



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

Doctorado en Ciencias Biomédicas

Instituto de Investigaciones Biomédicas

Caracterización de dos proteínas con similitud a TATA-Binding Protein en  
*Trichomonas vaginalis*.

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctorado en Ciencias Biomédicas

PRESENTA:

M. en C. Lluvia Leticia Rosas Hernández

TUTOR PRINCIPAL:

Dra. María Imelda López Villaseñor

Instituto de Investigaciones Biomédicas

MIEMBROS DEL COMITÉ TUTOR:

Dra. Alicia González Manjarrez

Dr. Roberto Coria Ortega

Instituto de Fisiología Celular

CD. MX., Mayo del 2021



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

DOCTORADO EN CIENCIAS BIOMÉDICAS  
Instituto de Investigaciones Biomédicas

Caracterización de dos proteínas con similitud a las TATA-binding protein en  
*Trichomonas vaginalis*.

Tesis  
Para obtener el grado de:  
Doctor en Ciencias Biomédicas

PRESENTA:  
M. en C. Lluvia Leticia Rosas Hernández

Directora de tesis:  
Dra. María Imelda López Villaseñor

CDMX, mayo del 2021.

*Para Amelia, Bruno y Ernesto.*

## **AGRADECIMIENTOS**

A la Universidad Nacional Autónoma de México, al Instituto de Investigaciones Biomédicas y al Posgrado de Ciencias Biomédicas. Gracias por su compromiso con la ciencia y por ser la base de programas de excelencia para la educación en México.

Al CONACyT por el apoyo con la beca durante mis estudios.

A la Dra. Imelda López Villaseñor por todo su apoyo y por todas sus enseñanzas durante mi estancia en el laboratorio. Imelda, fue un placer haberte tenido como tutora de este proyecto, gracias.

A los miembros de mi comité tutorial, Dra. María Imelda López Villaseñor, Dra. Alicia González Manjarrez y Dr. Roberto Coria Ortega. Comité, gracias por su tiempo, observaciones y apoyo durante todo mi doctorado.

A los miembros del jurado para el examen de grado por la revisión de esta tesis y las observaciones tan oportunas realizadas.

Al Dr. Roberto Hernández Fernández, a la Dra. Ana María Cevallos y a su grupo de trabajo. Gracias por su apoyo y por todas las reflexiones realizadas a mi trabajo.

A la M. en C. Olivia Parra Marín, por la realización de los experimentos complementarios a este trabajo, y a la M. en C. Karla López Pacheco por el apoyo técnico académico en el laboratorio.

A mis compañeros de laboratorio: Olivia Parra, Karla López, Bernardo Franco, Erick Amador, Nataly Morales, Andrea Vizcaino y Ernesto Guerrero. Gracias por aportar excelentes observaciones a mi trabajo en los seminarios, por su apoyo con los experimentos y por su amistad.

A mi familia, gracias por todo su apoyo y por todo el amor que siempre recibo de ustedes.

## RESUMEN.

El parásito protozoario *Trichomonas vaginalis* es un patógeno humano común de uno de los linajes eucariotas más antiguos. A nivel transcripcional, el elemento Inr altamente conservado rodea el sitio de inicio de la transcripción y es reconocido por la proteína de unión a Inr (IBP39) de los genes transcritos de RNA Polimerasa II. No se ha encontrado una caja TATA típica en los promotores de *T. vaginalis*; sin embargo, análisis BLAST del genoma de *T. vaginalis* han demostrado la presencia de dos genes putativos que codifican proteínas de unión a cajas TATA (en este documento denominadas TvTBP). El objetivo de este trabajo es determinar si estas proteínas tienen una función similar a una TBP canónica. Hicimos análisis *in silico* y varios análisis experimentales que muestran que ambos genes TvTBP se expresan en *T. vaginalis* y que ambas proteínas, teóricamente, pueden adoptar la estructura en forma de “silla de montar” distintiva de las TBPs. Por otro lado, una de las proteínas TvTBP interactúa con la proteína TvBRF1 putativa *in vivo*, aunque ninguna de las TvTBPs pudo complementar la función en una levadura que carecía de TBP. Además, en las condiciones utilizadas en este trabajo no se observó unión de las proteínas TvTBPs a los promotores de *T. vaginalis* utilizados en ensayos EMSA. Sin embargo, estudios posteriores con diferentes condiciones experimentales demostraron la habilidad de las TvTBPs para interactuar con el DNA, en una forma no específica de secuencia. Y también se observó la interacción de TvTBP1 con IBP39. En conclusión, estos ensayos sugieren que las TvTBPs podrían formar parte del complejo de inicio de la transcripción en *T. vaginalis*. La importancia de este estudio radica en contribuir con conocimiento para seguir comprendiendo el mecanismo transcripcional del patógeno humano *T. vaginalis*, con el fin de dilucidar este proceso en uno de los organismos que divergió temprano en la evolución del linaje Eucariota.

## **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 surrounds the transcription start site and is recognized by the Inr Binding Protein (IBP39) of RNA Pol II transcribed genes, whereas a typical TATA box is not found in *T. vaginalis* promoters. Nevertheless, BLAST analyses of the *T. vaginalis* genome have shown the presence of two genes encoding putative TATA binding proteins (herein referred to as TvTBPs). The goal of this work is to determine whether these proteins have a similar function as a canonical TBP. We did an *in silico* and several experimental analyses showing that both genes are expressed in *T. vaginalis* and they can, theoretically, adopt the saddle-shaped structure distinctive to TBPs. On the other hand, one of the TvTBPs interacts with the putative TvBRF1 *in vivo*, but none of them were able to complement the function in a TBP-lacking yeast. Additionally, under the conditions used in this work, none of the TvTBPs proteins were able to bind to the promoters tested. However, subsequent studies with different experimental conditions demonstrated the ability of TvTBPs to interact with DNA, in a sequence-independent manner. Furthermore, the interaction of TvTBP1 with IBP39 was also observed. In conclusion, these findings suggest that TvTBPs could be part of the transcription initiation complex in *T. vaginalis*. The importance of this study lies in contributing with knowledge to continue understanding the transcriptional mechanism of the human pathogen *T. vaginalis*, in order to elucidate this process in one of the organisms that diverged early in the evolution of the Eukaryotic lineage.

## ABREVIATURAS

<b>AC</b>	Active Domain – Dominio de activación
<b>BRF1</b>	TFIIB-related factor 1 – Factor relacionado a TFIIB 1
<b>BD</b>	Binding Domain – Dominio de unión a DNA
<b>EMSA</b>	Electrophoretic mobility-shift assays – Ensayo de cambio de movilidad electroforética
<b>ETS</b>	Enfermedad de transmisión sexual
<b>IBP39</b>	Initiator binding protein 39 – Proteína de unión al iniciador, 39
<b>PIC</b>	Preinitiation complex – Complejo de preiniciación.
<b>ScTBP</b>	Proteína TBP canónica de <i>S. cerevisiae</i> .
<b>TBP</b>	TATA binding protein – Proteína de unión a la caja TATA
<b>TBPs</b>	TATA binding proteins – Proteínas de unión a la caja TATA
<b>TFIIB</b>	Transcription factor IIB – Factor de transcripción IIB
<b>TSS</b>	Transcription Start Site – Sitio de inicio de la transcripción
<b>TvTBP</b>	Proteína de <i>T. vaginalis</i> con similitud a las TBPs
<b>TvTBP1</b>	TvTBP, proviene de la secuencia TVAG_291560
<b>TvTBP1*</b>	TvTBP1 modificada
<b>TvTBP2</b>	TvTBP, proviene de la secuencia TVAG_285070



# INDICE

<b>Resumen.</b>	5
<b>Abstract.</b>	6
<b>Abreviaturas.</b>	7
<b>I. Introducción.</b>	12
1. Biología de <i>Trichomonas vaginalis</i> .	12
2. El genoma de <i>Trichomonas vaginalis</i> .	14
3. Modelo de los factores involucrados en el inicio de la transcripción.	16
4. Estructura de la TBP	18
5. Factores relacionados a TBP.	20
6. TBPs en parásitos protozoarios.	20
<b>II. Planteamiento del problema</b>	22
<b>III. Antecedentes.</b>	23
<b>IV. Hipótesis</b>	25
<b>V. Objetivos</b>	26
<b>VI. Resultados</b>	27
1. Análisis bioinformáticos de las proteínas con similitud a TBP en <i>T. vaginalis</i> .	27
2. Los genes que codifican las TvTBPs se expresan en <i>T. vaginalis</i> .	35
3. Las TBPs de <i>T. vaginalis</i> no complementan la función de una levadura mutante en TBP.	37
4. Las TvTBPs no se unieron a los promotores de DNA analizados en este trabajo.	41
5. TvTBP1 es capaz de interactuar con TvTBFRF1 <i>in vivo</i> , a diferencia de TvTBP2.	52
6. Generación de anticuerpos anti TvTBP1.	55

<b>VII. Discusión</b>	57
<b>VIII. Conclusiones.</b>	60
<b>IX. Perspectivas.</b>	61
<b>X. Materiales y Métodos</b>	62
1. Análisis Bioinformáticos	62
2. Cultivo Celular	62
3. RT-PCR	62
4. Complementación en levadura	63
5. Ensayo <i>Western blot</i> , HisProbe-HRP	65
6. Ensayos EMSA (Electrophoretic Mobility Shift Assays).	65
7. Ensayo de Doble Híbrido.	67
8. Preparación de anticuerpo TvTBP1	68
<b>XI. Anexos.</b>	69
<b>XII. Referencias.</b>	75
<b>XIII. Artículo.</b>	80

## LISTA DE FIGURAS

*Figura 1. Representación de una célula de T. vaginalis con sus estructuras más sobresalientes.*

*Figura 2. Árbol representativo de los eucariotes.*

*Figura 3. Modelo del control del inicio de la transcripción en eucariotes.*

*Figura 4. Estructura conservada de la TBP.*

*Figura 5. Aminoácidos importantes en el reconocimiento de la caja TATA.*

*Figura 6. Modelaje de la estructura terciaria de las TvTBPs realizado con el servidor RaptorX.*

*Figura 7. Árbol de identidad realizado con algunas secuencias de aminoácidos de proteínas TBP.*

*Figura 8. Alineamiento y aminoácidos importantes en las TBPs de diversos organismos.*

*Figura 9. Alineamiento de secuencias TBPs, TBP-like y TvTBPs.*

*Figure 10. Red de interacción teórica proteína-proteína de las TvTBPs realizado en la base de datos STRING.*

*Figura 11. Determinación de la expresión de los genes TvTBP por RT-PCR.*

*Figura 12. Secuencias TvTBPs obtenidas a partir de la secuenciación de los productos RT-PCR.*

*Figura 13. Disección de tétradas para el ensayo de complementación de TBP.*

*Figura 14. Análisis de complementación de función, crecimiento de células haploides en un medio mínimo suplementadas con G418.*

*Figura 15. Mutagénesis dirigida TvTBP1.*

*Figura 16. Regiones promotoras usadas como sondas para los ensayos EMSA.*

*Figura 17. Expresión de la proteína recombinante 6xHis-TvTBP1.*

*Figura 18. Expresión de las proteínas recombinantes 6xHis-TvTBP1\*.*

*Figura 19. Expresión de la proteína recombinante 6xHis-TvTBP2.*

*Figura 20. Expresión de la proteína recombinante 6xHis-IBP39.*

*Figura 21. Purificación y expresión de la proteína recombinante 6xHis-SPT15.*

*Figura 22. Análisis para evaluar la interacción de las TvTBPs con el DNA.*

*Figura 23. Análisis para evaluar la interacción de TvTBPs, TvTBP1\* y el DNA con la presencia de TvTFIIB y TvBRF1.*

*Figura 24. Análisis para evaluar la interacción de la proteína ScTBP con el promotor de S. cerevisiae y los promotores de T. vaginalis.*

*Figura 25. Evaluación de la interacción entre las TvTBPs con factores transcripcionales putativos de T vaginalis, por ensayo de doble híbrido.*

*Figura 26. Evaluación de la interacción entre IBP39 con los controles para el ensayo de doble híbrido.*

*Figura 27. Purificación y cuantificación de la proteína recombinante TvTBP1 para generar anticuerpos.*

*Figura 28. Cuantificación relativa de proteínas recombinantes para ensayos EMSA.*

## **LISTA DE TABLAS**

*Tabla 1. Factores generales de transcripción de T. vaginalis.*

*Tabla 2. Matriz de porcentaje de identidad entre proteínas TBP.*

*Tabla 3. Proteínas de la red de interacción TvTBP1 con el proteoma de T. vaginalis.*

*Tabla 4. Descripción de las proteínas recombinantes usadas para los ensayos EMSA.*

# I. INTRODUCCIÓN

## 1. *Biología de Trichomonas vaginalis.*

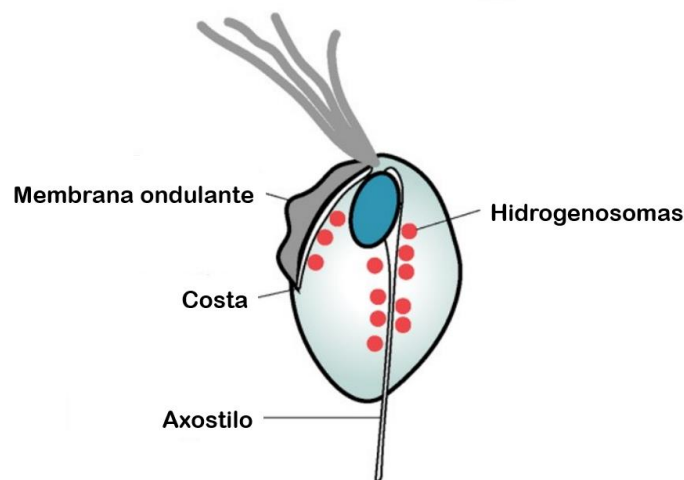
*Trichomonas vaginalis* es un protozoo parásito de humanos que causa la enfermedad de transmisión sexual (ETS), no viral, más común en el mundo. Es más habitual en mujeres; sin embargo, se transmite fácilmente entre parejas sexuales, incluso si la pareja infectada es asintomática. Mientras que algunos pacientes jamás han experimentado síntomas, otras mujeres han reportado secreción vaginal y disuria, así como hombres han reportado disuria y secreción uretral (*Van Gerwen and Muzny, 2019*). *T. vaginalis*, también está ligado al incremento de VIH, posicionado la infección como un problema grande de salud pública (*McClelland y Col., 2007*). La Organización Mundial de la Salud estimó en el 2016, 156 millones de casos de *T. vaginalis* en el mundo, representando casi la mitad de la incidencia mundial de ETS ese año (*Van Gerwen and Muzny, 2019*).

Para establecer la infección del tracto genitourinario conocida como tricomoniasis, el parásito debe mediar el contacto con las células epiteliales del huésped, evadir el ataque de la respuesta inmune y enfrentarse a la microbiota vaginal (*Mercer and Johnson, 2018*). Si bien más del 95% de las infecciones por *T. vaginalis* se pueden eliminar con el medicamento metronidazol 5-nitroimidazol, las cepas resistentes a los medicamentos están en aumento (*Mercer and Johnson, 2018*).

El tamaño de *T. vaginalis* varía de 10 a 20  $\mu\text{m}$ , aproximadamente el mismo tamaño que un linfocito o más pequeño que un neutrófilo (*Garber, 2005*). El diagnóstico de tricomoniasis tradicionalmente dependía de la observación del parásito móvil en muestras vaginales o cervicales y en secreciones uretrales o prostáticas. Sin embargo, el cultivo del parásito en medio Diamond's TYI, con tiempos de incubación de 2 a 7 días con  $\text{CO}_2$ , ha sido el estándar de oro para el diagnóstico de este organismo anaeróbico. La detección de DNA, por PCR, a partir de regiones diversas del genoma también ha sido una técnica utilizada como diagnóstico. Sin embargo, dependiendo de los estudios realizados, la mayor sensibilidad de esta técnica es inferior al 90% y se han obtenido falsos positivos. El examen de *T. vaginalis* se realiza con mayor frecuencia como parte de la investigación del flujo vaginal y vaginitis. La evaluación del pH puede ayudar a diferenciar rápidamente *T. vaginalis* de una infección por levadura. El pH vaginal que normalmente es 4.5, no se ve alterado por la infección por levaduras, pero está elevado en la vaginosis bacteriana y a menudo se eleva por encima pH 6 en tricomoniasis florida (*Garber, 2005*).

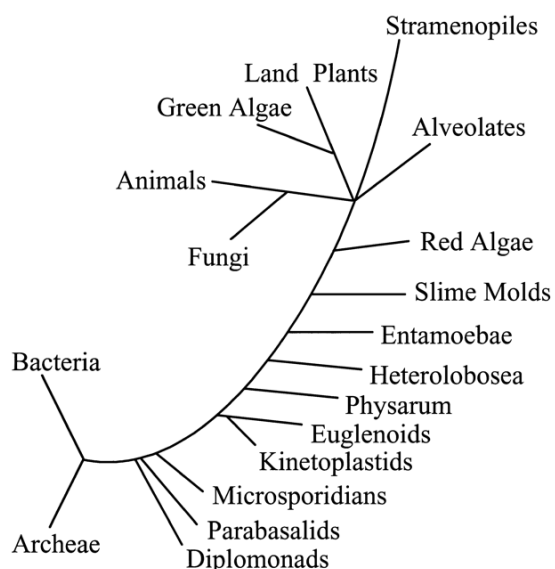
*T. vaginalis* carece de mitocondrias y peroxisomas. En lugar de mitocondria, *T. vaginalis* contiene unos organelos redondeados de 1 a 2  $\mu\text{m}$  de diámetro llamados hidrogenosomas (Figura 1). Estos hidrogenosomas carecen de DNA y producen ATP por fosforilación a nivel de sustrato con producción concomitante de hidrógeno. Cerca de cientos de hidrogenosomas están localizados a lo largo del axóstilo y la costa, estructuras citoesqueléticas únicas, que pueden consumir ATP generado por los hidrogenosomas (Voleman y Dolezal, 2019; Kulda, 1999).

## ***Trichomonas vaginalis***



**Figura 1.** Representación de una célula de *T. vaginalis* con sus estructuras más sobresalientes: 4 flagelos libres y un 5to flagelo que forma una membrana ondulatoria; paralelo a dicha membrana se dispone en el interior de la célula, un haz de microtúbulos denominado costa (estructura no contráctil que le de soporte); axostilo; e hidrogenosomas (Modificado de Voleman y Dolezal, 2019).

*T. vaginalis* pertenece al grupo de los parabasálidos, dentro del reino excavata, protistas flagelados, la mayoría de los cuales son endosimbiontes, comensales o parásitos de animales. Son anaerobios y carecen de mitocondrias. Estudios filogenéticos basados en secuencias de RNA ribosomal, posicionan a *Trichomonas* entre los primeros linajes de ramificación en el árbol evolutivo de los eucariotas. Muchos autores sugieren que se separaron del linaje eucariota principal antes del origen endosimbiótico de las mitocondrias (Gunderson y Col., 1995; Roger y Col., 1996), (Figura 2).



**Figura 2.** Árbol representativo de los eucariotes. Es una versión redibujada del árbol de eucariotas basada en secuencias de la subunidad pequeña del RNA ribosomal. Este árbol muestra la divergencia temprana del grupo de los parabasalidos, protistas flagelados, que incluyen a *T. vaginalis* (modificada de *Dacks and Doolittle, 2001*).

## 2. El genoma de *Trichomonas vaginalis*.

La secuenciación del genoma de *T. vaginalis* ha permitido análisis bioinformáticos extensos. Presenta un genoma grande, inusual para los parásitos, 175 Mb aproximadamente que están distribuidos en seis cromosomas, con un conjunto de 60,000 genes aproximadamente que codifican para proteínas. Se ha identificado la presencia de intrones en únicamente 65 genes (*Carlton y Col., 2007*).

A nivel genómico, los genes codificadores de proteínas, transcritos por la RNA Polimerasa II (RNA Pol II), están compuestos por promotores bipartitos, con una región promotora central conservada y elementos reguladores distales específicos del gen (*Vanacova y Col., 2003; Smith y Johnson, 2011*). Se ha descrito que en la mayoría de los promotores de los genes que codifican proteínas en *T. vaginalis* se encuentra un motivo, alrededor del sitio de inicio de la transcripción similar al elemento iniciador (Inr) descrito en metazoarios cuyo consenso es T/C C A+1 Py T/A. Además, se encontró una proteína nueva, exclusiva en *T. vaginalis* que es capaz de reconocer y unirse a este motivo, nombrada IBP39 (*Liston y Col., 2001*). No se han encontrado elementos TATA en estos promotores que dirijan el inicio de la transcripción (*Liston y Col., 1999*); sin embargo, se han identificado dos genes TBP cuya función se desconoce (TFIID: TBP en la Tabla 1) (*Carlton y Col., 2007*).

Únicamente se han encontrado dos genes en *T. vaginalis*, transcritos por la RNA Polimerasa III, que contienen en la región promotora una caja TATA clásica, el gen U6 snRNA que codifica para un componente del spliceosoma (*Simoes-Barbosa y Col., 2008*) y una variante del gen rRNA 5S (*Torres-Machorro y Col., 2006*).

**Tabla 1.** Factores basales generales de transcripción de *T. vaginalis*. Se indica la presencia o ausencia (+/-) de los factores generales de transcripción comparada con *P. falciparum*, Arquea, Metazoa, levadura y *H. sapiens* como se indica (*Carlton y Col., 2007*).

<b>Basal Transcription Factors</b>	<i>T. vaginalis</i>	<i>P. falciparum</i>	Archaea	Metazoa
TFIID: TBP	+	—	+	+
TFIIB	+	—	+	+
TFIIE: alpha	+	—	+	+
TFIIE: beta	+	—	—	+
TFIIF: alpha	—	+	—	+
TFIIF: beta	+	+	—	+
TFIIH: ERCC3/XPB	+	—	—	+
TFIIH: ERCC2/XPD	+	—	—	+
TFIIH: Cdk7	+	—	—	+
TFIIH: Cyclin H	+	—	—	+
TFIIH: MAT	—	—	—	+
TFIIH: p62	—	+	—	+
TFIIH: p52	+	—	—	+
TFIIH: p44	+	—	—	+
TFIIH: p34	+	—	—	+
TFIIA: large subunit	—	+	—	+
TFIIA: small subunit	—	+	—	+

<b>TBP-Associated Factors</b>	<i>T. vaginalis</i>	<i>P. falciparum</i>	Yeast	<i>H. sapiens</i>
TAF1	+	+	+	+
TAF1 L	+	+		+
TAF2	+	—	+	+
TAF3	+	—	+	+
TAF4	+	—	+	+
TAF5	+	—	+	+
TAF6	+	—	+	+
TAF7	+	+	+	+
TAF8	—	—	+	+
TAF9	+	—	+	+
TAF10	+	+	+	+
TAF11	—	—	+	+
TAF12	+	—	+	+
TAF13	—	—	+	+
TAF14	—	—	+	
TAF15	—	—		+
MOT1/BTAF1	+	—	+	+





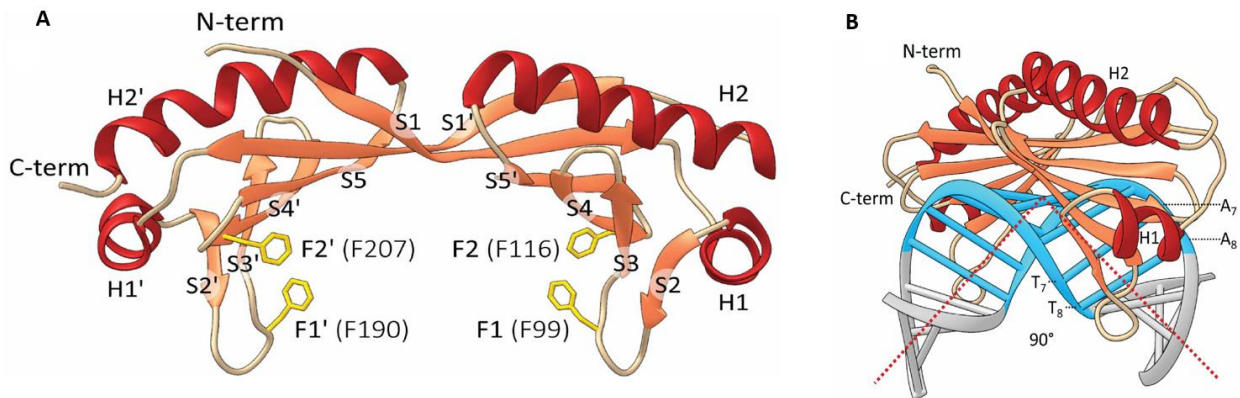
En conclusión, los mecanismos generales de regulación transcripcional muestran similitudes en organismos de los distintos dominios de la vida y las principales diferencias se relacionan con el ajuste fino y el comportamiento dinámico de las estructuras de cromatina y de los complejos TBP / TFIID (Koster y Col., 2015).

TBP es un factor importante para la transcripción eucariota y arqueobacteriana (Revisiones: White and Jackson, 1992; Hernandez, 1993; Vanini y Col., 2013). Clásicamente, en los promotores que contienen una caja TATA, TBP se une a la secuencia TATA (TATAa/tAa/t) en el surco menor y distorsiona drásticamente la estructura del DNA (Kim y Burley, 1994; Juo y Col., 1996). Esto da como resultado el reclutamiento del PIC hacia el promotor, que posiciona a la RNA Pol II sobre el TSS (Hahn, 2004). Muchos elementos en *cis* se han descrito en los promotores, exponiendo la alta complejidad y variabilidad en el proceso de formación del complejo de preiniciación (Danino y Col., 2015).

La proteína TBP se identificó por primera vez como un polipéptido que posee actividad de unión a la caja TATA, subunidad de TFIID, que sugirió una función exclusiva en la transcripción de promotores de RNA Pol II que contienen TATA (Hernández, 1993; Thomas y Col., 2006). Sin embargo, TBP está involucrada en los complejos de iniciación formados en promotores que contienen TATA y también en los promotores que carecen de ella (Hernández, 1993). TBP es un componente crucial de los tres complejos diferentes de iniciación de la transcripción: SL1, TFIID y TFIIB, y está involucrado en la transcripción por las RNA polimerasas I, II y III, respectivamente. Por lo tanto, TBP se recluta para los promotores a través de diferentes estrategias que implican la unión directa al DNA en el caso de los promotores que contienen TATA, o interacciones proteína-proteína en el caso de los promotores sin caja TATA (Rigby, 1993, Hahn, 2004 y Vannini, 2013). Típicamente, TBP interacciona con los factores de transcripción TFIIA y TFIIB. Las tres RNA polimerasas requieren factores relacionados, estructural y funcionalmente, a TFIIB para el inicio de la transcripción: TFIIB, factor relacionado a TFIIB 1 (BRF1) y Rrn7 (o TAF1B en humanos). La TBP está asociada con estos factores similares a TFIIB en los tres complejos de iniciación de la transcripción, que son necesarios para el reclutamiento de polimerasa para el promotor (Vannini, 2012). Cabe destacar que en el genoma de *T. vaginalis* no se encontraron proteínas relacionadas a TFIIA, pero si a TFIIB (Carlton y Col., 2007).

#### 4. Estructura de la TBP

La proteína TBP consiste en una región carboxilo terminal conservada de 180 aminoácidos y una extensión N-terminal que varía en longitud, secuencia y funcionalidad. La región conservada (el “core”) está formada por dos dominios simétricos, cada uno consiste en 60 residuos aproximadamente, arreglados en 5 cadenas antiparalelas láminas  $\beta$ . Estos dos dominios forman una superficie cóncava, en forma de “silla de montar”, e hidrofóbica que se une fuertemente al surco menor de la doble hélice de DNA. En adición, dos  $\alpha$ -hélices se forman en la superficie convexa de la “silla de montar” y tienen el potencial de mediar interacciones proteína-proteína (Figura 4) (Kramm, 2019).



**Figura 4.** Estructura conservada de la TBP. A) Estructura cristalográfica de rayos X de la TBP de *Saccharomyces cerevisiae*; se muestran las láminas  $\beta$  hidrofobas de cinco cadenas (en naranja, S1–5, S1’–5’), las hélices  $\alpha$  (en rojo, H1–2, H1’–2’) y las fenilalaninas conservadas (en amarillo, F1–2, F1’–2’), La extensión N-terminal se truncó para la cristalización. B) Flexión del promotor de DNA por la TBP unida a la caja TATA (en azul). El ángulo de flexión del DNA (gris) se traza con una línea punteada en rojo (modificado de Kramm, 2019).

La región *core* contienen el dominio de la superfamilia de las TBP\_TLF (cd00652) que está conservado en la familia de las proteínas TBP y las proteínas TBP-like, encontrada en el extremo carboxilo de las proteínas. Dentro de esta región se ubican las superficies de interacción con el DNA y con las proteínas TFIIA, TFIIIB y NC2 (Cofactor negativo 2) (Altschul y Col., 1990).

TBP es una proteína arquitectónica: al unirse al surco menor de la caja TATA, TBP distorsiona drásticamente el DNA. Abre el surco menor e inserta un par de residuos de fenilalanina en los primeros y últimos pares de bases (pb) de los 8 pb de la secuencia de reconocimiento (Figura 5). La inserción de los residuos induce al DNA a desenrollarse y doblarse aproximadamente 90°. La formación del PIC es facilitado en gran medida debido a la deformación del DNA tras la unión de TBP, por lo que la preservación de esta característica estructural es de gran importancia en el proceso de transcripción. Análisis de la secuencia revela que TBP no se une a regiones que contienen secuencias ricas en G-C (*Juo y Col., 1996, Millán-Pacheco y Col., 2009*).

Las posiciones de unión más fuertes corresponden al centro de la caja TATA, como se espera por la presencia de un par de asparaginas (N59, N159) y treoninas (T124, T215) que se unen en enlaces de hidrógeno con el borde de ranura menor de las bases (*Juo y Col., 1996*) (Figura 5).



**Figura 5.** Aminoácidos importantes en el reconocimiento de la caja TATA. Los números corresponden a la TBP de levadura (Modificado de *Millán-Pacheco y Col., 2009*).

