 2011; Koster et al. 2015). Among these TBPs, the orthologues from yeast, human and *E. histolytica* bind TATA box promoters (Kim and Burley 1994; Nikolov et al. 1996; Castañon-Sanchez et al. 2010). In promoters from trypanosomes, no TATA elements have been identified (Gomez et al. 2010), and there is no experimental evidence of a direct interaction between trypanosome TBPs and TATA boxes. Likewise, a classic TATA box has not been described for *T. vaginalis* RNA pol II promoters (Liston and Johnson 1999; Smith and Johnson 2011). Therefore, a question posed here is whether the TvTBPs are expressed *in T. vaginalis* and can bind to a particular DNA sequence.

Both TvTBP1 and TvTBP2 genes are expressed in *T. vaginalis*

As described above, two genes for TvTBPs are annotated in the TrichDB database (https://trichdb.org/trichdb/) but only ESTs for TvTBP2 (GT107997, TvI107B09) have been reported. To determine whether both Tv*tbp* genes are expressed, we used a reverse transcription polymerase chain reaction assay with oligonucleotides specific for TvTBP1 and TvTBP2 and an AUAP-B primer, as described in the 'Materials and methods' section. A nested PCR amplification was subsequently performed. Amplicons were cloned and automatically sequenced (Fig. 3a). The obtained cDNA sequences demonstrate that both Tv*tbp1* and Tv*tbp2* genes are expressed, and that transcripts are polyadenylated at the expected site downstream to the previously described *T. vaginalis* polyadenylation signal (Fuentes et al. 2012). Additionally, a nuclear protein enrichment suggests that TvTBP1, which is better recognised by the anti-TvTBP1 antibody (Fig. 3b), is expressed in the *T. vaginalis* nucleus (Fig. 3c).

TvTBP1 does not complement a *S. cerevisiae* strain lacking TBP

To determine whether TvTBP1 has a similar basic function as a canonical TBP, we performed yeast complementation assays in a *S. cerevisiae* diploid strain lacking TBP (Fig. 4). In this strain, one of the two alleles for the *SPT15* gene (which encodes ScTBP) was replaced with the *KANMX* cassette, which confers resistance to the G418 antibiotic. This strain was transformed with plasmids coding for TvTBP1, TvTBP1mut or ScTBP (as a positive control). TvTBP1mut was constructed by site-directed mutagenesis on TvTBP1-pYES2 to generate a TvTBP1 version containing the amino acids involved in DNA binding in ScTBP (Table 1). Complementation was analysed by tetrad dissection as described in the 'Materials and methods' section. Only two spores grew in YPGal medium when the *S. cerevisiae* mutant strain was transformed with plasmids coding for TvTBP1, TvTBP1mut, or with the empty vector, while all four spores grew when complemented with ScTBP. Figure 4 a shows one tetrad dissection of each as an example. Transformants that complement the lack of ScTBP are viable haploid cells resistant to the G418 antibiotic. None of the haploid cells transformed with TvTBP1 or TvTBP1mut grew in YPGal medium with the G418 antibiotic (Fig. 4b). Expression of the *T. vaginalis* TvTBP1 and TvTBP1mut recombinant proteins in the *S. cerevisiae* mutant was corroborated by western blot with the anti-TvTBP1 antibody (Fig. 4c). These results show that TvTBP1 and TvTBP1mut cannot replace (as a viable phenotype) ScTBP in the *S. cerevisiae* mutant. With this approach, we cannot determine if TvTBPs bind to promoters of *S. cerevisiae* genes. If they bind, they are apparently unable to recruit the PIC to initiate transcription in the yeast. Hence, we focused next on determining the DNA binding specificity of the TvTBPs in vitro.

Recombinant TvTBP proteins form dimers

The HisProbe-HRP recognised both TvTBP1 and TvTBP2 recombinant proteins (approximately 27 kDa) in a western blot assay (Fig. 5a). Interestingly, in addition to the monomeric form (+BME lanes), we observed a second signal of approximately 54 kDa that might correspond to recombinant TvTBP dimers in lanes lacking BME (Fig. 5a). No signal was detected in the protein extract from M15 *E. coli* transformed with the pQE30 vector. The anti-TvTBP1 antiserum was tested against enriched nuclear extracts from *T. vaginalis*

