UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

"Caracterización de Algunos Factores del Inicio de la Transcripción en *Trichomonas vaginalis*"

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:
M. en C. Olivia Parra Marín

TUTOR PRINCIPAL:
Dra. María Imelda López Villaseñor
Instituto de Investigaciones Biomédicas

MIEMBROS DEL COMITÉ TUTOR:
Dr. Roberto Coria Ortega
Instituto de Fisiología Celular

Dr. Ismael Bustos Jaimes
Facultad de Medicina

Ciudad de México. Abril, 2021

Universidad Nacional Autónoma de México
UNAM
Biblioteca Central
Dirección General de Bibliotecas de la UNAM

UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Este trabajo se realizó en el
Departamento de Biología Molecular y Biotecnología del
Instituto de Investigaciones Biomédicas de la
Universidad Nacional Autónoma de México bajo la
Dirección de la Dra. María Imelda López Villaseñor y la asesoría
del Dr. Ismael Bustos Jaimes y del Dr. Roberto Coria Ortega.

AGRADECIMIENTOS

A la Dra. Imelda López Villaseñor, quién dirigió el proyecto de investigación, mi más sincera gratitud por todo el apoyo brindado, por su orientación y la sabiduría con la que resuelve los asuntos del laboratorio, sus conocimientos me han resultados muy enriquecedores.

A la M. en C. Karla C. López Pacheco, por todo el apoyo técnico y experimental brindado para la realización de este proyecto. Le agradezco especialmente por su ayuda en el diseño y generación de las mutantes de levadura, su apoyo en la planeación y realización del experimento de doble híbrido y los ensayos de activad β-galactosidasa. Así como también por la obtención del plásmido pEG202-TvBRF1.

Al comité tutoral conformado por el Dr. Roberto Coria Ortega y el Dr. Ismael Bustos Jaimes por escuchar con paciencia la descripción de los resultados obtenidos semestre tras semestre y por brindar su valiosa asesoría.

Al Dr. Bernardo Franco y a la M. en C. Lluvia Rosas Hernández, quienes iniciaron el estudio de las TBPs en *T. vaginalis,* al Dr. Bernardo Franco por la obtención del plásmido pQE30-TvTBP1 y a la M. en C. Lluvia Rosas Hernández por la construcción de los plásmidos pQE30-TvTBP2, pQE30-IBP39, pEG202-TvTBP1, pEG202-TvTBP2 y pEG202-IBP39 que han sido ampliamente usados para el desarrollo de esta tesis.

A la Dra. Alexandra Ibáñez Escribano, quién obtuvo a la triple mutante de TvTBP1 y al plásmido pGEX4T1-TvBRF1.

Al Dr. Roberto Hernández Fernández y a la Dra. Ana María Cevallos Gaos y a su equipo de trabajo por las observaciones dadas al proyecto en cada seminario de grupo.

A la M. en C. Patricia de la Torre por la secuenciación de los plásmidos obtenidos en este trabajo.

Al M.V. Jorge Omar García Rebollar por su asistencia en la inmunización de ratones para la obtención del anticuerpo anti-IBP39.

A los integrantes del jurado de tesis, el Dr. Ruy Enrique Pérez Montfort, la. Dra. Marcela Lizano Soberón, la Dra. Rosa Laura Camarena Mejía, la Dra. Tzvetanka Dimitrova Dinkova y la Dra. María Soledad Funes Arguello, por el tiempo invertido en la revisión de esta tesis.

Al Programa de Apoyo a los Estudios de Posgrado (PAEP) por los apoyos que me permitieron asistir al XXX Congreso Nacional de Bioquímica del 2 al 8 de noviembre de 2014 en Guadalajara, Jalisco, realizar el curso de Seguridad y Protección Radiológica a Nivel Personal Ocupacionalmente Expuesto, con duración de 40 horas, impartido por el Centro Nuclear "Dr. Nabor Carrillo Flores" del 23 al 27 de Marzo de 2015 y al Congreso Internacional "*Mechanism of Eukaryotic Transcription*" en la universidad *Cold Spring Harbor* del 29 de Agosto al 2 de Septiembre de 2017.

Finalmente, al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo otorgado con No. 302910 para los estudios de doctorado.

DEDICATORIA

A mi familia, quienes me han apoyado incondicionalmente,
pero especialmente a mi padre, que siempre estuvo orgulloso de mí.
Siempre te voy a extrañar…

INDICE

ABREVIATURAS 1

RESUMEN 3

ABSTRACT 4

INTRODUCCIÓN 5

TRICHOMONAS VAGINALIS 5
EPIDEMIOLOGÍA 6
MORFOLOGÍA Y CICLO DE VIDA DE *T. VAGINALIS* 6
POSICIÓN FILOGENÉTICA DE *T. VAGINALIS* 7
GENOMA Y TRANSCRIPCIÓN DE *T. VAGINALIS* 8
RNA POLIMERASAS DE *T. VAGINALIS* 8
ELEMENTOS EN LOS PROMOTORES DE *T. VAGINALIS* 9
FACTORES DE TRANSCRIPCIÓN EN *T. VAGINALIS* 11
IBP39 COMO UN FACTOR DE TRANSCRIPCIÓN ESPECIALIZADO DE *T. VAGINALIS* 12
EXPRESIÓN REGULADA POR HIERRO 13

OBJETIVO GENERAL 15

OBJETIVOS PARTICULARES 15

MATERIALES Y MÉTODOS 16

CULTIVO DE *T. VAGINALIS* 16
EXTRACCIÓN DE DNA GENÓMICO DE *T. VAGINALIS* 16
MODELADO DE LAS ESTRUCTURAS TERCIARIAS 17
OBTENCIÓN DE EXTRACTOS PROTEICOS DE *T. VAGINALIS* 17
LOCALIZACIÓN CELULAR 18
FORMACIÓN DE DÍMEROS 18
COPRECIPITACIÓN MEDIANTE LA ETIQUETA DE GST 18
ENSAYO DE RETARDO DE LA MOVILIDAD ELECTROFORÉTICA (EMSA) 19
1. PREPARACIÓN DE LA SONDA 19
2.- PREPARACIÓN DE LAS PROTEÍNAS 19
3.-PREPARACIÓN DE LA REACCIÓN DE UNIÓN 20
GENERACIÓN DE LAS CEPAS MUTANTES 20
1.- OBTENCIÓN DEL CASETE *KANMX*: 20
2.- TRANSFORMACIÓN DE LA CEPA DE LEVADURA 21
3.- EXTRACCIÓN DE DNA GENÓMICO DE LEVADURA 21
4.- COMPROBACIÓN DE LA CEPA 22
INDUCCIÓN DE LA ESPORULACIÓN DE LAS CEPAS MUTANTES 23
DOBLE HÍBRIDO 23
OBTENCIÓN DE LOS PLÁSMIDOS 23
OBTENCIÓN DE LAS CEPAS DE LEVADURA TRANSFORMADAS 24
GOTEOS 25
ENSAYO DE ACTIVIDAD β-GALACTOSIDASA 25

SOUTHWESTERN BLOT **26**
GENERACIÓN DEL ANTICUERPO ANTI-IBP39 **26**

CAPÍTULO I: LA DIVERSIDAD DE LAS PROTEÍNAS TBPS EN PROTISTAS **28**

ESTRUCTURA DE TBP **29**
OBTENCIÓN DE LOS MODELOS DE TBPS DE PROTISTAS **30**
MECANISMOS DE UNIÓN AL DNA **31**
CONSERVACIÓN DE LOS AMINOÁCIDOS INVOLUCRADOS EN LA UNIÓN AL DNA DE PROTISTAS **32**
LAS TBPS DE PROTISTAS **34**
AMEBOZOA 34
APICOMPLEXA 34
EXCAVATA 34
DINOFLAGELLATA 35
EL POTENCIAL ELECTROSTÁTICO **36**
FACTORES RELACIONADOS A TBP **37**
TRF1 38
TRF2 39
TRF3 39
TRF4 40
CONSERVACIÓN EVOLUTIVA **40**

CAPÍTULO II: LAS TVTBPS COMO PARTE DE LA MAQUINARIA TRANSCRIPCIONAL DE *T. VAGINALIS* **42**

ANTECEDENTES **42**
RESULTADOS: **43**
SIMILITUD DE SECUENCIA Y ESTRUCTURA TERCIARIA DE LAS TVTBPS 43
LOCALIZACIÓN CELULAR DE TVTBP1 45
FORMACIÓN DE DÍMEROS DE TBP 46
UNIÓN DE LAS TVTBPS A IBP39 48
UNIÓN DE LAS TVTBPS AL DNA 49
UNIÓN DE LAS TVTBPS A PROMOTORES TRANSCRITOS POR LAS TRES RNA POLIMERASAS 50
TVTBP1 Y TVTBP2 SE UNEN AL DNA DE MANERA SECUENCIA-INDEPENDIENTE 52
CONCLUSIONES **55**
PERSPECTIVAS **55**

CAPÍTULO III: EL COMPLEJO TFIIIB EN *T. VAGINALIS* **56**

INTRODUCCIÓN **56**
TFIIIA 57
TFIIIB 58
TFIIIC 59
EL COMPLEJO TFIIIB EN *T. VAGINALIS* 60
RESULTADOS: **60**
TVBRF1 TIENE DOMINIOS CARACTERÍSTICOS DE LA FAMILIA RELACIONADA CON TFIIB 60
TVBRF1 NO COMPLEMENTA UNA CEPA MUTANTE DE LEVADURA 62
IDENTIFICACIÓN DE TVBDP1 64
INTERACCIÓN PROTEÍNA-PROTEÍNA EVALUADA POR DOBLE HÍBRIDO 66
ENSAYOS DE COPRECIPITACIÓN DE LAS PROTEÍNAS DEL COMPLEJO TFIIIB 70
EL COMPLEJO TFIIIB SE UNE A LOS PROMOTORES 5S Y U6 73

CONCLUSIONES 80
PERSPECTIVAS 80

CAPÍTULO IV: OTROS FACTORES ASOCIADOS A TBP EN COMPLEJOS QUE SE UNEN A PROMOTORES TRANSCRITOS POR LA RNA POLIMERASA I Y II EN *T. VAGINALIS* 82

INTRODUCCIÓN 82
IDENTIFICACIÓN DE TFIIB EN *T. VAGINALIS* 85
RESULTADOS: 86
IDENTIFICACIÓN DE RRN7 EN *T. VAGINALIS* 86
DOMINIOS DE LOS FACTORES RELACIONADOS A TFIIB 87
OBTENCIÓN DE LAS CEPAS MUTANTES DE *S. CEREVISIAE*: *SUA7*/*SUA7*Δ::*KANMX*, *RRN7*/*RRN7*Δ::*KANMX* Y *BRF1*/*BRF1*Δ::*KANMX* 89
ENSAYO DE COMPLEMENTACIÓN DE UNA MUTANTE DE LEVADURA DE TFIIB 89
ENSAYOS DE INTERACCIÓN POR DOBLE HÍBRIDO PARA TFIIB Y RRN7 91
ENSAYOS DE INTERACCIÓN POR COPRECIPITACIÓN (*GST PULLDOWN)* 92
TFIIB INTERACTÚA CON IBP39 94
UNIÓN DE TFIIB AL DNA 96
ENSAYO DE *SOUTHWESTERN BLOT* 97
ENSAYOS EMSA PARA TFIIB 98
TVRRN7 NO RECONOCE SECUENCIAS EN EL DNA 102
CONCLUSIONES 102

CONSIDERACIONES FINALES 103

BIBLIOGRAFÍA 105

ANEXO 1: AMORTIGUADOR DE UNIÓN PROBADOS EN LOS ENSAYOS DE RETARDO DE LA MOVILIDAD ELECTROFORÉTICA (EMSA) 115

ANEXO 2: DATOS DE LA ACTIVIDAD β-GALACTOSIDASA 117

ANEXO 3: ARTÍCULOS DERIVADOS DEL PRESENTE TRABAJO 119

ABREVIATURAS

AD	Dominio de Activación (*Activation domain*)
BD	Dominio de Unión (*Binding domain*)
BDP1	*B Double Prime 1*
BME	β-mercaptoetanol
BRE	Elemento de reconocimiento de TFIIB (*TFIIB-recognition element)*
BRF1	*B-Related Factor 1*
CF	*Core factor*
CFA	Adyuvante completo de Freund (*Complete Freund's adjuvant*)
CTD	Dominio carboxilo terminal (*Carboxi-terminal domain*)
DMSO	Dimetilsulfóxido
DO	Densidad óptica
DSE	Elemento de secuencia distal (*Distal sequence element*)
DTT	Ditiotreitol
EMSA	Ensayo de retardo de la movilidad electroforética (*Electrophoretic mobility shift assay*)
EST	Etiqueta de secuencia expresada (*Expressed sequence tag*)
IBD	Dominio de union al Inr (*Inr binding domain*)
IBP39	Proteína de unión al Inr de 39 kDa (*Inr binding protein of 39 kDa*)
ICE	Elemento de control interno (*Internal control element*)
IE	Elemento intermedio (*Intermediate element*)
IFA	Adyuvante incompleto de Freund (*Incomplete Freund's adjuvant*)
Inr	Iniciador (*Initiator*)
M3BP	Proteína de unión al motivo 3 (*Motif 3 binding protein*)
MRE	Elemento de reconocimiento a Myb (*Myb recognition element*)
ONPG	Orto-nitrofenil-β-galactósido
PIC	Complejo de preinicio (*Preinititation complex*)
PSE	Elemento de secuencia proximal (*Proximal sequence element*)
RT-PCR	Reacción en cadena de la polimerasa – Transcripción inversa (*Reverse transcription polymerase chain reaction*)
SDS	Duodecilsulfato de sodio (*Sodium dodecyl sulfate*)
SL	Cadena líder (*Splice leader*)
SL1	Factor selective 1 (*Selective factor 1*)
SNAPc	Complejo proteico de activación del SnRNA (*SnRNA activating protein complex*)
SSE	Elemento del sitio de inicio (*Start site element*)
TAE	Tris-Acetato-EDTA
TAF	Factor asociado a TBP (*TBP-associated factor*)
TBE	Tris-Borato-EDTA
TBP	Proteína de unión a caja TATA (*TATA-box binding protein*)
TBPL1	TBP tipo 1 (*TBP-like 1*)
TFIIB	Factor de transcripción IIB (*Transcription factor II B*)

TG	Tris-glicina
TLF	Factor tipo TBP (*TBP-like factor*)
TLP	Proteína tipo TBP (*TBP-like protein*)
TRF1	Factor 1 relacionado a TBP (*TBP-related factor 1*)
TRP	Factor relacionado a TBP (*TBP-related factor*)
TSS	Sitio de inicio de la transcripción (*Transcription start site*)
UAF	Factor de actividad rio arriba de unión al UAS (*UAS-binding upstream activity factor*)
UCE	Elemento de control rio arriba (*Upstream control element*)
UPE	Elemento del promotor rio arriba (*Upstream promoter element*)
UTR	Región no traducida (*Untranslated region*)
WT	Silvestre (*Wild type*)
X-gal	5-bromo-4-cloro-3-indolil-β-D-galactosidasa
YPD	Extracto de levadura, peptona y dextrosa (*Yeast extract*, *peptone*, *dextrose*)
YPGAL	Extracto de levadura, peptona y galactosa (*Yeast extract*, *peptone*, *galactose*)

RESUMEN

La síntesis de RNA a partir de DNA es un proceso conocido como transcripción y es considerado como el primer paso de la expresión génica. En eucariontes este proceso es llevado a cabo por al menos tres RNA polimerasas y consta de tres etapas: el inicio, la elongación y la terminación. La regulación de la expresión génica es fundamental para el desarrollo de las células y es el inicio de la transcripción el principal punto de control. Las RNA polimerasas no tienen la capacidad de reconocer el sitio de inicio de la transcripción y requieren de factores transcripcionales que las recluten y orienten adecuadamente al promotor, entre ellos TBP (*TATA-box binding protein*). Este factor transcripcional participa en el complejo de inicio de la transcripción de las tres RNA polimerasas, convirtiéndolo en un factor clave para la formación de diversos complejos. Adicionalmente, existen factores estructural y funcionalmente relacionados que participan específicamente en cada una las tres maquinarias transcripcionales, estos son los factores relacionados a TFIIB, entre los cuales se encuentran Rrn7, TFIIB y Brf1 que interaccionan con TBP y participan en la transcripción de la RNA polimerasa I, II y III, respectivamente.

Trichomonas vaginalis es el agente causal de la tricomoniasis, el parásito que origina la enfermedad de transmisión sexual no viral, más importante en el mundo. Este organismo es de divergencia evolutiva temprana y junto a otros protistas, engloban la mayor diversidad del reino eucarionte. A pesar de ello, los miembros de este grupo taxonómico son los menos estudiados a nivel molecular. En este trabajo se hace una caracterización de algunos de los factores de inicio de la transcripción en *T. vaginalis* y se divide en cuatro capítulos, en el primero se discuten las características atípicas de las TBPs de protistas, incluidas las de *T. vaginalis*. En el segundo, se realizó una descripción de dos TBPs en *T. vaginalis*. En el tercer capítulo, se profundizó sobre la unión de las TBPs con Brf1 y su posible participación en un complejo conocido como TFIIIB en *T. vaginalis*. Finalmente, en el cuarto capítulo se hizo una descripción breve de los factores Rrn7 y TFIIB..

ABSTRACT

The synthesis of RNA from DNA is called transcription and is considered the first step in gene expression. In eukaryotes, this process is carried out by at least three RNA polymerases and consists of three stages: initiation, elongation, and termination. Regulation of gene expression is essential for cell development and transcription initiation is considered as the main checkpoint. These RNA polymerases are not capable to recognize the transcription start point by themselves and require the so-called transcription factors to recruit and position them to the promoter, among these is the TBP (TATA-box binding protein). This transcription factor participates in the assembly of the transcription initiation complexes of the three RNA polymerases, making it a key factor in the formation of highly diverse complexes. Furthermore, additional structural and functionally related factors have been described that specifically participate in each of the three transcriptional machineries, these factors are the so called TFIIB-related factors. Among them, Rrn7, TFIIB and Brf1 interact with TBP and participate in the transcription of RNA polymerase I, II and III, respectively.

Trichomonas vaginalis is the causative agent of trichomoniasis, the parasite that causes the most important non-viral sexually transmitted disease in the world. This early divergent organism, together with other protists, encompasses the greatest diversity in the eukaryotic kingdom. Yet, the members of this taxonomic group are still among the least studied at the molecular level.

In this work, a brief characterization of the transcription initiation factors in *T. vaginalis* is made and it is divided into four chapters: the first chapter, the atypical characteristics of protists TBPs, including those of *T. vaginalis*, are discussed. In the second chapter is the characterization of two TBPs from *T. vaginalis*. The third chapter, the binding of the TvTBPs to Brf1 and their possible participation in a complex known as TFIIIB in *T. vaginalis*, were studied in depth. Finally, in the fourth chapter, focuses on a brief description of the Rrn7 and TFIIB factors.

INTRODUCCIÓN

El genoma de las células eucarióticas se transcribe principalmente por tres RNA polimerasas: la RNA polimerasa I que transcribe el RNA ribosomal (28S, 18S y el 5.8S); la RNA polimerasa II que transcribe la mayoría de los genes codificantes de proteínas y algunos RNAs no codificantes y la RNA polimerasa III que transcribe pequeños RNAs como los tRNAs, la subunidad 5S del ribosoma (rRNA 5S) y el snRNA U6 (Thomas y Chiang, 2006). Sin embargo, las RNA polimerasas no tienen la capacidad para reconocer el sitio de inicio de la transcripción y requieren de factores de transcripción, que las orienten al promotor. Los factores de transcripción junto con la RNA polimerasa conforman el complejo de preinicio de la transcripción (PIC).

En la transcripción llevada a cabo por la RNA polimerasa II, la formación del PIC comienza con la unión de TFIID al promotor. TFIID es un complejo que consiste de TBP y 13 factores asociados a TBP (TAFs) (Juven-Gershon y Kadonaga, 2010). La proteína TBP es un componente clave debido a que participa en los complejos de inicio de la transcripción de las tres RNA polimerasas (Cormack y Struhl, 1992).

TBP es una proteína altamente conservada desde arqueas hasta eucariontes superiores (Hernandez, 1993) y participa en el reconocimiento del promotor a través de la unión con las cajas TATA, un elemento de 8 pb rico en adeninas y timinas presente en el promotor basal y cuya secuencia es TATAWAAR, usualmente localizado a -30 pb del sitio de inicio de la transcripción (TSS) (Patikoglou *et al.*, 1999).

La regulación transcripcional de los genes codificantes de proteínas ha sido estudiada ampliamente sobre todo en eucariontes superiores, lo que ha permitido la identificación de secuencias de DNA conservadas que permiten el inicio de la transcripción a través del reclutamiento de factores generales de la transcripción de manera concertada. Sin embargo, en eucariontes inferiores, dado su origen evolutivo, estos procesos son poco conocidos. Con la liberación de la secuenciación del genoma de *T. vaginalis* en 2007, se proporcionó una plataforma importante para las investigaciones moleculares y celulares, entre ellos los de transcripción.

Trichomonas vaginalis

Trichomonas vaginalis es un parásito protozoario flagelado y microaerofílico que se adhiere al tracto urogenital de los humanos causando tricomoniasis, la mayor infección de transmisión sexual no viral en el mundo (WHO, 2011).

Epidemiología

T. vaginalis infecta el tracto genitourinario inferior de mujeres, así como la próstata y uretra de los hombres. *T. vaginalis* tiene un período de incubación de cinco a veintiocho días. Las mujeres con tricomoniasis a menudo se quejan de un flujo vaginal amarillo o verde maloliente, dispareunia, frecuencia urinaria, disuria, prurito o eritema vulvar. En los hombres a menudo no causa síntomas. Si un hombre es sintomático, el síntoma más común es la uretritis. Con menos frecuencia, los hombres también pueden desarrollar prostatitis y epididimitis (Bouchemal *et al.*, 2017). La infección por *T. vaginalis* puede tener severas complicaciones, tales como partos prematuros, bajo peso al nacer, infertilidad, enfermedad inflamatoria pélvica atípica, aumento de la predisposición a la infección por el virus VIH e incremento del riesgo de cáncer cervical (Harp y Chowdhury, 2011).

Morfología y ciclo de vida de *T. vaginalis*

Las células de *T. vaginalis* son típicamente piriformes, pero en contacto con las células del epitelio vaginal, el parásito cambia su apariencia a una forma ameboide (Petrin *et al.*, 1998). El tamaño de este organismo puede variar entre 9-23 x 7 µM (en promedio 13 µm) (Harp y Chowdhury, 2011). Estructuralmente, posee cuatro flagelos anteriores y un quinto flagelo que se encuentra asociado a una membrana ondulante. Esta estructura es soportada por la costa (Benchimol, 2004). El núcleo se localiza en la parte anterior del cuerpo celular. Otras estructuras como el axostilo (que recorre longitudinalmente el cuerpo celular) y el pelto se encuentran formados por microtúbulos (Figura 1) (Benchimol, 2004).

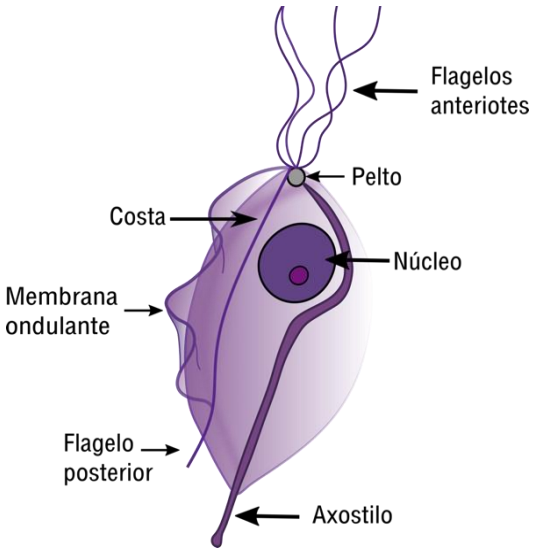


Figura 1.- Morfología de *Trichomonas vaginalis.*

La reproducción de este parásito es a través de fisión binaria mediante un tipo especial de mitosis que se le conoce como cerrada o primitiva, debido a que los cromosomas no hacen contacto directo con los husos mitóticos durante la segregación cromosómica. Este proceso es llevado a cabo sin la desaparición de la membrana nuclear (Ribeiro *et al.*, 2002).

El ciclo de vida de *T. vaginalis* consiste en un solo estadio, el trofozoíto (forma replicativa) sin la formación de quistes. Sin embargo, la presencia de formas esféricas con internalización del flagelo, denominados pseudoquistes (formas no proliferativas), han sido comúnmente observados bajo

condiciones de estrés, tales como la depleción de nutrientes; aunque también se han observado naturalmente en el hospedero, sugiriendo que puede coexistir con las formas replicativas (Pereira-Neves *et al.*, 2003). Los pseudoquistes son formas no móviles que carecen de una pared quística verdadera. Adicionalmente, *T. vaginalis* carece de mitocondrias y peroxisomas. En lugar de las mitocondrias, el parásito obtiene energía de un organelo conocido como hidrogenosoma, una estructura de membrana doble típico de los amitocondriados (Kulda, 1999). Los hidrogenosomas carecen de citocromos y de las enzimas responsables para la cadena respiratoria mitocondrial, pero fermentan carbohidratos y producen hidrógeno molecular (Harp y Chowdhury, 2011), a su vez, no contienen DNA. El ser humano es el único hospedero de *T. vaginalis* y su transmisión es por contacto sexual.

Posición filogenética de *T. vaginalis*

T. vaginalis pertenece al super grupo de Excavata. Estudios filogenéticos usando a la subunidad grande y pequeña del RNA ribosomal indican que este organismo se encuentra entre los eucariontes de divergencia evolutiva temprana (Sogin, 1991; Cavalier-Smith y Chao, 1996; Sogin y Silberman, 1998), convirtiendo a *T. vaginalis* en un organismo de interés para el entendimiento de la evolución eucariótica. La transcripción es un proceso ampliamente estudiado en animales, plantas y hongos; sin embargo, estos organismos no representan la gran diversidad de los eucariontes. Como se observa en la figura 2, el árbol filogenético demuestra que *T. vaginalis* es un organismo que divergió antes que el resto de los eucariontes, tales como hongos, plantas y animales (Quon *et al.*, 1994). Dada la posición filogenética de *T. vaginalis* es un buen modelo para comprender la transcripción en organismos de evolución temprana.

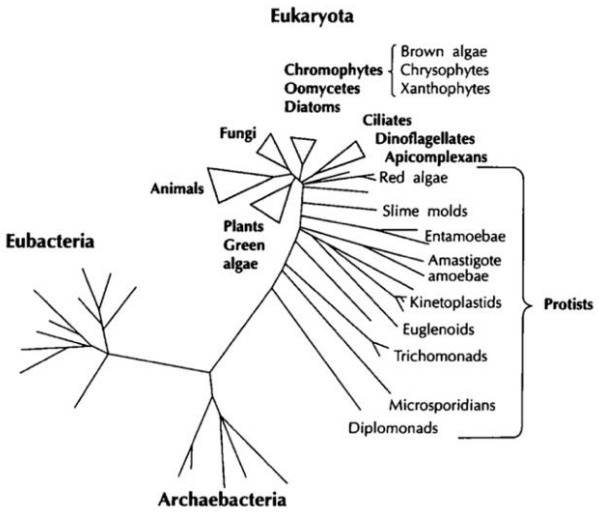


Figura 2.- Árbol filogenético de tres reinos (Eubacteria, Arqueobacteria y Eucariotas) inferidos a partir de la similitud de secuencia de la subunidad 16S y 18S del rRNA (Sogin, 1991).

Genoma y transcripción de *T. vaginalis*

El genoma de *T. vaginalis* fue secuenciado por Carlton y colaboradores en 2007, describiendo que este parásito tiene una gran capacidad codificante; se estimaron aproximadamente 60,000 genes en aproximadamente 160 Mpb que se organizan en 6 cromosomas (Carlton *et al.*, 2007). Sin embargo, actualmente se estiman 46,000 genes; este dato se obtuvo después de eliminar las secuencias de genes que parecen ser sólo fragmentos de otros (Smith y Johnson, 2011). A pesar de esto, *T. vaginalis* posee uno de los genomas más largos entre otros protistas de importancia médica (Singh *et al.*, 2011). Los datos del transcriptoma actualmente disponibles indican que solo cerca de la mitad de los genes codificantes parecen ser expresados (Gould *et al.*, 2013).

La presencia de intrones solo se ha encontrado en 65 genes (Carlton *et al.*, 2007). Los genes que codifican para una maquinaria de *splicing*, U1, U2, U4, U5 y U6 también han sido identificados (Simoes-Barbosa *et al.*, 2008). Adicionalmente, el genoma de *T. vaginalis* contiene un gran número de genes que codifican histonas, entre las cuales podemos identificar a H2A, H2B, H3 y H4 (Cong *et al.*, 2010). La secuencia de aminoácidos de las histonas H3 y H4 son divergentes de sus homólogos en eucariontes, no solo de organismos multicelulares sino también de otros protistas (Marinets *et al.*, 1996).

RNA polimerasas de *T. vaginalis*

La RNA polimerasa I consta de 14 subunidades, la RNA polimerasa II de 12 subunidades y la RNA polimerasa III de 17 subunidades. En los análisis *in silico* realizados por el grupo de Carlton en 2007, se identificaron diversos genes que codifican para varias de las subunidades de las tres RNA polimerasas (Carlton *et al.*, 2007), indicando su funcionalidad. Adicionalmente, experimentos de inhibición de la transcripción usando a la toxina α-amanitina en *T. vaginalis* proveen evidencia de que hay transcripción llevada a cabo por las tres distintas RNA polimerasas (Vanácová *et al.*, 2001). Estas tres RNA polimerasas muestran una sensibilidad diferencial a la toxina α-amanitina. En metazoarios, la RNA polimerasa II es altamente sensible, mostrando una inhibición del 50 % en concentraciones de 2-20 µg/ml, curiosamente la RNA polimerasa II de *T. vaginalis* es resistente a estas concentraciones y alcanza el 50 % de inhibición con 250 µg/ml de esta toxina (Quon *et al.*, 1996). El sitio de unión de la α-amanitina se encuentra en el dominio conocido como "*bridge*" de la subunidad RPB1 de la RNA polimerasa II (Bushnell *et al.*, 2002); sin embargo, no todos los residuos de este dominio están conservados en la subunidad RBP1 de *T. vaginalis*. Adicionalmente, el CTD (dominio carboxilo terminal) de RPB1 de eucariontes superiores, el cual es fosforilado y

participa a lo largo del proceso de la transcripción, se caracteriza por contener un heptapéptido YSPTSPS repetido 26 veces en levadura y 52 veces en mamíferos, mientras que el dominio C-terminal de RPB1 de *T. vaginalis* consta de un dipéptido de serina-prolina repetido 19 veces (Vanacova *et al.*, 2003). El número de repetidos generalmente refleja la complejidad del organismo (Hsin *et al.*, 2014). A pesar de esto, se sugiere que, debido a la presencia de serinas, el CTD puede ser fosforilado durante el inicio de la transcripción, al igual que en eucariontes superiores (Vanacova *et al.*, 2003).

Elementos en los promotores de *T. vaginalis*

Los promotores de *T. vaginalis* tienen una estructura bipartita; es decir, tienen elementos conservados en el promotor basal (*core promoter*) y elementos regulatorios río arriba (Figura 3).

Promotores transcritos por la RNA polimerasa I

Las secuencias codificantes de las subunidades 18S, 5.8S y 28S del ribosoma se encuentran en unidades repetidas en tándem y constituyen el cistrón ribosomal. Cada unidad es separada por una secuencia intergénica que varía en longitud entre especies. El promotor del cistrón ribosomal de *T. vaginalis* es una región compacta de aproximadamente 180 pb (-128/+50 pb), que contiene dos elementos; el elemento de control río arriba (UCE, por sus siglas en inglés) que se localiza a aproximadamente a -80 pb del sitio de inicio de la transcripción (TSS, por sus siglas en inglés) y el elemento central (*core*) a -27 pb. Ambos elementos son secuencias repetidas.(Franco *et al.*, 2012). El elemento de terminación de la transcripción también fue identificado a 34 pb río abajo de la secuencia codificante de la subunidad 28S. La función de este elemento depende de la polaridad, así como de la presencia del tracto de uridinas y de la estructura en forma de horquilla en el transcrito (Franco *et al.*, 2012).

Promotores transcritos por la RNA polimerasa II

Las cajas TATA canónicas alrededor de -30 pb parecen estar ausentes en los promotores codificantes de proteínas de *T. vaginalis* (Quon *et al.*, 1994); sin embargo, las regiones 5'UTR de los genes son inusualmente ricas en A+T (Liston y Johnson, 1998). En su lugar, el único elemento descrito del promotor basal es el elemento Inr *(Initiator),* el cual está altamente conservado tanto estructural como funcionalmente respecto a su contraparte en eucariontes superiores y parece ser

el responsable del reconocimiento adecuado del promotor (Liston *et al.*, 1999). El Inr de *T. vaginalis* es un elemento de 13 nucleótidos, cuya secuencia consenso es TCA^{+1}YTWYTCATTA, donde la adenosina representa la posición +1 (Quon *et al.*, 1994) y se localiza alrededor de 20 nucleótidos río arriba del codón de inicio. Este elemento parece estar presente en ~75 % de los promotores de genes codificantes de proteínas (Carlton *et al.*, 2007).

La señal de poliadenilación en los mRNAs de *T. vaginalis* es un tetranucleótido (UAAA), a diferencia del hexanucléotido de eucariontes superiores (AAUAAA). Adicionalmente esta señal, muchas veces se encuentra sobrepuesta con el codón de paro (TAA). Asimismo, se identificó el sitio de procesamiento Y↓(A)2-5UU y la presencia de la región rica en adeninas en el 3'UTR de la secuencia codificante (Fuentes *et al.*, 2012).

Promotores transcritos por la RNA polimerasa III

Mediante análisis *in silico* por homología se identificaron algunos elementos del promotor del gen 5S de *T. vaginalis*: el elemento de control interno (ICE, por sus siglas en inglés) que consta de la caja A [5'-N(G/C)(C/T)(C/T)AANCNNNNNNN-3'], el elemento intermedio [IE: 5'-(C/G)NN(G/A)(G/A)N-3'] y la caja C [5'NNG(G/A)TGGGNG(T/A)CCN(C/T)NNG-3'] (Torres-Machorro *et al.*, 2006). Adicionalmente, se identificaron elementos regulatorios río arriba similares al elemento del sitio de inicio (SSE:5'-AAACTATC-3') y al elemento del promotor río arriba (UPE: 5'-TGCG-3') comunes en las secuencias del 5S de la levadura *Saccharomyces cerevisiae*. La transcripción del gen 5S termina con una secuencia de timinas (Torres-Machorro *et al.*, 2006).

El promotor de U6 presenta un elemento en la posición -28 con similitud a una caja TATA canónica y 6 residuos de timina en la secuencias 3'UTR (Simoes-Barbosa *et al.*, 2008).

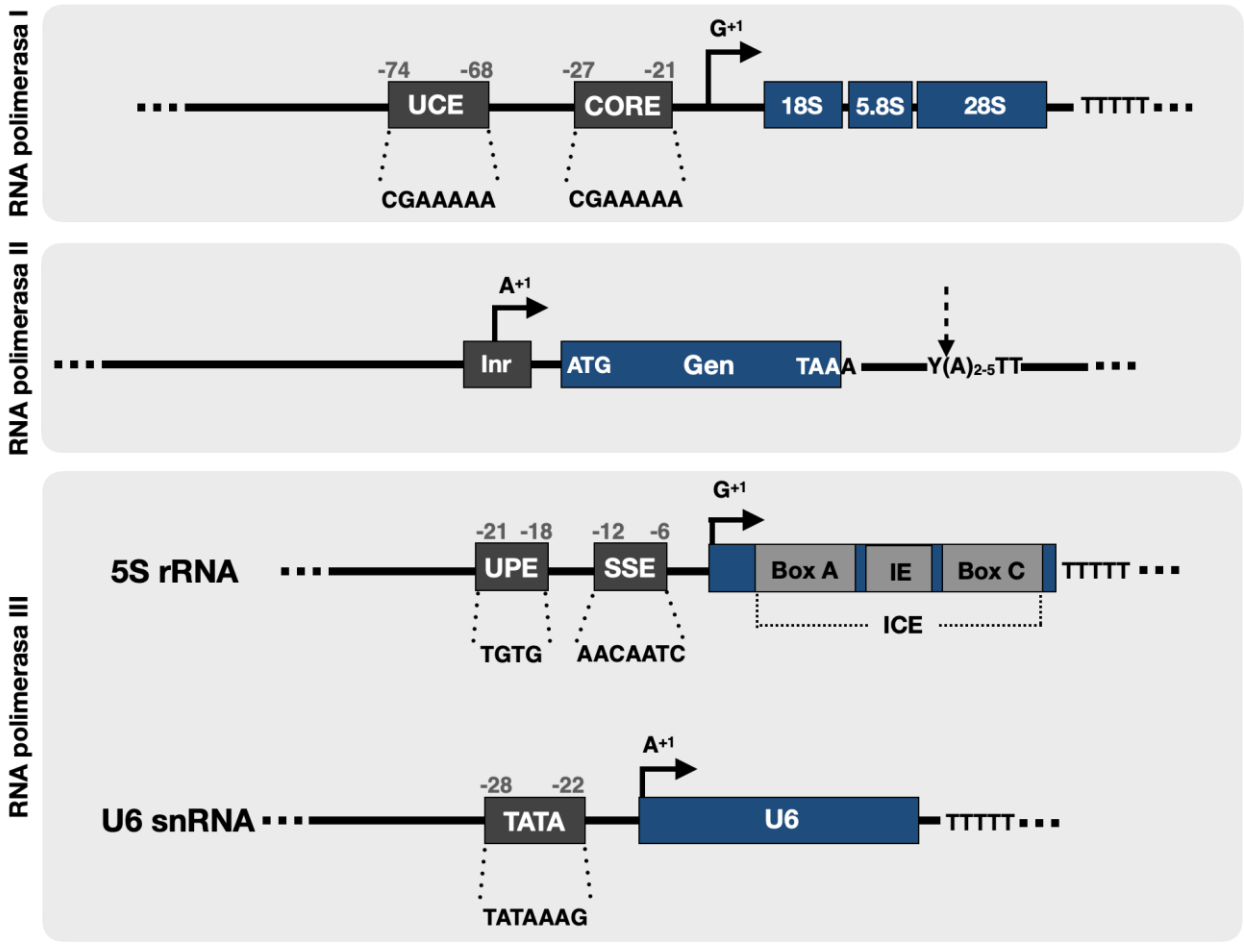


Figura 3.- Promotores transcritos por la RNA polimerasa I, II y III de *T. vaginalis*. El promotor del cistrón ribosomal transcrito por la RNA polimerasa I contiene a los elementos UCE (por sus siglas en inglés de *Upstream Control Element*) a -74 pb y el *core* a -27 pb; el sitio de termino de la transcripción se encuentra a 34 pb de la secuencia codificante 28S. Los promotores de genes codificantes de proteínas transcritos por la RNA polimerasa II, contienen un elemento Inr alrededor del sitio de inicio de la transcripción. Adicionalmente se señala el codón de inicio (ATG), el codón de paro (TAA) que coincide con la señal de poliadenilación (UAAA) en el mRNA. El sitio de procesamiento del mRNA Y↓(A)2-5UU, también se encuentra señalado. Los promotores de los genes que codifican para el 5S y el U6 son transcritos por la RNA polimerasa III, el promotor 5S tiene dos elementos río arriba: UPE (por sus siglas en inglés de *Upstream Promoter Element*) y el SSE (por sus siglas en inglés de *Star Site Element*), adicionalmente se han identificado un elemento de control interno (ICE) conformado por la caja A, un elemento intermedio (IE) y la caja C. El promotor de U6 contiene una caja TATA putativa a -28 pb del sitio de inicio de la transcripción que se encuentra indicado con una flecha en cada caso.

Factores de transcripción en *T. vaginalis*

En el genoma de *T. vaginalis* se encuentran anotados dos genes putativos que codifican para TBP (Carlton *et al.*, 2007), a pesar de ello, la ausencia de cajas TATA canónicas y los análisis sobre los

promotores sugieren que la caja TATA no es usada para dirigir la transcripción de los promotores codificantes de proteínas de *T. vaginalis* (Liston y Johnson, 1999). Adicionalmente, en el genoma se han identificado otros factores generales de la transcripción, entre ellos, la mayoría de los factores asociados a TBP, los genes que codifican para los factores transcripcionales TFIIB, TFIIE, TFIIF y algunas subunidades de TFIIH, TFIIA se encuentran aparentemente ausente (Carlton *et al.*, 2007). El mediador es un complejo multiproteico requerido para la transcripción de la RNA polimerasa II. El mediador tiene interacción con activadores o represores de la transcripción y con factores generales de la transcripción. Mediante análisis del genoma de *T. vaginalis* se han identificado 44 subunidades putativas del mediador (Bourbon, 2008).

IBP39 como un factor de transcripción especializado de *T. vaginalis*

El elemento Inr en otros organismos es reconocido por las proteínas TAF1 y TAF2 que forman parte del complejo TFIID (Smale y Kadonaga, 2003), mientras que la proteína IBP39 (*39 kDa Initiator Binding Protein*) de *T. vaginalis*, de la cuál no se han encontrado ortólogos, se identificó como una proteína que se une a este elemento (Liston *et al.*, 2001). IBP39 es una proteína nuclear que consta de dos dominios conectados por una región sensible a proteasas (Figura 4A) (Liston *et al.*, 2001; Lau *et al.*, 2003). El dominio N-terminal o IBD (*Inr Binding Domain*) de 14.5 kDa es el responsable del reconocimiento del elemento Inr (Figura 4A y 4B) (Liston *et al.*, 2001; Lau *et al.*, 2003; Schumacher *et al.*, 2003), mientras que el dominio C-terminal de 24.5 kDa es responsable del reclutamiento de la RNA polimerasa II al sitio de inicio de la transcripción, mediante su interacción con el CTD de la RNA polimerasa II (Figura 4A y 4C) (Schumacher *et al.*, 2003; Lau *et al.*, 2006). La estructura de IBP39 ha sido reportada en presencia y ausencia del elemento Inr (Schumacher *et al.*, 2003). La estructura del IBD tiene características híbridas entre el dominio de unión a DNA procarionte hélice-vuelta-hélice (HTH, *Helix-Turn-Helix*) y el dominio de unión a DNA eucarionte (ETS, *E26 Transformation-Specific*) sin tener identidad de secuencia con estos dominios de unión a DNA (Schumacher *et al.*, 2003). En el proteoma de *T. vaginalis* se han identificado al menos 100 proteínas con el dominio IBD (Iyer *et al.*, 2008). Estos datos sugieren que el IBD es un motivo de unión a DNA específico de linaje. La presencia de IBP39 hace posible que en *Trichomonas* exista un mecanismo único de reconocimiento del promotor.

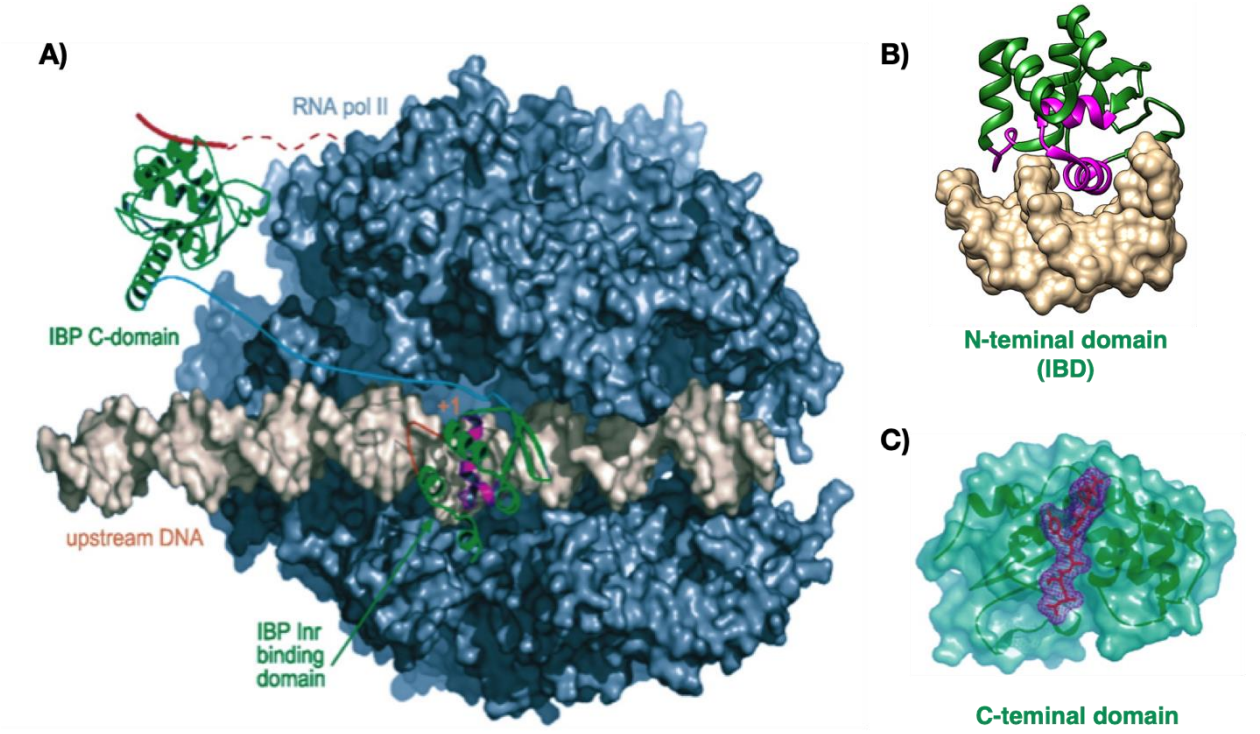


Figura 4.- Estructura de IBP39. A) Modelo de IBP39 (verde) unido al DNA (café), ambos dominios de IBP39 se encuentran unidos por una secuencia espaciadora (*linker)* sensible a proteasas (azul), a su vez se remarca el CTD de la polimerasa (rojo) y al núcleo (*core*) de la RNA polimerasa II (gris). B) Estructura del N-terminal de IBP39 (verde), remarcando los motivos (rosa) involucrados en el reconocimiento al DNA (café). C) Estructura del dominio C-terminal de IBP39 (verde), unida al dominio C-terminal (rosa) de la RNA polimerasa II (Tomada de Marmorstein, 2003).

Expresión regulada por hierro

Para *T. vaginalis*, el hierro es esencial para el metabolismo y la presencia de este elemento en el medio que habita, determina la expresión de fenotipos múltiples de virulencia, tales como la citoadherencia (Lehker *et al.*, 1991), mecanismo por el cual múltiples proteínas son expresadas en la superficie celular de *T. vaginalis*, permitiendo la adhesión a las células del epitelio vaginal.

Estudios basados en la expresión de genes ante la presencia de hierro en el medio de cultivo, han permitido determinar el papel del gen *ap65-1*, el cual codifica una enzima málica de 65 kDa, involucrada en la citoadherencia del parásito. El gen de *ap65-1* se regula por la unión de proteínas tipo Myb (Myb1 y Myb2) a dos elementos en el promotor: MRE-1/MRE-2r y MRE-2f (Tsai *et al.*, 2002). Adicionalmente, una búsqueda de elementos en *cis* en los promotores de *T. vaginalis*, permitió identificar algunos motivos en diversos promotores: el motivo 1 (TCAYTTYY), cuya secuencia coincide con el elemento Inr, el motivo 3 (DRCSGYTD) y el motivo 5 (GDCCTTYY)

(Smith *et al.*, 2011). El motivo 5 (M5) se encuentra principalmente en genes ribosomales, mientras que los motivos 1 y 3 (M1 y M3), se encuentran en promotores de una variedad de genes. Una proteína nuclear se une específicamente al M3, por lo cual se le ha llamado M3BP (*Motif 3 Binding Protein*) (Figura 5) (Smith *et al.*, 2011). El genoma de *T. vaginalis* codifica cerca de 400 genes con dominios de unión a DNA con un plegamiento Myb (Iyer *et al.*, 2008). A la fecha, no se ha identificado la proteína de unión del elemento M5.

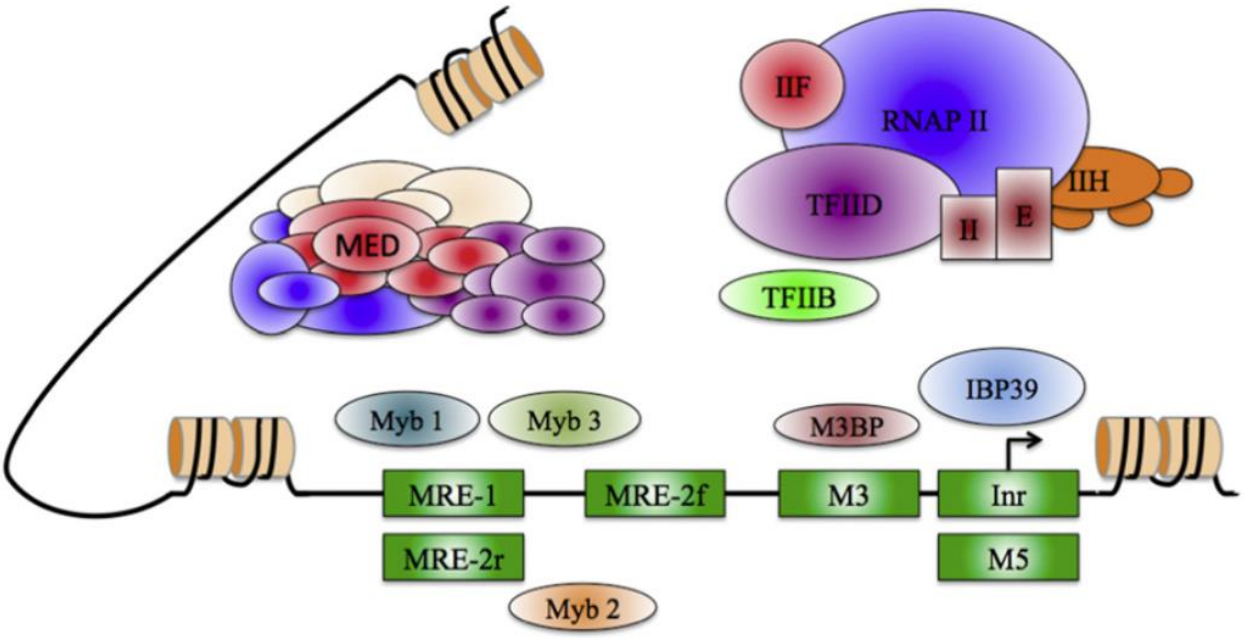


Figura 5.- Elementos del promotor y proteínas de *T. vaginalis* involucradas en la transcripción de la RNA polimerasa II. Los factores de transcripción, los complejos transcripcionales, la posición de los nucleosomas y las secuencias en los promotores se determinaron mediante ensayos bioquímicos y bioinformáticos. Abreviaturas: RNAPII= RNA polimerasa II, IIF= TFIIF, IIH= TFIIH, IIE= TFIIE, MED= Mediador (Tomado de Smith and Johnson, 2011).

OBJETIVO GENERAL

Identificar a las proteínas que conforman a los complejos de inicio de la transcripción en *T. vaginalis.*

OBJETIVOS PARTICULARES

1. Identificar mediante análisis *in silico* a posibles factores transcripcionales que formen parte de los complejos de preinicio de la transcripción en *T. vaginalis.*
2. Analizar las interacciones proteína-proteína y DNA-proteína de los posibles factores transcripcionales identificados.
3. Estudiar la función de algunos de estos posibles factores transcripcionales por complementación de mutantes de levadura.

MATERIALES Y MÉTODOS

Cultivo de *T. vaginalis*

Una alícuota de *T. vaginalis* proveniente del aislado CNCD147 (Alvarez-Sánchez *et al.*, 2000) se cultivó en medio TYM (20 % de Bacto^{TM} triptona, 10 % de extracto de levadura, 5 % de maltosa, 1 % de L-cisteína, 0.8 % de KH_2PO_4, 0.8 % de K_2HPO_4 y 0.2 % de ácido ascórbico) pH 6.2, suplementado con suero de caballo inactivado al 10 %. Los parásitos se crecieron a 37 ºC según se ha descrito previamente (Diamond, 1957; Clark y Diamond, 2002).

Extracción de DNA genómico de *T. vaginalis*

25 ml de parásitos se cultivaron hasta fase logarítmica (entre 2 a 2.5 x 10^6 células/ml aproximadamente). Las células se recolectaron por centrifugación a 3000 rpm, a 4 ºC durante 10 minutos y se lavaron con 20 ml de PBS 1X (137 mM de NaCl, 2.7 mM de KCl, 4.3 mM de Na_2HPO_4, 1.47 mM de KH_2PO_4) frío. Las células se resuspendieron en 4 ml de PBS 1X y se transfirieron a 4 tubos de 1.5 ml. Posteriormente, los tubos se centrifugaron a 14,000 rpm por 1 min para retirar el sobrenadante. A cada tubo se le agregaron 300 µl de amortiguador de extracción (0.2 % de SDS, 100 mM de NaCl, 10 mM de EDTA, 10 mM de Tris-HCl, pH 8) y 300 µl de fenol y se agitaron en el vórtex a velocidad media durante 30 segundos. Posteriormente, los tubos se centrifugaron a máxima velocidad durante 5 min. La fase superior se transfirió a otro tubo de 1.5 ml adicionando 300 µl de fenol y 300 µl de cloroformo. Los tubos se agitaron en vórtex y centrifugaron a la misma velocidad. El sobrenadante se pasó a otro tubo y se le agregó 500 µl de cloroformo, se agitó nuevamente en el vórtex y se centrifugó. El sobrenadante se transfirió a otro tubo para precipitar el DNA con 1 ml de etanol absoluto y 40 µl de acetato de sodio 3 M. El DNA se recolectó por centrifugación a 14,000 rpm durante 15 min. El etanol restante se retiró y el botón se lavó con etanol al 70 % frío. El tubo se centrifugó nuevamente para retirar los restos de etanol. Finalmente, el DNA se resuspendió en amortiguador TE (10 mM de Tris-HCl, pH 8; 0.1 mM de EDTA) y se agregaron 0.5 µl de RNasa (2 mg/ml) y se incubó durante 30 min a 37 ºC. El DNA se conservó a -20 ºC hasta su uso. De esta extracción se obtuvieron 4 alícuotas y se usaron para la obtención de todas las amplificaciones requeridas de *T. vaginalis*.

Modelado de las estructuras terciarias

Los modelos estructurales de las TBPs del capítulo I corresponden a los cristales 1YTB y 4B0A obtenidos de la base de datos RCSB PDB.

Los modelos de las TBPs de protistas corresponden a las secuencias C-terminal deducidas por alineación de las secuencias de *Acanthamoeba castellanii* (AcTBP, P26354), *Entamoeba histolytica* (EhTBP1: P52653, EhTRF1: C4M7H7), *Plasmodium falciparum* (P32086), *Cryptosporidium parvum* (Q6SEL4), *Trichomonas vaginalis* (TvTBP1: A2DQT7, TvTBP2: A2G4N7), *Giardia lamblia* (Q86GG9), *Crypthecodinium cohnii* (Q945B3), *Trypanosoma brucei* (Q387R9), *Trypanosoma cruzi* (TcTBP_E, haplotipo esmeraldo: Q4D013, TcTBP_NE, haplotipo no esmeraldo: Q4DK05), *Leishmania major* (Q4QD71) y *Leishmania tarentolae* (Q2TCV3). Las secuencias se analizaron en el programa HHPred de la plataforma"*Bioinformatics Toolkit*", donde se dedujo que el mejor templado correspondía al cristal 4B0a de *S. cerevisiae*. Los modelos se obtuvieron mediante modelado por homología mediante el programa Modeller incluido en esta plataforma, usando como templado la estructura obtenida por cristalografía de rayos X de *S. cerevisiae* (4B0A)

Todos los modelos mostrados en esta tesis se visualizaron, orientaron y colorearon usando el programa UCSF *Chimera*, excepto los modelos que muestran el potencial electrostático de las TBPs de protistas que se obtuvieron mediante el programa PyMOL.

Obtención de Extractos proteicos de *T. vaginalis*

A partir de 50 ml de cultivo de *T. vaginalis* (~ 2 millones de parásitos) se cosecharon las células por centrifugación a 3,000 rpm durante 10 min a 4 ºC. Las células se lavaron dos veces con PBS 1X frío y se resuspendió en 1 ml de amortiguador de lisis (10 mM de HEPES-KOH, pH 7.9; 1.5 mM de $MgCl_2$, 10 mM de KCl, 0.2 mM de PMSF, 0.5 mM de DTT y 1 X de inhibidores de proteasas). Las células se transfirieron a un homogenizador de vidrio tipo Dounce (Wheaton) y se incubaron en hielo por 30 min). Las células se lisaron con 20 golpes usando el pistilo "*tight*" procurando no generar espuma. El lisado se centrifugó a 14,000 rpm durante 10 min a 4 ºC. El sobrenadante (fracción citoplasmática) se transfirió a un tubo nuevo, mientras que el pellet se resuspendió en 500 µl de amortiguador de extracción (20 mM de HEPES-KOH pH 7.9; 400 mM de NaCl, 1 mM de EDTA pH 8, 1 mM de EGTA pH 8, 1 mM de PMSF, 1 mM de DTT y 1 X de inhibidores de proteasas) y se agitó a velocidad máxima en el vórtex durante 30 min a 4 ºC. La

fracción nuclear se obtuvo mediante centrifugación a 14,000 rpm durante 30 min a 4 ºC. Los extractos se usaron inmediatamente.

Localización celular

Previamente se obtuvieron extractos de T. vagina lis en dos fracciones, la fracción citosólica y la fracción nuclear, las cuales se cargaron en un gel de poliacrilamida al 12 %. El gel se transfirió a una membrana de PVDF. Un inmunblot se realizó utilizando el anticuerpo anti-TvTBP1 con un título 1:500, posteriormente se realizó un inmunoblot usando el anticuerpo anti-H3 1:1,000.

Formación de Dímeros

Cantidades similares de las proteínas recombinantes purificadas 6xHis-TvTBP1 y 6xHis TvTBP2 se trataron con amortiguador de carga sin β-mercaptoetanol (250 mM de Tris-HCl pH 6.8, 20 % de glicerol, 0.4 % de SDS y azul de bromofenol) y con β-mercaptoetanol (250 mM de Tris-HCl pH 6.8, 20 % de glicerol, 0.4 % de SDS, 286 mM de β-mercaptoetanol y azul de bromofenol), cargadas en un gel de poliacrilamida al 12 % con 0.4 % de SDS y corridas en amortiguador de corrida 1X (25 mM de Tris, 192 mM de glicina y 0.1 % de SDS). Adicionalmente, 15 µl de extractos nucleares de *T. vaginalis* se trataron con amortiguador sin β-mercaptoetanol y corridos bajo las mismas condiciones antes mencionadas. El gel se transfirió a una membrana de PDVF y se realizó un inmunoblot tipo *western blot* usando el anticuerpo anti-TvTBP1 con una dilución 1:500 para los extractos de *T. vaginalis* y la sonda HisProbe-HRP con una dilución 1:3,000 para las proteínas recombinantes.

Coprecipitación mediante la etiqueta de GST

A partir de 100 ml de cultivo inducido se obtuvieron los paquetes celulares, los cuales se resuspendieron en 4 ml de amortiguador de resuspensión (20mM de Tris-HCl, pH8; 1 mM de EDTA). La lisis se inició agregando 40 µl de lisozima (1mg/ml) y 0.5 µl de nucleasas benzonasa (250 U). La suspensión se dejó incubando 30 min a 4 ºC. Posteriormente se agregaron 500 µl de amortiguador de unión (PBS 10 X, pH 7.4) y 500 µl de agua bidestilada estéril. La suspensión se incubó 30 min a temperatura ambiente. Los lisados se obtuvieron centrifugando a 10,000 x g a 4 ºC. Los lisados que expresan a la proteína recombinante con la etiqueta de GST (GST-IBP39, GST-TvBRF1, GST-RRN7, GST-TFIIB o GST solo) se incubaron con perlas de glutatión sefarosa por

30 minutos a temperatura ambiente. Seguidos por incubación con el lisado celular que expresa la proteína recombinante con etiqueta de histidinas (6xHis-IBP39, 6xHis-TvTBP1, 6xHis-TvTBP2 o 6xHis-TvBDP1) por 30 minutos a temperatura ambiente. Después de cada incubación, las perlas se lavaron tres veces con 500 µl de PBS 1X, un cuarto lavado se realizó en 50 µl de PBS 1X. Los complejos proteicos se eluyeron de las perlas mediante incubación con 10 mM de glutatión reducido por 10 min. La detección se realizó con el anticuerpo anti-GST y HisProbe-HRP (Pierce).

Ensayo de retardo de la movilidad electroforética (EMSA)

1. Preparación de la sonda

La reacción de amplificación se realizó en 100 µl (10 µl de Amortiguador 10X (-$MgCl_2$), 3 µl de $MgCl_2$ 50mM, 2 µl de dCTP 10 mM, 2 µl de dGTP 10 mM, 2 µl de dTTP 10 mM, 2 µl de dATP 1 mM, 10µCi de $\alpha dATP^{32}$, 200 ng de cada oligonucléotido, 0.5 µl de enzima Taq polimerasa, c.b.p de agua bd). Las condiciones de amplificación fueron las siguientes: desnaturalización inicial: 94ºC por 5 min, 30 ciclos con desnaturalización a 94ºC por 45s, alineamiento a 54ºC (o dependiento de la temperatura de alineamiento requerida) por 30s y extensión a 72ºC por 30s, extensión final a 72ºC por 5 min. La purificación de la sonda se realizó usando una columna empacada con Sephadex G50 en amortiguador TE: se le hicieron pequeños agujeros a un tubo de 0.5 ml en el fondo. Posteriormente, se colocó fibra de vidrio estéril en el fondo del tubo. Se llenó el tubo con Sephadex G50 y se colocó sobre un tubo de 1.5 ml. El tubo se centrifugó a 4000 rpm por 2 minutos y se descartó el líquido. La resina en la columna debe lucir uniforme. La sonda marcada se cargó sobre la columna contenido en un tubo limpio de 1.5 ml y se procedió a centrifugar nuevamente a 4000 rpm por 2 min. Se recuperó la sonda del fondo del tubo y se procedió a cuantificarla. Se usaron aproximadamente 10,000 cpm por reacción de unión.

Las sondas no radioactivas se obtuvieron mediante amplificación por PCR y oligonucleótidos específicos. Todas las sondas se purificaron a través de columnas de sephadex G50 y cuantíficadas. Se usaron de 200 a 500 ng de DNA para cada ensayo.