## **5. Factores relacionados a TBP.**

Muchos eucariotas metazoos contienen múltiples parálogos de TBP (factores relacionados con TBP, TRF) con funciones específicas y distintas *in vivo* (Revisiones Davidson, 2003; Akhtar y Col., 2011). TRF1, específico para insectos, puede reemplazar la TBP en la transcripción, unirse a la caja TATA junto con TFIIA y TFIIB (Hansen y Col., 1997) y asociarse con BRF1 en *Drosophila melanogaster*. (Takada y Col., 2000). TRF2, específico de metazoos, también conocido como TLP (proteína similar a TBP), TLF (factor similar a TBP) o TBPL1 (similar a TBP 1) es el parólogo más distante. A diferencia de TBP, TRF2 carece de la capacidad de unirse a la caja TATA debido a la pérdida de tres de las cuatro fenilalaninas requeridas para el reconocimiento de la caja TATA (Dantonel y Col., 1999, Duttke y Col., 2014). Se ha propuesto que TFR2, que ha evolucionado a partir del gen TBP duplicado y ha mutado las tres Phe a lo largo del tiempo, retuvo otras propiedades clave de TBP como las interacciones con TFIIA y TFIIB (Duttke y Col., 2014). TRF2 puede dirigir a la RNA Pol II a subconjuntos de promotores sin TATA (Zehavi y Col., 2015). TRF3 es específico de vertebrados, también conocido como TBP2 (proteína de unión a TATA 2) o TBPL2 (proteína similar a TBP 2) es el más relacionado con TBP; se une a las cajas TATA y se asocia con TFIIA y TFIIB (Bartfai y Col., 2004). La TBP2 es abundante en los ovocitos y puede reemplazar la TBP para la transcripción en este tipo de células. A diferencia de la TBP, el TRF3 tras la maduración meiótica se degrada después del cierre global de la transcripción y los niveles residuales persisten en el óvulo y el embrión temprano (Jallow y Col., 2004; Akhtar y Col., 2011).

## **6. TBPs en parásitos protozoarios.**

Los protistas son organismos que representan la mayor parte de la diversidad eucariota y son a menudo considerados como eucariotas ancestrales. Uno de ellos es el "supergrupo" Excavata, que comprende flagelados unicelulares de diversos estilos de vida y contiene especies de importancia médica, como *Trichomonas*, *Giardia*, *Naegleria*, *Trypanosoma* y *Leishmania*. Hay poca investigación sobre el mecanismo molecular de regulación de la expresión génica en los parásitos protozoarios; sin embargo, recientemente se ha generado información importante sobre sus componentes transcripcionales. Las TBP de protistas están llamando la atención debido a su interesante divergencia (Koster y Col., 2015).

Algunas de estas TBP no retienen los cuatro residuos de fenilalanina críticos para la interacción DNA-TBP, ni los dos residuos de asparaginas ubicados en el centro del complejo DNA-TBP y conservados en las proteínas de la familia TBPs (Juo y Col., 1996). Al igual que en *T. vaginalis*, no se han identificado elementos TATA en ningún promotor de genes de tripanosomas (Gomez y Col., 2010). La TBP de *T. brucei*, llamada también factor 4 relacionado con TBP (TbTRF4) (Ruan y Col., 2004), y su homólogo en *T. cruzi* (TcTBP) son reclutados hacia los promotores del gen *splice-leader* RNA (SL RNA). TbTRF4 es esencial para la transcripción y su eliminación reveló una participación esencial en la transcripción de la RNA Pol I, II y III (Ruan y Col., 2004). Sin embargo, ni el ensayo de un híbrido ni los ensayos de cambio de movilidad electroforética (EMSA) con proteína recombinante pudieron demostrar una interacción directa de TcTBP con el DNA del promotor del gen SL RNA (Cribb and Serra, 2009). No obstante, se encontró que *in vitro* TcTBP muestra preferencia por regiones de DNA ricas en secuencias C/G (Cribb y Col. 2010). El homólogo de *L. tarentolae* TBP (LtTBP) es un componente del complejo *Small Nuclear Activating Protein Complex* (SNAP<sub>50</sub>) que interactúa con el elemento sin caja TATA -60 del promotor del gen SL RNA. *L. major* TBP (LmTBP) y SNAP<sub>50</sub> parecen unirse en complejo a todos los promotores de RNA Pol II y III (Thomas y Col., 2006 y Thomas y Col., 2009). Hasta ahora, no se conocen ensayos que indiquen alguna actividad en la transcripción del altamente divergente TBP de *G. lamblia* TBP (GITBP) (Best y Col., 2004). En contraste, los elementos reguladores en *cis* de *E. histolytica* contienen cajas TATA. Además, contiene dos homólogos de TBP (EhTBP y EhTRF1) más similares a una TBP canónica y, como es de esperar, se unen a su caja TATTTAAA *in vitro*, además de variantes de TATA y a cajas GAAC (Castañon-Sanchez y Col., 2010; Narayanasamy y Col., 2018).

## II. Planteamiento del problema

En el genoma de *T. vaginalis* no se han observado cajas TATA canónicas; sin embargo, se encuentran dos genes putativos cuyas secuencias proteicas presentan el dominio de la superfamilia de las TBPs y teóricamente podrían adoptar la estructura característica de “silla de montar” de las TBPs, indicando que estas proteínas tienen similitud a las TBPs canónicas, como ya lo han mencionado otros autores (*Castañon-Sanchez y Col., 2010; Smith and Johson, 2011; Franco 2012, Tesis doctoral; Koster, 2015*). Sin embargo, no existen estudios que hayan demostrado el papel funcional de estas dos proteínas en la expresión génica de *T. vaginalis*. De tal manera que el objetivo del presente trabajo es determinar si las TBPs putativas de *T. vaginalis* tienen una función similar a las TBPs canónicas y si pueden interactuar con factores transcripcionales putativos presentes en este parásito.

### III. Antecedentes

La regulación de la expresión de genes es un proceso fundamental en las células, que está dirigida por una maquinaria de transcripción compleja en eucariotas. Los complejos transcripcionales y los elementos regulatorios en el DNA han sido estudiados principalmente en sistemas complejos como hongos, plantas y animales; y poco se conoce en organismos eucariotas de una sola célula. Los mecanismos transcripcionales en este tipo de organismos, por ejemplo *T. vaginalis*, permanecen poco estudiados. *T. vaginalis* ofrece un sistema alternativo interesante y emocionante para estudiar la expresión de genes eucarióticos, ya que diverge de los sistemas de metazoarios y hongos que han recibido mayor atención (Johnson y Smith, 2011).

El grupo de trabajo de la Dra. Imelda López Villaseñor está interesado en estudiar la biología básica de *T. vaginalis*, enfocado a la regulación de la expresión genética en este organismo, en particular la transcripción por las RNA polimerasas y las señales para el procesamiento y poliadenilación de los RNA mensajeros. Como parte del trabajo de doctorado del Dr. Bernardo Franco, que se centró en el estudio de la transcripción del cistron ribosomal de *T. vaginalis* transcrito por la RNA Polimerasa I, se comenzaron los estudios de dos genes, TVAG\_291560 (TvTBP1) y TVAG\_285070 (TvTBP2), con dominios específicos de las proteínas TBP. Ambos genes contienen un posible elemento Inr en la región 5' UTR, y posibles señales para llevar a cabo el corte y procesamiento del mRNA una vez transcrito el gen, en la región 3'UTR; lo que sugiere se trata de genes que se expresan. Se demostró la transcripción del gen TvTBP1 en *T. vaginalis* por medio de la técnica RT-PCR (Franco 2012, Tesis doctoral).

Con el fin de iniciar el estudio de TBP en la transcripción de la RNA Polimerasa I en *T. vaginalis*, el Dr. Bernardo Franco generó una proteína recombinante de 25 kDa a partir de la clonación del marco de lectura abierto del gen TVAG\_291560, que tiene mayor % de identidad con la TBP de *S. cerevisiae* que el gen TVAG\_285070. En la construcción, este gen está fusionado con una etiqueta de seis histidinas en el extremo amino terminal que servirá para estudios posteriores de identificación y purificación. Está clonada en el vector de expresión pQE30 (QIAGEN) cuyo promotor es regulable y depende de la adición de IPTG para su expresión. Se transformó en diferentes cepas de *Escherichia coli*, después de analizar la cantidad de proteína obtenida en geles de SDS-PAGE se observó la mayor expresión en la cepa M15 (pREP4) (Franco 2012, Tesis doctoral).



Asimismo, el Dr. Bernardo Franco realizó un análisis estructural teórico, utilizando el programa Phyre, modelando las subunidades y factores transcripcionales relacionados con la RNA Pol I, demostrando que las TBPs putativas de *T. vaginalis* pueden adoptar la forma de “silla de montar” típica de las TBPs. Por último, realizó un alineamiento de varias secuencias de aminoácidos de proteínas TBPs, destacando la diversidad en la región amino terminal y observando en un cladograma que las TBPs putativas de *T. vaginalis* están alejadas filogenéticamente del resto de las TBPs (*Franco 2012, Tesis doctoral*).

Con el fin de continuar el estudio de las proteínas TBP hipotéticas en *T. vaginalis*, al inicio del trabajo de investigación que aquí se reporta se llevó a cabo un ensayo de *Western blot* que reveló la presencia de la etiqueta de histidinas en la proteína recombinante TvTBP1, herramienta que se utilizará para ensayos posteriores.

Con el presente trabajo pretendemos contribuir con conocimiento para comprender mejor un mecanismo básico para la sobrevivencia y proliferación del parásito, el mecanismo transcripcional de *T. vaginalis*. Los crecientes casos de resistencia del parásito al metronidazol (*Cudmore y Col. 2004, Graves y Col. 2020*), fármaco habitual de tratamiento, ponen de manifiesto la necesidad de continuar estudiando a este agente etiológico que es el responsable de la infección más común de transmisión sexual (no viral) en todo el mundo (*World Health Organization, 2011*).

#### **IV. Hipótesis**

Durante el proceso de transcripción de *Trichomonas vaginalis* las proteínas TBP participan en la formación del complejo de preiniciación (PIC) reconociendo regiones en el DNA y reclutando a la maquinaria transcripcional.

## V. Objetivos

### General:

Caracterizar a las proteínas con similitud a TBP en *Trichomonas vaginalis* y analizar su función.

### Específicos:

- Analizar la función de las proteínas con similitud a TBP en *T. vaginalis*, evaluando la capacidad de complementar una cepa de levadura mutante en TBP.
- Analizar la interacción de ambas proteínas con similitud a TBP con promotores reconocidos por las tres polimerasas de *T. vaginalis*.
- Determinar qué factores putativos, de la maquinaria de transcripción en *T. vaginalis*, interactúan con las proteínas con similitud a TBP.

## VI. RESULTADOS Y DISCUSIÓN

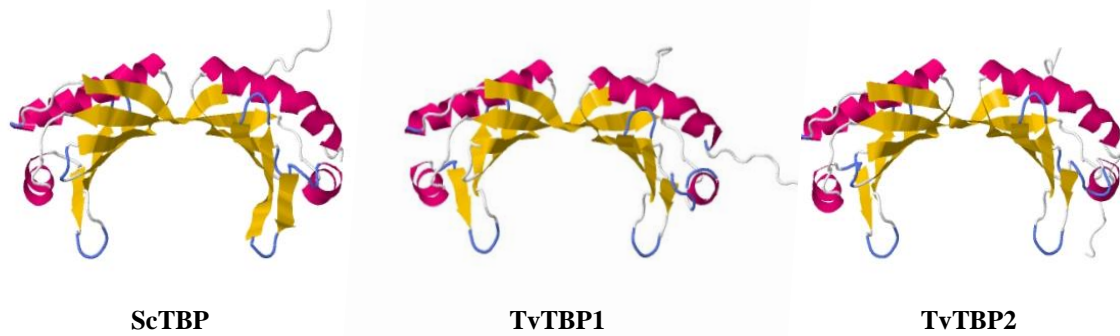
### 1. Análisis bioinformáticos de las proteínas con similitud a TBP en *T. vaginalis*.

#### 1.1 Las TvTBPs tienen dominios y una estructura teórica similar a una TBP canónica.

En *T. vaginalis* hay dos genes identificados como *transcription initiation factor tfiid putative* (TFIID:TBP) anotados como TVAG\_291560 y TVAG\_285070 (Carlton y Col., 2007), en este trabajo denominados TvTBP1 y TvTBP2 respectivamente o TvTBPs para ambos. Para validar la anotación, en un trabajo previo se reportó: 1) un análisis de alineamiento con las TBPs de otros organismos y las secuencias de aminoácidos putativas de las TvTBPs, resaltando regiones conservadas hacia el extremo carboxilo terminal donde se lleva a cabo la unión con el DNA. 2) Que las TvTBPs teóricamente pueden adoptar la estructura de “silla de montar” típica de las TvTBPs (Franco 2012, Tesis doctoral).

En el presente trabajo realizamos análisis bioinformáticos con ambos genes. Con la secuencia proteica de TBP de *Homo sapiens* y *Saccharomyces cerevisiae* realizamos un BLAST en el genoma de *T. vaginalis*. Ambas búsquedas resultaron en la identificación de únicamente dos proteínas putativas en *T. vaginalis*, que corresponden a aquellas identificadas como *transcription initiation factor TFIID putative* en la base de datos de Trichomonas (*TrichDB, Trichomonas Genomics Resource, Aurrecoechea C. y Col., 2009*), codificadas por los genes antes mencionados. Ambas proteínas putativas tienen en su extremo carboxilo terminal un dominio de la súper familia de las proteínas TBP, el dominio PF00352 de Pfam (Finn y Col., 2014). También se identificaron regiones teóricas para la unión con: DNA, con el cofactor negativo 2 (NC2) y con los factores transcripcionales TFIIA y TFIIB, aunque aparentemente en *T. vaginalis* no hay un ortólogo de TFIIA.

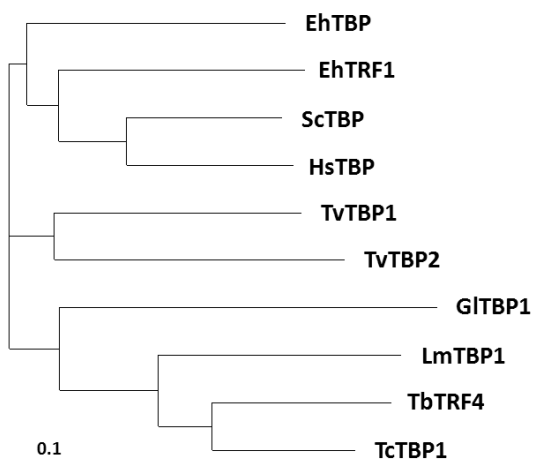
Asimismo, para corroborar su identidad como proteínas TBP realizamos un modelado de la estructura terciaria de las TvTBPs (Figura 6), realizado con el servidor *RaptorX* (Kallberg y Col., 2012). Como referencia de comparación incluimos la TBP de *S. cerevisiae* (denominada en este trabajo como ScTBP). Las estructuras obtenidas muestran que las proteínas adoptan claramente la estructura distintiva de “silla de montar” típica de las TBPs (Figura 4), como lo demostró anteriormente el Dr. Bernardo Franco con los programas *Phyre* y *Phyre2* (Franco 2012, Tesis doctoral).



**Figura 6.** Modelaje de la estructura terciaria de las TvTBPs realizado con el servidor RaptorX (Kallberg y Col., 2012). Teóricamente, las TvTBPs conservan la estructura de “silla de montar” típica de las TBPs. Como ejemplo y comparación, se modeló la proteína TBP de *S. cerevisiae* (denominada ScTBP, en este trabajo) que representa una TBP canónica. Las proteínas están coloreadas por estructura secundaria: en rosa las hélices  $\alpha$ , en amarillo las láminas  $\beta$  y en azul las vueltas.

## 1.2 Las TvTBPs se agrupan independientemente a otras TBPs.

Con la finalidad de inferir la relación evolutiva de las proteínas putativas con similitud a TBP de *T. vaginalis*, realizamos un árbol filogenético de identidad con TBPs de protozoarios parásitos, levadura y humano (Figura 7). El árbol mostró que las TBPs se agrupan en tres grupos: 1) las TBPs de *Homo sapiens*, *Saccharomyces cerevisiae* y *Entamoeba histolytica*; 2) las TBPs de protozoarios: *Giardia lamblia*, *Leishmania major*, *Trypanosoma brucei* y *Trypanosoma cruzi*; y 3) las TBPs de *T. vaginalis* de manera independiente. Dato interesante que mostró anteriormente Castañon-Sanchez, en un análisis con diversas TBPs de arqueas, eucariotas y otras proteínas similares a TBP (Castañon-Sanchez y Col., 2010), así como el Dr. Bernardo Franco en un análisis con TvTBP1, proteínas TBPs de mamíferos, entre otras (Franco 2012, Tesis doctoral).



**Figura 7.** Árbol de identidad realizado con algunas secuencias de aminoácidos de proteínas TBP. EhTBP: TBP de *Entamoeba histolytica* (P52653); EhTRF1: Factor relacionado a TBP1 de *Entamoeba histolytica* (EU052294); ScTBP: *Saccharomyces cerevisiae* (P13393); HsTBP: *Homo sapiens* (P20226); TvTBP1: TBP de *Trichomonas vaginalis* (EAY17204.1); TvTBP2 TBP de *Trichomonas vaginalis* (EAX87882.1); GITBP: TBP de *Giardia lamblia* (Q86GG9); LmTBP: TBP de *Leishmania major* (Q4QD71); TbTRF4: Factor relacionado a TBP *Trypanosoma brucei* (XP\_828065.1); TcTBP: TBP de *Trypanosoma cruzi* (Q4DK05). El alineamiento de las secuencias de realizo con el programa clustal W (Sievers y Col., 2011) y la edición del árbol con el programa TreeView (Win32) 1.6.6 (Page Roderic DM, 1996).

Asimismo, comparamos las secuencias de aminoácidos de las TBPs. El porcentaje de identidad entre TvTBP1 y TvTBP2 es de 41.59 %. El porcentaje de identidad entre las TvTBPs y otras TBPs de *G. lamblia*, *L. major*, *T. cruzi*, *T. brucei*, *E. histolytica*, *H. sapiens* y *S. cerevisiae* varia del 22.0 al 42.5 % de identidad, en orden ascendente. TvTBP1 tienen mayor porcentaje de identidad con la TBP de *S. cerevisiae*, seguido de la TvTBP2; TvTBP2 tiene mayor porcentaje de identidad con la TvTBP1, seguido de la TBP de *S. cerevisiae* (Tabla 2).

**Tabla 2.** Matriz de porcentaje de identidad entre proteínas TBP: *Giardia lamblia* (GITBP) (Q86GG9); *Leishmania major* (LmTBP) (Q4QD71); *Trypanosoma brucei* TBP-related factor (TbTRF4) (Q387R9); *Trypanosoma cruzi* (TcTBP)(Q4DK05); *Entamoeba histolytica* (EhTBP) (P52653); *E. histolytica* TBP-related factor 1 (EhTRF1) (EU052294); *Saccharomyces cerevisiae* (ScTBP) (P13393); *Homo sapiens* (HsTBP) (P20226); *Trichomonas vaginalis* (TvTBP1) (A2DQT7); *T. vaginalis* (TvTBP2) (A2G4NT). Creado con ClustalW (Sievers y Col., 2011). Sombreado se muestra el porcentaje de identidad de las TvTBPs con la ScTBP.

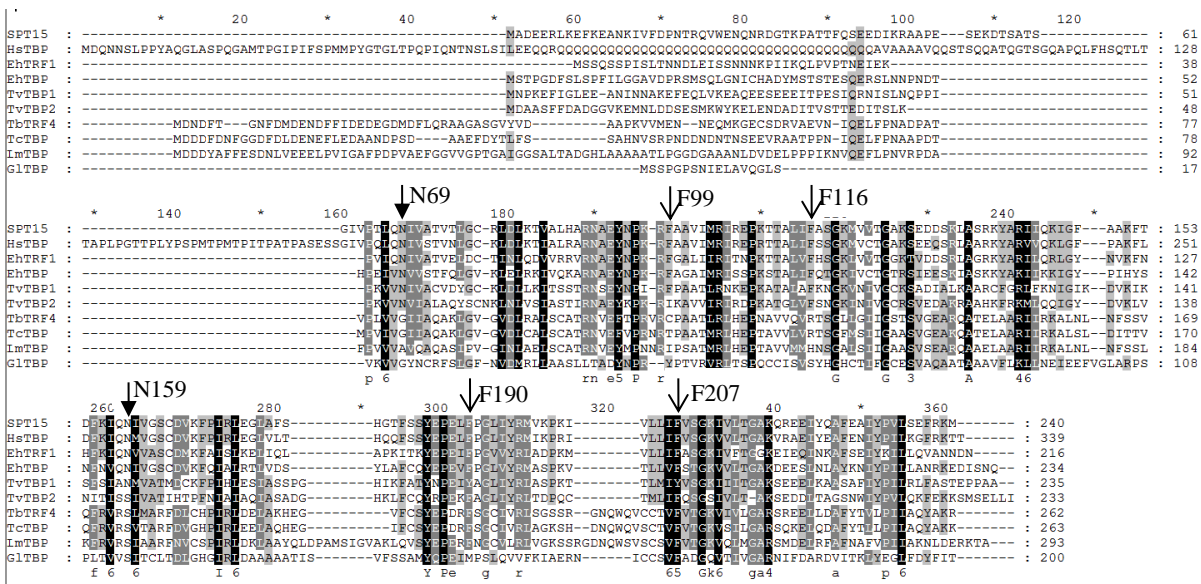
	GlTBP	LmTBP	TbTBP	TcTBP	EhTBP	EhTRF1	SPT15	HsTBP	TvTBP1	TvTBP2
1: GlTBP	100.00	21.94	25.39	26.42	21.88	17.65	20.86	20.32	22.56	22.40
2: LmTBP	21.94	100.00	48.09	51.33	27.80	23.47	26.05	23.62	22.77	24.20
3: TbTBP	25.39	48.09	100.00	65.50	28.10	26.24	27.95	25.00	28.64	25.48
4: TcTBP	26.42	51.33	65.50	100.00	31.13	28.14	31.14	27.80	29.30	25.71
5: EhTBP	21.88	27.80	28.10	31.13	100.00	41.35	49.01	46.12	34.82	35.16
6: EhTRF1	17.65	23.47	26.24	28.14	41.35	100.00	48.57	47.66	37.62	33.67
7: SPT15	20.86	26.05	27.95	31.14	49.01	48.57	100.00	64.44	42.64	39.27
8: HsTBP	20.32	23.62	25.00	27.80	46.12	47.66	64.44	100.00	40.48	34.30
9: TvTBP1	22.56	22.77	28.64	29.30	34.82	37.62	42.64	40.48	100.00	41.59
10: TvTBP2	22.40	24.20	25.48	25.71	35.16	33.67	39.27	34.30	41.59	100.00

### 1.3 Los aminoácidos involucrados en el reconocimiento e interacción con las cajas TATA no están conservados en las TvTBPs.

Realizamos un análisis detallado para identificar en las TBPs, arriba mencionadas, los aminoácidos involucrados en el reconocimiento e interacción de las cajas TATA. El extremo carboxilo terminal de las proteínas es el más conservado entre las TBPs (Figura 8a). Las proteínas que conservan los aminoácidos relevantes para la interacción de las cajas TATA son las TBPs de *S. cerevisiae*, *H. sapiens* y *E. histolytica*. Las TBPs de *T. vaginalis* conservan algunos aminoácidos relevantes, mientras que las TBPs de tripanosomátidos y de *G. lamblia* son las que presentan más substituciones en estos aminoácidos (Figura 8).

La TvTBP1 conserva los dos residuos de asparaginas, importantes para la interacción con las cajas TATA, que se localizan en el centro de las proteínas. De las cuatro fenilalaninas también importantes para la interacción, TvTBP1 conserva dos de ellas, mientras que las otras dos están substituidas por tirosina. TvTBP2 posee solamente una de las dos asparaginas conservadas, la segunda esta substituida por serina al igual que en las TBPs de Tripanosomátidos y de *G. lamblia*. Tres de las cuatro fenilalaninas están conservadas en TvTBP2, con solo una substituida por isoleucina, como en *L. major* (Figura 8).

a)



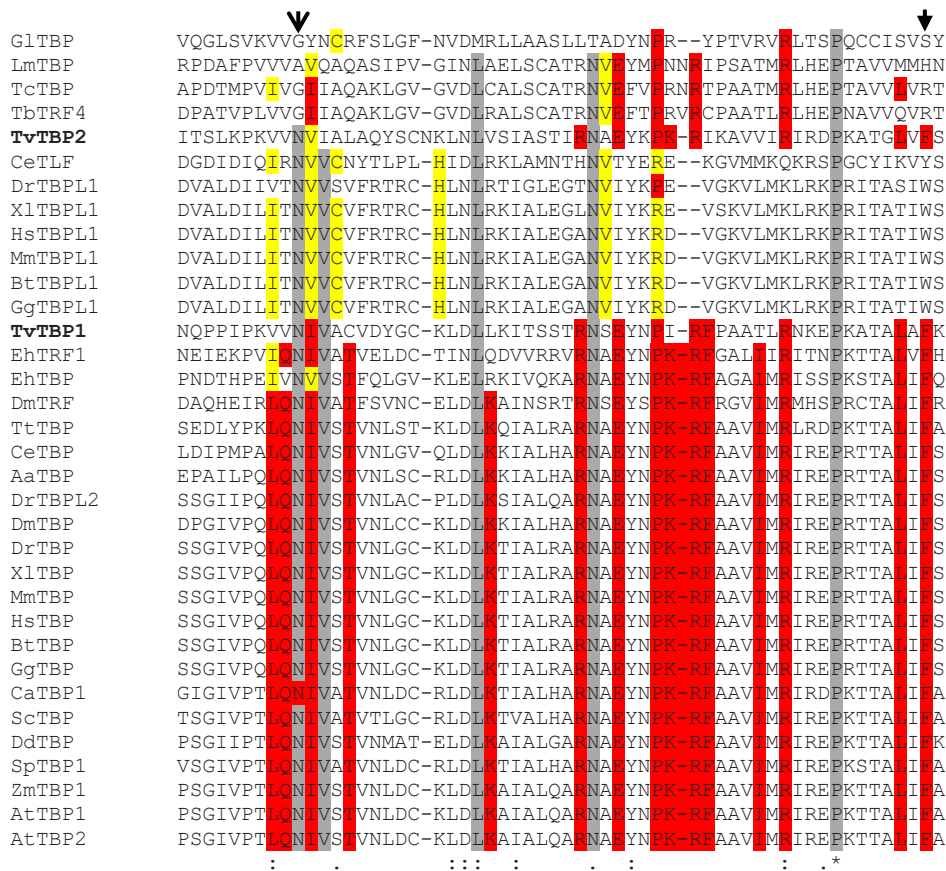
b)

	N69	V71	F99	F116	V122	T124	N159	V161	F190	V203	L205	F207	L214	T215
ScTBP	•	•	•	•	•	•	•	•	•	•	•	•	•	•
HsTBP	•	•	•	•	•	•	•	•	•	•	•	•	•	•
EhTRF1	•	•	•	•	•	•	•	•	•	•	•	•	F	•
EhTBP	•	•	•	•	•	•	•	•	•	T	•	•	•	•
TvTBP1	•	•	•	•	N	V	•	•	Y	T	M	Y	I	•
TvTBP2	•	I	I	•	N	V	S	•	•	T	•	•	•	•
TvTBP1*	•	•	•	•	•	•	•	•	•	•	•	•	I	•
TbTRF4	G	I	C	R	G	I	S	M	•	C	T	•	V	L
TcTBP	G	I	T	R	S	I	S	T	•	S	T	•	I	L
LmTBP	A	Q	I	H	S	I	S	A	•	S	S	•	•	M
GITBP	G	N	Y	S	T	F	S	T	M	C	S	•	I	V

**Figura 8.** Alineamiento y aminoácidos importantes en las TBPs de diversos organismos. **a)** Alineamiento de las secuencias proteicas de las TBP antes mencionadas en la tabla 1 con el programa GeneDoc (Nicholas KB, Nicholas HB Jr., 1997). Los números corresponden a la ScTBP (Millán-Pacheco y Col., 2009). Se muestran indicados con una flecha los aminoácidos más relevantes para la interacción con la caja TATA, dos asparaginas y cuatro fenilalaninas. **b)** Resumen de la identificación de algunos aminoácidos relevantes para la interacción con la caja TATA en varias proteínas TBP. Los puntos indican identidad en el aminoácido. Los cuadros sombreados indican conservación, los colores oscuros muestran mayor conservación.

## 1.4 Los aminoácidos propios de proteínas TBP-like no están conservados en las TvTBPs.

Realizamos un segundo alineamiento incluyendo proteínas TBP-like, con el fin de analizar si las TvTBPs conservan aminoácidos propios de estas proteínas (Figura 9). Como se observa, las TvTBPs tienen la mayoría de los aminoácidos conservados de la familia de las TBPs en general (en gris), que incluyen también a las TBP-like. Las TvTBPs no contienen la mayoría de los aminoácidos propios de las TBPs-like (en amarillo). Las TvTBPs comparten más aminoácidos propios de las TBPs (en rojo).