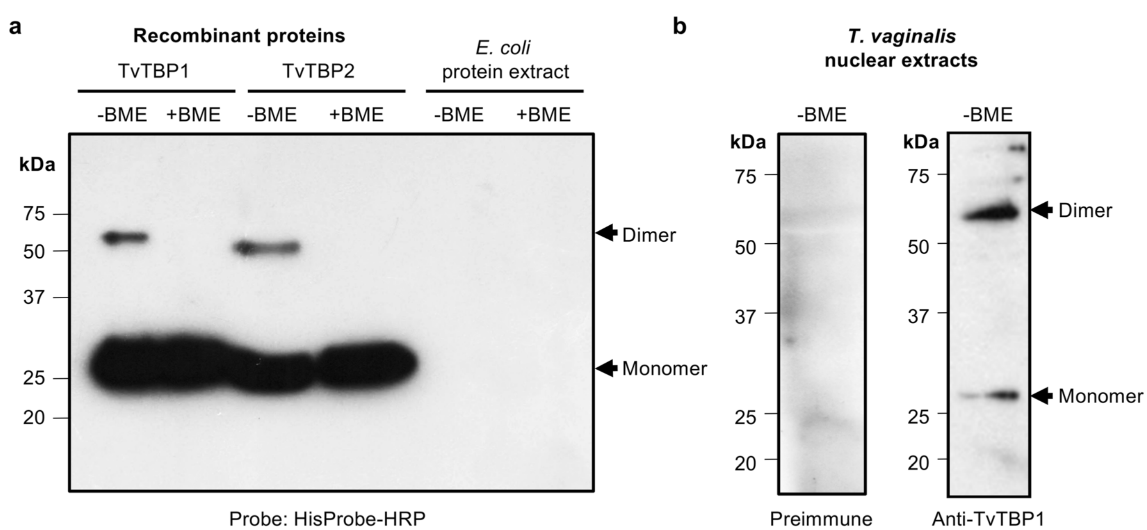


Fig. 5 TvTBPs form dimers both in vitro and in vivo. **a** The HisProbe-HRP recognises both TvTBP1 and TvTBP2 6xHis-tagged recombinant proteins. Two signals were present, and they could represent the monomer and the dimer of the recombinant TvTBPs, with an approximately weight of 27 kDa and 54 kDa, respectively. −BME, sample without β-mercaptoethanol in the sample buffer; +BME, sample with β-mercaptoethanol in the sample buffer. **b** Enriched nuclear extracts from *T. vaginalis* probed with the anti-TvTBP1 antiserum. In the absence of β-mercaptoethanol, two signals were detected representing an approximately weight of 26 kDa and 52 kDa, which could represent the monomer and dimer of the native TvTBP proteins in *T. vaginalis*

(Fig. 5b), and two predominant bands were again observed: one that could represent a TvTBP monomer (26 kDa), and one that could represent the dimer (52 kDa). In this case, the dimeric form was more abundant than the monomeric form. This experiment demonstrated the presence of a TvTBP among *T. vaginalis* proteins.

The presence of dimers of human TBPs in cellular extracts has already been reported in the absence of DNA, and these dimers must dissociate into monomers before binding to the TATA box (Coleman et al. 1995; Coleman and Pugh 1997). It has been proposed that the formation of dimers may play a regulatory role in transcription, preventing unregulated expression of genes transcribed by RNA pol II (Kou et al. 2003; Kou and Pugh 2004; Thomas and Chiang 2006).

TvTBP1 interacts with IBP39

To date, IBP39 is the only protein that has been described in *T. vaginalis* involved in the transcription initiation process, and it is responsible for binding the Inr region of the *T. vaginalis* pol II promoter (Liston et al. 2001). To test the ability of TvTBP1 to interact with IBP39 as part of the transcription initiation complex, GST pull-down assays were performed using the recombinant GST-IBP39 bound to the GST matrix and the recombinant 6xHis-TvTBP1 (Fig. 6 panel a). A band with the molecular weight of TvTBP1 was observed in the beads (matrix) and in the elution, evidencing an interaction between IBP39 and TvTBP1. Two negative controls were used: a non-related protein (TvCRK5) and the recombinant GST protein. As expected, no interaction was observed between IBP39 and TvCRK5 (panel b) or between TvTBP1 and the GST control (panel c), ruling out non-specific interactions among the recombinant proteins. The observed interaction between IBP39 and TvTBP1 suggests that this TBP might be part of the transcription initiation complex for transcription by RNA pol II in *T. vaginalis*. Additional transcription factors, yet to be defined, might also participate in this fundamental process.