2.- Preparación de las proteínas

Las proteínas se obtuvieron y cuantificaron el mismo día del ensayo. Las proteínas recombinantes se purificaron por columna de afinidad utilizando perlas de sepharosa recubiertas de glutatión (para proteínas con etiquetas de GST) o níquel (para proteínas con etiqueta de 6xHis).

Para el ensayo del complejo TFIIIB del capítulo III, las proteínas solas o en complejo se obtuvieron por coprecipitación utilizando la columna de afinidad de níquel. La proteína GST-BRF1, se une a la columna de níquel por si sola, por lo que para la obtención de esta proteína en estos ensayos, se usó indistintamente esta columna. Las proteínas se cuantificaron usando una curva de BSA como referencia. Los ensayos se realizaron con aproximadamente 500 ng de proteína

3.-Preparación de la reacción de unión

Concentraciones similares de proteínas recién purificada (se incubaron con la sonda en amortiguador de unión (10 mM de HEPES pH 7.9, 4 mM de $MgCl_2$, 5 mM sulfato de amonio, 8% (v/v) de glicerol, 2 % (w/v) de PEG, 55 mM de KCl, 5 mM de β-mercaptoethanol, 0.2 mM de EDTA y 0.2 mM de PMSF) durante 20 min a temperatura ambiente. Las muestras se cargaron en un gel nativo de acrilamida-bisacrilamida 29:1 al 4.5 % con 2.5 % (v/v) de glicerol. Se usó como amortiguador de corrida TBE al 0.5 X. Los geles se precorrieron durante 30 min antes de cargar las muestras. En algunos ensayos se usó 10,000 cpm de sonda radioactiva marcada por PCR usando [α-^{32}P] ATP y previamente purificada usando sephadex G50, mientras que para otros ensayos, el DNA se observó usando tinción con bromuro de etidio, en estos casos se usaron de 200 a 500 ng DNA.

Generación de las cepas mutantes

1.- Obtención del casete *KANMX*:

El análisis, diseño y generación de las mutantes estuvo supervisado por la técnico de laboratorio (M.C. Karla López Pacheco). Primero se revisó el número de anotación de los genes a estudiar en la base de datos de *S. cerevisiae* (SGD) y se obtuvo la secuencia genómica de los mismos con los extremos 5'UTR y 3'UTR. Los genes de *TFIIB* (*SUA7* / YPR086W), *BRF1* (YGR246C) y *RRN7* (YJL025W) se buscaron en la base de datos "*Saccharomyces Genome Deletion project*" donde se indica que los tres genes son esenciales, motivo por el cual se usó la cepa de levadura diploide BY4743 (*MAT*a/α *his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0*). La mutagénesis de uno de los alelos se realizó por recombinación homóloga. Para ello, se diseñaron los oligonucleótidos requeridos para amplificar la secuencia del casete de *KANMX* que además contenían en sus extremos 5' y 3' las secuencias 5'UTR y 3'UTR del gen a intercambiar. El DNA transformante se obtuvo mediante amplificación

usando como templado el plásmido pFA6a-kanMX4, que contiene el casete de *KANMX* y los oligonucleótidos antes mencionados. Los fragmentos de DNA se corrieron en un gel de agarosa para verificar el tamaño, posteriormente cada fragmento de DNA se purificó y cuantificó.

2.- Transformación de la cepa de levadura

Se inocularon 10 ml de YPD (*Yeast extract, peptone, dextroxe*) con una colonia de la cepa BY4743 y se incubó a 30 ºC con agitación constante durante 12 horas. El cultivo se cosechó por centrifugación a 3,000 rpm durante 5 min. Las células se resuspendieron en 10 ml de agua bidestilada estéril y se cuantificaron por densidad óptica. 50 ml de YPD se inocularon a partir del cultivo anterior, para tener una densidad óptica inicial de 0.1. El cultivo se incubó en agitación constante a 30 ºC hasta alcanzar una densidad óptica de 0.3. Posteriormente, las células se colectaron a 3,000 rpm durante 5 min, el sobrenadante se desechó y las células se resuspendieron en 1 ml de agua y se trasvasaron a un tubo de 1.5 ml. Las células se lavaron con 800 µl de solución TEL (Amortiguador TE 1 X, pH 7.5; 100 mM de acetato de litio). Posteriormente, las células se centrifugaron y se retiró el sobrenadante. Las células se resuspendieron en 200 µl de solución TEL. Se usaron 50 µl de la suspensión por cada evento de transformación. A cada tubo se le adicionó 50 µg de DNA transformante (casete de *KANMX* correspondiente), 10 µl de DNA de esperma de salmón (10 mg/ ml) (previamente hervido durante 5 minutos y enfriado) y 300 µl de la solución TELP (Amortiguador TE 1 X, pH 7.5; 100 mM de acetato de litio, 40 % de PEG 4,000). Las células se mezclaron por agitación suave en vórtex y se incubaron 30 min a 30 ºC en agitación (250 rpm). 40 µl de DMSO (Dimetilsulfóxido) se añadieron justo antes del choque térmico a 42 ºC por 15 minutos. Posteriormente, las células se centrifugaron a 14,000 rpm para retirar el sobrenadante y se resuspendieron en 200 µl de YPD. Las células se platearon en cajas de YPD + geneticina (200 µg/ml) e se incubaron a 30 ºC hasta ver crecimiento. Cada colonia obtenida se estrió 3 veces en cajas de YPD + geneticina. Se obtuvieron 5 colonias de la cepa *brf1*Δ::*KANMX*, 4 colonias de la cepa *tiib*Δ::*KANMX* y 1 de la cepa *rrn7*Δ::*KANMX*. Finalmente, alícuotas de cada cultivo se almacenaron a -70 ºC. Se usó la cepa marcada como 1 en cada caso.

3.- Extracción de DNA genómico de levadura

Una colonia aislada se inoculó en 10 ml de medio YPD (*Yeast extract, peptone, dextrose*) y se incubó a 30 ºC toda la noche. Las células se colectaron por centrifugación a 6,000 rpm por 10 min

y se lavaron en 2 ml de agua bidestilada estéril. Las células se resuspendieron con 200 µl de amortiguador TE (10 mM de Tris-HCl, pH 8; 0.1 mM de EDTA, pH 8) y 40 µl de RNasa A (10 mg/ml) y se incubaron a 50 ºC por 15 min. Posteriormente, las células se recolectaron por centrifugación y se resuspendieron en 200 µl de solución A (2 % de Tritón X-100, 1 % de SDS, 100 mM de NaCl, 10 mM de Tris-HCl, pH 8; 1 mM de Na-EDTA), 200 µl de solución B recién preparada [Fenol: $CHCl_3$: alcohol isoamílico (25:24:1)] y 300 µl de perlas de vidrio (425-600 µm) a cada muestra. Las células se lisaron mediante agitación en vórtex a velocidad máxima durante 4 min. A la suspensión se le adicionó 200 µl de amortiguador TE y se agitó nuevamente a velocidad máxima durante 1 min. La fase acuosa se recuperó mediante centrifugación a 14,000 rpm por 5 min. Se agregaron 400 µl de solución B a la suspensión celular y se agitó en vórtex a máxima velocidad por 30 segundos. La muestra se centrifugó a 14,000 rpm por 5 min para recuperar la fase acuosa. Se agregaron 400 µl de cloroformo y se agitó en vórtex a máxima velocidad por 30 segundos. El DNA se precipitó durante 30 min en 1 ml de etanol absoluto y 10 µl de acetato de amonio 4 M. Se eliminó el sobrenadante de la muestra y se lavó con etanol al 70 %. El DNA se resuspendió en 50 µl agua bidestilada estéril y se guardó a -20 ºC.

4.- Comprobación de la cepa

Para demostrar la integración del casete *KANMX* al genoma de la levadura, se realizó una amplificación mediante PCR usando un oligonucleótido que se ancla en el extremo 5'UTR de cada gen y un oligonucleótido que se ancla de manera interna en el casete de *KANMX* y como templado DNA genómico extraído de cada cepa obtenida. Los tamaños de los amplicones esperados son los siguientes: para la cepa *rrn7*Δ::*KANMX* (614 pb), para la cepa de *tiib*Δ::*KANMX* (589 pb) y para la cepa de *brf1*Δ::*KANMX* (571 pb). Dado que los genes a estudiar son esenciales es necesario que solo uno de los alelos esté mutado, mientras que el otro debe estar íntegro para conservar la viabilidad de la cepa. Para demostrar que sólo uno de los alelos se intercambió por el casete de *KANMX* se realizó otra amplificación usando oligonucleótidos que se anclan en el extremo 5'UTR y el extremo 3'UTR, de manera que se deberán amplificar dos productos, el correspondiente al alelo con *KANMX* y el alelo sin mutar. Los tamaños esperados de los amplicones para cada cepa fueron los siguientes: para la cepa *rrn7*Δ::*KANMX* (alelo sin mutar: 2,121 pb, alelo mutado: 2,122 pb), para la cepa de *tiib*Δ::*KANMX* (alelo sin mutar: 1,693 pb, alelo mutado: 2,201 pb) y para la cepa de *brf1*Δ::*KANMX* (alelo sin mutar: 2,399 pb, alelo mutado: 2,154 pb). Debido a que en algunos casos no se pudieron observar dos diferentes amplicones por diferencia de tamaño para las

cepas *rrn7*Δ::*KANMX* y *brf1*Δ::*KANMX*. El producto obtenido de la amplificación de la cepa *rrn7*Δ::*KANMX* se digirió con la enzima *SpeI* y se corrió un gel de agarosa para visualizar los productos de la digestión, donde se observaron fragmentos de DNA con los siguientes tamaños esperados (1872 pb, 1105 pb, 1016 pb y 250 pb). El producto obtenido de la amplificación de la cepa *brf1*Δ::*KANMX* se digirió con la enzima *NdeI* y se corrió un gel de agarosa para visualizar los productos de la digestión, donde se observaron fragmentos de DNA con los siguientes tamaños esperados (2,154 pb, 1,680 pb y 719 pb). Todos estos datos corroboraron que se obtuvieron las cepas *tiib*Δ::*KANMX*, *brf1*Δ::*KANMX y rrn7*Δ::*KANMX*, con un alelo íntegro y con otro alelo que contiene el casete de *KANMX*.

Inducción de la esporulación de las cepas mutantes

Las cepas mutantes se estriaron en medio rico de pre-esporulación (8 g/l de extracto de levadura, 3 g/l peptona, 100 g/l de dextrosa anhidra, 20g/l de agar) y se incubaron a 30 ºC durante 3 a 5 días. Posteriormente se estriaron en medio mínimo de esporulación (10 g/l de acetato de potasio, 5 mg/l de histidina, 25 mg/l de leucina y 20g/l de agar) y se incubaron a 30 ºC hasta observar esporas. Las cepas alcanzan aproximadamente un 30 % de esporulación.

La disección de las tétradas se realizó mediante un micromanipulador. Las células se recuperaron en medio rico YPD a 30ºC. Posteriormente, las células se crecieron en medio rico YPD en presencia o ausencia de geneticina G418.

Doble Híbrido

Obtención de los plásmidos

Los plásmidos pEG2020-TvTBP1 y pEG202-TvTBP2 se clonaron previamente (M. en C. Lluvia Rosas Hernández), así como el plásmido pEG202-TvBRF1 (M. en C. Karla López Pacheco).

Para el ensayo de doble híbrido, el marco de lectura abierto de TvBRF1 se clonó en el plásmido pJG4-5, el cual expresa a TvBRF1 como una proteína recombinante fusionada al dominio de activación "*acid blob B42*" (AD) bajo el control de un promotor inducible *GAL1*. El plásmido pJG4-5 contiene un gen *TRP1* para la selección de las transformantes. TvBDP1 se clonó en el plásmido pJG4-5 y en el plásmido pEG202, que es un plásmido de expresión para la obtención de proteínas "carnada" que expresa proteínas como una fusión al dominio de unión de DNA de LexA (BD) bajo el control del promotor *ADH1*; este plásmido contiene el marcador de selección de *HIS3*.

Todas las construcciones se verificaron por secuenciación de DNA y se transformaron a la cepa EGY48 usando el protocolo de transformación con acetato de litio. La cepa EGY48 contiene un sistema de reportero el cual le confiere protrofía a leucina cuando el operador de LexA es activado a través de la interacción de dos proteínas (el gen de *LEU2* bajo el operador de LexAop). Adicionalmente, un plásmido reportero se co-transformó. El plásmido pSH18 que contiene el reportero de *lacZ* bajo el operador de LexA insertado en el promotor *GAL1*. Este plásmido contiene al gen *URA3* para la selección de las transformantes.

Obtención de las cepas de levadura transformadas

A partir de un precultivo de la cepa EGY48 (MATα, *his3*, *trp1*, *ura3*, *LeXAop(x6)-LEU2*), se inoculó un cultivo en 50 ml en medio rico YPD y se creció en agitación constante a 30 ºC hasta alcanzar una D.O. entre 0.3 y 0.4 a 600 nm. Alcanzado este punto, el cultivó se trasvasó a un tubo estéril de 50 ml y se centrifugó a 6,000 rpm durante 10 min a 4 ºC. Se retiró el sobrenadante y se lavaron las células dos veces con 10 ml de agua estéril. El botón celular se resuspendió en 1 ml de agua estéril y se transfirió a un tubo de 1.5 ml estéril. Se retiró el líquido excedente mediante centrifugación a 14,000 rpm por 10s. Posteriormente, se adicionó 400 µl de agua, 50 µl de acetato de litio y 50 µl de amortiguador TE y se resuspendieron las células. Se usaron 50 µl de esta suspensión por cada evento de transformación. A cada tubo se le agregó 10 µl de esperma de salmón (10 mg / ml) previamente hervido durante 5 min, 300 µl de PEG al 40 % y 200 ng del DNA transformante (el plásmido reportero pSH18 y el plásmido pJG4-5 o pEG202 según correspondiera). Las células se incubaron por 30 min a 30 ºC. Después de la incubación, las células se incubaron a 42 ºC por 15 min. Se centrifugaron a 14, 000 rpm durante 10 s y se desechó el sobrenadante. Las células se resuspendieron en 180 µl de agua estéril y se platearon en cajas con medio según correspondiera.

Las células transformantes positivas de levadura se seleccionaron en medio mínimo complementado con dextrosa en ausencia de triptófano, uracilo e histidina y leucina (para la selección de los plásmidos pEG202 pJG4-5 y pSH18) o según correspondiera. Se realizaron tres pases de cada cepa transformada.

Goteos

Se inocularon 10 ml de medio de cultivo (según correspondiera) de tres clonas de cada cepa transformada y se dejaron creciendo a 30ºC hasta alcanzar una densidad óptica de 0.6 a 600 nm. A partir de estos cultivos se realizaron diluciones seriadas 1:5, 1:50 y 1:125 y se mantuvieron en hielo todo el tiempo. 3 µl de cada dilución se gotearon en cajas con medio mínimo suplementado con galactosa (para las células transformantes que contienen los tres plásmidos), medio mínimo suplementado con galactosa e histidina (para las transformantes sin el plásmido pEG202) o medio mínimo galactosa y triptófano (para las células transformantes sin el plásmido pJG4-5) y cajas con medio mínimo galactosa suplementado con leucina y X-gal (para las células transformantes que contienen los tres plásmidos), medio mínimo galactosa suplementado con leucina, histidina y X-gal (para las transformantes sin el plásmido pEG202) y medio mínimo galactosa suplementado con leucina, triptófano y X-gal (para las células transformantes sin el plásmido pJG4-5). Las interacciones positivas proteína-proteína se evaluaron a través de dos diferentes condiciones: (1) la expresión de *LEU2* permite que las transformantes crezcan en medio mínimo suplementado con galactosa en ausencia de leucina. (2) la expresión del reportero *lacZ* permite el desarrollo de colonias azules en presencia de 40 µg/ml de 5-bromo-4-cloro-3-indolil-β-D-galactósido (X-Gal). Adicionalmente, la afinidad de la interacción puede ser cuantificada mediante la medición de la actividad de β-galactosidasa.

Ensayo de actividad β-galactosidasa

Los cultivos celulares de cada cepa transformante se crecieron en medio mínimo suplementado con galactosa al 2 % y rafinosa al 0.1 %, en ausencia de triptófano, histidina y uracilo, hasta alcanzar una D.O. de 0.6. Posteriormente las células se recolectaron por centrifugación a 6,000 rpm durante 10 min y se lavaron con agua bidestilada estéril y fría. El pellet de células obtenido se resuspendió en amortiguador de extracción pH 7.5 (100 mM de amortiguador de fosfatos pH 7.5 y 1 mM de EDTA) complementado con 1 M de ditiotreitol (DTT) y se agitaron en vórtex con perlas de vidrio en rondas de un minuto (6 en total), dejando reposar en hielo 1 min entre cada ronda. El sobrenadante (que contiene el extracto proteico) se recuperó por centrifugación. 100 µl de extracto de proteínas se complementó con 900 µl de amortiguador Z pH 7 (60 mM de $Na_2HPO_4.7H_2O$, 40 mM de $NaH_2PO_4.H_2O$, 10 mM de KCl y 1 mM de $MgSO_4.7H_2O$) y se incubaron 5 min a 30 ºC. La reacción se inició mediante la adición de 200 µl de 4 mg/ml 2-nitrofenil β-D-galactopiranosido

(ONPG) y se contó el tiempo hasta que la reacción se detuvo a través de la adición de 500 µl de Na_2CO_3 1 M. La actividad específica se obtuvo con la siguiente fórmula:

$$\text{Actividad\ específica (nmol / min / mg de proteína)} = \frac{(\text{D.O.}_{420})(1.7)}{(0.0045)(\text{Vol. de extracto})(\text{tiempo})}$$

La cuantificación de proteínas se obtuvo mediante el método de Bradford. Se realizaron al menos tres réplicas por clona y los datos de obtenidos se encuentran en el anexo 2.

Southwestern blot

Diluciones con 50 ng, 100 ng, 200 ng y 400 ng de proteína en PBS 1X se adsorbieron sobre una membrana de nitrocelulosa. La membrana se bloquó con leche baja en grasa al 5 % en amortiguador SWB (10 mM de HEPES-KOH, pH 7.9; 50 mM de NaCl, 10 mM de $MgCl_2$, 0.1 mM de EDTA, 1 mM de DTT, 50 µM de $ZnSO_4$ y 0.01 % de Tween-20) durante 1 hora a temperatura ambiente. La membrana se lavó tres veces con bufer SWB durante 10 min. Posteriormente, la membrana se incubó con 2,000,000 cpm de la sonda radioactiva en 10 ml de amortiguador SWB y 0.25 % de BSA durante 4 horas a temperatura ambiente en agitación suave. Finalmente, la membrana se lavó tres veces con 20 ml de amortiguador SWB, 10 min cada vez. La membrana se expuso a una placa radiográfica y se reveló.

Generación del anticuerpo anti-IBP39

El anticuerpo policlonal anti-IBP39 se obtuvo por inmunización de 25 µg de 6xHis-IBP39, previamente purificada por columna de afinidad a 16 ratones hembra de la cepa BALB-c por vía intraperitoneal. Se usaron dos esquemas de inmunización:

Grupo 1 (Ratones 1-8)
Día 0: Preinmune
Día 9: Primera inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de IFA)
Día 23: Segunda inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 34: Tercera inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 69: Cuarta inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 84: Sangrado

Grupo 2 (Ratones 9-16)
Día 0: Preinmune (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de IFA)
Día 16: Primera inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 43: Segunda inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 84: Tercera inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 98: Cuarta inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 112: Quinta inmunización (25 µg de proteína en 50 µl de PBS 1 X + 50 µl de CFA)
Día 126: Sangrado

La sangre colectada se dejó coagular por tres horas, las muestras se guardaron toda la noche a 4 ºC. El suero se obtuvo por centrifugación a 3000 rpm por 30 min a 4 ºC. Las muestras se separaron en alícuotas y se guardaron a -70 ºC hasta su uso. El protocolo se llevó a cabo de acuerdo con las normas del Comité Interno para el Cuidado y Uso de Animales de Laboratorio (CICUAL) del Instituto de Investigaciones Biomédicas y cuenta con el número de aprobación 129. Los ratones se inmunizaron por el personal del bioterio.

CAPÍTULO I: La diversidad de las proteínas TBPs en protistas

Parra-Marín O, López-Pacheco K, Hernández R, López-Villaseñor I. The highly diverse TATA box-binding proteins among protists: A review. Mol Biochem Parasitol. 2020 Aug 6;239:111312. doi: 10.1016/j.molbiopara.2020.111312. Epub ahead of print. PMID: 32771681.

Uno de los mecanismos más estudiados en transcripción es la formación del complejo de inicio de la transcripción por la RNA polimerasa II, que inicia con el reconocimiento del promotor mediado por el complejo TFIID a través de TBP y el reclutamiento secuencial de otros factores generales de transcripción y la RNA polimerasa II de la siguiente manera: TFIIA, TFIIB, TFIIF/RNA polimerasa II, TFIIE y TFIIH (Figura 6). Primero, el complejo de preinicio de la transcripción (PIC) adopta un estado inactivo o "cerrado" que es incapaz de iniciar la transcripción. Posterior a la apertura del DNA y la fosforilación del dominio carboxilo terminal (CTD) de la polimerasa se forma un complejo activo o "abierto", en este paso se sintetizan algunas bases de RNA, permitiendo el escape del promotor, precedido por un proceso de transcripción abortiva, en el cual se sintetizan productos de RNA cortos.

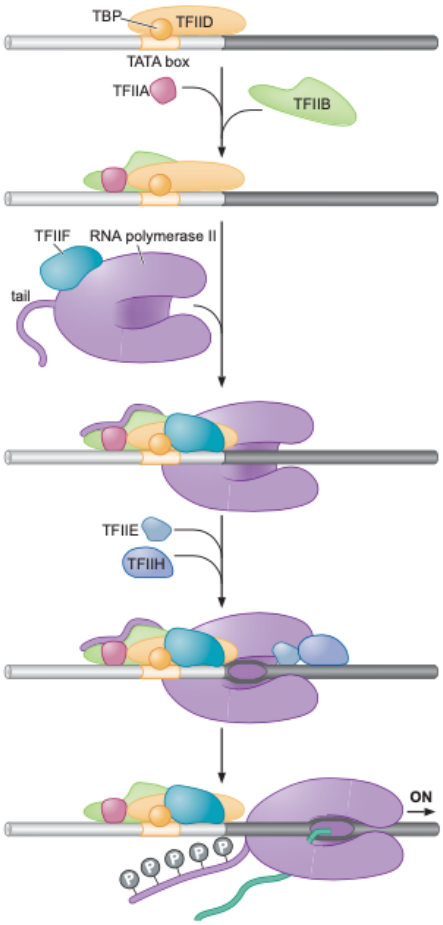


Figura 6.- Modelo del inicio de la transcripción por la RNA polimerasa II. (Watson JD, 2004)

La descripción de este modelo, aunque bueno para entender de manera rápida como funciona la transcripción, es, a su vez, un modelo bastante reduccionista. Con el descubrimiento de TRF1 (*TBP-related factor 1*) y otros factores relacionados a TBP, la variabilidad de los complejos de inicio de la transcripción quedó en evidencia, demostrando que la formación de estos complejos es de hecho más complejo de lo que se describe en los modelos de libros de texto (Duttke, 2015).

Estos datos son sin tomar en cuenta, que la mayoría de los estudios de transcripción se han realizado en un número limitado de organismos modelos: animales, plantas y hongos, dejando en evidencia que aún falta mucho por estudiar y entender.

Los protistas son un grupo de eucariontes microscópicos, unicelulares y con una gran variedad morfológica. La gran mayoría de estos organismos son de divergencia evolutiva temprana (Sogin

y Silberman, 1998), lo cual los hace un grupo de organismos de interés para su estudio debido a la poca información disponible en transcripción.

Hay más de 50,000 especies actualmente clasificadas en seis clados con base en su morfología. La mayoría de los protistas causantes de enfermedades se encuentran clasificados en dos clados: 1) Sarcomastigophora, donde podemos encontrar clasificados a *Acanthamoeba*, *Entamoeba*, *Trichomonas*, *Leishmania*, *Trypanosoma* y *Giardia*, 2) Apicomplexa, donde se encuentra agrupados *Plasmodium* y *Cryptosporidium* (Yaeger et al., 1996). Los protistas están presentes en cada hábitat posible. Algunos de ellos contienen ciclos complejos de vida que involucran a más de un hospedero (por ejemplo, los tripanosomátidos), mientras que otros como *Trichomonas*, tienen un solo huésped.

El objetivo de este capítulo es describir brevemente las particularidades de la TBP en estos organismos. Para este trabajo, se hizo una búsqueda bibliográfica para identificar a los protistas en los que, por lo menos, existiera un artículo que mencionara este factor transcripcional.

Estructura de TBP

La región C-terminal de la proteína TBP de aproximadamente 180 aminoácidos, constituye la región central de la proteína (*core*). A partir de los primeros estudios por cristalografía de rayos X de TBP, se revelaron las características estructurales de esta proteína, las cuales se encuentran altamente conservadas a través de la evolución (Figura 7).

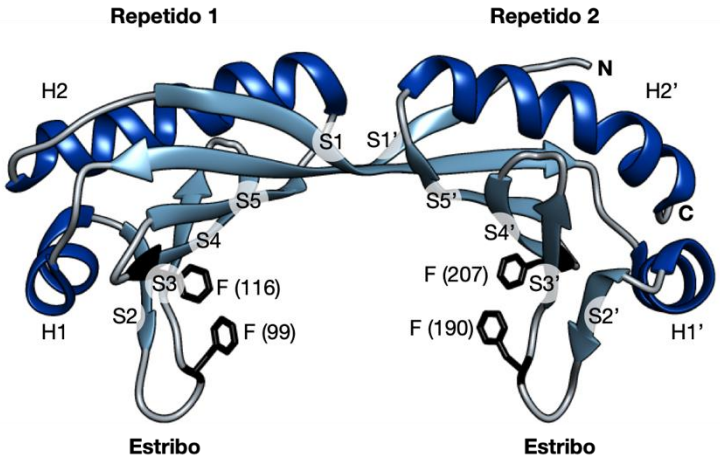


Figura 7.Estructura cristalográfica de TBP de levadura. La estructura obtenida por cristalograífa de rayos X de *S. cerevisiae* (1YTB) se tomó como modelo. Las hojas plegadas β (S1 a S5 y de S1' a S5') se muestran en azul claro y las hélices α (H1-2 y H1'-2') se muestran en azul oscuro. Las fenilalaninas importantes para la unión al DNA se encuentran resaltadas en negro.

El dominio C-terminal de TBP está conformado por dos dominios simétricos constituidos por 5 hojas plegadas β y dos hélices α. La conformación estructural de la TBP genera una superficie cóncava y una superficie convexa (Nikolov *et al.*, 1992). La superficie convexa presenta residuos importantes para la interacción proteína-proteína con factores y reguladores de la transcripción, mientras que la superficie cóncava contiene residuos que permiten la unión al surco menor del DNA. A nivel estructural, las TBPs de arqueas son muy similares a las TBPs de eucariontes, a pesar de que la identidad de secuencia es menor al 40 % (DeDecker *et al.*, 1996).

Obtención de los modelos de TBPs de protistas

Se obtuvieron los modelos de las estructuras terciarias para las TBPs de protistas usando como templado el cristal de TBP de *S. cerevisiae* (4B0A). De manera general, se observa que todas las estructuras tienen la forma típica de silla de montar (Figura 8).

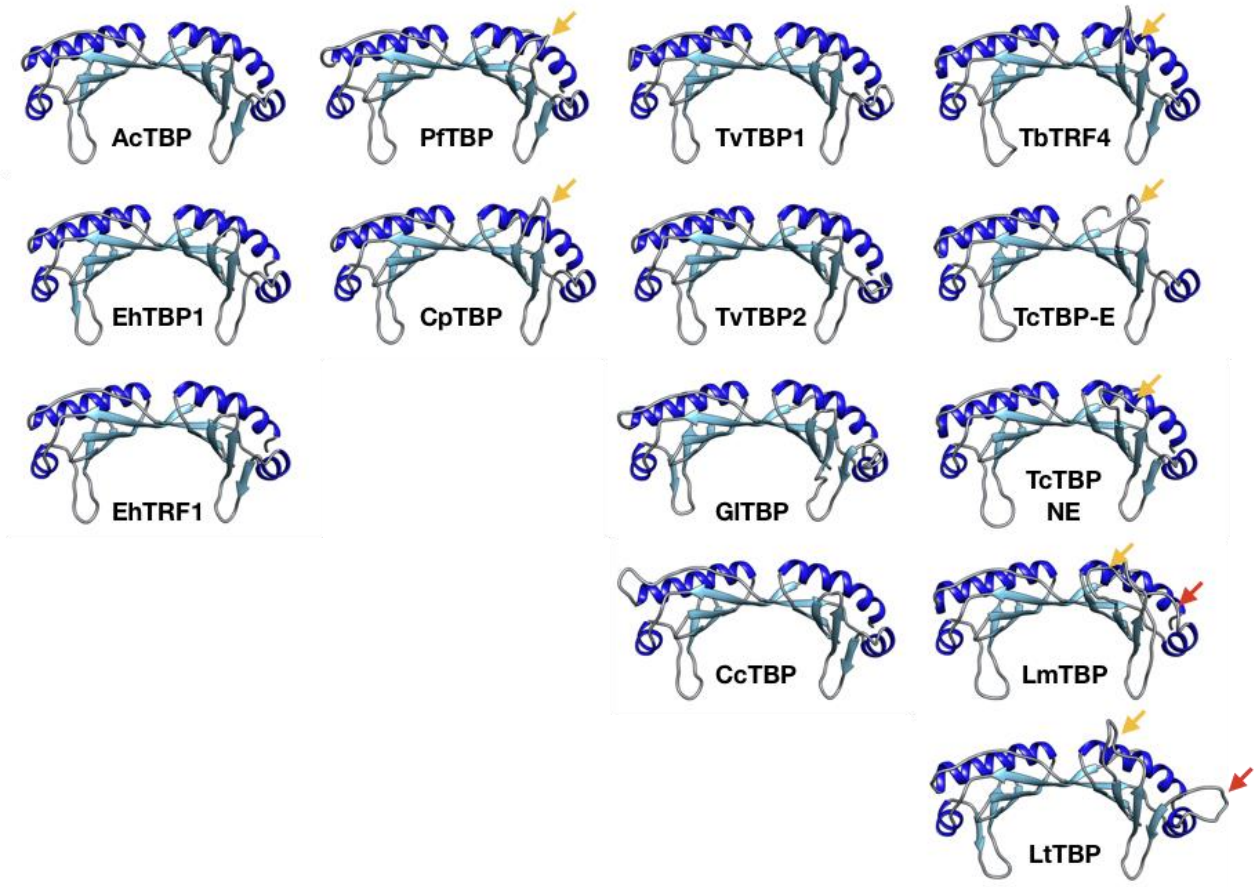


Figura 8.- Modelos de las estructuras terciarias de TBPs de protistas. Las secuencias C-terminal deducidas por alineación de secuencia de *A. castellanii* (AcTBP, P26354), *E. histolytica* (EhTBP1: P52653, EhTRF1: C4M7H7), *P. falciparum* (P32086), *C. parvum* (Q6SEL4), *T.*

vaginalis (TvTBP1: A2DQT7, TvTBP2: A2G4N7), *G. lamblia* (Q86GG9), *C. cohnii* (Q945B3), *T. brucei* (Q387R9), *T. cruzi* (TcTBP_E, haplotipo esmeraldo: Q4D013, TcTBP_NE, haplotipo no esmeraldo: Q4DK05), *L. major* (Q4QD71) y *L. tarentolae* (Q2TCV3) se modelaron con respecto a las secuencias de humano y levadura. Las hélices α, se muestran en color azul oscuro y las hojas plegadas β en azul claro. Las flechas indican inserciones entre las hojas plegadas β3' y β4' (amarillo) y entre la hélice α1' y la hoja plegada β2' (rojo). Los modelos se obtuvieron mediante modelado por homología usando como templado la estructura obtenida por cristalografía de rayos X de *S. cerevisiae* (4B0A) y se visualizaron en el programa *Chimera*.

Sin embargo, si se observa con más detalle, se pueden encontrar algunas diferencias; la más evidente es la ausencia de la hélice α2' en la TBP de *T. cruzi* (alelo esmeraldo) debido a que la secuencia de aminoácidos es más corta en el extremo C-terminal y no contiene la secuencia que corresponde a esta región. Adicionalmente, las TBPs de *Leishmania* presentan dos inserciones: una entre la hélice α1' y la hoja plegada β2' y la segunda entre las hojas plegada β3' y β4'. Mientras que las TBPs de los tripanosomátidos, de *P. falciparum* y *C. parvum* presentan solo una inserción entre las hojas plegadas β3' y β4' y son más cortas que las de *Leishmania*. Los aminoácidos presentes en estas inserciones pueden contener dominios importantes para la interacción con proteínas regulatorias de la transcripción, ya que se encuentran en la superficie convexa de la TBP. Por otro lado, estructuralmente no podemos identificar grandes diferencias, quedando en evidencia que a nivel de estructura terciaria se encuentran muy conservadas.

Mecanismos de unión al DNA

La caja TATA fue el primer elemento descubierto del promotor basal de eucariontes, consiste en un motivo rico en A/T (TATAWAWN), donde W representa a las bases A/T y N es cualquier base. Esta secuencia está flanqueada por una secuencia rica en G/C, que se localiza alrededor de -30 pb río arriba del sitio de inicio de la transcripción (Thomas y Chiang, 2006). Al igual que TBP, el elemento TATA se encuentra presente desde arqueas hasta eucariontes superiores. Sin embargo, la caja TATA está en aproximadamente el 35 % de los genes de humano y entre el 20 al 46 % en levadura (Dikstein, 2011). TBP se une al DNA a través de un mecanismo de ajuste inducido (Figura 9) (Akhtar y Veenstra, 2011). Tras la unión de TBP, el DNA se curva. Esta deformación del DNA es causada por la inserción de dos residuos de fenilalanina (F284 / F301 de humano) al surco menor del DNA entre las bases 1 y 2 de la caja TATA, produciendo un doblez del DNA (complejo DNA-proteína inestable). Posteriormente, la inserción de otro par de fenilalaninas (F193 / F210 de humano) entre las bases 7 y 8 de la caja TATA, genera otro cambio conformacional que regresa al

DNA a su conformación B (complejo DNA-proteína estable). Adicionalmente, otros residuos ubicados en la superficie cóncava de TBP tienen contacto con el DNA, mediante interacciones electrostáticas y fuerzas de Van der Waals.

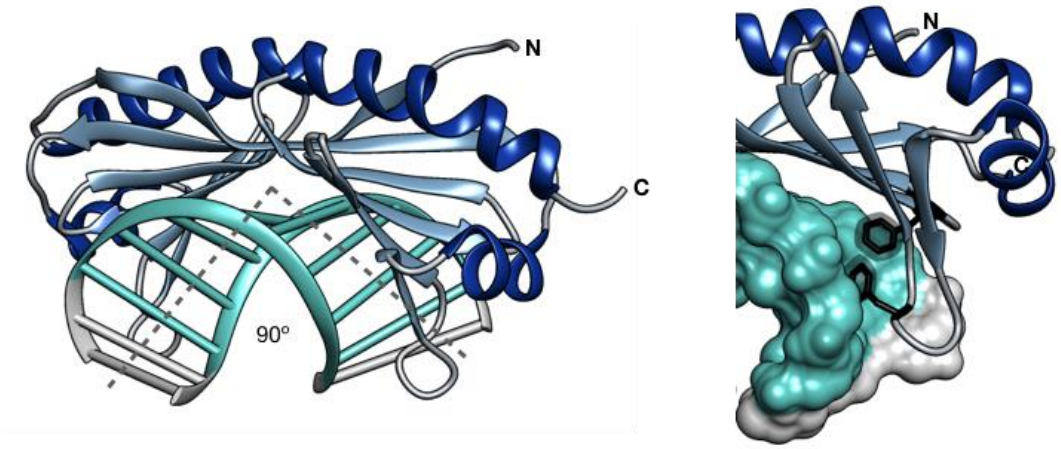


Figura 9.- Unión de TBP al DNA. A) Torsión de 90º del DNA generada por la unión de TBP. B) Intercalación de las fenilalaninas entre la caja TATA (verde) entre las primeras bases de timina y adenina (Akhtar y Veenstra, 2011; Kramm *et al.*, 2019).

Conservación de los aminoácidos involucrados en la unión al DNA de protistas

Mediante un alineamiento de las secuencias de aminoácidos de las TvTBPs y las TBPs de lavadura y otros protistas, observamos que muchos de los residuos involucrados en el reconocimiento del DNA de estos organismos, no estaba conservada. Para hacer más claros estos datos, los hemos puesto en la siguiente Tabla, donde se incluyen los residuos que tienen contacto con los grupos fosfato y las bases de DNA.

Tabla 1 .- Conservación de los aminoácidos involucrados en la unión al DNA en las TBPs de protistas. Se incluye los números de aminoácidos para las secuencias de TBP de levadura y las TBPs de protistas. El color gris oscuro indica identidad y el gris claro similitud. En la parte superior de la tabla se encuentra señala la posición donde se encuentra cada aminoácido: con flechas se indican las hojas plegada β (S1-S5 y S1'-S5'), así como la región conocida como estribo (*stirrup*).

		S1						S3			S4					S5			
						Estribo													
	ScTBP	**Q68**	**N69**	**V71**	**T73**	**R98**	**F99**	**I103**	**R105**	**K110**	**T112**	**L114**	**F116**	**S118**	**K120**	**V122**	**T124**	**G125**
Repetido 1	AcTBP	Q	N	V	T	R	F	I	R	K	T	L	F	S	K	V	T	G
	EhTRF1	Q	N	V	T	R	F	I	R	K	T	L	F	S	K	V	T	G
	EhTBP1	V	N	V	T	R	F	I	R	K	T	L	F	T	K	V	T	G
	EhTBP2	V	N	V	T	R	F	I	R	K	T	L	F	T	K	V	T	G
	TvTBP1	V	N	V	C	R	F	T	R	K	T	L	F	N	K	N	V	G
	TvTBP2	V	N	I	L	R	I	V	R	K	T	L	F	N	K	N	V	G
	CpTBP	Q	N	I	S	K	V	V	R	K	T	L	F	S	R	M	T	G
	PfTBP	H	N	I	S	K	I	I	R	Q	T	L	F	N	R	M	T	G
	LmTBP	V	A	Q	Q	R	I	T	R	T	V	M	H	S	A	S	I	G
	LtTBP	V	A	Q	Q	R	I	T	R	T	V	M	H	S	A	S	I	G
	TcTBP_E	V	G	I	Q	R	T	T	R	T	V	L	R	S	F	S	I	G
	TcTBP_NE	V	G	I	Q	R	T	T	R	T	V	L	R	S	F	S	I	G
	TbTRF4	V	G	I	Q	R	C	T	R	N	V	Q	R	S	L	G	I	G
	CcTBP	F	T	M	K	K	H	T	R	R	T	M	R	S	V	S	T	G
	GlTBP	V	G	N	R	-	Y	R	R	Q	C	S	S	H	H	T	F	G

		S1'				Estribo		S3'			S4'					S5'			
	ScTBP	**Q158**	**N159**	**V161**	**L189**	**F190**	**P191**	**I194**	**R196**	**K201**	**V203**	**L205**	**F207**	**S209**	**K211**	**V213**	**T215**	**G216**
Repetido 2	AcTBP	Q	N	V	L	F	P	I	R	K	V	L	F	S	K	V	T	G
	EhTRF1	Q	N	V	I	F	P	V	R	K	V	L	F	S	K	V	T	G
	EhTBP1	Q	N	V	V	F	P	V	R	K	T	L	F	T	K	V	T	G
	EhTBP2	Q	N	V	V	F	P	V	R	K	T	L	F	T	K	V	T	G
	TvTBP1	A	N	V	I	Y	A	I	R	K	T	M	Y	S	K	I	T	G
	TvTBP2	S	S	V	K	F	A	I	R	Q	T	L	F	S	S	V	T	-
	CpTBP	E	N	I	L	F	P	V	R	K	V	L	F	S	K	I	T	G
	PfTBP	E	N	I	L	F	A	V	R	K	V	L	F	S	K	I	T	G
	LmTBP	R	S	A	R	F	N	V	R	S	S	S	F	T	K	Q	M	G
	LtTBP	R	S	A	R	F	N	V	R	S	S	S	F	T	K	Q	M	G
	TcTBP_E	R	S	T	R	F	S	I	R	Q	S	T	F	T	K	S	L	G
	TcTBP_NE	R	S	T	R	F	S	I	R	Q	S	T	F	T	K	S	L	G
	TbTRF4	R	S	M	R	F	S	I	R	Q	C	T	F	T	K	I	L	G
	CcTBP	C	S	L	L	Y	C	I	R	K	T	L	T	G	K	M	S	G
	GlTBP	V	S	T	I	M	P	Q	V	N	C	S	F	D	Q	T	V	G

Las TBPs de protistas

Amebozoa

Las TBPs analizadas de este *phylum* son las más parecidas a las TBPs canónicas, tanto estructuralmente como a nivel de secuencia. *A. castellanii* codifica para una sola TBP (AcTBP) y es la más similar a HsTBP con 84 % de identidad, incluso mayor que la identidad que se observa entre las TBPs de levadura y humano. Esta proteína tiene todos los aminoácidos conservados para la unión a DNA y en ensayos *in vitro* se une a un elemento similar a una caja TATA canónica y es capaz de reemplazar a la HsTBP en el promotor de adenovirus (Wong *et al.*, 1992).

E. histolytica es una ameba, aunque su posición filogenética no es del todo clara, debido a la ausencia de mitocondrias (Purdy et al., 1996). El genoma de este parásito codifica para tres TBPs (EhTBP1, EhTBP2 y EhTRF1). EhTBP1 y EhTBP2 son idénticas en la secuencia del C-terminal y difieren porque la secuencia del N-terminal de EhTBP1 es más larga. La expresión de EhTBP2 se silencia mediante metilación de la secuencia codificante (Ali *et al.*, 2007). EhTBP1, con 54% de identidad, y EhTRF1, con 52% de identidad con respecto a HsTBP, reconocen a cajas TATA canónicas y variantes de esta (Castañon-Sanchez *et al.*, 2010; Narayanasamy *et al.*, 2018).

Apicomplexa

C. parvum codifica para una sola TBP (CpTBP) con 52 % de identidad con respecto a la TBP de humano. CpTBP se une a una secuencia con similitud a cajas TATA (Millership *et al.*, 2004), lo cual era de esperarse, debido a que tiene conservados la mayoría de los aminoácidos importantes para la unión al DNA. La TBP de *P. falciparum* (PfTBP) tiene solo 38 % de identidad con respecto a la TBP de humano. En EMSAs, PfTBP reconoce una secuencia con mucha similitud a una caja TATA (TGTAA) que se encuentra a -186 pb del sitio de inicio de transcripción (Ruvalcaba-Salazar *et al.*, 2005).

Excavata

T. vaginalis expresa dos TBPs que conservan algunos de los residuos de unión al DNA, pero tienen algunos cambios en las fenilalaninas. TvTBP1 y TvTBP2 tienen 45 y 38% de identidad respectivamente con la TBP de humano. Las TvTBPs no muestran ninguna preferencia de unión secuencia-específica (Parra-Marín *et al.*, 2019). *Giardia lamblia* codifica para un TBP que es 20

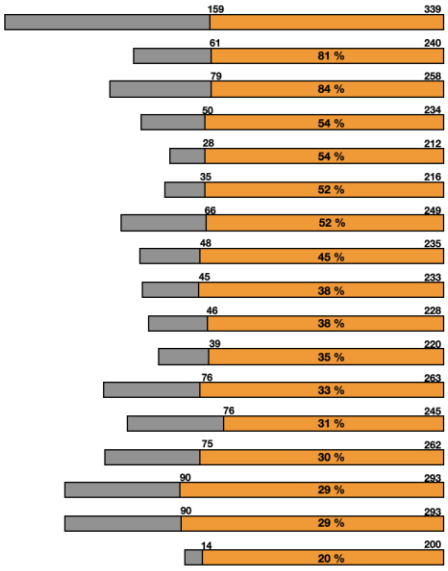
% idéntica a la TBP de humano y no conserva ninguno de los residuos importantes para la unión al DNA; es probablemente la TBP más divergente dentro de los protistas analizados.

Los tripanosomátidos codifican para una sola TBP y no contienen varios de los residuos importantes para la interacción al DNA, incluidas las fenilalaninas. En diversos estudios de inmunoprecipitación de la cromatina (ChIP), las TBPs se encuentran en interacción con los promotores del SL (*splice leader*) (Ruan *et al.*, 2004), donde interaccionan con el complejo SNAPc (*snRNA activiting protein complex*). La TBP es esencial para la transcripción de las tres RNA polimerasas. Sin embargo, los promotores de tripanosomátidos no son regiones ricas en adeninas y timinas (Martínez-Calvillo *et al.*, 2010) y, en su lugar, son regiones ricas en guaninas.

Dinoflagellata

C. cohnii es el único organismo analizado en este trabajo que no es un parásito. Este organismo tiene importancia industrial por la producción de DHA (Ácido docosahexaenoico). La TBP de este organismo (CcTBP) tiene 35 % de identidad con respecto a la TBP de humano. Esta proteína no tiene las fenilalaninas características para el reconocimiento de las cajas TATA y la gran mayoría de los aminoácidos de unión al DNA no se encuentran conservados. Mediante ensayos EMSAs, CcTBP reconoce una secuencia de poli-T con mayor afinidad que una secuencia con una caja TATA conservada (Guillebault *et al.*, 2002).

Los datos recabados muestras que las TBPs de protistas reconocen secuencias de DNA a pesas de poseer diversos cambios en los aminoácidos de unión a DNA (Figura 10). Sin embargo, estas secuencias no son las típicas cajas TATA. Algunas de ellas, como las TvTBPs no reconocen una secuencia específica y probablemente requieran de un factor transcripcional adicional, que promueva el reconocimiento del promotor de manera específica. Curiosamente, los cambios de aminoácidos importantes para la unión al DNA también parecen estar relacionados con la ausencia de cajas TATA canónicas dentro de las secuencias promotoras de estos organismos. Los mecanismos de unión al DNA pueden darse simplemente, por interacciones electrostáticas.



Organismo	Nombre	Unión a DNA	Sequencia	Referencias
Homo sapiens	HsTBP	Sí	Caja TATA	Nikolov et al., 1996
Saccharomyces cerevisiae	ScTBP	Sí	Caja TATA	Kim et al., 1993
Acanthamoeba castellanii	AcTBP	Sí	TATATAA	Wong & Bateman, 1994
Entamoeba histolytica	EhTBP1	Sí	TATTTAAAA y variantes	Castañón-Sánchez et al., 2010; Dios-Bravo et al., 2005, Naravanasamy et al.,
Entamoeba histolytica	EhTBP2	ND	-	-
Entamoeba histolytica	EhTRF1	Sí	TATTTAAAA y variantes	Castañón-Sánchez et al., 2010; Naravanasamy et al 2018
Cryptosporidium parvum	CpTBP	Sí	GCAGAGCATATAAGGTGAGGTAGGA	Millership et al., 2004
Trichomonas vaginalis	TvTBP1	Sí	Promotores rDNA, actina, 5S y U6	Parra-Marín et al., 2019
Trichomonas vaginalis	TvTBP2	Sí	Promotores rDNA, actina, 5S y U6	Parra-Marín et al., 2019
Plasmodium falciparum	PfTBP	Sí	TGTAA	Ruvalcaba-Salazar et al., 2005
Crypthecodinium cohni	CcTBP	Sí	TTTTTTTT	Guillebault et al., 2004
Trypanosoma cruzi	TcTBP-NE	Sí	Secuencia rica en G	Cribb et al., 2010
Trypanosoma cruzi	TcTBP-E	ND	-	-
Trypanosoma brucei	TbTRF4	Sí	Promotor del gen SL RNA (ChIP)	Schimanski et al., 2005; Das et al., 2005
Leishmania major	LmTBP	Sí	Promotor del gen SL RNA, tRNA, snRNA y 5S rRNA (ChIP)	Thoma et al., 2009
Leishmania tarentoloae	LtTBP	Sí	Promotor del gen SL RNA	Thomas et al., 2006
Giardia lamblia	GlTBP	ND	-	-

Figura 10.- Porcentaje de identidad de las secuencias de TBPs de protistas y su unión a DNA. Las secuencias de aminoácidos del C-terminal (naranja) se alinearon para obtener una matriz de identidad con respecto a la TBP de *Homo sapiens*. Las secuencias N-terminal (gris) no fueron alineadas. En la siguiente columna se muestra el nombre de la TBP y el organismo al que pertenece y se muestra la secuencia a la que se unen.

El potencial electrostático

La proteína TBP tiene una superficie cargada positivamente que le permite la interacción con el DNA con carga negativa otorgada por los grupos fosfatos del DNA. La superficie molecular de la HsTBP está cargada positivamente excepto en el segundo estribo (*stirrup)*, debido a la presencia de residuos acídicos (por ejemplo; E280 y E282) (Adachi *et al.*, 2008). Adicionalmente, existen varios residuos de lisina y arginina, que son aminoácidos con carga positiva, capaces de interaccionar mediante fuerzas electrostáticas, al ser el DNA una molécula de carga negativa (conferida por los grupos fosfato del DNA).

En este trabajo, se ha calculado el potencial electrostático para las TBPs de protistas haciendo uso de la herramienta anexa en el programa de PYMOL (Figura 11).

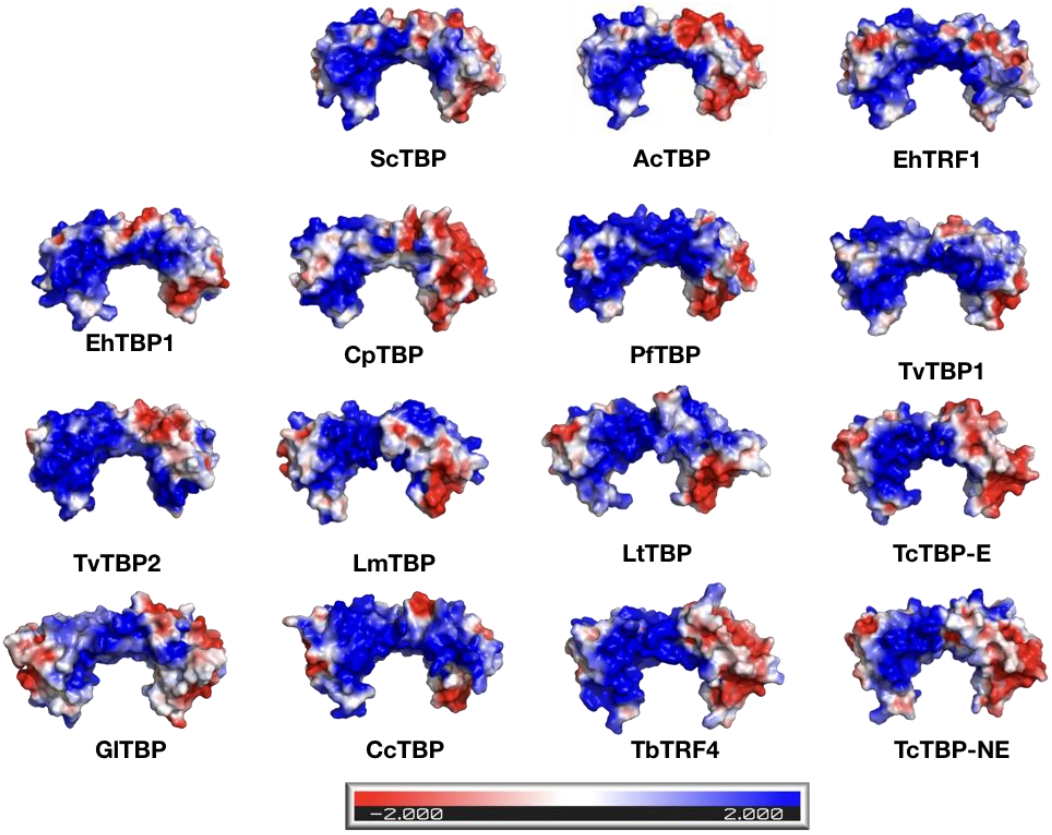


Figura 11.- Potencial electrostático de las TBPS de protistas. El color azul indica la superficie con carga positiva, mientras que el rojo indica carga negativa.

De manera general, se observa que las TBPs conservan una superficie cóncava con una mayor cantidad de residuos positivos que favorecerían la unión al DNA mediante cargas electrostáticas. GlTBP es una de las proteínas que tiene menos residuos cargados positivamente en la superficie cóncava, pero no podemos saber cómo afecta esto a la unión al DNA, ya que no hay estudios sobre esta proteína.

A diferencia de otras proteína de unión a DNA, la interfaz de unión entre la TBP y el DNA parece darse predominantemente por contactos hidrofóbicos y menos por contactos secuencia específica (Coleman *et al.*, 1995).

Factores relacionados a TBP

Hace algunos años, se llamó a TBP el factor de transcripción universal, sin embargo, en años recientes, se ha descrito un grupo de proteínas con similitud a TBP que pueden formar parte de los complejos de inicio de la transcripción de las tres RNA polimerasas. Estas proteínas parecen tener

un papel específico en la regulación de la transcripción: gene-específico o tejido específico, participando en procesos como la gametogénesis, etc. (Figura 12).

Estas proteínas han recibido distintos nombres, pero por sus características, se pueden clasificar en cuatro grupos principales: 1) TRF1, 2) TRF2, 3) TRF3 y 4) TRF4. En *Drosophila* se ha descrito un gen que codifica para una proteína similar a TBP, distinta a las anteriormente descritas; a la cual han llamado como TRF5, pero debido a falta de información no se hablará de ella en este capítulo.

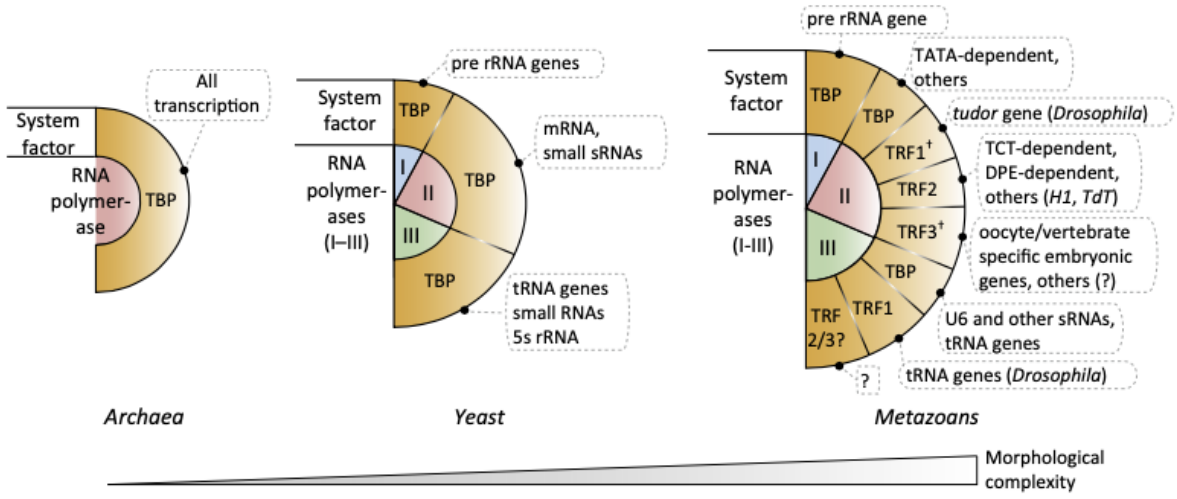


Figura 12.- Diversificación de la maquinaria transcripcional. De derecha a izquierda se muestra la maquinaria basal requerida para la transcripción de arqueas, levaduras y metazoarios, la cual aumenta en complejidad. Arqueas solo requiere de una RNA polimerasa y una TBP; mientras que, en levaduras son tres RNA polimerasas las encargadas de realizar la transcripción y una TBP. En el caso de metazoarios, no solo se requiere tres RNA polimerasas, sino que proteínas relacionadas a TBP también participan en la transcripción de algunos genes (Duttke, 2015).

TRF1

TRF1 fue identificada en *D. melanogaster* (Crowley *et al.*, 1993) y es específica de insectos. Está involucrada en la transcripción de un subconjunto de genes específicos del sistema nervioso central, en las células germinales de machos y en los cromosomas politénicos (Hansen *et al.*, 1997).

TRF1 comparte 63 % de identidad con respecto a las TBP canónicas (Zehavi *et al.*, 2015). En ensayos de transcripción *in vitro*, TRF1 puede unirse cajas TATA, reemplazando eficientemente a TBP (Hansen *et al.*, 1997), esto se explica fácilmente ya que TRF1 tiene conservados los residuos de unión al DNA. Adicionalmente, TRF1 puede asociarse a los factores de transcripción TFIIA y TFIIB (Hansen *et al.*, 1997), además de unirse a Brf1 y ser capaz de mediar la transcripción de promotores transcritos por la RNA polimerasa III (Takada *et al.*, 2000). La inmunodepleción de TRF1 de extractos nucleares, abate la transcripción de diversos transcritos de la RNA polimerasa

III (tRNAs y el 5S) (Takada *et al.*, 2000). Sin embargo, TBP, y no TRF1, es necesario para la transcripción de U6 (Verma *et al.*, 2013). Estos resultados indican que TRF1es una forma especializada de TBP, participando en la transcripción de algunos genes transcritos por la RNA polimerasa III.

TRF2

TRF2, también conocida como TLP (*TBP-like protein*), TLF (*TBP-like factor*), TRP (*TBP-related factor*) y TBPL1 (*TBP-like 1*). TRF2 está ampliamente distribuida en metazoarios de nematodos a vertebrados, incluido *H. sapiens*, pero no se encuentra en plantas ni hongos.

TRF2, a diferencia de las otras TRFs, es la menos conservada con respecto a TBP canónica (~ 40 % de identidad). Adicionalmente, TRF2 carece de los aminoácidos involucrados en la unión al DNA, incluidas tres de las cuatro fenilalaninas. Esta proteína, no se une a cajas TATA (Moore *et al.*, 1999) y no parece tener ninguna actividad de unión secuencia-específica al DNA (Dantonel *et al.*, 1999). Sin embargo, se puede asociar a TFIIA y TFIIB (Moore *et al.*, 1999; Rabenstein *et al.*, 1999). TRF2 está involucrada en la transcripción de la RNA polimerasa II y se ha sugerido que TRF2 puede nuclear el PIC en genes sin cajas TATA. Consistente con esta hipótesis, en *Drosophila* TRF2 se encuentra asociada en promotores *TATA-less* transcritos por la RNA polimerasa II (Zehavi *et al.*, 2015).

TRF3

TRF3, también conocida como TBP2(*TATA-binding protein 2*) o TBPL2 (*TBP-like protein 2*), es específica de vertebrados y es la más cercana en identidad a TBP canónica (95 %). El C-terminal está altamente conservado, incluyendo los aminoácidos involucrados en la unión a DNA. TFIIA y TFIIB (Persengiev *et al.*, 2003), y puede unirse a cajas TATA e interactuar con TFIIA y TFIIB (Bártfai *et al.*, 2004). Los datos disponibles para TRF3 sugieren que es el resultado de un evento de duplicación génica (Persengiev *et al.*, 2003; Bártfai *et al.*, 2004). TRF3 es predominantemente nuclear, sin embargo, durante la citocinesis se localiza en el citoplasma (Persengiev *et al.*, 2003). TRF3 se requiere para el desarrollo embrionario de *Zebrafish* y *Xenopus* (Jallow *et al.*, 2004; Bártfai *et al.*, 2004).

TRF4

TRF4 fue descrita recientemente en *D. melanogaster* y se encuentra conservada entre las especies de *Drosophila*.

TRF4 es una proteína que se localiza en el citoplasma y se concentra en el retículo endoplásmico, donde interacciona con TER94, una proteína del retículo endoplásmico que participa en la maquinaria de degradación (Kurshakova *et al.*, 2019). TRF4 es una proteína con una nueva función, tras la pérdida de la señal de localización nuclear (Kurshakova *et al.*, 2019).

Conservación evolutiva

Algunos artículos han llamado a las TBPs de protistas como factores relacionados a TBP, debido a que poseen características intermedias entre TBP y TRF (Guillebault *et al.*, 2002) o porque no encajaban en la descripción de las TRFs descritas al momento de su estudio (Ruan *et al.*, 2004). Entre estas características se encuentran una menor conservación del C-terminal y cambios importantes en los residuos de unión al DNA, incluidas la ausencia de las fenilalaninas importantes para el reconocimiento de las cajas TATA. Tal como se ha descrito aquí, las TBPs de protistas no pueden ser clasificadas como TBPs canónicas ni TRFs (según la clasificación presentada hasta ahora).

En un estudio en el cual la diversidad de secuencias de TBPs se analizó entre diversas especies del reino Eucaria, se observó que algunos residuos variantes parecen estar conservados desde el ancestro común de las TBPs de eucariontes, excepto en protistas, demostrando su gran diversidad biológica (Kawakami *et al.*, 2017). Curiosamente algunos protistas como *E. histolytica* y *T. vaginalis* (organismos unicelulares) expresan más de una TBP, sin embargo, a la fecha se desconoce si tienen una función diferencial.

Por otro lado, llama la atención que la baja conservación de aminoácidos involucrados en la unión al DNA está directamente relacionada con la ausencia de cajas TATA canónicas en los promotores de estos organismos. Para evidenciar este hecho, se realizó el siguiente gráfico, ilustrando los residuos importantes para la unión al DNA conservados en azul y los residuos no conservados en rojo, junto con árbol filogenético de las TBPs (Figura 13). Se puede observar que los protistas que no presentan cajas TATA canónicas en sus promotores y tienen un mayor número de modificaciones en estos residuos. Todos estos datos, sugieren que las TBPs de protistas y eucariontes superiores divergieron en distintas vías.