▼

G1TBP HCHCTIFGCE SVAQAATAAAVFLKLLNEIEEFV-GLARPSPLTVV SITCLT LGHGIRLD  
LmTBP SGALSI IGAASVSEARQAEEAARI IRKALNL---NFSSLKFRVRS IAARFNVCS FIRLD  
TcTBP SGFM SII GAASVGEAKQATELAARI IRKALS L---DITTVQFRVRS V TARF VGHFIRLE  
TbTRF4 SGLLGI I GSTVSGEARQATELAARI IRKALNL---NFSSVQFRVRS LMARF LCHFIRLD  
**TvTBP2** NGKINIVGCRS VEDAKRAAH FRKMLQQIGYD----VKLVNITISSI ATIH TPFNIAIA  
CeTLF SGKVYIVGCRS EADCKRAARS IARHVQVRVMGKTKERVS IRNRYVNNV L ATCLRFPGIKIE  
DrTBPL1 SGKIICTGATSEEEAKLGARRLARCLQKMGFK----VRFSDFKVVNV L AVCSMPFQIRLI  
XlTBPL1 SGKIICTGATSEEEAKVGARRLARS LQKLGFO----VKFTEFKVVNV L AVCTMPFEIRLN  
HsTBPL1 SGKIICTGATSEEEAKFGARRLARS LQKLGFO----VIFTFDKVVNV L AVCNMPFEIRLP  
MmTBPL1 SGKIICTGATSEEEAKFGARRLARS LQKLGFO----VIFTFDKVVNV L AVCNMPFEIRLP  
BtTBPL1 SGKIICTGATSEEEAKFGARRLARS LQKLGFO----VIFTFDKVVNV L AVCNMPFEIRLP  
GgTBPL1 SGKVICTGATSEEEAKFGARRLARCLQKLGFO----VIFTFDKVVNV L AVCNMPFEIRLP  
**TvTBP1** NGKVNIVGCKS ADIALKAARCFGR LFKNIGIK---DVKIKSFSIANM ATMCKF IHLE  
EhTRF1 SGKLVVTVGKTVDDSR LAGRYARI LQRLGYN---VKFNHFKIONV IASC MKFAISLK  
EhTBP TGKIVCTGTRS IEESKIASKYAKI IKKIGYP---IHYSNFVNI NGSC VVKFQIALR  
DmTRF TGKVICTGARN IEADIGSR FARI LQKLGFP---VKFMEYK LONIVATV LRF FIRLE  
TtTBP SGKMVCTGAKTEEDSNRAAR YAKI IQKIGFP---VQFKDFKIONI GSTVVKFEINLD  
CeTBP SGKMVCTGAKSE EASRLAAR YARIVQKLGFO----AKFTEFMVONM NGSC VRF F IQLE  
AaTBP SGKMVCTGAKSE EDSRLAAR YARI IQKLGFT---AKFLDFKIONM NGSC VRF F IRLE  
DrTBPL2 SGKMVCTGAKSE EQSRLAAR YARVVQKLGFP---AKFLDFKIONM NGSC VCF F IRLE  
DmTBP SGKMVCTGAKSE EDSRLAAR YARI IQKLGFP---AKFLDFKIONM NGSC VVKF F IRLE  
DrTBP SGKMVCTGAKSE EQSRLAAR YARVVQKLGFP---AKFLDFKIONM NGSC VVKF F IRLE  
XlTBP SGKMVCTGAKSE EQSRLAAR YARVVQKLGFP---AKFLDFKIONM NGSC VVKF F IRLE  
MmTBP SGKMVCTGAKSE EQSRLAAR YARVVQKLGFP---AKFLDFKIONM NGSC VVKF F IRLE  
HsTBP SGKMVCTGAKSE EQSRLAAR YARVVQKLGFP---AKFLDFKIONM NGSC VVKF F IRLE  
BtTBP SGKMVCTGAKSE EQSRLAAR YARVVQKLGFP---AKFLDFKIONM NGSC VVKF F IRLE  
GgTBP SGKMVCTGAKSE EQSRLAAR YARVVQKLGFP---AKFLDFKIONM NGSC VVKF F IRLE  
CaTBP1 SGKMVCTGAKSE EDDSKLASR YARI IQKLGFN---AKFCDFKIONI GSTVVKFAIRLE  
ScTBP SGKMVVTGAKSE EDDSKLASR YARI IQKIGFA---AKFTDFKIONI NGSC VVKF F IRLE  
DdTBP SGKMVCTGAKSE DASRFAAR YARI IQKLDFP---ARFTDFKIONI NGSC VVKF F IKLE  
SpTBP1 SGKMVVLGGKSE EDDSKLASR YARI IQKLGFN---AKFTDFKIONI NGSC VVKF F IRLE  
ZmTBP1 SGKMVCTGAKSE QSKLAAR YARI IQKLGFP---AKFKDFKIONI NGSC VVKF F IRLE  
AtTBP1 SGKMVCTGAKSE DFSKMAAR YARIVQKLGFP---AKFKDFKIONI NGSC VVKF F IRLE  
AtTBP2 SGKMVCTGAKSE HLSKLAAR YARIVQKLGFP---AKFKDFKIONI NGSC VVKF F IRLE

\* \* . . . : . . . : . . \* :

▼

G1TBP AAAAAATIS-----VFSSAM YQPE IMPSLQVVFKIAERN-----ICCSVFADGQVT  
LmTBP KLAAYQLDPAMSIGVAKLQVSYEPERNGC--VLR LVGKSSRGDNQWSVCSVFTGIVQ  
TcTBP ELAQH-----EGITCSYEPDRSGC--IVRLAGKS--HDNQVQVSC TVFVTGIVS  
TbTRF4 ELAKH-----EGVTFSSYEPDRSGC--IVRLSGSS--RGNQWQVCCTVFTGIVVI  
**TvTBP2** QIASADG-----HKLCQYRPEKTAGL--IYRLTDPQ-----CTM LIFQSGSIV  
CeTLF EVAAK-----YPSESTYEP ELSVGL--VWRSVTPK-----ATLR LHTTGSIT  
DrTBPL1 EFTKN-----NRPIAS YEP ELPHPAA--SYRIKNLR-----STVQV FSTGNIT  
XlTBPL1 EFTKQ-----NRPHAS YEP ELPHPAV--CYRIKSLR-----ATLQIFSTGSIT  
HsTBPL1 EFTKN-----NRPHAS YEP ELPHPAV--CYRIKSLR-----ATLQIFSTGSIT  
MmTBPL1 EFTKN-----NRPHAS YEP ELPHPAV--CYRIKSLR-----ATLQIFSTGSIT  
BtTBPL1 EFTKN-----NRPHAS YEP ELPHPAV--CYRIKSLR-----ATLQIFSTGSIT  
GgTBPL1 EFTKN-----NRPHAS YEP ELPHPAV--CYRIKTLR-----ATLQIFSTGSIT  
**TvTBP1** SIASSPG-----HIKATY NPEIYAGL--IYRLASPK-----TTLMIYVSGHII  
EhTRF1 ELIQL-----APKITKYEP EIPGV--VYRLADPK-----M L LIFASGHIV  
EhTBP TLVDS-----YLAFCQYEP EIPGL--VYRMASPK-----VTLV FSTGIVV  
DmTRF NLNHV-----HQQFSSYEP EIPGL--IYRMVKPR-----IVLLIFVSGHIV  
TtTBP HLEQD-----HKKEVQYEP EIPGK--IYREFNTK-----IVLLIFVSGHIV  
CeTBP GLCIT-----HSQFSTYEP EIPGL--IYRMVKPR-----VLL LIFVSGHIV  
AaTBP GLVLT-----HGK FSSYEP EIPGL--IYRMVKPR-----IVLLIFVSGHIV  
DrTBPL2 GLVLT-----HQQFSSYEP EIPGL--IYRMVKPR-----IVLLIFVSGHIV  
DmTBP GLVLT-----HCN FSSYEP EIPGL--IYRMV RPR-----IVLLIFVSGHIV  
DrTBP GLVLT-----HQQFSSYEP EIPGL--IYRMIKPR-----IVLLIFVSGHIV  
XlTBP GLVLT-----HQQFSSYEP EIPGL--IYRMIKPR-----IVLLIFVSGHIV  
MmTBP GLVLT-----HQQFSSYEP EIPGL--IYRMIKPR-----IVLLIFVSGHIV  
HsTBP GLVLT-----HQQFSSYEP EIPGL--IYRMIKPR-----IVLLIFVSGHIV  
BtTBP GLVLT-----HQQFSSYEP EIPGL--IYRMIKPR-----IVLLIFVSGHIV  
GgTBP GLVLT-----HQQFSSYEP EIPGL--IYRMIKPR-----IVLLIFVSGHIV  
CaTBP1 GLAFA-----HGT FSSYEP EIPGL--IYRMVKPK-----IVLLIFVSGHIV  
ScTBP GLAFS-----HGT FSSYEP EIPGL--IYRMVKPK-----IVLLIFVSGHIV  
DdTBP LLHNA-----HTSFTNYEP EIPGL--IYKMIQPK-----VLL LIFVSGHIV  
SpTBP1 GLAYS-----HGT FSSYEP EIPGL--IYRMVKPK-----VLL LIFVSGHIV  
ZmTBP1 GLAYS-----HGAFSSYEP EIPGL--IYRMKQPK-----IVLLIFVSGHIV  
AtTBP1 GLAYS-----HAAFSSYEP EIPGL--IYRMKVPK-----IVLLIFVSGHIV  
AtTBP2 GLAYS-----HSA FSSYEP EIPGL--IYRMKLPK-----IVLLIFVSGHIV

\* . \* : . : . . . : . . \* . . :

```

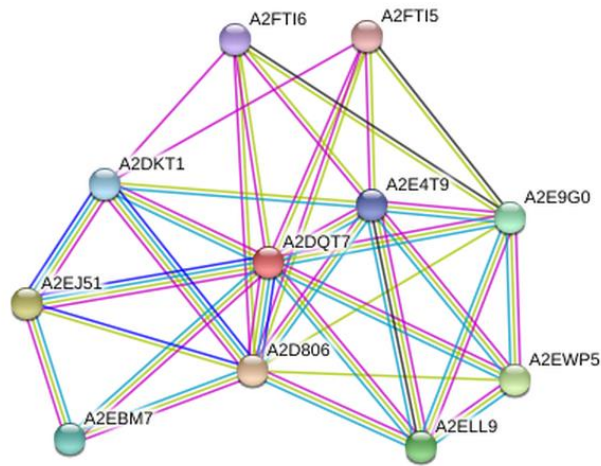
GI TBP      IVGARNIFDARDVITKLYEGFDYFIT-----
LmTBP      LMGARSMDELRFANAFVPIIAKNLDERKTA-----
TcTBP      ILGARSQKELQDAFYTLLEPIAQYAKK-----
TbTRF4     VLGARSREELLDAFYTVLPIIAQYAKR-----
TvTBP2    LTA-SEDDLTAGSNWYYPVQKFEKKSMSSELLI-----
CeTLF      VTGAQSEADVLEVLISKIYPIVLEFRCLERAKGNVAAQK-----
DrTBPL1    VTGPNVQ-SVASAVEEITYPLIFECRKTLS-----
XlTBPL1    VTGPDVK-SVASAIEQIYPPVFESRKTIL-----
HsTBPL1    VTGPNVK-AVATAVEQIYPPVFESRKEIL-----
MmTBPL1    VTGPNVK-AVATAVEQIYPPVFESRKEIL-----
BtTBPL1    VTGPNVK-AVATAVEQIYPPVFESRKEIL-----
GgTBPL1    VTGPNVK-AVASAVEQIYPPVFESRK-----
TvTBP1    ITGASSEELKAASAFIYPIIRLFASTEPPAA-----
EhTRF1     FTGGTEIEQINKAFSEIYKILQVANNDN-----
EhTBP      LTGASDEESLNLAYKNIYPIILANRKEDISNQ-----
DmTRF      FTGASRKDIMDCLEAISPILSFRKT-----
TtTBP      LTGASRENINKAFQKIYVWLYNYQKKDYRGANLHNQNLNIKPSIKN
CeTBP      ITGASTKRDIDEAFGQIYPIIKGFKK-----
AaTBP      LTGASVRQEIIDAFDNIYPIIKSFKKQ-----
DrTBPL2    LTGASERSEIYEAFENIYPIIKGFRKQ-----
DmTBP      LTGASVRQEIIDAFDKIFPIIKKFKKQS-----
DrTBP      LTGASVRGEIYEAFENIYPIIKGFRKTS-----
XlTBP      LTGASVRAEIIYEAFENIYPIIKGFRKTT-----
MmTBP      LTGASVRAEIIYEAFENIYPIIKGFRKTT-----
HsTBP      LTGASVRAEIIYEAFENIYPIIKGFRKTT-----
BtTBP      LTGASVRAEIIYEAFENIYPIIKGFRKTT-----
GgTBP      LTGASVRAEIIYEAFENIYPIIKGFRKTT-----
CaTBP1     LTGASKREEIIDAFESTIYPPVNEFRKN-----
ScTBP      LTGASQREEIYQAFEAIYPPVSEFRKM-----
DdTBP      LTGASVREEIYEAFENIYPPVSAFKKVNAITQ-----
SpTBP1     LTGASVREEIYQAFEAIYPPVSEFRKH-----
ZmTBP1     LTGASVREETYTAFENIYPPVAEFRKVQQ-----
AtTBP1     ITGASMRDETYKAFENIYPPVSEFRKIQQ-----
AtTBP2     ITGASMREETYTAFENIYPPVREFRKVQQ-----

```

**Figura 9.** Alineamiento de secuencias TBP, TBP-like y TvTBPs. Se muestra la secuencia carboxil terminal de las proteínas. Aminoácidos conservados en todas las proteínas se muestran sombreados en gris. Aminoácidos conservados en las TBPs, pero no encontrados en la posición equivalente en las TBP-like, se muestran en rojo. Aminoácidos conservados en las TBP-like, pero no encontrados en la posición equivalente en las TBPs, se muestran en amarillo. Las flechas indican la posición de las fenilalaninas conservadas en las TBPs que se intercalan en la secuencia de las cajas TATA. *Giardia intestinalis* TBP Q86GG9 (GITBP), *Leishmania major* TBP Q4QD71 (LmTBP), *Trypanosoma cruzi* TFIID-like protein Q4DK05 (TcTBP), *Trypanosoma brucei* TFIID-like protein Q387R9 (TbTRF4), *Trichomonas vaginalis* TBP1 A2DQT7 (TvTBP1), *Trichomonas vaginalis* TBP2 A2G4N7 (TvTBP2), *Caenorhabditis elegans* TBP P32085 (CeTBP), *Caenorhabditis elegans* TBP-like factor Q9XZQ7 (CeTLF), *Aedes aegypti* TBP Q16ME4 (AaTBP), *Danio rerio* TBP Q7SXL3 (DrTBP), *Danio rerio* TBP-like 1 Q5U385 (DrTBPL1), *Danio rerio* TBP-like 2 Q1JPY4 (DrTBPL2), *Xenopus laevis* TBP P27633 (XlTBP), *Xenopus laevis* TBP-like Q9W6Z2 (XlTBPL1), *Homo sapiens* TBP P20226 (HsTBP), *Homo sapiens* TBP-like1 P62380 (HsTBPL1), *Mus musculus* TBP P29037 (MmTBP), *Mus musculus* TBP-like P62340 (MmTBPL1), *Bos taurus* TBP Q2HJ52 (BtTBP), *Bos taurus* TBP-like Q32LB1 (BtTBPL1), *Gallus gallus* TBP O13270 (GgTBP), *Gallus gallus* TBP-like Q9YGV8 (GgTBPL1), *Entamoeba histolytica* TBP1 P52653 (EhTBP), *Entamoeba histolytica* TBP-like A7UFC2 (EhTBPL1), *Drosophila melanogaster* TBP P20227 (DmTBP), *Drosophila melanogaster* TBP-related factor Q27896 (DmTRF), *Tetrahymena thermophila* TBP Q27850 (TtTBP), *Arabidopsis thaliana* TBP P28147 (AtTBP), *Candida albicans* TBP O43133 (CaTBP), *Saccharomyces cerevisiae* TBP P13393 (ScTBP), *Dictyostelium discoideum* TBP P26355 (DdTBP), *Schizosaccharomyces pombe* TBP P17871 (SpTBP), *Zea mays* TBP1 P50158 (ZmTBP1).

### 1.5 Las TvTBPs podrían interactuar con otras proteínas putativas encontradas en *T. vaginalis*.

Para identificar posibles proteínas que se unen a las TvTBPs en *T. vaginalis*, realizamos un análisis teórico de interacción proteína-proteína en la base de datos STRING (Franceschini A., 2013) (Figura 10) que se basa en interacciones conocidas y reportadas de un número grande de organismos. El resultado de este análisis sugiere que ambas TvTBPs pueden interactuar con diversas proteínas de *T. vaginalis* que incluyen: el factor transcripcional TFIIIB putativo (TvTFIIIB), el factor transcripcional BRF1 putativo (TvBRF1), factores asociados a TBP (TAFs), subunidad de la RNA Pol II/I, entre otras (Tabla 3).



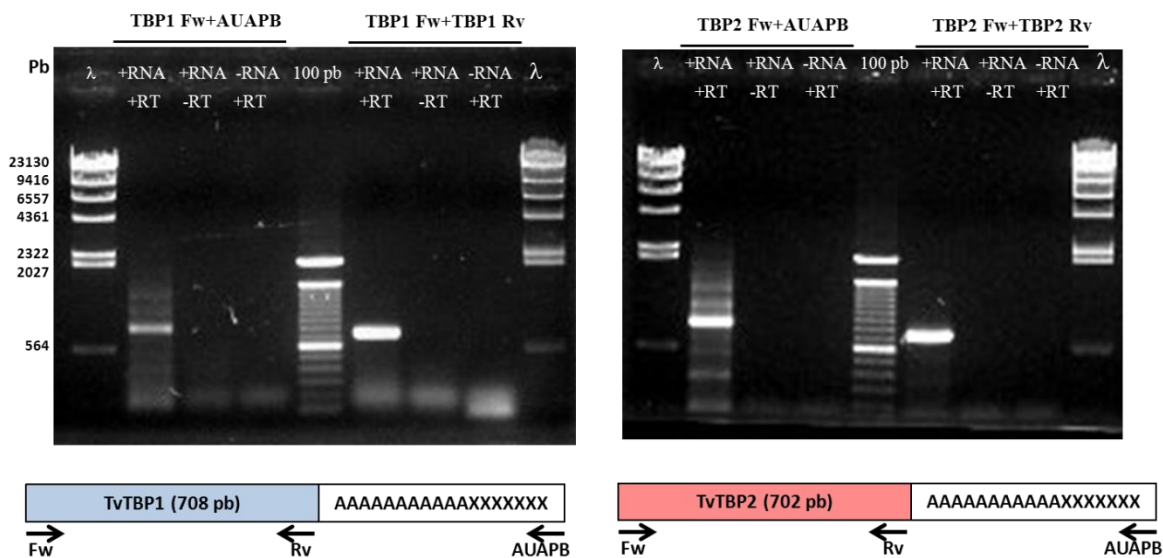
**Figura 10.** Red de interacción teórica proteína-proteína de las TvTBPs realizado en la base de datos STRING (Franceschini A., 2013). a) Resultado de la interacción TvTBP1 (A2DQT7) contra el proteoma de *T. vaginalis*. El resultado de TvTBP2 (A2G4N7) es el mismo, no mostrado. Los círculos simbolizan proteínas, y las líneas los tipos de evidencia en los que se basa la asociación. Líneas: en azul oscuro "Coocurrencia" familias de genes cuyos patrones de ocurrencia a través de genomas muestran similitudes; en azul claro "Base de datos", información relevante transferida de otras especies; en amarillo "Mención en texto", búsqueda de proteínas que son frecuentemente mencionadas juntas; en púrpura "Experimentos", datos experimentales relevantes de interacción proteína-proteína transferidos de otras especies.

**Tabla 3.** Proteínas de la red de interacción TvTBP1 con el proteoma de *T. vaginalis*.

<b>UniProtKB</b>	<b>Gen ID TrichDB</b>	<b>TrichDB</b>	<b>BLAST</b>	<b>En este trabajo</b>
A2DQT7	TVAG_291560	<i>Transcription initiation factor tfiid, putative</i>	TBP1	TvTBP1
A2G4N7	TVAG_285070	<i>Transcription initiation factor tfiid, putative</i>	TBP2	TvTBP2
A2D806	TVAG_071030	<i>Transcription factor iib, putative</i>	TFIIB	TvTFIIB
A2EJ51	TVAG_139840	<i>Transcription initiation factor brf1, putative</i>	BRF1	TvBRF1
A2DKT1	TVAG_247430	<i>DNA-directed RNA polymerase I largest subunit, putative</i>	<i>RNA polymerase II/I subunit</i>	-
A2EBM7	TVAG_482180	<i>Transcription factor tflIIB B component, putative</i>	<i>Myb-like DNA-binding domain containing protein</i>	-
A2FT16	TVAG_142600	<i>Lymphoid specific helicase, putative</i>	<i>SNF2 family N-terminal domain containing protein</i>	-
A2FT15	TVAG_142590	<i>Conserved hypothetical protein</i>	TAF	-
A2E4T9	TVAG_160840	<i>Conserved hypothetical protein</i>	TAF12	-
A2E9G0	TVAG_364490	<i>tfiid 150 kD subunit, putative</i>	TAF2	-
A2EWP5	TVAG_040830	<i>Transcription initiation factor TFIIID subunit, putative</i>	TAF7	-
A2ELL9	TVAG_149520	<i>tafi32, putative</i>	TAF9	-

## **2. Los genes que codifican para las proteínas TBP putativas se expresan en *T. vaginalis*.**

Realizamos un ensayo por RT-PCR para determinar la expresión de los genes TvTBP1 y TvTBP2. A partir de RNA total de *T. vaginalis* se sintetizó DNA complementario (cDNA) con la enzima transcriptasa reversa (RT). Se incluyeron dos reacciones de control en la síntesis de cDNA: una reacción que incluye RNA, pero no RT y otra reacción que no incluye RNA, pero si incluye RT. Para cada gen, se realizaron amplificaciones con dos parejas de oligonucleótidos: 1) Con un oligonucleótido *forward* que alinea con el extremo 5' del ORF de cada gen y con un oligonucleótido AUABP *reverse* que alinea con la etiqueta (representada con 7 Xs en la Figura 11) encontrada después de la cola de poly(A), que se añadió al sintetizar el cDNA; 2) Con un oligonucleótido *forward* que alinea con el extremo 5' del ORF de cada gen y con un oligonucleótido *reverse* que alinea con el extremo 3' del ORF de cada gen. Con cada pareja de oligonucleótidos se llevaron a cabo tres reacciones independientes: con el cDNA obtenido (+RNA +RT), con el control sin RT (+RNA -RT) y con el control sin RNA (-RNA+RT) (Figura 11).



**Figura 11.** Determinación de la expresión de los genes TvTBP por RT-PCR. Cada gen se amplificó con dos parejas de oligonucleótidos: 1) TBP Fw + AUAPB, el oligonucleótido *forward* alinea en con el extremo 5' de cada gen, y el oligonucleótido AUAPB alinea con la etiqueta encontrada después de la cola de poly(A) (representada con siete Xs), y 2) TBP Fw + TBP Rv, el oligonucleótido *forward* alinea en con el extremo 5' de cada gen, y el oligonucleótido *reverse* alinea con el extremo 3' de cada gen. Con cada pareja de oligonucleótidos se realizaron tres amplificaciones a partir del cDNA sintetizado con (+) o sin (-) RNA y con (+) o sin (-) RT. En el extremo inferior de la figura se muestra una representación del alineamiento de los oligonucleótidos, no está a escala.

Los amplicones más abundantes de la reacción con los oligonucleótidos TBP Fw + AUAPB los clonamos en el vector pGEM Easy-T vector (Promega) y enviamos una muestra para secuenciar, con el propósito de determinar la identidad del amplicón y analizar el extremo 3' UTR de los genes. Un análisis de las secuencias obtenidas mostró que corresponden a los genes TvTBP1 y TvTBP2 (Figura 12). Al comparar la secuencia de la TBP (TVAG\_291560), depositada en la base de datos de *Trichomonas*, con la proveniente de la de la secuenciación de TvTBP1, observamos que hay dos cambios silenciosos: en el nucleótido 180 T por C y en el nucleótido 369 A por G.

La secuenciación del extremo 3' nos permite identificar la posible señal de poliadenilación; además, nos indica el sitio de corte del mRNA para la poliadenilación. La región 3' del gen en la base de datos nos muestra la secuencia de nucleótidos después del sitio de corte, donde se observa una pequeña región rica en timinas también necesaria para el procesamiento (Figura 12a) (*Fuentes y Col. 2012*). El mismo análisis se realizó para el gen TvTBP2; la secuenciación mostró que no hay cambios en la secuencia de nucleótidos comparada con la secuencia depositada en la base de datos (TVAG\_285070), asimismo, se determinó el sitio de corte y las posibles señales de procesamiento (Figura 12b). Este resultado indica que ambos genes se expresan a nivel de RNA mensajero en *T. vaginalis*.

**a) TVTBP1(708 pb)**

```

1   ATGAATCCAA AAGAATTTAT TGGCTTAGAA GAAGCTAATA TAAATAATGC AAAAGAATTC
61  GAACAGCTTG TTAAGAAGGC TCAAGAGGAA TCAGAGGAAG AAATTACACC AGAATCAATA
121 CAGAGAATA TAAGTTTAAA TCAACCTCCC ATTCCCAAAG TTGTCAACAT TGTTCATGC
181 GTAGATTACG GATGCAAACT TGACCTTCTT AAGATTACAT CTTCAACAAG AAATTCAGAA
241 TATAATCCGA TTCGTTTCCC TGCTGCTACG CTAAGGAACA AAGAACCAA AGCCACAGCT
301 TTAGCTTTTA AAAACGGAAA AGTTAATATC GTAGGATGTA AAAGCGCTGA TATAGCATTA
361 AAAGCTGCGC GTTGCTTTGG TAGATTATTC AAAAATATCG GCATTAAAGA CGTTAAAAATC
421 AAATCATTCT CAATAGCTAA CATGGTTGCA ACAAATGGACT GCAAATCCC GATACATCTT
481 GAATCAATTG CTTTCATCGCC TGGTCATATT AAATTTGCTA CATACAATCC TGAAATTTAT
541 GCAGGTTTGA TTTATAGATT GGCTAGTCCA AAGACGACTT TAATGATATA CGTTTCAGGT
601 AAAATTATTA TCACTGGAGC TAAATCAGAA GAAGAATTA AAGCCGCATC TGCATTTATT
661 TATCCTATTC TTAGATTGTT TGCTAGCACA GAACCGCCAG CTGCTTAAAT TAATCCAATT
721 TTGTTTATTA AATTTCTATT GTTTCACAAA AAAAAAAAAA AAGTACTAGT CGACGCGTGG
781 CC

```

**Región 3' de la base de datos**

```

661 TATCCTATTC TTAGATTGTT TGCTAGCACA GAACCGCCAG CTGCTTAAAT TAATCCAATT
721 TTGTTTATTA AATTTCTATT GTTTCCTTTA TAAATACTAT TTTCCACTAT AGAGTATATC
AAGAGCCAAT TTCATTCATT TTTAGTAT

```

**b) TvTBP2 (702 pb)**

```

1   ATGGATGCTG CGAGTTTCTT TGATGCTGAT GCGGGTGTGA AAGAAATGAA TCTTGATGAT
61  TCAGAATCTA TGAAATGGTA TAAAGAATTA GAGAATGATG CTGATATCAC AGTATCAACA
121 ACAGAAGATA TAACATCATT AAAACCAAAA GTTGTAAATG TTATCGCTCT TGCACAATAT
181 TCTTGCAATA AACTCAATCT TGTCAGTATT GCCTCTACAA TTCGTAATGC TGAATACAAG
241 CCAAAGAGAA TTAAGCCCGT AGTTATTTCGT ATTCGTGACC CAAAGGCTAC AGGATTAGTA
301 TTCTCAAATG GCAAATCAA CATTGTTGGT TGCCGTTCTG TTGAGGATGC TAAGCGTGCT
361 GTTCACAAGT TTAGAAGAT GCTTCAGCAG ATTTGGATAG ATGTCAAGCT CGTCAATATT
421 ACTATATCAT CAATTGTTGC TACAATACAT ACACCATTCA ACATCGCTAT TGCACAAATA
481 GCCAGTGCTG ATGGCCACAA ACTATTTTGT CAATATAGAC CAGAAAAATT CGCTGGTTTG
541 ATTTACAGGC TTACAGACCC TCAATGCACA ATGCTCATCT TCCAATCAGG AAGTATTGTT
601 CTTACTGCAA AATCCGAGGA TGACCTTACA GCAGGATCAA ACTGGATATA CCCAGTTCTA
661 CAGAAGTTCG AGAAGAAGTC TATGTCGGAA TTACTTATAT AAATCCGCAA TAAAATTGTT
721 GTATTTTGA AATATTATTC TTTGTAAAGC TCTAAATTTT CTC TAAAAA AAAAAAAAAA
781 AAGTACTAGT CGACGCGTGG CC

```

**Región 3' de la base de datos**

```

661 CAGAAGTTCG AGAAGAAGTC TATGTCGGAA TTACTTATAT AAATCCGCAA TAAAATTGTT
721 GTATTTTGA AATATTATTC TTTGTAAAGC TCTAAATTTT CTC TAACCCA TCTTTATTTT
CTTGGGTAGT CCCCTCTTAA

```

**Figura 12.** Secuencias TvTBPs obtenidas a partir de la secuenciación de productos RT-PCR. También se muestran las regiones 3' después del sitio corte de los mRNAs teórico obtenidas de la base de datos de *T. vaginalis* (TrichDB, *Aurecochea C.*, y *Col. 2009*). a) TvTBP1, clona 2 y b) TvTBP2, clona 7. En negritas, codones ATG y TAA. En cursivas, nucleótidos de las señales potenciales de procesamiento: señal de poliadenilación, sitio de corte para la poliadenilación y elemento rico en U, río abajo del sitio de corte. Subrayado, los nucleótidos que alinean con los oligonucleótidos *forward* específicos para cada gen y AUAPB *reverse*. En cursiva y subrayado, los sitios de restricción *EcoRI* (55) y *EcoRV* (105) de TvTBP1 y TvTBP2, respectivamente, diferenciales en cada gen. En rojo, nucleótidos en TvTBP1 que difieren del gen TvTBP1 reportado en la base de datos, que no modifica el aminoácido codificado.

**3. Las TBPs de *T. vaginalis* no complementan la función de una levadura mutante en TBP.**

Como primera aproximación para caracterizar a las TvTBPs determinamos si éstas eran funcionales *in vivo*. Realizamos un ensayo de complementación en levadura para evaluar si las TvTBPs tenían la misma función que la ScTBP (codificada por el gen *SPT15*), una

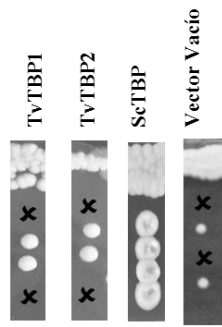
proteína ampliamente estudiada y con características canónicas. Utilizamos una cepa comercial de *S. cerevisiae*, diploide, heterocigota para algunos genes y mutante en uno de los alelos de TBP por reemplazo con el *cassette* de Kan que confiere resistencia al antibiótico G418 (Genotipo: BY4743, *MATa/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 MET15/met15Δ0 ura3Δ0/ura3Δ0*). Esta cepa la transformamos con vectores que llevaban los genes TvTBP1 o TvTBP2 (TvTBP1-pYES2 o TvTBP2-pYES2). La determinación de complementación se realizó por disección de tétradas, como se describe en materiales y métodos, con la premisa de que aquellas células que complementen la ausencia de ScTBP serán levaduras haploides y resistentes al antibiótico G418. Como controles del experimento se transformó la misma cepa parental con el vector vacío (pYES2; control negativo) y con el vector que lleva la ScTBP (ScTBP-pYES2; control positivo).