TvTBP1 and TvTBP2 bind to promoter regions of *T. vaginalis* transcribed by the three RNA polymerases

To test the DNA-binding ability of TvTBP1 and TvTBP2, we performed EMSAs probing five different promoter sequences from *T. vaginalis* that are known to be recognised by the three RNA polymerases: the first was the ribosomal RNA gene promoter recognised by RNA polymerase I (Franco et al. 2012); the second and third were the β subunit of succinyl CoA synthetase gene promoter (Lahti et al. 1992) and an actin gene promoter (Espinosa et al. 2001), both recognised by RNA polymerase II; and the fourth and fifth were two DNA regions involved in transcription by RNA polymerase III: the 5S rRNA (Torres-Machorro et al. 2006) and the U6 small nuclear RNA (snRNA) (Simoes-Barbosa et al. 2008) (Fig. 7a). As shown in Fig. 7b, TvTBP1 and TvTBP2 bound all the tested probes. This result is not completely unexpected because a putative TATA box (−28 nt) has been identified in the U6 snRNA (Simoes-Barbosa et al. 2008), and all the other probes contain regions rich in adenine and thymine. On the other hand, an apparent difference between TvTBP1 and TvTBP2 was observed regarding their binding ability, which may be related to the affinity and stability with which the DNA-TvTBP complexes are formed.

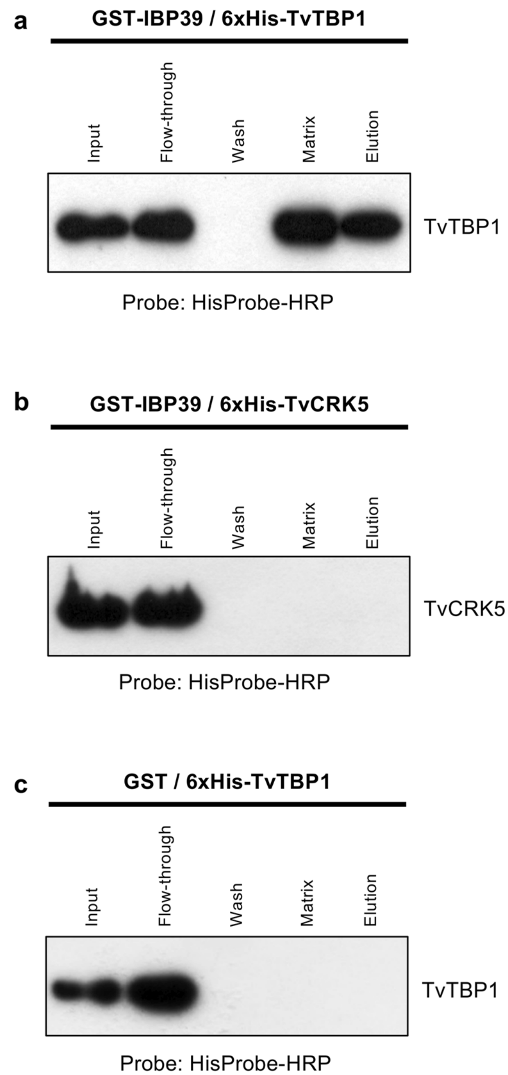


Fig. 6 TvTBP1 interacts with IBP39. Western blot detection with HisProbe-HRP of pull-down assays of GST-IBP39 with 6xHis-TvTBP1 (panel **a**), 6xHis-TvCRK5 (panel **b**) and GST with 6xHis-TvTBP1 (panel **c**). Input: Cell lysates from *E. coli* expressing 6xHis-TvTBP1 (panels **a** and **c**) or 6xHis-TvCRK5 (panel **b**). Flow-through: unbound protein. Wash: Last wash of the beads before elution of the complexes. Matrix: proteins attached to Glutathione Sepharose beads. Elution: proteins eluted with reduced glutathione