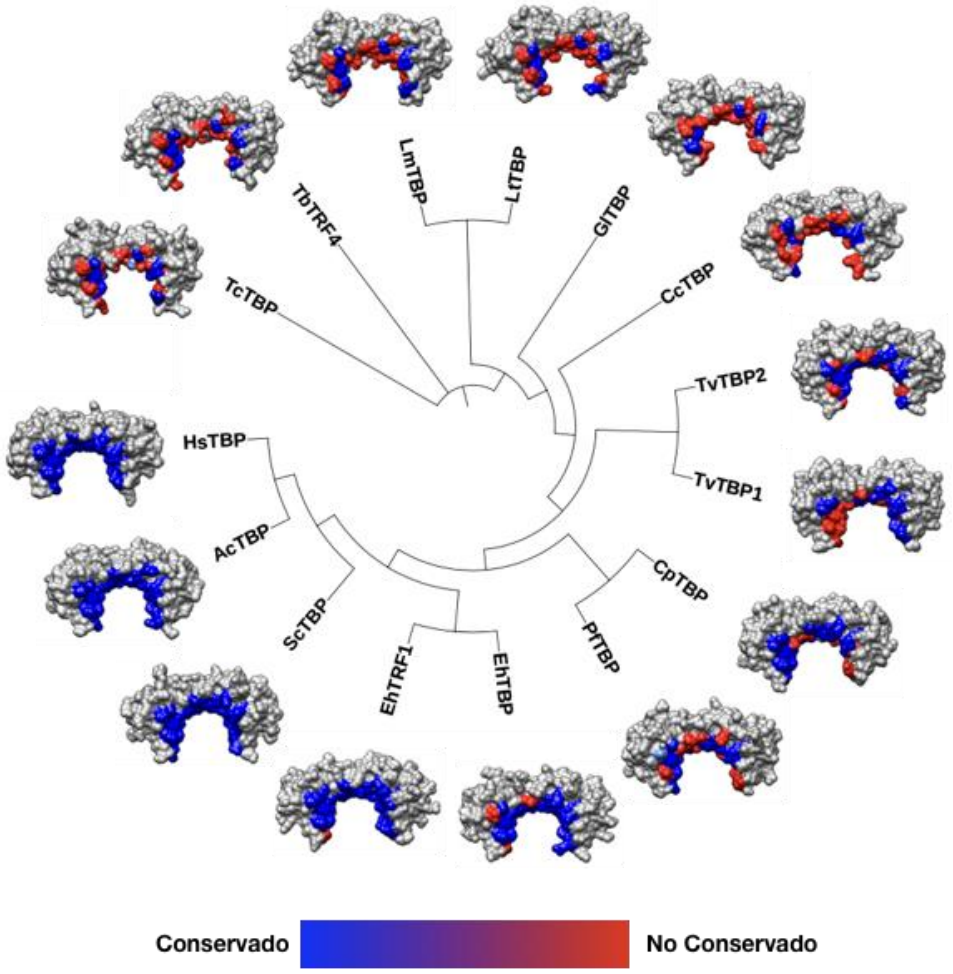


Figura 13.- Conservación evolutiva de los aminoácidos involucrados en la unión de las TBPs. La conservación de aminoácidos involucrados en la unión al DNA con respecto a dos TBPs canónicas (HsTBP, 1CDW y ScTBP, 4b0A) se muestra en azul y los aminoácidos no conservados en rojo. El cristal de TBP de levadura se usó como templado para obtener lo modelos de las TBPS de protistas y se visualizaron con *Chimera*. El árbol filogenético se obtuvo mediante el programa de ANCESCON.

CAPÍTULO II: Las TvTBPS como parte de la maquinaria transcripcional de *T. vaginalis*

Parra-Marín, O., Rosas-Hernández, L., López-Pacheco, K., Franco, B., Ibáñez-Escribano, A., Hernández, R., & López-Villaseñor, I. (2019). An in vitro characterisation of the *Trichomonas vaginalis* TATA box-binding proteins (TBPs). *Parasitology Research*, *118*(10), 3019–3031. https://doi.org/10.1007/s00436-019-06438-z)

Antecedentes

Durante el trabajo de tesis doctoral de Bernardo Franco en nuestro laboratorio se estudiaron los promotores reconocidos por la RNA polimerasa I, así como los sitios de terminación de la transcripción y se realizó una búsqueda bioinformática en la base de datos del genoma de *T. vaginalis* (TrichDB) para identificar algunos componentes de la maquinaria transcripcional de la RNA polimerasa I, entre ellos a TBP. De esta manera, se logró identificar a dos genes que codifican para proteínas con similitud a TBP. Estas proteínas tienen un 36.4 % de similitud y 39.75 % de identidad entre ellas, si se compara la secuencia completa (Franco Bárcenas, 2012), mientras que si se compara únicamente el extremo C-terminal (más conservado), el porcentaje de identidad es del 45.16 %.

TBP forma parte del complejo de inicio de la transcripción por las tres RNA polimerasas; sin embargo, solo algunos promotores transcritos por la RNA polimerasa II contienen cajas TATA.

Este dato llamó la atención debido a que los promotores transcritos por la RNA polimerasa II de *T. vaginalis*, carecen de cajas TATA (Liston *et al.*, 1999).

Los marcos abiertos de lectura se clonaron en un vector de expresión para obtener proteínas recombinantes fusionadas a una etiqueta de histidinas (Franco Bárcenas, 2012; M. en C. Lluvia Rosas Hernández, tesis doctoral no publicada). A estas proteínas se les llamó TvTBP1 (TVAG_291560) y TvTBP2 (TVAG_285070) debido a su similitud con la secuencia de TBP de levadura (Spt15). La expresión de ambos genes fue corroborada a nivel de mRNA mediante RT-PCR (por sus siglas en inglés de *reverse transcription polymerase chain reaction*) (M. en C. Lluvia Rosas Hernández, tesis doctoral no publicada). Adicionalmente, se llevó a cabo un ensayo de complementación de una mutante de levadura (*spt15*Δ), donde se observó que TvTBP1 y TvTBP2 no complementan (M. en C. Lluvia Rosas Hernández, tesis doctoral no publicada; M. en C. Karla López Pacheco). Por lo que parte de este trabajo se enfocó a estudiar algunas de las características de ambas TvTBPs.

Resultados:

Similitud de secuencia y estructura terciaria de las TvTBPs

Estructuralmente, TBP se conforma de dos dominios: 1) el N-terminal, no conservado en tamaño y secuencia entre especies y 2) el C-terminal, con un alto grado de conservación tanto, a nivel de secuencia, como de estructura terciaria (Hernandez, 1993). El N-terminal es una región desordenada y no se tiene una estructura cristalográfica de esta región, mientras que el C-terminal consiste en una estructura bipartita formada por dos mitades cuasi-simétricas que adoptan una estructura típica de silla de montar. Cada mitad se compone de 5 hojas plegadas β antiparalelas y dos hélices α.

La estructura terciara del C-terminal de las proteínas TvTBP1 y TvTBP2 se obtuvieron mediante modelado por homología de secuencia, usando como templado un cristal de TBP de *S. cerevisiae* (4b0a, cadena A) (Figura 14). Adicionalmente, la calidad de los modelos se verificó mediante el programa Verify 3D (Bowie *et al.*, 1991; Lüthy *et al.*, 1992).

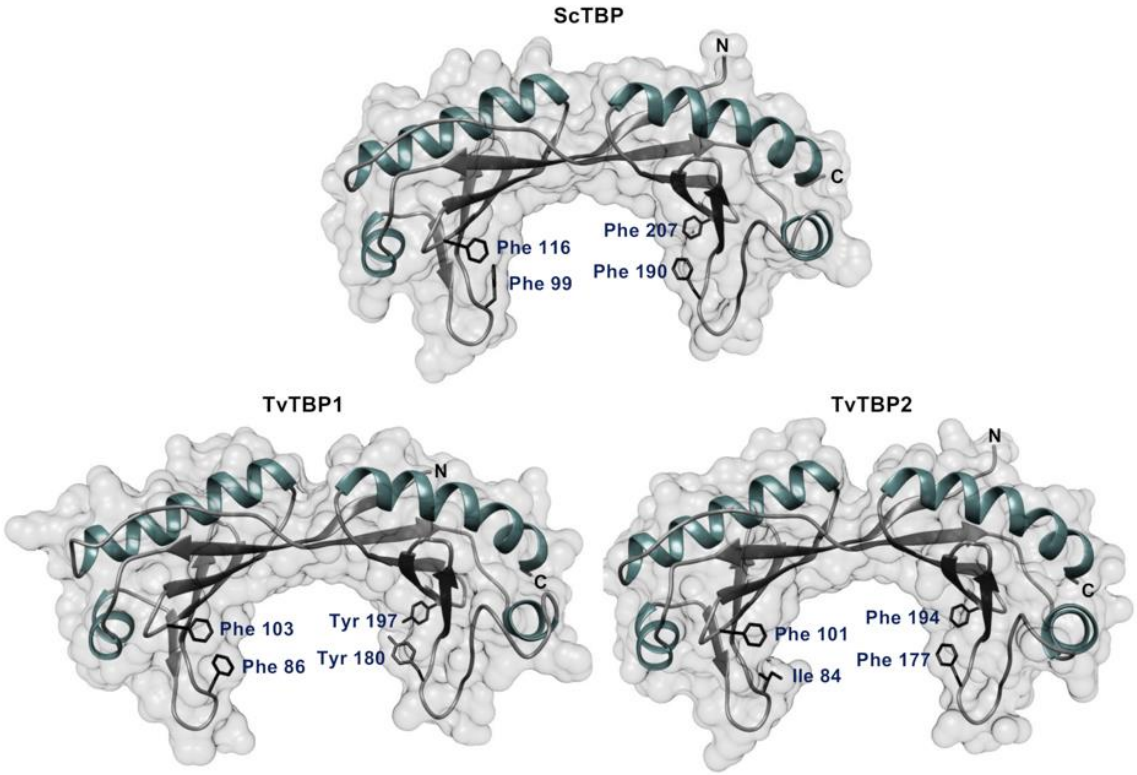


Figura 14.- Modelos de la estructura terciaria de las TvTBPs. Las secuencias C-terminal de ambas TvTBPs se modelaron por homología usando como templado la estructura obtenida por cristalografía de rayos X de *S. cerevisiae* (4b0a). Las hélices α se muestran en verde y las hojas plegadas β en gris oscuro. Los residuos de fenilalanina importantes para el reconocimiento del

DNA se encuentran indicados y numerados de acuerdo con la secuencia. La superficie molecular se muestra en gris claro. Los modelos se obtuvieron con el programa de Modeller y visualizados en *Chimera*.

Los modelos de las estructuras obtenidas para las TvTBPs adoptan la típica estructura de silla de montar, demostrando que las TvTBPs tienen características similares a las TBPs canónicas. Adicionalmente, en la superficie cóncava de las TBPs canónicas se encuentran dos pares de fenilalaninas, las cuales se intercalan en el surco menor del DNA entre el primero y los dos últimos pares de bases de la caja TATA. La interacción de estos residuos con el DNA genera una torsión de aproximadamente 90º. Sin embargo, en las TvTBPs estos residuos no se encuentran conservados. En TvTBP1 dos de estos aminoácidos son tirosinas (Tyr180, Tyr197) y en TvTBP2 hay un cambio por isoleucina (Ile84).

Adicionalmente, las secuencias de aminoácidos de las TvTBPs se compararon con secuencias de TBPs canónicas (*H. sapiens* y *S. cerevisiae*) y TBPs de otros protistas (*E. histolytica*, *G. lamblia*, *Trypanosoma cruzi*, *T. brucei* y *Leishmania major*). La estructura secundaria se muestra en el alineamiento con base en la TBP de humano (HsTBP), aunque se encuentra bastante conservada en las TBPs de protistas. Adicionalmente, los aminoácidos importantes para el reconocimiento del DNA se muestran con asteriscos negros y solo las fenilalaninas se encuentran señalados (asterisco rojo) (Figura 15). Nótese que estos residuos se encuentran principalmente en las hojas plegadas β debido a que estas son las que conforman la superficie de unión a DNA. Curiosamente, los protistas no tienen todos los aminoácidos importantes para la unión al DNA conservados.

Mediante una matriz de identidad obtenida únicamente con las secuencias del C-terminal, sabemos que TvTBP1 y TvTBP2 comparten 45.16 % de identidad entre ellas y que presentan ligeramente mayor conservación con respecto a la secuencia de ScTBP (46.67 % y 40.22 %, respectivamente) que con HsTBP (44.75 % y 37.78 %, respectivamente). La identidad de secuencia entre las TBPs de *Trichomonas* y otros protistas es del 22.35 % al 46.67 %.

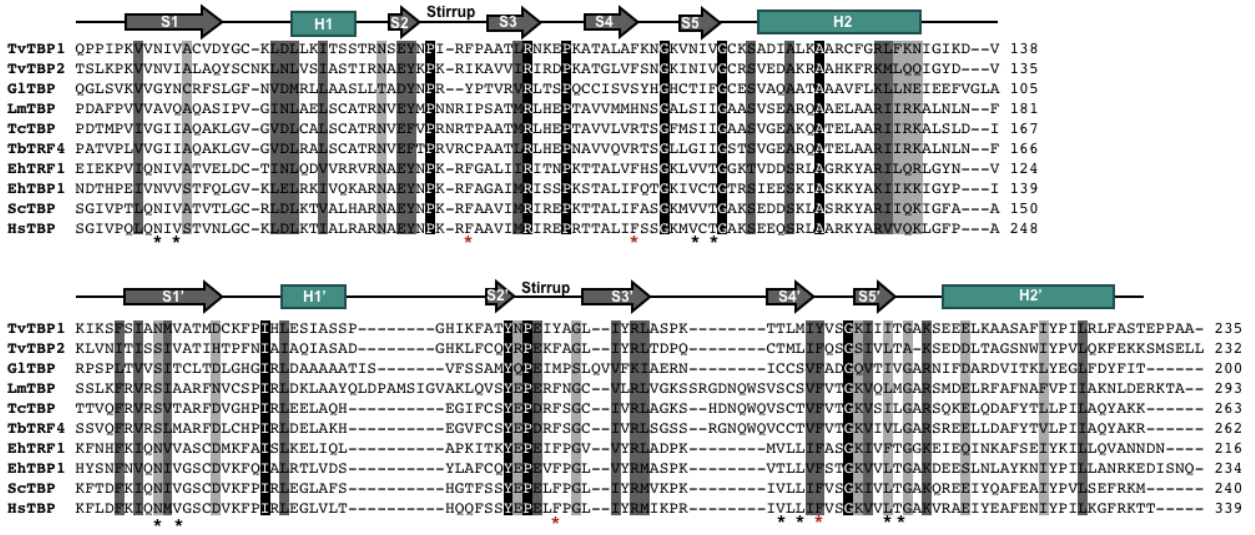


Figura 15. Alineamiento de secuencias de las TBPs. Las secuencias de la región C-terminal de las TBPs de *T. vaginalis* (TvTBP1, A2DQT7) y (TvTBP2, A2G4N7), *G. lamblia* (GlTBP, Q86GG9), *L. major* (LmTBP, Q4QD71), *T. cruzi* (TcTBP, Q4DK05), *T. brucei* (TbTRF4, Q387R9), *E. histolytica* (EhTBP1, P52653) y (EhTRF1, A7UFC2), *S. cerevisiae* (ScTBP, P13393) y *H. sapiens* (HsTBP, P20226) se alinearon usando el programa Clustal Omega. Los números de acceso corresponden a la base de datos de Uniprot. La predicción de la estructura secundaria se basa en la TBP de humano (H, hélices alfa; S, hojas beta plegada). Los residuos de fenilalanina (asteriscos rojos) y otros aminoácidos (asteriscos negros) involucrados en la unión al DNA están indicados.

Localización celular de TvTBP1

En el laboratorio se contaba con un anticuerpo que se obtuvo inmunizando la proteína recombinante 6xHis-TvTBP1 (M. en C. Lluvia Rosas Hernández, tesis doctoral no publicada), y que no había sido probado previamente. Por lo tanto, primero se determinó si el anticuerpo reconocía a la proteína TvTBP1 y si presentaba reacción cruzada con TvTBP2 debido al alto grado de similitud en secuencia y estructura entre ambas. Una cantidad similar de las proteínas recombinantes 6xHis-TvTBP1 y 6xHis-TvTBP2 se resolvieron en un gel de poliacrilamida y se transfirieron a una membrana de nitrocelulosa. Posteriormente, se realizó una inmunodetección tipo *western blot* usando el anticuerpo anti-TvTBP1 con un título de 1:10,000. Bajo las condiciones probadas, el anticuerpo reconoce a TvTBP1, pero no a TvTBP2 (Figura 16A). Posteriormente, para determinar la localización celular de TvTBP1, se obtuvieron extractos de proteínas de *T. vaginalis* en dos fracciones, una fracción citoplásmica y una fracción con enriquecimiento nuclear y se realizó una inmunodetección usando el anticuerpo anti-TvTBP1 con un título de 1:500 (Figura 16B). Una señal con el peso molecular esperado para TvTBP1, aparece en la fracción nuclear, pero no en el

citoplasma. Para corroborar que la señal es nuclear, se realizó una inmunodetección utilizando un anticuerpo anti-H3 que reconoce a la histona H3 (marcador nuclear) y se detectó una señal correspondiente con el peso molecular de la histona H3, en la fracción nuclear pero no en la fracción citoplásmica. Estos resultados demuestran que al menos TvTBP1 se expresa a nivel de proteína y su localización es nuclear.

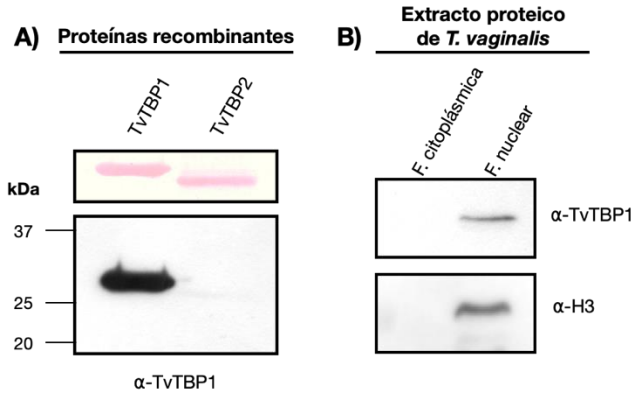


Figura 16. TvTBP1 es una proteína nuclear. A) Inmunodetección tipo *western blot* utilizando un anticuerpo anti-TvTBP1 donde se observa que el anticuerpo reconoce a TvTBP1, pero no a TvTBP2. **B)** TvTBP1 es detectada en la fracción de proteínas nucleares. Un anticuerpo anti-histona H3 se usó como marcador nuclear.

Formación de dímeros de TBP

Los estudios bioquímicos y de cristalografía de rayos X revelaron que TBP se une al DNA como un monómero (Nikolov *et al.*, 1992). Sin embargo, en ausencia de DNA, TBP dimeriza a través del dominio de unión a DNA. Los dímeros de TBP son bastante estables en solución (Coleman *et al.*, 1999), sugiriendo que la dimerización es un mecanismo para la autorregulación negativa de su actividad de unión al DNA. Los dímeros de TBP existen en condiciones fisiológicas (Coleman *et al.*, 1995). *In vivo*, la dimerización de TBP previene la expresión génica no regulada y su degradación (Jackson-Fisher *et al.*, 1999).

Para probar si las TvTBPs forman dímeros, se obtuvieron las proteínas recombinantes 6xHis-TvTBP1 y 6xHis-TvTBP2 mediante purificación en columna de níquel y se realizó una electroforesis utilizando dos condiciones distintas en el amortiguador de carga: en presencia o ausencia de β-mercaptoetanol (+BME y -BME, respectivamente). Este compuesto es un agente reductor que desnaturaliza a las proteínas y es muy utilizado para disociar proteínas en subunidades individuales. Existen reportes de proteínas que no se disocian bajo algunas condiciones desnaturalizantes (Grigorian *et al.*, 2005) y de proteínas que se disocian en presencia de BME a pesar de que sus secuencias no contienen cisteínas (Einhoff *et al.*, 1987). Las muestras se cargaron en un gel desnaturalizante (0.4 % de SDS), se transfirió a una membrana de nitrocelulosa y se realizó una inmunodetección usando la sonda HisProbe-HRP para revelar la presencia de las proteínas 6xHis-TvTBPs.

Tal como se observa en la imagen (Figura 17A), bajo ambas condiciones aparece una señal arriba de 25 kDa que corresponde con el tamaño del monómero de 6xHis-TvTBP1 y 6xHis-TvTBP2 (ambas con peso aproximado de 27 kDa). En el carril sin BME aparece una señal arriba de 50 kDa que corresponde a un dímero de TvTBP (con un peso estimado en 54 kDa), esta señal desaparece ante la presencia de β-mercaptoetanol, sugiriendo la formación de dímeros *in vitro*. Para demostrar que la señal del dímero no es un artefacto, se usó como control negativo un extracto total de proteínas de *Escherichia coli* transformado con el vector pQE30 vacío que es donde se clonaron los genes que codifican para las proteínas recombinantes 6xHis-TvTBPs y no hay reconocimiento por la sonda en estos carriles, confirmando que la señal detectada corresponde a dímeros de TvTBP.

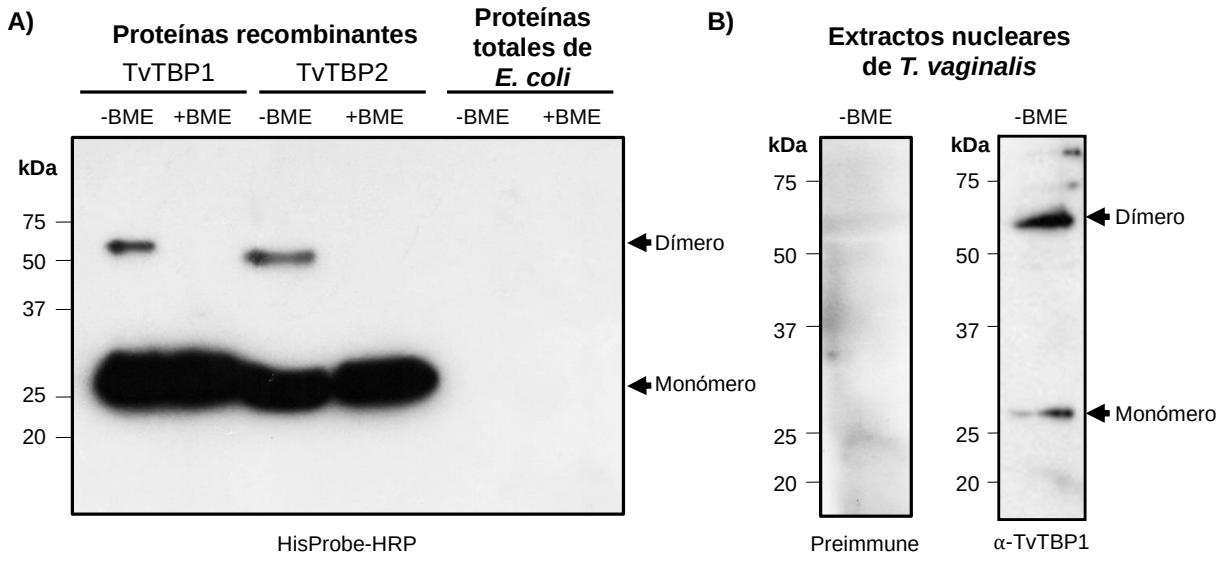


Figura 17.- Las TvTBPs forman dímeros *in vitro* e *in vivo*. A) Inmunodetección tipo *western blot* revelado con HisProbe-HRP que reconoce a las proteínas recombinantes 6xHis-TvTBP1 y 6xHis-TvTBP2. Un extracto total de proteínas de *E. coli* transformadas con pQE30 se usó como control negativo. Las muestras se trataron bajo dos condiciones: -BME, muestras sin β-mercaptoetanol en el amortiguador de carga; +BME, muestras con 1X de β-mercaptoetanol en el amortiguador de carga. **B)** Inmunodetección anti-TvTBP1 y preinmune de extractos nucleares de *T. vaginalis* tratados sin β-mercaptoetanol.

Para probar la formación de dímeros de TvTBP *in vivo*, se tomaron extractos nucleares de *T. vaginalis* tratados sin β-mercaptoetanol y se analizaron por inmunodetección utilizando el anticuerpo anti-TvTBP1 (Figura 17B). Se observaron dos señales, una que concuerda en tamaño con el monómero de TvTBP1 (~26 kDa) y la segunda con el dímero (~52 kDa). De manera interesante, en los extractos nucleares la señal de dímero es más abundante que la del monómero a diferencia de los dímeros observados *in vitro*. Estos experimentos demuestran que TvTBP1 y TvTBP2 son capaces de formar dímeros *in vitro* y que TvTBP1 forma dímeros *in vivo*.

En levadura, la disociación de dímeros de TBP se favorece por la presencia de factores como TFIIA o Brf1, aumentando la disponibilidad de TBP para llevar a cabo la transcripción (Coleman *et al.*, 1999; Alexander *et al.*, 2004). En *T. vaginalis* no se ha encontrado al ortólogo de TFIIA, pero si de TvBRF1, por lo que sería interesante probar si los dímeros de TvTBP disminuyen *in vitro* en presencia de esta proteína.

Unión de las TvTBPs a IBP39

La proteína IBP39 es responsable del reconocimiento del sitio de inicio de la transcripción en los promotores transcritos por la RNA polimerasa II en *T. vaginalis*. Por lo tanto, la interacción entre esta proteína y alguna de las TvTBPs nos puede sugerir que las TvTBPs son parte de este complejo transcripcional. Para probar si TvTBP1 y TvTBP2 pueden interactuar con IBP39, se realizó un ensayo de coprecipitación en columnas de afinidad. Para este ensayo se usaron las siguientes construcciones: 1) IBP39 fusionada con la proteína GST (Glutatión-S-transferasa) en el extremo N-terminal, la cual funciona como carnada y 2) TvTBP1 y TvTBP2 fusionadas con una etiqueta de seis histidinas (6xHis) en el extremo N-terminal, las cuales funcionan como presas.

Para el ensayo de coprecipitación se obtuvieron los lisados proteicos de *E. coli* que expresan las proteínas GST-IBP39 y GST. Por separado, cada lisado se incubó con perlas de sefarosa unida a glutatión. Posteriormente, las perlas se lavaron para eliminar las proteínas no unidas a la resina.

En un segundo paso, las perlas se incubaron con los lisados proteicos de *E. coli* que expresan las proteínas recombinantes 6xHis-TvTBP1 (1 volumen) y 6xHis-TvTBP2 (2 volúmenes) de manera independiente. Después de la incubación, las perlas se lavaron y finalmente las proteínas unidas a resina se eluyeron mediante el uso de glutatión reducido. Las proteínas eluídas se separaron por electroforesis desnaturalizante y posteriormente se transfirieron a una membrana de nitrocelulosa. La presencia de las proteínas recombinantes 6xHis-TvTBPs se reveló usando la sonda HisProbe-HRP. Por ensayos previos sabemos que 6xHis-TvTBP2 se expresa en menor cantidad que 6xHis-TvTBP1 a pesar de usar distintas condiciones de inducción. Para tratar de igualar la cantidad de proteína con respecto a 6xHis-TvTBP1, se usó el doble de cultivo para 6xHis-TvTBP2. Esto se puede observar en la Figura 18A en el carril correspondiente al extracto total donde existe una mayor carga de proteínas para los extractos que contienen 6xHis-TvTBP1 que para 6xHis-TvTBP2. Sin embargo, como se observa en la inmunodetección (Figura 18B), la cantidad de 6xHis-TvTBP2 sigue siendo menor. En el gel teñido con azul de Coomassie (Figura 18A), se puede observar la cantidad de GST usada en el ensayo (de 26 kDa, señalado con un triángulo rojo), así como la

cantidad de GST-IBP39 (de aproximadamente 65 kDa, señalado con un triángulo negro). En la imagen de la inmunodetección, una señal de aproximadamente 27 kDa que corresponde al tamaño de 6xHisTvTBP1 aparece en la muestra coprecipitada con GST-IBP39, pero no con GST sola (Figura 18B, triángulo azul). No se encontró interacción entre IBP39 y 6xTvTBP2 bajo las condiciones probadas (Figura 18B, triángulo verde).

La interacción observada por coprecipitación entre 6xHis-TvTBP1 y GST-IBP39 sugiere que TvTBP1 podría formar parte del complejo de inicio de la transcripción transcrito por la RNA polimerasa II, junto a IBP39.

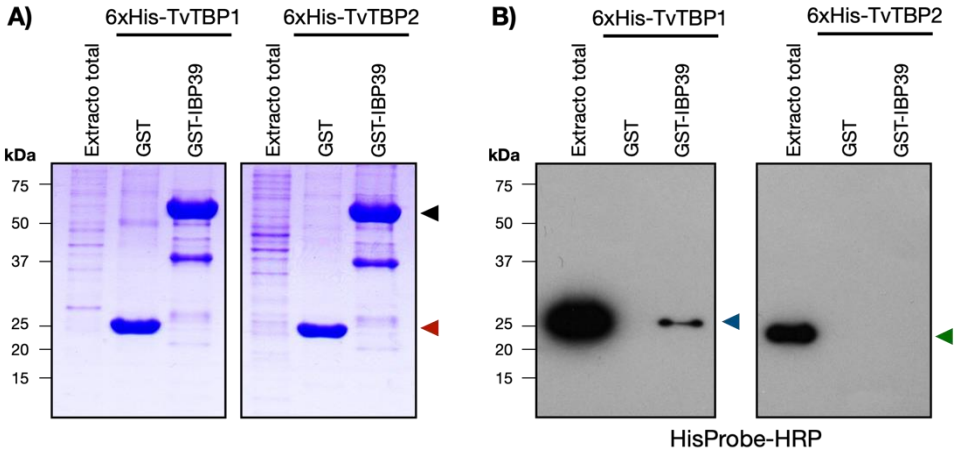


Figura 18.- GST-IBP39 interactúa con 6xHis-TvTBP1. A) Las muestras del ensayo de coprecipitación entre las proteínas GST-IBP39/6xHis-TvTBP1 y GST-IBP39/6xHis-TvTBP2 se cargaron en un gel de poliacrilamida desnaturalizante al 12 % y se tiñó con azul de Coomassie. El extracto total corresponde a 1/10 del lisado celular de *E. coli* que expresa a la proteína recombinante 6xHis-TvTBP1 y 2/10 de 6xHis-TvTBP2. GST se expresó y usó como control negativo en el ensayo de coprecipitación. GST-IBP39 de ~ 65 kDa se encuentra señalado con un triángulo negro, GST de ~ 26 kDa se encuentra señalado con un triángulo rojo. B) Inmunodetección usando la sonda HisProbe-HRP para revelar la presencia de las proteínas 6xHis-TvTBP1 (triángulo azul) y 6xHis-TvTBP2 (triángulo verde) de ~ 27 kDa.

Unión de las TvTBPs al DNA

En un trabajo previo de nuestro laboratorio, se realizaron ensayos de retardo de la movilidad electroforética (EMSA) para determinar si las TvTBPs pueden reconocer y unirse a secuencias de DNA, específicamente a las cajas TATA. Sin embargo, bajo las condiciones probadas no se observó ningún retardo (M. en C. Lluvia Rosas, tesis doctoral no publicada). No obstante, la capacidad de unión al DNA es una de las características más emblemáticas de la proteína TBP,

motivo por el cuál, como parte de este trabajo, se decidió corroborar y probar otras condiciones en el ensayo.

En ensayos EMSA se usan comúnmente tres tipos de amortiguadores de corrida: TAE (Tris-Acetato-EDTA), TBE (Tris-Borato-EDTA) y TG (Tris-Glicina). El amortiguador de corrida TG 1X (usado en los ensayos previos) ofrece una mayor fuerza iónica durante el tiempo de corrida de las muestras que el amortiguador TBE 0.5X (Roder y Schweizer, 2001), provocando que complejos DNA-proteína no estables se disocien fácilmente. Asimismo, las condiciones fisicoquímicas en el amortiguador de unión son decisivas para observar retardos. En esta sección se decidió probar diferentes condiciones para evaluar la unión de TBP-DNA y la estabilidad del complejo durante la electroforesis haciendo uso de las mismas secuencias como sondas con las que ya se contaba y algunas adicionales (ver condiciones en el anexo I).

Unión de las TvTBPs a promotores transcritos por las tres RNA polimerasas

Para probar la capacidad de unión al DNA de TvTBP1 y TvTBP2, se realizó un ensayo EMSA usando cinco diferentes sondas derivadas de las secuencias promotoras reconocidas por las tres RNA polimerasas de *T. vaginalis* (Figura 19A):

1) La sonda rDNA (164 pb) que incluye toda la región promotora del gen RNA ribosomal y que contiene los elementos de control río arriba (UCE y el *core*) previamente descritos para este promotor (Franco *et al.*, 2012).
2) La sonda SCS (82 pb) contiene la secuencia promotora del gen que codifica la subunidad β de la succinil CoA sintetasa y contiene un elemento Inr (Lahti *et al.*, 1992).
3) La sonda de actina (102 pb), que corresponde al promotor de un gen de actina (TVAG_172680), la cual contiene al elemento Inr y un elemento con similitud a una caja canónica en la posición -59 (Espinosa *et al.*, 2001).
4) La sonda 5S (186 pb) que contiene algunos elementos regulatorios cercanos al sitio de inicio de la transcripción: el elemento río arriba del promotor *(*UPE, *upstream promotor element*), el elemento que incluye el elemento del sitio de inicio (SSE, *star site element*) y al motivo 5 (M5) (Torres-Machorro *et al.*, 2006).
5) La sonda U6 (120 pb) que contiene una caja putativa a los -28 pb (Simoes-Barbosa *et al.*, 2008).

Las sondas se incubaron a temperatura ambiente con las proteínas recombinantes 6xHis-TvTBP1 y 6xHis-TvTBP2 en amortiguador de unión. Las muestras se cargaron a un gel de acrilamida-bisacrilamida al 4.5 % en condiciones nativas. Posteriormente los geles se tiñeron con bromuro de etidio para visualizar el DNA. El retardo observado en presencia de TvTBP1 y TvTBP2 indicó que ambos factores se unen a todas las sondas probadas (Figura 19B). Este resultado no fue del todo inesperado, ya que el tamaño de las sondas probadas es grande, contienen muchos elementos y son regiones ricas en adeninas y timinas. De modo interesante, TvTBP1 parece unirse con una mejor afinidad al DNA que TvTBP2, adicionalmente en los retardos con TvTBP2 se alcanza a ver un barrido a lo largo del carril, indicando que TvTBP2 forma un complejo DNA-proteína inestable.

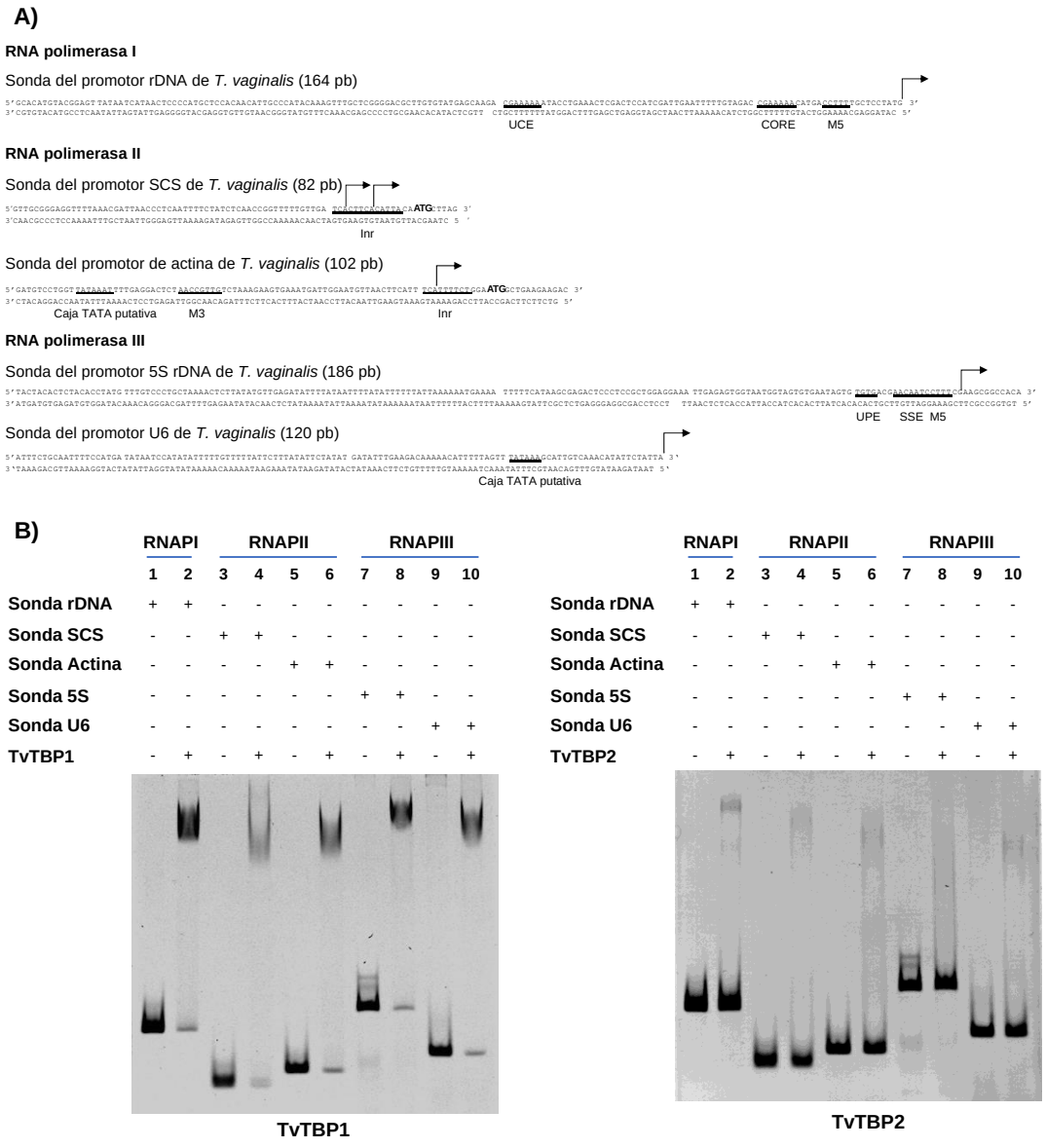


Figura 19.- Las TvTBPs pueden unirse a diferentes secuencias promotoras de *T. vaginalis*. A) Secuencias de los promotores usados como sonda. Cada elemento en la secuencia se la sonda se

encuentra subrayado. El sitio de inicio de la transcripción es señalado con una flecha. Abreviaturas: Elemento de control río arriba (UCE, *Upstream control element*), Motivo 5 (M5, *Motif 5)*; Motivo 3 (M3, *Motif 3)*; Elemento del promotor río arriba (UPE, *Upstream promoter element*) y Elemento del sitio de inicio (SSE, *Start site element*). El sitio de inicio de la transcripción está indicado con una flecha. **B)** EMSAS para TvTBP1 y TvTBP2. Los geles se tiñeron con bromuro de etidio y se fotografiaron bajo luz UV. Los colores de las imágenes mostradas están invertidos.

Las diferencias de unión observadas con TvTBP1 y TvTBP2 son considerables. Sin embargo, estos resultados deben tomarse con precaución, ya que como se mencionó anteriormente, TvTBP2 se induce en menor cantidad que TvTBP1 en la bacteria y para igualar las concentraciones de ambas proteínas usadas en el ensayo, se requiere usar un mayor volumen de la elución de TvTBP2 obtenida al final de la purificación. Como resultado, los ensayos con TvTBP2 tienen ligeramente una mayor concentración de sales en la reacción final que podría afectar la interacción de TvTBP2 con el DNA.

TvTBP1 y TvTBP2 se unen al DNA de manera secuencia-independiente

Con el objetivo de definir un sitio de unión de las TvTBPs en el DNA, se decidió usar la sonda del promotor de actina (102 pb) como secuencia modelo, debido que presenta un elemento con similitud a una caja TATA (TATAAATT) entre los nucleótidos -66 a -59. Algunas adeninas de esta secuencia se cambiaron por citosinas (TCTCACTT) para probar si este elemento es determinante en la unión de TvTBP1 (Figura 20A). Para este ensayo se usaron 25, 75 y 100 ng de TvTBP1 y 75 ng de la proteína TBP de levadura (Spt15) como control de unión. Las proteínas se incubaron con una sonda radioactiva. TvTBP1 se une a ambas sondas, aunque parece haber ligeramente menos retardo con la sonda mutada que con la sonda silvestre (WT, por sus siglas en inglés) (Figura 20B). Curiosamente, Spt15 se une igualmente a ambas sondas y se pueden ver al menos tres complejos DNA-proteína, uno de ellos claramente enriquecido. Estos resultados indican que el elemento considerado como una caja TATA putativa en la sonda de actina no es relevante para la unión de TvTBP1 al DNA, por lo que altamente probable que otra región del DNA esté involucrada en el reconocimiento de TvTBP1.

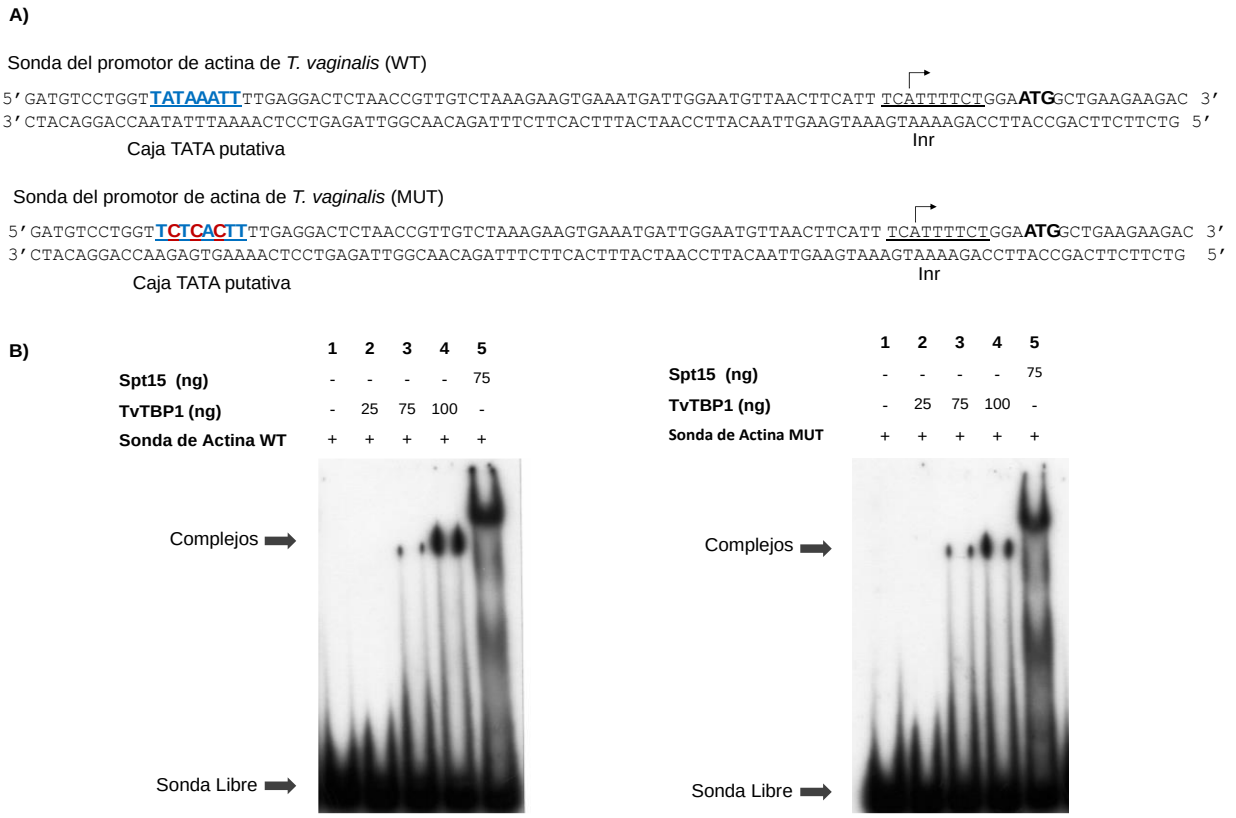


Figura 20.- TvTBP1 se une a una sonda con la caja TATA mutada. A) Secuencias de las sondas de actina silvestre (WT) y mutada (MUT). **B)** Ensayos de retardo para TvTBP1 y la proteína TBP de *S. cerevisiae* (Spt15). Para estos ensayos se usaron sondas marcadas radioactivamente.

La siguiente estrategia consistió mapear la región del DNA que es reconocida por TvTBP, para ello, se obtuvieron sondas más cortas a partir de la sonda de actina de 102 pb mediante amplificación. La caja TATA putativa está contenida en las sondas A y B, mientras que, C, D y E no presentan este elemento. Las sondas B y D son más cortas del extremo 3' y por tanto no contienen al elemento Inr completo (Figura 21A). Cantidades similares de proteína se incubaron con las sondas, mientras que la cantidad de sondas se intentó igualar, no siempre fue el caso. Las muestras se cargaron al gel en condiciones nativas y fue teñido con bromuro de etidio, por lo que hay que considerar la intensidad de cada señal también es debido al tamaño de la sonda. Adicionalmente, para este ensayo se incluyó como control negativo de unión una proteína cinasa dependiente de ciclina de *T. vaginalis* (TvCRK5) (Amador *et al.*, 2017), esta es una proteína no relacionada.

Con TvTBP1 se observa retardo con cada una de las sondas usadas, lo mismos para TvTBP2, aunque las señales de los retardo son muy débiles. TvCRK5 no mostró ninguna interacción con el DNA, confirmando que las señales de retardo observados para las TvTBPs son reales. Los

resultados generales sugieren que las TBPs de *Trichomonas* se unen al DNA independientemente de secuencia.

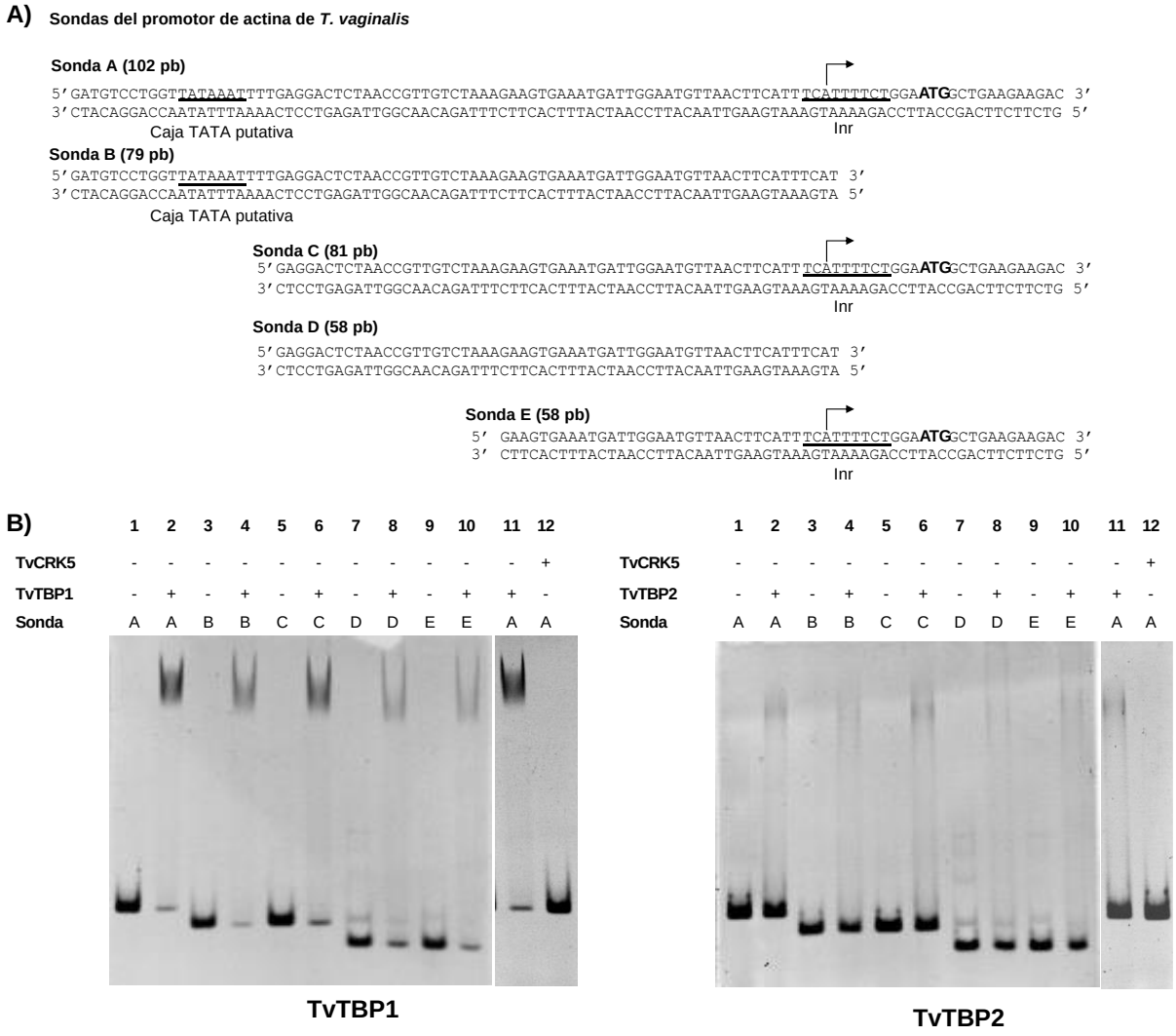


Figura 21.- Unión de las TvTBPs al promotor de actina. Panel superior: diferentes regiones del promotor de actina que se usaron como sonda. Panel inferior: ensayos tipo EMSA utilizando cada una de las sondas indicadas y las proteínas TvTBP1 y TvTBP2. Los geles se tiñeron con bromuro de etidio y fotografiados bajo luz UV. Los colores de las imágenes mostradas están invertidos.

En un ensayo adicional, se usó una sonda no relacionada, proveniente del sitio de multiclonación del plásmido pUC19, sin embargo, ambas TvTBPs retardan esta sonda. Los datos mostrados sugieren que las TvTBPs son capaces de unirse al DNA y que probablemente requieren de otro factor transcripcional que les dé especificidad de unión.

Conclusiones

1. Los datos obtenidos hasta ahora demuestran que ambas TvTBPs son capaces de formar dímeros *in vitro*.
2. TvTBP1 es una proteína nuclear y forma dímeros *in vivo*.
3. TvTBP1 es capaz de interactuar con IBP39, una proteína que participa en el reconocimiento de promotores transcritos por la RNA polimerasa II, sugiriendo que TvTBP1 podría formar parte del complejo de inicio de la transcripción junto con IBP39.
4. Ambas TvTBPs reconocen al DNA y se unen a él de manera inespecífica, por lo que es probable que *in vivo* requieran de otros factores adicionales para el reconocimiento adecuado del promotor.

Perspectivas

1.- Evaluar si las TvTBPs forman parte de los complejos de inicio de la transcripción *in vivo*. Para contestar esta pregunta se sugiere realizar un ensayo tipo ChIP. En el laboratorio contamos con un anticuerpo anti-IBP39, obtenido durante este trabajo, que nos permitiría conocer si alguna de las TvTBP se encuentra formando parte del PIC transcrito por la RNA polimerasa II.

2.- Inferir si ambas proteínas se expresan al mismo nivel. Mediante ensayos usando RT-qPCR se podrían conocer los niveles de expresión de ambos genes y así darnos una idea si las proteínas codificadas por estos podrían tener una expresión diferencial. *T. vaginalis* cuenta son un solo estadio, aunque dos formas son conocidas, como trofozoíto y con forma ameboide cuando se encuentra en presencia de hierro en el medio de cultivo.

3.- Determinar el papel funcional de ambas TvTBPs. *T. vaginalis* es un protozoario unicelular que expresa dos TvTBPs. En humanos y *D. melanogaster*, se ha reportado una TBP y otros factores relacionados a TBP, la diferencia entre estas proteínas es la participación en la transcripción de distintos genes.

CAPÍTULO III: El complejo TFIIIB en *T. vaginalis*

Introducción

La transcripción por la RNA polimerasa III está especializada en la síntesis de pequeñas moléculas de RNA, involucradas principalmente en la traducción y el procesamiento del RNA (Turowski y Tollervey, 2016). Un total de 17 proteínas componen esta polimerasa: cinco subunidades se encuentran compartidas entre las tres RNA polimerasas (ABC27, ABC23, ABC14.5, ABC10α y ABC10β), dos subunidades se comparten con la RNA polimerasa I (AC19 & AC40), cuatro subunidades más son homólogas a subunidades de la RNA polimerasa I y la RNA polimerasa II (C160, C128, C25 y C11) y seis subunidades más son únicas de la RNA polimerasa III (C82, C53, C37, C34, C31 y C17) (Turowski y Tollervey, 2016).

Los promotores reconocidos por la RNA polimerasa III se pueden clasificar en tres tipos (Figura 22).

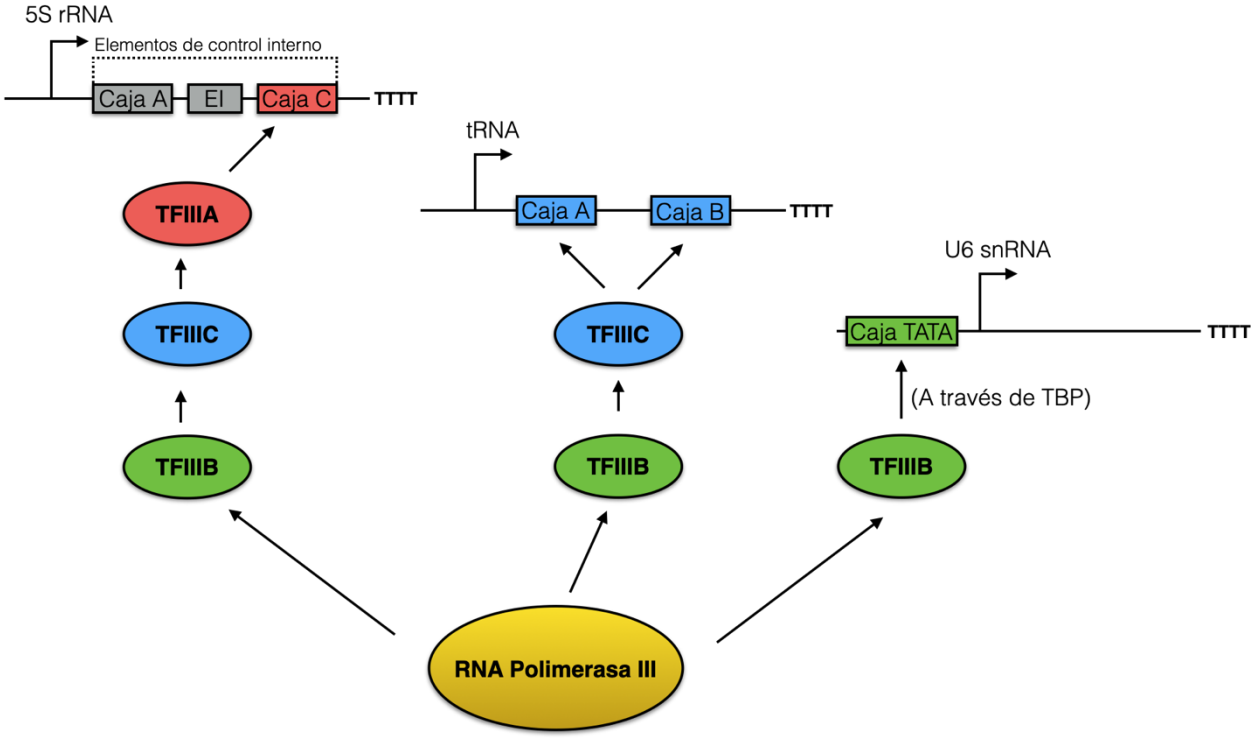


Figura 22.- Transcripción por la RNA polimerasa III. Se ejemplifican los tres tipos de promotores transcritos por la RNA polimerasa III: El promotor de tipo I del gen 5S que contiene una región de control interno que consta de tres elementos: la caja A, un elemento intermedio (EI) y la caja C. El promotor de tipo II del gen tRNA (clase II) que consta de una caja A y una caja B. El promotor de tipo III del gen U6 (clase III) que tiene elementos río arriba, como la caja TATA. TFIIIA es esencial para el reconocimiento del promotor 5S a través de la caja C. El complejo TFIIIC es responsable del reconocimiento del promotor de tRNA a través de las cajas A y C.

Adicionalmente, TFIIIC estabiliza el complejo TFIIIA-DNA en el promotor del 5S. El complejo TFIIIB participa en el reconocimiento del promotor de U6 a través de TBP y es esencial para la formación de los complejos de inicio de la transcripción de los tres tipos de promotores de la RNA Pol III.

Los promotores de la clase I son característicos del gen 5S, contienen elementos de control interno compuestos por la caja A, un elemento intermedio (IE) y la caja C, este último elemento es reconocido por el factor transcripcional TFIIIA, que a su vez recluta a TFIIIC, seguido por TFIIIB y al resto de la maquinaria transcripcional.

Los promotores de la clase II son típicos de los genes tRNAs, contienen dos elementos intragénicos conservados: la caja A y la caja B. Estos coinciden con las características estructurales conservadas en el cuerpo del tRNA maduro, denominado bucle D y bucle T (Turowski y Tollervey, 2016). Ambos elementos son reconocidos por el complejo TFIIIC para posteriormente reclutar al resto de la maquinaria transcripcional.

Los promotores de la clase III son típicos del gen U6. Estructuralmente se parecen más a los promotores transcritos por la RNA polimerasa II, tienen elementos externos del gen río arriba del sitio de inicio de la transcripción: el elemento de secuencia proximal (PSE), un elemento de secuencia distal (DSE) y una caja TATA, la cual es reconocida por la proteína TBP y forma parte de un complejo conocido como TFIIIB, junto a las proteínas Brf1 y Bdp1. Este complejo también participa en la transcripción de los otros tipos de promotores, pero solo en éstos es encargado del reconocimiento del DNA y del posterior reclutamiento de la maquinaria transcripcional.

Los tres tipos de promotores contienen una corrida de timinas al final de la secuencia del gen que es importante para finalizar la transcripción.

La transcripción por la RNA polimerasa III se regula principalmente durante el inicio de la transcripción (Abascal-Palacios *et al.*, 2018), por tres factores de transcripción que son claves en este proceso: TFIIIA, TFIIIB y TFIIIC (Figura 22).

TFIIIA

TFIIIA es una proteína de ~40 kDa que se caracteriza por tener nueve motivos de dedos de zinc consecutivos, permitiéndole unirse con una gran afinidad y especificidad al DNA. Sin embargo, TFIIIA es poco conservado en secuencia y las similitudes entre este tipo de proteínas está limitado a la presencia de este dominio. TFIIIA participa en la transcripción del promotor 5S a través del reconocimiento del elemento de control interno (ICE) (Layat *et al.*, 2013).

TFIIIB

TFIIIB es un complejo compuesto por tres subunidades: TBP, Brf1 (*B-related factor 1*) y Bdp1 (*B double prime* o B’’). TBP y Brf1 forman un complejo muy estable llamado B’, mientras que Bdp1 (B’’) se encuentra débilmente asociado a estos factores en ausencia de DNA y es separable por cromatografía. Estudios bioquímicos indican que TBP y Brf1 se unen cooperativamente al DNA; sin embargo, la presencia de Bdp1 contribuye a la formación de un complejo TFIIIB-DNA altamente estable (Gouge *et al.*, 2017). La unión de TFIIIB al DNA es un paso importante para el posterior reclutamiento de la RNA polimerasa III y el inicio de la transcripción (Khoo *et al.*, 2014). En levaduras, TFIIIB puede unirse al DNA por dos vías: a través de la unión a cajas TATA mediada por TBP o a través del complejo TFIIIC (Kassavetis y Geiduschek, 2006).

En levaduras, TBP participa en la transcripción de los tres tipos de promotores transcritos por la RNA polimerasa III. Sin embargo, en años recientes se han identificado proteínas con similitud a TBP. TRF1 es un factor transcripcional relacionado a TBP identificado en *Drosophila* y es específico de insectos (Crowley *et al.*, 1993). El extremo C-terminal de TRF1 tiene un alto grado de similitud con TBP y puede unirse a TFIIA, TFIIB y a DNA y es capaz de sustituir a TBP en ensayos de transcripción *in vitro* (Hansen *et al.*, 1997). Experimentos bioquímicos muestran que TRF1 interacciona con Brf1 y puede sustituir a TBP del complejo TFIIIB (Kang *et al.*, 2016). Adicionalmente, TRF1 participa en la transcripción del gen 5S y diversos genes de tRNA (Takada *et al.*, 2000). Sin embargo, TBP es necesario para la transcripción de U6 (Verma *et al.*, 2013).

El dominio N-terminal de Brf1 contiene un plegamiento de dedos de zinc y dos plegamientos tipo ciclina que constan de 5 hélices α. Este dominio es homólogo a los factores de transcripción TFIIB y Rrn7 del sistema transcripcional de la RNA polimerasa II y la RNA polimerasa I, respectivamente. Mientras que el C-terminal de Brf1 le otorga afinidad de unión por la RNA polimerasa III. En levaduras, Brf1 contiene tres bloques de homología de secuencia, el bloque I (entre los aminoácidos 287 a 304), II (entre los aminoácidos 461 a 515) y III (entre los aminoácidos 570 a 596) que se encuentran conservados en *Schizosaccharomyces pombe*, *Candida albicans*, *Kluyveromyces lactis* y *Homo sapiens* (Khoo *et al.*, 2014). Análisis estructurales del complejo Brf1-TBP-DNA indican que el bloque de homología II está posicionado a lo largo de la superficie convexa y lateral de TBP, este bloque también interacciona con Bdp1. Los bloques de homología están separados por dos regiones no conservadas, llamadas C-*linker* 1 y 2 (Khoo *et al.*, 2014). En humanos, una variante de BRF1, llamada BRF2 es responsable de la transcripción del gen U6

(Schramm *et al.*, 2000). Brf1 interactúa con algunas subunidades del complejo TFIIIC (Tfc1, Tfc4 y Tfc8) y de la RNA polimerasa III (C160, C128, C34 y C17) (Khoo *et al.*, 2014).

La proteína Bdp1 fue descrita por primera vez en levadura (Kassavetis *et al.*, 1995; Rüth *et al.*, 1996; Roberts *et al.*, 1996). Bdp1 se caracteriza por contener un dominio SANT altamente conservado, llamado así porque se identificó en proteínas reguladoras de la cromatina y en reguladores transcripcionales (SWI3, ADA2, N-CoR y TFIIIB B''). El dominio SANT es un motivo de aproximadamente 50 aminoácidos que se pliegan formando una hélice-vuelta-hélice conformado por tres hélices alfa, que a diferencia del dominio Myb, tiene como función el contacto proteína-proteína (Aasland *et al.*, 1996). Bdp1 no tiene parálogos en los complejos transcripcionales de la RNA polimerasa I ni la RNA polimerasa II. En levadura, la región central de Bdp1, que contiene al dominio SANT, es esencial para la transcripción junto con el N-terminal de Brf1. En levadura, los residuos 241 a 285 de Bdp1, tienen una estructura predicha de hoja beta plegada que es esencial para la apertura del promotor. Mientras que, en humano, el gen *TFNR* codifica para la proteína BDP1 que tiene una longitud de 2254 aminoácidos y contiene 9 repetidos de un motivo de 55 aminoácidos con función desconocida. La secuencia codificante de HsBDP1 está conformada por 32 exones que da origen a diversas isoformas por *splicing* alternativo que van desde los 55 kDa a los 250 kDa (Kelter *et al.*, 2000). Los ortólogos de esta proteína se encuentran poco conservados en secuencia y solo comparten el dominio SANT.