Las tétradas observadas en el microscopio se separaron y se colocó cada una de las 4 esporas en medio agar esperando crecimiento, si hay complementación se esperaría el crecimiento de todas. El resultado de la transformación con TvTBP1 o TvTBP2 mostró el crecimiento de solo 2 colonias. Los controles crecieron como era de esperarse, 4 colonias en general de todas las tétradas disectadas en las levaduras transformadas con el control ScTBP y 2 colonias en las levaduras transformadas con el vector vacío (Figura 13). No obstante, todas las células que fueron viables después de la disección se analizaron.

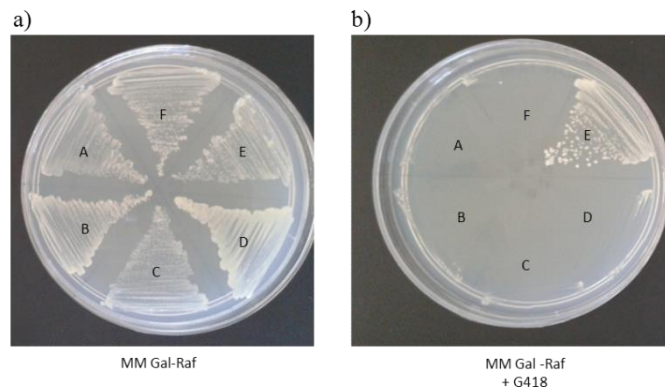
En total obtuvimos y analizamos 31 esporas viables de la transformación con TvTBP1-pYES2, 38 de la transformación con TvTBP2-pYES2, 22 de la transformación con SPT15-pYES2 y 18 de la transformación con pYES2 vacío. De todas ellas corroboramos su haploidía analizando la segregación de los genes heterocigotos *LYS2/lys2Δ0* y *MET15/met15Δ0* por crecimiento en diversas cajas de medio suplementadas con lisina o metionina y en cajas de medio sin estos aminoácidos. Así mismo, se analizó el crecimiento de estas colonias en cajas con G418.

Únicamente aquellas transformadas con ScTBP fueron capaces de crecer en un medio con G418 (Figura 14), no así las transformadas con TvTBP1 y TvTBP2. Este resultado indica que TvTBP1 y TvTBP2 no pueden complementar la ausencia de TBP en la levadura mutante, como lo hace ScTBP.





**Figura 13.** Disección de tétradas para el ensayo de complementación de TBP. Se muestra como ejemplo el crecimiento de las colonias después de la disección de una tétrada para cada transformación con los plásmidos TvTBP1-pYES2, TvTBP2-pYES2, ScTBP-pYES2 y pYES2. Únicamente se observó el crecimiento de cuatro esporas en algunas colonias transformadas con ScTBP-pYES2, indicando complementación. Las X indican el lugar donde se colocó una espora y no hubo crecimiento.



**Figura 14.** Análisis de complementación de función, crecimiento de células haploides en un medio mínimo suplementadas con G418. a) Todas las colonias haploides pueden crecer en un medio mínimo suplementado con galactosa. Como ejemplo, se muestra el crecimiento de dos esporas viables para cada ensayo de transformación: A, B para TvTBP1 y C, D para TvTBP2. Y el crecimiento de una espora para cada control: E para ScTBP y F para el vector vacío. b) Solamente colonias transformadas con ScTBP crecieron en un medio mínimo suplementado con galactosa y G418.

Adicionalmente, por medio de la técnica mutagénesis dirigida *in vitro* generamos una proteína TvTBP1 modificada en su secuencia de aminoácidos (TvTBP\*), sustituyendo en la secuencia de TvTBP1 clonada previamente en pYES2: N109 por V, V111 por T, Y180 por F, T193 por V, M195 por L y Y197 por F (Figura 15). Esta modificación con el fin de analizar si esta nueva proteína, al contener los aminoácidos más importantes involucrados en la unión con las cajas TATA, pudiera ser capaz de complementar la mutación de SPT15 en levadura. De la misma manera, este vector se transformó en la cepa mutante en TBP de *S. cerevisiae* y se determinó si complementaba la ausencia de TBP. No obstante, ninguna de las células analizadas fue capaz de crecer en un medio suplementado con G418, indicando que la TvTBP\* no es capaz de complementar a una



levadura mutante en TBP. El ensayo de mutagénesis dirigida, a partir de TvTBP1-pYES2 construido en el presente trabajo, y el análisis de complementación en levadura lo realizó Alexandra Ibáñez Escribano, estudiante de doctorado de la Universidad Complutense de Madrid, quien realizó una estancia en nuestro laboratorio, como parte de sus estudios de doctorado.

a)

	N69	V71	F99	F116	V122	T124	N159	V161	F190	V203	L205	F207	L214	T215
ScTBP	•	•	•	•	•	•	•	•	•	•	•	•	•	•
TvTBP1	•	•	•	•	N	V	•	•	Y	T	M	Y	I	•
TvTBP1*	•	•	•	•	•	•	•	•	•	•	•	•	I	•
TvTBP2	•	I	I	•	N	V	S	•	•	T	•	•	•	•

b)

```

1   ATGAATCCAAAAGAATTTATTGGCTTAGAAGAAGCTAATATAAATAATGCAAAAAGAATTCGAACAGCTTGTTAAAGAAGCTCAAGAGGAA
1   M N P K E F I G L E E A N I N N A K E F E Q L V K E A Q E E

91  TCAGAGGAAGAAATTACACCAGAATCAATACAGAGAAATATAAGTTTAAATCAACCTCCCATCCCAAAGTTGTCAACATTGTTGCATGT
31  S E E E I T P E S I Q R N I S L N Q P P I P K V V N I V A C

181 GTAGATTACGGATGCAAACCTTGACCTTCTTAAGATTACATCTTCAACAAGAAATTCAGAATATAATCCGATTTCGTTTCCCTGCTGCTACG
61  V D Y G C K L D L L K I T S S T R N S E Y N P I R F P A A T

                                     5´-CGGAAAAGTTGTTATCACAGGATGTAAAAGC-3´

271 CTAAGGAACAAAGAACCAAAAAGCCACAGCTTTAGCTTTTAAAAACGGAAAAGTTAAATATCGTAGGATGTAAAAGCGCTGATATAGCATT
91  L R N K E P K A T A L A F K N G K V N I V G C K S A D I A L
                                     V T

361 AAAGCTGCACGTTGCTTTGGTAGATTATTCAAAAATATCGGCATTAAAGACGTTAAAATCAATCATCTCAATAGCTAACATGGTTGCA
121 K A A R C F G R L F K N I G I K D V K I K S F S I A N M V A

                                     5´-CCTGAAATTTT

451 ACAATGGACTGCAAATTCGGATACATCTTGAATCAATTGCTTCATCGCCTGGTCATATTAATTTGCTACATACAATCCTGAAATTTAT
151 T M D C K F P I H L E S I A S S P G H I K F A T Y N P E I Y
                                     F

GCAGGTTTGATTTATAG-3´   5´-CCAAGACGGTTTTACTGATATTCGTTTCAGG-3´
541 GCAGGTTTGATTTATAGATTGGCTAGTCCAAAGACGACTTTAATGATATACGTTTCAGGTTAAATTTATATCACTGGAGCTAAATCAGAA
181 A G L I Y R L A S P K T T L M I Y V S G K I I I T G A K S E
                                     V L F

631 GAAGAATTAAGCCGCATCTGCATTTATTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT
211 E E L K A A S A F I Y P I L R L F A S T E P P A A *
```

**Figura 15.** Mutagénesis dirigida TvTBP1. a) Tabla que resume algunos de los aminoácidos importantes involucrados en el reconocimiento de las cajas TATA de ScTBP y la comparación con TvTBP1, TvTBP1\* y TvTBP2. Los números corresponden a los aminoácidos de ScTBP. b) Secuencia de nucleótidos y aminoácidos de TvTBP1. Se muestran el alineamiento con los tres oligonucleótidos utilizados para sustituir los siguientes aminoácidos: N109 por V, V111 por T, Y180 por F, T193 por V, M195 por L y Y197 por F. En rojo se indica el nucleótido para llevar a cabo la sustitución del aminoácido. Subrayado se indica el codón que se modifica.

#### **4. Las TvTBPs no se unieron a los promotores de DNA analizados en este trabajo.**

Para determinar si las proteínas TvTBPs tienen la capacidad de reconocer un motivo, ya sea una caja TATA canónica o una región no canónica en los promotores de *T. vaginalis* realizamos un ensayo de cambio en la corrida electroforética, EMSA (por sus siglas en inglés *electrophoretic mobility shift assay*). Esta técnica se basa en juntar fragmentos de DNA marcados (sondas) y proteínas, posteriormente esta mezcla se corre en un gel de acrilamida de tal forma que, si hay interacción de la sonda con las proteínas, el patrón electroforético migrará diferente a la sonda sola, en particular, se retrasará. A continuación, se describe la construcción de las sondas y de las proteínas utilizadas para este ensayo.

#### **Sondas**

Seleccionamos cinco regiones de promotores de genes anteriormente caracterizados, transcritas por las tres RNA polimerasas (Figura 16). Las regiones de DNA fueron amplificadas por PCR (ver lista de oligonucleótidos en anexos), clonadas en un vector y verificadas por secuenciación. Posteriormente marcadas radioactivamente para los ensayos EMSA.

- 1) **ProHis4Sc.** Fragmento de 172 pb de un promotor del gen *HIS4* transcrito por la RNA Pol II de *S. cerevisiae* (Fishburn y Col., 2012; Donahue y Col., 1982), contiene una caja TATA canónica funcional.
- 2) **ProrDNATv.** Fragmento de 164 pb del promotor del cistrón ribosomal de *T. vaginalis* transcrito por la RNA Pol I. Esta región no contiene una caja TATA aparente, anteriormente caracterizado en nuestro laboratorio por el Dr. Bernardo Franco (Franco y Col., 2012).
- 3) **ProSCSTv.** Región promotora de 81 pb del gen que codifica para la subunidad  $\beta$  de la succinil CoA sintetasa (Lahti y Col., 1992; Quon y Col., 1994), transcrita por la RNA Pol II de *T. vaginalis*. Contienen dos elementos Inr y no contiene una caja TATA aparente.
- 4) **Pro5STv.** Fragmento de 186 pb del promotor del gen ribosomal 5S de *T. vaginalis*, transcrito por la RNA Pol III. Caracterizado anteriormente en nuestro laboratorio por la Dra. Ana Lilia Torres (Torres-Machorro y Col., 2006), contiene una región rica en TA.
- 5) **ProU6Tv.** Fragmento de 120 pb del promotor del gen RNA nuclear pequeño U6 (Simoës-Barbosa y Col., 2008), transcrito por la RNA Pol III de *T. vaginalis*. Contiene una caja TATA putativa no caracterizada (TATAAA).

### 1) *S. cerevisiae* His4

GAATTCGATTGGATATGACTATGAACAGTAGTATACTGTGTATATAATAGATATGGAACGTTATATTCACCTCCGATG  
TGTGTTGTACATACATAAAAAATATCATAGCACAACTGCGCTGTGTAATAGTAATACAATAGTTTACAAAATTTTTTTCTGAATAA  
TGGTTTTGCCGATTCTACAATCACTAGTGAATTC

### 2) *T. vaginalis* rDNA

GAATTCAGTAGTGATTGCACATGTACGGAGTTATAATCATAACTCCCCATGCTCCACAACATTGCCCATACAAAGTTTGCTCGGGG  
ACGCTTGTGTATGAGCAAGACGAAAAATACCTGAAACTCGACTCCATCGATTGAATTTTTGTAGACCGA  
AAAAACATGACCTTTTGCTCCTATGAATCGAATTC

### 3) *T. vaginalis* SCS

GAATTCGATTTGCGGGAGGTTTTAAACGATTAACCCTCAATTTTCTATCTCAACCGTTTTTTGTTGATCACTTCACATTACAATGC  
TTAGAATCACTAGTGAATTC

### 4) *T. vaginalis* 5S rDNA

GAATTCGATTTACTACACTCTACACCTATGTTTGTCCCTGCTAAAACCTTTATATGTTGAGATATTTTATAATTTTATATTTTTTA  
TTAAAAAATGAAAATTTTTTCATAAGCGAGACTCCCTCCGCTGGAGGAAATTGAGAGTGGTAATGGTAGTG  
TGAATAGTGTGTGACGAACAATCCTTTTGAAGCGGCCACAAATCACTAGTGAATTC

### 5) *T. vaginalis* U6 snRNA

GAATTCAGTAGTGATTATTTCTGCAATTTTCCATGATATAATCCATATATTTTTGTTTTATTCTTTATATTCTATAT  
GATATTTGAAGACAAAAACATTTTAGTTTATAAAGCATTGTCAAACATATTCTATTAAATCGAATTC

**Figura 16.** Regiones promotoras usadas como sondas para los ensayos EMSA. 1) Fragmento del gen *HIS4* de *S. cerevisiae*, transcrito por la RNA Pol II. 2) Fragmento del cistron ribosomal de *T. vaginalis*, transcrito por la RNA Pol II. No contiene una caja TATA canónica. 3) Fragmento del gen que codifica la subunidad  $\beta$  del gen de la succinil CoA sintetasa, transcrito por la RNA Pol II de *T. vaginalis*. No contiene una caja TATA canónica. 4) Fragmento de promotor del gen ribosomal 5S transcrito por la RNA Pol III de *T. vaginalis*, contiene una región rica en TA. 5) Fragmento del promotor del gen U6 snRNA transcrito por la RNA Pol III de *T. vaginalis*, contiene una caja TATA putativa no caracterizada. Los fragmentos de DNA fueron amplificados por PCR, clonados en pGEM-T Easy, digeridos con *EcoRI* y rellenados en sus extremos con [ $\alpha$ - $^{32}$ P]dATP. En *italicas*: sitios de restricción para *EcoRI* del vector pGEM-T Easy. En *gris*, secuencia del vector pGEM-T Easy. En *negrita*, codón ATG. *Subrayado*, caja TATA, caja TATA putativa o elemento Inr. Se muestra el sitio de inicio de la transcripción +1.

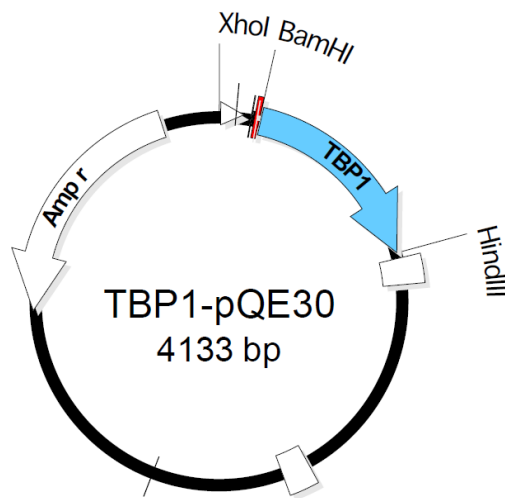
## Proteínas recombinantes

Para los ensayos EMSA se utilizaron las siguientes proteínas recombinantes, descritas en la tabla 4: 6xHis-TvTBP1 (Figura 17), 6xHis-TvTBP1\* (Figura 18), 6xHis-TvTBP2 (Figura 19), 6xHis-IBP39 (Figura 20), 6xHis-ScTBP (Figura 21), GST-TFIIB y GST-BRF1. En el presente trabajo se estandarizó el primer ensayo de *Western blot* realizado en el laboratorio, el ensayo se llevó a cabo como se indica en materiales y métodos. Los resultados de la secuenciación de las proteínas recombinantes TvTBPs, IBP39 y ScTBP se pueden consultar en los anexos de este trabajo (Anexos 2).

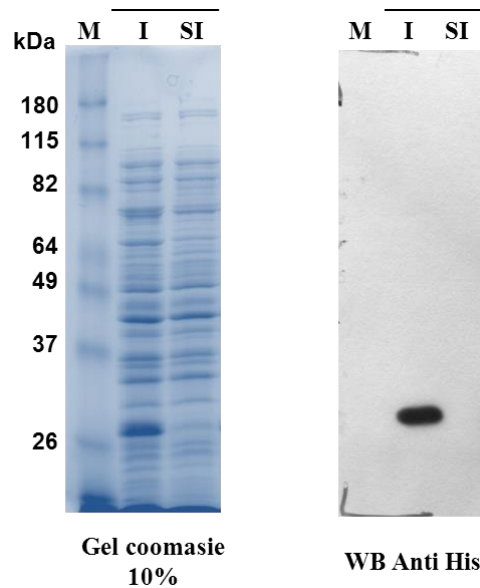
**Tabla 4. Descripción de las proteínas recombinantes usadas para los ensayos EMSAS.** Se indica el nombre usado este trabajo, el nombre del gen, el gen ID de la base de datos TrichDB (*Aurrecoechea C. y Col., 2009*) y para SPT15 de la base de datos SDG (*Saccharomyces Genome Database*), el peso de la proteína recombinante y el nombre de la persona contribuyó con la construcción de la proteína recombinante.

Proteína recombinante	Gen	Gen ID	Proteína kDa	Construido por
6xHis-TvTBP1	TvTBP1	TVAG_291560	27.22	Bernardo Franco Barcenas
6xHis-TvTBP1*	TvTBP1	TVAG_291560	27.22	Alexandra Ibáñez Escribano
6xHis-TvTBP2	TvTBP2	TVAG_285070	27.14	Lluvia Rosas Hernández
6xHis-IBP39	IBP39	TVAG_455080	40.6	Lluvia Rosas Hernández
6xHis-ScTBP	SPT15	YER148W	27.97	Lluvia Rosas Hernández
GST-TFIIB	TvTFIIB	TVAG_071030	60.7	Olivia Parra Marín
GST-BRF1	TvBRF1	TVAG_139840	71.6	Alexandra Ibáñez Escribano

## 6xHis-TvTBP1



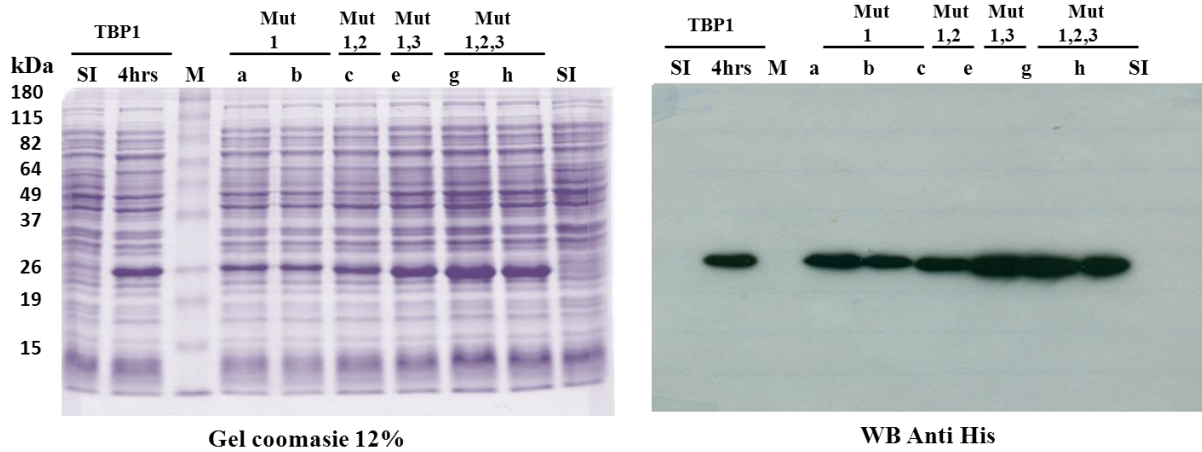
## Expresión



**Figura 17.** Expresión de la proteína recombinante 6xHis-TvTBP1 (27.22 kDa) en la cepa M15 (pREP4) de *E. coli*, se muestra la clona sin inducir y 1 clona inducida por 4 horas con IPTG 1 mM. El *Western blot*, con una sonda anti His revela la presencia de la proteína. (Pag. 126, bitácora #1 Lluvia). Este ensayo representa el primer ensayo de *Western blot* realizado en nuestro laboratorio, elaborado durante en el presente trabajo. La proteína recombinante 6xHis-TvTBP1 fue construida por el Dr. Bernardo Franco (Franco 2012, Tesis doctoral).

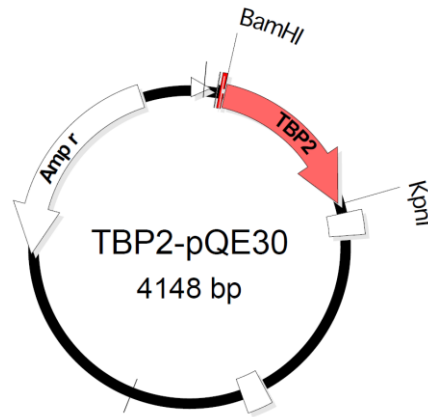
**6xHis-TvTBP1\*  
(Modificada)**

**Expresión**

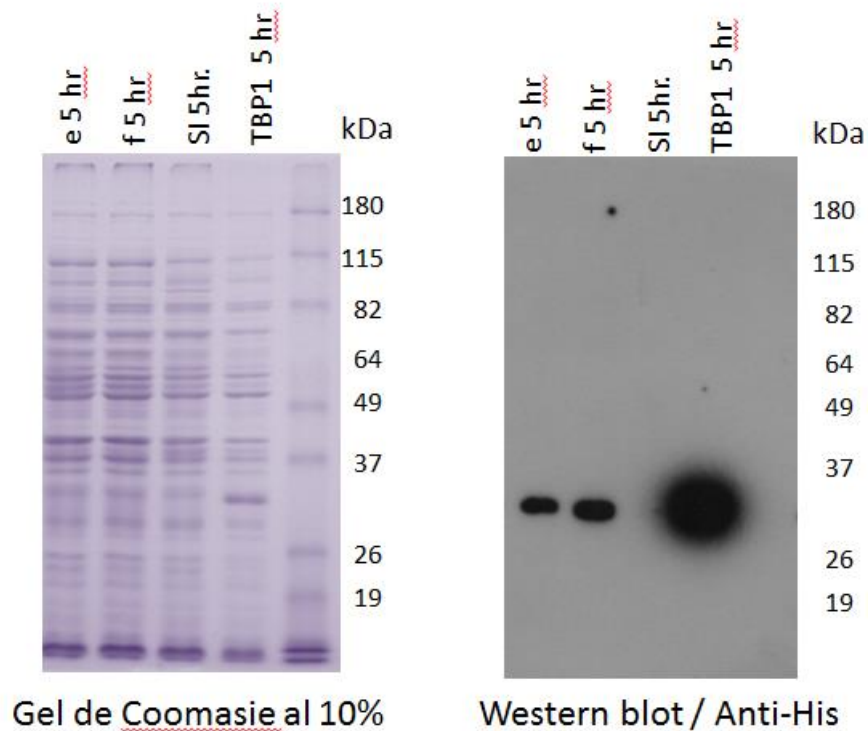


**Figura 18.** Expresión de las proteínas recombinantes 6xHis-TvTBP1\* (27.22 kDa) en la cepa M15 (pREP4) de *E. coli*. Como control se muestra la expresión de 6xHis-TvTBP1 sin inducir e inducida 4 horas. Se muestra la expresión de las proteínas modificadas: 6xHis-TvTBP1 Mut1, clonas “a” y “b”; 6xHis-TvTBP1 Mut1,2, clona “c”; 6xHis-TvTBP1 Mut1,3, clona “e” y 6xHis-TvTBP1 Mut1,2,3, clonas “g” y “h”. Y como control sin inducir la clona 6xHis-TvTBP1 Mut1, clona “a”. Las inducciones se realizaron por 4 1/2 horas con IPTG 1 mM. El *Western blot* con una sonda anti His revela la presencia de las proteínas. (Pag. 121a Bitácora #2 Lluvia). La proteína recombinante 6xHis-TvTBP1\* fue construida por Alexandra Ibáñez Escribano.

## 6xHis-TvTBP2



## Expresión

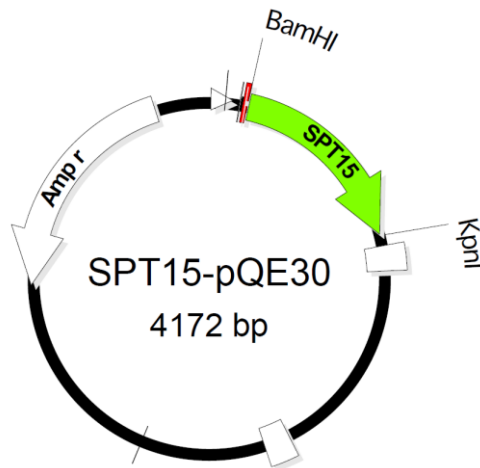


**Figura 19.** Expresión de la proteína recombinante 6xHis-TvTBP2 (27.14kDa), en la cepa M15 (pREP4) de *E. coli*. Se muestra la clona sin inducir y clonas e y f inducidas por 5 horas con IPTG 1 mM. Como control, se muestra la proteína 6xHIS-TvTBP1. El *Western blot*, con una sonda anti His revela la presencia de las proteínas.

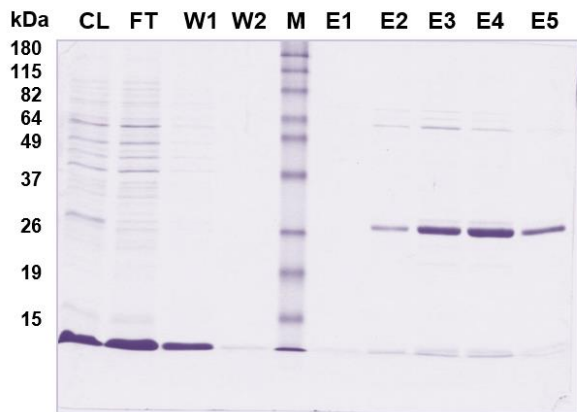




## 6xHis-SPT15



## Expresión



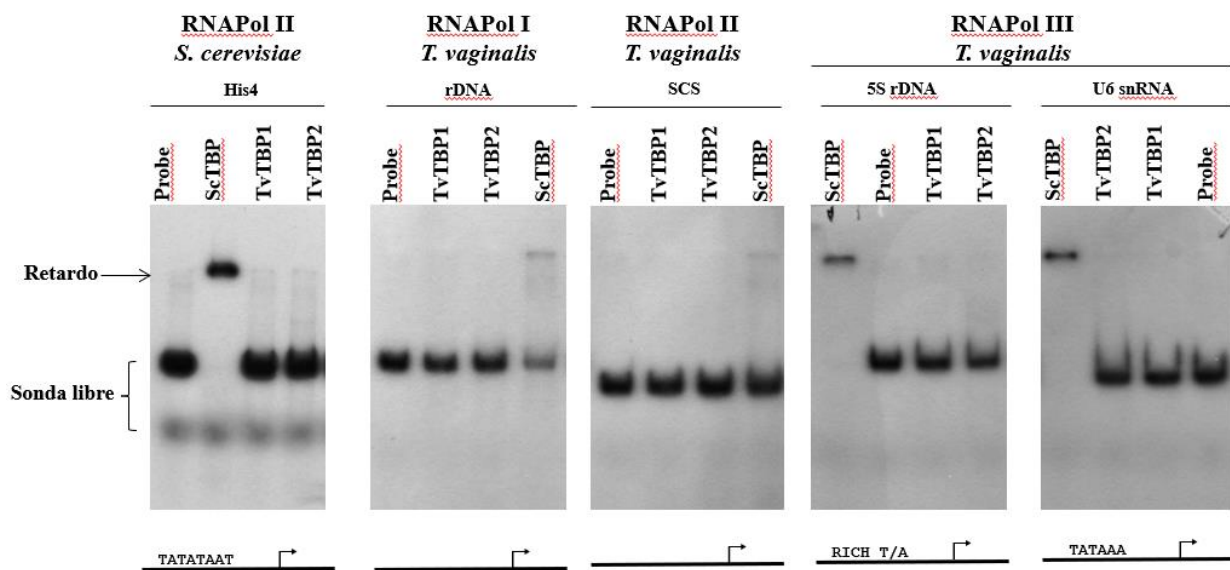
Gel coomassie



Western blot

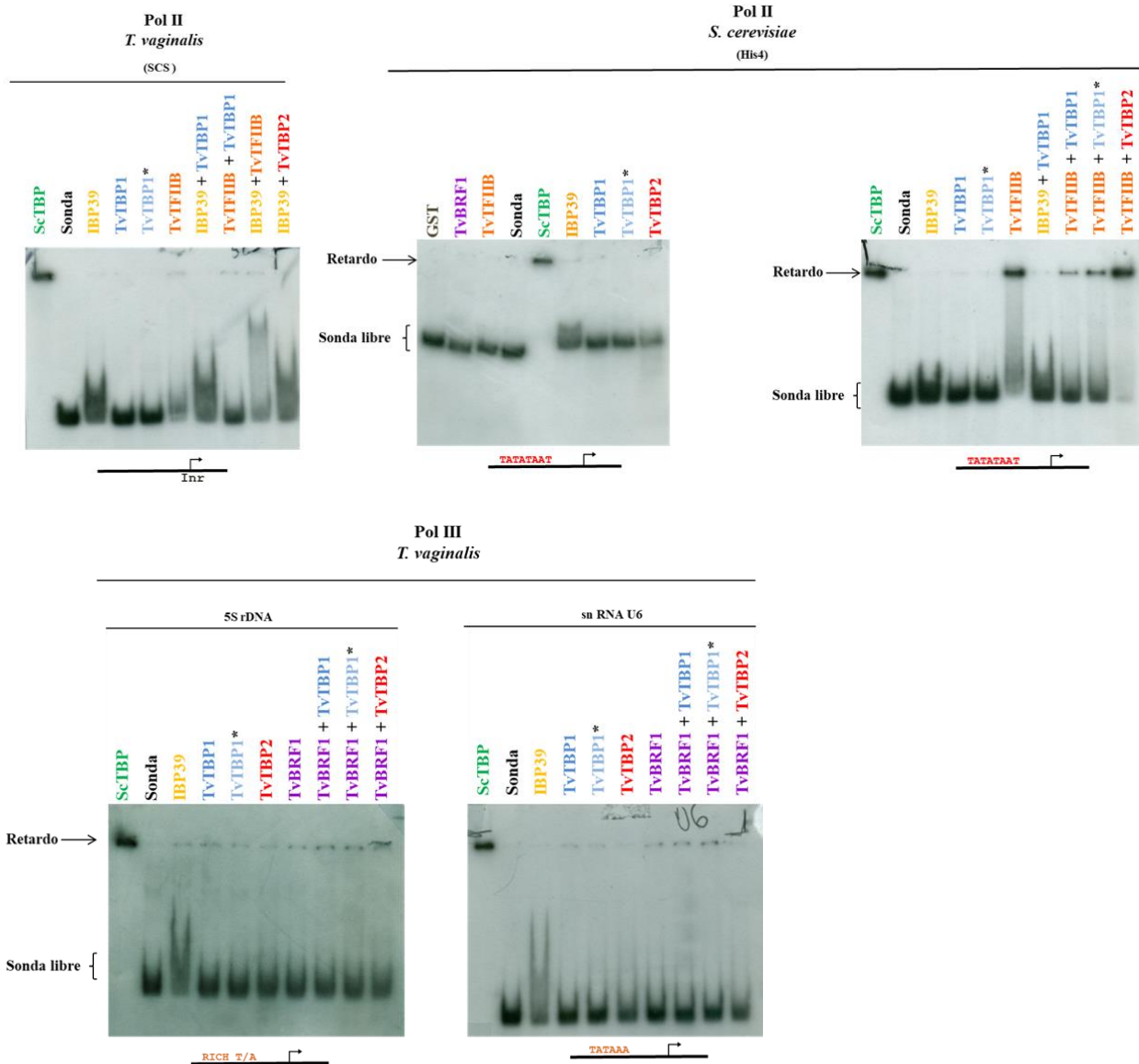
**Figura 21.** Purificación y expresión de la proteína recombinante 6xHis-SPT15 (ScTBP). Cromatografía de afinidad en columnas de níquel. Inducción de 250 mL de cultivo, 1 mM de IPTG. CL: Células lisadas, FT: Primer fluido no unido, W1 y W2: Lavados 1 y 2, M: Marcador preteñido, E1, E2, E3, E4 y E5: Eluidos. El ensayo *Western blot* con una sonda anti His revela la presencia de las proteínas. Peso esperado 27.97 kDa.