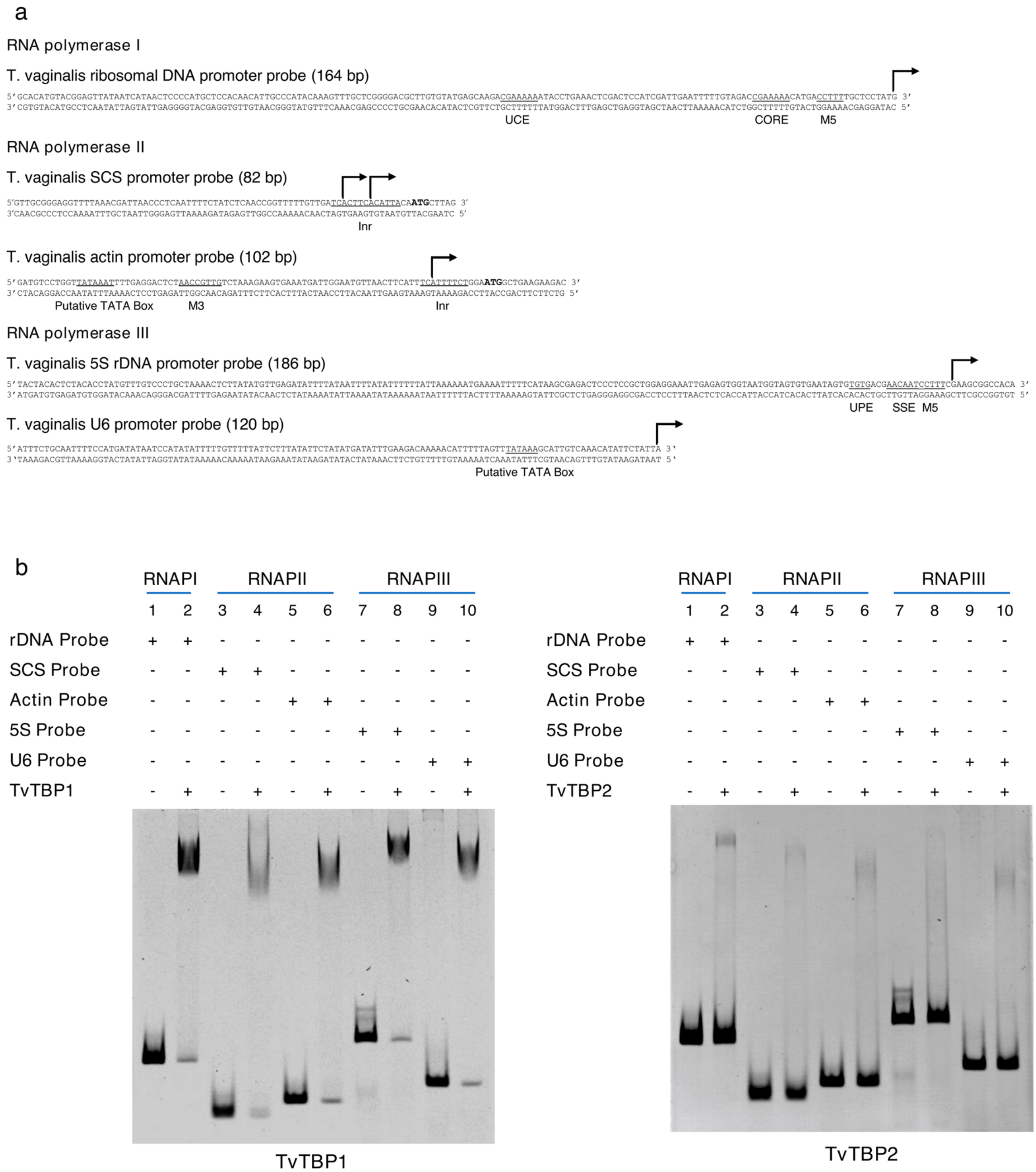


Fig. 7 TvTBPs bind different *T. vaginalis* promoter sequences. **a** Sequences of the tested promoter regions from *T. vaginalis*: (1) the ribosomal RNA gene promoter (rDNA), transcribed by RNA pol I. (2) Promoter region from the β subunit of succinyl CoA synthetase (β-SCS) and (3) actin promoter region from TVAG_172680, both transcribed by RNA pol II. (4) 5S ribosomal RNA gene promoter (5S rDNA) and (5) U6 small nuclear RNA gene promoter (U6 snRNA), both transcribed by RNA pol III. In bold: the ATG codon; underlined: Upstream Control Element (UCE), core element, putative TATA box, Inr element, Upstream Promoter Element (UPE), Start Site Element (SSE), motif 3 (M3) and motif 5 (M5). The +1 transcription start site is shown by an arrow. **b** Electrophoretic mobility shift assay (EMSA) for the TvTBP1 and TvTBP2 recombinant proteins interacting with probes described in panel (**a**). The gels were visualised by staining with ethidium bromide and photographed under UV light. The images are shown inverted

TvTBP1 and TvTBP2 bind DNA in a non-specific manner

To define a possible binding motif for the TvTBPs, we used the actin promoter probe (102 bp) as a model due to the presence of the TATAAATT sequence from position − 66 to − 59. This motif corresponds to a consensus TATA box (TATAWAWN). Smaller regions of this probe were obtained by PCR amplification and the smaller amplicons were used to test their interactions with the TvTBPs (Fig. 8a). The consensus TATA box sequence is present only in probes A and B, while probes C, D and E do not possess this element. The Inr