TFIIIC

En levaduras, TFIIIC es un complejo compuesto por 6 subunidades Tfc1, Tfc3, Tfc4, Tfc6, Tfc7 y Tfc8 (TFIIIC63, TFIIIC220, TFIIIC102, TFIIIC110, TFIIIC35 y TFIIIC90 en humanos). Estas subunidades se encuentran organizadas en dos dominios globulares conocidos como τ_A y τ_B, debido a que se unen a las cajas A y B del promotor del gen tRNA, respectivamente. Ambos dominios se encuentran conectados por una región sensible a proteólisis. τ_A está compuesto por Tfc1 (74 kDa), Tfc4/Pcf1 (120 kDa) y Tfc7 (49 kDa), constituyendo el dominio de unión río arriba. τ_B está compuesto por las subunidades Tfc3 (132 kDa), Tfc6 (75 kDa) y Tfc8 (68 kDa) (Geiduschek y Kassavetis, 2001). Tfc8 funciona de enlace entre ambos dominios, mientras que Tfc6 y Tfc3 participan en la unión al DNA. El complejo TFIIIC participa en la transcripción de los promotores de las clases I y II (Vannini, 2013). TFIIIC se une y estabiliza a complejos preformados de TFIIIA-DNA sobre el promotor del 5S. La función principal de TFIIIC se ubica principalmente durante el inicio de la transcripción, en el reconocimiento del promotor y el reclutamiento de TFIIIB. A pesar

de esto, se entiende poco sobre los mecanismos moleculares que subyacen las funciones transcripcionales de TFIIIC en otras etapas de la transcripción (Vannini, 2013).

El complejo TFIIIB en *T. vaginalis*

Dada la importancia que parece tener el complejo TFIIIB para el inicio de la transcripción de la RNA polimerasa III y que TBP es parte de este complejo, nos preguntamos si existe un ortólogo del complejo TFIIIB en *T. vaginalis* y cuál de las proteínas TvTBPs que se expresan en este organismo es la que participa en la transcripción de la RNA polimerasa III.

El ortólogo de Brf1 en *T. vaginalis* (TvBRF1) se identificó mediante un análisis *in silico* de interactoma obtenido a través de la base de datos STRING (*Search Tool for the Retrieval of Interacting Genes*) que genera redes de asociación funcionales entre proteínas usando las secuencias de las TvTBPs en la búsqueda (M. en C. Lluvia Rosas Hernández) y el gen correspondiente se clonó en el plásmido pGEX4T1 (Dra. Alexandra Ibáñez Escribano), que permite obtener una proteína recombinante fusionada a GST en el extremo N-terminal. Esta proteína se usó para realizar ensayos de coprecipitación (*GST-pulldown)*. Adicionalmente, se contaban con algunos plásmidos para probar la interacción mediante el ensayo de doble híbrido en levaduras: TvTBP1 y TvTBP2 en el plásmido pEG202 (M. en C. Lluvia Rosas Hernández), que se usaron en este trabajo.

El ortólogo de Bdp1 en *T. vaginalis* (TvBDP1), se identificó durante el desarrollo de este trabajo.

Resultados:

TvBRF1 tiene dominios característicos de la familia relacionada con TFIIB

En el genoma de *T. vaginalis* hay un gen anotado como *factor del inicio de la transcripción brf1, putativo* de 1209 pb que codifica para una proteína de 409 aminoácidos (TVAG_139840) (Carlton *et al.*, 2007). Una etiqueta de secuencia expresada (EST) está anotada en TrichDB, sugiriendo la expresión a nivel de mRNA de TvBRF1. La proteína Brf1 tiene homólogos funcionales que participan en la transcripción por la RNA polimerasa I (Rrn7 en levaduras o TAF1B en humanos) y la RNA polimerasa II (TFIIB). Para descartar que la proteína identificada como TvBRF1 putativa corresponde a un homólogo de Brf1, se realizó un árbol filogenético usando las secuencias de los factores pertenecientes a esta familia de humano y levaduras. TvBRF1 se agrupa junto a sus ortólogos demostrando que realmente se trata de Brf1 (Figura 23A). Adicionalmente, un

alineamiento de las secuencias proteicas de BRF1 de humano, *S. cerevisiae* y *S. pombe* nos permitieron identificar los dominios más importantes de TvBRF1 (Figura 23B). El extremo N-terminal de TvBRF1 contiene dos dominios característicos: los dedos de zinc (entre los aminoácidos 2 a 32) y dos plegamientos de ciclina (entre los aminoácidos 74 a 164 y 173 a 264). A su vez, estos dominios se conectan por un espaciador en el dominio N-terminal (*N-linker*), que en levaduras presenta dos estructuras claras: el *B-reader* (BR) y el *B-linker* (BL). El extremo C-terminal de TvBRF1 es claramente más corto y presenta poca identidad de secuencia con las proteínas de levaduras y humanos. TvBRF1 se encuentra relativamente más cercano a HsBRF1 (33 % de identidad) que a ScBrf1 (28 %). Finalmente, el N-terminal de TvBRF1 se plegó por homología de secuencia usando como templado la estructura tridimensional obtenida por cristalografía de rayos X de ScBrf1 (Figura 23C), la proteína TvBRF1 adopta una arquitectura similar a la de levadura, conservando los plegamientos típicos de ciclina y los dedos de zinc. En estas estructuras se puede observar que TvBRF1 presenta el *B-reader*, pero no el *B-linker*.

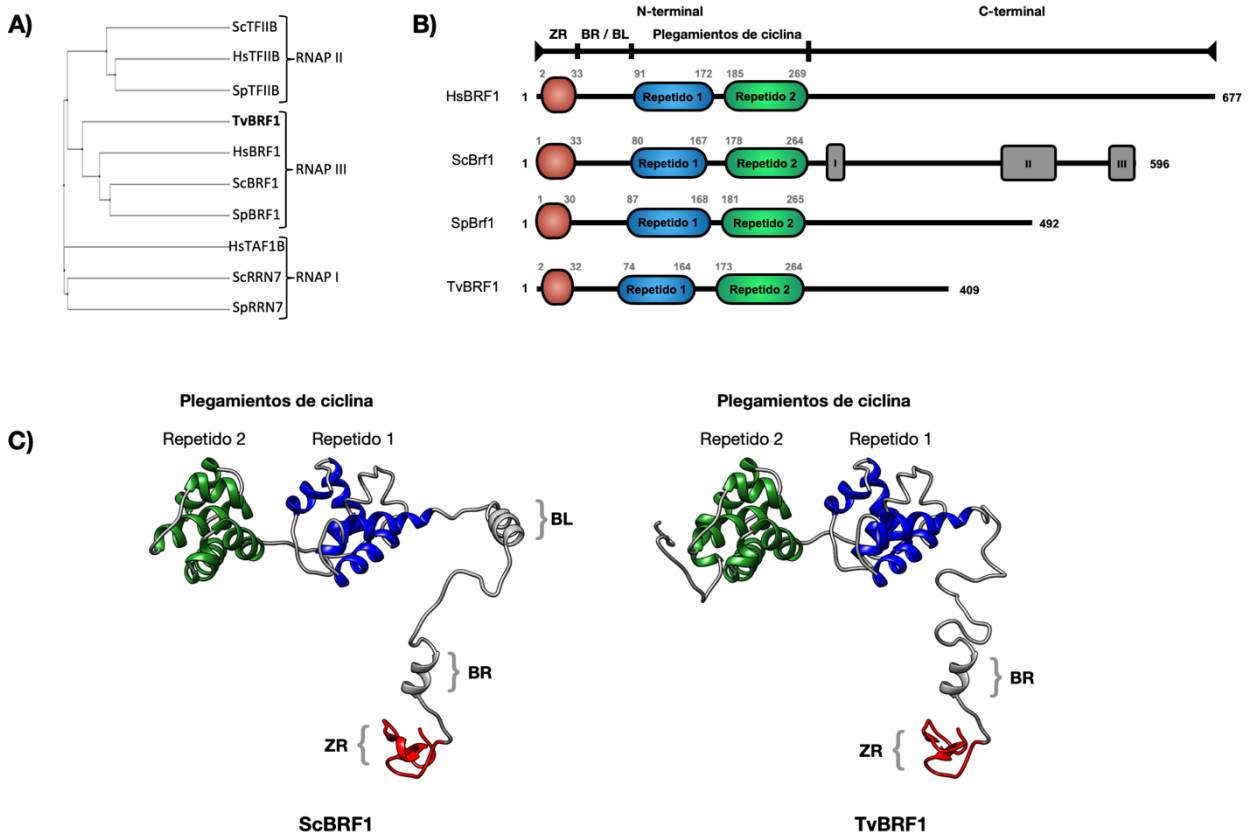


Figura 23.- Dominios y estructura predicha de TvBRF1. A) Árbol filogenético de TFIIB y los factores relacionados a TFIIB (HsTAF1B: Q53T94, HsTFIIB: Q00403, and HsBRF1: Q92994), *S. cerevisiae* (ScRrn7: P40992, ScTFIIB: P29055 y ScBrf1:P29056), *S. pombe* (SpRrn7: Q9UST5, SpTFIIB: O13749, and SpBrf1:Q9P6R0), y *T. vaginalis* (TvBRF1: A2EJ51). El árbol filogenético

se obtuvo por el método de *Neighbor-joining*. B) Dominios de BRF1: los dedos de zinc, ZR (rojo), el espaciador del amino terminal compuesto a su vez por el BR (*B-Reader*) y el BL (*B-Linker*), los plegamientos de ciclina (azul y verde) y los tres bloques de homología descritos para Brf1 de levadura (gris). C) Estructura predicha del N-terminal de TvBRF1 usando a Brf1 de *S. cerevisiae* como modelo: El dominio de dedos de zinc (rojo) y los plegamientos de ciclina (azul y verde) se indican en la figura.

TvBRF1 no complementa una cepa mutante de levadura

Como un método para determinar la función de TvBRF1, se realizó un ensayo de complementación en una cepa mutante de *BRF1* de levadura. Para este ensayo, se obtuvo a la cepa mutante de *BRF1* de levadura usando una cepa diploide, debido a que *BRF1* es un gen esencial. La mutación se realizó mediante recombinación homóloga. Uno de los alelos que contiene el gen *BRF1* se reemplazó por un casete de *KANMX*, el cual le confiere resistencia al antibiótico G418 y permitió la selección de las clonas mutadas. Adicionalmente se obtuvieron los plásmidos TvBRF1-pYES2 y ScBRF1-pYES2 en los que los genes Tv*brf1* y Sc*BRF1*, respectivamente, se encuentran bajo el control del promotor *GAL1*. Los plásmidos se transformaron a la cepa de levadura *brf1*Δ::*KANMX* y se puso en medio de esporulación para obtener esporas con cuatro tétradas (dos que contienen el alelo silvestre y dos con el alelo que contiene el casete de *KANMX*.

La complementación se analizó mediante disección de tétradas (Figura 24). Se observa crecimiento de cuatro tétradas cuando la cepa se transformó con el plásmido ScBRF1-pYES2 (control positivo) y sólo dos tétradas cuando la cepa se transformó con el plásmido pYES2 vacío (control negativo). La cepa transformada con TvBRF1-pYES2 muestra únicamente crecimiento de dos esporas, por lo que es posible concluir que la proteína de *T. vaginalis* no complementa la función del gen *BRF1* de levadura (Figura 24A). Todas las tétradas crecen en medio rico en galactosa (YPGAL); sin embargo, en presencia de geneticina, sólo pueden crecer las esporas que contienen el alelo con el casete de *KANMX*, para la cepa transformada con el plásmido ScBRF1-pYES2 este resultado corrobora que las esporas corresponden a una misma tétrada, mientras que para las cepas transformadas con el plásmido TvBRF1-pYES2 o el plásmido vacío, nos indica que las dos esporas que crecen en YPGAL contienen el gen endógeno y no el alelo reemplazado con *KANMX*. Para una segunda comprobación, se extrajo el DNA de cada una de las células que crecieron en el medio YPGAL y se realizó una amplificación por PCR usando un par de oligonucleótidos, los cuales se anclan rio arriba y río abajo del gen Sc*BRF1* y del casete de KANMX integrado en el mismo contexto genómico. El producto obtenido de la amplificación del gen Sc*BRF1* tiene un tamaño esperado de 2,399 pb y del casete de *KANMX* de 2,154 pb.

De los amplicones obtenidos del control positivo, se puede observar que dos corresponden en tamaño con el alelo con el casete de *KANMX* y dos con el alelo silvestre (flechas verdes), mientras que los amplicones obtenidos para el control negativo y para TvBRF1 tienen el tamaño del alelo silvestre (flechas azules y amarillo, respectivamente) (Figura 24B). Estos resultados demuestran que TvBRF1 no complementa la cepa mutante de levadura.

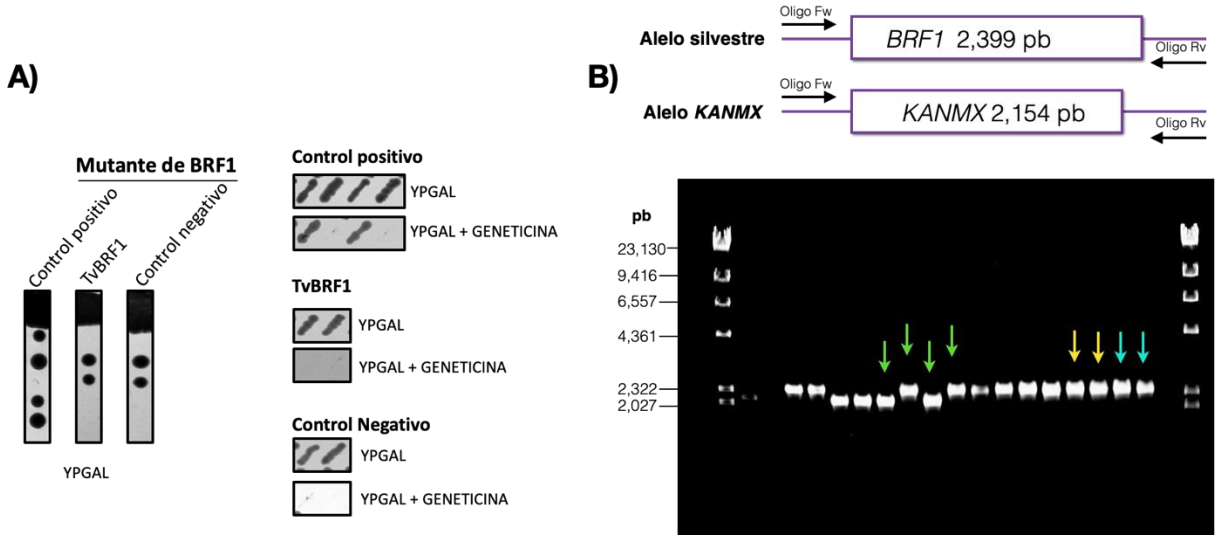


Figura 24.- TvBRF1 no complementa una cepa mutante *brf1*Δ::*KANMX* de levadura. A) La cepa *brf1Δ::KANMX* fue transformada con los plásmidos TvBRF1-pYES2, ScBRF1-pYES2 (control positivo) y pYES2 vacío (control negativo) y crecidas en medio mínimo hasta que se observó esporulación. Las esporas se trataron con liticasa, se separaron por disección de tétradas y se crecieron en medio YPGAL. Posteriormente, cada célula que creció en este medio se inoculó en los medios YPGAL y en YPGAL con geneticina. B) Comprobación mediante amplificación por PCR usando un par de oligonucleótidos que se anclan rio arriba y rio abajo del locus del gen *BRF1* en levadura. Las flechas verdes corresponden al control positivo, las amarillas a TvBRF1 y las azules al control negativo.

El modelo estructural de TvBRF1 indica que los dominios más representativos del extremo N-terminal están conservados con respecto a ScBrf1. Por otro lado, la porción C-terminal de ScBrf1 es esencial para su función y contiene residuos importantes para la interacción con la RNA polimerasa III y las subunidades TBP y Bdp1 del complejo TFIIIB (Knutson y Hahn, 2011; Khoo *et al.*, 2014) que no se encuentran conservados en TvBRF1, debido a que la secuencia de aminoácidos es más corta para la proteína de *Trichomonas* (Figura 25). Adicionalmente, el CTD de ScBrf1 contiene tres dominios de homología de secuencia que probablemente tenga sitios de interacción con otras proteínas propias de levadura que no pueden ser complementadas por la proteína de *Trichomonas*.

Para determinar si los dominios del N-terminal de TvBRF1 son funcionales, sería interesante obtener una construcción que nos permita tener una fusión del N-terminal de TvBRF1 y el C-terminal de ScBrf1. Si no hubiera complementación con la fusión, significaría que los residuos importantes para llevar a cabo la interacción con algunas subunidades de la RNA polimerasa III (C17, C160, C34 y C128) en ScBrf1, son esenciales, ya que no se encuentran del todo conservados en TvBRF1 (Figura 25).

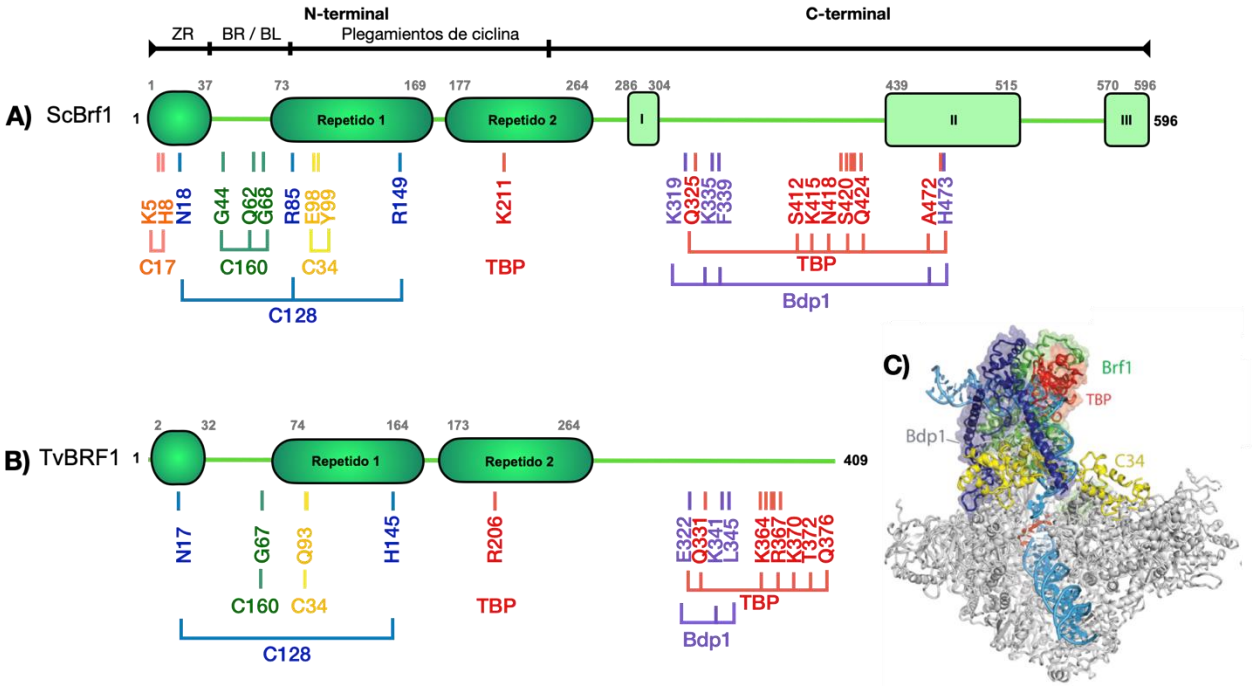


Figura 25.- Aminoácidos de ScBrf1 involucrados en la interacción con la RNA polimerasa III, TBP y Bdp1 en comparación con TvBRF1. A) Las líneas de colores indican la posición en ScBrf1 de los aminoácidos involucrados en la interacción con las subunidades de la RNA polimerasa III: C17 (naranja), C128 (azul), C160 (verde) C34 (amarillo), TBP (rojo) y Bdp1 (morado). Las abreviaturas ZR indican los dedos de zinc, el espaciador N-terminal está indicado por las letras BR/BL (*B-Reader* y *B-Linker*) y los tres bloques de homología de secuencia (I, II y III) se encuentran señalados. B) Conservación de los aminoácidos en TvBRF1 siguiendo el mismo código de color con respecto a levadura. C) Estructura de la RNA polimerasa III de levadura (gris), Bdp1 (morado), TBP (rojo), C34 en amarillo y Brf1 (verde).

Identificación de TvBDP1

Una búsqueda por homología en la base de datos de TrichDB usando las secuencias de BDP1 de humano y levadura permitieron identificar a un ortólogo de Bdp1 (TVAG_482180) en *T. vaginalis* anotado como *Factor de transcripción TFIIIb B* (Carlton *et al.*, 2007) y que planteamos que debe corresponder a TvBDP1. La expresión de TvBDP1 se verificó mediante la búsqueda de EST en la

base de datos TrichDB. El gen de Tv*bdp1* codifica para una proteína de 289 aminoácidos con un peso teórico de 33 kDa.

La proteína Bdp1 es una proteína altamente divergente en secuencia y tamaño (2624 aminoácidos para la secuencia canónica HsBDP1 y 594 aminoácidos para ScBdp1). Estas proteínas tienen un único dominio conservado conocido como dominio SANT, el cual está constituido por tres hélices α (Aasland *et al.*, 1996), aunque recientemente se ha descrito el dominio SANT extendido que incluye a las tres hélices α originales y las regiones C-terminal y N-terminal que flanquean este dominio, las cuales se conforman de otras dos hélices α y son esenciales en *S. cerevisiae* (Gouge *et al.*, 2017).

Un alineamiento de secuencia múltiple usando las secuencias de BDP1 de *Homo sapiens*, *S. cerevisiae, S. pombe* y *T. vaginalis,* muestran una región única de similitud que comprende el dominio SANT. Posteriormente, un análisis mediante la base de datos de INTERPRO, mostró que el dominio SANT se encuentra entre los residuos 131 y 179 de TvBDP1 (Figura 26A). Adicionalmente, el modelaje por homología para obtener la estructura terciaria de TvBDP1 usando como templado la estructura tridimensional obtenida por cristalografía de rayos X de levadura, muestra que esta proteína contiene el plegamiento típico que adopta el dominio SANT (Figura 26B). Curiosamente, TvBDP1 presenta un plegamiento de hélice α en el extremo C-terminal que se encuentra conservado en la estructura cristalográfica de levadura. Debido a la poca similitud en secuencia entre las proteínas Bdp1, no se realizó un ensayo de complementación para TvBDP1.

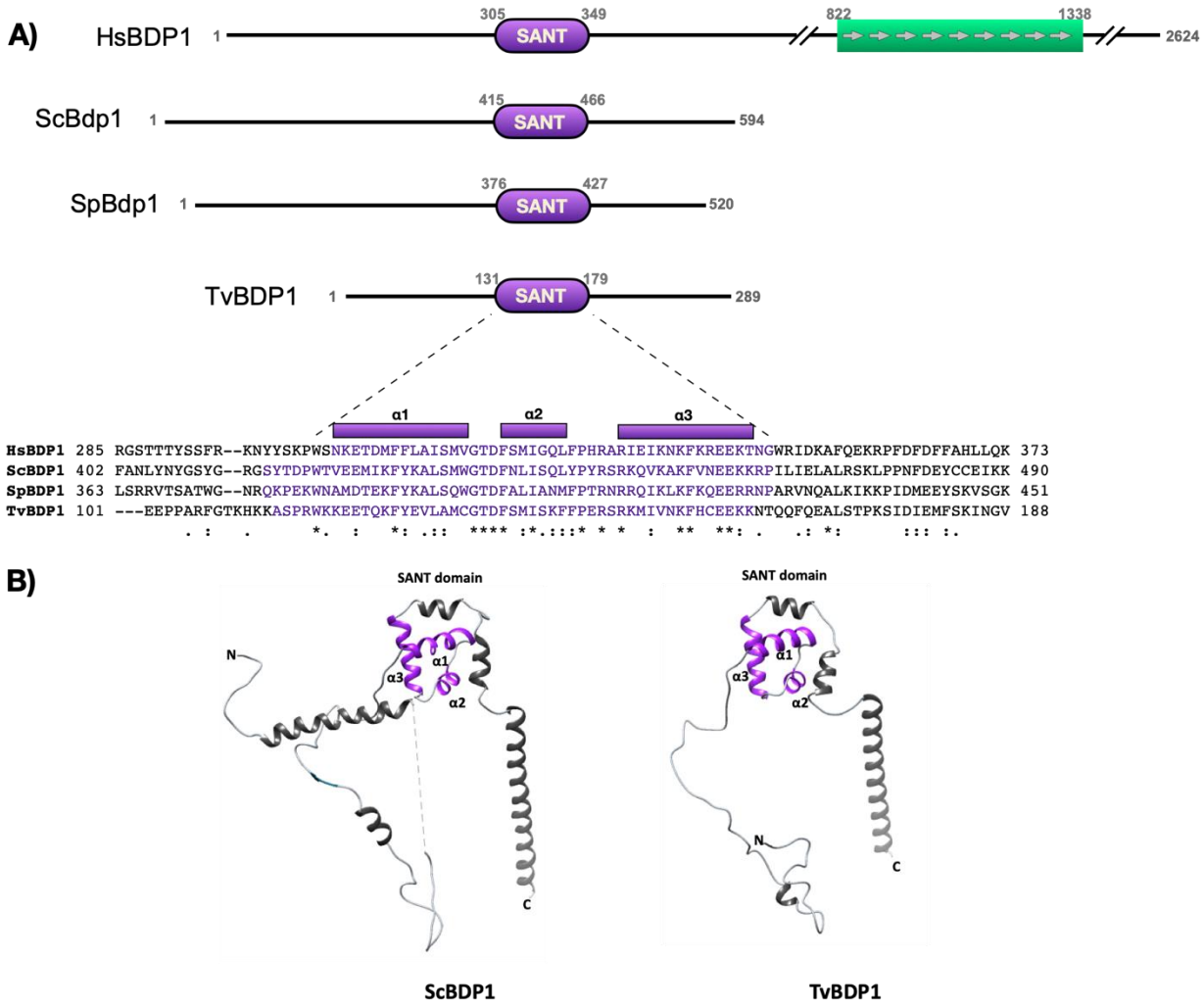


Figura 26.- Dominios y estructura predicha de TvBDP1. A) Dominios de BDP1: *H. sapiens* HsBDP1 (Q38AI0), *S. cerevisiae* ScBdp1 (P46678), *S. pombe* SpBdp1 (O94481) y *T. vaginalis* TvBDP1 (A2EBM7). El dominio SANT se representa en morado. HsBDP1 presenta 9 repetidos de secuencia (verde). B) Estructura predicha para TvBDP1 usando a ScBdp1 como modelo, el dominio SANT se representa en morado.

Interacción proteína-proteína evaluada por doble híbrido

Para probar la interacción entre las proteínas TvTBPs, TvBDP1 y TvBRF1, se realizó un análisis usando el método de doble híbrido en levadura. La interacción proteína-proteína se evaluó mediante la activación de los reporteros *LEU* y *lacZ*.

Brevemente, TvTBP1, TvTBP2 (M en C. Lluvia Rosas Hernández) y TvBDP1 se clonaron en el vector pEG202 que expresa a la proteína como una fusión al dominio de unión al DNA de LexA (BD, por sus siglas en inglés de *Binding Domain*) bajo el promotor constitutivo del gen *ADH1*

(alcohol deshidrogenasa 1). Adicionalmente, este plásmido contiene al gen *HIS3* que permite la selección de las células transformantes.

Por otro lado, TvBRF1 y TvBDP1 se clonaron en el vector pJG4-5 que expresa a la proteína como una fusión al dominio de activación B42 (AD, por sus siglas en inglés de *Activation Domain*) bajo el promotor *GAL1* (por lo que requiere que el medio contenga galactosa). Este plásmido contiene el gen *TRP1* para la selección de las transformantes.

Los plásmidos se transformaron junto con pSH18-34 (que contiene el gen reportero *lacZ*) a la cepa de levadura EGY48 [Mat *α*, *trp1*, *his3*, *ura3*, (lexAops-*LEU2*)], en la cual las secuencias de activación río arriba del gen *LEU2* (requerido para la biosíntesis de leucina) se ha reemplazado con 6 operadores *LexA*. El plásmido pSH18-34 contiene al gen *URA3* para la selección de las transformantes. La selección de las clonas transformantes se realizó en medio mínimo en ausencia de histidina, triptófano y uracilo según aplicara.

Para probar que el sistema de doble híbrido funciona adecuadamente se usaron como control positivo dos proteínas de levadura: α y β fusionadas al dominio de unión (BD) y al dominio de activación (AD) respectivamente. Como controles negativos se usaron: 1) endoquitinasa como proteína no relacionada, 2) los plásmidos pEG202-TvTBP1, pEG202-TvTBP2, pEG202-TvBDP1 en ausencia del activador transcripcional y los plásmidos pJG4-5-TvBDP1 y pJG4-5-TvBRF1 en ausencia del dominio de unión al DNA, los cuales no deberían producir colonias viables en ausencia de leucina y crecerán colonias blancas en el medio que contiene X-gal. Este control es importante para identificar si las proteínas de *Trichomonas* presentan autoactivación transcripcional.

Las células transformadas se cultivaron hasta alcanzar una densidad óptica de 0.6 y a partir de éstas se hicieron diluciones seriadas de 1:5, 1:25, 1:125. Se tomaron 3 µl de cada dilución para realizar los goteos en dos medios distintos: 1) Medio mínimo galactosa, en el cual se evalúa el reportero de leucina. En este medio se observa crecimiento cuando hay interacción proteína-proteína. 2) Medio mínimo galactosa, leucina y X-gal, en el que se evalúa el gen reportero *lacZ*. En este medio todas las células pueden crecer, pero las colonias son azules sólo cuando hay interacción proteína-proteína. En este medio también se puede observar la viabilidad de las células transformantes.

En las interacciones probadas para TvTBP1 se observa que TvTBP1/TvBRF1 activan el promotor de leucina y el promotor de *lacZ*, mientras que TvTBP1/TvBDP1 activa únicamente el promotor de leucina y no el de *lacZ* (Figura 27A).

Por otro lado, en los ensayos para TvTBP2, TvTBP2/TvBRF1 y TvTBP2/TvBDP1 activan solo el reportero de leucina y el crecimiento de las células transformantes es menor si se compara con el crecimiento obtenido con TvTBP1 (Figura 27B). TvTBP1(BD), TvTBP2(BD), TvBRF1(AD) y TvBDP1(AD) no presentan autoactivación transcripcional (Figura 27D y E). Estos resultados sugieren que tanto TvTBP1 como TvTBP2 pueden interactuar con TvBRF1 y TvBDP1; sin embargo, la interacción con TvTBP1 es más estable.

De los ensayos con TvBDP1 se pudo observar que esta proteína tiene actividad transcripcional por si sola (Figura 27C y D). A pesar de ello, en presencia de TvBRF1 hay mayor crecimiento de las células transformantes en medio mínimo galactosa, sugiriendo que existe una interacción TvBRF1-TvBDP1 (Figura 27C). Debido a la activación que presenta TvBDP1 junto al dominio de unión al DNA, se intentó probar la interacción inversa, es decir, TvBRF1en el plásmido pEG202 y TvBDP1 en el plásmido pJG4-5. Sin embargo, esta combinación no se pudo usar debido a que TvBRF1 junto al dominio de unión al DNA, presenta una mayor actividad transcripcional (datos no mostrados).

Como dato curioso de estos ensayos, se observa que TvTBP1/TvBRF1 activa el reportero de *lacZ* pero otras interacciones como TvTBP1/TvBDP1 que activan el reportero de leucina, no lo hacen con el reportero *lacZ*. Los sistemas reporteros usados aquí no son del todo equivalentes en el sentido de que uno se encuentra integrado en el genoma (reportero de leucina), mientras que el otro es episomal (reportero de *lacZ*). Adicionalmente, el reportero de leucina en la cepa EGY48 contiene seis operadores de *LexA*, mientras que el reportero de *lacZ* posee ocho (Estojak *et al.*, 1995), sin embargo, para algunas proteínas, el reportero de *LEU2* en la cepa EGY48 es más sensible que el reportero *lacZ* del plásmido pSH18-34 (Golemis *et al.*, 2000), por lo que las diferencias observadas en la activación de ambos reporteros no son del todo inesperadas.

La activación del reportero *lacZ* es cuantificable, pudiendo de este modo, determinar la afinidad de las interacciones proteína-proteína probadas en este ensayo. El análisis se realizó mediante un ensayo colorimétrico a través de la cuantificación del ONPG (Orto-nitrofenil-β-galactósido) hidrolizado por la acción de la β-galactosidasa. Los extractos de proteínas se obtuvieron de las células transformantes crecidas a una densidad óptica de 0.6. Un volumen determinado del extracto se puso en presencia del ONPG y se incubó a 30ºC. Los datos obtenidos de este ensayo se muestran en la tabla del anexo II. Una fuerte actividad de la β-galactosidasa se detectó en las células transformantes con TvTBP1-TvBRF1, indicando que la interacción entre estas proteínas es muy afín (Figura 27).

Los ensayos de doble híbrido y la medición de la actividad de β-galactosidasa sugieren que TvTBP1 interacciona con mayor afinidad con las proteínas del complejo TFIIIB (TvBRF1 y TvBDP1) que TvTBP2 y que por tanto tiene más probabilidad TvTBP1 de formar parte del complejo TFIIIB en *T. vaginalis*.

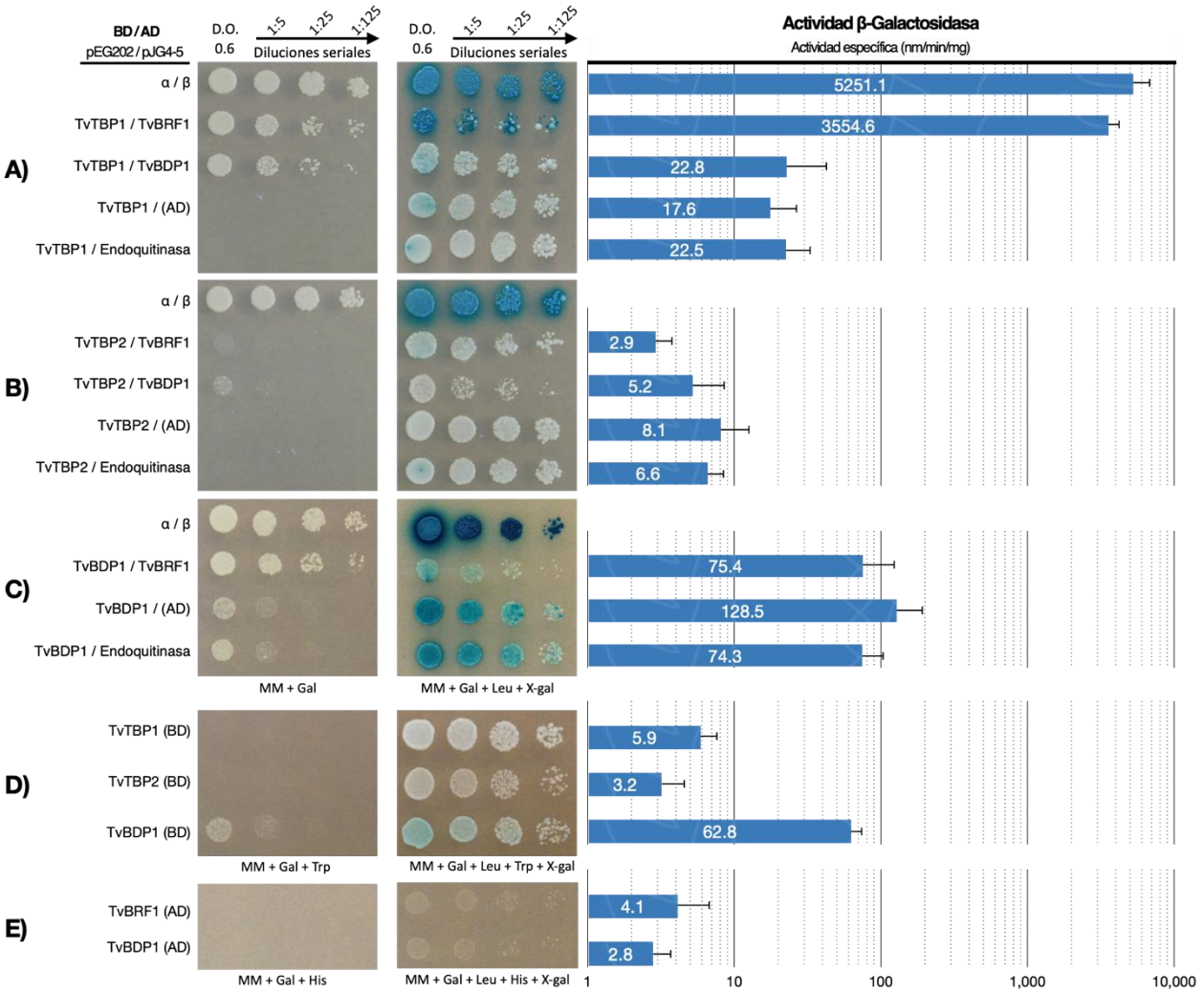


Figura 27.- Ensayo de doble híbrido para probar la interacción entre las TvTBPs, TvBDP1 y TvBRF1. A) Ensayo para TvTBP1. B) Ensayo para TvTBP2. C) Ensayo para TvBDP1. D) Controles de TvTBP1, TvTBP2 y TvBDP1 en pEG202 sin el plásmido pJG4.5. E) Controles de TvBRF1 y TvBDP1 en pJG4.5 sin pEG202. Las proteínas de levadura α y β fusionadas al dominio de unión (BD) y al dominio de activación (AD) se usaron como control positivo. Endoquitinasa se usó como control negativo. Los goteos se realizaron por triplicado y se muestra un ejemplo representativo de cada interacción. Las barras de la gráfica reflejan la actividad específica de la β-galactosidasa usando ONPG como sustrato. Los valores de la actividad específica de la β-galactosidasa son expresados en nm/min/mg de proteína y representan el promedio de al menos tres ensayos independientes, la barra indica la desviación estándar.

Ensayos de coprecipitación de las proteínas del complejo TFIIIB

El ensayo por doble híbrido permitió tener un panorama general de las interacciones proteicas del complejo TFIIIB, no obstante, una confirmación por otro método era necesaria. De modo complementario, el siguiente ensayo de coprecipitación permitió analizar las interacciones entre dos y tres proteínas al mismo tiempo.

Para este ensayo, las proteínas GST-BRF1 o GST sola (como control negativo), se acoplaron a la matriz de glutatión sefarosa para probar su interacción con 6xHis-TvTBP1, 6xHis-TvTBP2 y 6xHis-TvBDP1 tal como se muestra en el esquema (Figura 28A). Para ello, se requirieron realizar 12 ensayos de manera simultánea. Las muestras del lisado celular (extracto total, líneas 1 y 4) y de la coprecipitación se analizaron por SDS-PAGE, tinción con azul de Coomassie y por inmunoblot usando una sonda (HisProbe-HRP) contra la etiqueta de histidinas.

El ensayo de coprecipitación para probar la interacción entre TvBDP1 y TvBRF1 se muestra en ambos geles (Figura 28B líneas 1, 2 y 3 y C líneas 9, 10 y 11). En las líneas 1 y 9 del inmunoblot (que corresponde al extracto total de TvBDP1), se observan dos señales, la de mayor peso molecular corresponde en peso a la proteína completa de TvBDP1 de 71.6 kDa. Una señal de menor peso molecular (aproximadamente 20 kDa), parece ser un producto de la degradación de TvBDP1. A pesar de que la interacción directa entre las proteínas TvBDP1 y TvBRF1 fue observada por ensayos de doble híbrido (Figura 27C), por este método no se obtuvo el mismo resultado (Figura 28B línea 3 y 20C, línea 11). En *S. cerevisiae* y humanos, Bdp1 se encuentra débilmente asociado con TBP y Brf1, esta unión se estabiliza solo en presencia de DNA (Schramm *et al.*, 2000), por lo que es probable que las diferencias observadas en la interacciones entre TvBDP1 y TvBRF1 se deba a la ausencia de DNA en los ensayos de coprecipitación. Los resultados sugieren que estas proteínas se comportan de manera similar a las de eucariontes superiores.

En los ensayos de interacción entre TvBRF1-TvTBP1 (Figura 28B, líneas 4, 5 y 6) y TvBRF1-TvTBP2 (Figura 28C, líneas 12, 13 y 14) se usó mayor concentración del lisado para los ensayos con TvTBP2, razón por lo cual se observa una mayor carga de proteínas en esta línea (Figura 28B línea 4 y C, línea 12). No obstante, la señal de aproximadamente 26 kDa correspondiente a la coprecipitación es más intensa para TvTBP1 que para TvTBP2 (Figura 28B línea 6 y C, línea 14). Los resultados sugieren que la interacción entre TvTBP1/TvBRF1 es más afín que para TvTBP2/TvBRF1. Estos datos concuerdan con los obtenidos por doble híbrido.

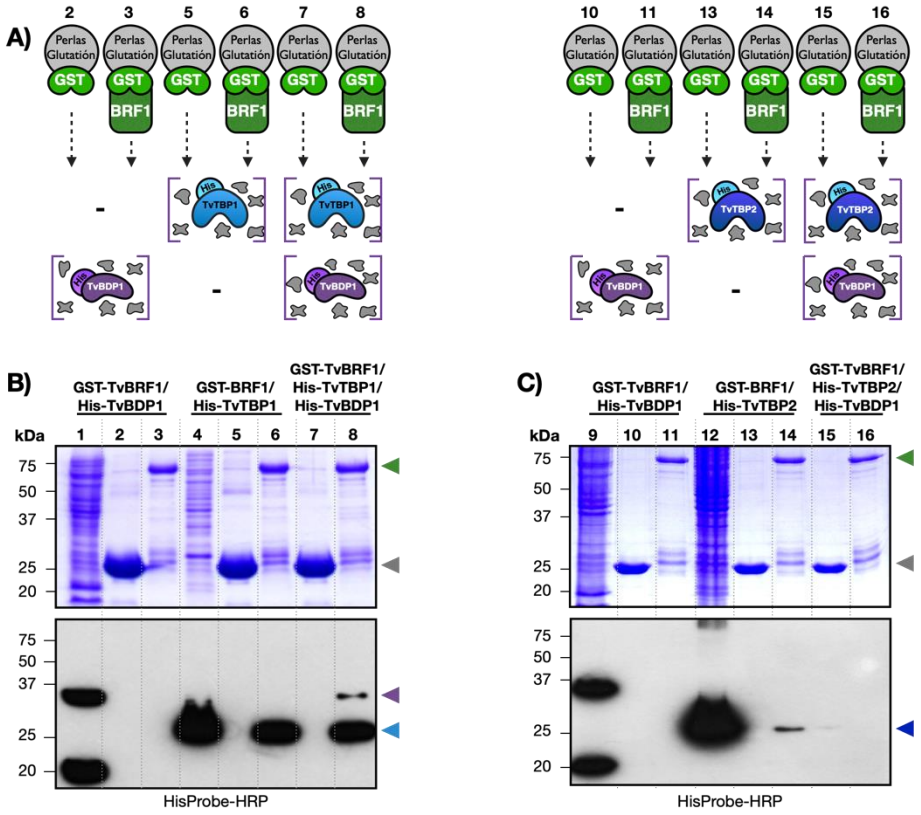


Figura 28.- Ensayo de coprecipitación para probar la formación del complejo TFIIIB. A) Representación esquemática de los ensayos de coprecipitación. La proteína recombinante GST-BRF1 o GST sola (control negativo) se usó como carnada para interactuar con los lisados de *E. coli* expresando las proteínas recombinantes 6xHis-TvTBP1, 6xHis-TVTBP2 o 6xHis-BDP1 solas o en combinación. Los números en el esquema indican el número del carril donde se cargaron las muestras. B) Ensayos de coprecipitación para TvTBP1. El orden de los ensayos es como sigue: Coprecipitación GST-BRF1 con 6xHisBDP1: el extracto de proteínas de *E. coli* expresando a 6xHis-BDP1 (línea 1) se incubó con las perlas unidas a GST (línea 2) o GST-BRF1 (línea 3). Coprecipitación GST-BRF1 con 6xHisTvTBP1: el extracto de proteínas de E. coli expresando a 6xHis-TvTBP1 (línea 4) se incubó con las perlas unidas a GST (línea 5) o GST-BRF1 (línea 6). Coprecipitación GST-BRF1 con 6xHis-TvTBP1 y 6xHis-TvBDP1: los extractos de proteínas de *E. coli* expresando a 6xHis-TvTBP1 y 6xHis-TvBDP1 se incubaron con las perlas unidas a GST (línea 7) o GST-BRF1 (línea 8). C) Ensayos de coprecipitación para TvTBP2. El orden de los ensayos es como sigue: Coprecipitación GST-BRF1 con 6xHisBDP1: el extracto de proteínas de *E. coli* expresando a 6xHis-BDP1 (línea 1) se incubó con las perlas unidas a GST (línea 2) o GST-BRF1 (línea 3). Coprecipitación GST-BRF1 con 6xHisTvTBP2: el extracto de proteínas de *E. coli* expresando a 6xHis-TvTBP2 (línea 4) se incubó con las perlas unidas a GST (línea 5) o GST-BRF1 (línea 6). Coprecipitación GST-BRF1 con 6xHis-TvTBP2 y 6xHis-TvBDP1: los extractos de proteínas de *E. coli* expresando a 6xHis-TvTBP2 y 6xHis-TvBDP1 se incubaron con las perlas unidas a GST (línea 7) o GST-BRF1 (línea 8). Se muestran los geles teñidos con azul de Coomassie (panel superior) y el ensayo de inmunoblot con la sonda HisProbe-HRP que revela la presencia de las proteínas con etiqueta de histidina (panel inferior). La posición de las proteínas se encuentra

señalada en los geles según su peso molecular; GST-BRF1 de71.6 kDa (triángulo verde), GST (26 kDa), 6xHis-TvTBP1 (triángulo azul claro) y 6xHis-TvTBP1 (triángulo azul oscuro) de aproximadamente 26 kDa y 6xHis-TvBDP1 de 32.9 kDa (triángulo morado).

Finalmente, para probar la formación de un complejo entre las tres proteínas; TvBRF1 se incubó al mismo tiempo con los extractos celulares que contenían a las TvTBPs y a TvBDP1 (Figura 28B, líneas 7 y 8; C, líneas 15 y 16). Dos señales, que corresponden en peso molecular con TvTBP1 y TvBDP1 se observan en el ensayo de coprecipitación con TvBRF1 (Figura 28B, línea 8). Particularmente con TvTBP2, no se observa ninguna señal (Figura 28C, línea 16). La interacción entre TvTBP2 y TvBRF1 es de por si débil (Figura 28C, línea 14), por lo que es probable que, al adicionar una mayor cantidad de volumen, correspondiente al lisado de TvBDP1, la señal se haya perdido. Los resultados de esta parte del ensayo resultan muy interesantes; por un lado, observamos lo mismo que en los dobles híbridos, la interacción entre TvTBP1-TvBRF1 es más afín que la que se puede dar entre TvTBP2-TvBRF1. Por otro lado, se observa una interacción entre las proteínas TvTBP, TvBRF1 y TvBDP1, lo que sugiere la formación de un complejo TFIIIB en *T. vaginalis* y la participación de las tres proteínas sugiere una misma función biológica. Adicionalmente, la interacción entre estas tres proteínas parece estar dada a través de TvTBP. Recientemente se obtuvo la estructura cristalográfica del complejo TFIIIB-DNA de *S. cerevisiae* (Figura 29), en esta se puede observar que TBP (rosa) interactúa con los repetidos de ciclina (verde) y el C-terminal (naranja) de ScBrf1 por un lado y por el otro con ScBdp1 (morado) a través del dominio SANT. Esta estructura coindice con los resultados observados para el complejo TFIIIB de *T. vaginalis* y explica por qué se requiere de TBP para observar a las tres proteínas en el ensayo de coprecipitación, ya que la interacción Brf1-Bdp1, es débil y poco estable.

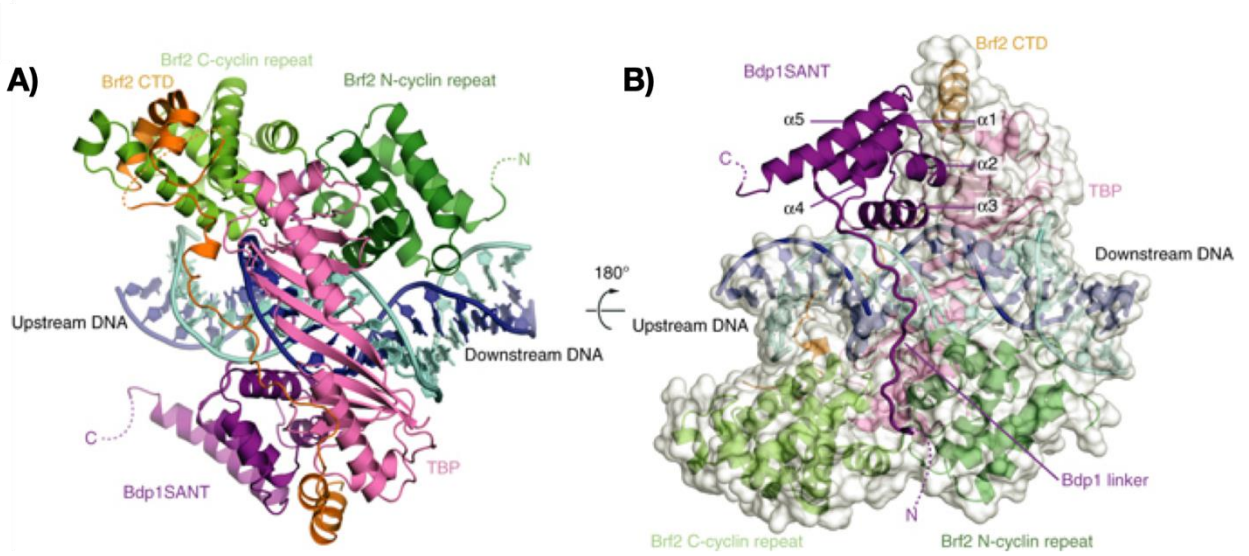


Figura 29.- Estructura cristalográfica del complejo Brf2-TBP-Bdp1-DNA de *S. cerevisiae*.A) vista superior. B) Vista inferior. La hebra templado y no templado se ilustran en color azul oscuro y azul claro respectivamente. TBP (rosa), Bdp1 (morado), los repetidos de ciclina (verde) y el CTD (naranja) de Brf1 (Gouge et al., 2017).

El complejo TFIIIB se une a los promotores 5S y U6

Diversos estudios han demostrado que el complejo TFIIIB se ensambla sobre las tres clases de promotores transcritos por la RNA polimerasa III y que Bdp1 es requerido para estabilizar el complejo al DNA. En esta parte, nos enfocamos en estudiar la interacción de los diferentes componentes del complejo TFIIIB al DNA mediante ensayos de retardo de la movilidad electroforética.

Las proteínas individuales (GST-BRF1, 6xHis-TvTBP1, 6xHis-TvTBP2 y 6xHis-TvBDP1) se purificaron por columna de afinidad. Mientras que, las distintas combinaciones de proteínas (GST-BRF1/6xHis-TvTBP1, GST-BRF1/6xHis-TvTBP2, GST-BRF1/6xHis-TvBDP1, 6xHis-TvBDP1/6xHis-TvTBP1, 6xHis-TvBDP1/6xHis-TvTBP2 y GST-BRF1/6xHis-TvTBP1/6xHis-TvBDP1, GST-BRF1/6xHis-TvTBP2/6xHis-TvBDP1) se obtuvieron mediante coprecipitación usando la matriz de níquel. Las proteínas recién purificadas se incubaron con las sondas 5S y U6 de *T. vaginalis* marcadas radioactivamente por PCR usando [α-^{32}P]ATP.

La sonda del promotor 5S de *T. vaginalis* de 186 pb contiene elementos río arriba del sitio de inicio de la transcripción, específicamente el UPE (*Upstream Promoter Element*) y un elemento con similitud al SSE (*Start Site Element*) (Figura 30A). En levaduras, estos elementos participan en la unión del factor TFIIIB (Lee *et al.*, 1997). A su vez, el promotor del gen U6 de *T. vaginalis* contiene

un elemento con similitud a una caja TATA cercano al sitio de inicio de la transcripción en la posición 28 (Figura 30B) (Simoes-Barbosa *et al.*, 2008). Ambos promotores son regiones ricas en A/T, por lo que no es extraño encontrar secuencias con similitud a una caja TATA rio arriba de los elementos descritos.

En general, los resultados del ensayo para la sonda 5S y U6 son bastante similares y se resumen en los siguientes puntos: 1) TvTBP1 forma un retardo estable con ambas sondas (5S y U6) (Figura 30, línea 2), mientras que TvTBP2 retarda ambas sondas, pero su unión con el DNA no parece estable, ya que se observa un barrido a través del carril (Figura 30, línea 10). Este resultado es el mismo que observamos en el capítulo anterior. 2) TvBRF1 no se une al DNA, al menos por sí sola (Figura 30, líneas 3 y 11). La principal función de Brf1 es la de formar una conexión entre TFIIIB y la RNA polimerasa III. Aunque, recientemente se describió que un dominio de ScBrf1 conocido como la hélice "*lid*" que se encuentra entre el bloque de homología I y II tiene elementos de unión al DNA (Vorländer *et al.*, 2018), está hélice no está conservada en TvBRF1. 3) TvBDP1 se une al DNA. Sin embargo, al igual que TvTBP2 con el DNA, parece formar un complejo inestable DNA-proteína ya que se observa un barrido en este carril (Figura 30, líneas 4 y 12). Aunque este barrido sólo ocurre en la mitad del carril, a diferencia de TvTBP2, no tiene que ver con el peso molecular de las proteínas (TvBDP1 = 32.9 kDa y TvTBP2 = 26 kDa) y puede deberse a otros efectos como la carga o la conformación del complejo DNA-proteína (Lane *et al.*, 1992). Este resultado también puede deberse a que TvBDP1 no es estable en solución y se degrada rápidamente. 4) La unión del complejo TvTBP1-TvBRF1 al DNA genera lo que parece ser un superretardo (Figura 30, línea 5), ya que la señal se observa arriba de la obtenida con TvTBP1-DNA (Figura 30, línea 2) y es más intensa. A su vez, la unión del complejo TvTBP2-TvBRF1 al DNA parece ser más estable que cuando está TvTBP2 sola (Figura 30, línea 13). 5) La unión de TvBDP1 a cualquiera de las proteínas TvTBPs es suficiente para formar un complejo estable, ya que se lleva toda la sonda libre y la señal del retardo es la más intensa (Figura 30, líneas 6, 7 y 15). Nota: en la línea 14 correspondiente a TvBDP1-TvTBP2, no se ve ningún retardo, ni sonda libre. Una posibilidad es que la muestra no se cargara adecuadamente en ambos pozos, sin embargo, no es tan probable. Otra posibilidad es que durante la purificación de TvTBP2, en un intento por concentrar la muestra, se copurificara alguna impureza que degradó la muestra de DNA. 6) Para la línea correspondiente a TvBRF1 y TvBDP1 (Figura 30, líneas 8 y 16) no se observa diferencia con respecto a TvBDP1-DNA (Figura 30, líneas 4 y 12), respaldando los resultados obtenidos por coprecipitación.

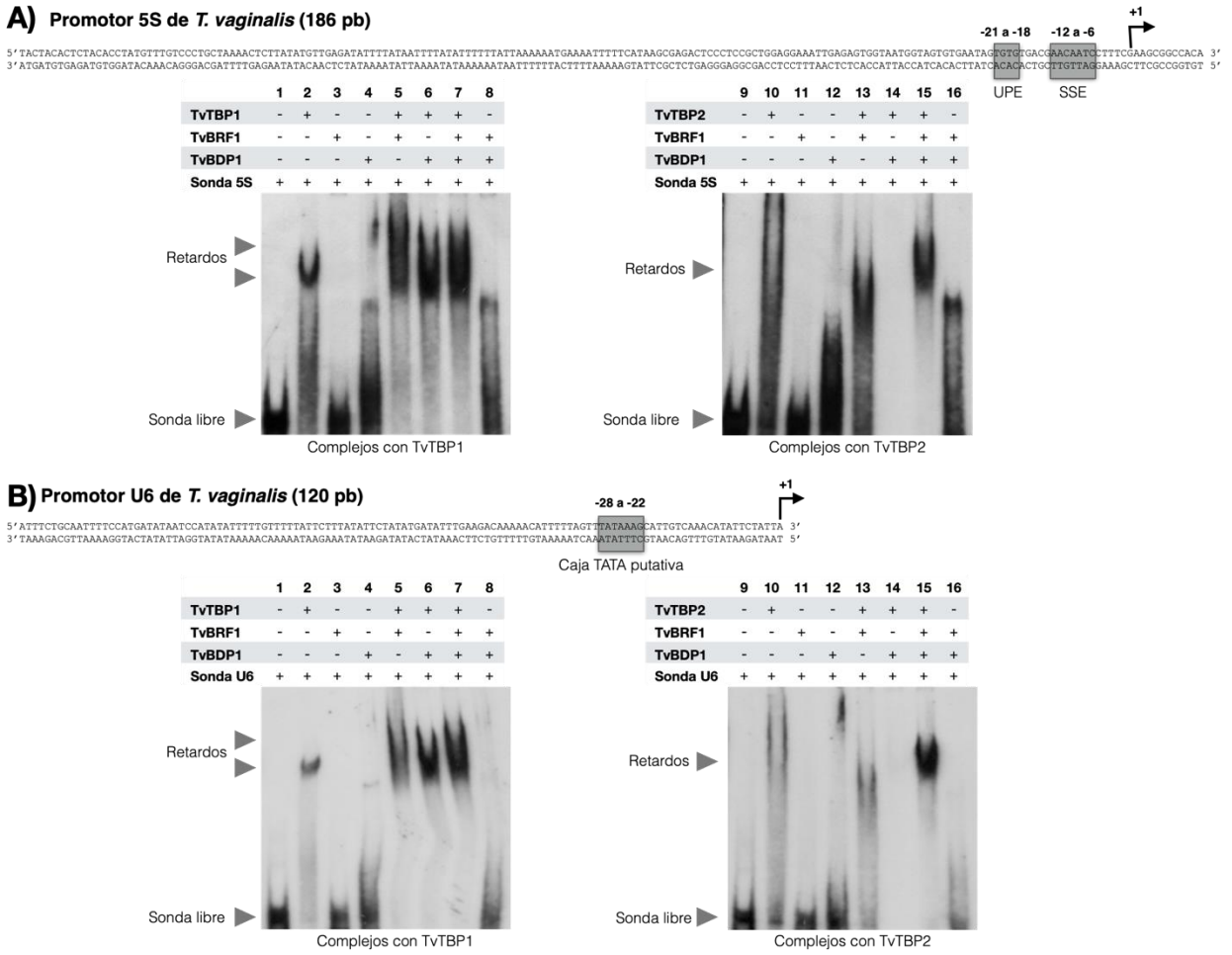


Figura 30.- Ensayos EMSA para probar la unión del complejo TFIIIB al promotor 5S y U6 de *T. vaginalis.* A) En el panel superior se muestra la secuencia de la sonda del promotor 5S de 186 pb que contiene los elementos UPE (*Upstream Promoter Element*) y el SSE (*Start Site Element*) marcados en gris. El sitio de inicio de la transcripción se muestra con una flecha. Los ensayos EMSA realizados con la sonda 5S se muestran en el panel inferior. B) En el panel superior se muestra la sonda del promotor U6 de 120 pb que contiene una caja putativa en la posición -28 (gris). Los ensayos EMSA realizados con la sonda U6 se muestran en el panel inferior. Cada carril es indicado con un número. La presencia de proteína se indica con el símbolo (+) y la ausencia con el símbolo (-).

Para probar el efecto de la estabilidad en los complejos formados, se decidió repetir el ensayo con la sonda U6 bajo las mismas condiciones, excepto que las proteínas purificadas el día anterior se usaron en el ensayo (Figura 30). En esta figura se observa que el retardo formado por TvTBP1 se ve bastante afectado (línea 2). TBP suele formar dímeros en solución, motivo por el cual, limita la disponibilidad de la proteína libre para unirse al DNA. Mientras que la señal del retardo formado por TvBDP1 no se ve afectado en intensidad, pero se ve más debajo de lo esperado (línea 4). TvBDP1 se degrada (Figura 31) y es probable que uno de los fragmentos se siga uniendo a la sonda.

Los complejos TvBRF1-TvTBP1-TvBDP1, TvBRF1-TvTBP1 y TvBDP1-TvTBP1 (líneas 5, 6 y 7), se conservan casi idénticos, siendo este último el más estable. Estos datos sugieren, de manera indirecta, que realmente se está formando un complejo y que la interacción entre estas proteínas les da cierta estabilidad en solución.

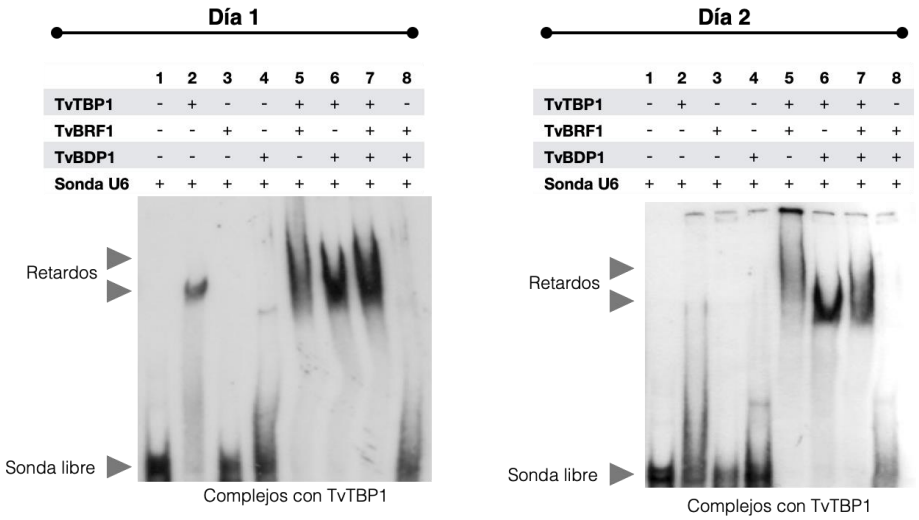


Figura 31.- Estabilidad del complejo TFIIIB con un promotor U6. Las proteínas recién extraídas (Día 1) y con un día de antigüedad (Día 2) se compararon en ensayos de retardo con la sonda U6.