El resultado del ensayo EMSA mostró que las proteínas TvTBP1 y TvTBP2, en las condiciones de unión descritas en este trabajo, no fueron capaces de unirse a un promotor que contiene una caja TATA, tampoco a ningún otro promotor de *T. vaginalis* aquí probados (Figura 22). Como control positivo del ensayo se observa que, como es de esperarse, ScTBP reconoce y se une al promotor que contiene la caja TATA de *S. cerevisiae*. Se observa también un ligero desplazamiento (o retardo) de la sonda con ScTBP en los promotores de *T. vaginalis* reconocidos por la RNA Pol I y II, y se observa un desplazamiento total, con esta proteína control, en las sondas de los promotores de *T. vaginalis* reconocidos por la RNA Pol III, aparentemente por la presencia de regiones ricas en AT (Figura 22).



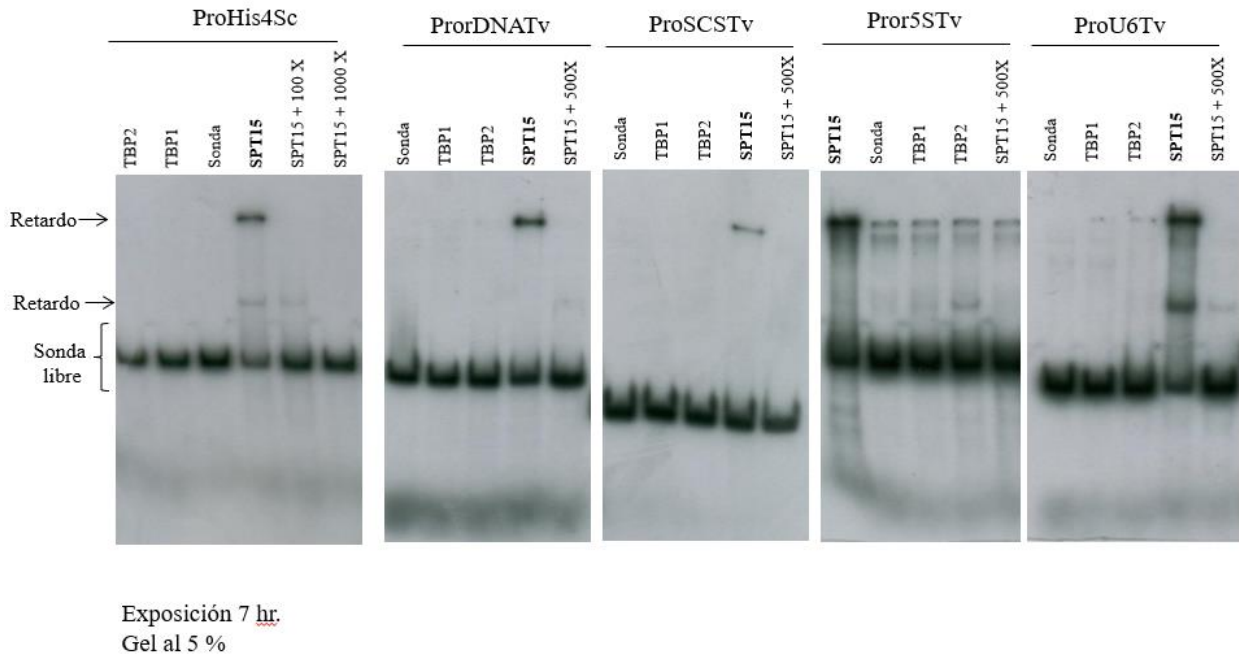
**Figura 22.** Análisis para evaluar la interacción de las TvTBPs con el DNA. El ensayo se llevó a cabo con las proteínas TvTBP1, TvTBP2 y ScTBP, analizando su interacción con sondas de 5 promotores distintos, uno de ellos conteniendo con una caja TATA canónica de *S. cerevisiae* y los promotores restantes de *T. vaginalis* sin caja TATA. Sondas: 1) La región promotora del gen *HIS4*, reconocida por la RNA Pol II de *S. cerevisiae* (His4), 2) Región promotora del gen RNA ribosomal, reconocido por la RNA Pol I de *T. vaginalis* (rDNA), 3) Región promotora de la subunidad  $\beta$  de la succinil CoA sintetasa (SCS), 4) Región promotora del gen ribosomal 5S (5S rDNA) y 5) región promotora del gen RNA pequeño nuclear U6 (U6 snRNA), ambos reconocidos por la RNA Pol III de *T. vaginalis*. El contexto de la secuencia TATA se muestra al final de los paneles.

Como una estrategia adicional se añadió a la interacción proteína – DNA, los factores transcripcionales putativos TFIIB (referido aquí como TvTFIIB) y BRF1 (referido aquí como TvBRF1) de *T. vaginalis*. Sin embargo, no se observó cambio en el resultado del ensayo, las TvTBPs no fueron capaces de unirse a las sondas analizadas (Figura 23). Adicionalmente, también se analizó la interacción de la proteína TvTBP1\* (Figura 15) y no se observó interacción con las regiones promotoras probadas (Figura 23).



**Figura 23.** Análisis para evaluar la interacción de TvTBPs, TvTBP1\* y el DNA con la presencia de TvTFIIB y TvBRF1, factores transcripcionales putativos de *T. vaginalis*. Cinco regiones promotoras fueron probadas: 1) La región promotora del gen HIS4, reconocida por la RNA Pol II de *S. cerevisiae* (His4), 2) Región promotora del gen RNA ribosomal, reconocido por la RNA Pol I de *T. vaginalis* (rDNA), 3) Región promotora de la subunidad β de la succinil CoA sintetasa (SCS). 4) Región promotora del gen ribosomal 5S (5S rDNA) y 5) región promotora del gen RNA pequeño nuclear U6 (U6 snRNA), ambos reconocidos por la RNA Pol III de *T. vaginalis*.

Finalmente, las uniones observadas del control SPT15 con el DNA se corroboraron por competencia con sondas frías (Figura 24).



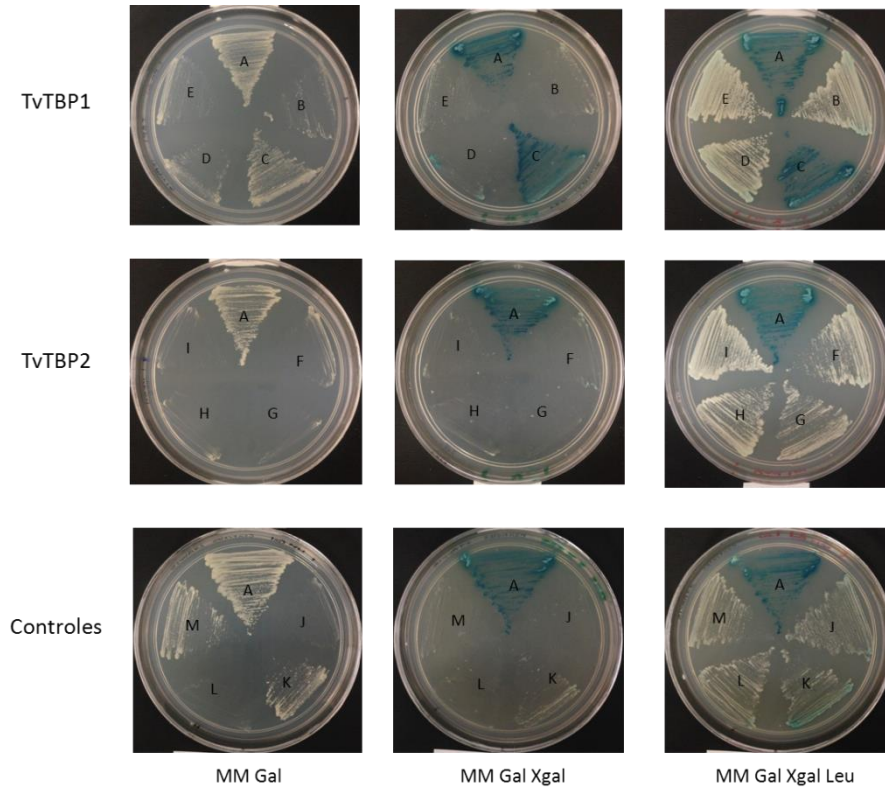
**Figura 24.** Análisis para evaluar la interacción de la proteína ScTBP con el promotor de *S. cerevisiae* y los promotores de *T. vaginalis*. Ensayos EMSA con sonda marcada y sin marcar (sonda fría). Se añadió a la reacción sonda sin marcar 100X, 500X o 1000X. La presencia de la sonda fría compite con la interacción ScTPB-sonda marcada. Cinco regiones promotoras fueron probadas. 1) La región promotora del gen HIS4, reconocida por la RNA Pol II de *S. cerevisiae* (His4). 2) Región promotora del gen RNA ribosomal, reconocido por la RNA Pol I de *T. vaginalis* (rDNA). 3) Región promotora de la subunidad  $\beta$  de la succinil CoA sintetasa (SCS). 4) Región promotora del gen ribosomal 5S (5S rDNA) y 5) región promotora del gen RNA pequeño nuclear U6 (U6 snRNA), ambos reconocidos por la RNA Pol III de *T. vaginalis*.

## **5. *TvTBP1* es capaz de interactuar con *TvTBRF1* in vivo, a diferencia de *TvTBP2*.**

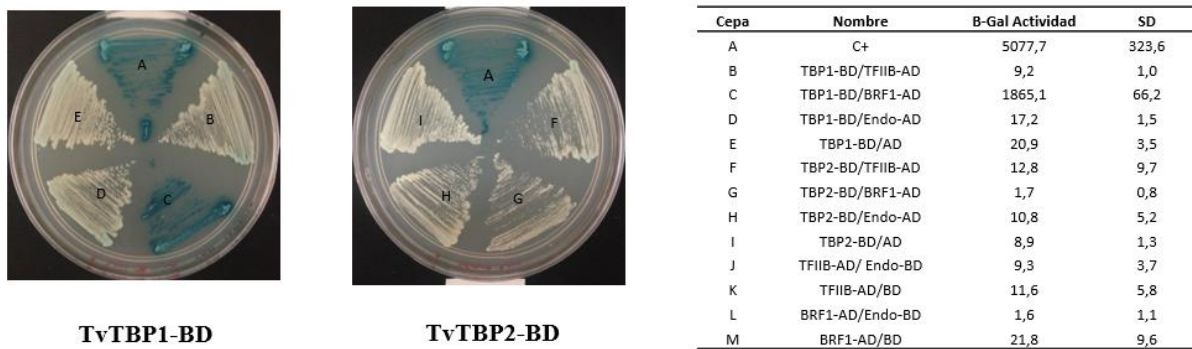
El análisis bioinformático de interacción proteína-proteína (Figura 10) y diversos artículos publicados (*Hahn and Young, 2011; Vannini and Cramer, 2012; Gietl y Col., 2014*) nos sugieren que TBP es capaz de interactuar con los factores de transcripción TFIIB, en los promotores transcritos por la RNA Pol II, y con BRF1, en los promotores transcritos por la RNA Pol III. Con el fin de analizar estas interacciones realizamos un ensayo *in vivo* de Doble Híbrido en levadura. Este ensayo se basa en analizar la expresión de un gen reportero regulado por un promotor que requiere reconstruir un factor transcripcional uniendo sus dominios de activación y de unión de DNA a través de la interacción de dos proteínas. El sistema de doble híbrido empleado en este ensayo evalúa la interacción registrada a través de la activación del operador LexA de dos genes reporteros *LEU2* y *LacZ* en una levadura auxótrofa para leucina. Si la interacción entre dos proteínas es positiva, las levaduras podrán crecer en un medio sin leucina y con la presencia de XGal se observarán colonias azules. Los primeros ensayos de Doble Híbrido aquí reportados se realizaron en colaboración con el Biólogo Erik Amador Gaytan.

Para llevar a cabo el experimento anterior, se clonaron los genes *TvTBP1* y *TvTBP2* con el dominio de unión a DNA (BD, por sus siglas en inglés *Binding Domain*) de LexA. Paralelamente se clonaron los genes putativos *TvTFIIB* y *TvBRF1*, con el dominio de activación (AD, por sus siglas en inglés *Active Domain*) de LexA. Como control positivo del ensayo se monitoreó la interacción entre dos subunidades de la proteína G ( $\alpha$  y  $\beta$ ), como control negativo se evaluó la interacción de las *TvTBPs* con una proteína no relacionada de *S. cerevisiae*, la endoquitinasa (*CTS1*). Además, se evaluó la interacción de cada construcción con únicamente el dominio opuesto. Se evaluó el crecimiento de las cepas de levadura que llevan estos plásmidos en diferentes medios de crecimiento: 1) sin leucina para evaluar el reportero de *LEU2*, 2) sin leucina y con XGal para evaluar ambos reporteros, y 3) con leucina y XGal para evaluar el reportero de *LacZ* y la viabilidad de las colonias (Figura 25).

i)



ii)

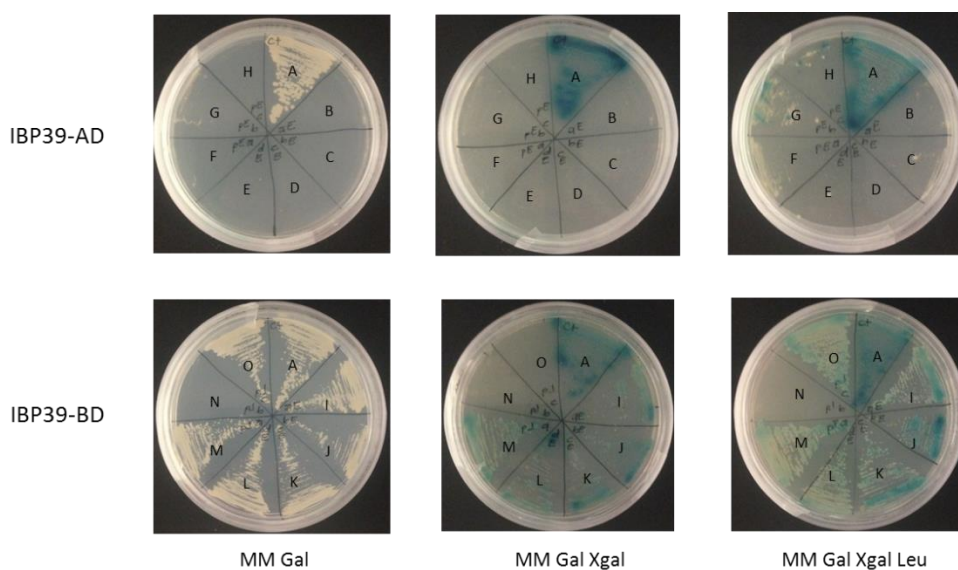


**Figura 25.** Evaluación de la interacción entre las TvTBPs con factores transcripcionales putativos de *T vaginalis*, por ensayo de doble híbrido. i) Se evaluó el crecimiento de las siguientes construcciones: A) Control positivo,  $\alpha$ -BD/ $\beta$ -AD subunidades de la proteína G, B) TvTBP1-BD/TvTFIIB-AD, C) TvTBP1-BD/TvBRF1-AD, D) TvTBP1-BD/Endoquitinasa-AD. E) TvTBP1-BD/AD, F) TvTBP2-BD/TvTFIIB-AD, G) TvTBP2-BD/TvBRF1-AD, H) TvTBP2-BD/ Endoquitinasa-AD, I) TvTBP2-BD/AD, J) TvTFIIB-AD/Endoquitinasa-BD, K) TvTFIIB-AD/BD, L) TvBRF1-AD/ Endoquitinasa-BD, M) TvBRF1-AD/BD. El reportero *LEU2* se evalúa en los medios mínimos sin leucina (MM Gal y MM Gal Xgal) y el reportero de *LacZ* y la viabilidad de las colonias se evalúan en el medio mínimo GAL suplementado con Xgal y leucina. ii) Cuantificación de la actividad de  $\beta$ -Gal. El ensayo de  $\beta$ -galactosidasa muestra el promedio de dos clones independientes. El crecimiento y el color azul de las colonias de la construcción C) TvTBP1-BD/TvBRF1-AD indican interacción, así como el aumento de la actividad de  $\beta$ -galactosidasa.



La figura 25 muestra únicamente una interacción de las TvTBPs, se observa que la cepa que lleva las construcciones TvTBP1-BD/TvBRF1-AD (cepa C) crece bien en un medio sin leucina y las colonias presenta coloración azul en los medios que contienen XGal, es decir hay expresión de los reporteros, indicando que TvTBP1 y TvBRF1 pueden interactuar *in vivo*. Para determinar cuantitativamente la interacción se llevaron a cabo ensayos para medir la actividad de  $\beta$ -Galactosidasa, demostrando un incremento en la actividad que lleva dichos plásmidos. También, se evaluó el poco crecimiento que se observa en los medios sin leucina con los controles (Figura 25: i, K y M) y se determinó que es mínima la interacción, la cual se considera fondo de interacción. Por otro lado, como se observa, no se encontró interacción de TvTBP2 con BRF1 o TFIIB (Figura 25); sin embargo, también se realizó la cuantificación de la actividad de  $\beta$ -Galactosidasa en todas las construcciones, corroborando que no hay interacción.

De la misma manera, para analizar la interacción de las TvTBPs con IBP39 se realizó una construcción para clonar a IBP39 junto con el dominio de activación de LexA; sin embargo, las colonias no fueron viables (resultado no mostrado). Tampoco fueron viables las cepas que llevan la construcción IBP39-AD/Endoquitinasa-BD e IBP39-AD/BD. Sin embargo, las colonias que llevan la construcción IBP39-BD/Endoquitinasa-AD ó IBP39-BD/AD crecen en un medio sin leucina y son azules con la presencia de Xgal (Figura 26), es decir hay expresión de los reporteros, lo que indica interacción y la posibilidad de que IBP39 esté activando la transcripción en *S. cerevisiae*.

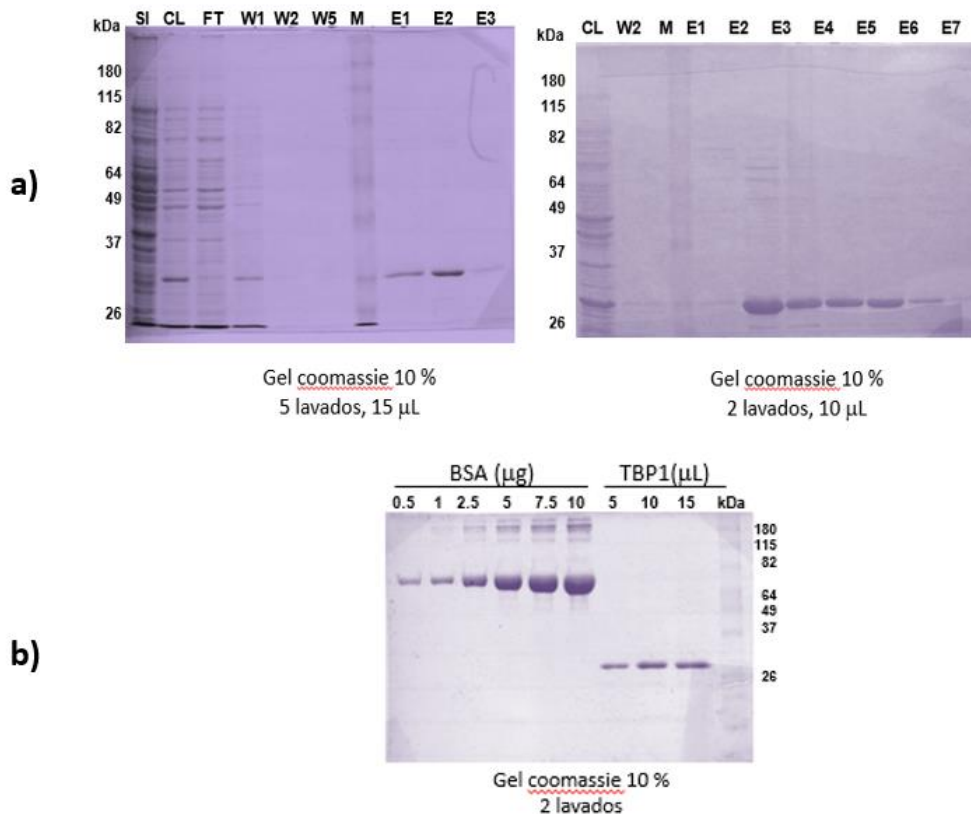


**Figura 26.** Evaluación de la interacción entre IBP39 con los controles para el ensayo de doble híbrido. Se evaluó el crecimiento de las siguientes interacciones: A) Control positivo,  $\alpha$ -BD/ $\beta$ -AD. B, C, D, y E) cuatro clones de IBP39-AD/Endoquitinasa-BD. F, G y H) tres clones IBP39-AD/BD. I, J, K y L) tres clones de

IBP39-BD/Endoquitinasa-AD. M, N y O) tres clonas IBP39-BD/AD. El reportero *LEU2* se evalúa en los medios mínimos sin leucina (MM Gal y MM Gal Xgal). El reportero de *LacZ* y la viabilidad de las colonias se evalúan en el medio mínimo Gal suplementado con Xgal y leucina. Se observa que aquellas colonias que llevan la construcción IBP39 con el dominio de activación no crecen en el medio suplementado con leucina, es decir, no son viables. Aquellas construcciones de IBP39 que llevan el dominio de unión (BD) al DNA crecen en un medio sin leucina (a excepción de la clona N que no es viable) y son colonias azules con la presencia de Xgal, lo que indica que son capaces de activar la transcripción de los genes reporteros.

## 6. Generación de anticuerpo TvTBP1.

Generamos un anticuerpo contra la proteína TvTBP1 con el fin de obtener una herramienta molecular que nos permitiera continuar con el estudio de las TvTBPs.



**Figura 27.** Purificación y cuantificación de la proteína recombinante TvTBP1 para la generación de anticuerpos. a) Purificación de la proteína recombinante TvTBP1 en columnas de níquel, se compara realizando de 2 a 5 lavados de la columna. SI: Sin inducir, CL: Células lisadas, FT: Fluido no unido a la columna, M: Marcador preteñido, W: Lavados, E: Eluidos. b) Cuantificación de la proteína recombinante TvTBP1 (27 kDa) para la inmunización de conejos y generación de anticuerpos. La cuantificación se realizó utilizando como patrón estándar una curva de calibración con albúmina sérica bovina (BSA) que tiene un peso de 66.4 kDa.

Se indujo la expresión de la proteína TvTBP1-6xHis, contenida en la cepa M15 de *Escherichia coli*, por 5 horas con 1 mM de IPTG (Franco 2012, Tesis doctoral). Posteriormente se realizó la purificación de la proteína recombinante por cromatografía



de afinidad utilizando columnas de níquel que permiten la purificación por unión a la etiqueta de histidinas en el extremo amino terminal de la proteína. Se determinó que dos lavados eran óptimos para obtener mayor cantidad de proteína recombinante (Figura 27-a). Para la inmunización se usó una mezcla de la segunda elución de cuatro purificaciones independientes que provenían de 250 mL de cultivo cada una. Por medio de una curva de calibración con albúmina sérica bovina (BSA) se realizó la cuantificación comparativa y se determinó que en 10  $\mu$ L del eluido de TvTBP1 recombinante hay aproximadamente 1  $\mu$ g de proteína (Figura 27-b).

Con la proteína recombinante TvTBP1-6xHis purificada se inmunizaron dos conejos New Zealand, de ocho semanas de edad. Se inyectaron tres inmunizaciones de ~250  $\mu$ g de la proteína recombinante por vía subcutánea con intervalos de dos meses, la primera con adyuvante completo de Freund y las posteriores con adyuvante incompleto (*Ausubel y Col., 1995*). Se obtuvieron 59.3 mL de suero inmunizado y 1 mL de suero pre-inmune del primero conejo; 60.1 mL de suero inmunizado y 200  $\mu$ L de suero pre-inmune del segundo conejo. Este anticuerpo se utilizó para corroborar la expresión de la proteína TvTBP1 en extractos nucleares de *T. vaginalis*, en estudios posteriores en el laboratorio por la M. en C. Olivia Parra Marín, descritos en el artículo adjunto.

## VII. DISCUSIÓN.

El estudio *in silico* de las proteínas hipotéticas TBP de *T. vaginalis* sugiere características particulares. Por un lado, el estudio apunta a que pertenecen a la familia de las proteínas TBP (*Carlton y Col., 2007 y Franco 2012, Tesis doctoral*); y, por otro lado, información adicional sugieren que las TvTBPs cuentan con características particulares, propias de estas proteínas. Se destaca las diferencias observadas en la región de unión con el DNA donde encontramos sustitución de algunos aminoácidos relevantes para la interacción con las cajas TATA (Figura 8). Así como se ha visto en otros protistas, incluyendo *G. intestinalis*, *T. brucei*, *T. cruzi* y *L. major*, donde hay reemplazo de algunas de las fenilalaninas críticas para la interacción con el DNA, estos organismos usan diferentes estrategias de ensamble del PIC, que dependen de TBP, pero no de interacciones con la caja TATA. Y la interacción con el DNA puede depender de la presencia de otros factores transcripcionales (*Koster y Col., 2015*). También es interesante la manera independiente en que las TvTBPs se agrupan en el árbol de identidad (Figura 7). Esto aunado a que la transcripción de *T. vaginalis* difiere a lo descrito en organismos canónicos como lo refleja la presencia de *Inr* e *IBP39* (*Liston y Col., 2001*) y la ausencia de cajas TATA en los promotores (*Liston y Col., 1999*), pone de manifiesto la particularidad de las TvTBPs.

En el presente trabajo se describen los primeros estudios experimentales *in vitro* realizados para la búsqueda de interacción entre las TvTBPs con secuencias de DNA, mostrando que las TvTBPs no se unen a estas bajo las condiciones aquí probadas. En este ensayo utilizamos sondas marcadas con radioactividad en un ensayo EMSA que permitió observar la interacción de la proteína control ScTBP de *S. cerevisiae* con DNA de levadura que contenía una caja TATA canónica y con otras dos sondas de *T. vaginales*, sin cajas TATA, pero con la presencia de regiones ricas en TA (Figura 22). Aunque el resultado del control (ScTBP) sugiere que las condiciones en el ensayo EMSA presentaban un ambiente propicio para la interacción, en este ensayo no se observó unión de las TvTBPs con el DNA, aún con la presencia de los factores transcripcionales putativos TvTFIIB y TvBRF1 (Figura 23). No obstante, la elección de una estrategia diferente en un ensayo con sondas sin marcaje radioactivo revelado con bromuro de etidio y visualizado con luz UV, el aumento significativo en la cantidad de proteína usada y el uso de componentes distintos en las soluciones de interacción permitió que, en experimentos posteriores a este trabajo realizados por mi compañera M. en C. Olivia

Parra, se demostrara una clara interacción TvTBPs–DNA independiente de secuencia (Parra-Marin y Col., 2019).

La unión TvTBPs-DNA requirió de condiciones específicas de interacción, que no se obtuvieron en el primer ensayo, mostrando que bajo algunas condiciones pudiera existir el reconocimiento y la unión, por lo menos *in vitro*. Siendo así, este último ensayo muestra un resultado preliminar para dilucidar la interacción *in vivo* de las TvTBPs con regiones particulares en el DNA, que pudiera llevarse a cabo a través de un proceso no canónico. Aunado a que ahora sabemos que las TvTBPs no tienen la misma función que una TBP canónica, como lo mostramos aquí, TvTBP1 y TvTBP2 no pudieron sustituir la ausencia de TBP en levadura (Figura 14). Sin embargo, con este ensayo de complementación no podemos determinar si las proteínas fueron capaces de unirse a secuencias en *cis* en el DNA, o si reconocieron motivos, pero no pudieron reclutar a la maquinaria de transcripción.

Adicionalmente, la clara interacción de TvTBP1 con TvTBRF1 revelada con los ensayos de doble híbrido en levadura (Figura 25) sugiere la participación de TvTBP1 en la transcripción de aquellos genes cuyos promotores son transcritos por la RNA Pol III. Por otro lado, no se encontró interacción de TvTBP1 con TvTFIIB, tampoco de TvTBP2 con TvTBRF1 o con TvTFIIB, poniendo de manifiesto la probable participación diferencial en la transcripción de ambas TvTBPs (Figura 25).

Un hallazgo muy interesante se observó cuando tratamos de realizar los ensayos de doble híbrido para probar la interacción de las TvTBPs con IBP39. Cabe resaltar que IBP39 es un factor transcripcional específico de *T. vaginalis*, con dominios de activación y unión al DNA propios, y se resalta que no hay ortólogos aparentes de IBP39 en levadura ni en otros organismos (Liston y Col., 2001). Al llevar a cabo los ensayos control, donde se analizan las posibles interacciones de las proteínas con los dominios solos, se observó que IBP39 podría estar interaccionando con factores transcripcionales de *S. cerevisiae*. El ensayo de IBP39-BD con los controles reveló la expresión de los genes reporteros (Figura 26), es probable que IBP39 esté activando la transcripción en este sistema heterólogo reclutando a la maquinaria transcripcional de la levadura. Y, por otro lado, quizá alterando la transcripción por unión al DNA (con su propio dominio) cuando lleva el AD, ya que esta construcción alteró la viabilidad de las colonias (Figura 26). Estos resultados podrían deberse a la capacidad que tiene IBP39 como un importante factor transcripcional, reconocido en *T. vaginalis*.