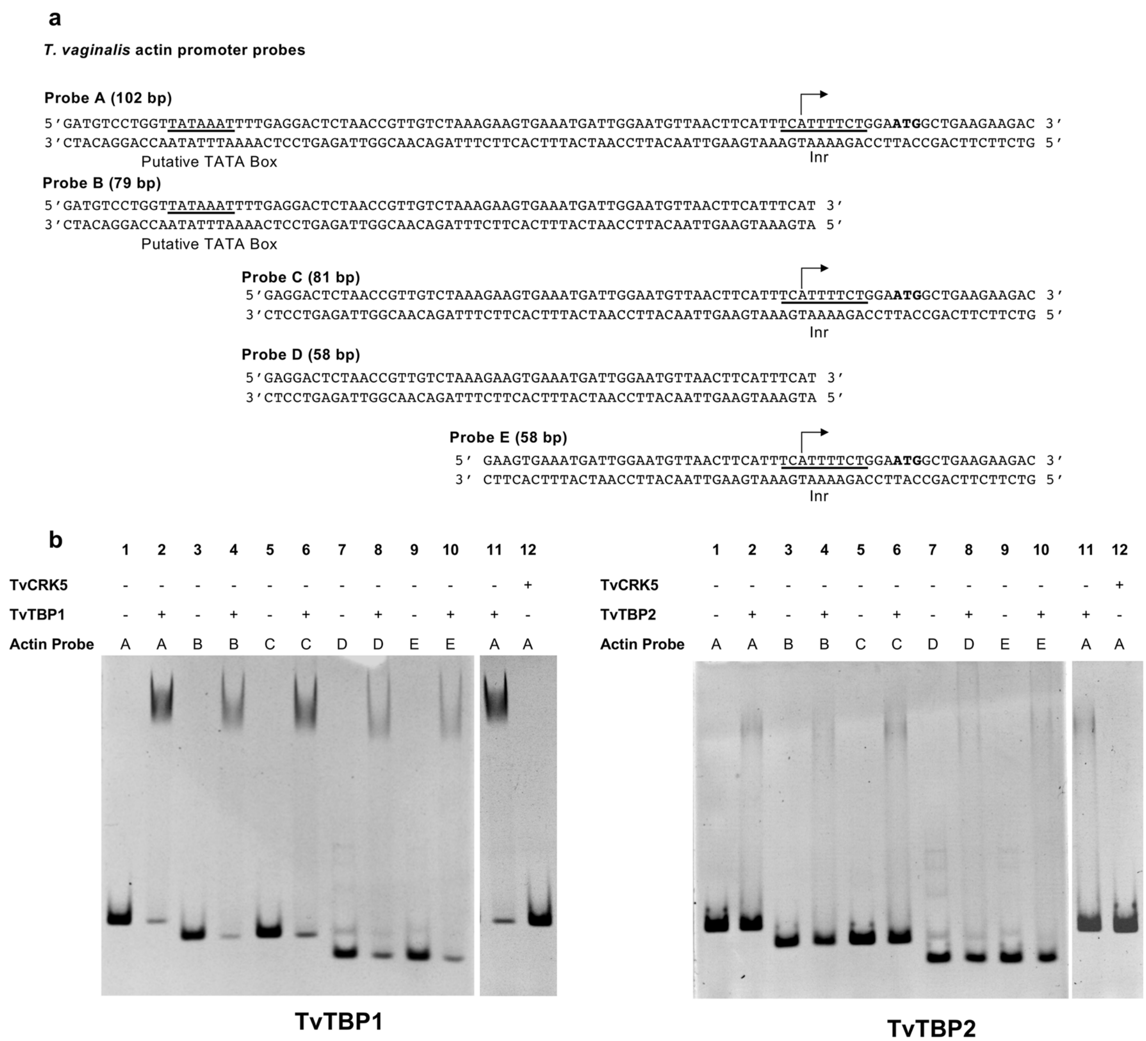


Fig. 8 TvTBPs bind DNA in a non-specific manner. **a** Different sections of the *T. vaginalis* actin gene promoter used as probes in EMSAs with the TvTBPs. **b** 500 ng of TvTBP1 or 400 ng of TvTBP2 recombinant protein was incubated with the indicated probes, as described in the 'Materials and methods' section. The gels were visualised by staining with ethidium bromide and photographed under UV light. The images are shown inverted

element is absent from probes B and D. A shift was observed in all cases with both TvTBPs (Fig. 8b, lanes 1–10). A non-related protein (TvCRK5) was used as a negative control for DNA binding, and no shift was observed (Fig. 8b, lane 12). Additionally, a pUC19 multiple cloning site DNA fragment was used as a non-related DNA sequence control in an EMSA, in which a shift was also observed with both TvTBPs (data not shown). These results show that both TvTBPs bind to double-stranded DNA in a non-specific manner and independently of the presence of a TATA box.