En el capítulo anterior, observamos que las TBPs de *T. vaginalis* se unen al DNA de manera independiente de secuencia, por lo que es muy probable que otros factores sean los responsables del adecuado reconocimiento del sitio de inicio de la transcripción. Adicionalmente, los resultados con la proteína TvBDP1 nos sugieren que la interacción con TvTBP resulta en un complejo estable. Por consiguiente, nos preguntamos si TvTBP-TvBDP1 o el complejo TFIIIB (TvBDP1-TvTBP-TvBRF1) son capaces de reconocer específicamente un sitio en el promotor.

Para probar esta hipótesis, realizamos un ensayo EMSA dividiendo las regiones más importantes del promotor 5S de *T. vaginalis* en sondas de 60-80 pb (Figura 32A). La sonda A contiene una región rica en A/T, regiones reguladoras de levadura (UPE, SSE) y el sitio de inicio de la transcripción. La sonda B contiene el sitio de unión a TFIIIC, la caja A y el elemento intermedio y finalmente la sonda C contiene la caja C y la secuencia de terminación de la transcripción. Las con Las proteínas TvTBP1, TvTBP2, TvBDP1 individuales o el complejo TFIIIB (TvBRF1-TvTBP-TvBDP1) obtenido por coprecipitación, se incubaron con las sondas. TvTBP1 se une a todas las sondas probadas, aunque con la sonda A no se alcanza a definir una señal, casi no queda sonda libre (Figura 32, línea 2). Para TvTBP2 no se define una señal, pero se alcanza a ver menos sonda

libre, sugiriendo que TvTBP2 se une a todas las sondas. Es probable que por la cantidad de DNA usada en el ensayo no se alcance a ver el barrido a través del carril como sucede con esta proteína en anteriores ensayos (Figura 32, línea 6). Una señal bien definida correspondiente a la unión de TvBDP1 al DNA se observa en todas las sondas probadas (Figura 32, carril 3 y 7). La presencia de TvBRF1 en el complejo TFIIIB, tampoco hace un cambio, ya que TFIIIB retarda las tres sondas, ya sea que el complejo esté formado con TvTBP1 o con TvTBP2(Figura 32, carril 4 y 8). Estos datos sugieren que la presencia de TvBDP1 o TvBRF1 en el complejo no le otorgan especificidad de unión al DNA.

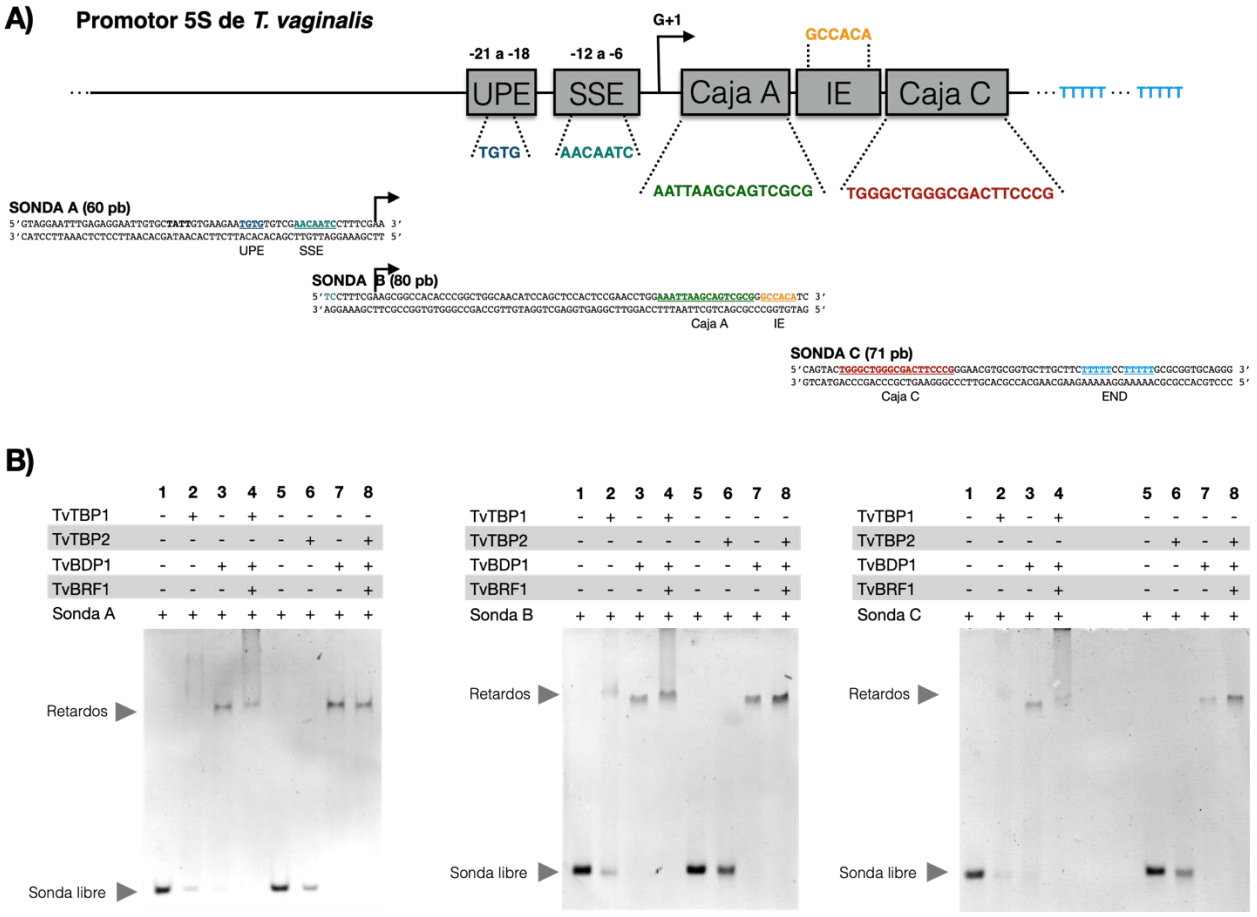


Figura 32.- Unión del complejo TFIIIB a distintas regiones del promotor 5S. El promotor 5S se muestra en el panel superior y contiene el elemento UPE, el elemento SSE, la caja A, el elemento intermedio (IE) y la caja C. La sonda A de 60 pb contiene a los elementos UPE y SSE, la sonda B contiene a la caja A y al elemento intermedio y la sonda C contiene a la caja C y el elemento de terminación. En el panel inferior, se muestran los EMSAS con las proteínas TvBDP1, TvTBP1 y TvTBP2 individualmente o en complejo. Cuando se indica la presencia de las tres proteínas significa que se agregó el complejo obtenido por coprecipitación y no las proteínas individuales. El gel se tiñó con bromuro de etidio y los colores de la imagen se muestran invertidos.

En un último intento por identificar el sitio de unión del complejo TFIIIB en el promotor 5S, se obtuvieron sondas de 40 pb mediante aliniamiento de oligonucleótidos. Uno de los oligonucleótidos sin alinear fue usado como control negativo, ya que en teoría, TBP no se une a DNA de cadena sencilla. La sonda completa del promotor 5S de 200 pb se usó como control positivo (Figura 33A). Por otro lado se obtuvo el complejo TFIIIB (TvBDP1-TvTBP1-TvBRF1) mediante coprecipitación y se incubó con las sondas. La sonda que contiene a todos los elementos del promotor 5S es retardada por el complejo TFIIIB (Figura 33B, línea2), mientras que la sonda de cadena sencilla no se retarda (Figura 33B, línea 10). El ensayo de retardo muestra que el complejo TFIIIB se una a todas las sondas de doble cadena de 40 pb (Figura 33B, líneas 4 a 9). Por este método fue imposible determinar el sitio de unión de TFIIIB al promotor 5S de *T. vaginalis* y es probable que se requiera de otros factores, como el complejo TFIIIC, para el adecuado reconocimiento del sitio de inicio de la transcripción. Los ensayos usando extractos nucleares de *T. vaginalis* no son una opción, ya que este ensayo se ha intentado en un trabajo previo usando distintas condiciones de unión, sin éxito (Torres Machorro, 2010). El problema de usar extractos totales o nucleares de *T. vaginalis* consiste en que este organismo contiene una gran cantidad de proteasas y nuclesas, haciendo muy dificil obtener extractos íntegros a pesar de las precauciones tomadas. Por otro lado, utilizando ensayos de transcripción *in vitro*, se pudo determinar que el promotor del gen 5S es funcional. Posteriormente, se intentó probar las diferentes regiones del promotor 5S mediante mutagénesis dirigida, pero esta tarea no se puso completar por problemas técnicos (Torres Machorro, 2010).

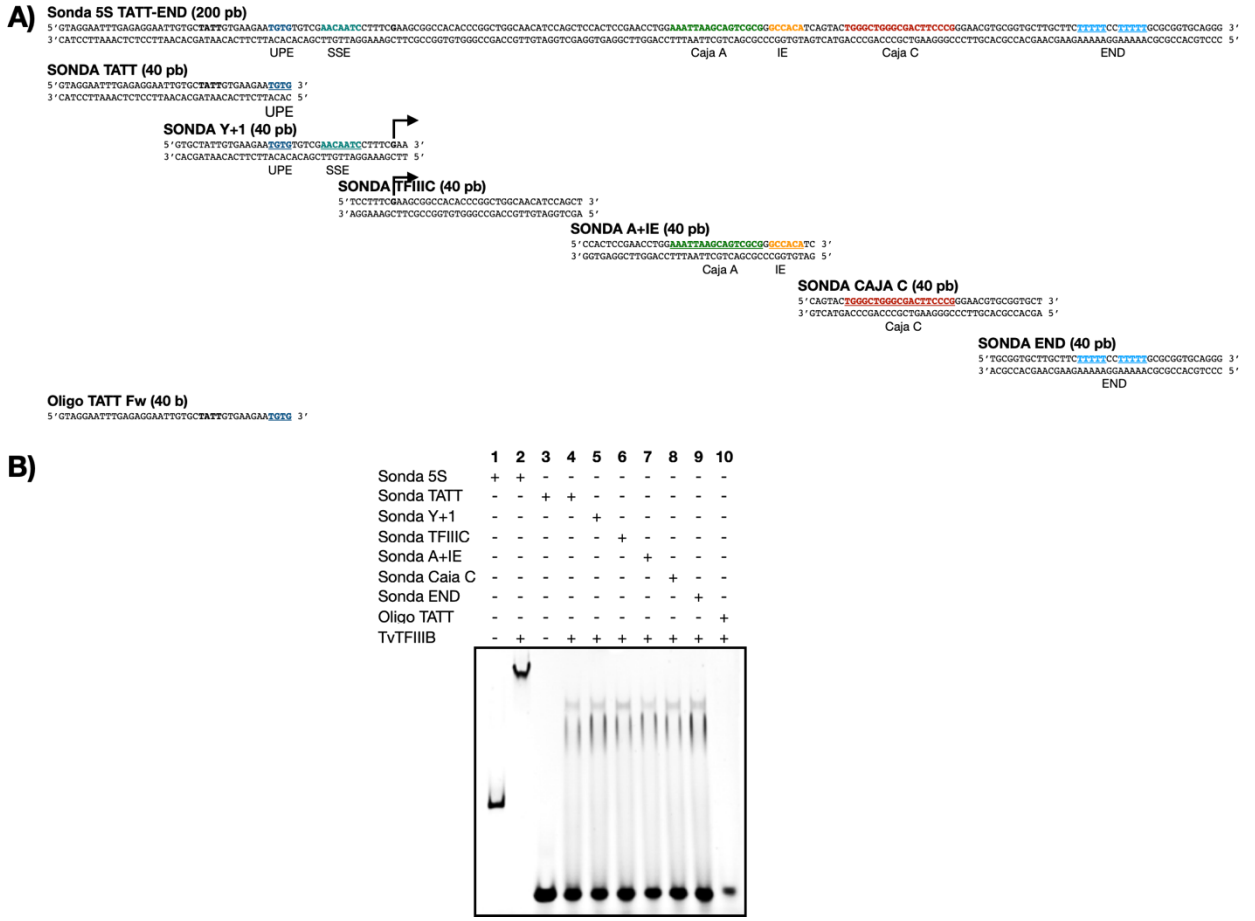


Figura 33.- Mapeo de la unión de TFIIIB al promotor 5S. A) Secuencias de las sondas. La sonda 5STATT-END de 200 pb contiene todos los elementos del promotor 5S, la sonda TATT de 40 pb que contiene una región rica en A/T y al elemento UPE, la sonda Y+1 de 40 pb que contiene los elementos UPE y SSE, la sonda TFIIIC de 40 pb que contiene el sitio de inio de la transcripción y el elemento de unión del complejo TFIIIC, la sonda A+IE de 40 pb de contienen a la caja A y al elemento intermedio, la sonda caja C de 40 pb, la sonda END de 40 pb y una secuencia de cadena sencilla de 40 pb (el oligonucleótido TATT). En el panel inferior se muestran los EMSAS con el complejo TFIIIB (TvBDP1-TvTBP1-TvBRF1) obtenido por coprecipitación. El gel se tiñó con bromuro de etidio y los colores de la imagen se muestran invertidos.

Conclusiones

Los resultados obtenidos sugieren la formación de un complejo TFIIIB en *T. vaginalis*. En el cual, las proteínas TvBRF1 y TvBDP1 se asocian preferencialmente con TvTBP1. La formación del complejo con TvTBP2 es posible; sin embargo, parece ser una interacción más débil. Es importante determinar cuál de las dos proteínas forma parte del complejo *in vivo* en *T. vaginalis*.

Por otro lado, la presencia de TvTBP es necesaria para la formación del complejo, debido a que, es TvTBP quien forma interacciones con TvBDP1 y TvBRF1. Los datos sugieren que, en ausencia de esta proteína, no se puede formar el complejo.

La unión del complejo TFIIIB a los promotores transcritos por la RNA polimerasa III está dada por TvTBP y TvBDP1, ya que en los ensayos de retardo observamos que TvBRF1 no muestra ninguna actividad de unión. Sin embargo, el sitio de unión de TFIIIB aún se debe determinar. En la figura 34, se muestra un modelo de los resultados obtenidos. El promotor U6 ha sido usado como ejemplo, sin embargo, esta unión también puede darse con el promotor del gen 5S.

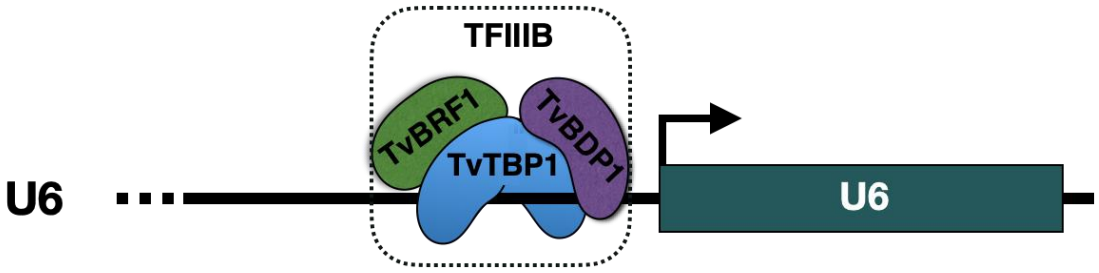


Figura 34.- Modelo hipotético de la formación del complejo TFIIIB en *T. vaginalis* ensamblado sobre el promotor U6. TvTBP1 interactúa con TvBRF1 y con TvBDP1 para ensamblar al complejo TFIIIB. Las interacciones proteína-proteína se dedujeron de los ensayos de doble híbrido y coprecipitación, y las interacciones con el DNA, se dedujeron por ensayos EMSA.

Perspectivas

1.- Determinar si el complejo TFIIIB se une de manera secuencia-específica a los promotores de los genes 5S y U6. Por ensayos de retardo, no fue posible determinar el sitio de unión para el promotor 5S. Según lo descrito en la literatura, el promotor 5S es previamente reconocido por TFIIIA y por TFIIIC. Por lo que existe la posibilidad de que el reconocimiento secuencia específica para el promotor 5S esté otorgado por otra proteína.

Por otro lado, en levadura, el complejo TFIIIB es el responsable del reconocimineto del sitio de inicio de la transcripción en el promotor del gen U6, por lo que sería interesante intentar mapear por ensayos EMSA el sitio de unión de este complejo.

Un ensayo de *southwestern*, nos permitiría confirmar los datos obtenidos por EMSA para el promotor 5S y a la vez mapear la interacción de TFIIIB con el promotor del gen U6. Otras técnicas también pueden ser empleadas, sin embargo, este método ya se ha probado en el laboratorio y es relativamente sencillo de realizar.

2.- Estudiar *in vivo* si TFIIIB participa en las tres clases de promotores transcritos por la RNA polimerasa III. Para este ensayo se sugiere realizar un ChIP, usando un anticuerpo contra cada una de las proteínas del complejo TFIIIB. De esta manera sería posible conocer si las tres proteínas están unidas a la mima región de DNA.

3.- Determinar *in vivo*, si TvTBP1 o TvTB2 forman parte del complejo TFIIIB. Ya que existen dos panoramas de este punto: 1) *in vivo*, sólo una de las proteínas forma parte del complejo TFIIIB y 2) de manera similar a *Drosophila*, existe una una pariticipación diferencial entre ambas proteínas, en la que TvTBP1 participa en la transcripción de algunos promotores y TvTBP2 en otros.

CAPÍTULO IV: Otros factores asociados a TBP en complejos que se unen a promotores transcritos por la RNA polimerasa I y II en *T. vaginalis*

Introducción

El inicio de la transcripción por la RNA polimerasa II, depende de TFIIB. Esta proteína forma un puente entre la RNA polimerasa II y el promotor (Kostrewa *et al.*, 2009). La unión de TFIIB al promotor unido con TBP, resulta en un complejo ternario TFIIB-TBP-DNA más estable. TFIIB tiene un papel importante para el reclutamiento de la RNA polimerasa y la selección del sitio de inicio de la transcripción (Thomas y Chiang, 2006). En humanos, TFIIB es una proteína de 316 aminoácidos, mientras que en levaduras es una proteína de 345 aminoácidos codificada por el gen *SUA7*. TFIIB se encuentra conservada evolutivamente entre especies desde arqueas.

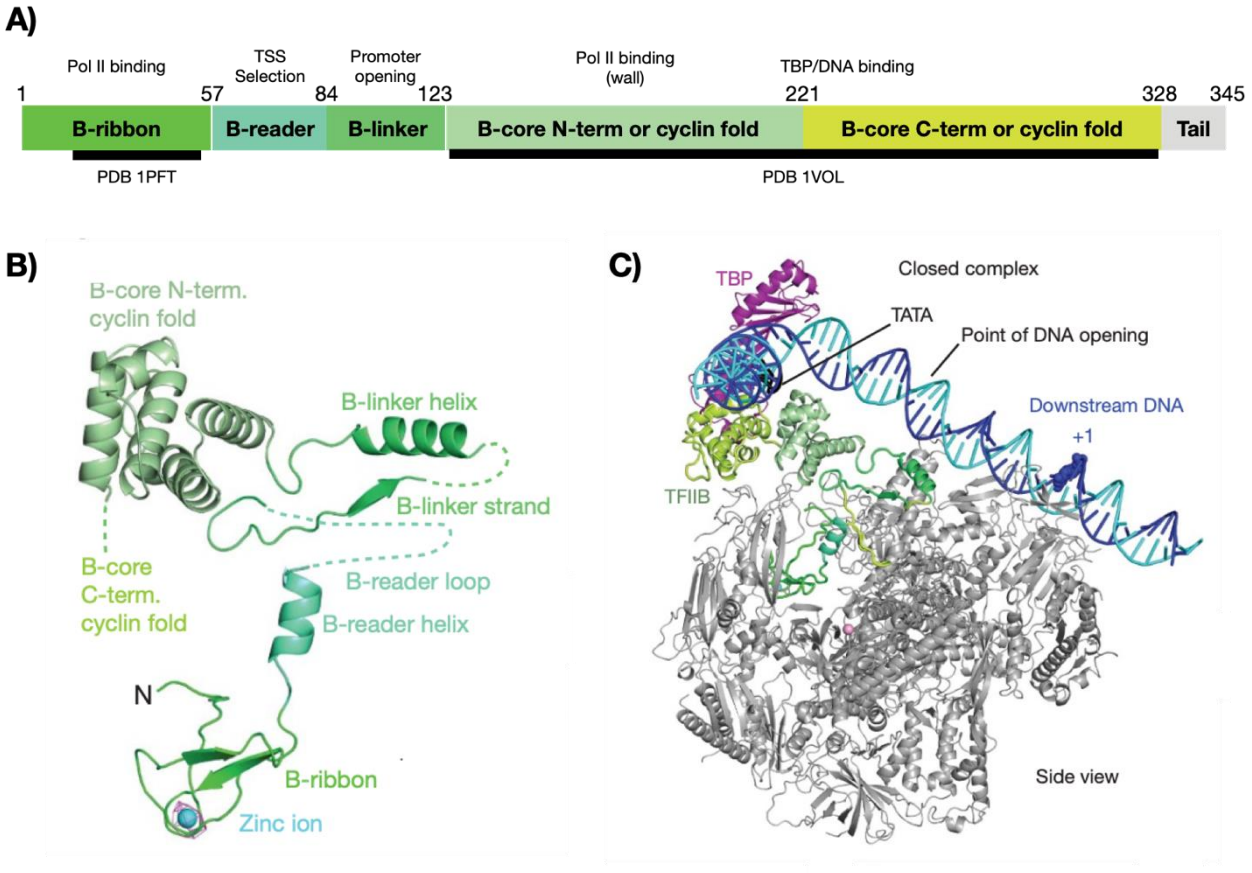


Figura 35.- Estructura de TFIIB. A) Dominios de TFIIB de *S. cerevisiae*. **B)** Modelo de la estructura terciaria de TFIIB. **C)** Modelo del complejo de inicio de la transcripción de la RNA polimerasa II. La caja TATA se muestra en color negro, TBP en morado, TFIIB en verde, el DNA en azul y parte del *core* de la RNA polimerasa II en gris (Kostrewa *et al.*, 2009).

TFIIB contacta al promotor a través de un elemento de reconocimiento a TFIIB (BRE) localizado río arriba y río debajo de la caja TATA. Estructuralmente, TFIIB contiene diversos dominios, el N-terminal se compone de un plegamiento de dedos de zinc (*B-ribbon*) y dos plegamientos de ciclina (*B-core*) que consisten en 5 hélices α (BH1, BH2, BH3, BH4 y BH5 en el primer repetido y BH1', BH2', BH3', BH4' y BH5' en el segundo repetido) (Figura 35). Los dedos de zinc y los plegamientos de ciclina se encuentran conectados por un espaciador que contiene dos hélices α (el *B-reader* y el *B-linker)* (Thomas y Chiang, 2006).

Los dedos de zinc de TFIIB unen a la RNA polimerasa II, mientras que los dominios *B-reader* y *B-linker* se ubican en el sitio activo de la RNA polimerasa II donde participan en la formación del complejo abierto y la selección del sitio de inicio de la transcripción y los plegamientos de ciclina se unen a TBP, a la RNA polimerasa II y al promotor (Kostrewa *et al.*, 2009).

Las subunidades de las RNA polimerasas eucarióticas y de arqueas se encuentran estructuralmente relacionadas y requieren de componentes similares para el inicio de la transcripción que permiten el reconocimiento del promotor, la orientación de la RNA polimerasa y la apertura del DNA (Vannini y Cramer, 2012). TBP está presente en cada uno de estos complejos, mientras que otros factores están estructural o funcionalmente relacionados, como TFIIB que comparte homología con BRF1 (de la maquinaria transcripcional de la RNA polimerasa III). Sin embargo, no se conocían homólogos para TFIIB entre los factores auxiliares de la transcripción por la RNA polimerasa I. A pesar de ello, Rrn7, una subunidad del complejo CF (*Core factor*) comparte muchas actividades asociadas a TFIIB. En años recientes se demostró que Rrn7 (levaduras) y TAF1B (humanos) involucrados en la transcripción de la RNA polimerasa I son homólogos de TFIIB (Naidu *et al.*, 2011; Knutson y Hahn, 2011).

En levadura, el inicio de la transcripción por la RNA polimerasa I requiere de cuatro factores generales de transcripción: el factor UAF (*Upstream activitating factor*), la proteína TBP, el complejo CF y del factor regulatorio Rrn3. A su vez, el complejo CF consta de las subunidades Rrn6, Rrn7 y Rrn11. Los complejos UAF y CF interaccionan con TBP. En humanos, el complejo análogo al CF es el factor SL1 (*Selective factor 1*), compuesto de cuatro factores asociado a TBP: TAF1A, TAF1B, TAF1C, TAF1D y TAFII12, en el cual los factores TAF1A, TAF1B y TAF1C son ortólogos de Rrn6, Rrn7 y Rrn11 (Figura 36) (Knutson y Hahn, 2013).

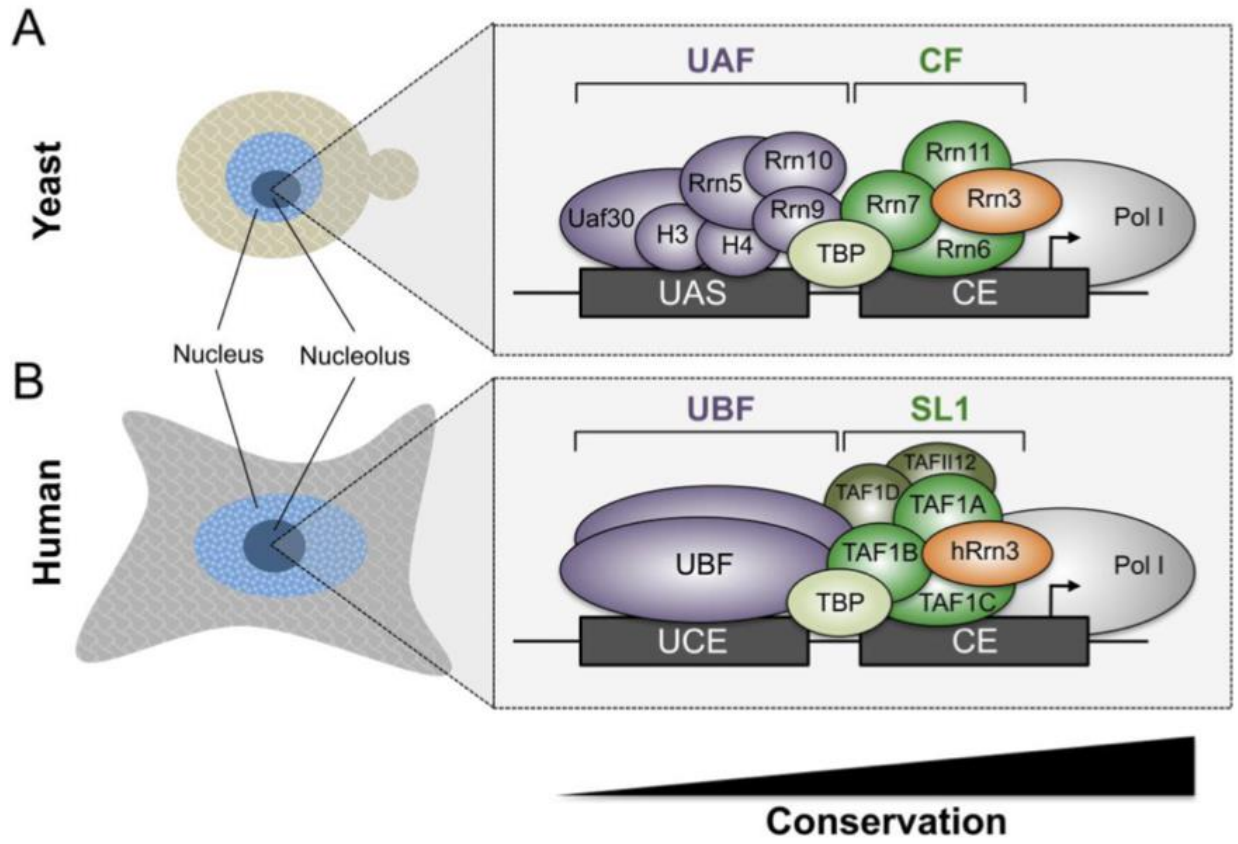


Figura 36 Complejos de preinicio de la transcripción de Pol I en levadura y humano. A) Levadura: El promotor de rDNA contiene dos elementos, la secuencia de activación río arriba (UAS) y el elemento central (CE). B) Humano: el promotor de rDNA de humano contiene dos elementos: el elemento de control río arriba (UCE) y el elemento central (CE). Los ortólogos de humano y levadura están representados con el mismo color. La conservación evolutiva relativa de los factores se indica debajo de los modelos. (Knutson y Hahn, 2013).

Las subunidades de la RNA polimerasa I tienen una conservación de secuencia de aminoácidos baja en comparación con sus contrapartes en las RNA polimerasas II y III, donde Rrn7 y TAF1B no son la excepción (Knutson y Hahn, 2011). Rrn7 y TAF1B muestran una conservación de secuencia que va desde el 8 % al 16 % de identidad con respecto a los miembros de la familia de TFIIB y entre sí (Knutson y Hahn, 2011). A pesar de ello, Rrn7 y TAF1B contienen los mismos dominios encontrados en las proteínas tipo TFIIB: los dedos de zinc, el *B-reader*, el *B-linker* y los plegamientos de ciclina. BRF1, Rrn7 y TAF1B contienen una extensión hacia el C-terminal, sin embargo, esta extensión no está relacionada entre proteínas. El C-terminal de Rrn7 es específico de levadura, mientras que el C-terminal de TAF1B está conservado entre metazoarios, pero no están relacionados con el C-terminal de BRF1 (Figura 37).

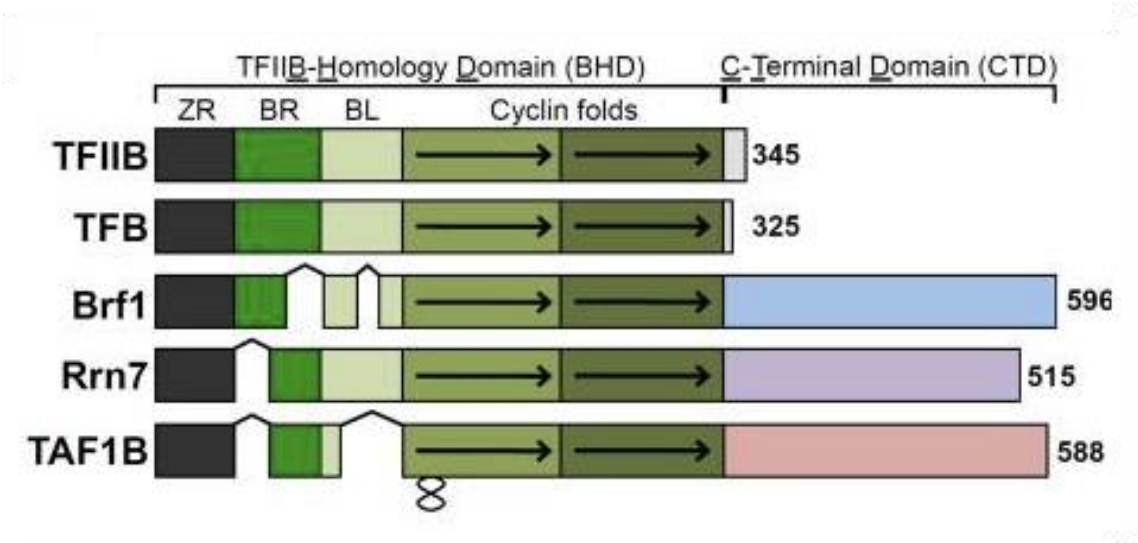


Figura 37.- Dominios de homología de los factores relacionados a TFIIB de la RNA polimerasa I, RNA polimerasa II y la RNA polimerasa III. Representación esquemática de los dominios de la familia de TFIIB: ZR (plegamientos de zinc), BR (*B-reader*), BL (*B-linker*) y las flechas indican los plegamientos de ciclina. Las inserciones son marcadas con líneas negras (Knutson y Hahn, 2011).

Identificación de TFIIB en *T. vaginalis*

Como parte de un proyecto anterior, en el cual se intentó caracterizar a las proteínas TvTBPs, se realizó una búsqueda *in silico* usando las secuencias de aminoácidos de TvTBP1 y TvTBP2 en el programa STRING 9.1 con el objetivo de identificar otras proteínas que ayudaran a proponer un papel diferencial entre ambas TvTBPs (M. en C. Lluvia Rosas Hernández). Este programa es una base de datos de interacciones proteicas conocidas y predichas, las cuales incluyen asociaciones directas e indirectas identificadas de cuatro fuentes: contexto genómico, experimentos de alto rendimiento, coexpresión y conocimiento previo (derivado de la bibliografía). No se encontraron diferencias en las búsquedas con ambas secuencias de las TvTBPs; sin embargo, como resultado de esta búsqueda se identificaron a dos proteínas putativas: TFIIB y BRF1 de *T. vaginalis*. Un ejemplo de los análisis obtenidos se muestra en la figura 38. Cada nodo corresponde a una proteína putativa de *T. vaginalis* identificada mediante búsquedas en bases de datos como INTERPRO y PFAM. TvTBP1 y TvTBP2 de color interaccionan con TvBRF1 (descrita en el capitulo anterior) y TvTFIIB también marcadas en rojo.

TvTFIIB se encuentra anotada en el genoma de *T. vaginalis* como *transcription initiation factor iib, putative* con el número de anotación (TVAG_071030). El gen Tv*tfiib* de 945 pb, codifica para una proteína de 314 aminoácidos y un peso teórico de 34.78 kDa.

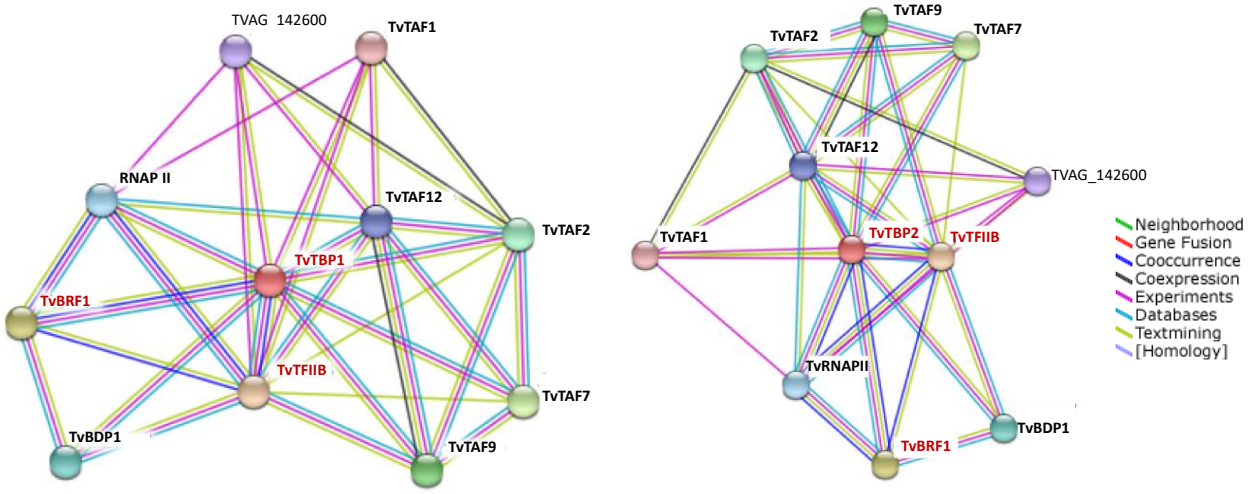


Figura 38.- Identificación de TvTFIIB mediante un análisis de interactoma con TvTBP1 y TvTBP2. Cada nodo indica una proteína putativa de *T. vaginalis*, las cuales se nombraron según el tipo de dominio que presentan. Las líneas indican interacciones conocidas o predichas de distintas fuentes de información .El análisis se realizó en el programa STRING 9.1. La distancia entre cada nodo no tiene ningún significado. Las proteínas identificadas son TvTAF1 (A2FTI5), TvTAF2 (A2E9G0), TvTAF7 (A2EWP5), TvTAF9 (A2ELL9), TvTAF12 (A2E4T9), TvBDP1 (A2EBM7), Rpb1 subunidad de la RNA polimerasa II (A2DKT1), TvTFIIB (A2D806), TvBRF1(A2EJ51) y una proteína no identificada con el número de anotación TVAG_142600 que tiene un dominio SNF2.

Resultados:

Identificación de RRN7 en *T. vaginalis*

Por otro lado, RRN7 de *T. vaginalis* fue difícil de identificar debido a que los factores que participan en la transcripción de la RNA polimerasa I son más distantes evolutivamente que los factores de la RNA polimerasa II y III (Knutson y Hahn, 2013). Primero se realizó un análisis BLAST utilizando las secuencias de aminoácidos del N-terminal de Rrn7 (levadura) y TAF1B (humano) ya reportadas. En cada resultado se identificó al gen con el identificador TVAG_045550 que codifica para una proteína de 440 aminoácidos y que contiene en su extremo N-terminal un dominio tipo Rrn7 correspondiente al plegamiento de dedos de zinc; sin embargo, el valor E en cada caso era poco significativo. Cuando se buscó en la literatura acerca de esta proteína se mencionaba que Rrn7 es en realidad bastante divergente y que tienen poca similitud de secuencia, haciendo complicado su identificación (Knutson y Hahn, 2011). Por lo que se realizó un análisis utilizando el programa HHpred (Söding *et al.*, 2005), que es capaz de identificar si proteínas

distantes son homólogas y que fue utilizado para detectar a Rrn7 y TAF1B como ortólogos de TFIIB (Knutson y Hahn, 2013). Este análisis nos mostró que la proteína anotada con el número TVAG_0455500 se trata de un factor transcripcional de la familia de TFIIB con una probabilidad del 73.4 %. Adicionalmente, se realizó un análisis mediante un árbol filogenético usando diversas secuencias de aminoácidos de Rrn7, TAF1B y MEE12 (como se le conoce en plantas), donde se observa que la TvRrn7 de *T. vaginalis* junto con la proteína AtMEE12 de *Arabidopsis thaliana* son de los ortólogos más divergentes (Figura 39).

TvRRN7 (TVAG_045550) está anotada en la base de datos de TrichDB como una proteína hipotética de 440 aminoácidos y un peso teórico de 50.5 kDa.

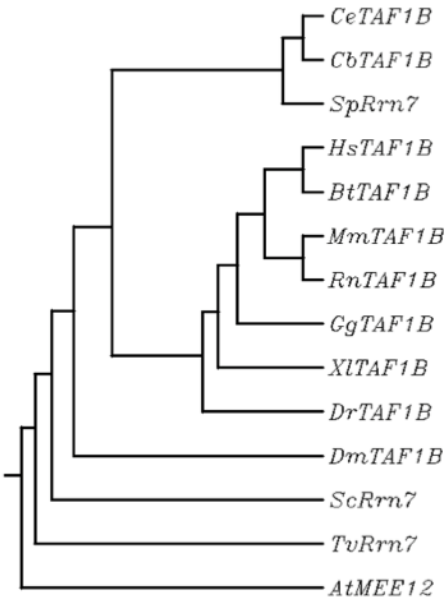


Figura 39.- Árbol filogenético de las proteínas relacionadas a TFIIB de la RNA polimerasa I. ScRrn7 (*S. cerevisiae*), SpRrn7 (*S. pombe*), CeTAF1B (*Caenorhabditis elegans*), CbTAF1B (*Caenorhabditis briggae*), HsTAF1B (*H. sapiens*), BtTAF1B (*Bos taurus*), MmTAF1B (*Mus musculus*), AtMEE12 (*A. thaliana*), GgTAF1B (*Gallus gallus*), RnTAF1B (*Rattus norvegicus*), DmTAF1B (*D. melanogaster*), DrTAF1B (*Danio rerio*) y XlTAF1B (*Xenopus laevis*).

Dominios de los factores relacionados a TFIIB

Con el propósito de averiguar si estas proteínas contienen los dominios típicos de las proteínas relacionadas a TFIIB, se realizaron análisis usando las bases de datos de INTERPRO y PFAM. Las proteínas Rrn7 y TFIIB se caracterizan por contener los dominios de dedos de zinc (*B-ribbon*) y los plegamientos de ciclina (*B-core*) en el N-terminal. Se usó el programa I-TASSER para obtener los modelos de TvRRN7 y TvTFIIB. La secuencia de aminoácidos de TvRRN7 no tiene mucha identidad con las secuencias de levadura o humano. Sin embargo, la estructura terciaria de TvRRN7 adopta un plegamiento similar al cristal de ScRrn7 reportado recientemente (6rui, cadena Q) (Figura 40). Este cristal muestra que la estructura de ScRrn7 no es tan similar a TFIIB como se creyó en un principio, ya que el plegamiento de ciclina II, se interrumpe por una inserción que

forma seis hélices α (Engel *et al.*, 2017). Los dominios de TvRRN7 se han determinado con base en el modelo de la estructura terciara, determinando que esta proteína también contiene una inserción que interrumpe el segundo plegamiento de ciclina. Por otro lado, TvTFIIB se encuentra bastante conservada con respecto al ortólogo de levadura y contiene los mismos plegamientos reportados para ScTFIIB (Figura 40). Los datos obtenidos de este análisis sugieren fuertemente que TvRRN7 y TvTFIIB son los ortólogos de ScRrn7 y ScTFIIB, respectivamente.

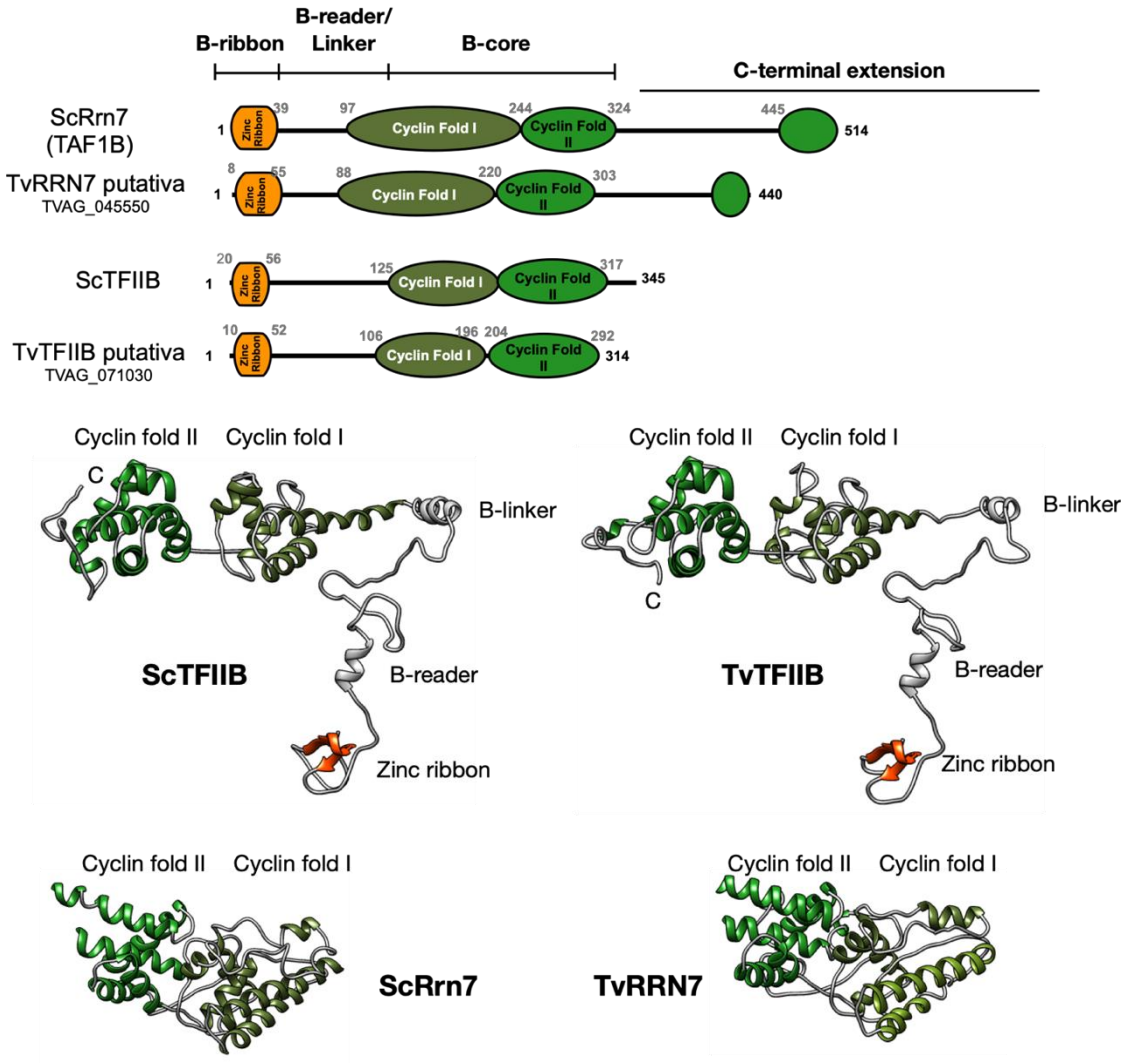


Figura 40.- Estuctura y dominios de TFIIB y RRN7.-Los dedos de zinc (naranja) y los plegamientos de ciclina (verde). RRN7 se modeló usando I-TASSER usando como templado la estructura obtenida por cristalografía de rayos X de *S. cerevisiae* (6rui). El plegamiento de ciclina II de RRN7 se interrumpe por la secuencia del C-terminal.

Obtención de las cepas mutantes de *S. cerevisiae*: *SUA7*/*sua7*Δ::*KANMX*, *RRN7*/*rrn7*Δ::*KANMX* y *BRF1*/*brf1*Δ::*KANMX*

Las cepas mutantes de *S. cerevisiae* (*SUA7*/*sua7*Δ::*KANMX*, *RRN7*/*rrn*7Δ::*KANMX* y *BRF1*/*brf1*Δ::*KANMX*) se obtuvieron usando la cepa diploide BY4743 para probar la función de TvRRN7, TvTFIIB y TvBRF1 mediante ensayos de complementación. La mutación se realizó por recombinación homologa usando un casete de resistencia a kanamicina (*KANMX*) que contenía en sus extremos 5' y 3' las secuencias río arriba y río abajo del locus del gen a eliminar. La integración del casete en el locus correspondiente se verificó por PCR usando oligonucleótidos que se anclan río arriba y río abajo del gen de cada alelo y obteniendo amplicones, en cada caso que corresponden en tamaño para el alelo que se reemplazó por el casete *KANMX* y el alelo silvestre.

Ensayo de complementación de una mutante de levadura de TFIIB

El ensayo de complementación usando la cepa *SUA7*/*sua7*Δ::*KANMX* se utilizó para probar la función del gen Tv*fiib*. El genotipo de la cepa BY4743 es *MAT*a*/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0,* por lo tanto, la cepa es auxótrofa a histidina, leucina y uracilo. El gen Tv*fiib* se clonó en el plásmido pYES2 que contiene como marcador de selección al gen *URA2*, que participa en la biosíntesis de uracilo. Por lo tanto, la selección de las colonias transformantes se realizó en medio mínimo complementado con histidina y leucina.

Como control positivo del ensayo se utilizó la construcción pYES2-ScTFIIB (*SUA7*) de levadura y como control negativo únicamente al plásmido pYES2 vacío transformado en la cepa mutante. El ensayo de complementación se realizó por la técnica de disección de tétradas que consiste en separar cada una de las células que conforman la tétrada, dos de las cuales contendrán el alelo silvestre y por lo tanto pueden crecer sin dificultad en medio YPGAL (*yeast extract*, *peptone*, *galactose*); mientras que las otras dos células portarán en su lugar el casete *KANMX* y solo podrán crecer si el locus expresado desde el plásmido bajo el control del promotor de galactosa (*GAL1*) es capaz de complementar la función de la proteína silvestre.

TvTFIIB no fue capaz de rescatar a la mutante de levadura, a pesar de que es estructuralmente similar a ScTFIIB y conserva los dominios característicos de esta proteína; sin embargo, en secuencia solo tiene aproximadamente un 20 % de identidad con respecto a su ortólogo en levadura (Figura 41). Para confirmar que la disección proviene de una misma tétrada, cada una de las clonas obtenidas se creció en medio YPGAL (todas las clonas crecen) y en medio YPGAL + GENETICINA (donde sólo crecen las esporas que contienen el casete *KANMX* y por lo tanto son

resistentes a geneticina). Este análisis demuestra que dos de las cuatro células del control positivo contienen el alelo mutado, tal como se esperaba, mientras que para TvTFIIB solo hay crecimiento en el medio YPGAL, pero no en el medio con geneticina, confirmando que estas dos clonas contienen el alelo silvestre y que las otras dos células de la tétrada que no pudieron crecer desde la disección contenían el alelo mutado. Finalmente, como última comprobación del ensayo se obtuvo DNA genómico a partir de cada una de las clonas obtenidas y mediante PCR utilizando unos oligonucleótidos que se anclan río arriba y río abajo del gen, se obtienen los amplicones del tamaño esperado para el alelo silvestre de 1693 pb o para el alelo que contiene el casete *KANMX* de 2,201 pb, confirmando una vez más que las dos clonas que crecieron para TFIIB contienen el alelo silvestre. El hecho de que TvTFIIB no complemente a pesar de tener un plegamiento teóricamente parecido al de levaduras se puede deber a que TvTFIIB debe tener ausentes en su secuencia algunos aminoácidos necesarios para establecer las interacciones proteína-proteína necesarias.

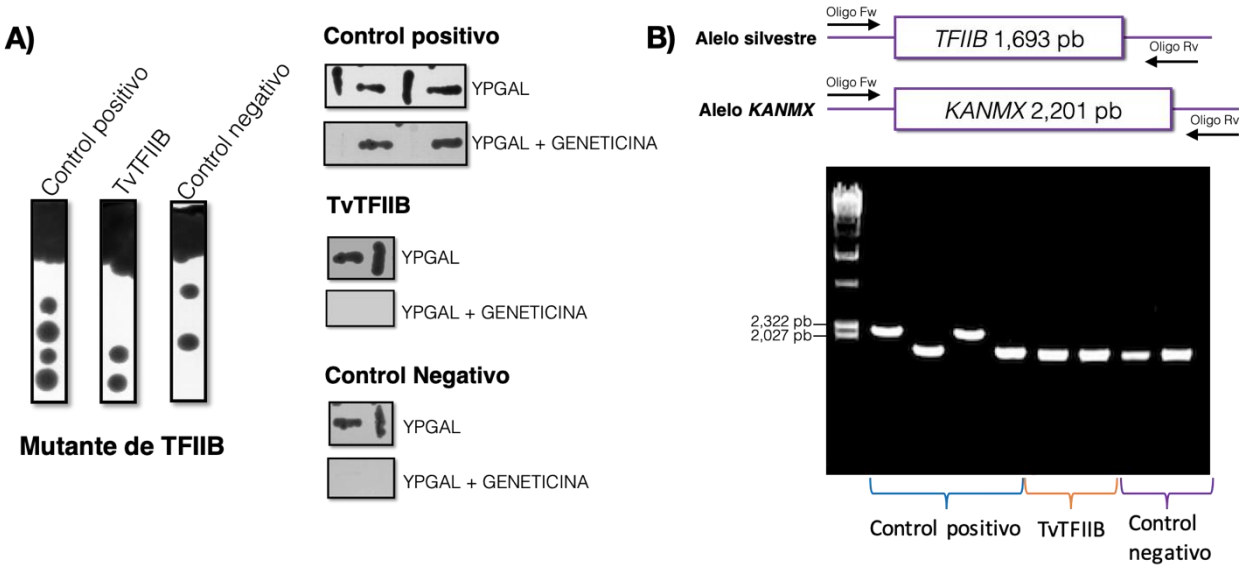


Figura 41.- Ensayo de complementación de una mutante de TFIIB. A) Disecciones realizadas con cada una de las construcciones transformadas en la cepa mutante de levadura. Cada una de las clonas se tomó y se sembró en medio YPGAL y YPGAL+ GENETICINA como método pata probar el tipo de alelo que contienen. B) Comprobación por amplificación. Se extrajo DNA genómico de cada colonia para poder realizar la reacción de PCR, utilizando oligonucleótidos que se anclan río arriba y río abajo del gen, obteniendo un amplicón de aproximadamente 1,693 pb para el alelo silvestre y 2,201 pb para el alelo que contiene el casete de *KANMX*.

El ensayo de complementación usando a la cepa mutante *BRF1*/*brf1*Δ::*KANMX* se describió en un capitulo anterior, mientras que el ensayo de complementación para TvRRN7 usando la cepa *RRN7*/*rrn7*Δ::*KANMX* no se realizó, debido a que TvRRN7 es la proteína con menor similitud que

TvBRF1 y TvTFIIB quienes no fueron capaces de complementar sus mutantes respectivas en levadura.

Ensayos de interacción por doble híbrido para TFIIB y RRN7

Para probar la interacción de TvTFIIB y TvRRN7 con TvTBP1, TvTBP2 e IBP39 se realizaron ensayos de doble híbrido. Los cultivos con las células transformantes se crecieron a una densidad de 0.6 y posteriormente se hicieron diluciones de 1:5, 1:25 y 1:50. Cada dilución se goteó en cajas con medio mínimo galactosa (para ver la actividad del reportero de leucina) y cajas con medio mínimo, galactosa, leucina y X-gal (para ver la actividad del reportero *lacZ*). Todas las células transformantes son viables ya que se observa crecimiento en el medio mínimo, galactosa, leucina y X-gal. Sin embargo, por este método no se logró ver interacción entre TvTBPs/TvTFIIB y TvTBPs/TvRRN7 (Figura 42A).

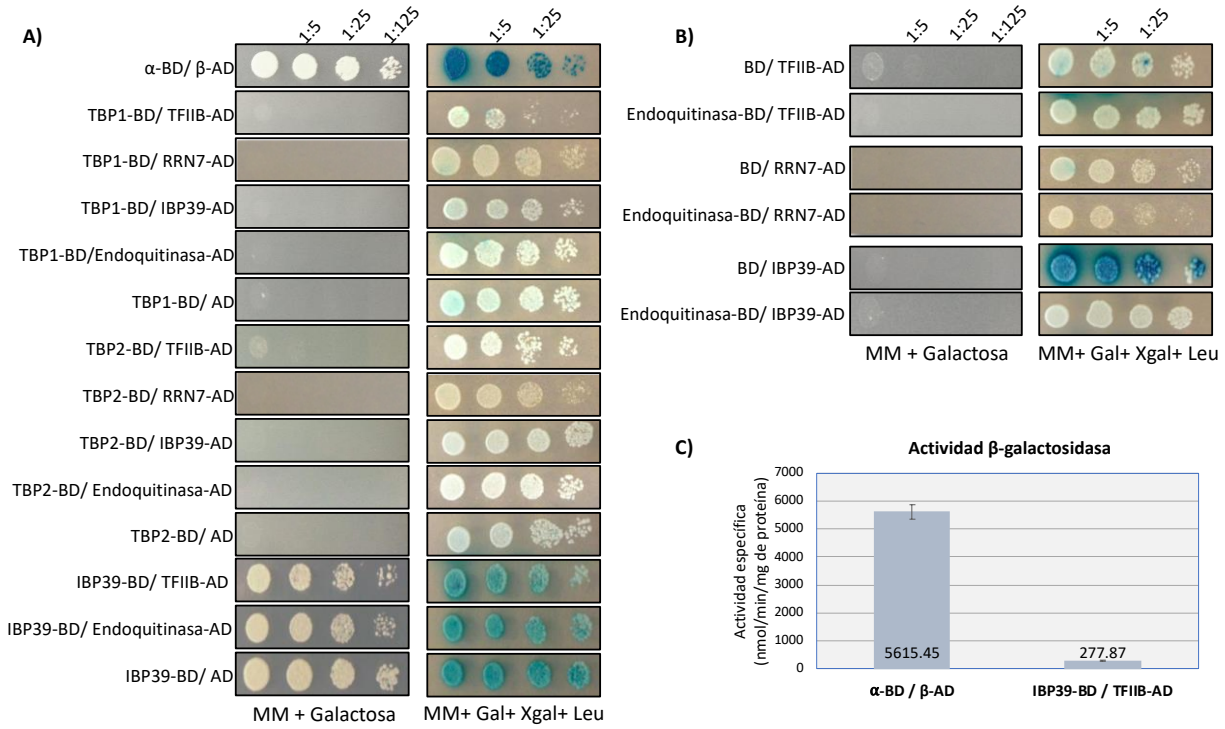


Figura 42.- Ensayos de doble híbrido para TFIIB, IBP39 y las TvTBPs. A) TvTBP1, TvTBP2 e IBP39 en pEG2002, permiten la expresión de las proteínas unidas al dominio de unión a DNA (BD). TvTFIIB, TvRRN7 e IBP39 en pJG4-5, permiten la expresión de las proteínas unidas al dominio de activación (AD). Los cultivos se obtvieron con una densidad óptica inicial de 0.6, a partir del cual se hicieron diluciones seriadas. Los goteos se realizaron con 3 µl de cada muestra sobre cajas MM + Galactosa (para evaluar el reportero de leucina) y MM+Gal+Xgal+Leu (para evaluar el reportero de *lacZ*). **B)** Controles de interacción para evaluar que TvTFIIB, TvRRN7 e IBP39 en el dominio de activación, no activan la transcripción. Los goteos se realizaron por

triplicado. **C)** Ensayos de β-galactosidasa para probar la activación del reportero *lacZ* mediante la interacción de IBP39-BD/ TFIIB-AD.

Por otro lado, IBP39 unido al dominio de unión al DNA (BD) permite el crecimiento de las células en ausencia de leucina y las colonias son azules en presencia del X-gal, indicando una autoactivación de la transcripción. Para cuantificar esta activación, se tomaron a las células transformadas con IBP39-BD/TvTFIIB-AD y se midió la actividad de la β-galactosidasa (Figura 42C). Sin embargo, la actividad obtenida por IBP39 está muy por debajo de la actividad del control positivo. Adicionalmente, IBP39 unido al dominio de activación (AD), activa el promotor *lacZ* (Figura 42B). Por lo tanto, la interacción entre IBP39 y TvTFIIB no se pudo probar por este método. Las interacciones proteína-proteína por este método pueden dar falsos negativos, debido a motivos como localización celular, modificaciones post-traduccionales, entre otros (Brückner *et al.*, 2009).

Para descartar falsos negativos, se decidió probar la interacción proteína-proteína mediante ensayos de coprecipitación.

Ensayos de interacción por coprecipitación (*GST pulldown)*

TFIIB y Rrn7 participan en los complejos de inicio de la transcripción de las RNA polimerasas II y I, respectivamente. Los plegamientos de ciclina de TFIIB y Rrn7 participan en la interacción con TBP. Para demostrar que TvTFIIB y TvRRN7 son parte del complejo de inicio de la transcripción, se realizó el siguiente ensayo de coprecipitación.

El marco de lectura abierto del gen Tv*tfiib* se clonó en el plásmido pGEX4T1 para obtener a la proteína recombinante fusionada a una etiqueta de GST en el extremo N-terminal, esto con el propósito de probar la interacción de TvTFIIB con las proteínas TvTBP1, TvTBP2 e IBP39 que se encuentran clonadas en el plásmido pQE30, con el cual se pueden expresar a las proteínas recombinantes fusionadas a una etiqueta de histidinas. GST-TvTFIIB fue unida a la matriz de glutatión sefarosa y se incubó posteriormente con los lisados de *E. coli* que expresan a las proteínas recombinantes 6xHisTvTBP1 y 6xHisTvTBP2 (extracto total). Las perlas se lavaron tres veces para eliminar las proteínas no unidas, una muestra del lavado se cargó al gel para probar que no quedaran proteínas no unidas (lavado). Posterior a este paso, se tomó un volumen de perlas las cuales se trataron con β-mercaptoetanol y se cargaron al gel (matriz). Finalmente, las perlas restantes se trataron con glutatión reducido para favorecer el intercambio con las proteínas unidas

a las perlas (elución). El gel fue transferido para realizar un ensayo de inmunodetección usando a la sonda His-Probe que nos permite detectar a las proteínas unidas a histidinas. Una señal del tamaño correspondiente a TvTBP1 y TvTBP2 aparece en la matriz, pero en la elución, sólo se observa señal para TvTBP1. Para el ensayo GST-TvTFIIB/6xHis-TvTBP2 se observa una señal débil desde la matriz, por lo que en la muestra del eluato, si hay proteína TvTBP2, es difícil detectar la señal, debido a que debe estar más diluida. Estos ensayos sugieren una interacción entre las proteínas TvTFIIB y TvTBP1 (Figura 43A).

Por otro lado, el marco de lectura abierto del gen de Tv*rrn7* se clonó igualmente en el vector de expresión pGEX4T1 para obtener a la proteína recombinante TvRRN7 fusionada a GST en el extremo N-terminal. El ensayo se realizó de manera similar al descrito anteriormente, excepto que en este ensayo se cargó una muestra de las proteínas no unidas a la matriz. Este ensayo es preliminar debido a que no se ha realizado por triplicado y se sugiere probar otras condiciones. Los resultados de este ensayo muestran una señal apenas visible en la matriz con TvTBP1, sugiriendo que TvRRN7 podría interaccionar de manera débil con esta proteína, pero no con TvTBP2 (Figura 43B). Entre las proteínas relacionadas a TFIIB de *T. vaginalis*, TvRRN7 parece ser la más alejada evolutivamente de sus contrapartes en otros organismos y los datos que tenemos a la fecha sobre esta proteína son insuficientes para demostrar su participación en la formación de los complejos de inicio de la transcripción por la RNA polimerasa I.