Por otro lado, la presencia de dos proteínas TBPs en *T. vaginalis* no es un hallazgo sorprendente. Como se ha visto, muchos organismos eucariotes codifican para múltiples parálogos de TBP (proteínas TBP, proteínas TBP-like o TRFs), algunos de ellos son capaces de unirse a cajas TATA clásicas y otros perdieron su capacidad de unión tal vez debido a la pérdida de las dos fenilalaninas requeridas para el reconocimiento de la caja TATA (Koster y Col., 2015). Y como explico Koster y Col., la divergencia en la estructura, expresión y función de los homólogos de TBP podría explicar su retención evolutiva (Koster y Col., 2015). Aunque aparentemente la estructura y la expresión de ambas TvTBPs es similar, esperaríamos que la retención de ambas proteínas en este organismo se debiera a que tendrán funciones distintas en la regulación de la transcripción de *T. vaginalis*. Asimismo, nuestros experimentos indican que la presencia de los aminoácidos más importantes para la unión con una caja TATA canónica, en la proteína TvTBP1 modificada que construimos, no es suficiente para observar interacción con el DNA (Figura 23).

En resumen, en conjunto estos ensayos y la información que tenemos hoy en día del proceso de transcripción de *T. vaginalis* nos lleva a pensar que las proteínas TvTBPs sí participan en la formación del PIC en este organismo, aún sin la presencia de cajas TATA, probablemente como parte del complejo de proteínas de andamiaje y quizá en algunas ocasiones reconociendo regiones, aún no descritas, en el DNA para reclutar al resto de la maquinaria transcripcional. Este trabajo representa el estudio de una de las proteínas más estudiadas e importantes para la transcripción, las TBPs, pero enfocado en un organismo interesante que divergió temprano en la evolución y que tiene características únicas como lo es *T. vaginalis*.

## VIII. CONCLUSIONES.

Experimentalmente, demostramos que los genes de TvTBPs se expresan en *T. vaginalis* como RNAm.

En este trabajo se realizaron los primeros ensayos de interacción TvTBPs-DNA, determinando que las TvTBPs no se unen a una caja TATA típica descrita para levadura, tampoco a las secuencias promotoras de DNA de *T. vaginalis* probadas en este ensayo, al menos en las condiciones aquí experimentadas. Este ensayo preliminar permitió realizar experimentos posteriores empleando una estrategia de unión diferente y mostrando que ambas TvTBPs son capaces de unirse a regiones de DNA de manera inespecífica e independientemente de la presencia de una caja TATA. Este ensayo se realizó en nuestro laboratorio por la M. en C. Olivia Parra Marín, con las sondas y proteínas descritas en el presente trabajo (rDNA, SCS, 5S, U6 y TvTBP2), publicado en *Parasitology Research* (2019) y en el cual comparto autoría. El artículo está incluido al final en esta tesis.

Las TvTBPs no fueron capaces de complementar una mutante de TBP en *S. cerevisiae*, por lo que se considera que no poseen la misma función que una TBP canónica.

TvTBP1 es capaz de unirse al factor transcripcional TvBRF1 *in vivo*, pero no a TFIIB. Esto sugiere que la TvTBP1 participa en la transcripción de genes transcritos por la RNA Pol III en *T. vaginalis*. No obstante, no se observó interacción de TvTBP2 con TvBRF1 ni tampoco con TFIIB.

Finalmente, los anticuerpos anti TvTBP1 generados en el presente trabajo permitieron que se demostrara la expresión de la proteína TvTBP1 directamente en *T. vaginalis*, experimento que se llevó a cabo posteriormente en el laboratorio por la M. en C. Olivia Parra.

## IX. PERSPECTIVAS

Hoy en día, continuamos con preguntas sin resolver interesantes para continuar el estudio de la transcripción en *T. vaginalis*: ¿Las TvTBPs se requieren para la transcripción de *T. vaginalis*? ¿Por qué hay dos TBPs en *T. vaginalis*? ¿Hay regiones específicas en el genoma de *T. vaginalis* que las TvTBPs pueden reconocer y unirse?

Consideramos que en este trabajo la selección de los promotores elegidos para el estudio de unión TBP-DNA es apropiada. Seleccionamos un promotor canónico de *S. cerevisiae* y cuatro promotores de *T. vaginalis* de genes transcritos por las tres RNA polimerasas, cuya transcripción ha sido demostrada. Sin embargo, el número de genes en *T. vaginalis* es amplio y continúa latente la posibilidad de que las TvTBPs puedan unirse a otras regiones en el DNA de *T. vaginalis*, inclusive regiones sin una caja TATA canónica, como se ha observado en estudios recientes que ponen en evidencia que los mecanismos de inicio de la transcripción son más diversos que el modelo del libro (*Ahn y Col., 2012*). Por lo tanto, es necesario continuar con el estudio de la interacción TvTBPs-DNA.

Un siguiente enfoque para continuar con el estudio de las TvTBPs podría estar direccionado a estudios *in vivo* para analizar su interacción con el genoma de *T. vaginalis*, y así para poder dilucidar su participación en la transcripción. El ensayo de inmunoprecipitación de la cromatina (ChIP) es un método usado para determinar la localización en el genoma de histonas y proteínas *in vivo* (*Mundade y Col. 2014*) que se podría utilizar para encontrar interacciones TBP-DNA en *T. vaginalis*. Tal como se describió anteriormente para la TBP de humanos, donde se encontraron, además de los esperados, sitios de unión diferentes (*Denissov y Col. 2017*).

Asimismo, se ha puesto de manifiesto la utilización de una de las herramientas moleculares más importantes de nuestra época, en el 2018 se publicó la primera referencia del uso de CRISPR/Cas9 para realizar *knock out* en genes de *T. vaginalis* (*Janssen 2018*). Esta herramienta puede ser de utilidad para estudiar la función de los genes en *T. vaginalis* que codifican factores de transcripción putativos aún no descritos.

## X. MATERIALES Y MÉTODOS

### ***Análisis Bioinformáticos.***

Para obtener las secuencias de las TvTBPs en la base de datos de NCBI realicé una búsqueda en el genoma de *T. vaginalis* con el programa BLAST (por sus siglas en inglés: *Basic Local Alignment Search Tool*) (Altschul y Col., 1990) de la base de datos NCBI (por sus siglas en inglés: *National Center for Biotechnology Information*), con la secuencia de proteínas TBP de *Homo sapiens* (NCBI Reference Sequence: NP\_003185.1/UniProtKB: P20226.2) y *Saccharomyces cerevisiae* (NCBI Reference Sequence: NP\_011075.3/UniProtKB: P1339.3). Ambas búsquedas resultan en la identificación de dos proteínas putativas en el genoma de *T. vaginalis*: 1) NCBI Reference Sequence: XP\_001329427.1/ UniProtKB: A2DQT7 y 2) NCBI Reference Sequence XP\_001300812.1/UniProtKB: A2G4NT. El modelaje de la estructura terciaria de las TvTBPs se realizó con el programa Phyre2 (Kelley y Col., 2009) (datos no mostrados) y RaptorX (Kallberg y Col., 2012), ambos programas concuerdan en la estructura teórica. Para realizar el árbol de identidad, las secuencias se alinearon con el programa Clustal W (Sievers y Col., 2011) y la edición del árbol con el programa TreeView (Win32) 1.6.6 (Page, 1996). La matriz del porcentaje de identidad entre las TBPs fue creada con ClustalW (Sievers y Col., 2011). Las secuencias se alinearon con el programa GeneDoc (Nicholas KB, Nicholas HB Jr., 1997).

### ***Cultivo celular***

El cultivo celular de *T. vaginalis* utilizado en este trabajo para la extracción de DNA y RNA total, proviene del aislado CNCD147. El cultivo de parásitos se realizó a 37 °C, en medio TYM (Trypticasa-Extracto de levadura-Maltosa) suplementado con suero de caballo inactivado por calor (GIBCO) (Diamond, 1957).

### ***RT-PCR***

Para determinar la expresión de los genes TVAG\_291560 y TVAG\_285070 (TvTBP1 y TvTBP2, respectivamente) se realizó una reacción de RT-PCR (del inglés: *Reverse Transcription Polymerase Chain Reaction*). Crecimos las células a una densidad aproximada de  $1.5 \times 10^6$  células/mL. Se extrajo RNA total de acuerdo con protocolo de extracción con Trizol (Invitrogen). La fracción de RNA poly(A) se separó utilizando el kit

*illustra™ mRNA Purification, GE Healthcare* que contiene oligodeoxiribonucleótido (oligo)-dT. Para generar cDNA, se usaron 500 ng aproximadamente de RNA poly(A), 200 ng del oligonucleótido AP (que es complementario a la cola de poly(A) y añade una etiqueta al extremo 5') y una mezcla de dNTPs 1mM. La reacción se incubó a 65 °C por 5 min. y 1 min. en hielo. Enseguida se añadieron 100 unidades de *SuperScript™ III reverse transcriptase* (Invitrogen), *first strand buffer* 1X y DTT 0.05 M y se incubó a 42 °C por 15 min., seguido de una segunda incubación a 50 °C por 30 min. La reacción se detuvo incubando a 70 °C. Se añadió 1 unidad de RNase H (Invitrogen) y la reacción se incubó a 37 °C por 35 min. Dos reacciones paralelas como control se llevaron a cabo de la misma manera y con los mismos reactivos, excepto una de ellas sin RNA total y la segunda sin *SuperScript™ III reverse transcriptase*.

Con cada una de las tres muestras anteriores realizamos reacciones de PCR con dos parejas de oligonucleótidos. Un oligonucleótido específico *forward* que alinea con la región 5' de cada TvTBP y el oligonucleótido AUAPB *reverse* que es complementario a la etiqueta, que añadió el oligo AP, encontrada después de la cola de poly(A) en el cDNA sintetizado. Un oligonucleótido específico *forward* que alinea con la región 5' de cada TvTBP y un oligonucleótido específico *reverse* que alinea con la región 3' de cada TvTBP. La ausencia en los controles indica que los amplicones observados derivan de transcritos de RNA (Figura 11). Los amplicones que provienen de la amplificación con el oligonucleótido AUAPB los clonamos en pGEM Easy-T y se determinó su secuencia. Las secuencias obtenidas (Figura 12), corresponden a los genes TvTBP1 y TvTBP2, respectivamente (ver oligonucleótidos en material suplementario).

### **Complementación en levadura.**

TBP es un gen esencial en levadura, para realizar el ensayo de complementación se requirió trabajar con una cepa comercial de *S. cerevisiae*, diploide y heterocigota. Lleva un *knock out* para TBP en uno de los alelos, derivada de la colección *Yeast Knock-out collection of OpenBiosystems, catalog No. YSC1021-674345, Thermo Scientific Life Science Research Products*. Esta cepa se construyó reemplazando un alelo de TBP con el cassette de KanMX que confiere resistencia al antibiótico G418. Esta mutación se realizó en la cepa comercial BY4743, *MATa/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 MET15/met15Δ0 ura3Δ0/ura3Δ0*.



Amplificamos por PCR los genes de TvTBP1, TvTBP1\*, TvTBP2 y *SPT15* (ScTBP) a partir de DNA de *T. vaginalis* y de DNA de *S. cerevisiae*, respectivamente, con oligonucleótidos específicos (Ver anexos) para clonarlos en el vector de expresión pYES2 (2  $\mu$ m, *URA3*) inducible con galactosa, contiene el gen que confiere resistencia a ampicilina para la selección de transformantes positivas en *E. coli*. De manera independiente transformamos cada vector recombinante y el vector vacío como control negativo, en la cepa JM109 de *E. coli* para secuenciar, propagar y mantener el plásmido construido. Posteriormente lo transformamos en la cepa de levadura diploide mutante para TBP, descrita anteriormente. El crecimiento y manipulación de las levaduras se llevó a cabo de acuerdo con los procedimientos estándares. La selección de transformantes de levadura se realizó por prototrofia con uracilo, en cajas SC -U (*synthetic minimal plates without uracil*) que contenían galactosa al 2%, rafinosa al 0.2% suplementadas con histidina al 0.005 % y leucina 0.01%.

Indujimos a las levaduras transformantes a pre-esporular, en un medio de pre-esporulación (Bacto-extracto de levadura al 0.8%, Bacto-peptona al 0.3%, Dextrosa al 10%, Bacto-agar al 2%) por dos días, y posteriormente a esporular en un medio mínimo de esporulación el cual contiene únicamente acetato de potasio al 1 %. Durante la esporulación las células se dividen por meiosis, generando cuatro células haploides. Examinamos el proceso de esporulación por microscopía, para confirmar la producción de tétradas. Una vez que se obtuvimos suficientes tétradas, después de 1 semana aproximadamente, proseguimos a la disección de tétradas. Con la solución Sigma  $\beta$ -Glucuronidasas rompimos el asca que cubre las tétradas. Realizamos la disección de tétradas con un micromanipulador en una superficie de agar al 2%, posteriormente la trasladamos a una caja con medio YPG que fue incubada por 2 días a 30 °C. Las colonias viables se analizaron, determinamos su haploidía creciéndolas en diferentes cajas de medio mínimo SC, suplementadas con histidina, leucina, y algunas con metionina y lisina. Aproximadamente 30 colonias haploides de cada transformación se crecieron en un medio mínimo suplementado con el antibiótico G418.

El genotipo determinado de las colonias estriadas en la figura 14 es el siguiente: a) *his3 $\Delta$ 1 leu2 $\Delta$ 0 lys2 $\Delta$ 0 met15 $\Delta$ 0 [pYES-TvTBP1]*, b) *his3 $\Delta$ 1 leu2 $\Delta$ 0 [pYES-TvTBP1]*, c) *his3 $\Delta$ 1 leu2 $\Delta$ 0 lys2 $\Delta$ 0 met15 $\Delta$ 0 [pYES-TvTBP2]*, d) *his3 $\Delta$ 1 leu2 $\Delta$ 0 [pYES-TvTBP2]*, e) *his3 $\Delta$ 1 leu2 $\Delta$ 0 lys2 $\Delta$ 0 met15 $\Delta$ 0 spt15::Kan [pYES-SPT15]*, f) *his3 $\Delta$ 1 leu2 $\Delta$ 0 lys2 $\Delta$ 0 met15 $\Delta$ 0 [pYES]*.

### **Ensayo Western blot, HisProbe-HRP.**

Después de la electroforesis se realizó la transferencia de las proteínas a una membrana de nitrocelulosa por 1 hora a 350 mA, en un buffer de transferencia que contenía 840 mg de bicarbonato de sodio (NaHCO<sub>3</sub>), 318 mg de carbonato de sodio (NaCO<sub>3</sub>) y 200 mL de metanol, contenido en 1000 mL. Posteriormente, la membrana se bloqueó con leche *Svelty* al 5 % en TTBS (TTBS, 1000 mL: 2.42 g de Tris y 29.2 g de NaCl a un pH 7.5, más 1.0 mL de Tween 20 (0.1%)). Se enjuagó 2 veces con 20 mL aproximadamente de TTBS y se incubó en la misma solución 1 vez por 15 min. y 2 veces por 10 min., realizando cambios de la solución de TTBS en cada enjuague. A continuación, se incubó 1 hora con 10 mL de *HisProbe-HRP* 1:5000 en TTBS. Al finalizar la incubación se realizaron lavados a la membrana, se enjuagó 2 veces con TTBS y se incubó 3 veces por 10 min., cada vez, con 20 mL de la misma solución. Finalmente, la interacción de la sonda His (*Thermo Scientific™ HisProbe™-HRP Conjugate*, catálogo: 15165) con el tag de histidinas en las proteínas recombinantes se reveló con una solución 1:1 de Pico West de Pierce, por 5 min. a temperatura ambiente y exponiendo la membrana en una película de rayos X a diferentes tiempos. Protocolo proporcionado por el laboratorio del Dr. Georges Dreyfus, IFC UNAM.

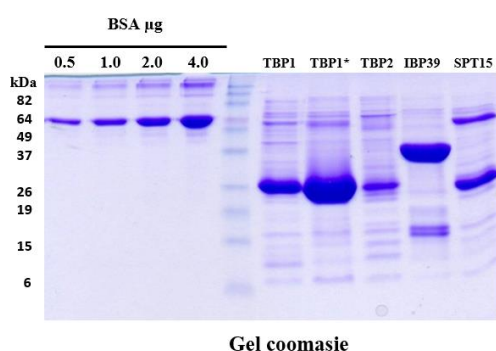
### **Ensayos EMSA (Electrophoretic Mobility Shift Assays).**

Las regiones promotoras seleccionadas para ser usadas como sondas en los ensayos EMSA fueron amplificadas por PCR, clonadas en el vector pGEM®-T Easy Vector System (Promega) y la secuencia verificada por secuenciación. Los fragmentos clonados fueron posteriormente liberados por digestión con la enzima *EcoRI* (Invitrogen) y purificados. Para generar las sondas, estos fragmentos de DNA se marcaron [ $\alpha$ -<sup>32</sup>P]dATP en sus extremos por medio de la reacción *fill-in* con el fragmento Klenow de la DNA Polimerasa I (Invitrogen).

Sonda	Tamaño	Secuencia promotora	Organismo	Transcrito por polimerasa	Caja TATA Contexto
Pro-rDNA	164 pb	Cistrón ribosomal (18S, 5.8S y 28S).	<i>T. vaginalis</i>	RNA Pol I	No definida
Pro-SCS	82 pb	Succinil CoA sintetasa.	<i>T. vaginalis</i>	RNA Pol II	No
Pro-5S	186 pb	RNA ribosomal 5S.	<i>T. vaginalis</i>	RNA Pol III	No definida
Pro-U6	120 pb	snRNA U6 (Small nuclear RNA U6).	<i>T. vaginalis</i>	RNA Pol III	Si
Pro-His4	172 pb	His4	<i>S. cerevisiae</i>	RNA Pol II	Si

Tabla 4. Las regiones promotoras seleccionadas para ser usadas como sondas en los ensayos EMSA.

El ensayo EMSA se realizó usando las sondas descritas anteriormente y ~100 ng (Figura 28) de las proteínas recombinantes purificadas TvTBP1, TvTBP1\*, TvTBP2, IBP39 y ScTBP (como control). Se incubaron 20  $\mu$ L de reacción, por 30 min. a temperatura ambiente en un buffer de unión que contiene: 2  $\mu$ g de suero de albumina bovina (BSA), glicerol al 4%, KCl 60 mM, Tris-Cl 5 mM pH 8, MgCl<sub>2</sub> 5 mM, DTT 1 mM, y 10,000 cpm de sonda marcada. Los complejos proteína-DNA fueron analizados por electroforesis a través de condiciones no desnaturizantes en geles de poliacrilamida al 5 % (acrilamida 5 % a partir de un stock acrilamida:bisacrilamida 20%:0.33%, glicerol 2.5 %, glicina 190 mM, acetato de magnesio 10 mM, Tris 2.5 mM pH 8.3, EDTA 1 mM y DTT 0.5 mM) y buffer de corrida TG (Tris-Cl 2.5 mM pH 8.3, glicina 190 mM, EDTA 1 mM, y acetato de magnesio 5 mM) a 100 V por ~ 90 min. Los geles fueron pre-corridos a 100 V por ~ 1 hora antes de ser utilizados. El colorante azul de bromofenol fue añadido a los carriles que contenían DNA únicamente, porque es conocido que este reactivo puede interferir con la unión DNA-Proteína. Los geles se secaron y se expusieron a *Phosphor Imager screen* por 6 horas aproximadamente (*modificado de Sprouse y Col., 2009*).



**Figura 28.** Cuantificación relativa de proteínas recombinantes para ensayo EMSA. En los primeros 4 carriles del gel se presenta una curva de cuantificación con cantidades crecientes de BSA que va de 0.5 a 4  $\mu$ g. El carril 5 muestra el marcador de peso de proteínas. En los carriles de 6 al 10 se observa la expresión de las proteínas usadas para el ensayo EMSA, la banda más abundante de aproximadamente 27 kDa representa a las proteínas recombinantes TBPs, y de 40.6 kDa a IBP39. Para cada ensayo se emplearon 100 ng aproximadamente.

### ***Ensayo de Doble Híbrido.***

El sistema de dos híbridos de levadura LexA-B42 se usó para probar la interacción entre proteínas. Los plásmidos utilizados fueron pEG202, que porta el dominio de unión al DNA (BD) de la proteína LexA, y pJG4-5 que porta el dominio activo de transcripción (AD). Los productos de PCR de los genes BamHI-XhoI TvTBP1, EcoRI-NotI TvTBP2, EcoRI-XhoI TvTFIIB y EcoRI-XhoI TvBRF1 se clonaron en pGEM®-T Vector System (Promega) y luego se subclonaron en los plásmidos para el sistema de interacción de dos híbridos. TvTBP1 y TvTBP2 se clonaron en el plásmido pEG202 (BD), y TvTFIIB y TvBRF1 se clonaron en pJG4-5 (AD). Se usó *E. coli* XL1 blue para propagar los plásmidos. Las subunidades  $\alpha$  y  $\beta$  de la proteína G (LexA-G $\alpha$  con B42-G $\beta$ ) de *S. cerevisiae* se utilizaron como control positivo (Kawasaki y Col., 2005). El gen endoquitinasa de *S. cerevisiae* (CTS1) clonado en pGE202 (Ongay-Larios y Col., 2000) o pJG4-5 (Kawasaki y Col., 2005) y los plásmidos parentales se usaron como controles negativos. La cepa EGY48 de *S. cerevisiae* (MAT $\alpha$ , his3, trp1, ura3-52, leu2::pLeu-LexAop6) (Golemis y Col., 1997) se utilizó como receptor para los experimentos de interacciones de proteínas y se analizó con el sistema de doble híbrido LexA-B42, determinado por la expresión del gen indicador LacZ ubicado en el plásmido pSH18-34 (Ongay-Larios y Col. 2000). La interacción de las proteínas analizadas fue revelada por la presencia de colonias azules; cinco clones de cada ensayo se sembraron en medio SGal mínimo (0,67% de base de nitrógeno de levadura sin aminoácidos (Difco, Detroit, MI) y 2% de galactosa) suplementado con 0,1 mg / ml de leucina y 40  $\mu$ g / ml de X-Gal (5- bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosido). Las placas se incubaron durante 72 horas a 30 ° C. La figura 25 muestra solo un clon de cada interacción. La actividad de la  $\beta$ -galactosidasa se cuantificó como se describe Ongay-Larios y Col., 2000.

### ***Preparación del anticuerpo TvTBP1.***

El anticuerpo policlonal de conejo anti TvTBP1 se produjo por inmunización de dos conejos New Zealand con la proteína recombinante purificada TvTBP1-6xHis. Los conejos llegaron al bioterio con 8 semanas de edad, 2 semanas posteriores se obtuvo el suero pre inmune y posterior a 2 semanas más (a la edad de los conejos de 12 semanas) se comenzó el esquema de inmunización. Se inyectaron tres inmunizaciones de ~250 µg de cada una de las proteínas recombinantes por vía subcutánea con intervalos de dos meses, la primera con adyuvante completo de Freund y las posteriores con adyuvante incompleto (*Ausubel y Col., 1995*). Después de la tercera inmunización, en el bioterio se recolectó la sangre de los conejos y se permitió coagular durante 3 horas, se incubó toda la noche a una temperatura de 4 °C, se centrifugó a 4000 rpm durante 20 min a 4 ° C y realizaron alícuotas del suero que se mantuvieron a una temperatura de -70 ° C.

## VIII. ANEXOS

### 1. Lista de oligonucleótidos

Nombre	Secuencia	T <sub>m</sub> °C	Para el plásmido
ScTBP-BamHI-Fw	5' TATGGATCCATGGCCGATGAGGAACGTTTAAAG 3'	60.4	pYES2
ScTBP-EcoRI-Rv	5' CGCGCGAATTCTCACATTTTCTAAATTCACCTAGC 3'	59.2	pYES2
TBP1Fw (BamHI)	5' CCCGGATCCATGAATCCAAAAGAATTTATTGGC 3'	59.1	pQE30 and pYES2
TBP1-691-XhoI-Rv	5' CTAICTCGAGTTAAGCAGCTGGCGGTTTC 3'	60.5	pYES2
TBP2-1-BamHI-Fw	5' GATAGGATCCATGGATGCTGCGAGTTTCTTTG 3'	60.4	pYES2
TBP2-680-NotI-Rv	5' CTAGCGGCCGCTTATATAAGTAATTCGACATAG 3'	60.3	pYES2
AP primer	5' TTTTTTTTTTTTTTTTTCATGATCAGCTGCGCACCGG 3'	58.1	RT-PCR
AUAP-B primer	5' CATGATCAGCTGCGCACCGGCTAGG 5'	63.6	RT-PCR
TBP1-VTmut-Fw	5' CGGAAAAGTTGTTATCACAGGATGTAAGAAGC 3'	56.4	TvTBP1* (VT)
TBP1-VTmut-Rv:	5' GCTTTTACATCCTGTGATAACAACCTTTCCG 3'	56.4	TvTBP1* (VT)
TBP1-Fmut-Fw:	5' CCTGAAATTTTGCAGGTTTGATTTATAG 3'	51.9	TvTBP1* (F)
TBP1-Fmut-Rv:	5' CTATAAATCAAACCTGCAAAAATTTTCAGG 3'	51.9	TvTBP1* (F)
TBP1-VLFmut-Fw:	5' CCAAAGACGGTTTACTGATATTCGTTTCAGG 3'	57.8	TvTBP1* (VLF)
TBP1-VLFmut-Rv:	5' CCTGAAACGAATATCAGTAAAACCGTCTTTGG 3'	57.8	TvTBP1* (VLF)
TBP1 Rv (HindIII)	5' CCAAAGCTTTTAAGCAGCTGGCGGTTCTGTGC 3'	64.2	pQE30
TBP2-3-BamHI-Fw	5' GATAGGATCCGACGCTGCGAGTTTC 3'	58.9	pQE30
TBP2-680-KpnI-Rv	5' CGCGGTACCTTATATAAGTAATTCGACATA 3'	56.4	pQE30
IBP39ORF-Fw-BamHI	5' TATGGATCCGATCCCAACGACCTTGAAGC 3'	59°C	pQE30 and pGEX4T-1
IBP39ORF-Rv-KpnI	5' GCGGTACCTTACATTGGAGCGAAAGTAG 3'	59°C	pQE30
ScTBP-BamHI-Fw	5' TATGGATCCATGGCCGATGAGGAACGTTTAAAG 3'	60.4	pQE30
ScTBP-KpnI-RV	5' CGCGGTACCTCACATTTTCTAAATTCACCTAGC 3'	59.1	pQE30
TvTFIIB-BamHI-Fw	5' ATCGGATCCATGTCTCTTAAAGCAAAGAGG 3'	59.0	pGEX4T-1
TvTFIIB-SmaI-Rv	5' GCCCGGGTTATCTCTTTCATTTCTTCATCTTTC 3'	59.1	pGEX4T-1
IBP39ORF-XhoI-Rv	5' GATACTCGAGTTACATTGGAGCGAAAGTAGG 3'	59.1	pGEX4T-1
Pro-SchHis4-Fw	5' GGATATGACTATGAACAGTAGTATAC 3'	51.0	pGEM-T Easy
Pro-SchHis4-Rv	5' GTAGAATCGGCAAACCATTTATTCAGA 3'	52.2	pGEM-T Easy
Pro-TvrRNA-Fw	5' GCACATGTACGGAGTTATAATC 3'	49.4	pGEM-T Easy
Pro-TvrRNA-Rv	5' CATAGGAGCAAAAGGTCATG 3'	48.4	pGEM-T Easy
Pro-TvSCS-Fw	5' GTTGCGGGAGGTTTTAAACGATTAAC 3'	54.2	pGEM-T Easy
Pro-TvSCS-Rv	5' CTAAGCATTGTAATGTGAAGTGATCAAC 3'	53.1	pGEM-T Easy
Pro-Tv5S-Fw	5' TACTACACTCTACACCTATGTTTGTCCCTGC 3'	59.1	pGEM-T Easy
Pro-Tv5S-Rv	5' TGTGGCCGCTTCGAAAGGATTGTTTCG 3'	58.9	pGEM-T Easy
Pro-TvU6-Fw	5' ATTTCTGCAATTTTCCATGATATAATCC 3'	50.2	pGEM-T Easy
Pro-TvU6-Rv	5' TAATAGAATATGTTTGACAATGCTTTATAAAC 3'	50.1	pGEM-T Easy

## 2. Secuenciación de los plásmidos utilizados para expresar las proteínas recombinantes.

### 6xHis-TvTBP1

#### Secuencia

```
1  ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAATCCAAAAGAATTTATTGGC
1  M R G S H H H H H G S M N P K E F I G

61  TTAGAAGAAGCTAATATAAATAATGCAAAAGAATTCGAACAGCTTGTTAAAGAAGCTCAA
21  L E E A N I N N A K E F E Q L V K E A Q

121 GAGGAATCAGAGGAAGAAATTACACCAGAATCAATACAGAGAAATATAAGTTTAAATCAA
41  E E S E E E I T P E S I Q R N I S L N Q

181 CCTCCCATTCCCAAAGTTGTCAACATTGTTGCATGCGTAGATTACGGATGCAAACCTTGAC
61  P P I P K V V N I V A C V D Y G C K L D

241 CTTCTTAAGATTACATCTTCAACAAGAAATTCAGAATATAATCCGATTCGTTTCCCTGCT
81  L L K I T S S T R N S E Y N P I R F P A

301 GCTACGCTAAGGAACAAAGAACCAAAAGCCACAGCTTTAGCTTTTAAAAACGGAAAAGTT
101 A T L R N K E P K A T A L A F K N G K V

361 AATATCGTAGGATGTAAGCGCTGATATAGCATTAAAAGCTGCCGTTGCTTTGGTTAGA
121 N I V G C K S A D I A L K A A R C F G R

421 TTATTCAAAAATATCGGCATTAAGACGTTAAAATCAAATCATTCTCAATAGCTAACATG
141 L F K N I G I K D V K I K S F S I A N M

481 GTTGCAACAATGGACTGCAAATTCCCGATACATCTTGAATCAATTGCTTCATCGCCTGGT
161 V A T M D C K F P I H L E S I A S S P G

541 CATATTAAATTTGCTACATACAATCCTGAAATTTATGCAGGTTTGATTTATAGATTGGCT
181 H I K F A T Y N P E I Y A G L I Y R L A

601 AGTCCAAAGACGACTTTAATGATATACGTTTCAGGTAAAATATTATCACTGGAGCTAAA
201 S P K T T L M I Y V S G K I I I T G A K

661 TCAGAAGAAGAATTAAGCCGCATCTGCATTTATTTATCCTATTCTTAGATTGTTTGCT
221 S E E E L K A A S A F I Y P I L R L F A

721 AGCACAGAACCGCCAGCTGCTTAA
241 S T E P P A A *
```

**6xHis-TvTBP1.** Secuencia de TvTBP1 (TVAG\_291560) clonada en el vector pQE30 que adiciona una etiqueta de 6 histidinas en el extremo amino terminal. Integrada *Bam*HI-*Hid*III. En negritas el codón de inicio (ATG) y el codón de paro (TAA); en rojo la secuencia que codifica la etiqueta de 6 histidinas; en verde el sitio de restricción para *Bam*HI; en azul cambios silenciosos T/C y A/G en la secuencia de TvTBP1 de nuestra cepa. Construcción realizada por el Dr. Bernardo Franco (*Franco 2012, Tesis doctoral*).