Altogether, the EMSAs suggest that TvTBP1 and TvTBP2 bind to DNA in a sequence-independent manner. Nevertheless, we cannot exclude that the TvTBPs might bind to *T. vaginalis* promoter regions through non-canonical processes as described in other organisms. TBPs in the protist *E. histolytica* recognise TATA box variants (De Dios-Bravo et al. 2005), while *Crypthecodinium cohnii*, a marine unicellular organism, expresses an atypical TBP with properties between TBP and TRF which recognises TTTT box sequences (Guillebault et al. 2002).

Conclusions

The data presented here show that *T. vaginalis*, an early divergent unicellular organism, expresses two proteins similar to TBP that bind differently to DNA and in a non-specific manner. TvTBP1 recombinant protein interacts with the recombinant IBP39 *T. vaginalis* specific transcription factor, which is responsible for Inr binding in this organism. The apparent absence of classical TATA boxes in *T. vaginalis*, together with the ability of TvTBP1 to bind IBP39 and the ability of TvTBP1 and TvTBP2 to interact with DNA in a non-specific manner, suggests that TvTBPs might be part of a transcription initiation complex which components remain to be defined. Some basal transcription factors such as TFIIB, BRF1 and TBP-associated factors (TAFs) are annotated in the *T. vaginalis* genome (Carlton et al. 2007), but their participation in the transcription process remains to be determined.

Acknowledgements We thank Dr. Roberto Coria Ortega for support with the yeast complementation system.

Funding information This work was financially supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT-79983); Dirección General de Asuntos del Personal Académico (DGAPA-PAPIIT-IN205813). Olivia Parra-Marín is a doctoral student from Programa de Maestría y Doctorado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México (UNAM). Lluvia Rosas-Hernández is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM). OPM and LRH were supported by scholarships from CONACyT, Mexico.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Abbreviations TBP, TATA box-binding protein; TvTBP, TBP from *T. vaginalis*; IBP39, Initiator-binding protein 39; EMSA, Electrophoretic mobility shift assay

References

Akhtar W, Veenstra GJC (2011) TBP-related factors: a paradigm of diversity in transcription initiation. Cell Biosci 1:23. https://doi.org/10.1186/2045-3701-1-23

Alvarez-Sánchez ME, Avila-González L, Becerril-García C et al (2000) A novel cysteine proteinase (CP65) of *Trichomonas vaginalis* involved in cytotoxicity. Microb Pathog 28:193–202. https://doi.org/10.1006/mpat.1999.0336

Amador E, López-Pacheco K, Morales N et al (2017) Characterization of cyclin-dependent kinases and Cdc2/Cdc28 kinase subunits in *Trichomonas vaginalis*. Parasitology 144:571–582. https://doi.org/10.1017/S0031182016002195

Bowie JU, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science (80-) 253:164–170. https://doi.org/10.1126/science.1853201

Carlton JM, Hirt RP, Silva JC et al (2007) Draft genome sequence of the. Science (80-) 315:207–212. https://doi.org/10.1126/science.1132894

Castañon-Sanchez CA, Luna-Arias JP, de Dios-Bravo MG, Herrera-Aguirre ME, Olivares-Trejo JJ, Orozco E, Hernandez JM (2010) *Entamoeba histolytica*: a unicellular organism containing two active genes encoding for members of the TBP family. Protein Expr Purif 70:48–59. https://doi.org/10.1016/j.pep.2009.12.007

Clark CG, Diamond LS (2002) Methods for cultivation of luminal parasitic protists of clinical importance. Clin Microbiol Rev 15:329–341. https://doi.org/10.1128/CMR.15.3.329-341.2002

Coleman RA, Pugh BF (1997) Slow dimer dissociation of the TATA binding protein dictates the kinetics of DNA binding. Proc Natl Acad Sci 94:7221–7226. https://doi.org/10.1073/pnas.94.14.7221

Coleman RA, Taggart AKP, Benjamin LR, Pugh BF (1995) Dimerization of the TATA binding protein. J Biol Chem 270:13842–13849. https://doi.org/10.1074/jbc.270.23.13842

Danino YM, Even D, Ideses D, Juven-Gershon T (2015) The core promoter: at the heart of gene expression. Biochim Biophys Acta 1849:1116–1131. https://doi.org/10.1016/j.bbagrm.2015.04.003

Davidson I (2003) The genetics of TBP and TBP-related factors. Trends Biochem Sci 28:391–398. https://doi.org/10.1016/S0968-0004(03)00117-8

De Dios-Bravo G, Luna-Arias JP, Riverón AM et al (2005) *Entamoeba histolytica* TATA-box binding protein binds to different TATA variants in vitro. FEBS J 272:1354–1366. https://doi.org/10.1111/j.1742-4658.2005.04566.x

Diamond LS (1957) The establishment of various trichomonads of animals and man in axenic cultures. J Parasitol 43:488–490. https://doi.org/10.2307/3274682