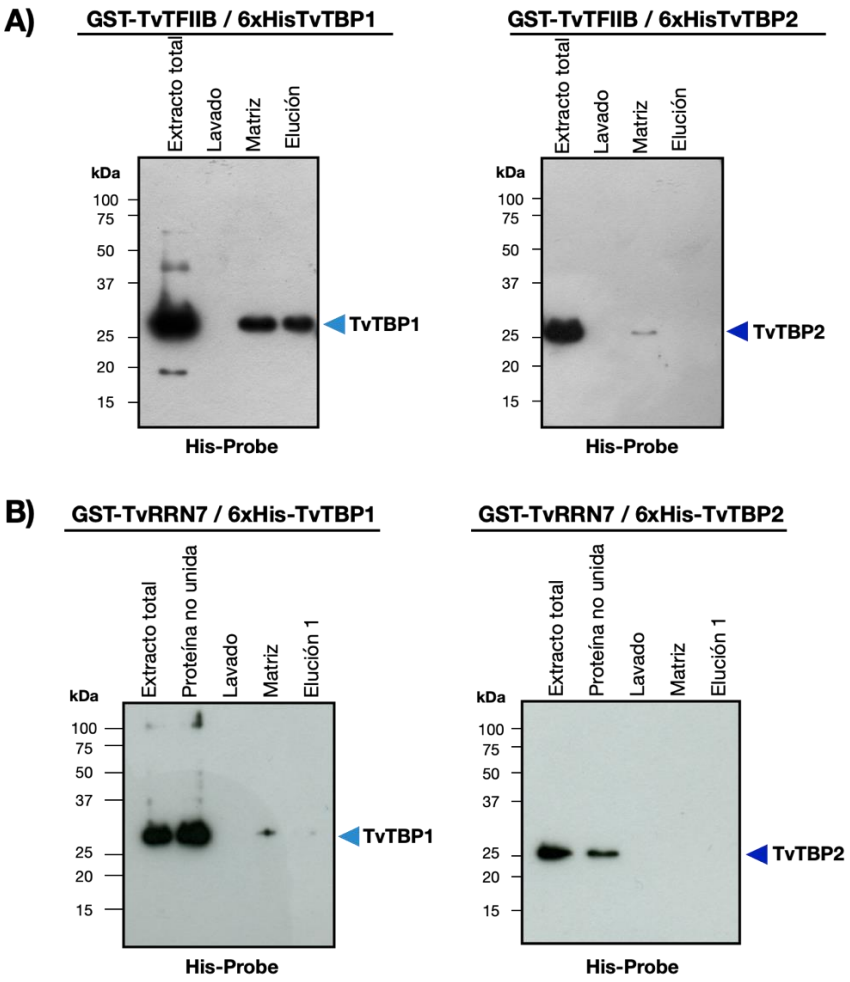


Figura 43.- Ensayo de interacción entre TvTFIIB/TvTBPs y TvRRN7/TvTBPs. A) Ensayos de interacción de TvTFIIB con TvTBP1 y TvTBP2. B) Ensayos de interacción de TvRRN7 con TvTBP1 y TvTBP2. El extracto total de *E. coli* que expresa a la proteína recombinante 6xHis-TvTBP1 o 6xHisTBP2 se incubaron con las proteínas GST-TvTFIIB o GST-TvRRN7. Para el ensayo con TvRRN7 se cargó adicionalmente una muestra de la proteína no unida a la matriz que contiene la proteína recombinante 6xHisTvTBP excedente. El lavado corresponde a una muestra del último lavado realizado en 50 µl. La matriz es una muestra de las perlas lavadas después de la interacción y la elución es una muestra del eluato después de tratar a las perlas con 10 mM de glutatión reducido. Se realizó una inmunodetección usando a la sonda HisProbe-HRP para detectar la presencia de las proteínas recombinantes 6xHis-TvTBP1 o 6xHis-TvTBP2.

TFIIB interactúa con IBP39

Una forma de analizar si TvTFIIB está involucrado en la transcripción por la RNA polimerasa II en *T. vaginalis*, es averiguar si TvTFIIB interacciona con IBP39, un proteína que participa en la transcripción de la RNA polimerasa II de *T. vaginalis*. Se realizó un ensayo de coprecipitación usando como carnada a la proteína recombinante GST-TvTFIIB, la cual se incubó con las perlas de glutatión. Las perlas se lavaron para eliminar las proteínas no unidas. Estas perlas se incubaron

con un lisado de *E. coli* que expresa a la proteína recombinante 6xHis-IBP39, una muestra de este extracto fue cargado en el gel. Las perlas se centrifugaron para retirar el exceso del volumen de lisado, una muestra de este lisado se cargó en el gel (proteína no unida). Posteriormente las perlas son lavadas para asegurar que no queden restos de proteínas sin unir a las perlas (lavado). Un volumen de las perlas ya lavadas, se trataron con β-mercaptoetanol y se cargaron al gel (matriz). El volumen restante de perlas se trató con glutatión reducido para eluir las proteínas unidas (elución); en este ensayo se realizó este paso tres veces. Todas las muestras se cargaron en un gel y se transfirió para realizar un ensayo de inmunodetección usando la sonda His-Probe que nos permite detectar la presencia de la proteína 6xHis-IBP39 (Figura 44). En el ensayo de inmunodetección se observa una señal de aproximadamente 15 kDa y no de 39 kDa que corresponde al peso teórico de IBP39. Esta proteína consta de dos dominios conectados por un espaciador sensible a proteasas, el dominio N-terminal tiene un peso aproximado de 14.5 kDa y el dominio C-terminal de aproximadamente 25 kDa (Liston *et al.*, 2001). Lo que sugiere que la señal que detectamos en el ensayo de inmunodetección corresponde al dominio N-terminal de IBP39 que contiene la etiqueta de histidinas. El dominio C-terminal de IBP39 no se puede detectar por inmunodetección, pero se observa una señal del peso esperado para este dominio cuando se tiñe la membrana con rojo Ponceau. Desconocemos si el corte de la proteína se dio desde la obtención del extracto total, lo que es altamente probable debido a la presencia de otras proteínas en el lisado. Este hecho resulta interesante, ya que el dominio N-terminal de IBP39 es el que reconoce el elemento iniciador (Inr), mientras que el dominio C-terminal interacciona con la RNA polimerasa y dependiendo con que dominio interaccione TvTFIIB, nos puede sugerir un papel de este proteína. Este resultado muestra que TvTFIIB puede interaccionar con la proteína IBP39, sugiriendo que TvTFIIB puede participar al igual que IBP39 en la transcripción por la RNA polimerasa II.

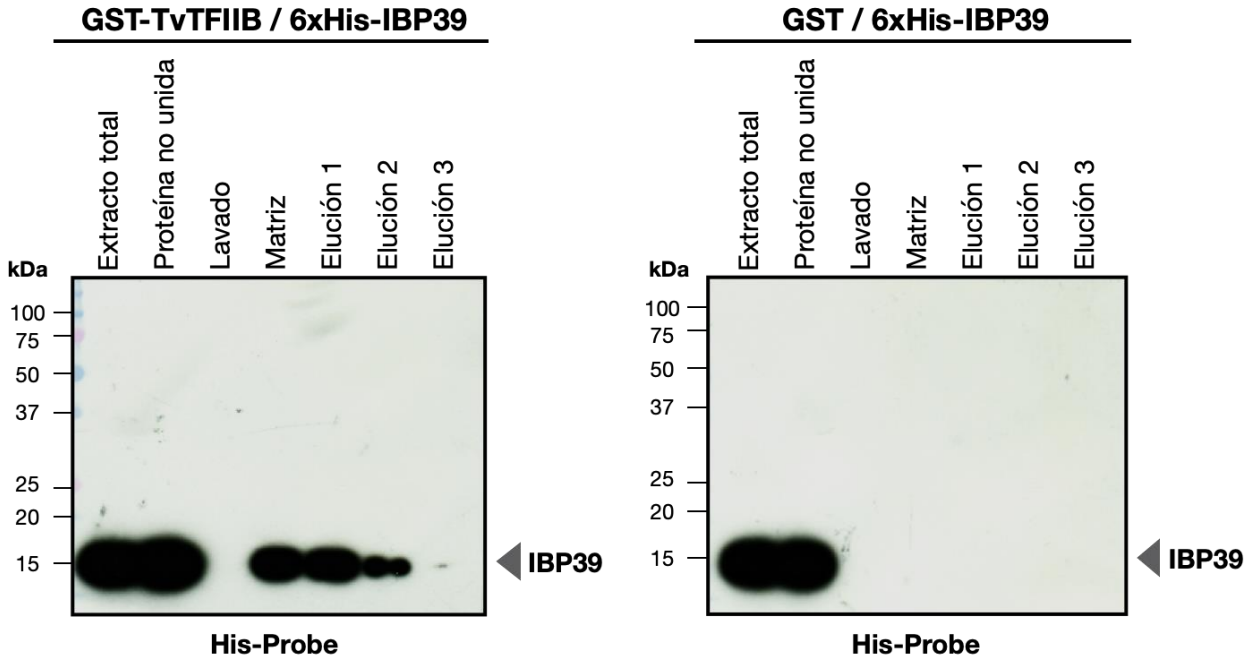


Figura 44.- TvTFIIB interactúa con IBP39.- El Lisado de *E. coli* que expresa a la proteína recombinante 6xHis-IBP39 (extracto total), se incubó con las perlas unidas a GST-TvTFIIB. La proteína no unida corresponde a la fracción de proteínas *E. coli* y 6xHis-IBP39 que no alcanzó unirse a la matriz después de la incubación. La muestra de lavado, corresponde a una fracción del cuarto lavado de las perlas en 50 µl de PBS 1X después de la incubación. Se tomó 1/10 de las perlas lavadas y se trataron con β-mercaptoetanol, esta muestra corresponde a la línea marcada como matriz. El resto de las perlas se trataron con 10 mM de glutatión reducido para eluir las proteínas unidas. La presencia de 6xHis-IBP39 se reveló por inmunodetección tipo *western blot* usando la sonda HisProbe-HRP.

Unión de TFIIB al DNA

Una de las funciones de TFIIB es asegurar a la proteína TBP al promotor a través de la unión de TFIIB al elemento BRE (*B Recognition Element*), que en eucariontes metazoarios se encuentra río arriba y río abajo de la caja TATA, a estos elementos de unión se le conoce como caja BREu y caja BREd respectivamente. Las secuencias de estos elementos en eucariontes son 5’ G/C-G/C-G/A-C-G-C-C 3’ para la caja BREu y 5’ G/A-T-T/G/A-T/G-G/T-T/G-T/G 3’ para BREd (Butler y Kadonaga, 2002). Sin embargo, estas secuencias no han sido identificadas en los promotores de *T. vaginalis*. Para probar la capacidad de unión de TvTFIIB al DNA se usaron dos estrategias, los ensayos por *Southwestern blot* y EMSAs.

Ensayo de *Southwestern blot*

Este ensayo se hizo con el objetivo de demostrar la unión de TFIIB a un promotor transcrito por la RNA polimerasa II. En este momento se contaba ya con sondas de diversos tamaños para el promotor de actina, pero debido a que no se conoce el sitio de unión de TvTBP, se decidió usar la sonda de mayor tamaño (102 pb) que contiene un elemento parecido a una caja TATA en la posición -64 y un elemento Inr alrededor del sitio de inicio de la transcripción el cual es reconocido por IBP39, por lo que esta proteína se usó como control positivo del ensayo. De manera breve, las proteínas recombinantes y purificadas se gotearon en concentraciones conocidas sobre una membrana de nitrocelulosa, la cual se hibridó con una sonda de doble cadena marcada radioactivamente con ATP^{32} obtenida por PCR, después de tres lavados. La membrana se colocó en un casete y se expusieron placas para revelar las señales. Las proteínas recombinantes GST y CRK5 se usaron como controles negativos. Con este método observamos interacción de la sonda con IBP39, cuya unión al DNA es bastante estable e intensa, sin embargo, no se detectó ninguna señal para TvTBPs o TvTFIIB (Figura 45). Al aumentar el tiempo de exposición, la señal obtenida por IBP39 es tan fuerte que opaca las regiones donde se encuentran las otras proteínas, lo que lamentablemente nos impidió ver si hay una señal de unión por mínima que sea para TvTFIIB. El ensayo funciona muy bien y se recomienda realizarse nuevamente con algunos cambios: 1) dejar el carril de IBP39 apartado del resto de las proteínas de interés, 2) realizar menos lavados o por menos tiempo, 3) aumentar el tiempo de hibridación y 4) cambiar algunos de los componentes del amortiguador de hibridación.

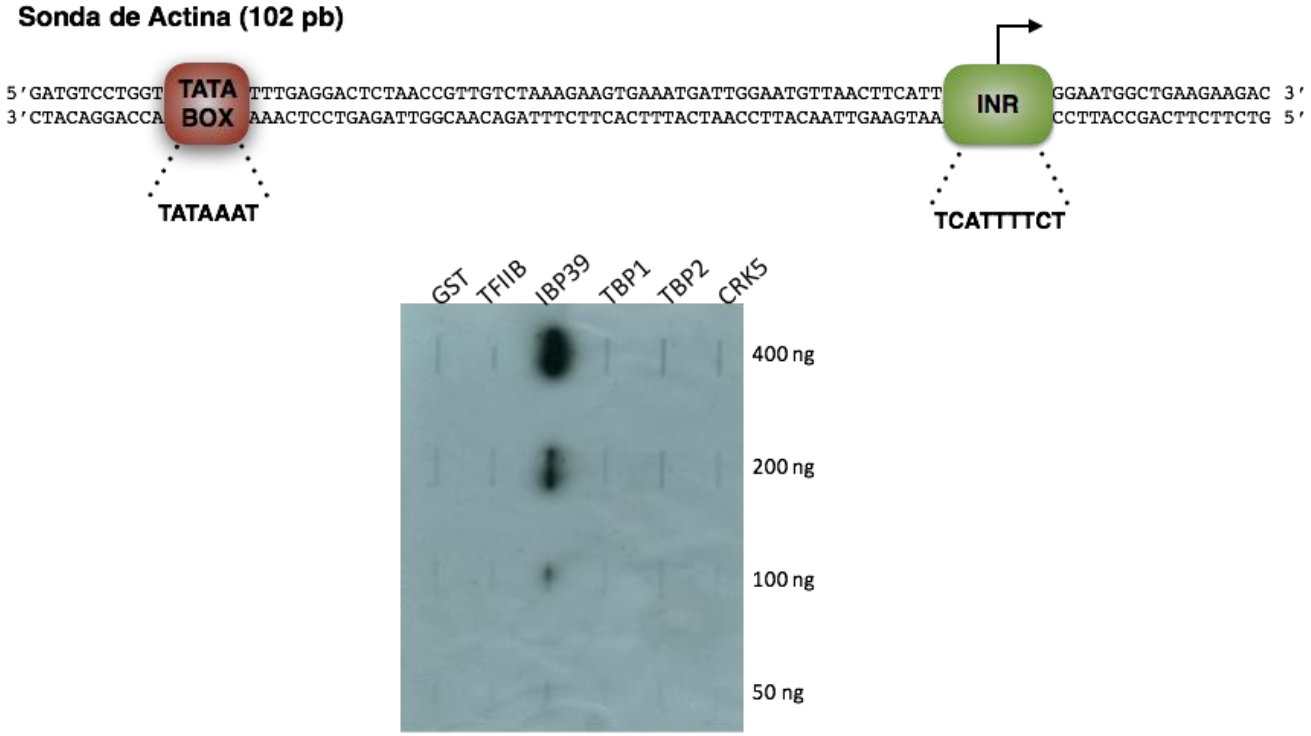


Figura 45.- ***Southwestern blot*** **para probar la unión de TFIIB al DNA.** Panel superior, secuencia de la sonda radioactiva marcada con ATP^{32}. Panel inferior, *Southwestern blot.* Nota importante: La placa se escaneó encima de una membrana de muestra teñida con *amido black* para revelar la presencia de las proteínas.

Ensayos EMSA para TFIIB

Para probar si TvTFIIB reconoce una secuencia promotora de la RNA polimerasa II de *T. vaginalis* de manera independiente o si requiere a TvTBP1 para unirse a una secuencia promotora, realizamos ensayos tipo EMSA. Como un primer paso, se intentaron mejorar las condiciones de unión para TvTFIIB, probando distintos amortiguadores de unión, cuya composición se detallan en el anexo 1. El amortiguador de unión marcado como III, es el amortiguador de referencia, ya que es el que usamos en todos nuestros ensayos al inicio. En el carril 5 se observa un retardo con TvTBP1, en el carril 6 también se alcanza a ver un retardo con TvTFIIB, lo que nos sugiere que TvTFIIB no requiere a TvTBP1 para unirse al DNA y en el carril 7, donde esperábamos ver un superretardo, no se alcanza a ver diferencias con respecto al carril 5. Para el resto de los amortiguadores de unión se usaron ambas proteínas (TvTBP1 y TvTFIIB). El amortiguador de unión I no presenta diferencias respecto al amortiguador III, pero con los amortiguadores de unión II, IV y V se observa un retardo mejor definido y menor cantidad de sonda libre (Figura 46). Finalmente se decidió usar el amortiguador de unión II.

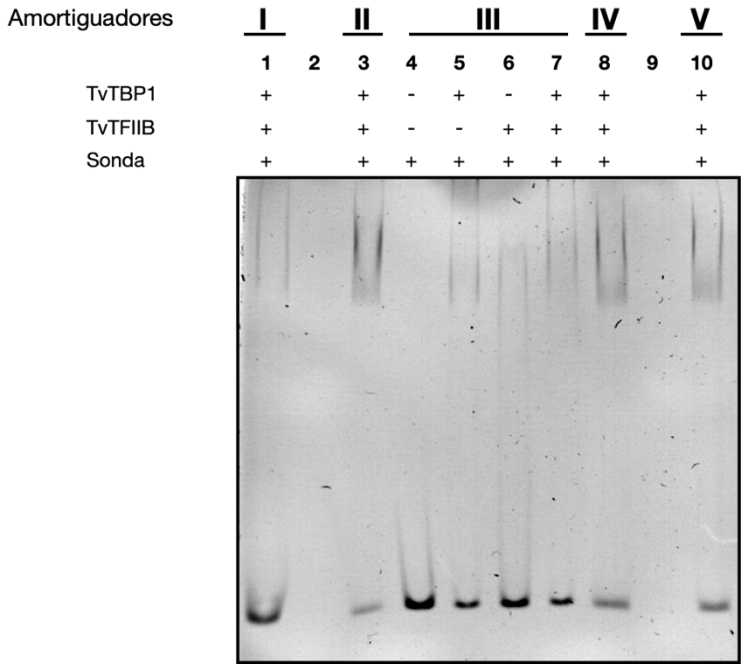


Figura 46.- Condiciones de unión probadas para TvTFIIB en ensayos EMSA. Las proteínas TvTBP1 y TvTFIIB se incubaron con la sonda de actina de 102 pb, usando distintos amortiguadores de unión (anexo 1). El gel se tiñó con bromuro de etidio y los colores de la imagen se muestra invertidos.

La sonda derivada del promotor de actina de *T. vaginalis* de 102 pb usada en el ensayo de *Southwestern blot*, también se empleó como sonda en los ensayos EMSA. Las proteínas recombinantes 6xHis-TvTBP1 y GST-TvTFIIB purificadas se incubaron con la sonda en dos ensayos independientes, utilizando a la sonda sin marcar y marcada radioactivamente (Figura 47). En ambos casos se observa un retardo con TvTFIIB, sugiriendo que esta proteína se une al DNA sin la unión previa de TvTBP1, tal como se observó en el ensayo con los diferentes amortiguadores de unión. Se observa un ligero superretardo en el ensayo con la sonda radioactiva, indicando que ambas proteínas se encuentran unidas a la sonda,.

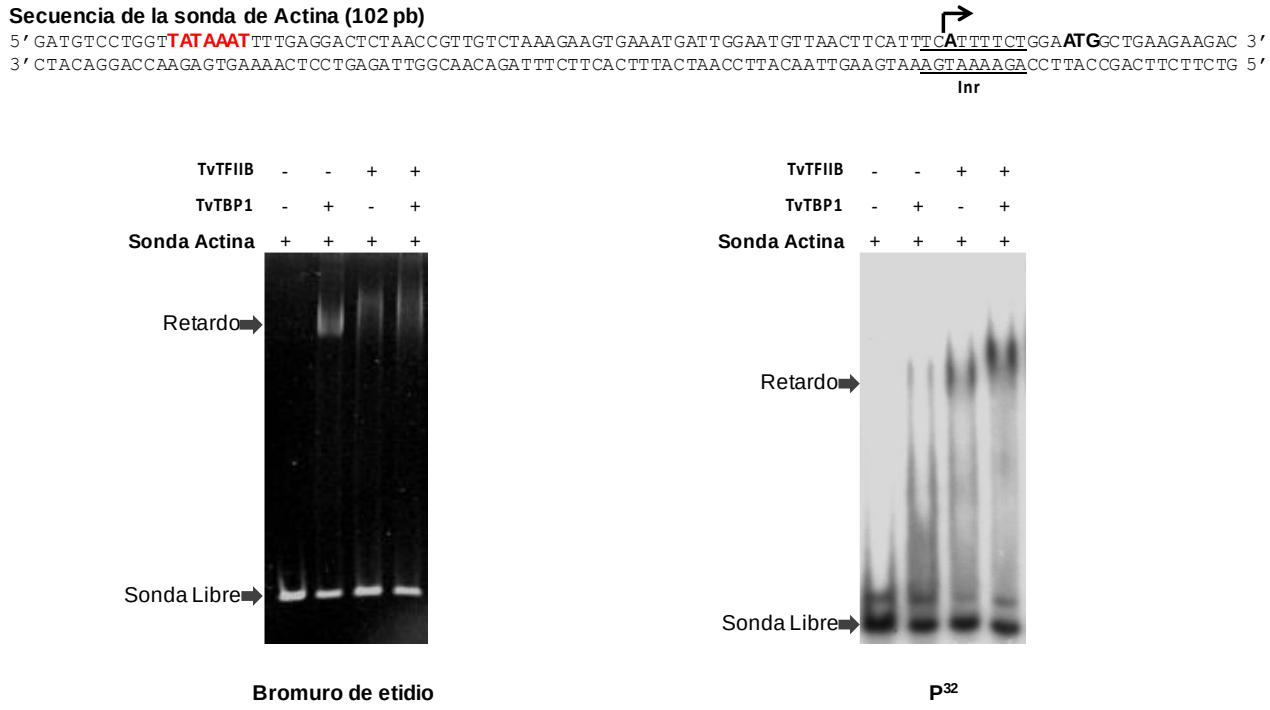


Figura 47.- Ensayos EMSA para probar la unión de TFIIB al DNA. La secuencia de la sonda de actina se indica arriba. El ensayo de la izquierda se tiñó con de bromuro de etidio y en el ensayo de la derecha se usó una sonda marcada radioactivamente.

En el ensayo anterior, no pudimos ver la formación de un superretardo, creemos que si disminuimos el porcentaje del gel de 5 % a 4 %, el complejo migrará mejor, por lo que se realizó el siguiente ensayo EMSA con marca radioactiva usando como sonda el promotor de Actina de 102 pb.

TvTFIIB retarda la sonda de actina, sin embargo, en lugar de ver una señal definida, se observa un barrido, lo cual puede deberse a que el complejo TvTFIIB-DNA no es estable. Por otro lado, cuando TvTFIIB se encuentra junto a la proteína IBP39 se observa un ligero superretardo con respecto a la señal observada con IBP39 sola, indicando que ambas proteínas interactúan al DNA al mismo tiempo (Figura 48). De manera similar, cuando la sonda se incubó con TvTFIIB y TvTBP1 se observa un ligero aumento de la migración con respecto a TvTBP1 sola, lo que indicaría que ambas proteínas se encuentran unidas al DNA simultáneamente o una está interaccionando con la otra. Este resultado no es tan claro con TvTBP2.

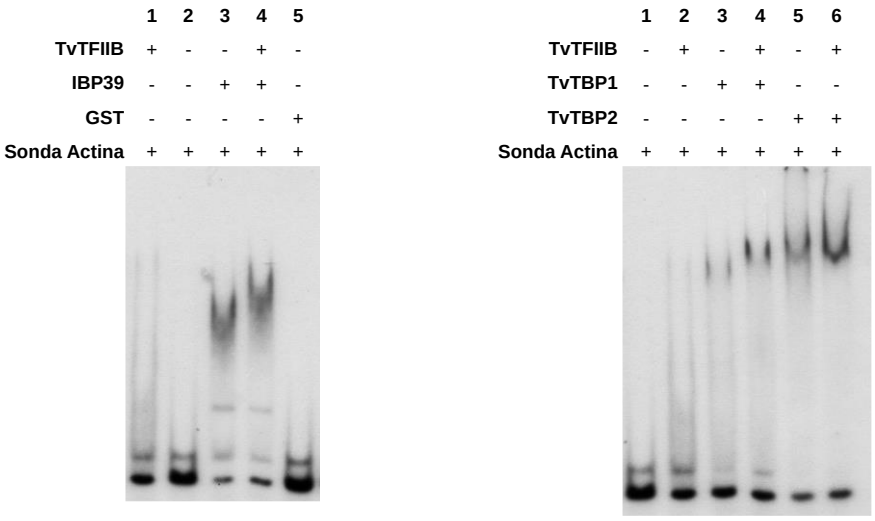


Figura 48.- EMSA radioactivo de TvTFIIB. La sonda usada corresponde al promotor de actina de 102 pb. Se usaron concentraciones similares de TvTFIIB, TvTBP1, TvTBP2 e IBP39 para el ensayo. Se usó GST como control negativo.

Observar superretardos no es tan claro como se esperaba, por lo que se recomienda cambiar la resolución del gel, cambiando la proporción de acrilamida-bisacrilamida o aumentar el tamaño de los geles. También hay que considerar que, al ser un gel nativo, el complejo no migra exclusivamente por tamaño, sino que el desplazamiento en la matriz del gel está influenciado por diversos factores, entre ellos la carga neta y la forma que adopte el complejo. En los ensayos EMSA realizados con TvBRF1, se realizó primero un ensayo de coprecipitación para obtener a las proteínas en complejo antes de ser incubadas con la sonda, mejorando los resultados obtenidos, por lo que sería interesante probar este método para TvTFIIB.

Con el propósito de determinar si TFIIB pertenece a los complejos del inicio de la transcripción *in vivo* de *T. vaginalis*, se obtuvo el anticuerpo anti-IBP39 para realizar ensayos de co-inmunoprecipitación seguidos por secuenciación de proteínas para determinar la identidad de las proteínas que interaccionan con IBP39 en los complejos formados para la transcripción de la RNA polimerasa II. Sin embargo, no se alcanzó a realizar.

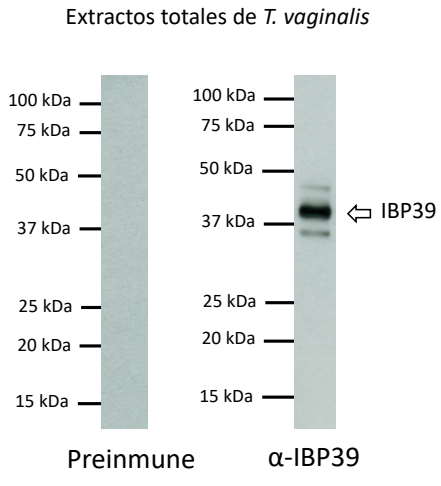


Figura 49.- Validación del anticuerpo α-IBP39. El anticuerpo policlonal anti-IBP39 se obtuvo inmunizando la proteína recombinante 6xHis-IBP39. La inmunodetección tipo *western blot* se realizó usando extractos totales de *T. vaginalis* y un título de 1:500. Una señal por arriba de los 37

kDa corresponde con el peso esperado para IBP39, el cual no se detecta en el *western blot* usando el suero preinmune.

TvRRN7 no reconoce secuencias en el DNA

Rrn7 forma parte del complejo CF de levadura, junto a las proteínas Rrn3, Rrn6 y Rrn11. Para determinar si TvRRN7 se une a una secuencia de DNA, se realizó un ensayo tipo EMSA, usando como sonda una secuencia que corresponde al cistrón ribosomal, previamente probada con TvTBP1. Como control positivo de unión se usó a la proteína TvTBP1. Las proteínas se incubaron de manera independiente con la sonda en concentraciones similares. La proteína TvRRN7 no retarda la sonda (Figura 49). Cuando se incubaron TvTBP1 y TvRRN7 con la sonda, no se observa un superretardo, sino al contrario, parece que la señal del retardo con TvTBP1, desaparece. Un modelado por homología, usando a las proteínas del CF de levadura sobre el DNA, en el que Rrn7 se subrepuso con la estructura de TFIIB y muestra que Rrn7 no puede unirse al DNA del mismo modo que TFIIB (Engel *et al.*, 2017); esto es debido en gran parte a la inserción que presenta esta proteína entre el plegamiento de ciclina II.

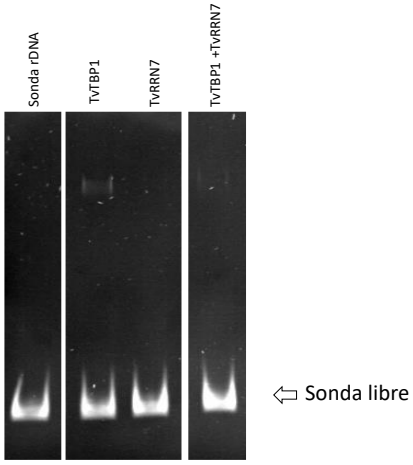


Figura 50.- Ensayo EMSA para TvRRN7.- La sonda del cistrón ribosomal se incubó con las proteínas recombinates TvTBP1 y TvRRN7. El gel se tiñó con bromuro de etidio.

Conclusiones

Los datos de interacción mostrados aquí sugieren que TvTFIIB podría formar parte de los complejos de inicio de la transcripción de la RNA polimerasa II en *T. vaginalis*, ya que interacciona con IBP39 y TvTBP1. Adicionalmente, TvTFIIB tiene capacidad de unión con el DNA. Por otro lado, la proteína putativa TvRRN7, parece ser un ortólogo distante de la familia relacionada a TFIIB, que tiene baja identidad de secuencia y estructura con respecto a ScRrn7. TvRRN7, muestra una interacción débil con TvTBP1 y no se une a secuencias en un promotor transcrito por la RNA polimerasa I.

Consideraciones finales

La proteína TBP juega un papel central en la formación de los complejos de inicio de la transcripción de las tres RNA polimerasas, por tal motivo, estudiar a esta proteína era un buen punto de partida. En el caso de *T. vaginalis*, se expresan dos proteínas con similitud a TBP, sin embargo, los modelos estructurales obtenidos por modelado de homología no muestran diferencias. Las proteínas relacionadas a TFIIB: TvRRN7 de la RNA polimerasa I, TvTFIIB de la RNA polimerasa II y TvBRF1 de la RNA polimerasa III, nos permitieron poner en contexto los tres principales complejos de inicio de la transcripción. Las interacciones de estas proteínas con las TvTBPs nos muestran una diferencia clara entre la proteína TvTBP1 yTvTBP2. Sugiriendo un papel de TvTBP1 en la formación de al menos los complejos transcritos por la RNA polimerasa II y III. La participación de cualquiera de las dos TvTBPs en los complejos de la RNA polimerasa I, es incierto y requiere de más evidencias. Adicionalmente, las proteínas TvBRF1, TvBDP1 y TvTBP1 conforman al complejo TFIIIB que forma parte de los complejos transcripcionales de la RNA polimerasa III. El estudio de las interacciones proteína-proteína probadas en este trabajo se resumen en la figura 51. Las proteínas TvTBPs, TvTFIIB y TvBDP1 reconocen DNA, sin embargo, se desconoce a qué región de los promotores se une.

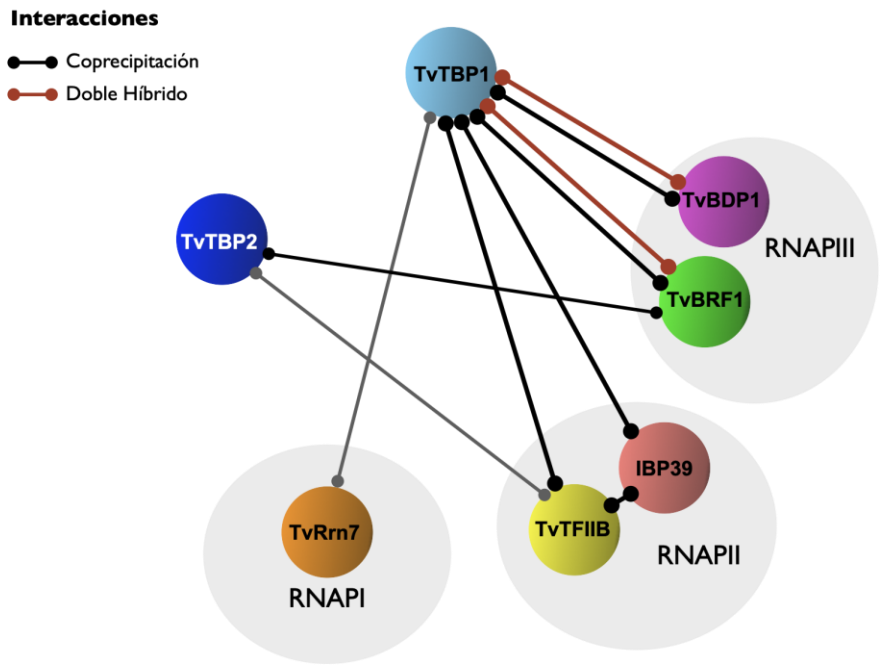


Figura 51.- Interacciones proteína-proteína de los tres complejos de inicio de la transcripción en *T. vaginalis*. TvRRN7 representa al complejo formado para la transcripción de la RNA polimerasa I; TvTFIIB e IBP39 forman parte de los complejos transcripcionales de la RNA polimerasa II y las proteínas TvBRF1 y TvBDP1 son parte de los complejos transcripcionales de

la RNA polimerasa III. Las líneas rojas son interacciones probadas por doble híbrido, miestras que las líneas negras y grises muestran as interacciones probadas por ensayos de coprecipitación. Las líneas grises se han representado en este color para indicar una interacción parcial, debido a que la proteína se observó solamente unida a las perlas (matriz), pero no se detectó en la elución, sugiriendo interacciones débiles y que deben ser probadas por otro método.

BIBLIOGRAFÍA

Aasland, R., Stewart, A. F. y Gibson, T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends in biochemical sciences* **21**, 87–8. doi: 10.1016/S0968-0004(96)30009-1.

Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E. y Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. *Nature* **553**, 301–306. doi: 10.1038/nature25441.

Adachi, N., Senda, M., Natsume, R., Senda, T. y Horikoshi, M. (2008). Crystal structure of Methanococcus jannaschii TATA box-binding protein. *Genes to Cells* **13**, 1127–1140. doi: 10.1111/j.1365-2443.2008.01233.x.

Akhtar, W. y Veenstra, G. J. C. (2011). TBP-related factors: A paradigm of diversity in transcription initiation. *Cell and Bioscience* **1**,. doi: 10.1186/2045-3701-1-23.

Alexander, D. E., Kaczorowski, D. J., Jackson-Fisher, A. J., Lowery, D. M., Zanton, S. J. y Pugh, B. F. (2004). Inhibition of TATA binding protein dimerization by RNA polymerase III transcription initiation factor Brf1. *The Journal of biological chemistry* **279**, 32401–6. doi: 10.1074/jbc.M405782200.

Ali, I. K. M., Ehrenkaufer, G. M., Hackney, J. A. y Singh, U. (2007). Growth of the protozoan parasite Entamoeba histolytica in 5-azacytidine has limited effects on parasite gene expression. *BMC genomics* **8**, 7. doi: 10.1186/1471-2164-8-7.

Alvarez-Sánchez, M. E., Avila-González, L., Becerril-García, C., Fattel-Facenda, L. V., Ortega-López, J. y Arroyo, R. (2000). A novel cystein proteinase (CP65) of Trichomonas vaginalis involved in cytotoxicity. *Microbial Pathogenesis* **28**, 193–202. doi: 10.1006/mpat.1999.0336.

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

Bártfai, R., Balduf, C., Hilton, T., Rathmann, Y., Hadzhiev, Y., Tora, L., Orbán, L. y Müller, F. (2004). TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish. *Current biology : CB* **14**, 593–8. doi: 10.1016/j.cub.2004.03.034.

Benchimol, M. (2004). Trichomonads under Microscopy. *Microscopy and Microanalysis* **10**, 528–550. doi: 10.1017/S1431927604040905.

Bouchemal, K., Bories, C. y Loiseau, P. M. (2017). crossm of Trichomonas vaginalis Infections. **30**, 811–825.

Bourbon, H. M. (2008). Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Research* **36**, 3993–4008. doi: 10.1093/nar/gkn349.

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

Brückner, A., Polge, C., Lentze, N., Auerbach, D. y Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Sciences* **10**, 2763–2788. doi: 10.3390/ijms10062763.

Bushnell, D. A., Cramer, P. y Kornberg, R. D. (2002). Structural basis of transcription: -Amanitin-RNA polymerase II cocrystal at 2.8 A resolution. *Proceedings of the National Academy of Sciences* **99**, 1218–1222. doi: 10.1073/pnas.251664698.

Butler, J. E. F. y Kadonaga, J. T. (2002). The RNA polymerase II core promoter: A key component in the regulation of gene expression. *Genes and Development* **16**, 2583–2592. doi: 10.1101/gad.1026202.

Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., Wortman, J. R., Bidwell, S. L., Alsmark, U. C. M., Besteiro, S., Sicheritz-Ponten, T., Noel, C. J., Dacks, J. B., Foster, P. G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G. J., Westrop, G. D., Muller, S., Dessi, D., Fiori, P. L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F. D., Simoes-Barbosa, A., Brown, M. T., Hayes, R. D., Mukherjee, M., Okumura, C. Y., Schneider, R., Smith, A. J., Vanacova, S., Villalvazo, M., Haas, B. J., Pertea, M., Feldblyum, T. V, Utterback, T. R., Shu, C.-L., Osoegawa, K., de Jong, P. J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S.-B., Logsdon, J. M., Henze, K., Gupta, A., Wang, C. C., Dunne, R. L., Upcroft, J. A., Upcroft, P., White, O., Salzberg, S. L., Tang, P., Chiu, C., Lee, Y.-S., Embley, T. M., Coombs, G. H., Mottram, J. C., Tachezy, J., Fraser-Liggett, C. M. y Johnson, P. J. (2007). Draft Genome Sequence of the Sexually Transmitted Pathogen Trichomonas vaginalis. *Science* **315**, 207–212. doi: 10.1126/science.1132894.

Castañon-Sanchez, C. A., Luna-Arias, J. P., de Dios-Bravo, M. G., Herrera-Aguirre, M. E., Olivares-Trejo, J. J., Orozco, E. y Hernandez, J. M. (2010). Entamoeba histolytica: A unicellular organism containing two active genes encoding for members of the TBP family. *Protein Expression and Purification* **70**, 48–59. doi: 10.1016/j.pep.2009.12.007.

Cavalier-Smith, T. y Chao, E. E. (1996). Molecular phylogeny of the free-living archezoan Trepomonas agilis and the nature of the first eukaryote. *Journal of Molecular Evolution* **43**, 551–562. doi: 10.1007/BF02202103.

Clark, C. G. y Diamond, L. S. (2002). Methods for cultivation of luminal parasitic protists of clinical importance. *Clinical Microbiology Reviews* **15**, 329–341. doi: 10.1128/CMR.15.3.329-341.2002.

Coleman, R. A., Taggart, A. K. P., Benjamin, L. R. y Pugh, B. F. (1995). Dimerization of the TATA binding protein. *The Journal of biological chemistry* **270**, 13842–9. doi: 10.1074/jbc.270.23.13842.

Coleman, R. A., Taggart, A. K. P., Burma, S., Chicca, J. J. y Pugh, B. F. (1999). TFIIA regulates TBP and TFIID dimers. *Molecular Cell* **4**, 451–457. doi: 10.1016/S1097-2765(00)80453-0.

Cong, P., Luo, Y., Bao, W. y Hu, S. (2010). Genomic organization and promoter analysis of the Trichomonas vaginalis core histone gene families. *Parasitology International* **59**, 29–34. doi: 10.1016/j.parint.2009.08.010.

Cormack, B. P. y Struhl, K. (1992). The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell* **69**, 685–696. doi: 10.1016/0092-8674(92)90232-2.

Crowley, T. E., Hoey, T., Liu, J., Jan, Y. N., Jan, L. Y. y Tjian, R. (1993). A new factor related to TAT A- restricted expression patterns in Drosophila. **361**, 557–561.

Dantonel, J. C., Wurtz, J. M., Poch, O., Moras, D. y Tora, L. (1999). The TBP-like factor: An alternative transcription factor in Metazoa? *Trends in Biochemical Sciences* **24**, 335–339. doi: 10.1016/S0968-0004(99)01436-X.

DeDecker, B. S., O'Brien, R., Fleming, P. J., Geiger, J. H., Jackson, S. P. y Sigler, P. B. (1996). The crystal structure of a hyperthermophilic archaeal TATA-box binding protein. *Journal of Molecular Biology* **264**, 1072–1084. doi: 10.1006/jmbi.1996.0697.

Diamond, L. S. (1957). The Establishment of Various Trichomonads of Animals and Man in Axenic Cultures Author (s): Louis S . Diamond Published by : The American Society of Parasitologists Stable URL : http://www.jstor.org/stable/3274682 Your use of the JSTOR archive indicates. *The Journal of Parasitology* **43**, 488–490.

Dikstein, R. (2011). The unexpected traits associated with core promoter elements. *Transcription* **2**, 201–206. doi: 10.4161/trns.2.5.17271.

Duttke, S. H. C. (2015). Evolution and diversification of the basal transcription machinery. *Trends in Biochemical Sciences* **40**, 127–129. doi: 10.1016/j.tibs.2015.01.005.

Einhoff, W., Kummer, H. y Rüdiger, H. (1987). Mercaptoethanol assists in cleaving a cysteine/cystine-free protein into its subunits. *Biological chemistry Hoppe-Seyler* **368**, 405–7. doi: 10.1515/bchm3.1987.368.1.405.

Engel, C., Gubbey, T., Neyer, S., Sainsbury, S., Oberthuer, C., Baejen, C., Bernecky, C. y Cramer, P. (2017). Structural Basis of RNA Polymerase I Transcription Initiation. *Cell* **169**, 120-131.e22. doi: 10.1016/j.cell.2017.03.003.

Espinosa, N., Hernández, R., López-Griego, L., Arroyo, R. y López-Villaseñor, I. (2001). Differences between coding and non-coding regions in the Trichomonas vaginalis genome: an actin gene as a locus model1Data deposition: The sequence reported in this paper has been deposited in the GeneBank database under the accession no. AF237734.1. *Acta Tropica* **78**, 147–154. doi: 10.1016/s0001-706x(00)00180-7.

Estojak, J., Brent, R. y Golemis, E. A. (1995). Correlation of two-hybrid affinity data with in vitro measurements. *Molecular and Cellular Biology* **15**, 5820–5829. doi: 10.1128/mcb.15.10.5820.

Franco, B., Hernández, R. y López-Villaseñor, I. (2012). Trichomonas vaginalis ribosomal RNA: Identification and characterisation of the transcription promoter and terminator sequences. *Molecular and Biochemical Parasitology* **185**, 1–9. doi: 10.1016/j.molbiopara.2012.05.004.

Franco Bárcenas, B. (2012). Caracterización de la región promotora y terminadora de la transcripción del cistrón ribosomal de Trichomonas vaginalis.

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

Geiduschek, E. P. y Kassavetis, G. A. (2001). The RNA polymerase III transcription apparatus. *Journal of Molecular Biology* **310**, 1–26. doi: 10.1006/jmbi.2001.4732.

Golemis, E. A., Serebriiskii, I., Finley, R. L., Kolonin, M. G., Gyuris, J. y Brent, R. (2000). Interaction Trap/Two-Hybrid System to Identify Interacting Proteins. *Current Protocols in Cell Biology* **8**, 1–7. doi: 10.1002/0471143030.cb1703s08.

Gouge, J., Guthertz, N., Kramm, K., Dergai, O., Abascal-Palacios, G., Satia, K., Cousin, P., Hernandez, N., Grohmann, D. y Vannini, A. (2017). Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation. *Nature Communications* **8**, 1–10. doi: 10.1038/s41467-017-00126-1.

Gould, S. B., Woehle, C., Kusdian, G., Landan, G., Tachezy, J., Zimorski, V. y Martin, W. F. (2013). Deep sequencing of Trichomonas vaginalis during the early infection of vaginal epithelial cells and amoeboid transition. *International Journal for Parasitology* **43**, 707–719. doi: 10.1016/j.ijpara.2013.04.002.

Grigorian, A. L., Bustamante, J. J., Hernandez, P., Martinez, A. O. y Haro, L. S. (2005). Extraordinarily stable disulfide-linked homodimer of human growth hormone. *Protein science : a publication of the Protein Society* **14**, 902–13. doi: 10.1110/ps.041048805.

Guillebault, D., Sasorith, S., Derelle, E., Wurtz, J. M., Lozano, J. C., Bingham, S., Tora, L. y 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. *Journal of Biological Chemistry* **277**, 40881–40886. doi: 10.1074/jbc.M205624200.

Hansen, S. K., Takada, S., Jacobson, R. H., Lis, J. T. y Tjian, R. (1997). Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* **91**, 71–83. doi: 10.1016/S0092-8674(01)80010-6.

Harp, D. F. y Chowdhury, I. (2011). Trichomoniasis: Evaluation to execution. *European Journal of Obstetrics and Gynecology and Reproductive Biology* **157**, 3–9. doi: 10.1016/j.ejogrb.2011.02.024.

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

Hsin, J.-P., Xiang, K. y Manley, J. L. (2014). Function and Control of RNA Polymerase II C-Terminal Domain Phosphorylation in Vertebrate Transcription and RNA Processing. *Molecular and Cellular Biology* **34**, 2488–2498. doi: 10.1128/mcb.00181-14.

Iyer, L. M., Anantharaman, V., Wolf, M. Y. y Aravind, L. (2008). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *International Journal for Parasitology* **38**, 1–31. doi: 10.1016/j.ijpara.2007.07.018.

Jackson-Fisher, A. J., Chitikila, C., Mitra, M. y Pugh, B. F. (1999). A role for TBP dimerization in preventing unregulated gene expression. *Molecular Cell* **3**, 717–727. doi: 10.1016/S1097-

2765(01)80004-6.

Jallow, Z., Jacobi, U. G., Weeks, D. L., Dawid, I. B. y Veenstra, G. J. C. (2004). Specialized and redundant roles of TBP and a vertebrate-specific TBP paralog in embryonic gene regulation in Xenopus. **101**, 13525–13530.

Juven-Gershon, T. y Kadonaga, J. T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Developmental Biology* **339**, 225–229. doi: 10.1016/j.ydbio.2009.08.009.

Kang, J. J., Kang, Y. S. y Stumph, W. E. (2016). TFIIIB subunit locations on U6 gene promoter DNA mapped by site-specific protein-DNA photo-cross-linking. *FEBS Letters* **590**, 1488–1497. doi: 10.1002/1873-3468.12185.

Kassavetis, G. A. y Geiduschek, E. P. (2006). Transcription factor TFIIIB and transcription by RNA polymerase III: Figure 1. *Biochemical Society Transactions* **34**, 1082–1087. doi: 10.1042/bst0341082.

Kassavetis, G. A., Nguyen, S. T., Kobayashi, R., Kumar, A., Geiduschek, E. P. y Pisano, M. (1995). Cloning, expression, and function of TFC5, the gene encoding the B" component of the Saccharomyces cerevisiae RNA polymerase III transcription factor TFIIIB. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 9786–90. doi: 10.1073/pnas.92.21.9786.

Kawakami, E., Adachi, N., Senda, T. y Horikoshi, M. (2017). Leading role of TBP in the Establishment of Complexity in Eukaryotic Transcription Initiation Systems. *Cell Reports* **21**, 3941–3956. doi: 10.1016/j.celrep.2017.12.034.

Kelter, A. R., Herchenbach, J. y Wirth, B. (2000). The transcription factor-like nuclear regulator (TFNR) contains a novel 55-amino-acid motif repeated nine times and maps closely to SMN1. *Genomics* **70**, 315–26. doi: 10.1006/geno.2000.6396.

Khoo, S.-K., Wu, C.-C., Lin, Y.-C., Lee, J.-C. y Chen, H.-T. (2014). Mapping the protein interaction network for TFIIB-related factor Brf1 in the RNA polymerase III preinitiation complex. *Molecular and cellular biology* **34**, 551–9. doi: 10.1128/MCB.00910-13.

Knutson, B. A. y Hahn, S. (2011). Yeast Rrn7 and human TAF1B are TFIIB-related RNA polymerase I general transcription factors. *Science (New York, N.Y.)* **333**, 1637–40. doi: 10.1126/science.1207699.

Knutson, B. A. y Hahn, S. (2013). TFIIB-related factors in RNA polymerase I transcription. *Biochimica et biophysica acta* **1829**, 265–73. doi: 10.1016/j.bbagrm.2012.08.003.

Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M. y Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature* **462**, 323–330. doi: 10.1038/nature08548.

Kramm, K., Engel, C. y Grohmann, D. (2019). Transcription initiation factor TBP: Old friend new questions. *Biochemical Society Transactions* **47**, 411–423. doi: 10.1042/BST20180623.

Kulda, J. (1999). Trichomonads, hydrogenosomes and drug resistance. *International journal for parasitology* **29**, 199–212. doi: 10.1016/s0020-7519(98)00155-6.

Kurshakova, M. M., Nabirochkina, E. N., Georgieva, S. G. y Kopytova, D. V. (2019). TRF4, the novel TBP-related protein of Drosophila melanogaster, is concentrated at the endoplasmic reticulum and copurifies with proteins participating in the processes associated with endoplasmic reticulum. *Journal of Cellular Biochemistry* **120**, 7927–7939. doi: 10.1002/jcb.28070.

Lahti, C. J., D'Oliveira, C. E. y Johnson, P. J. (1992). β-Succinyl-coenzyme A synthetase from Trichomonas vaginalis is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. *Journal of Bacteriology* **174**, 6822–6830.

Lane, D., Prentki, P. y Chandler, M. (1992). Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiological reviews* **56**, 509–28.

Lau, A. O. T., Liston, D. R., Vanacova, S. y Johnson, P. J. (2003). Trichomonas vaginalis initiator binding protein, IBP39, contains a novel DNA binding motif. *Molecular and Biochemical Parasitology* **130**, 167–171. doi: 10.1016/S0166-6851(03)00172-5.

Lau, A. O. T., Smith, A. J., Brown, M. T. y Johnson, P. J. (2006). Trichomonas vaginalis initiator binding protein (IBP39) and RNA polymerase II large subunit carboxy terminal domain interaction. *Molecular and Biochemical Parasitology* **150**, 56–62. doi: 10.1016/j.molbiopara.2006.06.008.

Layat, E., Probst, A. V. y Tourmente, S. (2013). Structure, function and regulation of Transcription Factor IIIA: From Xenopus to Arabidopsis. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1829**, 274–282. doi: 10.1016/j.bbagrm.2012.10.013.

Lee, Y., Wong, W. M., Guyer, D., Erkine, A. M. y Nazar, R. N. (1997). In vivo analyses of upstream promoter sequence elements in the 5 S rRNA gene from Saccharomyces cerevisiae. *Journal of Molecular Biology* **269**, 676–683. doi: 10.1006/jmbi.1997.1071.

Lehker, M. W., Arroyo, R. y Alderete, J. F. (1991). The regulation by iron of the synthesis of adhesins and cytoadherence levels in the protozoan Trichomonas vaginalis. *Journal of Experimental Medicine* **174**, 311–318. doi: 10.1084/jem.174.2.311.

Liston, D. R. y Johnson, P. J. (1998). Gene transcription in Trichomonas vaginalis. *Parasitology Today* **14**, 261–265. doi: 10.1016/S0169-4758(98)01264-2.

Liston, D. R. y Johnson, P. J. (1999). Analysis of a Ubiquitous Promoter Element in a Primitive Eukaryote: Early Evolution of the Initiator Element. *Molecular and Cellular Biology* **19**, 2380–2388. doi: 10.1128/mcb.19.3.2380.

Liston, D. R., Carrero, J. C. y Johnson, P. J. (1999). Upstream regulatory sequences required for expression of the Trichomonas vaginalis α-succinyl CoA synthetase gene. *Molecular and Biochemical Parasitology* **104**, 323–329. doi: 10.1016/S0166-6851(99)00137-1.

Liston, D. R., Lau, A. O. T., Ortiz, D., Smale, S. T. y Johnson, P. J. (2001). Initiator Recognition in a Primitive Eukaryote: IBP39, an Initiator-Binding Protein from Trichomonas vaginalis. *Molecular and Cellular Biology* **21**, 7872–7882. doi: 10.1128/mcb.21.22.7872-7882.2001.

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

Marinets, A., Müller, M., Johnson, P. J., Kulda, J., Scheiner, O., Wiedermann, G. y Duchêne, M. (1996). The sequence and organization of the core histone H3 and H4 genes in the early branching amitochondriate protist Trichomonas vaginalis. *Journal of molecular evolution* **43**, 563–71. doi: 10.1007/bf02202104.

Marmorstein, R. (2003). Transcription initiation at its most basic level. *Cell* **115**, 370–2. doi: 10.1016/s0092-8674(03)00888-2.

Martínez-Calvillo, S., Vizuet-De-Rueda, J. C., Florencio-Martínez, L. E., Manning-Cela, R. G. y Figueroa-Angulo, E. E. (2010). Gene expression in trypanosomatid parasites. *Journal of Biomedicine and Biotechnology* **2010**,. doi: 10.1155/2010/525241.

Millership, J. J., Waghela, P., Cai, X., Cockerham, A. y Zhu, G. (2004). Differential expression and interaction of transcription co-activator MBF1 with TATA-binding protein (TBP) in the apicomplexan Cryptosporidium parvum. *Microbiology* **150**, 1207–1213. doi: 10.1099/mic.0.26891-0.

Moore, P. A., Ozer, J., Salunek, M., Jan, G., Zerby, D., Campbell, S. y Lieberman, P. M. (1999). A Human TATA Binding Protein-Related Protein with Altered DNA Binding Specificity Inhibits Transcription from Multiple Promoters and Activators. **19**, 7610–7620.

Naidu, S., Friedrich, J. K. y Russell, J. (2011). TAF1B Is a TFIIB-Like Component. *Science* **333**, 2010–2012.

Narayanasamy, R. K., Castañón-sanchez, C. A., Luna-arias, J. P., García-rivera, G., Avendaño-borromeo, B., Labra-barrios, M. L., Valdés, J., Herrera-aguirre, M. E. y Orozco, E. (2018). The Entamoeba histolytica TBP and TRF1 transcription factors are GAAC-box binding proteins , which display differential gene expression under different stress stimuli and during the interaction with mammalian cells. 1–16.

Nikolov, D. B., Hu, S. H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N. H., Roeder, R. G. y Burley, S. K. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* **360**, 40–46. doi: 10.1038/360040a0.

Parra-Marín, O., Rosas-Hernández, L., López-Pacheco, K., Franco, B., Ibáñez-Escribano, A., Hernández, R. y López-Villaseñor, I. (2019). An in vitro characterisation of the Trichomonas vaginalis TATA box-binding proteins (TBPs). *Parasitology research* **118**, 3019–3031. doi: 10.1007/s00436-019-06438-z.

Patikoglou, G. A., Kim, J. L., Sun, L., Yang, S. H., Kodadek, T. y Burley, S. K. (1999). TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes & development* **13**, 3217–30. doi: 10.1101/gad.13.24.3217.

Pereira-Neves, A., Ribeiro, K. C. y Benchimol, M. (2003). Pseudocysts in Trichomonads – New Insights. *Protist* **154**, 313–329. doi: 10.1078/143446103322454095.

Persengiev, S. P., Zhu, X., Dixit, B. L., Maston, G. A., Kittler, E. L. W. y Green, M. R. (2003). Vertebrate-Specific and Widely Expressed. *Sciences-New York* **100**,.

Petrin, D., Delgaty, K., Bhatt, R. y Garber, G. (1998). Clinical and Microbiological Aspects of Trichomonas vaginalis. **11**, 300–317.

Quon, D. V., Delgadillo, M. G., Khachi, A., Smale, S. T. y Johnson, P. J. (1994). Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 4579–83. doi: 10.1073/pnas.91.10.4579.

Quon, D. V. K., Delgadillo, M. G. y Johnson, P. J. (1996). Transcription in the early diverging eukaryote Trichomonas vaginalis: an unusual RNA polymerase II and alpha-amanitin-resistant transcription of protein-coding genes. *Journal of molecular evolution* **43**, 253–62. doi: 10.1007/bf02338833.

Rabenstein, M. D., Zhou, S., Lis, J. T. y Tjian, R. (1999). TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 4791–6. doi: 10.1073/pnas.96.9.4791.

Ribeiro, K. C., Pereira-Neves, A. y Benchimol, M. (2002). The mitotic spindle and associated membranes in the closed mitosis of trichomonads. *Biology of the Cell* **94**, 157–172. doi: 10.1016/S0248-4900(02)01191-7.

Roberts, S., Miller, S. J., Lane, W. S., Lee, S. y Hahn, S. (1996). Cloning and functional characterization of the gene encoding the TFIIIB90 subunit of RNA polymerase III transcription factor TFIIIB. *The Journal of biological chemistry* **271**, 14903–9. doi: 10.1074/jbc.271.25.14903.

Roder, K. y Schweizer, M. (2001). Running-buffer composition influences DNA–protein and protein–protein complexes detected by electrophoretic mobility-shift assay (EMSA). *Biotechnology and Applied Biochemistry* **33**, 209. doi: 10.1042/ba20010007.

Ruan, J. -p., Arhin, G. K., Ullu, E. y Tschudi, C. (2004). Functional Characterization of a Trypanosoma brucei TATA-Binding Protein-Related Factor Points to a Universal Regulator of Transcription in Trypanosomes. *Molecular and Cellular Biology* **24**, 9610–9618. doi: 10.1128/mcb.24.21.9610-9618.2004.

Rüth, J., Conesa, C., Dieci, G., Lefebvre, O., Düsterhöft, A., Ottonello, S. y Sentenac, A. (1996). A suppressor of mutations in the class III transcription system encodes a component of yeast TFIIIB. *The EMBO Journal* **15**, 1941–1949. doi: 10.1002/j.1460-2075.1996.tb00545.x.

Ruvalcaba-Salazar, O. K., Ramírez-Estudillo, M. D. C., Montiel-Condado, D., Recillas-Targa, F., Vargas, M. y Hernández-Rivas, R. (2005). Recombinant and native Plasmodium falciparum TATA-binding-protein binds to a specific TATA box element in promoter regions. *Molecular and Biochemical Parasitology* **140**, 183–196. doi: 10.1016/j.molbiopara.2005.01.002.

Schramm, L., Pendergrast, P. S., Sun, Y. y Hernandez, N. (2000). Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes and Development* **14**, 2650–2663. doi: 10.1101/gad.836400.

Schumacher, M. A., Lau, A. O. T. y Johnson, P. J. (2003). Structural Basis of Core Promoter Recognition in a Primitive Eukaryote. **115**, 413–424.

Simoes-Barbosa, A., Meloni, D., Wohlschlegel, J. A., Konarska, M. M. y Johnson, P. J. (2008).

Spliceosomal snRNAs in the unicellular eukaryote Trichomonas vaginalis are structurally conserved but lack a 5'-cap structure. *Rna* **14**, 1617–1631. doi: 10.1261/rna.1045408.

Singh, S., Singh, G., Singh, A. K., Gautam, B., Farmer, R., Lodhi, S. S. y Wadhwa, G. (2011). Prediction and analysis of paralogous proteins in Trichomonas vaginalis genome. *Bioinformation* **6**, 31–34. doi: 10.6026/97320630006031.

Smale, S. T. y Kadonaga, J. T. (2003). The RNA Polymerase II Core Promoter. *Annual Review of Biochemistry* **72**, 449–479. doi: 10.1146/annurev.biochem.72.121801.161520.

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

Smith, A. J., Chudnovsky, L., Simoes-Barbosa, A., Delgadillo-Correa, M. G., Jonsson, Z. O., Wohlschlegel, J. A. y Johnson, P. J. (2011). Novel Core Promoter Elements and a Cognate Transcription Factor in the Divergent Unicellular Eukaryote Trichomonas vaginalis. *Molecular and Cellular Biology* **31**, 1444–1458. doi: 10.1128/mcb.00745-10.

Söding, J., Biegert, A. y Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* **33**, 244–248. doi: 10.1093/nar/gki408.

Sogin, M. L. (1991). Early evolution and the origin of eukaryotes. *Current Opinion in Genetics and Development* **1**, 457–463. doi: 10.1016/S0959-437X(05)80192-3.

Sogin, M. L. y Silberman, J. D. (1998). Evolution of the protists and protistan parasites from the perspective of molecular systematics. *International journal for parasitology* **28**, 11–20. doi: 10.1016/s0020-7519(97)00181-1.

Takada, S., Lis, J. T., Zhou, S. y Tjian, R. (2000). A TRF1:BRF complex directs Drosophila RNA polymerase III transcription. *Cell* **101**, 459–469. doi: 10.1016/S0092-8674(00)80857-0.

Thomas, M. C. y Chiang, C. M. (2006). The general transcription machinery and general cofactors. *Critical reviews in biochemistry and molecular biology* **41**, 105–178. doi: 10.1080/10409230600648736.

Torres-Machorro, A. L., Hernández, R., Sánchez, J. y López-Villaseñor, I. (2006). The 5S ribosomal RNA gene from the early diverging protozoa Trichomonas vaginalis. *Molecular and Biochemical Parasitology* **145**, 269–273. doi: 10.1016/j.molbiopara.2005.10.009.

Torres Machorro, A. L. (2010). Obtención de un sistema de análisis para genes transcritos por RNA polimerasa III en Trichomonádidos. Estudios del gen RNAr 5S.

Tsai, C.-D., Liu, H.-W. y Tai, J.-H. (2002). Characterization of an iron-responsive promoter in the protozoan pathogen Trichomonas vaginalis. *The Journal of biological chemistry* **277**, 5153–62. doi: 10.1074/jbc.M110234200.

Turowski, T. W. y Tollervey, D. (2016). Transcription by RNA polymerase III: insights into mechanism and regulation. *Biochemical Society Transactions* **44**, 1367–1375. doi: 10.1042/bst20160062.

Vanacova, S., Liston, D. R., Tachezy, J. y Johnson, P. J. (2003). Molecular biology of the amitochondriate parasites, Giardia intestinalis, Entamoeba histolytica and Trichomonas

vaginalis. *International Journal for Parasitology* **33**, 235–255. doi: 10.1016/S0020-7519(02)00267-9.

Vanácová, S., Tachezy, J., Ullu, E. y Tschudi, C. (2001). Unusual diversity in alpha-amanitin sensitivity of RNA polymerases in trichomonads. *Molecular and biochemical parasitology* **115**, 239–47. doi: 10.1016/s0166-6851(01)00294-8.

Vannini, A. (2013). A structural perspective on RNA polymerase I and RNA polymerase III transcription machineries. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* **1829**, 258–264. doi: 10.1016/j.bbagrm.2012.09.009.

Vannini, A. y Cramer, P. (2012). Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries. *Molecular Cell* **45**, 439–446. doi: 10.1016/j.molcel.2012.01.023.

Verma, N., Hung, K. H., Kang, J. J., Barakat, N. H. y Stumph, W. E. (2013). Differential utilization of TATA box-binding protein (TBP) and TBP-related factor 1 (TRF1) at different classes of RNA polymerase III promoters. *Journal of Biological Chemistry* **288**, 27564–27570. doi: 10.1074/jbc.C113.503094.

Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. H. y Müller, C. W. (2018). Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553**, 295–300. doi: 10.1038/nature25440.

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

Wong, J. M., Liu, F. y Bateman, E. (1992). Cloning and expression of the Acanthamoeba castellanii gene encoding transcription factor TFIID. *Gene* **117**, 91–97. doi: 10.1016/0378-1119(92)90494-A.

Zehavi, Y., Kedmi, A., Ideses, D. y Juven-Gershon, T. (2015). TRF2: TRansForming the view of general transcription factors. *Transcription* **6**, 1–6. doi: 10.1080/21541264.2015.1004980.

ANEXO 1: Amortiguador de unión probados en los ensayos de retardo de la movilidad electroforética (EMSA)

En este trabajo se probaron las siguientes condiciones:

Amortiguador de corrida: TBE al 0.5 % debido a que presenta una menor fuerza iónica que el amortiguador TG.

Concentración del gel: Anteriormente se usó geles al 6 % de acrilamida-bisacrilamida 29:1, sin embargo, en algunos casos, la muestra se quedaba en el pozo, por tal motivo, se decidió bajar la concentración al 4 – 4.5 %.

Amortiguador de unión: Este amortiguador debe contener las concentraciones de iones y pH necesarios para llevar a cabo la interacción, por tal motivo, se decidió probar cuatro distintos amortiguadores de unión, como se muestra a continuación:

Amortiguador de unión 1: Wiesler & Weinzierl, 2010			
Componentes:	[5X]	Usar:	[Final= 1X]
HEPES 1 M pH 7.9	200 mM	100 µl	40 mM
NaCl 5 M	1250 mM	125 µl	250 mM
EDTA 0.5 M	0.5 mM	0.5 µl	0.1 mM
$MgCl_2$ 0.5 M	12.5 mM	12.5 µl	2.5 mM
Glicerol	25 %	125 µl	5 %
β-mercaptoetanol 0.5 M	25 mM	25 µL	5 mM
Agua bidestilada	c.b.p. 500 µl	112 µl	

Amortiguador de unión 2: Glossop et al., 2004			
Componentes:	[5X]	Usar:	[Final= 1X]
HEPES 1 M pH 7.9	50 mM	25 µl	10 mM
$MgCl_2$ 0.5 M	20 mM	20 µl	4 mM
Sulfato de amonio 1 M	25 mM	12.5 µl	5 mM
Glicerol	40 %	200 µl	8 %
PEG 40 %	10 %	125 µl	2 %
KCl 2 M	275 mM	68.75 µl	55 mM
β-mercaptoetanol 0.5 M	25 mM	25 µl	5 mM
EDTA 1 M	1 mM	0.5 µl	0.2 mM
PMSF 100 mM	1 mM	5 µl	0.2 mM
Agua bidestilada	c.b.p. 500 µl	18.25 µl	

Amortiguador de unión 3: Amortiguador con Zinc			
Componentes:	[5X]	Usar:	[Final= 1X]
Tris-HCl 1 M pH 8	100 mM	50 µl	20 mM
NaCl 5 M	500 mM	50 µl	100 mM
Glicerol	25 %	125µl	5 %
BSA	0.5	25 µl	0.1 µg/µl
$MgCl_2$	2.5 mM	2.5 µl	0.5 mM
$ZnS0_4$	2.5 mM	1.25 µl	0.5 mM
Agua bidestilada	c.b.p. 500 µl	246.25 µl	

Amortiguador de unión 4:			
Componentes:	[5X]	Usar:	[Final= 1X]
HEPES 1 M pH 7.9	100 mM	50 µl	20 mM
KCl 2 M	325 mM	81.25 µl	65 mM
EDTA 1 M	5 mM	2.5 µl	1 mM
$MgCl_2$ 500 mM	40 mM	40 µl	8 mM
$ZnS0_4$	2.5 mM	2.5 µl	0.5 mM
Glicerol	30 %	150 µl	6 %
NP40 10 %	0.05 %	2.5 µl	0.01 %
DTT 1 M	10 mM	5 µl	2 mM
Agua bidestilada	c.b.p. 500 µl	166.25 µl	

Amortiguador de unión 5:			
Componentes:	[5X]	Usar:	[Final= 1X]
HEPES 1 M pH 7.9	100 mM	50 µl	20 mM
KCl 2 M	500 mM	125 µl	100 mM
EDTA 1 M	5 mM	2.5 µl	1 mM
$MgCl_2$ 500 mM	25 mM	25 µl	5 mM
Glicerol	25 %	125 µl	5 %
BSA	500 ng/µl	25 µl	100 ng/µl
DTT	5 mM	2.5 µl	1 mM
Agua bidestilada	c.b.p. 500 µl	145 µl	

ANEXO 2: Datos de la actividad β-galactosidasa

Ensayo de Actividad Beta-galactosidasa														
No. de Tubo	Muestra	Clona	OD 420 nm	Volumen de reacción (ml)	Coef. de Ext. ONPG	[mg/ml]	Muestra analizada (ml)	Tiempo de reacción (min)	Act. Específica (nmoles/min /mg de proteína)	Act. Esp. Promedio	SD por clona	Act. Esp. Promedio Total	SD por interacción	Error Estándar
1	Alfa-Beta	A	0.401	1.7	0.0045	2.7	0.025	0.5	4411.3	4840.5	607.1	5251.1	1606.2	927.3
			0.32	1.7	0.0045	2.7	0.05	0.167	5269.8					
		B	0.463	1.7	0.0045	1.9	0.025	1	3743.6	3890.1	207.3			
			0.666	1.7	0.0045	1.9	0.05	0.667	4036.7					
		C	0.913	1.7	0.0045	4.1	0.025	0.5	6799.4	7022.8	315.8			
			1.296	1.7	0.0045	4.1	0.05	0.333	7246.1					
2	TBP1-BRF1	A	0.602	1.7	0.0045	6.0	0.025	0.567	2664.2	2794.8	184.7	3554.6	689.8	398.2
			0.777	1.7	0.0045	6.0	0.05	0.333	2925.4					
		B	0.454	1.7	0.0045	2.3	0.025	0.833	3621.2	3727.5	150.3			
			0.577	1.7	0.0045	2.3	0.05	0.5	3833.7					
		C	0.802	1.7	0.0045	4.2	0.025	0.667	4319.1	4141.4	251.3			
			1.472	1.7	0.0045	4.2	0.05	0.667	3963.7					
3	TBP1-BDP1	A	0.095	1.7	0.0045	6.3	0.025	25	9.2	9.1	0.0	22.8	20.0	11.5
			0.189	1.7	0.0045	6.3	0.05	25	9.1					
		B	0.129	1.7	0.0045	2.9	0.025	15	44.2	45.7	2.2			
			0.276	1.7	0.0045	2.9	0.05	15	47.3					
		C	0.078	1.7	0.0045	2.9	0.025	30	13.6	13.6	0.1			
			0.157	1.7	0.0045	2.9	0.05	30	13.7					
4	TBP1-pJG4.5	A	0.096	1.7	0.0045	7.7	0.025	20	9.5	10.0	0.7	17.6	9.1	5.3
			0.213	1.7	0.0045	7.7	0.05	20	10.5					
		B	0.067	1.7	0.0045	2.3	0.025	30	14.5	15.2	1.0			
			0.147	1.7	0.0045	2.3	0.05	30	15.9					
		C	0.146	1.7	0.0045	2.2	0.025	30	34.2	27.7	9.1			
			0.182	1.7	0.0045	2.2	0.05	30	21.3					
5	TBP1-Endo	B	0.067	1.7	0.0045	2.3	0.025	30	14.9	14.9	0.1	22.5	10.8	7.7
			0.133	1.7	0.0045	2.3	0.05	30	14.8					
		C	0.176	1.7	0.0045	2.6	0.025	30	33.8	30.2	5.0			
			0.278	1.7	0.0045	2.6	0.05	30	26.7					
6	TBP1-X	A	0.054	1.7	0.0045	3.7	0.025	30	7.3	6.7	0.9	5.9	1.8	1.0
			0.089	1.7	0.0045	3.7	0.05	30	6.0					
		B	0.047	1.7	0.0045	3.8	0.025	25	7.4	7.3	0.2			
			0.091	1.7	0.0045	3.8	0.05	25	7.1					
		C	0.012	1.7	0.0045	2.4	0.025	25	3.0	3.9	1.2			
			0.038	1.7	0.0045	2.4	0.05	25	4.8					
7	TBP2-BRF1	A	0.014	1.7	0.0045	3.8	0.025	25	2.2	2.2	0.1	2.9	0.9	0.5
			0.029	1.7	0.0045	3.8	0.05	25	2.3					
		B	0.008	1.7	0.0045	1.7	0.025	30	2.4	2.5	0.2			
			0.018	1.7	0.0045	1.7	0.05	30	2.7					
		C	0.026	1.7	0.0045	2.7	0.025	30	4.8	3.9	1.2			
			0.033	1.7	0.0045	2.7	0.05	30	3.1					
8	TBP2-BDP1	A	0.019	1.7	0.0045	4.8	0.025	25	2.4	2.6	0.3	5.2	3.4	2.0
			0.045	1.7	0.0045	4.8	0.05	25	2.8					
		B	0.038	1.7	0.0045	2.1	0.025	30	9.1	9.1	0.0			
			0.076	1.7	0.0045	2.1	0.05	30	9.1					
		C	0.013	1.7	0.0045	1.8	0.025	30	3.6	3.9	0.4			
			0.03	1.7	0.0045	1.8	0.05	30	4.2					
9	TBP2-pJG4.5	A	0.03	1.7	0.0045	5.9	0.025	25	3.1	2.7	0.5	8.1	4.6	2.7
			0.047	1.7	0.0045	5.9	0.05	25	2.4					
		B	0.077	1.7	0.0045	3.6	0.025	30	10.9	11.2	0.5			
			0.164	1.7	0.0045	3.6	0.05	30	11.6					
		C	0.133	1.7	0.0045	9.4	0.025	20.333	10.5	10.2	0.4			
			0.252	1.7	0.0045	9.4	0.05	20.5	9.9					
10	TBP2-Endo	B	0.038	1.7	0.0045	3.0	0.025	30	6.5	5.3	1.7	6.6	1.9	1.3
			0.048	1.7	0.0045	3.0	0.05	30	4.1					
		C	0.068	1.7	0.0045	4.1	0.025	30	8.4	7.9	0.7			
			0.121	1.7	0.0045	4.1	0.05	30	7.5					
11	TBP2-X	A	0.005	1.7	0.0045	0.5	0.025	30	4.9	4.2	1.0	3.2	1.4	0.8
			0.007	1.7	0.0045	0.5	0.05	30	3.4					
		B	0.016	1.7	0.0045	2.8	0.025	25	3.5	3.8	0.4			
			0.037	1.7	0.0045	2.8	0.05	25	4.0					
		C	0.007	1.7	0.0045	2.8	0.025	25	1.5	1.6	0.2			
			0.016	1.7	0.0045	2.8	0.05	25	1.7					
12	BDP1-BRF1	A	0.131	1.7	0.0045	5.3	0.025	26	14.5	15.8	1.8	75.4	48.9	21.9
			0.273	1.7	0.0045	5.3	0.05	23	17.1					
		B	0.128	1.7	0.0045	2.9	0.025	15	45.1	44.5	0.9			
			0.249	1.7	0.0045	2.9	0.05	15	43.9					
		C	0.238	1.7	0.0045	4.7	0.025	10	76.9	74.6	3.2			
			0.448	1.7	0.0045	4.7	0.05	10	72.3					
		A	0.225	1.7	0.0045	2.9	0.05	6	96.1	99.8	5.2			
			0.323	1.7	0.0045	2.9	0.1	4	103.5					
		B	0.468	1.7	0.0045	2.5	0.05	10	143.8	142.2	2.2			
			0.458	1.7	0.0045	2.5	0.1	5	140.7					
13	BDP1-pJG4.5	A	0.305	1.7	0.0045	2.9	0.025	13	123.6	124.8	1.7	128.5	63.7	36.8
			0.271	1.7	0.0045	2.9	0.05	5.667	125.9					
		B	0.271	1.7	0.0045	5.0	0.025	12	67.9	66.8	1.5			
			0.525	1.7	0.0045	5.0	0.05	12	65.7					
		C	0.39	1.7	0.0045	2.9	0.025	10	202.6	194.0	12.1			
			0.714	1.7	0.0045	2.9	0.05	10	185.4					

Ensayo de Actividad Beta-galactosidasa														
No. de Tubo	Muestra	Clona	OD 420 nm	Volumen de reacción (ml)	Coef. de Ext. ONPG	[mg/ml]	Muestra analizada (ml)	Tiempo de reacción (min)	Act. Específica (nmoles/min /mg de proteína)	Act. Esp. Promedio	SD por clona	Act. Esp. Promedio Total	SD por interacción	Error Estándar
14	BDP1-Endo	A	0.138	1.7	0.0045	4.0	0.025	14.5	36.3	57.8	30.5	74.3	30.2	17.5
			0.236	1.7	0.0045	4.0	0.05	5.667	79.4					
		B	0.205	1.7	0.0045	4.4	0.025	13	53.6	55.9	3.2			
			0.445	1.7	0.0045	4.4	0.05	13	58.2					
		C	0.227	1.7	0.0045	2.9	0.025	11	109.2	109.2	0.0			
			0.454	1.7	0.0045	2.9	0.05	11	109.2					
15	BDP1-X	A	0.421	1.7	0.0045	4.0	0.025	20	79.2	80.6	2.0	62.8	12.1	5.4
			0.436	1.7	0.0045	4.0	0.05	10	82.0					
		B	0.124	1.7	0.0045	2.1	0.025	15	58.8	57.5	1.8			
			0.237	1.7	0.0045	2.1	0.05	15	56.1					
		C	0.169	1.7	0.0045	3.0	0.025	15	56.1	57.8	2.5			
			0.359	1.7	0.0045	3.0	0.05	15	59.6					
		A	0.273	1.7	0.0045	3.6	0.05	12.167	47.5	49.4	2.6			
			0.29	1.7	0.0045	3.6	0.1	6	51.2					
		B	0.382	1.7	0.0045	4.1	0.05	10.167	69.4	68.5	1.3			
			0.366	1.7	0.0045	4.1	0.1	5	67.6					
16	BRF1-BDP1	A	0.392	1.7	0.0045	5.1	0.05	4.667	124.2	286.0	228.9	221.0	54.3	24.3
			0.749	1.7	0.0045	1.5	0.1	4.333	447.9					
		B	0.378	1.7	0.0045	3.1	0.05	3	303.3	206.2	137.3			
			0.673	1.7	0.0045	6.4	0.1	3.667	109.1					
		C	0.448	1.7	0.0045	6.2	0.05	4.5	121.7	168.5	66.2			
			0.456	1.7	0.0045	3.4	0.1	2.333	215.3					
		A	0.363	1.7	0.0045	2.7	0.05	4	258.3	270.0	16.6			
			0.396	1.7	0.0045	2.7	0.1	2	281.8					
		B***	0.012	1.7	0.0045	2.1	0.05	15	2.9	2.6	0.4			
			0.019	1.7	0.0045	2.1	0.1	15	2.3					
		C	0.574	1.7	0.0045	4.2	0.05	6	172.7	174.3	2.3			
			0.585	1.7	0.0045	4.2	0.1	3	176.0					
17	BRF1-pJG4.5	A	0.669	1.7	0.0045	8.0	0.05	4.08	155.3	166.6	16.1	220.1	62.8	36.2
			0.744	1.7	0.0045	5.9	0.1	2.667	178.0					
		B	0.565	1.7	0.0045	5.8	0.05	2	369.5	289.2	113.5			
			0.518	1.7	0.0045	5.1	0.1	1.833	208.9					
		C	0.334	1.7	0.0045	1.5	0.05	5	346.2	204.6	200.3			
			0.305	1.7	0.0045	3.1	0.1	5.833	62.9					
18	BRF1-Endoquitinasa	A	0.654	1.7	0.0045	6.4	0.05	4	194.3	168.4	36.7	128.4	80.0	46.2
			0.738	1.7	0.0045	6.2	0.1	3.167	142.4					
		B	0.435	1.7	0.0045	3.4	0.05	3.833	250.1	180.4	98.5			
			0.702	1.7	0.0045	8.0	0.1	3	110.8					
		C	0.172	1.7	0.0045	5.9	0.05	6.167	35.6	36.3	1.0			
			0.33	1.7	0.0045	5.8	0.1	5.833	37.0					
19	BRF1-X	A	0.383	1.7	0.0045	2.5	0.05	4	286.6	296.3	13.8	252.7	61.7	43.6
			0.409	1.7	0.0045	2.5	0.1	2	306.0					
		B	0.402	1.7	0.0045	2.4	0.05	6.167	208.6	209.1	0.7			
			0.393	1.7	0.0045	2.4	0.1	3	209.6					
20	pEG202-BRF1	A	0.12	1.7	0.0045	4.0	0.025	28	16.2	16.8	0.9	22.7	6.7	3.9
			0.201	1.7	0.0045	4.0	0.05	21.667	17.5					
		B	0.096	1.7	0.0045	3.7	0.025	20	19.9	21.3	2.0			
			0.22	1.7	0.0045	3.7	0.05	20	22.8					
		C	0.193	1.7	0.0045	4.8	0.025	20	30.5	30.0	0.8			
			0.372	1.7	0.0045	4.8	0.05	20	29.4					
21	pEG202-BDP1	A	0.073	1.7	0.0045	2.7	0.025	32	12.5	12.8	0.3	10.3	4.8	2.8
			0.151	1.7	0.0045	2.7	0.05	32	13.0					
		B	0.019	1.7	0.0045	3.0	0.025	20	4.7	4.7	0.0			
			0.038	1.7	0.0045	3.0	0.05	20	4.7					
		C	0.077	1.7	0.0045	4.5	0.025	20	13.1	13.3	0.4			
			0.16	1.7	0.0045	4.5	0.05	20	13.6					
22	Endo-BRF1	A	0.013	1.7	0.0045	3.7	0.025	20	2.7	2.3	0.5	2.9	0.8	0.5
			0.019	1.7	0.0045	3.7	0.05	20	2.0					
		C	0.017	1.7	0.0045	4.1	0.025	20	3.1	3.4	0.4			
			0.04	1.7	0.0045	4.1	0.05	20	3.7					
23	Endo-BDP1	A	0.009	1.7	0.0045	3.6	0.025	31	1.2	1.1	0.2	2.4	1.9	1.3
			0.014	1.7	0.0045	3.6	0.05	31	0.9					
		C	0.02	1.7	0.0045	4.2	0.025	20	3.6	3.8	0.2			
			0.043	1.7	0.0045	4.2	0.05	20	3.9					
24	X-BRF1	A	0.028	1.7	0.0045	5.1	0.025	35	2.4	2.3	0.1	4.1	2.7	1.2
			0.052	1.7	0.0045	5.1	0.05	35	2.2					
		B	0.018	1.7	0.0045	5.0	0.025	30	1.8	1.9	0.1			
			0.038	1.7	0.0045	5.0	0.05	30	1.9					
		C	0.129	1.7	0.0045	7.8	0.025	30	8.3	8.6	0.4			
			0.277	1.7	0.0045	7.8	0.05	30	9.0					
		A	0.015	1.7	0.0045	2.6	0.05	15	2.9	3.8	1.3			
			0.049	1.7	0.0045	2.6	0.1	15	4.8					
		B	0.016	1.7	0.0045	2.1	0.05	15	3.9	3.9	0.0			
			0.032	1.7	0.0045	2.1	0.1	15	3.9					
25	X-BDP1	A	0.015	1.7	0.0045	4.0	0.025	35	1.6	1.7	0.1	2.8	0.9	0.4
			0.032	1.7	0.0045	4.0	0.05	35	1.7					
		B	0.024	1.7	0.0045	4.2	0.025	30	2.9	3.6	1.1			
			0.073	1.7	0.0045	4.2	0.05	30	4.4					
		C	0.023	1.7	0.0045	4.3	0.025	30	2.7	3.0	0.5			
			0.059	1.7	0.0045	4.3	0.05	30	3.4					
		A	0.007	1.7	0.0045	2.6	0.05	15	1.4	2.0	1.0			
			0.028	1.7	0.0045	2.6	0.1	15	2.7					
		B	0.012	1.7	0.0045	1.8	0.05	15	3.4	3.7	0.4			
			0.028	1.7	0.0045	1.8	0.1	15	4.0					

ANEXO 3: Artículos derivados del presente trabajo

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
© Springer-Verlag GmbH Germany, part of Springer Nature 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 T*vtbp* 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 presporulation 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

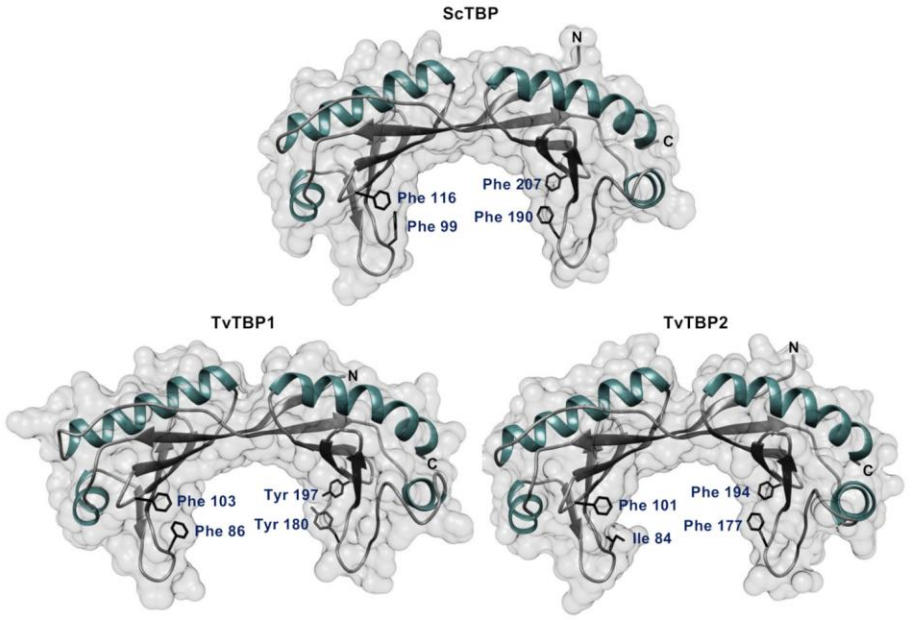


Fig. 1 Tertiary structure models for *T. vaginalis* TBPs. Theoretical structures of TvTBP1 and TvTBP2 adopt the typical saddle-shaped structure shown by a canonical TBP from *Saccharomyces cerevisiae* (ScTBP). Alpha helices are shown in green, and beta strands are shown in dark grey. Important amino 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.

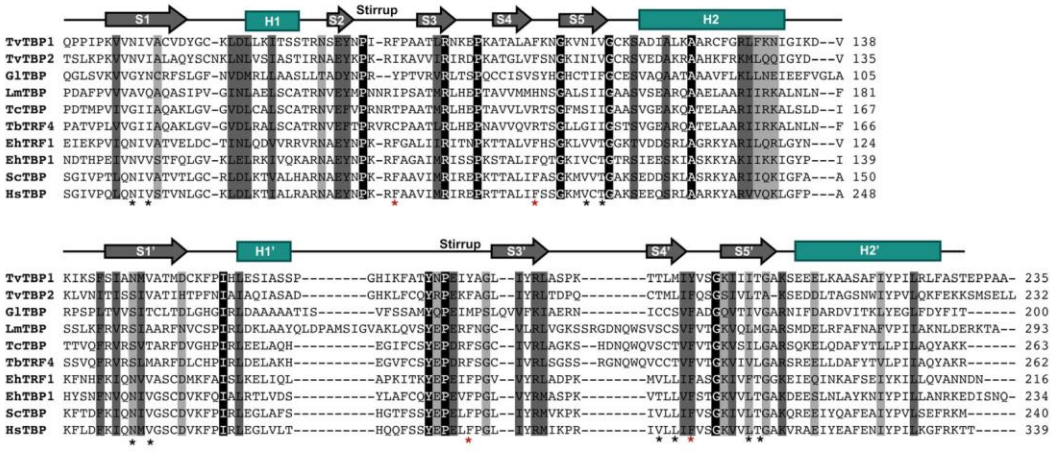


Fig. 2 Alignment of TBP amino acid sequences. Carboxi-terminal sequences of TBPs from *T. vaginalis* (TvTBP1 (A2DQT7) and TvTBP2 (A2G4N7)), *Giardia lamblia* (GlTBP (Q86GG9)), *Leishmania major* (LmTBP (Q4QD71)), *Trypanosoma cruzi* (TcTBP (Q4DK05)), *Trypanosoma brucei* (TbTRF4 (Q387R9)), *Entamoeba 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

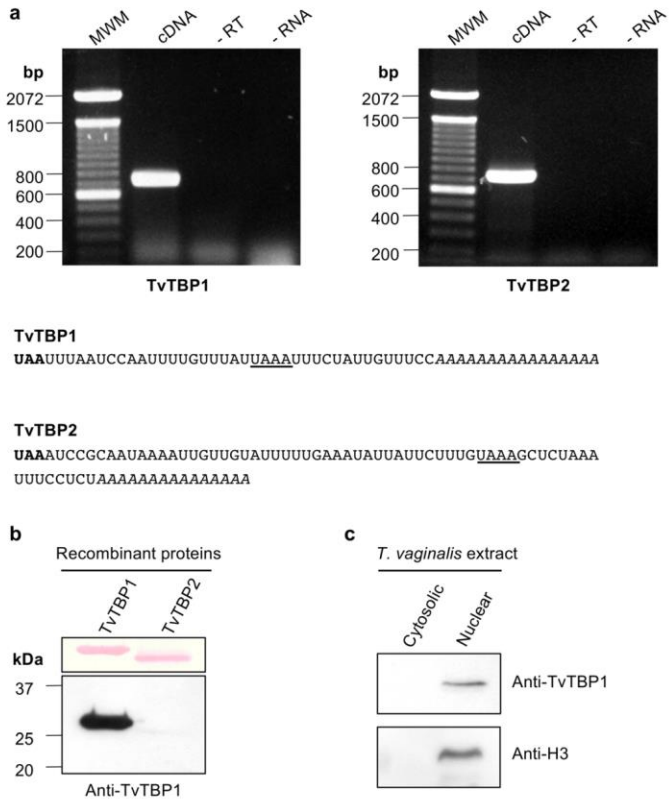


Fig. 3 **a** Both Tv*tbp* genes are expressed in *T. vaginalis*. Agarose gels stained with ethidium bromide show RT-PCRs amplified with specific primers for Tv*tbp1* and Tv*tbp2*; the 3′-UTR cDNA sequences are shown underneath. MWM, molecular weight markers; cDNA, the complete reaction; −RT, reaction 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

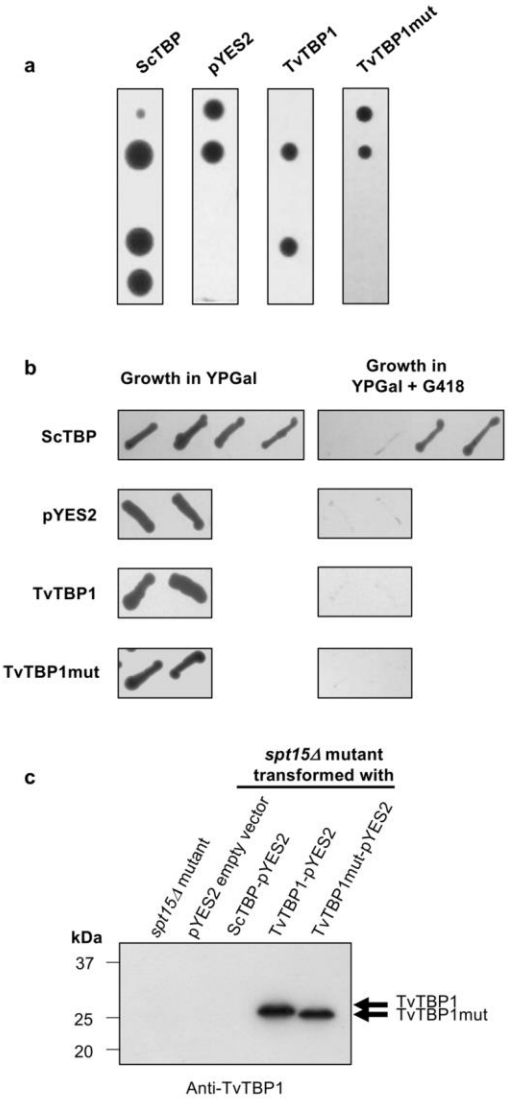
EMSAs were performed using purified recombinant 6xHis-TvTBP1, 6xHis-TvTBP2 and 6xHis-TvCRK5 (as negative control) proteins. The DNA probes were obtained by PCR amplification using specific primers and *T. vaginalis* genomic DNA as the template (ESM2 Table). The probes were purified by Sephadex G50 and 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*

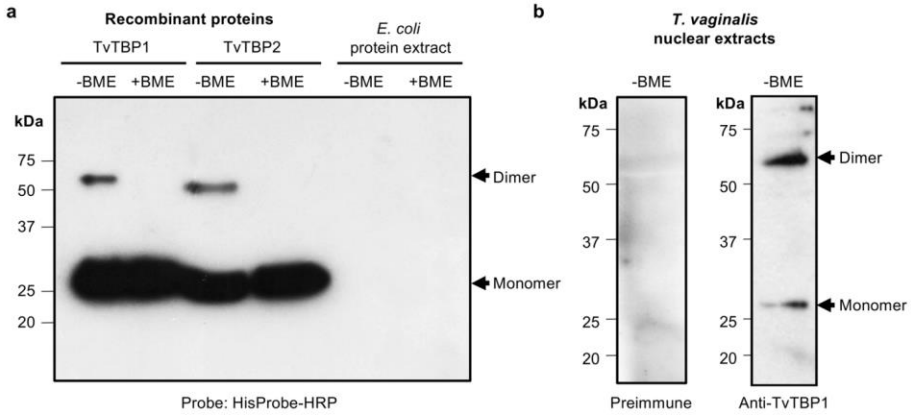


Fig. 5 TvTBPs form dimers both in vitro and in vivo. **a** The HisProbe-HRP recognises both TvTBP1 and TvTBP2 6xHis-tagged recombinant proteins. Two signals were present, and they could represent the monomer and the dimer of the recombinant TvTBPs, with an approximately weight of 27 kDa and 54 kDa, respectively. −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.

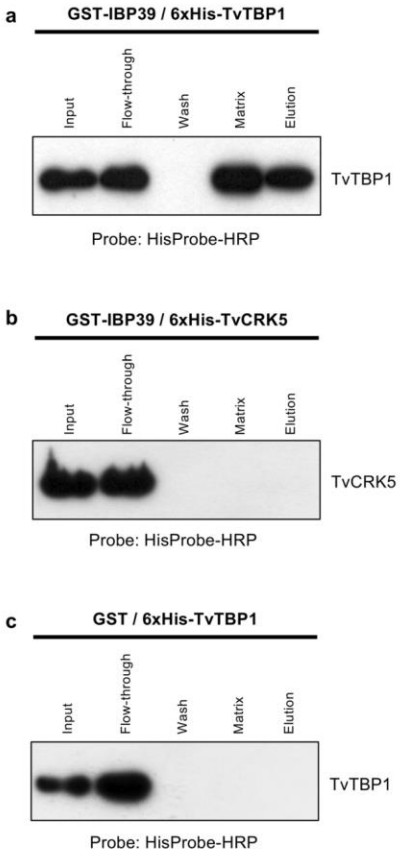


Fig. 6 TvTBP1 interacts with IBP39. Western blot detection with HisProbe-HRP of pull-down assays of GST-IBP39 with 6xHis-TvTBP1 (panel **a**), 6xHis-TvCRK5 (panel **b**) and GST with 6xHis-TvTBP1 (panel **c**). Input: Cell lysates from *E. coli* expressing 6xHis-TvTBP1 (panels **a** and **c**) or 6xHis-TvCRK5 (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

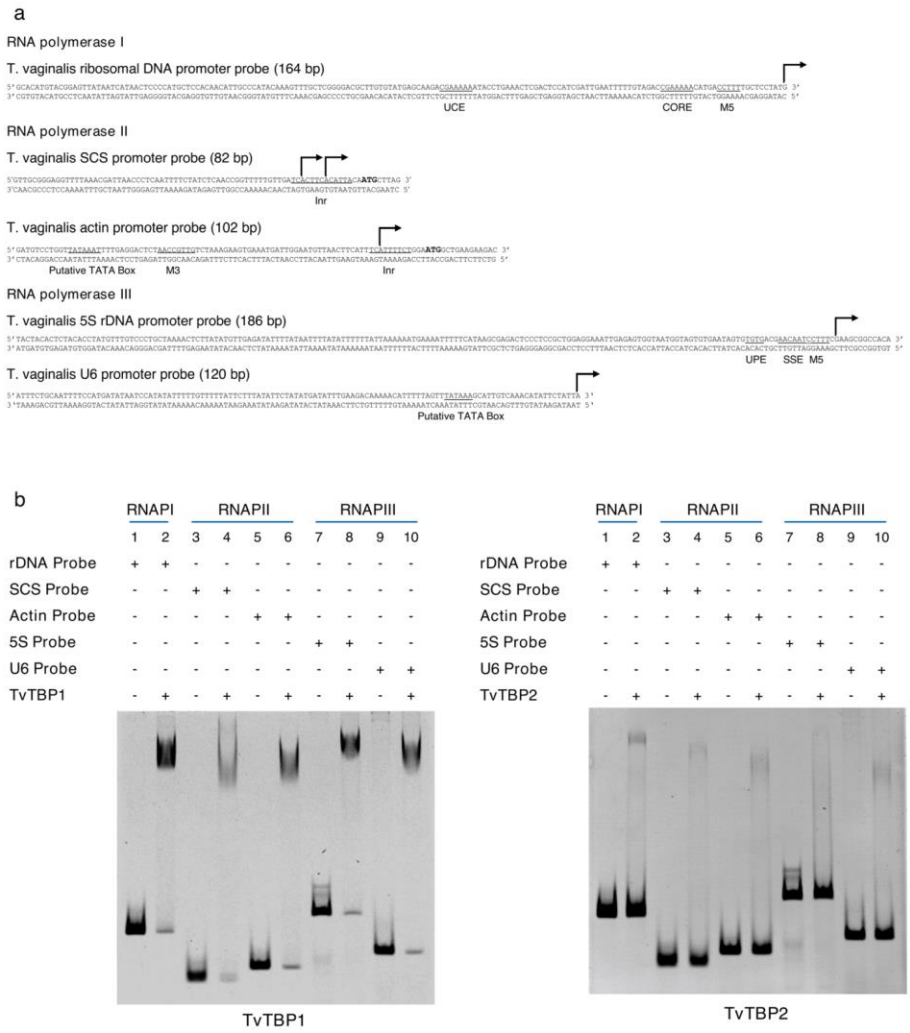


Fig. 7 TvTBPs bind different *T. vaginalis* promoter sequences. **a** Sequences of the tested promoter regions from *T. vaginalis*: (1) the ribosomal RNA gene promoter (rDNA), transcribed by RNA pol I. (2) Promoter region from the β subunit of succinyl CoA synthetase (β-SCS) and (3) actin promoter region 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

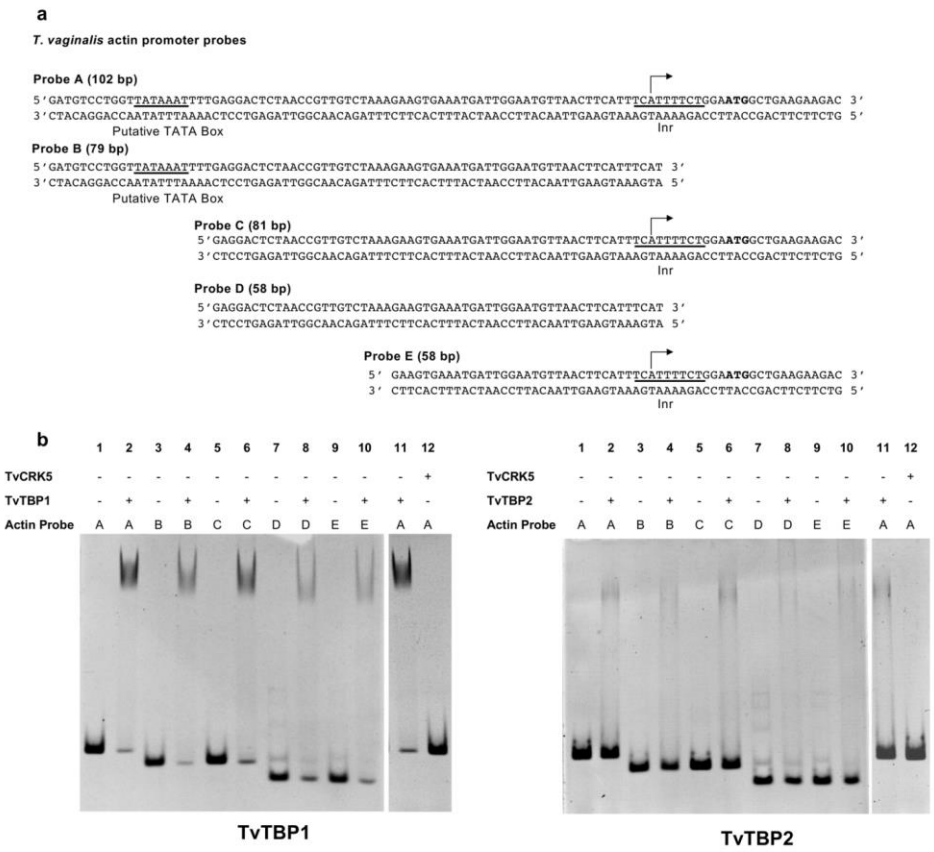


Fig. 8 TvTBPs bind DNA in a non-specific manner. **a** Different sections of the *T. vaginalis* actin gene promoter used as probes in EMSAs with the TvTBPs. **b** 500 ng of TvTBP1 or 400 ng of TvTBP2 recombinant protein was incubated with the indicated probes, as described in the 'Materials and methods' 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

Molecular & Biochemical Parasitology 239 (2020) 111312

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

journal homepage: www.elsevier.com/locate/molbiopara

Review

The highly diverse TATA box-binding proteins among protists: A review

Olivia Parra-Marín, Karla López-Pacheco, Roberto Hernández, Imelda López-Villaseñor*

Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Ciudad de México, Mexico

ARTICLE INFO

Keywords:
Protists
TBP
Protozoa
TATA-box

ABSTRACT

Transcription is the first step of gene expression regulation and is a fundamental mechanism for establishing the viability and development of a cell. The TATA box-binding protein (TBP) interaction with a TATA box in a promoter is one of the best studied mechanisms in transcription initiation. TBP is a transcription factor that is highly conserved from archaea to humans and is essential for the transcription initiated by each of the three RNA polymerases. In addition, the discovery of TBP-related factor 1 (TRF1) and other factors related to TBP shed light on the variability among transcription initiation complexes, thus demonstrating that the compositions of these complexes are, in fact, more complicated than originally believed. Despite these facts, the majority of studies on transcription have been performed on animal, plant and fungal cells, which serve as canonical models, and information regarding protist cells is relatively scarce. The aim of this work is to review the diversity of the TBPs that have been documented in protists and describe some of the specific features that differentiate them from their counterparts in higher eukaryotes.

1. Introduction

Transcription is a process by which the information contained in a gene is copied to RNA molecules called transcripts. This process represents the earliest event in gene expression. It is widely accepted that transcription is highly regulated, and the process has been divided into three stages for study: initiation, elongation, and termination. In eukaryotes, transcription is carried out by at least three RNA polymerases that have a common core [1] but synthesize specific sets of genes. Briefly, RNA polymerase I (RNA pol I) synthesizes the ribosomal cistron; RNA polymerase II (RNA pol II) transcribes mRNAs and some noncoding genes, and RNA polymerase III (RNA pol III) transcribes small RNAs such as 5S rRNA, tRNAs, and other small nuclear RNAs. RNA polymerases do not have the ability to recognize the transcription start site and thus a coordinated interaction among DNA and other proteins is required; i.e., transcription factors are needed [2]. A transcription factor and an RNA polymerase form a scaffold on a gene promoter called the preinitiation complex (PIC). TFIID is a multi-subunit transcription factor that plays a key role in the nucleation of RNA pol II promoters because its TBP subunit binds to DNA. TBP is highly conserved from archaea to eukaryotes and was first identified as a necessary component of TFIID in transcription initiated by RNA pol II. Moreover, TBP is an essential component of the machinery of all three yeast RNA polymerases [3]. TBP may interact with DNA through a sequence called the TATA box, an 8 bp element rich in A/T. The TATA box element and the TBP recognition of it probably constitute one of the most studied mechanisms in transcription. It is now known that the majority of genes lack a TATA box in the core promoter [4]; therefore, the TATA box is not a ubiquitous element in RNA pol II promoters. Moreover, in recent studies, new members of the TBP family, known as TBP-related factors (TRFs), have been identified in metazoans [5] in which they are involved in transcription. Altogether, these studies suggest that the mechanisms of transcription initiation in eukaryotes are more diverse than reflected by canonical textbook models [6,7].

Protists are a diverse group of unicellular eukaryotic organisms, many of which diverged early in evolution [8]. Information regarding the transcription mechanisms of protozoans is relatively scarce. This work analyses some characteristics of TBPs in some free-living and parasitic organisms.

1.1. The TBP structure

TBP was initially isolated from a human cell line as a component of TFIID [9], and by 1992, the first structure of TBP was determined in

Abbreviations: TBP, TATA box-binding protein; TRF, TBP-related factor; PIC, preinitiation complex; EMSA, electrophoretic mobility shift assay; ChIP, chromatin immunoprecipitation assay

* Corresponding author.
E-mail address: imelda@biomedicas.unam.mx (I. López-Villaseñor).

https://doi.org/10.1016/j.molbiopara.2020.111312
Received 20 February 2020; Received in revised form 28 June 2020; Accepted 22 July 2020
Available online 06 August 2020
0166-6851/ © 2020 Elsevier B.V. All rights reserved.

Arabidopsis thaliana by Nikolov and co-workers. Since its discovery, crystallographic structures of TBPs in different species have been obtained, which has allowed to conclude that the structure of TBP has been highly conserved throughout evolution. The structure of TBP has been widely analysed in a variety of studies: it has been found that, with minor variations, each TBP is composed of two major domains: 1) an N-terminal domain with divergent sequences and lengths across species indicating little or no conservation [10,11] and 2) a C-terminal core domain, which consists of two repeats that cause the typical folding of a saddle-shaped structure, with each repeat consisting of a five-strand β-sheet and two α-helices [10].

As stated above, the TBP N-terminal domain is different among species, and its function is poorly understood. It has been suggested that this domain may be necessary to interact with other regulatory transcription factors [12,13]. Some studies suggest that the N-terminal domain may negatively regulate TBP interactions with the TATA box sequence, with other transcription factors, and with itself [11]. In yeast, deletion of the N-terminal region has no effect on cell viability [11]. In protists, an extremely diverse group of organisms, the N-terminal portion of TBP is not conserved, a finding consistent with their broad classification [14].

The TBP C-terminal core domain comprises approximately 180 residues that are highly conserved among eukaryotic species. This region contains two structural repeats that fold into a saddle-like structure that is essential for interactions with the TATA box [15]. This structural conformation generates both convex and concave surfaces in the protein, which endows the protein with two capabilities: the convex surface causes DNA kinking as the protein binds to the minor groove of the TATA box element [10], and the concave surface exposes residues that interact with transcriptional factors such as TFIIA and TFIIB [16,17]. The TBPs in archaea are very similar to the TBPs in eukaryotes, although their sequence identity is less than 40 % [18].

1.2. The TATA box and the amino acid residues involved in DNA binding

The sequence elements in the core promoter serve as a platform for the assembly of transcription preinitiation complexes. The TATA box was the first element described in the core promoter. The motif consists of an A/T-rich short sequence (TATAWAWN) flanked by G/C-rich regions localized approximately -30 to -25 bp upstream from the transcription start site [19]). The TATA box motif is highly conserved. Initially, it was thought to be an essential sequence element, but surprisingly, this motif is present in only approximately 35 % of human genes and in 20–46 % of yeast promoters [20]. TBP binds to the minor groove of DNA in the TATA box sequence, bending it approximately 90°. The crystal structures of yeast [21], human [10] and *Arabidopsis* [22] TBPs in complex with DNA have allowed the identification of the amino acids involved in binding to DNA. The important residues for this interaction are in the concave TBP surface, including two pairs of phenylalanine residues (Phe 193/Phe 284 and Phe 210/Phe 301) that interact with the first and last base of the TATA box element, producing a DNA kink stabilized through van der Waals forces [10,23].

1.3. TBPs and their participation in all three transcription complexes

PIC formation is a step-limiting process of the start of transcription, and a highly diverse set of transcription factors participate in this process. Although TBP can recognize DNA by itself, other factors that position and stabilize it are required for DNA binding. It is then not surprising that TBP is present in several complexes and that it can interact with a wide diversity of transcription factors. TBP participates in the transcription of all three types of RNA polymerases and is present in at least three preinitiation complexes: selective factor 1 (SL1), TFIID and TFIIIB for RNA pol I, RNA pol II and RNA pol III, respectively.

TATA boxes have not been identified in the promoters recognized by RNA polymerase I. RNA polymerase I promoters require two transcription factors: SL1 and upstream binding factor (UBF). SL1 is formed by TBP and three other factors known as TBP-associated factors (TAFs), which interact with the rDNA promoter and direct transcription by RNA pol I [24].

TFIID is a multi-subunit complex composed of TBP and approximately 13 TAFs and is required for transcription by RNA pol II. TFIID serves as a scaffold for the other transcription factors that are part of the preinitiation complex transcribed by RNA pol II [19]. Some components of TFIID directly interact with the core promoter known as the initiator element (Inr) and other downstream core elements. In the transcription process initiated by RNA pol II, other factors are subsequently recruited upon the binding of TBP to DNA: TFIIA, TFIIB, TFIIE, TFIIF, TFIIG and TFIIH. TFIIA and TFIIB can directly contact TBP, and their main role is to recognize and stabilize the preformed DNA-TBP binding complex at the promoter.

Three classes of promoters are transcribed by RNA pol III, and the TFIIIB complex is necessary in all cases. TFIIIB is a trimeric complex formed by TBP, B-related factor 1 (BRF1) and B double prime 1 (BDP1). TBP and BRF1 form a very stable complex, while BDP1 is less tightly associated. In yeast, there are no TATA boxes in most of the promoters transcribed by RNA pol III, excluding a small number of gene promoters (e.g., U6 snRNA) [25].

1.4. TBP-related factors (TRFs)

Historically, TBP was considered a universal transcription factor [3]. However, in recent years, it was discovered that a group of proteins similar to TBPs can interact with the three complexes formed by RNA pol I, II and III [19] and that, in contrast to TBP, they have a specific role in the regulation of transcription: they are gene-specific or tissue-specific and have distinct functions *in vivo* [26,27]. It has been proposed that TRFs recognize noncanonical sequences since the consensus binding sequences for TRFs are still unknown in most cases [28].

TRFs can be classified into four groups: a) TRF1, b) TRF2 (TLF, TBPL1, TLP, and TRP), c) TRF3 (TBP2 and TBPL2) and d) TRF4, which is briefly described herein. Additionally, in *Drosophila*, TRF5 has been identified but has an unknown function [29].

1.4.1. TRF1

The first TBP-related factor to be discovered, now called TRFI, was identified in *Drosophila melanogaster* by a mutation that caused a leg-shaking phenotype when the fly was anaesthetized [30]. This protein is insect-specific and is involved in the regulation of specific gene subsets in the central nervous system, in male germ cells and in polytene chromosomes [31]. TRF1 has 63 % similarity with canonical TBPs [15] and can substitute for TBP in *in vitro* transcription assays [31]. Sequence comparisons of *Drosophila* TBP and TRF1 showed that the residues in the DNA-binding surfaces are conserved, explaining the ability of TRF1 to bind to a TATA box [32]. Additionally, TRF1 can bind to DNA and may associate with TFIIA and TFIIB [31]. In the *Drosophila* gene promoters transcribed by RNA pol III, TRF1 can replace TBP in the TFIIIB complex [33]. Immuno-depletion of TRF1 in a nuclear extract eliminated the transcription of several RNA pol III transcripts (tRNAs and 5S) [34]. However, TBP, but not TRF1, is used for U6 transcription by RNA pol III [35]. Additionally, BRF1 co-immunoprecipitates with TRF1 but not with TBP, supporting the hypothesis that BRF1 and TRF1 associate. Finally, BRF1 and TRF1 colocalized on polytene chromosomes at sites that correspond to clusters of tRNA and 5S rRNA genes [34]. These results indicate that TRF1 is a specialized form of TBP that participates in the transcription of promoters transcribed by RNA pol III. For example, in *D. melanogaster,* TRF1 (which is insect-specific) can replace TBP in the transcription process by binding to the TATA box together with TFIIA and TFIIB [31] and associating with BRF1 [34].

1.4.2. TRF2

TRF2 is also known as either TBP-like protein (TLP), TBP-like factor

(TLF), TBP-related protein (TRP), or TBP-like 1 (TBPL1). It has been found to be widely distributed in metazoans from nematodes to vertebrates [36], including *Homo sapiens* but is apparently absent in plants and fungi [25]. TRF2 is the most distant paralogue of TBP because it is the protein with the least similarity to it (only 40 % identity and 60 % similarity). In contrast to the ability of TBP, TRF1 and TRF3 to bind DNA, the TRF2 factor does not have DNA-binding activity. In fact, multiple amino acids involved in DNA binding are not conserved in TRF2, including three of the four Phe residues required for TATA box recognition [36–38]. However, there are very few differences in the amino acids involved in the interaction with TFIIA and TFIIB; therefore, it is not surprising that TRF2 binds both TFIIA and TFIIB *in vitro* [39–42]. It has been proposed that TRF2, having evolved from a duplicated TBP gene and acquiring mutations in the three Phe residues over time, retained other key properties of TBP, such as the ability to interact with TFIIA and TFIIB [38]. TRF2 is involved in RNA pol II transcription. These similarities and differences of TBPs and TRF2s suggest that TRF2 might nucleate the assembly of the PIC on a different class of genes with a promoter element distinct from the TATA box [43]. Consistent with this hypothesis, in *Drosophila,* TRF2 is mainly associated with TATA-less promoters transcribed by RNA pol II [15,33].

1.4.3. TRF3

TRF3, also known as TATA-binding protein 2 (TBP2) or TBP-like protein 2 (TBPL2), is present in vertebrates but not in lower metazoans and is the most closely related protein to TBP (95 %). Its C-terminal domain is highly conserved, including all important amino acids involved in binding to DNA, TFIIA and TFIIB [44]. Therefore, it is not a surprise that TRF3 binds to DNA and forms complexes with TFIIA and TFIIB factors [45]. On the other hand, the N-terminus is different among TRF3 orthologues and TBP [45]. The available data suggest that TRF3 is a result of a vertebrate-specific gene duplication event [44,45]. TRF3 is predominantly localized in the nucleus; however, during cytokinesis, it is localized in the cytoplasm, whereas TBP is not translocated from the nucleus [44]. TRF3 is widely expressed in various cell lines and displays tissue-specific expression [46]; nevertheless, it is required in the embryonic development of zebrafish and *Xenopus* [45,47]. A deletion/mutation of TRF3 leads to developmental defects, and accordingly, TRF3 and TAF3 are required for the initiation of haematopoiesis in mice [48,49].

1.4.4. TRF4

Recently, a new TBP-related protein, TRF4, was described in *D. melanogaster,* and it is conserved among *Drosophila* species. TRF4 is a protein of 331 amino acids. Interestingly, TRF4 is localized in the cytoplasm and is concentrated in the endoplasmic reticulum (ER), where it interacts with TER94, a component of the ER protein degradation machinery. TRF4 is an example of protein neofunctionalization caused by the loss of its nuclear localization [29].

1.5. TBPs in protists

Protists are generally considered any eukaryotic organism that is not an animal, plant or fungus. The protist kingdom is a paraphyletic group that comprises three major groups: algae, flagellate fungi and protozoa [50]. Since protozoa cover a broad range of microorganisms, some authors have considered them a phyla within the animal kingdom; however, protozoa are now considered to be within the protist kingdom as microorganisms with basic eukaryotic features [51,52]. A large majority of protozoa are early-branching eukaryotes, and their biology is distinct in many ways from most model eukaryotes [8]. There are more than 50,000 species currently classified into six phyla based on their morphology. The parasites best known for causing disease belong to two phyla: Sarcomastigophora (*Acanthamoeba*, *Leishmania*, *Trypanosoma*, *Giardia*, *Trichomonas* and *Entamoeba*) and Apicomplexa (*Plasmodium* and *Cryptosporidium*) [53]. Protozoa reside in every habitat. Some protozoa have complex life cycles requiring two different host species (e.g., trypanosomatids), while others have only one host (e.g., *Trichomonas vaginalis*).

Early phylogenetic analyses demonstrated that TBPs are better conserved among higher eukaryotes than they are in lower eukaryotes, including protists [54,55]. In general, these studies were presumed to be representative of all eukaryotes, but little is known about TBPs in protists. Therefore, we here aim to describe the available information regarding TBPs in representative protists, focusing on their canonical properties such as the conservation of amino acids in the concave surface and their ability to bind DNA.

1.5.1. Acanthamoeba castellanii

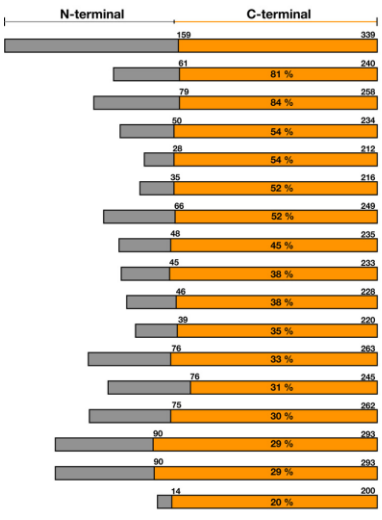
Acanthamoeba castellanii is a small free-living amoeba commonly found in soil and freshwater. This microorganism has a two-phase life cycle: cyst and trophozoite [56]. *A. castellanii* has emerged as a model organism because of the different processes in which it is involved, such as its ability to capture preys by phagocytosis and to produce serious infections in humans [57]. For instance, it is one of the aetiologic agents of keratitis, a severe infection that affects the cornea and usually causes eye pain accompanied by a characteristic symptomatology [58].

A. castellanii genome comprises approximately 42 Mb and encodes a wide variety of transcription factors, including general transcription factors such as AcTBP and specialized factors such as TBP promoter binding factor (TPBF), a tetrameric protein with a recently discovered DNA-binding motif that is regulated by phosphorylation [59]. *A. castellanii* encodes a single TBP, and TRFs have not been described for this organism. AcTBP was characterized in 1992 [60]. Its deduced sequence is 258 amino acids long, with 180 residues corresponding to the C-terminus that are highly conserved across organisms (with 84 % identity to HsTBP) (Fig. 1). Using *in vitro* transcription assays, Wong and collaborators found that the AcTBP recombinant protein (expressed in *E. coli)* can bind to a TATAAAA sequence and can replace HsTBP for transcription initiation in the adenovirus major late promoter [60]. In addition, it was demonstrated that AcTBP binds with higher affinity for TATATAA sequences than for other variants with enriched A/T sequences [61]. AcTBP is required by the three RNA polymerase systems and is a component of the polymerase I-specific transcription initiation factor (TIF-IB), also known as SL1 in humans [62]. Studies by Radebaugh and co-workers suggested that AcTBP is assembled in distinct ways in the promoters transcribed by the three RNA polymerases, since different oligonucleotides with a TATA box inhibited the *in vitro* transcription by the three RNA polymerases [62]. The transcription of the TBP gene is inhibited with anti-TBP antibodies [63]. The level of AcTBP gene expression is regulated by TPBF, which binds to TPE (TBP promoter element) in the TBP promoter [59].

1.5.2. Entamoeba histolytica

Entamoeba histolytica is a parasitic protozoan responsible for human amoebiasis worldwide. *E. histolytica* is an amitochondriate parasite; therefore, it is considered an archezoan protist [64,65], and its position in the phylogenetic tree is unclear due to the absence of mitochondria [66]. *E. histolytica* is classified as a member of the Amoebozoa super group, and its life cycle consists of two stages: cyst with four nuclei (infective stage) and trophozoite with one nucleus (proliferative stage).

The genome of this parasite is 25 Mb and its A/T content is 75.9 % [67]; therefore, it is not unusual to find sequences with similarity to a TATA box in *E. histolytica* promoters. The core promoters of protein-coding genes in *E. histolytica* have been characterized and are composed of some of the following elements: 1) the TCCAAAA element (similar to the CCAAT box in higher eukaryotes) located at -100 bp relative to the transcription start site, identified in the lectin gene promoter (*hgl2, gEh170*) by quantification of CAT activity [68]; 2) the GAAC element (GAACT), identified in the *hgl5* gene promoter [66,69], which controls the rate and site of transcription initiation independent of the TATA element [70]; 3) the start site AAAAATTCA and AAAAATCA sequences



Organism	Name	DNA binding	Sequence	References
Homo sapiens	HsTBP	YES	TATA Box	[10]
Saccharomyces cerevisiae	ScTBP	YES	TATA Box	[21]
Acanthamoeba castellanii	AcTBP	YES	TATATAA	[61]
Entamoeba histolytica	EhTBP1	YES	GAAC box, TATTTAAAA and variants	[73, 75, 76]
Entamoeba histolytica	EhTBP2	ND	-	-
Entamoeba histolytica	EhTRF1	YES	GAAC box, TATTTAAAA and variants	[73, 76]
Cryptosporidium parvum	CpTBP	YES	GCAGAGCATATAAGGTGAGGTAGGA	[89]
Trichomonas vaginalis	TvTBP1	YES	rDNA, actin, 5S and U6 promoters	[86]
Trichomonas vaginalis	TvTBP2	YES	rDNA, actin, 5S and U6 promoters	[86]
Plasmodium falciparum	PfTBP	YES	TGTAA	[96]
Crypthecodinium cohnii	CcTBP	YES	TTTTTTTT	[118]
Trypanosoma cruzi	TcTBP-NE	YES	G-rich sequence	[112]
Trypanosoma cruzi	TcTBP-E	ND	-	-
Trypanosoma brucei	TbTRF4	YES	SL RNA gene promoter (ChIP)	[114, 115]
Leishmania major	LmTBP	YES	SL RNA, tRNA, snRNA, 5S rRNA gene promoters (ChIP)	[105]
Leishmania tarentolae	LtTBP	YES	SL RNA gene promoter, tRNA and snRNA	[108]
Giardia lamblia	GlTBP	ND	-	-

Fig. 1. Sequence identity among the TBPs described in this paper. The C-terminal amino acid sequences (represented in scale as orange bars) of *Homo sapiens* (P20226), *Saccharomyces cerevisiae* (P13393), *Acanthamoeba castellanii* (P26354), *Cryptosporidium parvum* (Q6SEL4), *Entamoeba histolytica* (EhTBP1: P52653, EhTBP2: C4M7H7), *Trichomonas vaginalis* (TvTBP1: A2DQT7, TvTBP2: A2G4N7), *Plasmodium falciparum* (P32086), *Trypanosoma cruzi* (TcTBP_E, Esmeraldo haplotype: Q4D013, TcTBP_NE, Non-Esmeraldo haplotype: Q4DK05), *Trypanosoma brucei* (Q387R9), *Leishmania major* (Q4QD71), *Leishmania tarentolae* (Q2TCV3), *Crypthecodinium cohnii* (Q945B3) and *Giardia lamblia* (Q86GG9) were aligned to obtain identity values compared to the *H. sapiens* sequence, using the Sequence Identity And Similarity tool from the Immunomedicine Group, Universidad Complutense, Madrid (http://imed.med.ucm.es/Tools/sias.html). N-terminal sequences (shown in scale as grey bars) were not used for the alignment. The "Name" column corresponds to the TBP nomenclature that was used in this work. TBPs shown to bind DNA with the corresponding reported sequences are shown. ND represents not determined.

[71]; and 4) elements with similarity to TATA boxes such as TATTT AAA, TATATTTAAA and TATAAA; these elements have been localized close to the start site, approximately at -30 bp. [71].

Among the transcription factors characterized in *E. histolytica*, the only canonical member of the basal transcription machinery is an orthologue of TBP (no genes have been identified for TFIIA or TFIIB). Interestingly, in the *E. histolytica* genome, three loci encode three different proteins with similarity to TBP (EhTBP1, EhTBP2 and EhTRF1). Eh*tbp1* and Eh*trf1* are transcriptionally active. On the other hand, EhTBP2, with 212 amino acids, is a shorter version of EhTBP1, and it is expressed at a very low basal level and endogenously silenced by DNA methylation of the coding region [72].

The deduced sequences for EhTBP1 and EhTBP2 are identical in residues 23–234, and the only difference between these two sequences is in the extension of the N-terminal domain [73]. EhTBP1 was the first protein described with 234 amino acids and 26 kDa. This protein shares 54 % identity with HsTBP and is located in the nucleus and in the EhkO, a kinetoplast-like organelle [74]. EhTBP1 recognizes the canonical TATTTAAA-box, as well as different TATA box variants. Interestingly, in contrast to canonical TBPs, EhTBP1 has a higher binding affinity for TATA box variants, probably because of changes in the amino acids involved in TBP-DNA binding [75].

EhTRF1 is 216 amino acids long and shares 52 % identity with HsTBP (Fig. 1). EhTRF1 also binds the TATTTAAA-box and other variant sequences. The dissociation constant was calculated for EhTBP1 and EhTRF1, and the binding affinity for the TATTTAAA box was found to be higher for EhTBP1 than for EhTRF1 [73]. In a recent study, it was observed that EhTBP1 and EhTRF1 bound to GAAC boxes [76], demonstrating the variety of DNA sequences to which these TBPs can bind.

1.5.3. Trichomonas vaginalis

T. vaginalis is an early divergent protozoan and a member of the parabasalid subgroup of flagellated protists within the Excavata supergroup. *T. vaginalis* has a simple life cycle with only a trophozoite stage, and it is the causal agent of trichomoniasis, a sexually transmitted disease in humans [77]. No cyst stage has been found, although pseudocysts have been observed under stress conditions [78].

T. vaginalis has an unusually large genome for a protist (approximately 160 Mb), which was sequenced in 2007 [79]. The transcription process is not well characterized, as is the case for most organisms described herein. The Inr element in the core promoter of *T. vaginalis* is a functional homologue of its metazoan counterpart and plays a central role in *T. vaginalis* gene expression [80,81]. Additionally, new elements have been identified in the promoters of *T. vaginalis*, such as motif 3 (M3), which is recognized by a Myb protein, and motif 5 (M5), which overlaps with the transcription start site. Both motifs are positioned near the transcription start site and can be regulated by iron [82,83].