## 6xHis-TvTBP1\* (Modificada)

### Secuencia

```
1   ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAATCCAAAAGAATTTATTGGC
1   M R G S H H H H H H G S M N P K E F I G

61  TTAGAAGAAGCTAATATAAATAATGCAAAAAGAATTCGAACAGCTTGTTAAAGAAGCTCAA
21  L E E A N I N N A K E F E Q L V K E A Q

121 GAGGAATCAGAGGAAGAAATTACACCAGAATCAATACAGAGAAATATAAGTTTAAATCAA
41  E E S E E E I T P E S I Q R N I S L N Q

181 CCTCCCATTCCCAAAGTTGTCAACATTGTTGCATGCGTAGATTACGGATGCAAACCTTGAC
61  P P I P K V V N I V A C V D Y G C K L D

241 CTTCTTAAGATTACATCTTCAACAAGAAATTCAGAATATAATCCGATTCGTTTCCCTGCT
81  L L K I T S S T R N S E Y N P I R F P A

301 GCTACGCTAAGGAACAAAGAACCAAAAGCCACAGCTTTAGCTTTTAAAAACGGAAAAGTT
101 A T L R N K E P K A T A L A F K N G K V

361 GTTATCACAGGATGTAAGCGCTGATATAGCATTAAAAGCTGCGCGTTGCTTTGGTAGA
121 V I T G C K S A D I A L K A A R C F G R

421 TTATTCAAAAATATCGGCATTAAAGACGTTAAAATCAAATCATTCTCAATAGCTAACATG
141 L F K N I G I K D V K I K S F S I A N M

481 GTTGCAACAATGGACTGCAAATTCGCCATACATCTTGAATCAATTGCTTCATCGCCTGGT
161 V A T M D C K F P I H L E S I A S S P G

541 CATATTAAATTTGCTACATACAATCCTGAAATTTTTGCAGGTTTGATTTATAGATTGGCT
181 H I K F A T Y N P E I F A G L I Y R L A

601 AGTCCAAAGACGGTTTTACTGTATATTCGTTTCAGGTAAAATTATTATCACTGGAGCTAAA
201 S P K T V L L I F V S G K I I I T G A K

661 TCAGAAGAAGAATTTAAAGCCGCATCTGCATTTATTTATCCTATTCTTAGATTGTTTGCT
221 S E E E L K A A S A F I Y P I L R L F A

721 AGCACAGAACCGCCAGCTGCTTAA
241 S T E P P A A *
```

**6xHis-TvTBP1\* (Modificada).** Secuencia de TvTBP1\*-pYES2 Mut1,2,3 (construida por Alexandra Ibañez Escribano) subclonada en el vector pQE30 que adiciona una etiqueta de 6 histidinas en el extremo amino terminal. Integrada *Bam*HI-*Hind*III. En negritas el codón de inicio (ATG) y el codón de paro (TAA); en rojo la secuencia que codifica la etiqueta de 6 histidinas; en verde el sitio de restricción para *Bam*HI; en azul cambios silenciosos T-C y A-G en la secuencia de TvTBP1 de nuestra cepa; en azul claro, cambios realizados por mutagénesis dirigida en TvTBP1 para tener una TvTBP1 más parecida a la ScTBP. Mutagénesis: Mut1: Y por F (A/T), Mut2: N por V (AA/GT) y V por T (GT/AC), Mut3: T por V (C/T), M por L (A/C), Y por F (A/T). Se construyeron por separado los vectores Mut1, Mut1,2, Mut1,3 y Mut1,2,3, sólo se muestra la secuencia de Mut1,2,3.



## 6xHis-TvTBP2

### Secuencia

```
1  ATGAGAGGATCGCATCACCATCACCATCACGGATCGACGCTGCGAGTTTCTTTGATGCT
1  M R G S H H H H H H G S D A A S F F D A
61  GATGGCGGTGTGAAAGAAATGAATCCTTGATGATTCAGAATCTATGAAATGGTATAAAGAA
21  D G G V K E M N L D D S E S M K W Y K E
121  TTAGAGAATGATGCTGATATCACAGTATCAACAACAGAAGATATAACATCATTAACCA
41  L E N D A D I T V S T T E D I T S L K P
181  AAAGTTGTTAATGTTATCGCTCTTGACACAATATTCTTGCAATAAACTCAATCTTGTCAGT
61  K V V N V I A L A Q Y S C N K L N L V S
241  ATTGCCTCTACAATTCGTAATGCTGAATACAAGCCAAAGAGAATTAAGCCGTAGTTATT
81  I A S T I R N A E Y K P K R I K A V V I
301  CGTATTCGTGACCCAAAGGCTACAGGATTAGTATTCTCAAATGGCAAATCAACATTGTT
101  R I R D P K A T G L V F S N G K I N I V
361  GGTGCGGTTCTGTTGAGGATGCTAAGCGTGCTGCTCACAAGTTTAGAAAGATGCTTCAG
121  G C R S V E D A K R A A H K F R K M L Q
421  CAGATTGGATATGATGTCAAGCTCGTCAATATTACTATATCATCAATTGTTGCTACAATA
141  Q I G Y D V K L V N I T I S S I V A T I
481  CATAACCATTCAACATCGCTATTGCACAAATAGCCAGTGCTGATGGCCACAACTATTT
161  H T P F N I A I A Q I A S A D G H K L F
541  TGTCAATATAGACCAGAAAAATTCGCTGGTTTGATTTACAGGCTTACAGACCCCTCAATGC
181  C Q Y R P E K F A G L I Y R L T D P Q C
601  ACAATGCTCATCTTCCAATCAGGAAGTATTGTTCTTACTGCAAAATCCGAGGATGACCTT
201  T M L I F Q S G S I V L T A K S E D D L
661  ACAGCAGGATCAAACCTGGATATACCCAGTTCTACAGAAGTTCGAGAAGAAGTCTATGTCG
221  T A G S N W I Y P V L Q K F E K K S M S
721  GAATTACTTATATTAA
241  E L L I *
```

**6xHis-TvTBP2.** Secuencia de TvTBP2 (TVAG\_285070) clonada en el vector pQE30 que adiciona una etiqueta de 6 histidinas en el extremo amino terminal. Integrada *Bam*HI-*Not*I. Se eliminó el ATG inicial de TvTBP2 y se modificó el siguiente ATG (ATG por ACG) con el fin de dirigir la traducción al ATG del vector. En negritas el codón de inicio (ATG) y el codón de paro (TAA); en rojo la secuencia que codifica la etiqueta de 6 histidinas; en verde el sitio de restricción para *Bam*HI; en azul cambio de T por C modificado en esta proteína recombinante para eliminar el ATG. Clona secuenciada #10.

## 6xHis-IBP39

### Secuencia

```
1  ATGAGAGGATCGCATCACCATCACCATCACGGATCCAACGACCTTGAAGCAAGT
1  M R G S H H H H H H G S D S N D L E A S
61  TTTACATCTCGTCTCCACCGGAGATAGTGGCAGCTCTAAAAAGAAAATCTAGCCGTGAT
21  F T S R L P P E I V A A L K R K S S R D
121  CCAAACTCTCGCTTTCCAAGGAAGTTACACATGCTCCTCACTTATCTTGCTTCAAATCCA
41  P N S R F P R K L H M L L T Y L A S N P
181  CAACTCGAAGAGGAAATTGGTTTATCCTGGATAAGCGATACAGAGTTCAAATGAAGAAG
61  Q L E E E I G L S W I S D T E F K M K K
241  AAAAATGTTGCACTGGTTATGGGAATTAAACTCAACACTCTCAACGTTAATCTTCGTGAT
81  K N V A L V M G I K L N T L N V N L R D
301  CTTGCATTCGAGCAACTCCAACACGATAAAAGCGGATGGACCCAATGGAAACGTTCCGGC
101  L A F E Q L Q H D K G G W T Q W K R S G
361  TTTACCAGAAATAGTGTATTCGAAGACCCAACTCAAACGATTCCCCGATGCCAGTAAAT
121  F T R N S V F E D P T Q N D S P M P V N
421  AAAAAAGATCCAATGGTACGAAAAGAGTTGAATTTCCGACCACGAAGAAGTCGATGTGC
141  T K R S N G T K R V E F P T T K K S M C
481  ATTGGAAATTCACACCCAATGAACAAGAAACATTTAGGGCCAAAGTGGACGAAATCTGG
161  I G N S T P N E Q E T F R A K V D E I W
541  TTTAGGTTAACGCAAAAAACAGACGGAACTGTTATGCGAGACTTCCTAATTGAAAAGGCT
181  F R L T Q K T D G T V M R D F L I E K A
601  GCCGAATATTTCAAGCAGCCTGAACAACCAAAACAAAATGCGATAGAAGTAATTAGCGCC
201  A E Y F K Q P E Q P K Q N A I E V I S A
661  ATTATGGCTCCTCAAGAAGAACAGACAAAAGCAAGGCAGATCTTTATAAATTCTTAGCA
221  I M A P Q E E Q T K S K A D L Y K F L A
721  ATGTTTGGTCCTTATGAGACCATTATGCTTAAGATTGCTTCTTTACTTCTAATTTCGAAT
241  M F G P Y E T I M L K I A S L L L I S N
781  AACAAAGGCCATTGGCTTACCTTCGATCCTCAAGCTGAGAAGAACGCTAATAATCAACGC
261  N K G H W L T F D P Q A E K N A N N Q R
841  GATTCTATCTCAGGTTGGTTTGATCAAACGAGCCAAACTGTTTGATCTTGAAGACCCCG
281  D S I S G W F D Q N E P N C L I L K T P
901  ACTGGAATAAGGAAAATATGGAACAACCAATTAATCGAGGCTACCGGACAATACTTAATG
301  T G I R K I W N K P L I E A T G Q Y L M
961  GATGAAAATGGTGAAAATACGATTCTTGGGATAAATATTTTGAATGAAGCCAATTGAA
321  D E N G E K Y D S W D K Y F E M K P I E
1021  ACTTACCTCACAGCTTATCCTACTTTCGCTCCAATGTAA
341  T Y L T A Y P T F A P M *
```

**6xHis-IBP39.** Secuencia de IBP39 clonada en el vector pQE30 que adiciona una etiqueta de 6 histidinas en el extremo amino terminal. Integrada *Bam*HI-*Kpn*I, se eliminó el ATG inicial de IBP39 y se modificó el siguiente ATG (ATG por ACG) con el fin de dirigir la traducción al ATG del vector. En negritas el codón de inicio (ATG) y el codón de paro (TAA); en rojo la secuencia que codifica la etiqueta de 6 histidinas; en verde el sitio de restricción para *Bam*HI; en azul cambios silenciosos T-C modificado para esta proteína recombinante y G-T cambio en la secuencia de IBP39 de nuestra cepa.

## 6xHis-SPT15

### Secuencia

```
1  ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGGCCGATGAGGAACGTTTAAAG
1  M R G S H H H H H H G S M A D E E R L K

61  GAGTTTAAAGAGGCCAAACAAGATAGTGTGGATCCAAATACCAGACAAGTATGGGAAAAC
21  E F K E A N K I V F D P N T R Q V W E N

121  CAGAATCGAGATGGTACAAAACCAGCAACTACTTTCCAGAGTGAAGAGGACATAAAAAGA
41  Q N R D G T K P A T T F Q S E E D I K R

181  GCTGCCCCAGAATCTGAAAAAGACACCTCCGCCACATCAGGTATTGTTCCAACACTACAA
61  A A P E S E K D T S A T S G I V P T L Q

241  AACATTGTGGCAACTGTGACTTTGGGGTGCAGGTTAGATCTGAAAACAGTTGCGCTACAT
81  N I V A T V T L G C R L D L K T V A L H

301  GCCCCTAATGCAGAATATAACCCCAAGCGTTTTGCTGCTGTCATCATGCGTATTAGAGAG
101  A R N A E Y N P K R F A A V I M R I R E

361  CCAAAAACCTACAGCTTTAATTTTTGCCTCAGGGAAAATGGTTGTTACCGGTGCAAAAAGT
121  P K T T A L I F A S G K M V V T G A K S

421  GAGGATGACTCAAAGCTGGCCAGTAGAAAATATGCAAGAATTATCCAAAAAATCGGGTTT
141  E D D S K L A S R K Y A R I I Q K I G F

481  GCTGCTAAATTCACAGACTTCAAATAACAAAATATTGTCGGTTCGTGTGACGTTAAATTC
161  A A K F T D F K I Q N I V G S C D V K F

541  CCTATACGTCTAGAAGGGTTAGCATTTCAGTCATGGTACTTTCTCCTCCTATGAGCCAGAA
181  P I R L E G L A F S H G T F S S Y E P E

601  TTGTTTCTGGTTTGATCTATAGAATGGTGAAGCCGAAAATTTGTTGTTAATTTTTGTT
201  L F P G L I Y R M V K P K I V L L I F V

661  TCAGGAAAGATTGTTCTTACTGGTGCAAAGCAAAGGGAAGAAATTTACCAAGCTTTTGAA
221  S G K I V L T G A K Q R E E I Y Q A F E

721  GCTATATACCCTGTGCTAAGTGAATTTAGAAAAATTGA
241  A I Y P V L S E F R K M *
```

**6xHis-SPT15.** Secuencia de SPT15 de *S. cerevisiae* clonada en el vector pQE30 que adiciona una etiqueta de 6 histidinas en el extremo amino terminal. Integrada *Bam*HI-*Kpn*I. En negritas el codón de inicio (ATG) y el codón de paro (TGA); en rojo la secuencia que codifica la etiqueta de 6 histidinas; en verde el sitio de restricción para *Bam*HI.

## IX. REFERENCIAS

- [1] Ahn S, Huang CL, Ozkumur E, Zhang X, Chinnala J, Yalcin A, Bandyopadhyay S, Russek S, Selim UM, Charles D, and Iranij RJ. TATA Binding Proteins Can Recognize Nontraditional DNA Sequences. *Biophys J.* 2012;103(7) 1510–1517.
- [2] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J. Mol. Biol.* 1990;215(3):403-410.
- [3] Aurrecochea C, Brestelli J, Brunk BP, Carlton JM, Dommer J, Fischer S, Gajria b, Gao X, Glinge A, Grant G, Harb OS, Heiges M, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Miller JA, Morrison HG, Nayak V, Pennington C, Pinney DF, Roos DS, Ross C, Stoeckert Jr CJ, Sullivan S, Treatman C, Wang H. GiardiaDB and TrichDB; integrated genomic resources for eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis*. 2009;37(Database issue):D526-530 <http://trichdb.org/trichdb/>
- [4] Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. *Current Protocols in Molecular Biology.* John Wiley & Sons, Inc. 1995;2:11.12.1-11.12.6.
- [5] Akhtar W, Veenstra GJ. TBP-related factors: a paradigm of diversity in transcription initiation. *Cell Biosci.* 2011;1:23.
- [6] Bartfai R, Balduf C, Hilton T, Rathmann Y, Hadzhiev Y, Tora L, Orban L, Muller F. TBP2, a vertebrate-specific member of the TBP family, is required in embryonic development of zebrafish. *Curr Biol.* 2004;14(7):593-598.
- [7] Best AA, Morrison HG, McArthur AG, Sogin ML, Olsen GJ. Evolution of eukaryotic transcription: insights from the genome of *Giardia lamblia*. *Genome Res* 2004;14(8):1537–1547.
- [8] Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UCM, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C, Van de Peer Y, Miranda-Saavedra D, Barton GJ, Westrop GD, Müller S, Dessi D, Fiori PL, Ren Q, Paulsen I, ZhangbH, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes, RD, Mukherjee M, Okumura CY, Schnider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu CL, Osoegawa K, de Jong PJ, Hrdy I, Horvarthova L, Zubacova Z, Dolezal P, Shehre-Banoo M, Logsdon Jr JM, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu CH, Lee YS, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-Liggett CM, Johnson PJ. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 2007; 315(5809):207-212.
- [9] Castañon-Sanchez CA, Luna-Arias JP, Dios-Bravo MG, Herrera-Aguirre ME, Olivares-Trejo JJ, Orozco E, Hernandez JM. *Entamoeba histolytica*: A unicellular organism containing two active genes encoding for members of the TBP family. *Protein Expression and Purification.* 2010;70(1):48–59.
- [10] Cribb P and Serra E. One- and two-hybrid analysis of the interactions between components of the *Trypanosoma cruzi* spliced leader RNA gene promoter binding complex. *Int J Parasitol,* 2009;39(5):525-532.
- [11] Cribb P, Esteban L, Trochine A, Girardini J and Serra E. *Trypanosoma Cruzi* TBP Shows Preference for C/G-rich DNA Sequences *in Vitro.* *Exp Parasitol,* 2010; 124(3):346-349.
- [12] Cudmore SL, Delgaty KL, Hayward-McClelland SF, Petrin DP, and Garber GE. Treatment of Infections Caused by Metronidazole-Resistant *Trichomonas vaginalis*. *Clin Microbiol Rev.* 2004;17(4):783-793.

- [13] Dacks JB and Doolittle WF. Reconstructing/deconstructing the earliest eukaryotes: how comparative genomics can help. *Cell*, 2001;107(4):419–425.
- [14] Danino YM, Even D, Ideses D, Juven-Gershon T. The core promoter: At the heart of the gene expression. *Biochim Biophys Acta*. 2015;1849(8):1116-11131.
- [15] Dantoni JC, Wurtz JM, Poch O, Moras D, Tora L. The TBP-like factor: an alternative transcription factor in metazoa? *Trends Biochem Sci*. 1999;24(9):335-339.
- [16] Davidson I. The genetics of TBP and TBP-related factors. *Trends Biochem Sci*. 2003;28(7): 391-398.
- [17] Denissov S, van Driel M, Voit R, Hekkelman M, Hulsen T, Hernandez N, Grummt I, Wehrens R, Stunnenberg H. Identification of novel functional TBP-binding sites and general factor repertoires. *EMBO J*. 2007;21;26(4):944-954.
- [18] Diamon LS. The establishment of various trichomonads of animals and man in axenic cultures. *J Parasitol*. 1957;43(4):488-490.
- [19] Donahue TF, Farabaugh PJ, Fink GR. The nucleotide sequence of the HIS4 region of yeast. *Gene* 1982;18(1):47-59.
- [20] Duttke SHC, Doolittle RF, Wang YL, Kadonaga JT. TRF2 and the evolution of the bilateria. *Genes Dev*. 2014;28(19):2071-2076.
- [21] Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. The Pfam protein families database. *Nucleic Acids Res*. 2014;(Database issue);42:D222-D230 <http://pfam.xfam.org/>
- [22] Fishburn J, Hahn S. Architecture of the yeast RNA polymerase II open complex and regulation of activity by TFIIF. *Mol Cell Biol* 2012;32(1):12-25.
- [23] Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C and Jensen LJ. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*, 2013;41 (Database issue) D808–D815.
- [24] Franco B, Hernández R, López-Villaseñor I. *Trichomonas vaginalis* ribosomal RNA: Identification and characterization of the transcription promoter and terminator sequences. *Mol Biochem Parasitol* 2012;185(1):1-9.
- [25] Franco B and López-Villaseñor I. Caracterización de la región promotora y terminadora de la transcripción del cistron ribosomal de *Trichomonas vaginalis*. Tesis doctoral, 2012. <http://132.248.9.195/ptd2012/junio/0681443/Index.html>
- [26] Fuentes V, Barrera G, Sánchez J, Hernández R, López-Villaseñor I. Functional analysis of sequence motifs involved in the polyadenylation of *Trichomonas vaginalis* mRNAs. *Eukaryot Cell*. 2012;11(6):725-734.
- [27] Garber GE. The laboratory diagnosis of *Trichomonas vaginalis*. *Can J Infect Dis Med Microbiol* 2005;16(1):35-38.
- [28] Gietl A, Holzmeister P, Blombach F, Schulz S, Voith von Voithenberg L, Lamb DC, Werner F, Tinnefeld P and Grohmann D. Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. *Nucleic Acids Res*. 2014;42(10):6219–6231.
- [29] Golemis EA and Khazak V. Alternative yeast two-hybrid systems. The interaction trap and interaction mating. *Methods Mol Biol*. 1997;63:197-218.
- [30] Gomez C, Ramirez ME, Calixto-Galvez M, Medel O, Rodríguez MA. Regulation of gene expression in protozoa parasites. *J Biomed Biotechnol*. 2010;2010:726045.
- [31] Graves KJ, Novak J, Secor WE, Kissinger PJ, Schwebke JR, Muzny CA. A Systematic Review of the Literature on Mechanisms of 5-Nitroimidazole Resistance in *Trichomonas vaginalis*. *Parasitology*. 2020;147(13)1383-1391.

- [32] Gunderson J, Hinkle G, Leipe D, Morrison HG, Stickel SK, Odelson DA, Breznak JA, Nerad TA, Müller M, Sogin ML. Phylogeny of trichomonads inferred from small subunit rRNA sequences. *J Eukaryot Microbiol.* 1995; 42(4):411-415.
- [33] Hahn S. Structure and mechanism of the RNA Polymerase II transcription machinery. *Nat Struct Mol Biol.* 2004;11(5):394–403.
- [34] Hahn S and Young ET. Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics.* 2011;189(3):705–736.
- [35] Hansen SK, Takada S, Jacobson RH, Lis JT, Tjian R. Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell* 1997;91(1):71-83.
- [36] Hernandez N. TBP, a universal eukaryotic transcription factor? *Genes Dev.* 1993;7(7B):1291-1308.
- [37] Jallow Z, Jacobi UG, Weeks DL, Dawid IB, Veenstra GJC. Specialized and redundant roles of TBP and a vertebrate-specific TBP paralog in embryonic gene regulation in *Xenopus*. *Proc Natl Acad Sci.* 2004;14;101(37):13525-13530.
- [38] Janssen BD, Chen YP, Molgora BM, Wang SE, Simoes-Barbosa A and Johnson PJ. CRISPR/Cas9-mediated gene modification and gene knock out in the human-infective parasite *Trichomonas vaginalis*. *Sci Rep.* 2018;8(1):270.
- [39] Juo ZS, Chiu TK, Leiberman PM, Baikalov I, Berk AJ, Dickerson RE. How proteins recognize the TATA box. *J Mol Biol.* 1996;261(2):239-254.
- [40] Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc.* 2012;7(8):1511-1522. <http://raptorx.uchicago.edu/>
- [41] Kawasaki L, Saviñón-Tejeda AL, Ongay-Larios L, Ramírez J, Coria R. The Gbeta(KISte4p) Subunit of the Heterotrimeric G Protein Has a Positive and Essential Role in the Induction of Mating in the Yeast *Kluyveromyces Lactis*. *Yeast,* 2005;22(12):947-956.
- [42] Kelley LA and Sternberg MJE. Protein structure prediction on the web: a case study using the Phyre server. *Nat Protoc* 2009;4(3):363-371. <http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>
- [43] Kim JL and Burley SK. 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nat Struct Biol.* 1994;1(19):638-653.
- [44] Koster MJE, Snel B, Timmers HTM. Genesis of chromatin and transcription dynamics in the origin of species. *Cell,* 2015;161(4):724-736.
- [45] Kramm K, Engel C and Grohmann D. Transcription initiation factor TBP: old friend new questions. *Biochem Soc Trans.* 2019;28;47(1):411-423.
- [46] Kulda J. Trichomonads\ hydrogenosomes and drug resistance. *Int J Parasitol.* 1999;29(2)199-212.
- [47] Lahti CJ, D'Oliveira CE, Johnson PJ.  $\beta$ -Succinyl-Coenzyme A synthetase from *Trichomonas vaginalis* is a soluble hydrogenosomal protein with an Amino-terminal sequence that resembles mitochondrial presequences. *J Bacteriol* 1992;174(21):6822-6830.
- [48] Liston DR, Johnson PJ. Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. *Mol Cell Biol.* 1999;19(3):2380-2388.
- [49] Liston DR, Lau AO, Ortiz D, Smale ST and Johnson PJ. Initiator recognition in a primitive eukaryote: IBP39, an initiator-binding protein from *Trichomonas vaginalis*. *Mol Cell Biol.* 2001;21(22):7872-7882.



- [50] Louder RK, He Y, López-Blanco JR, Fang J, Chacón P, and Nogales E. Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature*. 2016;531(7596):604–609.
- [51] McClelland RS, Sangare L, Hassan WM, Lavreys L, Mandaliya K, Kiarie J, Ndinya-Achola J, Jaoko W, and Baeten JM. Infection with *Trichomonas vaginalis* increases the risk of HIV-1 acquisition. *J Infect Dis*. 2007;195(5): 698-702.
- [52] Mercer F. and Johnson PJ. *Trichomonas vaginalis*: Pathogenesis, Symbiont Interactions, and Host Cell Immune Responses. *TREPAR* 2018;34(8)683-693.
- [53] Millán-Pacheco C, Capistrán VM and Pastor N. On the consequences of placing amino groups at the TBP–DNA interface. Does TATA really matter? *J Mol Recognit*, 2009;22(6):453-464.
- [54] Mundade R, Ozer HG, Wei H, Prabhu L and Lu T. Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. *Cell Cycle* 2014;13(18):2847-2852.
- [55] Narayanasamy RK, Castañón-Sánchez CA, Luna-Arias JP, García-Rivera G, Avendaño-Borromeo B, Labra-Barrios ML, Valdés J, Herrera-Aguirre ME and Orozco E. 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*. 2018;11(1):153.
- [56] Nicholas KB, Nicholas HB Jr: GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors. <http://www.cris.com/ketchup/genedoc.shtml> (1997).
- [57] Ongay-Larios L, Saviñón-Tejeda AL, Williamson Jr MJ, Durán-Avelar Md, Coria R. The Leu-132 of the Ste4(G $\beta$ ) subunit is essential for proper coupling of the G protein with the Ste2 $\alpha$  factor receptor during the mating pheromone response in yeast. *FEBS Lett*, 2000;467(1):22-26.
- [58] Page RD. TreeView: An application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 1996 12(4):357-358.
- [59] Quon DV, Delgadillo MG, Khachi A, Smale ST and Johnson PJ. Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. *Proc Natl Acad Sci USA* 1994; 91(10):4579–4583.
- [60] Rigby PW. Three in one and one in three: It all depends on TBP. *Cell*. 1993;72(1):7-10.
- [61] Roger AJ, Clark CG and Doolittle WF. A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. *Proc Natl Acad Sci U S A*, 1996;93(25): 14618–14622.
- [62] Ruan JP, Arhin GK, Ullu E and Tschudi C. Functional characterization of a *Trypanosoma brucei* TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes. *Mol Cell Biol*. 2004;24(21):9610-9618.
- [63] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, and Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011;7:539 <http://www.ebi.ac.uk/Tools/msa/clustalo/>
- [64] Simoes-Barbosa A, Meloni D, Wohlschlegel JA, Konarska MM and Johnson PJ. Spliceosomal snRNAs in the unicellular eukaryote *Trichomonas vaginalis* are structurally conserved but lack a 59-cap structure. *RNA* 2008;14(8):1617-1631.
- [65] Smith A. and Johnson P. Gene expression in the unicellular eukaryote *Trichomonas vaginalis*. *Res Microbiol*. 2011;162(6):646-654.

- [66] Sprouse RO, Wells MN, and Auble DT. TATA-binding Protein Variants That Bypass the Requirement for Mot1 *in Vivo*. *J Biol Chem*, 2009;13,284(7):4525-4535.
- [67] Takada S, Lis JT, Zhou S and Tjian R. A TRF1:BRF Complex directs *Drosophila* RNA polymerase III transcription. *Cell*. 2000; 101(5):459-469.
- [68] Thomas MC and Chiang CM. The General transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol*. 2006;41(3):105-178.
- [69] Thomas S, Yu MC, Sturm NR, and Campbell DA. A non-universal transcription factor? The *Leishmania tarentolae* TATA box-binding protein LtTBP associates with a subset of promoters. *Int. J. Parasitol*. 2006;36(10-11):1217-1226.
- [70] Thomas S, Green A. Sturm NR, and Campbell DA. Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. *BMC Genomics*. 2009;8,10:152.
- [71] Torres-Machorro AL, Hernández R, Sánchez J, and López-Villaseñor I. The 5S ribosomal RNA gene from the early diverging protozoa *Trichomonas vaginalis* *Mol Biochem Parasitol* 2006;145(2):269–273.
- [72] Van Gerwen OT and Muzny CA. Recent advances in the epidemiology, diagnosis, and management of *Trichomonas vaginalis* infection. *F1000Res*. 2019;8 F1000 Faculty Rev-1666.
- [73] Vanacova S, Liston DR, Tachezy J, Johnson PJ. Molecular biology of the amitochondriate parasites, *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. *Int J Parasitol*. 2003;33(3):235-255.
- [74] Vannini A, Cramer P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol Cell*. 2012;45(4):439-446.
- [75] Vannini A. A structural perspective on RNA polymerase I and RNA polymerase III transcription machineries. *Biochim Biophys Acta*. 2013;1829(3-4):258-264.
- [76] Voleman L and Dolezal P. Mitochondrial dynamics in parasitic protists. *PLoS Pathog*. 2019;15(11): e1008008.
- [77] White RJ and Jackson SP. The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. *Trends Genet*. 1992;8(8):284-28.
- [78] World Health Organization (2011) Prevalence and incidence of selected sexually transmitted infections, *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *syphilis* and *Trichomonas vaginalis*: methods and results used by WHO to generate 2005 estimates. WHO 1–38.
- [79] Zehavi Y, Kedmi A, Ideses D, and Juven-Gershon T. TRF2: TRansForming the view of the general transcription factors. *Transcription*. 2015;6(1):1-6.