Dunn B, Wobbe CR (1993) Preparation of protein extracts from yeast. Curr Protoc Mol Biol 23:13.13.1–13.13.9. https://doi.org/10.1002/0471142727.mb1313s23

Espinosa N, Hernández R, López-Griego L, Arroyo R, López-Villaseñor I (2001) Differences between coding and non-coding regions in the *Trichomonas vaginalis* genome: an actin gene as a locus model. Acta Trop 78:147–154. https://doi.org/10.1016/S0001-706X(00)00180-7

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230. https://doi.org/10.1093/nar/gkt1223

Franco B, Hernández R, López-Villaseñor I (2012) *Trichomonas vaginalis* ribosomal RNA: identification and characterisation of the transcription promoter and terminator sequences. Mol Biochem Parasitol 185:1–9. https://doi.org/10.1016/j.molbiopara.2012.05.004

Fuentes V, Barrera G, Sánchez J, Hernández R, López-Villaseñor I (2012) Functional analysis of sequence motifs involved in the polyadenylation of *Trichomonas vaginalis* mRNAs. Eukaryot Cell 11:725–734. https://doi.org/10.1128/EC.05322-11

Glossop JA (2004) A conformational change in TFIIB is required for activator-mediated assembly of the preinitiation complex. Nucleic Acids Res 32:1829–1835. https://doi.org/10.1093/nar/gkh504

Gomez C, Rodríguez MA, Esther Ramirez M et al (2010) Regulation of gene expression in protozoa parasites. J Biomed Biotechnol:2010. https://doi.org/10.1155/2010/726045

Guillebault D, Sasorith S, Derelle E, Wurtz JM, Lozano JC, Bingham S, Tora L, Moreau H (2002) A new class of transcription initiation factors, intermediate between TATA box-binding proteins (TBPs) and TBP-like factors (TLFs), is present in the marine unicellular organism, the dinoflagellate *Crypthecodinium cohnii*. J Biol Chem 277:40881–40886. https://doi.org/10.1074/jbc.M205624200

Gunderson J, Hinkle G, Leipe D et al (1995) Phylogeny of trichomonads inferred from small-subunit rRNA sequences. J Eukaryot Microbiol 42:411–415. https://doi.org/10.1111/j.1550-7408.1995.tb01604.x

Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol 11:394–403. https://doi.org/10.1038/nsmb763

Hernandez N (1993) TBP, a universal eukaryotic transcription factor? Genes Dev 7:1291–1308. https://doi.org/10.1101/gad.7.7b.1291

Juo ZS, Chiu TK, Leiberman PM, Baikalov I, Berk AJ, Dickerson RE (1996) How proteins recognize the TATA box. J Mol Biol 261:239–254. https://doi.org/10.1006/jmbi.1996.0456

Kim JL, Burley SK (1994) 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. Nat Struct Biol 1:638–653. https://doi.org/10.1038/nsb0994-638

Koster MJE, Snel B, Timmers HTM (2015) Genesis of chromatin and transcription dynamics in the origin of species. Cell 161:724–736. https://doi.org/10.1016/j.cell.2015.04.033

Kou H, Pugh BF (2004) Engineering dimer-stabilizing mutations in the TATA-binding protein. J Biol Chem 279:20966–20973. https://doi.org/10.1074/jbc.M401535200

Kou H, Irvin JD, Huisinga KL, Mitra M, Pugh BF (2003) Structural and functional analysis of mutations along the crystallographic dimer interface of the yeast TATA binding protein. Mol Cell Biol 23:3186–3201. https://doi.org/10.1128/MCB.23.9.3186-3201.2003

Lahti CJ, D'Oliveira CE, Johnson PJ (1992) Beta-succinyl-coenzyme A synthetase from *Trichomonas vaginalis* is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. J Bacteriol 174:6822–6830. https://doi.org/10.1128/jb.174.21.6822-6830.1992

Lau AO, Liston DR, Vanacova S, Johnson PJ (2003) *Trichomonas vaginalis* initiator binding protein, IBP39, contains a novel DNA binding motif. Mol Biochem Parasitol 130:167–171. https://doi.org/10.1016/S0166-6851(03)00172-5

Lau AOT, Smith AJ, Brown MT, Johnson PJ (2006) *Trichomonas vaginalis* initiator binding protein (IBP39) and RNA polymerase II large subunit carboxy terminal domain interaction. Mol Biochem Parasitol 150:56–62. https://doi.org/10.1016/j.molbiopara.2006.06.008

Liston DR, Johnson PJ (1999) Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. Mol Cell Biol 19:2380–2388. https://doi.org/10.1128/MCB.19.3.2380