Several components of TFIID (TAFs) in the *T. vaginalis* genome have been annotated, such as TFIIB and BRF1, but their participation in transcription remains to be determined [79].

The initiator binding protein of 39 kDa (IBP39) has been identified only in *T. vaginalis*; therefore, no orthologues have been described in other organisms [84]. IBP39 interacts with the Inr element through its amino terminal portion [85], while the C-terminal domain of this protein interacts with the C-terminal domain (CTD) of RNA pol II [82].

Other canonical elements, such as a TATA box, have not been identified in the promoters of *T. vaginalis*; however, two TBPs are expressed: TvTBP1 and TvTBP2, with 45 % and 38 % shared identity with HsTBP, respectively (Fig. 1) [86]. These two proteins can bind to *T. vaginalis* promoter sequences, as shown by electrophoretic mobility shift assay (EMSA). However, no DNA sequence specificity is apparent for TvTBP binding, although TvTBP1 interacts with IBP39, which suggests its participation in transcription initiation [86].

1.5.4. Cryptosporidium parvum

Cryptosporidium parvum is an intracellular parasite belonging to the Apicomplexa phylum. *C. parvum* is considered to be an opportunistic pathogen in humans that causes acute diarrhoea in immunocompromised patients. The life cycle of *C. parvum* is complex, with both asexual and sexual stages and several developmental stages [87]. The genome of *C. parvum* sequenced in 2004 is 9.1 Mb, and approximately 70 % of its sequence consists of A/T; however, canonical TATA boxes have not been found [88]. As with many other protists, little is known about the regulation of gene transcription in *C. parvum*. This parasite has only one TBP named CpTBP1 (herein referred to as CpTBP), which is part of the basal transcription machinery [89]. Searching for sequences of canonical TBPs, Millership and collaborators identified a TBP gene homologue of 750 bp that encodes a 27.8 kDa protein with 53 % shared identity with HsTBP (Fig. 1). An EMSA of CpTBP with a biotinylated single oligonucleotide containing a TATA box consensus sequence revealed a binding signal [89].

Other general transcription factors, including TFIIA and TFIIB, have been identified in the *C. parvum* genome; however, no further studies related to these proteins have been published to date. On the other hand, a transcriptional coactivator known as multiprotein bridging factor type 1 (MBF1), which enhances transcription by facilitating the interconnections of distinct regulatory DNA-binding proteins to TBP, was identified in *C. parvum* [90]. MBF1 mediates the transcriptional activation of various genes in eukaryotes by interconnecting TBP with sequence-specific regulators. The interaction between CpMBF1 and CpTBP was observed by EMSA [89]. However, the genes transcribed by CpTBP and CpMBF1 have not been identified; therefore, the core promoter sequence to which they may bind remains unknown, and further studies are indeed needed to dissect the biological roles of CpTBP in the development of this parasite.

1.5.5. Plasmodium falciparum

Malaria is a disease transmitted by female mosquitoes of the *Anopheles* genus that can inject *Plasmodium* parasites through the human skin. *Plasmodium falciparum* is one of the six protozoan parasites that cause malaria in humans and is considered to be the most severe and mortal pathogenic agent [91]. The genome of *P. falciparum* (approximately 23 Mb) is distributed across 14 chromosomes and has more than 5300 genes that are expressed in a complex life cycle that involves defined morphological stages with stage-specific gene expression in both hosts: humans and mosquitoes [92]. The A/T richness of intergenic regions has made it difficult to identify the essential promoter elements critical for the fine-tuned regulation of its gene expression [92]. Additionally, a bioinformatic search using a hidden Markov model (HMM) was used to identify different general transcription factors involved in transcription initiated by RNA pol II. Consequently, homologues of TBP, TFIIB, TFIIEα and some components of TFIIH were identified, but no other components of the TFIID complex or subunits of TFIIA were found [93]. Through two-dimensional hydrophobic cluster analysis (HCA) and the profile-based search method, PSI-BLAST, another group predicted orthologues of the TFIIA large subunit, two candidates for the TFIIA small subunit, TFIIEβ, and some orthologues of the TFIID complex (TAF1, TAF2, TAF7 and TAF10). These TAFs do not contain the classical histone fold domain, suggesting an unusual architecture of this TFIID complex [94].

PfTBP was initially characterized by John Hyde and collaborators [95], who described it as a gene with 76 % A + T DNA content without introns that encodes a protein of 228 amino acids (26.164 kDa). The C-terminal core domain of PfTBP comprises 183 residues, which are highly conserved and are similar in length to other known TBP sequences. The N-terminal domain comprises 45 amino acids and does not share sequence similarities with other TBPs. The C-terminus of PfTBP has 38 % shared identity with HsTBP (Fig. 1). Additionally, most of the residues involved in DNA binding are conserved in PfTBP [95]; this protein is expressed in the nuclei of *P. falciparum* during its blood stage [95,96]. In EMSAs and chromatin immunoprecipitation assays (ChIPs), it was observed that PfTBP recognized both a TATAA sequence located -81 bp from the *kahrp* promoter region and a TATA box-like sequence, TGTAA, -186 bp from the *gbp-130* promoter region. ChIPs validated the association of PfTBP with both TATA box elements [96]. However, the biological relevance of the TATAA sequence requires additional experimental evidence [97]. ChIPs also revealed that PfTBP and PfTFIIE are bound in promoter sequences that are expressed in blood stages: histidine-rich protein 3 (*hrp3*) and skeleton-binding protein 1 (*sbp1*) promoters with ring-specific expression, and merozoite surface protein 1 (*msp1*) and rhoptry high-molecular weight protein 3 (*rhoph3*) promoters with trophozoite-specific expression. PfTBP and PfTFIIE are bound even when the gene is transcriptionally inactive, suggesting that these two transcription factors are pre-assembled at the promoter and that additional signals are required for promoter activation [98]. Recently, an in silico analysis showed the expression of a PfTBP-like protein in the gametocyte stage [99]; however, there are no additional studies related to the function of this protein.

1.5.6. Trypanosomatids

Trypanosomatids constitute a group of unicellular flagellated parasites that include the *Leishmania* and *Trypanosoma* genera, which cause important diseases in humans and animals. Trypanosomatids belong to the kinetoplastid order, for which the main characteristic is the presence of an organelle, known as kinetoplast, that contains mitochondrial DNA. These parasites have a complex life cycle, involving several developmental forms between two hosts: an invertebrate and a vertebrate.

Trypanosomatids are early-branching protists and have an atypical mechanism of gene expression. The protein-coding genes are arranged in polycistronic units [100], and the transcribed mRNAs are processed by the addition of a spliced leader sequence (SL) at the 5′ end of each mRNA by trans-splicing and by polyadenylation at the 3′ end. The SL RNA gene promoter is the best characterized promoter in kinetoplastids and is transcribed by RNA pol II from individual promoters. The SL RNA gene is under the control of upstream promoter elements and does not contain a TATA box. The promoter recruits transcription factors such as TBP, the snRNA-activating protein complex (SNAPc), and homologues of TFIIA and TFIIB [101].

1.5.6.1. Leishmania major. *Leishmania major* is the causative agent of leishmaniasis in humans. Leishmania infection is transmitted by the bite of female sand flies [102]. There are more than 20 *Leishmania* species that cause this disease in three main clinical forms: visceral, cutaneous and mucocutaneous leishmaniasis. *L. major* is responsible for cutaneous leishmaniasis and is characterized by ulcers on the face and extremities.

The genome of *L. major,* which is 32.8 Mb, is distributed across 36 chromosomes [103]. Through in silico analyses, orthologues for TBP, BRF1, SNAP50 and the three subunits of TFIIH have been identified [103]. BDP1, a subunit of TFIIIB, was recently characterized. LmBDP1 is essential for the viability of *L. major* [104]. LmTBP has 29 % shared identity with HsTBP (Fig. 1). ChIP-chip analyses were reported to show a high level of acetylated histone H3 at the 5′ ends of all polycistronic protein-coding gene clusters, which were found mainly in dividing promastigotes. Accordingly, with these findings, it was determined that the binding of TBP and SNAP50 showed almost identical patterns associated with tRNA, snRNA and SL RNA gene promoter regions [105].

1.5.6.2. Leishmania tarentolae. *Leishmania tarentolae* is a non-pathogenic trypanosomatid protozoan that parasitizes some reptile species. The life cycle of *L. tarentolae* is based in two hosts and has two main morphological stages: flagellated promastigotes (extracellular) in the insect vector and non-motile amastigotes (intracellular) in the lizard host [106]. The genome of *L. tarentolae* is distributed across 36 chromosomes and is 30.4 Mb [106]. *L tarentolae*

has been widely used as a model organism for studies on RNA editing [107].

The SL RNA genes in *L. tarentolae* are organized into two separate head-to-tail tandem arrays, *MINA* (approximately 60 gene copies) and *MINB* (approximately 40 gene copies). The transcribed regions in both cases are identical. The promoter is bipartite and contains two regions: the -60 element, which shows similarities to the proximal sequence element (PSE), and the -30 element [19].

In *L. tarentolae*, a non-canonical form of the TBP transcription factor was described. LtTBP has 29 % shared identity with HsTBP (Fig. 1). *L. tarentolae* encodes a protein (LtTBP) of 31.39 kDa that is similar to TBP. The N-termini of LtTBP and LmTBP are conserved, while only a small sequence is conserved among these transcription factors and TBP in *Trypanosoma brucei* [108]. Additionally, *Leishmania* TBPs contain two insertions: the first one consists of 10 amino acids between the α1′-helix and β2′-sheet and the second one, consisting of 8 amino acids, is between the β3′- and β4′-sheets. EMSAs showed that LtTBP binds to the -60 element and is also implicated in snRNA and tRNA gene transcription [108].

1.5.6.3. Trypanosoma cruzi. *Trypanosoma cruzi* is the aetiological agent of Chagas disease, also named American trypanosomiasis. *T. cruzi* is a protozoan in the Excavata supergroup and is a haemoflagellate intracellular parasite. The life cycle of *T. cruzi* is complex, with different stages among vectors (triatomine insect) and mammals (including humans and other species). There are three main stages: amastigotes (the proliferative form in the mammalian host), epimastigotes (the proliferative form in the vector) and trypomastigotes (the infective form) [109].

The *T. cruzi* genome (approximately 105 Mb) is distributed between 20 and 46 chromosomes [110]. In hybrid strains, there are two alleles encoding TBP: the esmeraldo (E) and the non-esmeraldo (NE) haplotypes. The first description of TBP corresponded to the NE haplotype. TcTBP-NE shares 33 % identity with HsTBP, and TcTBP-E shares 31 % identity with HsTBP (Fig. 1). A yeast two hybrid assay showed that TcTBP has the ability to interact with the SNAP complex [111]. Additionally, TcTBP has an *in vitro* preference for G-rich sequences [112].

1.5.6.4. Trypanosoma brucei. *T. brucei* is the cause of African sleeping sickness, also known as African trypanosomiasis. It is transmitted to humans by the bite of a tsetse fly. A single TBP homologous protein is expressed in *T. brucei* and was named TbTRF4 because its characteristics do not match the description of a canonical TBP. However, TbTRF4 is unrelated to the TRF4 protein in *D. melanogaster*. The N-terminal domain of TbTRF4 is not similar to that of TBP in other trypanosomatids, while the C-terminal domain shares approximately 31 % identity with HsTBP (Fig. 1) [113]. Additionally, essential residues involved in binding to the TATA box are not conserved; nevertheless, TbTRF4 is more closely related to canonical TBP than other members of the TBP family (e.g., TRFs). The Phe residues involved in DNA kinking are replaced with cysteine and arginine residues.

ChIPs showed that TbTRF4 recognizes the procyclic acidic repetitive gene promoter, the spliced leader promoter and snRNA promoters, transcribed by RNA pol I, II and III, respectively [113]. The results from iRNA assays showed that TbTRF4 is essential for transcription by the three RNA polymerases [113]. Curiously, there are no TATA elements identified in any of the promoters tested, and the mechanism by which TbTRF4 binds to DNA is unknown [114]. A few general transcription factors have been identified in *T. brucei*, including SNAPc, which consists of three subunits that participate in the transcription of the SL-RNA promoter in trypanosomatids, and TFIIA, which interacts with TbTRF4 and is associated with the SL RNA promoter [114,115].

1.5.7. Crypthecodinium cohnii

Crypthecodinium cohnii is a small marine dinoflagellate found as a free-living microorganism and has industrial importance for producing docosahexaenoic acid (DHA). One of the most important characteristics of dinoflagellates is their lack of histones and the unusual chromatin structure caused by the chromosomes remaining condensed throughout the cell cycle [116]. There is little information about transcription in this organism; to date, a small number of genes have been studied in *C. cohnii* and other dinoflagellates, whose upstream sequences apparently lack a TATA box and other regulatory elements [117,118]. Guillebault and collaborators isolated a cDNA encoding a protein with homology to canonical TBPs, which they named cTBP/CcTLF (herein named CcTBP), because the deduced amino acid sequence lacks the important residues for DNA binding, which is similar to the TLF protein, herein named TRF2. However, the CcTBP sequence is more closely related to TBP than to TRFs (TRF1, TRF2, TRF3 or TRF4). CcTBP is 35 % identical to HsTBP (Fig. 1). Phylogenetic analyses and bootstrap calculations did not classify CcTBP within the families of TBP or TRFs; therefore, it was proposed as a member of a new family of transcription factors related to protists [118]. Structurally, CcTBP contains the typical two-repeat signature of the C-terminal domain, which forms the saddle-like structure, whereas the N-terminal domain does not share homology with other TBPs, as usually described for this region. CcTBP lacks the two pairs of Phe residues that are involved in DNA binding through intercalation in the minor groove. Some of the key residues involved in the interaction with TFIIA and TFIIB are conserved.

Using EMSAs, it was observed that CcTBP binds to both TATAAAAA and TTTTTTTT sequences *in vitro*. Interestingly, the preference for the TTTTTTTT sequence was clearly greater than it was for the TATAAAAA sequence. However, these sequences have not been identified in upstream promoter regions in this organism [116]. The binding affinity of CcTBP for the TATAAAAA sequence did not increase when the protein was mutated to introduce the characteristic Phe residues. Furthermore, the presence of TFIIA did not modify the preference of CcTBP for this sequence. In the transcriptome of other dinoflagellates, such as *Alexandrium catenella* and *Lingulodinium polyedrum,* two proteins with similarity to the CcTBP are found [116,119].

1.5.8. Giardia lamblia

Giardia lamblia is a parasitic protist of the diplomonads group. It is the most common intestinal parasite of humans living in developed countries [120] and is commonly found in freshwater lakes. Giardia has a simple life cycle that consists of two stages: trophozoite and cyst.

The promoters in *G. lamblia* are short, containing fewer than 250 bp, due to the high degree of gene density. Two elements were identified in these promoters: an A/T-rich element was identified surrounding the transcription start site, and the *Giardia* CAAT box (g-CAB) [121]. The gene promoters encoding proteins do not possess canonical TATA boxes. Nevertheless, an A/T-rich sequence was identified in the sequence surrounding the transcription start site; however, these sequences were not found in the -30 bp region, where typical TATA boxes are found [122]. The A/T-rich sequence is often referred to as an Inr-like element due its position [123] although it is very different from canonical Inr.

A GlTBP candidate was identified in the *Giardia* genome, which is very different than that of other eukaryotes, even among the protists described herein (it is 21 % identical to HsTBP) (Fig. 1) [54]. Neither TFIIA, TFIIB, TAFs nor other transcription initiation factors were identified, perhaps due to the remarkable divergence of *Giardia*. RRN3 is the only transcription factor identified that is involved in transcription initiated by RNA pol I. BRF has also been identified in the *Giardia* genome [54]; nevertheless, it is unknown whether these factors participate together with GlTBP in the formation of the PIC.

2. Final considerations

Although animals, plants and yeast are the most studied eukaryotes, these organisms do not represent the diversity in the Eukaryota kingdom. In contrast, protists are a more biologically diverse group

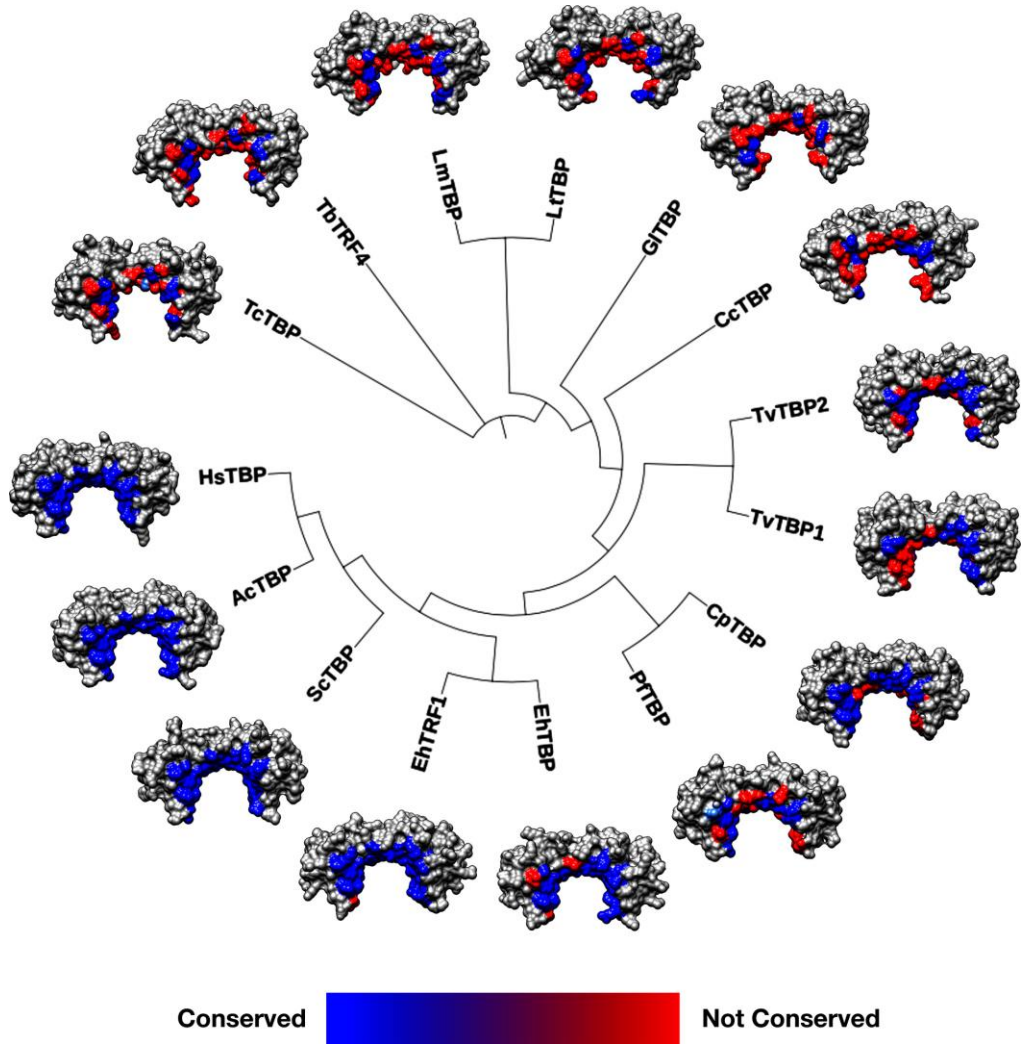


Fig. 2. Comparison of amino acids involved in the binding of TBPs to DNA. Amino acid conservation in protist TBPs is shown, including two canonical TBPs (HsTBP and ScTBP). Conserved amino acids with respect to the canonical TBPs are shown in blue, while non-conserved amino acids are shown in red. The structures of HsTBP and ScTBP were modelled using their own crystals (1cdw and 4b0A, respectively). The ScTBP crystal was used as a template to obtain the tertiary structures of protist TBPs as visualized with the Chimera program [128]. The phylogenetic tree was generated with the ANCESCON tool [126] in the Bioinformatics Toolkit platform [127].

with deeply divergent lineages [52].

For many years, protists were considered to be simple organisms [124]; nevertheless, they have to adapt to a wide variety of micro-environments, generating great variability in their molecular mechanisms, including the transcription process. In contrast to those of higher eukaryotes, the RNA polymerases of *E. histolytica, T. brucei*, *T. vaginalis*, and *G. lamblia* have a low sensitivity to α-amanitin, suggesting that the properties of their RNA polymerases are more diverse than those in metazoans [125]. Additionally, studies about gene expression in protists have revealed particular mechanisms such as RNA editing and *trans*-splicing in kinetoplastids [120]. Although general transcription factors for transcription initiation have been considered universally conserved, important differences in the number and composition of transcriptional regulators have now become apparent in unicellular and multicellular organisms [5].

TBP was originally described as a universal transcription factor. In eukaryotes, TBP participates in the process of transcription by each of the three types of RNA polymerases. However, in recent years, TBP-related factors have been described, and their roles in the formation of several transcription initiation complexes have been addressed. This group of proteins has received different names and, to date, attempts have been made to classify them into four subgroups (TRF1, TRF2, TRF3 and TRF4). TBP-related factors emerged in multicellular organisms as a form of specialization (e.g., tissue-specific) and are involved in

distinct biological processes [5]. Interestingly, *T. vaginalis* and *E. histolytica* are two unicellular protists that express more than one TBP. Protist TBPs cannot be classified as canonical TBPs or TRFs as described in metazoans because protist TBPs share characteristics of both TBP and TRF [73,118].

The diversity of protist TBPs and their DNA-binding ability is shown in Fig. 1, with their sequence identity referred to HsTBP. *H. sapiens* and yeast TBPs bind canonical TATA boxes; interestingly, as the sequence identity of protist TBPs is diminished, the binding of TBP to DNA seems to become more widespread, even as their specific recognition of conserved TATA sequences is progressively lost. It is noteworthy that *A. castellanii* TBP, which shares 84 % identity with HsTBP, recognizes a sequence that can be considered a canonical TATA box. EhTBP1 and EhTRF (approximately 50 % identical to HsTBP) bind variants of the canonical TATA box, while TvTBPs, with approximately 40 % identity to HsTBP, bind non-specifically to DNA sequences [86]. These data correlate with the absence of conserved TATA box elements in these organisms.

Fig. 2 shows the hypothetical tertiary structure of TBPs, modelled using the ScTBP crystal as the template and depicted following their phylogenetic position as obtained by the ANCESCON tool [126] from the Bioinformatics Toolkit platform [127]. It shows that all TBPs fold in the consensus saddle-shaped structure, and the conserved amino acids involved in binding to DNA are emphasized. When the data in Figs. 1 and 2 are compared, it is clear that the conservation of the amino acids involved in TATA box recognition clearly correlates with the presence of canonical TATA boxes. Accordingly, TBPs in protists that do not have canonical TATA boxes in their promoters contain more variations in the residues located at the concave surface of this protein. In a study evaluating the diversity of TBP sequences among species in Eukarya, the variant residues that appeared in common ancestors of the eukaryotic TBPs were observed to have been maintained during evolution in all Eukarya except protists, reflecting the biological diversity of these organisms [55]. In conclusion, TBPs are highly conserved among higher eukaryotes [3], and their conservation is lower among protists, a feature that correlates with the variability observed in the transcription machinery and regulatory elements among early divergent eukaryotes.

Funding

This work was supported by a grant to Imelda López Villaseñor from Dirección General de Asuntos del Personal Académico (DGAPA PAPIIT-IN211320), Universidad Nacional Autónoma de México (UNAM).

Acknowledgements

OPM is a doctoral student from Programa de Maestría y Doctorado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México (UNAM) and was supported by a scholarship from CONACyT, México. This article is a partial fulfilment for obtaining her PhD degree.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molbiopara.2020.111312.

References

[1] A. Vannini, P. Cramer, Conservation between the RNA polymerase I, II, and III transcription initiation machineries, Mol. Cell 45 (2012) 439–446, https://doi.org/10.1016/j.molcel.2012.01.023.
[2] P. Cramer, Eukaryotic transcription turns 50, Cell 179 (2019) 808–812, https://doi.org/10.1016/j.cell.2019.09.018.
[3] N. Hernandez, TBP, a universal eukaryotic transcription factor? Genes Dev. 7 (1993) 1291–1308, https://doi.org/10.1101/gad.7.7b.1291.
[4] C. Yang, E. Bolotin, T. Jiang, F.M. Sladek, E. Martinez, Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters, Gene 389 (2007) 52–65, https://doi.org/10.1016/j.gene.2006.09.029.
[5] S.H.C. Duttke, Evolution and diversification of the basal transcription machinery, Trends Biochem. Sci. 40 (2015) 127–129, https://doi.org/10.1016/j.tibs.2015.01.005.
[6] F. Müller, M.A. Demény, L. Tora, New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors, J. Biol. Chem. 282 (2007) 14685–14689, https://doi.org/10.1074/jbc.R700012200.
[7] S. Ahn, C.-L. Huang, E. Ozkumur, X. Zhang, J. Chinnala, A. Yalcin, S. Bandyopadhyay, S. Russek, M.S. Unlü, C. DeLisi, R.J. Irani, TATA binding proteins can recognize nontraditional DNA sequences, Biophys. J. 103 (2012) 1510–1517, https://doi.org/10.1016/j.bpj.2012.08.030.
[8] M.L. Sogin, J.D. Silberman, Evolution of the protists and protistan parasites from the perspective of molecular systematics, Int. J. Parasitol. 28 (1998) 11–20, https://doi.org/10.1016/S0020-7519(97)00181-1.
[9] T. Matsui, J. Segall, P.A. Weil, R.G. Roeder, Multiple factors required for accurate initiation of transcription by purified RNA polymerase II, J. Biol. Chem. 255 (1980) 11992–11996 http://www.ncbi.nlm.nih.gov/pubmed/7440580.
[10] D.B. Nikolov, H. Chen, E.D. Halay, A. Hoffman, R.G. Roeder, S.K. Burley, Crystal structure of a human TATA box-binding protein/TATA element complex, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 4862–4867, https://doi.org/10.1073/pnas.93.10.4862.
[11] M. Lee, K. Struhl, Multiple functions of the nonconserved N-terminal domain of yeast TATA-binding protein, Genetics 158 (2001) 87–93 http://www.ncbi.nlm.nih.gov/pubmed/11333220.
[12] M.G. Peterson, N. Tanese, B.F. Pugh, R. Tjian, Functional domains and upstream activation properties of cloned human TATA binding protein, Science (80-.) 248 (1990) 1625–1630, https://doi.org/10.1126/science.2363050.
[13] B.F. Pugh, R. Tjian, Mechanism of transcriptional activation by Sp1: evidence for coactivators, Cell 61 (1990) 1187–1197, https://doi.org/10.1016/0092-8674(90)90683-6.
[14] L.A. Stargell, M.A. Gorovsky, TATA-binding protein and nuclear differentiation in Tetrahymena thermophila, Mol. Cell. Biol. 14 (1994) 723–734, https://doi.org/10.1128/mcb.14.1.723.
[15] Y. Zehavi, A. Kedmi, D. Ideses, T. Juven-Gershon, TRF2: TRansForming the view of general transcription factors, Transcription. 6 (2015) 1–6, https://doi.org/10.1080/21541264.2015.1004980.
[16] D.B. Nikolov, H. Chen, E.D. Halay, A.A. Usheva, K. Hisatake, D.K. Lee, R.G. Roeder, S.K. Burley, Crystal structure of a TFIIB-TBP-TATA-element ternary complex, Nature 377 (1995) 119–128, https://doi.org/10.1038/377119a0.
[17] S. Tan, Y. Hunziker, D.F. Sargent, T.J. Richmond, Crystal structure of a yeast TFIIA/TBP/DNA complex, Nature 381 (1996) 127–151, https://doi.org/10.1038/381127a0.
[18] B.S. DeDecker, R. O'Brien, P.J. Fleming, J.H. Geiger, S.P. Jackson, P.B. Sigler, The crystal structure of a hyperthermophilic archaeal TATA-box binding protein, J. Mol. Biol. 264 (1996) 1072–1084, https://doi.org/10.1006/jmbi.1996.0697.
[19] M.C. Thomas, C.-M. Chiang, The general transcription machinery and general cofactors, Crit. Rev. Biochem. Mol. Biol. 41 (2006) 105–178, https://doi.org/10.1080/10409230600648736.
[20] R. Dikstein, The unexpected traits associated with core promoter elements, Transcription. 2 (2011) 201–206, https://doi.org/10.4161/trns.2.5.17271.
[21] J.L. Kim, D.B. Nikolov, S.K. Burley, Co-crystal structure of TBP recognizing the minor groove of a TATA element, Nature 365 (1993) 520–527, https://doi.org/10.1038/365520a0.
[22] Y. Kim, J.H. Geiger, S. Hahn, P.B. Sigler, Crystal structure of a yeast TBP/TATA-box complex, Nature 365 (1993) 512–520, https://doi.org/10.1038/365512a0.
[23] Z.S. Juo, T.K. Chiu, P.M. Leiberman, I. Baikalov, A.J. Berk, R.E. Dickerson, How proteins recognize the TATA box, J. Mol. Biol. 261 (1996) 239–254, https://doi.org/10.1006/jmbi.1996.0456.
[24] R.J. White, S.P. Jackson, The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III, Trends Genet. 8 (1992) 284–288, https://doi.org/10.1016/0168-9525(92)90255-3.
[25] E.P. Geiduschek, G.A. Kassavetis, The RNA polymerase III transcription apparatus, J. Mol. Biol. 310 (2001) 1–26, https://doi.org/10.1006/jmbi.2001.4732.
[26] I. Davidson, The genetics of TBP and TBP-related factors, Trends Biochem. Sci. 28 (2003) 391–398, https://doi.org/10.1016/S0968-0004(03)00117-8.
[27] W. Akhtar, G.J.C. Veenstra, TBP-related factors: a paradigm of diversity in transcription initiation, Cell Biosci. 1 (2011) 23, https://doi.org/10.1186/2045-3701-1-23.
[28] C. Millán-Pacheco, V.M. Capistrán, N. Pastor, On the consequences of placing amino groups at the TBP-DNA interface. Does TATA really matter? J. Mol. Recognit. 22 (2009) 453–464, https://doi.org/10.1002/jmr.963.
[29] M.M. Kurshakova, E.N. Nabirochkina, S.G. Georgieva, D.V. Kopytova, TRF4, the novel TBP-related protein of Drosophila melanogaster, is concentrated at the endoplasmic reticulum and copurifies with proteins participating in the processes associated with endoplasmic reticulum, J. Cell. Biochem. 120 (2018) 7927–7939, https://doi.org/10.1002/jcb.28070.
[30] T.E. Crowley, T. Hoey, J.K. Liu, Y.N. Jan, L.Y. Jan, R. Tjian, A new factor related to TATA-binding protein has highly restricted expression patterns in Drosophila, Nature 361 (1993) 557–561, https://doi.org/10.1038/361557a0.
[31] S.K. Hansen, S. Takada, R.H. Jacobson, J.T. Lis, R. Tjian, Transcription properties of a cell type-specific TATA-binding protein, TRF, Cell 91 (1997) 71–83, https://doi.org/10.1016/s0092-8674(01)80010-6.
[32] R. Hori, M. Carey, Transcription: TRF walks the walk but can it talk the talk? Curr.

Biol. 8 (1998) R124–R127, https://doi.org/10.1016/s0960-9822(98)70987-x.
[33] Y. Isogai, S. Takada, R. Tjian, S. Keleş, Novel TRF1/BRF target genes revealed by genome-wide analysis of Drosophila pol III transcription, EMBO J. 26 (2007) 79–89, https://doi.org/10.1038/sj.emboj.7601448.
[34] S. Takada, J.T. Lis, S. Zhou, R. Tjian, A TRF1:BRF complex directs Drosophila rna polymerase III transcription, Cell 101 (2000) 459–469, https://doi.org/10.1016/s0092-8674(00)80857-0.
[35] N. Verma, K.-H. Hung, J.J. Kang, N.H. Barakat, W.E. Stumph, Differential utilization of TATA box-binding protein (TBP) and TBP-related factor 1 (TRF1) at different classes of RNA polymerase III promoters, J. Biol. Chem. 288 (2013) 27564–27570, https://doi.org/10.1074/jbc.C113.503094.
[36] J.C. Dantonel, J.M. Wurtz, O. Poch, D. Moras, L. Tora, The TBP-like factor: an alternative transcription factor in metazoa? Trends Biochem. Sci. 24 (1999) 335–339, https://doi.org/10.1016/s0968-0004(99)01436-x.
[37] T. Ohbayashi, T. Kishimoto, Y. Makino, M. Shimada, T. Nakadai, T. Aoki, T. Kawata, S. Niwa, T. Tamura, Isolation of cDNA, chromosome mapping, and expression of the human TBP-like protein, Biochem. Biophys. Res. Commun. 255 (1999) 137–142, https://doi.org/10.1006/bbrc.1999.0159.
[38] S.H.C. Duttke, R.F. Doolittle, Y.-L. Wang, J.T. Kadonaga, TRF2 and the evolution of the bilateria, Genes Dev. 28 (2014) 2071–2076, https://doi.org/10.1101/gad.250563.114.
[39] M.D. Rabenstein, S. Zhou, J.T. Lis, R. Tjian, TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 4791–4796, https://doi.org/10.1073/pnas.96.9.4791.
[40] E. Maldonado, Transcriptional functions of a new mammalian TATA-binding protein-related factor, J. Biol. Chem. 274 (1999) 12963–12966, https://doi.org/10.1074/jbc.274.19.12963.
[41] P.A. Moore, J. Ozer, M. Salunek, G. Jan, D. Zerby, S. Campbell, P.M. Lieberman, A human TATA binding protein-related protein with altered DNA binding specificity inhibits transcription from multiple promoters and activators, Mol. Cell. Biol. 19 (1999) 7610–7620, https://doi.org/10.1128/mcb.19.11.7610.
[42] M. Teichmann, Z. Wang, E. Martinez, A. Tjernberg, D. Zhang, F. Vollmer, B.T. Chait, R.G. Roeder, Human TATA-binding protein-related factor-2 (hTRF2) stably associates with hTFIIA in HeLa cells, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 13720–13725, https://doi.org/10.1073/pnas.96.24.13720.
[43] A.J. Berk, TBP-like factors come into focus, Cell 103 (2000) 5–8, https://doi.org/10.1016/s0092-8674(00)00098-2.
[44] S.P. Persengiev, X. Zhu, B.L. Dixit, G.A. Maston, E.L.W. Kittler, M.R. Green, TRF3, a TATA-box-binding protein-related factor, is vertebrate-specific and widely expressed, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 14887–14891, https://doi.org/10.1073/pnas.2036440100.
[45] R. Bártfai, C. Balduf, T. Hilton, Y. Rathmann, Y. Hadzhiev, L. Tora, L. Orbán, F. Müller, TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish, Curr. Biol. 14 (2004) 593–598, https://doi.org/10.1016/j.cub.2004.03.034.
[46] Y. Yang, J. Cao, L. Huang, H.Y. Fang, H.Z. Sheng, Regulated expression of TATA-binding protein-related factor 3 (TRF3) during early embryogenesis, Cell Res. 16 (2006) 610–621, https://doi.org/10.1038/sj.cr.7310064.
[47] Z. Jallow, U.G. Jacobi, D.L. Weeks, I.B. Dawid, G.J.C. Veenstra, Specialized and redundant roles of TBP and a vertebrate-specific TBP paralog in embryonic gene regulation in Xenopus, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 13525–13530, https://doi.org/10.1073/pnas.0405536101.
[48] D.O. Hart, T. Raha, N.D. Lawson, M.R. Green, Initiation of zebrafish haematopoiesis by the TATA-box-binding protein-related factor Trf3, Nature 450 (2007) 1082–1085, https://doi.org/10.1038/nature06349.
[49] D.O. Hart, M.K. Santra, T. Raha, M.R. Green, Selective interaction between Trf3 and Taf3 required for early development and hematopoiesis, Dev. Dyn. 238 (2009) 2540–2549, https://doi.org/10.1002/dvdy.22083.
[50] M.A. Sleigh, Protozoa as Members of the Protista, in: Protozoa and Other Protists, Cambridge University Press, 1989, p. 342, https://doi.org/10.1017/S002531540003232X.
[51] T. Cavalier-Smith, Kingdom protozoa and its 18 phyla, Microbiol. Rev. 57 (1993) 953–994, https://doi.org/10.1128/mmbr.57.4.953-994.1993.
[52] F. Burki, A.J. Roger, M.W. Brown, A.G.B. Simpson, The new tree of eukaryotes, trends ecol, Evol. 35 (2020) 43–55, https://doi.org/10.1016/j.tree.2019.08.008.
[53] R.G. Yaeger, Protozoa: structure, classification, growth, and development. In: medical microbiology, in: S. Baron (Ed.), Med. Microbiol. 4th ed., University of Texas Medical Branch at Galveston, Galveston (TX), 1996p. Chapter 77. http://www.ncbi.nlm.nih.gov/pubmed/21413323.
[54] A.A. Best, H.G. Morrison, A.G. McArthur, M.L. Sogin, G.J. Olsen, Evolution of eukaryotic transcription: insights from the genome of Giardia lamblia, Genome Res. 14 (2004) 1537–1547, https://doi.org/10.1101/gr.2256604.
[55] E. Kawakami, N. Adachi, T. Senda, M. Horikoshi, Leading role of TBP in the establishment of complexity in eukaryotic transcription initiation systems, Cell Rep. 21 (2017) 3941–3956, https://doi.org/10.1016/j.celrep.2017.12.034.
[56] R. Siddiqui, N.A. Khan, Biology and pathogenesis of Acanthamoeba, Parasit. Vectors 5 (2012) 6, https://doi.org/10.1186/1756-3305-5-6.
[57] N.A. Khan, Acanthamoeba: biology and increasing importance in human health, FEMS Microbiol. Rev. 30 (2006) 564–595, https://doi.org/10.1111/j.1574-6976.2006.00023.x.
[58] J. Lorenzo-Morales, C.M. Martín-Navarro, A. López-Arencibia, F. Arnalich-Montiel, J.E. Piñero, B. Valladares, Acanthamoeba keratitis: an emerging disease gathering importance worldwide? Trends Parasitol. 29 (2013) 181–187, https://doi.org/10.1016/j.pt.2013.01.006.
[59] W. Huang, E. Bateman, Transcription of the Acanthamoeba TATA-binding protein gene. A single transcription factor acts both as an activator and a repressor, J. Biol. Chem. 272 (1997) 3852–3859, https://doi.org/10.1074/jbc.272.6.3852.
[60] J.M. Wong, F. Liu, E. Bateman, Cloning and expression of the Acanthamoeba castellanii gene encoding transcription factor TFIID, Gene 117 (1992) 91–97, https://doi.org/10.1016/0378-1119(92)90494-a.
[61] J.M. Wong, E. Bateman, TBP-DNA interactions in the minor groove discriminate between A:t and T:a base pairs, Nucleic Acids Res. 22 (1994) 1890–1896, https://doi.org/10.1093/nar/22.10.1890.
[62] C.A. Radebaugh, J.L. Matthews, G.K. Geiss, F. Liu, J.M. Wong, E. Bateman, S. Camier, A. Sentenac, M.R. Paule, TATA box-binding protein (TBP) is a constituent of the polymerase I-specific transcription initiation factor TIF-IB (SL1) bound to the rRNA promoter and shows differential sensitivity to TBP-directed reagents in polymerase I, II, and III transcription fac, Mol. Cell. Biol. 14 (1994) 597–605, https://doi.org/10.1128/mcb.14.1.597.
[63] F. Liu, E. Bateman, An upstream promoter element of the Acanthamoeba castellanii TBP gene binds a DNA sequence specific transcription activating protein, TPBF, Nucleic Acids Res. 21 (1993) 4321–4329, https://doi.org/10.1093/nar/21.18.4321.
[64] T. Cavalier-Smith, The origin of eukaryotic and archaebacterial cells, Ann. N. Y. Acad. Sci. 503 (1987) 17–54, https://doi.org/10.1111/j.1749-6632.1987.tb40596.x.
[65] T. Cavalier-Smith, Archaebacteria and archezoa, Nature 339 (1989), https://doi.org/10.1038/339100a0 100-01.
[66] J.E. Purdy, L.T. Pho, B.J. Mann, W.A. Petri, Upstream regulatory elements controlling expression of the Entamoeba histolytica lectin, Mol. Biochem. Parasitol. 78 (1996) 91–103, https://doi.org/10.1016/s0166-6851(96)02614-x.
[67] C.G. Clark, U.C.M. Alsmark, M. Tazreiter, Y. Saito-Nakano, V. Ali, S. Marion, C. Weber, C. Mukherjee, I. Bruchhaus, E. Tannich, M. Leippe, T. Sicheritz-Ponten, P.G. Foster, J. Samuelson, C.J. Noël, R.P. Hirt, T.M. Embley, C.A. Gilchrist, B.J. Mann, U. Singh, J.P. Ackers, S. Bhattacharya, A. Bhattacharya, A. Lohia, N. Guillén, M. Duchêne, T. Nozaki, N. Hall, Structure and content of the Entamoeba histolytica genome, Adv. Parasitol. 65 (2007) 51–190, https://doi.org/10.1016/S0065-308X(07)65002-7.
[68] H. Buss, C. Lioutas, S. Dobinsky, R. Nickel, E. Tannich, Analysis of the 170-kDa lectin gene promoter of Entamoeba histolytica, Mol. Biochem. Parasitol. 72 (1995) 1–10, https://doi.org/10.1016/0166-6851(95)00060-e.
[69] U. Singh, J.B. Rogers, B.J. Mann, W.A. Petri, Transcription initiation is controlled by three core promoter elements in the hgl5 gene of the protozoan parasite Entamoeba histolytica, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 8812–8817, https://doi.org/10.1073/pnas.94.16.8812.
[70] U. Singh, C.A. Gilchrist, J.M. Schaenman, J.B. Rogers, J.W. Hockensmith, B.J. Mann, W.A. Petri, Context-dependent roles of the Entamoeba histolytica core promoter element GAAC in transcriptional activation and protein complex assembly, Mol. Biochem. Parasitol. 120 (2002) 107–116, https://doi.org/10.1016/s0166-6851(01)00441-8.
[71] I. Bruchhaus, M. Leippe, C. Lioutas, E. Tannich, Unusual gene organization in the protozoan parasite Entamoeba histolytica, DNA Cell Biol. 12 (1993) 925–933, https://doi.org/10.1089/dna.1993.12.925.
[72] I.K.M. Ali, G.M. Ehrenkaufer, J.A. Hackney, U. Singh, Growth of the protozoan parasite Entamoeba histolytica in 5-azacytidine has limited effects on parasite gene expression, BMC Genomics 8 (2007) 7, https://doi.org/10.1186/1471-2164-8-7.
[73] C.A. Castañon-Sanchez, J.P. Luna-Arias, M.G. de Dios-Bravo, M.E. Herrera-Aguirre, J.J. Olivares-Trejo, E. Orozco, J.M. Hernandez, Entamoeba histolytica: a unicellular organism containing two active genes encoding for members of the TBP family, Protein Expr. Purif. 70 (2010) 48–59, https://doi.org/10.1016/j.pep.2009.12.007.
[74] J.P. Luna-Arias, R. Hernandez-Rivas, G. de Dios-Bravo, J. Garcia, L. Mendoza, E. Orozco, The TATA-box binding protein of Entamoeba histolytica: cloning of the gene and location of the protein by immunofluorescence and confocal microscopy, Microbiology 145 (Pt 1) (1999) 33–40, https://doi.org/10.1099/13500872-145-1-33.
[75] G. de Dios-Bravo, J.P. Luna-Arias, A.M. Riverón, J.J. Olivares-Trejo, C. López-Camarillo, E. Orozco, Entamoeba histolytica TATA-box binding protein binds to different TATA variants in vitro, FEBS J. 272 (2005) 1354–1366, https://doi.org/10.1111/j.1742-4658.2005.04566.x.
[76] R.K. Narayanasamy, C.A. Castañón-Sanchez, J.P. Luna-Arias, G. García-Rivera, B. Avendaño-Borromeo, M.L. Labra-Barrios, J. Valdés, M.E. Herrera-Aguirre, E. Orozco, The Entamoeba histolytica TBP and TRF1 transcription factors are GAAC-box binding proteins, which display differential gene expression under different stress stimuli and during the interaction with mammalian cells, Parasit. Vectors 11 (2018) 153, https://doi.org/10.1186/s13071-018-2698-7.
[77] D.F. Harp, I. Chowdhury, Trichomoniasis: evaluation to execution, Eur. J. Obstet. Gynecol. Reprod. Biol. 157 (2011) 3–9, https://doi.org/10.1016/j.ejogrb.2011.02.024.
[78] G. Dias-Lopes, L. Saboia-Vahia, E.T. Margotti, N. de S. Fernandes, C.L. de F. Castro, F.O. Oliveira, J.F. Peixoto, C. Britto, F.C.E. Silva, P. Cuervo, J.B. de Jesus, Morphologic study of the effect of iron on pseudocyst formation in Trichomonas vaginalis and its interaction with human epithelial cells, Mem. Inst. Oswaldo Cruz 112 (2017) 664–673, https://doi.org/10.1590/0074-02760170032.
[79] J.M. Carlton, R.P. Hirt, J.C. Silva, A.L. Delcher, M. Schatz, Q. Zhao, J.R. Wortman, S.L. Bidwell, U.C.M. Alsmark, S. Besteiro, T. Sicheritz-Ponten, C.J. Noel, J.B. Dacks, P.G. Foster, C. Simillion, Y. Van de Peer, D. Miranda-Saavedra, G.J. Barton, G.D. Westrop, S. Müller, D. Dessi, P.L. Fiori, Q. Ren, I. Paulsen, H. Zhang, F.D. Bastida-Corcuera, A. Simoes-Barbosa, M.T. Brown, R.D. Hayes, M. Mukherjee, C.Y. Okumura, R. Schneider, A.J. Smith, S. Vanacova, M. Villalvazo, B.J. Haas, M. Pertea, T.V. Feldblyum, T.R. Utterback, C.-L. Shu,

K. Osoegawa, P.J. de Jong, I. Hrdy, L. Horvathova, Z. Zubacova, P. Dolezal, S.-B. Malik, J.M. Logsdon, K. Henze, A. Gupta, C.C. Wang, R.L. Dunne, J.A. Upcroft, P. Upcroft, O. White, S.L. Salzberg, P. Tang, C. Chiu, Y. Lee, T.M. Embley, G.H. Coombs, J.C. Mottram, J. Tachezy, C.M. Fraser-Liggett, P.J. Johnson, Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis, Science (80-.) 315 (2007) 207–212, https://doi.org/10.1126/science.1132894.

[80] D.V.K. Quon, M.G. Delgadillo, A. Khachi, S.T. Smale, P.J. Johnson, Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 4579–4583, https://doi.org/10.1073/pnas.91.10.4579.

[81] D.R. Liston, P.J. Johnson, Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element, Mol. Cell. Biol. 19 (1999) 2380–2388, https://doi.org/10.1128/mcb.19.3.2380.

[82] M.A. Schumacher, A.O.T. Lau, P.J. Johnson, Structural basis of core promoter recognition in a primitive eukaryote, Cell 115 (2003) 413–424, https://doi.org/10.1016/s0092-8674(03)00887-0.

[83] A.J. Smith, L. Chudnovsky, A. Simoes-Barbosa, M.G. Delgadillo-Correa, Z.O. Jonsson, J.A. Wohlschlegel, P.J. Johnson, Novel core promoter elements and a cognate transcription factor in the divergent unicellular eukaryote Trichomonas vaginalis, Mol. Cell. Biol. 31 (2011) 1444–1458, https://doi.org/10.1128/MCB.00745-10.

[84] D.R. Liston, A.O. Lau, D. Ortiz, S.T. Smale, P.J. Johnson, Initiator recognition in a primitive eukaryote: IBP39, an initiator-binding protein from Trichomonas vaginalis, Mol. Cell. Biol. 21 (2001) 7872–7882, https://doi.org/10.1128/MCB.21.22.7872-7882.2001.

[85] A.O.T. Lau, A.J. Smith, M.T. Brown, P.J. Johnson, Trichomonas vaginalis initiator binding protein (IBP39) and RNA polymerase II large subunit carboxy terminal domain interaction, Mol. Biochem. Parasitol. 150 (2006) 56–62, https://doi.org/10.1016/j.molbiopara.2006.06.008.

[86] O. Parra-Marín, L. Rosas-Hernández, K. López-Pacheco, B. Franco, A. Ibáñez-Escribano, R. Hernández, I. López-Villaseñor, An in vitro characterisation of the Trichomonas vaginalis TATA box-binding proteins (TBPs), Parasitol. Res. 118 (2019) 3019–3031, https://doi.org/10.1007/s00436-019-06438-z.

[87] G.J. Leitch, Q. He, Cryptosporidiosis-an overview, J. Biomed. Res. 25 (2011) 1–16, https://doi.org/10.1016/S1674-8301(11)60001-8.

[88] M.S. Abrahamsen, T.J. Templeton, S. Enomoto, J.E. Abrahante, G. Zhu, C.A. Lancto, M. Deng, C. Liu, G. Widmer, S. Tzipori, G.A. Buck, P. Xu, A.T. Bankier, P.H. Dear, B.A. Konfortov, H.F. Spriggs, L. Iyer, V. Anantharaman, L. Aravind, V. Kapur, Complete genome sequence of the apicomplexan, Cryptosporidium parvum, Science (80-.) 304 (2004) 441–445, https://doi.org/10.1126/science.1094786.

[89] J.J. Millership, P. Waghela, X. Cai, A. Cockerham, G. Zhu, Differential expression and interaction of transcription co-activator MBF1 with TATA-binding protein (TBP) in the apicomplexan Cryptosporidium parvum, Microbiology 150 (2004) 1207–1213, https://doi.org/10.1099/mic.0.26891-0.

[90] G. Zhu, M.J. LaGier, S. Hirose, J.S. Keithly, Cryptosporidium parvum: functional complementation of a parasite transcriptional coactivator CpMBF1 in yeast, Exp. Parasitol. 96 (2000) 195–201, https://doi.org/10.1006/expr.2000.4574.

[91] A.F. Cowman, J. Healer, D. Marapana, K. Marsh, Malaria: biology and disease, Cell 167 (2016) 610–624, https://doi.org/10.1016/j.cell.2016.07.055.

[92] M.J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R.W. Hyman, J.M. Carlton, A. Pain, K.E. Nelson, S. Bowman, I.T. Paulsen, K. James, J.A. Eisen, K. Rutherford, S.L. Salzberg, A. Craig, S. Kyes, M. Chan, V. Nene, S.J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M.W. Mather, A.B. Vaidya, D.M.A. Martin, A.H. Fairlamb, M.J. Fraunholz, D.S. Roos, S.A. Ralph, G.I. McFadden, L.M. Cummings, G.M. Subramanian, C. Mungall, J.C. Venter, D.J. Carucci, S.L. Hoffman, C. Newbold, R.W. Davis, C.M. Fraser, B. Barrell, Genome sequence of the human malaria parasite Plasmodium falciparum, Nature 419 (2002) 498–511, https://doi.org/10.1038/nature01097.

[93] R.M.R. Coulson, N. Hall, C.A. Ouzounis, Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum, Genome Res. 14 (2004) 1548–1554, https://doi.org/10.1101/gr.2218604.

[94] I. Callebaut, K. Prat, E. Meurice, J.-P. Mornon, S. Tomavo, Prediction of the general transcription factors associated with RNA polymerase II in Plasmodium falciparum: conserved features and differences relative to other eukaryotes, BMC Genomics 6 (2005) 100, https://doi.org/10.1186/1471-2164-6-100.

[95] M.B. McAndrew, M. Read, P.F.G. Sims, J.E. Hyde, Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T-rich human malaria parasite Plasmodium falciparum, Gene 124 (1993) 165–171, https://doi.org/10.1016/0378-1119(93)90390-o.

[96] O.K. Ruvalcaba-Salazar, M. del Carmen Ramírez-Estudillo, D. Montiel-Condado, F. Recillas-Targa, M. Vargas, R. Hernández-Rivas, Recombinant and native Plasmodium falciparum TATA-binding-protein binds to a specific TATA box element in promoter regions, Mol. Biochem. Parasitol. 140 (2005) 183–196, https://doi.org/10.1016/j.molbiopara.2005.01.002.

[97] C.G. Toenhake, R. Bártfai, What functional genomics has taught us about transcriptional regulation in malaria parasites, Brief. Funct. Genomics 18 (2019) 290–301, https://doi.org/10.1093/bfgp/elz004.

[98] A.M. Gopalakrishnan, L.A. Nyindodo, M. Ross Fergus, C. López-Estraño, Plasmodium falciparum: preinitiation complex occupancy of active and inactive promoters during erythrocytic stage, Exp. Parasitol. 121 (2009) 46–54, https://doi.org/10.1016/j.exppara.2008.09.016.

[99] E. Bischoff, C. Vaquero, In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of Plasmodium falciparum, BMC Genomics 11 (2010) 34, https://doi.org/10.1186/1471-2164-11-34.

[100] P.J. Johnson, J.M. Kooter, P. Borst, Inactivation of transcription by UV irradiation of T. Brucei provides evidence for a multicistronic transcription unit including a VSG gene, Cell 51 (1987) 273–281, https://doi.org/10.1016/0092-8674(87)90154-1.

[101] S. Martínez-Calvillo, J.C. Vizuet-de-Rueda, L.E. Florencio-Martínez, R.G. Manning-Cela, E.E. Figueroa-Angulo, Gene expression in trypanosomatid parasites, J. Biomed. Biotechnol. 2010 (2010) 525241, , https://doi.org/10.1155/2010/525241.

[102] M.Z. Handler, P.A. Patel, R. Kapila, Y. Al-Qubati, R.A. Schwartz, Cutaneous and mucocutaneous leishmaniasis: clinical perspectives, J. Am. Acad. Dermatol. 73 (2015) 897–908, https://doi.org/10.1016/j.jaad.2014.08.051 quiz 909–910.

[103] A.C. Ivens, C.S. Peacock, E.A. Worthey, L. Murphy, G. Aggarwal, M. Berriman, E. Sisk, M. Rajandream, E. Adlem, R. Aert, A. Anupama, Z. Apostolou, P. Attipoe, N. Bason, C. Bauser, A. Beck, S.M. Beverley, G. Bianchettin, K. Borzym, G. Bothe, C.V. Bruschi, M. Collins, E. Cadag, L. Ciarloni, C. Clayton, R.M.R. Coulson, A. Cronin, A.K. Cruz, R.M. Davies, J. De Gaudenzi, D.E. Dobson, A. Duesterhoeft, G. Fazelina, N. Fosker, A.C. Frasch, A. Fraser, M. Fuchs, C. Gabel, A. Goble, A. Goffeau, D. Harris, C. Hertz-Fowler, H. Hilbert, D. Horn, Y. Huang, S. Klages, A. Knights, M. Kube, N. Larke, L. Litvin, A. Lord, T. Louie, M. Marra, D. Masuy, K. Matthews, S. Michaeli, J.C. Mottram, S. Müller-Auer, H. Munden, S. Nelson, H. Norbertczak, K. Oliver, S. O'neil, M. Pentony, T.M. Pohl, C. Price, B. Purnelle, M.A. Quail, E. Rabbinowitsch, R. Reinhardt, M. Rieger, J. Rinta, J. Robben, L. Robertson, J.C. Ruiz, S. Rutter, D. Saunders, M. Schäfer, J. Schein, D.C. Schwartz, K. Seeger, A. Seyler, S. Sharp, H. Shin, D. Sivam, R. Squares, S. Squares, V. Tosato, C. Vogt, G. Volckaert, R. Wambutt, T. Warren, H. Wedler, J. Woodward, S. Zhou, W. Zimmermann, D.F. Smith, J.M. Blackwell, K.D. Stuart, B. Barrell, P.J. Myler, The genome of the kinetoplastid parasite, Leishmania major, Science (80-.) 309 (2005) 436–442, https://doi.org/10.1126/science.1112680.

[104] F.C. Román-Carraro, L.E. Florencio-Martínez, G. Romero-Meza, T. Nepomuceno-Mejía, J.C. Carrero, R. Arroyo, J. Ortega-López, R.G. Manning-Cela, S. Martínez-Calvillo, TFIIIB subunit Bdp1 participates in RNA polymerase III transcription in the protozoan parasite leishmania major, Biomed Res. Int. 2019 (2019) 1425281, , https://doi.org/10.1155/2019/1425281.

[105] S. Thomas, A. Green, N.R. Sturm, D.A. Campbell, P.J. Myler, Histone acetylations mark origins of polycistronic transcription in Leishmania major, BMC Genomics 10 (2009) 152, https://doi.org/10.1186/1471-2164-10-152.

[106] F. Raymond, S. Boisvert, G. Roy, J.-F. Ritt, D. Légaré, A. Isnard, M. Stanke, M. Olivier, M.J. Tremblay, B. Papadopoulou, M. Ouellette, J. Corbeil, Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species, Nucleic Acids Res. 40 (2011) 1131–1147, https://doi.org/10.1093/nar/gkr834.

[107] L. Simpson, R. Aphasizhev, G. Gao, X. Kang, Mitochondrial proteins and complexes in Leishmania and Trypanosoma involved in U-insertion/deletion RNA editing, RNA 10 (2004) 159–170, https://doi.org/10.1261/rna.5170704.

[108] S. Thomas, M.C. Yu, N.R. Sturm, D.A. Campbell, A non-universal transcription factor? The Leishmania tarentolae TATA box-binding protein LtTBP associates with a subset of promoters, Int. J. Parasitol. 36 (2006) 1217–1226, https://doi.org/10.1016/j.ijpara.2006.04.004.

[109] K.M. Tyler, D.M. Engman, The life cycle of Trypanosoma cruzi revisited, Int. J. Parasitol. 31 (2001) 4724–4781, https://doi.org/10.1016/s0020-7519(01)00153-9.

[110] F. Callejas-Hernández, N. Gironès, M. Fresno, Genome sequence of trypanosoma cruzi strain Bug2148, Genome Announc. 6 (2018) 1–2, https://doi.org/10.1128/genomeA.01497-17.

[111] P. Cribb, E. Serra, One- and two-hybrid analysis of the interactions between components of the Trypanosoma cruzi spliced leader RNA gene promoter binding complex, Int. J. Parasitol. 39 (2009) 525–532, https://doi.org/10.1016/j.ijpara.2008.09.008.

[112] P. Cribb, L. Esteban, A. Trochine, J. Girardini, E. Serra, Trypanosoma cruzi TBP shows preference for C/G-rich DNA sequences in vitro, Exp. Parasitol. 124 (2010) 346–349, https://doi.org/10.1016/j.exppara.2009.11.003.

[113] J.-P. Ruan, G.K. Arhin, E. Ullu, C. Tschudi, Functional characterization of a Trypanosoma brucei TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes, Mol. Cell. Biol. 24 (2004) 9610–9618, https://doi.org/10.1128/MCB.24.21.9610-9618.2004.

[114] A. Das, Q. Zhang, J.B. Palenchar, B. Chatterjee, G.A.M. Cross, V. Bellofatto, Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression, Mol. Cell. Biol. 25 (2005) 7314–7322, https://doi.org/10.1128/MCB.25.16.7314-7322.2005.

[115] B. Schimanski, T.N. Nguyen, A. Günzl, Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in Trypanosoma brucei, Mol. Cell. Biol. 25 (2005) 7303–7313, https://doi.org/10.1128/MCB.25.16.7303-7313.2005.

[116] S. Roy, D. Morse, Transcription and maturation of mRNA in dinoflagellates, Microorganisms 1 (2013) 71–99, https://doi.org/10.3390/microorganisms1010071.

[117] L. Li, J.W. Hastings, The structure and organization of the luciferase gene in the photosynthetic dinoflagellate Gonyaulax polyedra, Plant Mol. Biol. 36 (1998) 275–284, https://doi.org/10.1023/a:1005941421474.

[118] D. Guillebault, S. Sasorith, E. Derelle, J.-M. Wurtz, J.-C. Lozano, S. Bingham, L. Tora, H. Moreau, 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 (2002) 40881–40886, https://doi.org/10.1074/jbc.M205624200.

[119] E. Toulza, M.-S. Shin, G. Blanc, S. Audic, M. Laabir, Y. Collos, J.-M. Claverie, D. Grzebyk, Gene expression in proliferating cells of the dinoflagellate

Alexandrium catenella (Dinophyceae), Appl. Environ. Microbiol. 76 (2010) 4521–4529, https://doi.org/10.1128/AEM.02345-09.
[120] S. Vanacova, D.R. Liston, J. Tachezy, P.J. Johnson, Molecular biology of the amitochondriate parasites, giardia intestinalis, entamoeba histolytica and trichomonas vaginalis, Int. J. Parasitol. 33 (2003) 235–255, https://doi.org/10.1016/s0020-7519(02)00267-9.
[121] J. Yee, A. Tang, W.-L. Lau, H. Ritter, D. Delport, M. Page, R.D. Adam, M. Müller, G. Wu, Core histone genes of Giardia intestinalis: genomic organization, promoter structure, and expression, BMC Mol. Biol. 8 (2007) 26, https://doi.org/10.1186/1471-2199-8-26.
[122] T. Bruderer, C. Wehrli, P. Köhler, Cloning and characterization of the gene encoding pyruvate phosphate dikinase from Giardia duodenalis, Mol. Biochem. Parasitol. 77 (1996) 225–233, https://doi.org/10.1016/0166-6851(96)02605-9.
[123] D.V. Holberton, J. Marshall, Analysis of consensus sequence patterns in Giardia cytoskeleton gene promoters, Nucleic Acids Res. 23 (1995) 2945–2953, https://doi.org/10.1093/nar/23.15.2945.
[124] F. Burki, The eukaryotic tree of life from a global phylogenomic perspective, Cold Spring Harb, Perspect. Biol. 6 (2014) a016147, , https://doi.org/10.1101/cshperspect.a016147.
[125] S. Vanácová, J. Tachezy, E. Ullu, C. Tschudi, Unusual diversity in alpha-amanitin sensitivity of RNA polymerases in trichomonads, Mol. Biochem. Parasitol. 115 (2001) 239–247, https://doi.org/10.1016/s0166-6851(01)00294-8.
[126] W. Cai, J. Pei, N.V. Grishin, Reconstruction of ancestral protein sequences and its applications, BMC Evol. Biol. 4 (2004) 33, https://doi.org/10.1186/1471-2148-4-33.
[127] L. Zimmermann, A. Stephens, S.-Z. Nam, D. Rau, J. Kübler, M. Lozajic, F. Gabler, J. Söding, A.N. Lupas, V. Alva, A Completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core, J. Mol. Biol. 430 (2018) 2237–2243, https://doi.org/10.1016/j.jmb.2017.12.007.
[128] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612, https://doi.org/10.1002/jcc.20084.