Liston DR, Lau AOT, Ortiz D, Smale ST, Johnson PJ (2001) Initiator recognition in a primitive eukaryote: IBP39, an initiator-binding protein from Trichomonas vaginalis. Mol Cell Biol 21:7872–7882. https://doi.org/10.1128/MCB.21.22.7872-7882.2001

Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356:83–85. https://doi.org/10.1038/356083a0

Millán-Pacheco C, Capistrán VM, Pastor N (2009) On the consequences of placing amino groups at the TBP-DNA interface. Does TATA really matter? J Mol Recognit 22:453–464. https://doi.org/10.1002/jmr.963

Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG, Burley SK (1996) Crystal structure of a human TATA box-binding protein/TATA element complex. Proc Natl Acad Sci 93:4862–4867. https://doi.org/10.1073/pnas.93.10.4862

Ong S-J, Huang S-C, Liu H-W, Tai J-H (2004) Involvement of multiple DNA elements in iron-inducible transcription of the ap65-1 gene in the protozoan parasite *Trichomonas vaginalis*. Mol Microbiol 52:1721–1730. https://doi.org/10.1111/j.1365-2958.2004.04088.x

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF chimera a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084

Quon DV, Delgadillo MG, Khachi A, Smale ST, Johnson PJ (1994) Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. Proc Natl Acad Sci 91:4579–4583. https://doi.org/10.1073/pnas.91.10.4579

Rigby PWJ (1993) Three in one and one in three: it all depends on TBP. Cell 72:7–10. https://doi.org/10.1016/0092-8674(93)90042-O

Ruvalcaba-Salazar OK, Ramírez-Estudillo MDC, Montiel-Condado D et al (2005) Recombinant and native *Plasmodium falciparum* TATA-binding-protein binds to a specific TATA box element in promoter regions. Mol Biochem Parasitol 140:183–196. https://doi.org/10.1016/j.molbiopara.2005.01.002

Schumacher MA, Lau AOT, Johnson PJ (2003) Structural basis of core promoter recognition in a primitive eukaryote. Cell 115:413–424. https://doi.org/10.1016/S0092-8674(03)00887-0

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2014) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539–539. https://doi.org/10.1038/msb.2011.75

Simoes-Barbosa A, Meloni D, Wohlschlegel JA, Konarska MM, Johnson PJ (2008) Spliceosomal snRNAs in the unicellular eukaryote *Trichomonas vaginalis* are structurally conserved but lack a 5′-cap structure. RNA 14:1617–1631. https://doi.org/10.1261/rna.1045408

Smith A, Johnson P (2011) Gene expression in the unicellular eukaryote *Trichomonas vaginalis*. Res Microbiol 162:646–654. https://doi.org/10.1016/j.resmic.2011.04.007

Smith AJ, Chudnovsky L, Simoes-Barbosa A, Delgadillo-Correa MG, Jonsson ZO, Wohlschlegel JA, Johnson PJ (2011) Novel core promoter elements and a cognate transcription factor in the divergent unicellular eukaryote *Trichomonas vaginalis*. Mol Cell Biol 31:1444–1458. https://doi.org/10.1128/MCB.00745-10

Thomas MC, Chiang C-M (2006) The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol 41:105–178. https://doi.org/10.1080/10409230600648736

Torres-Machorro AL, Hernández R, Sánchez J, López-Villaseñor I (2006) The 5S ribosomal RNA gene from the early diverging protozoa *Trichomonas vaginalis*. Mol Biochem Parasitol 145:269–273. https://doi.org/10.1016/j.molbiopara.2005.10.009

Vanacova S, Liston DR, Tachezy J, Johnson PJ (2003) Molecular biology of the amitochondriate parasites, *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. Int J Parasitol 33:235–255. https://doi.org/10.1016/S0020-7519(02)00267-9

Vannini A (2013) A structural perspective on RNA polymerase I and RNA polymerase III transcription machineries. Biochim Biophys Acta Gene Regul Mech 1829:258–264. https://doi.org/10.1016/j.bbagrm.2012.09.009

Vannini A, Cramer P (2012) Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Mol Cell 45:439–446. https://doi.org/10.1016/j.molcel.2012.01.023

Webb B, Sali A (2016) Comparative protein structure modeling using MODELLER. In: Current protocols in bioinformatics. Wiley, Hoboken, pp 5.6.1–5.6.37

White RJ, Jackson SP (1992) The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. Trends Genet 8:284–288. https://doi.org/10.1016/0168-9525(92)90255-3

World Health Organization (2011) Prevalence and incidence of selected sexually transmitted infections, *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, syphilis and *Trichomonas vaginalis*: methods and results used by WHO to generate 2005 estimates. WHO 1–38

Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its Core. J Mol Biol 430:2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007