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

**Olivia Parra-Marín, Lluvia Rosas-  
Hernández, Karla López-Pacheco,  
Bernardo Franco, Alexandra Ibáñez-  
Escribano, Roberto Hernández, et al.**

**Parasitology Research**  
Founded as Zeitschrift für  
Parasitenkunde

ISSN 0932-0113

Parasitol Res  
DOI 10.1007/s00436-019-06438-z



**Your article is protected by copyright and all rights are held exclusively by Springer-Verlag GmbH Germany, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**



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

Olivia Parra-Marín<sup>1</sup> · Lluvia Rosas-Hernández<sup>1</sup> · Karla López-Pacheco<sup>1</sup> · Bernardo Franco<sup>1,2</sup> · Alexandra Ibáñez-Escribano<sup>1,3</sup> · Roberto Hernández<sup>1</sup> · Imelda López-Villaseñor<sup>1</sup>

Received: 8 May 2019 / Accepted: 22 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 *Tvtbp* 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 archaeobacterial 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 (TATAAWAN) 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 TFIIB 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

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

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

<sup>3</sup> 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 aetiological 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 *Tvtpb* 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 CNC147 (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 *Tvtpb1* and *Tvtpb2* transcripts, respectively. Briefly, the cells were grown to a density of approximately  $1.5 \times 10^6$  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 *Tvtpb* 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  $\beta$ -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  $\mu$ g 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  $1 \times 10^6$  *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<sub>2</sub>, 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 cytoplasmic 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  $\beta$ -mercaptoethanol (BME) in the sample buffer) was performed. The polyclonal TvTBP1 antiserum was diluted 1:10,000 to test the recombinant proteins and 1:500 to test the enriched nuclear extracts from *T. vaginalis*.

### Yeast complementation analyses

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

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

The expression of the recombinant TvTBP1 and TvTBP1mut in the transformed yeast cultures was demonstrated by western blot using the anti-TvTBP1 antibody (Dunn and Wobbe 1993). Yeast *spt15Δ* mutant diploid cells were transformed with the pYES2 constructs and grown in induction medium (synthetic dropout medium lacking uracil and containing 2% galactose, 1% raffinose). Cultures were grown to mid-log and cells were collected by centrifugation. The pellet was washed

twice with ice-cold water. Cells were resuspended in three volumes of glass bead disruption buffer (20 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 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).

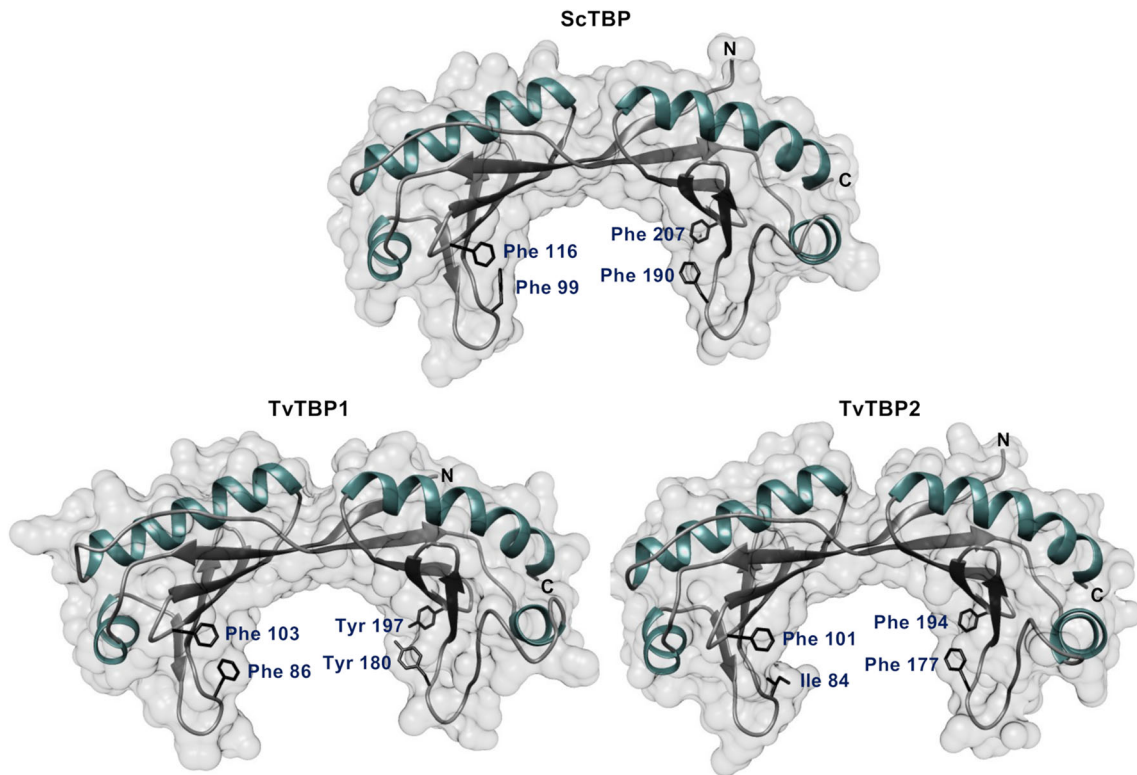
### Pull-down assays

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	● <sup>56</sup>	● <sup>58</sup>	● <sup>86</sup>	● <sup>103</sup>	N <sup>109</sup>	V <sup>111</sup>	● <sup>147</sup>	● <sup>149</sup>	Y <sup>180</sup>	T <sup>193</sup>	M <sup>195</sup>	Y <sup>197</sup>	I <sup>204</sup>	● <sup>205</sup>
TvTBP2	● <sup>53</sup>	I <sup>55</sup>	I <sup>84</sup>	● <sup>101</sup>	N <sup>107</sup>	V <sup>109</sup>	S <sup>144</sup>	● <sup>146</sup>	● <sup>177</sup>	T <sup>190</sup>	● <sup>192</sup>	● <sup>194</sup>	● <sup>201</sup>	● <sup>202</sup>
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
GITBP	G	N	Y	S	T	R	S	T	M	C	S	●	I	V

Dots indicate amino acid identity as compared with ScTBP. Shaded 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

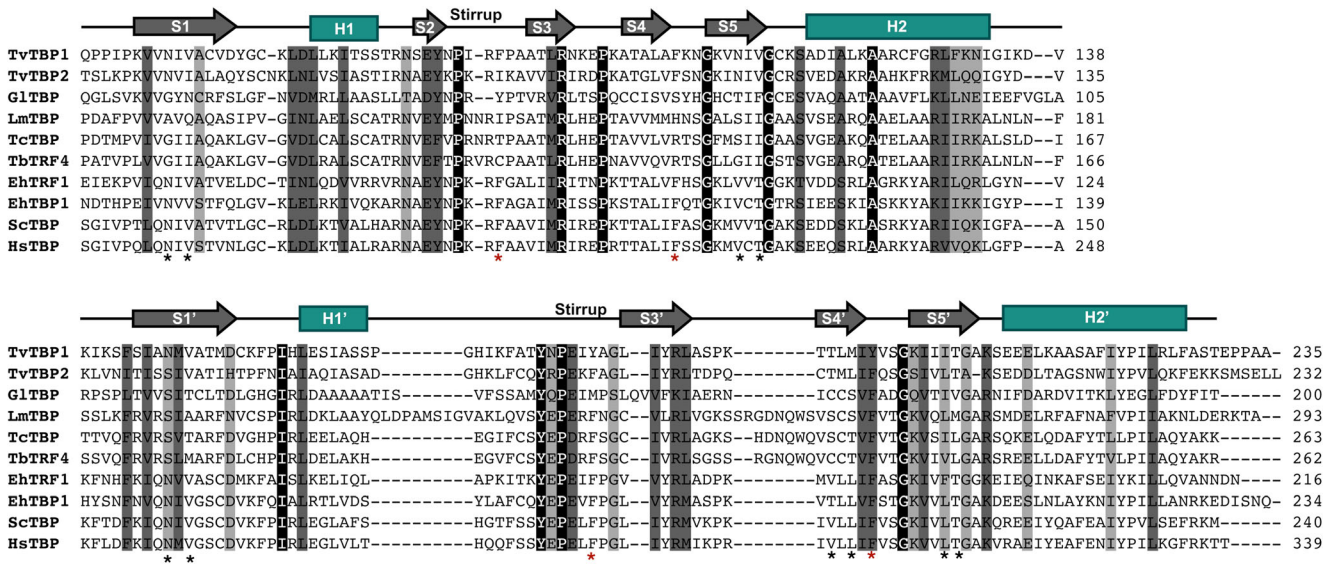


**Fig. 1** Tertiary structure models for *T. vaginalis* TBPs. Theoretical structures of TvTBP1 and TvTBP2 adopt the typical saddle-shaped structure shown by a canonical TBP from *Saccharomyces cerevisiae* (ScTBP). Alpha helices are shown in green, and beta strands are shown in dark

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.

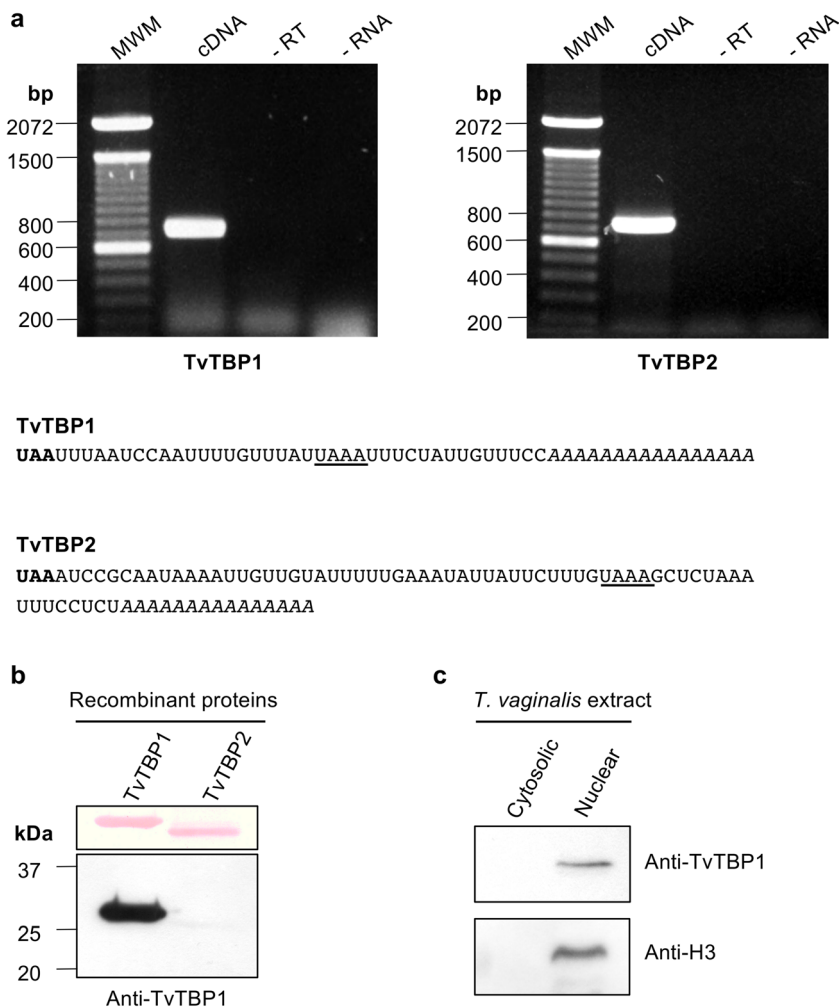


**Fig. 2** Alignment of TBP amino acid sequences. Carboxy-terminal sequences of TBPs from *T. vaginalis* (TvTBP1 (A2DQ7) and TvTBP2 (A2G4N7)), *Giardia lamblia* (G1TBP (Q86GG9)), *Leishmania major* (LmTBP (Q4QD71)), *Trypanosoma cruzi* (TcTBP (Q4DK05)), *Trypanosoma brucei* (TbTRF4 (Q387R9)), *Entamoeba 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,  $\alpha$  helix; S,  $\beta$  strand). Phenylalanine residues (red) and other residues (black) involved in DNA binding are marked with asterisks



**Fig. 3** **a** Both *Tvtbp* genes are expressed in *T. vaginalis*. Agarose gels stained with ethidium bromide show RT-PCRs amplified with specific primers for *Tvtbp1* and *Tvtbp2*; 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

EMSA 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<sub>2</sub>, 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

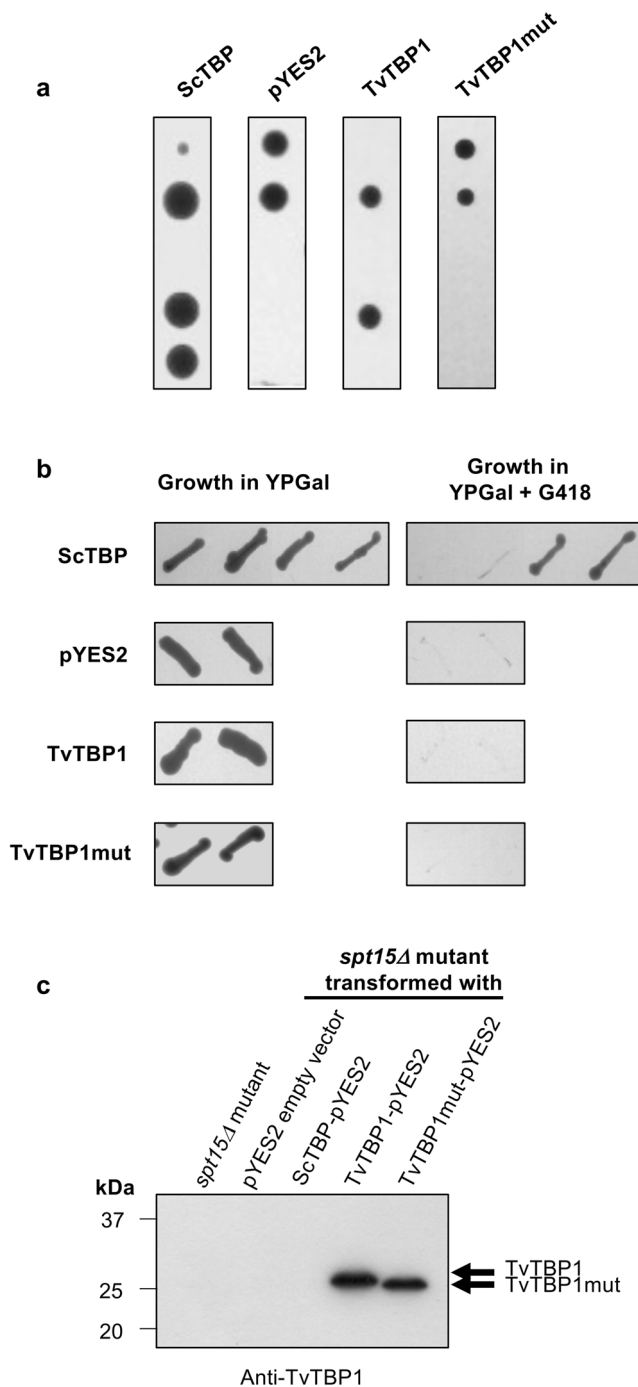
terminal domain of the TBP superfamily (Pfam: PF00352) (Finn et al. 2014), and the models of their tertiary structures clearly adopt the distinctive saddle-shaped structure (Fig. 1). These two genes are therefore herein referred to as *Tvtbp1* and *Tvtbp2*, respectively.

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, Tv1107B09) have been reported. To



determine whether both *Tvtp* 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 *Tvtp1* and *Tvtp2* 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).

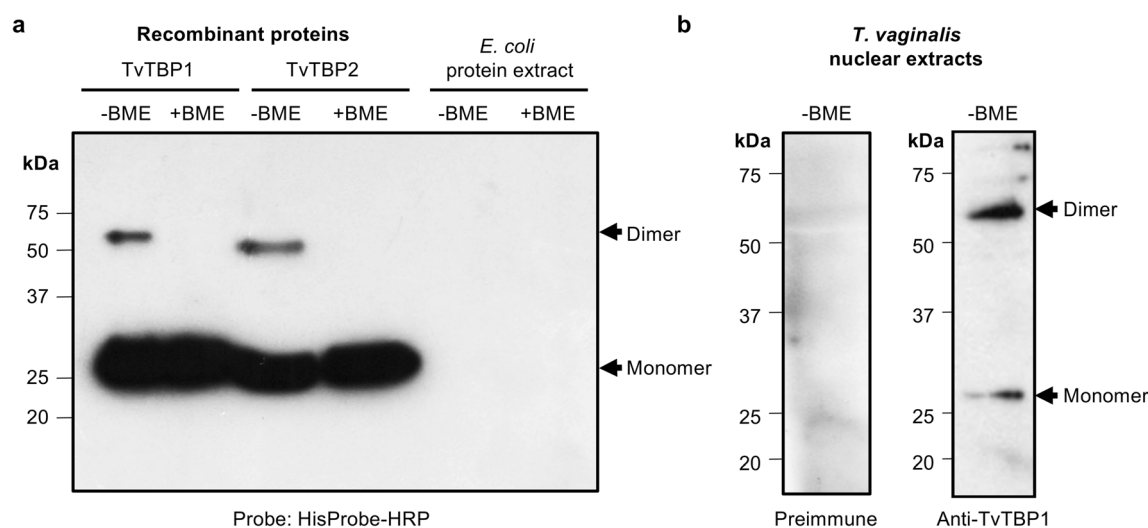
### TvTBP1 does not complement a *S. cerevisiae* strain lacking TBP

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 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  $\beta$ -mercaptoethanol in the sample buffer; +BME, sample with  $\beta$ -

mercaptoethanol in the sample buffer. **b** Enriched nuclear extracts from *T. vaginalis* probed with the anti-TvTBP1 antiserum. In the absence of  $\beta$ -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).

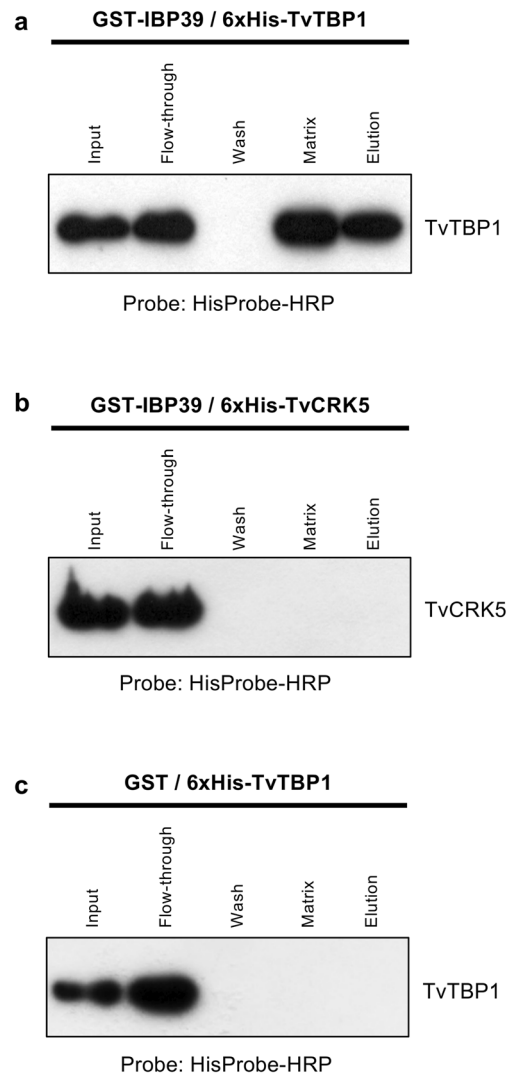
### TvTBP1 interacts with IBP39

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.

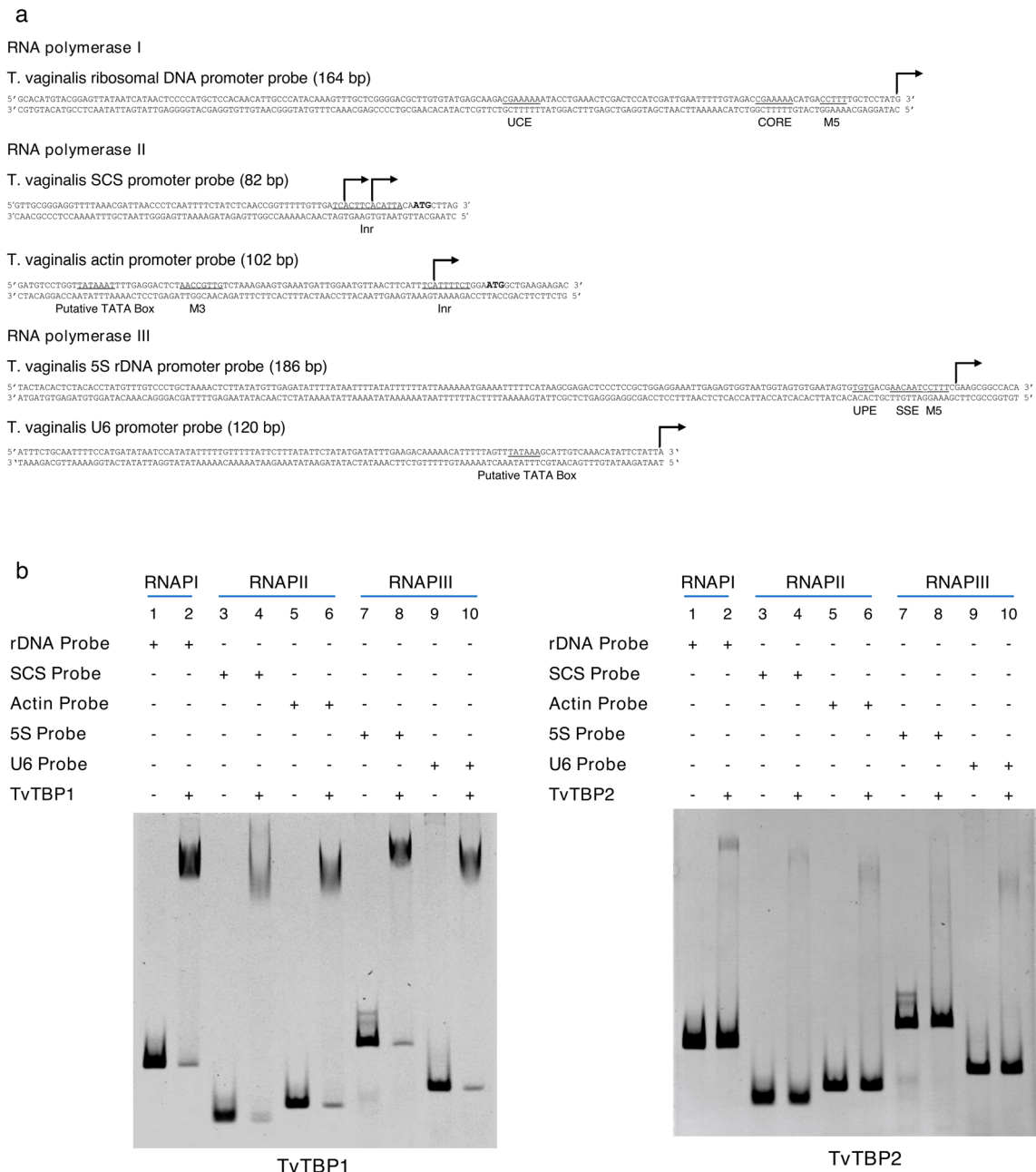
### TvTBP1 and TvTBP2 bind to promoter regions of *T. vaginalis* transcribed by the three RNA polymerases

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  $\beta$  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  $\beta$  subunit of succinyl CoA synthetase ( $\beta$ -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:

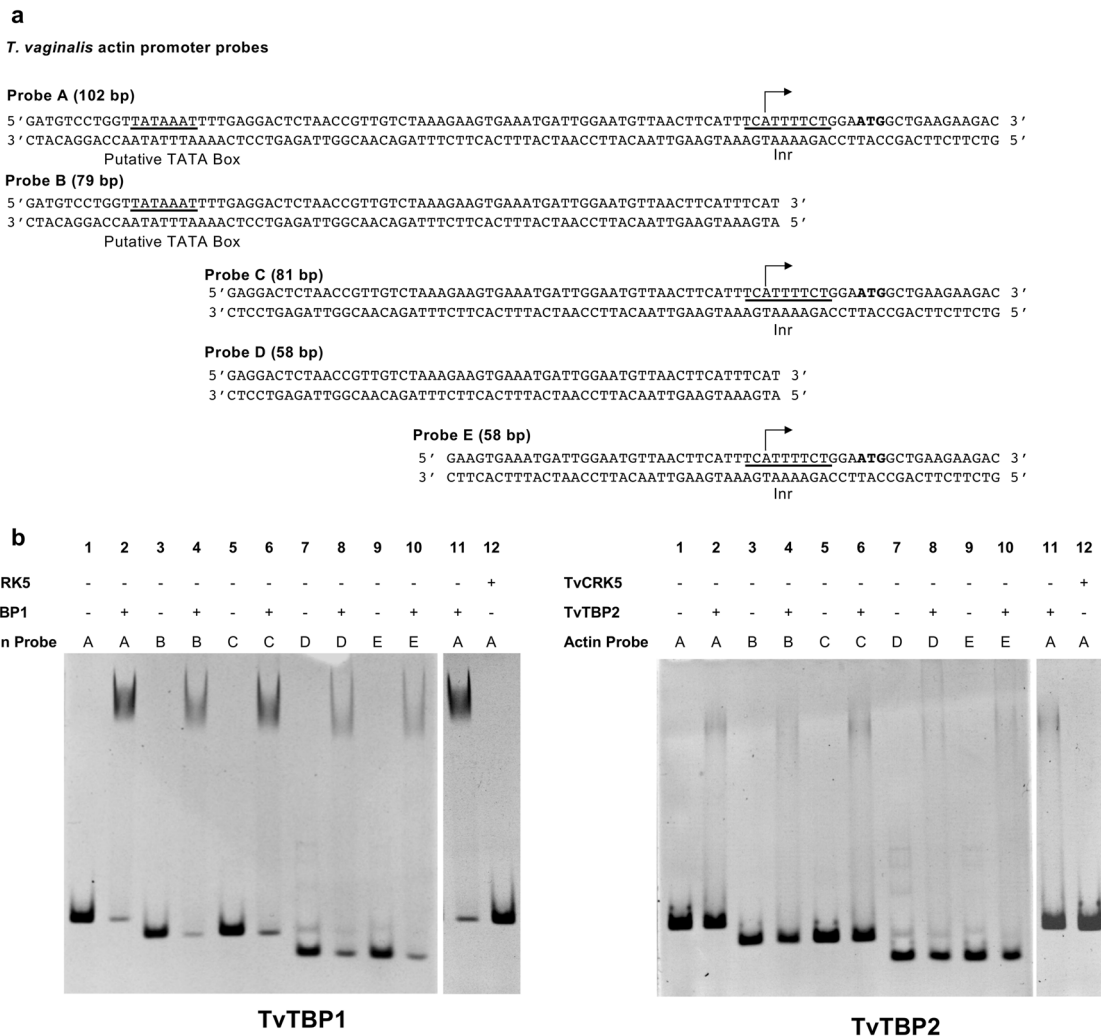
Upstream Control Element (UCE), core element, putative TATA box, Inr element, Upstream Promoter Element (UPE), Start Site Element (SSE), motif 3 (M3) and motif 5 (M5). The +1 transcription start site is shown by an arrow. **b** Electrophoretic mobility shift assay (EMSA) for the TvTBP1 and TvTBP2 recombinant 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 *Cryptosporidium parvum*, 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>
- Castañón-Sánchez 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.bbagnm.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](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](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 *Cryptophycinium 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, Baikov 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 TATAAAG. *Nat Struct Biol* 1:638–653. <https://doi.org/10.1038/nsb0994-638>

- Koster MJE, Snel B, Timmers HTM (2015) Genesis of chromatin and transcription dynamics in the origin of species. *Cell* 161:724–736. <https://doi.org/10.1016/j.cell.2015.04.033>
- Kou H, Pugh BF (2004) Engineering dimer-stabilizing mutations in the TATA-binding protein. *J Biol Chem* 279:20966–20973. <https://doi.org/10.1074/jbc.M401535200>
- Kou H, Irvin JD, Huisinga KL, Mitra M, Pugh BF (2003) Structural and functional analysis of mutations along the crystallographic dimer interface of the yeast TATA binding protein. *Mol Cell Biol* 23:3186–3201. <https://doi.org/10.1128/MCB.23.9.3186-3201.2003>
- Lahti CJ, D'Oliveira CE, Johnson PJ (1992) Beta-succinyl-coenzyme A synthetase from *Trichomonas vaginalis* is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. *J Bacteriol* 174:6822–6830. <https://doi.org/10.1128/jb.174.21.6822-6830.1992>
- Lau AO, Liston DR, Vanacova S, Johnson PJ (2003) *Trichomonas vaginalis* initiator binding protein, IBP39, contains a novel DNA binding motif. *Mol Biochem Parasitol* 130:167–171. [https://doi.org/10.1016/S0166-6851\(03\)00172-5](https://doi.org/10.1016/S0166-6851(03)00172-5)
- Lau AOT, Smith AJ, Brown MT, Johnson PJ (2006) *Trichomonas vaginalis* initiator binding protein (IBP39) and RNA polymerase II large subunit carboxy terminal domain interaction. *Mol Biochem Parasitol* 150:56–62. <https://doi.org/10.1016/j.molbiopara.2006.06.008>
- Liston DR, Johnson PJ (1999) Analysis of a ubiquitous promoter element in a primitive eukaryote: early evolution of the initiator element. *Mol Cell Biol* 19:2380–2388. <https://doi.org/10.1128/MCB.19.3.2380>
- Liston DR, Lau AOT, Ortiz D, Smale ST, Johnson PJ (2001) Initiator recognition in a primitive eukaryote: IBP39, an initiator-binding protein from *Trichomonas vaginalis*. *Mol Cell Biol* 21:7872–7882. <https://doi.org/10.1128/MCB.21.22.7872-7882.2001>
- Lüthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356:83–85. <https://doi.org/10.1038/356083a0>
- Millán-Pacheco C, Capistrán VM, Pastor N (2009) On the consequences of placing amino groups at the TBP-DNA interface. Does TATA really matter? *J Mol Recognit* 22:453–464. <https://doi.org/10.1002/jmr.963>
- Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG, Burley SK (1996) Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci* 93:4862–4867. <https://doi.org/10.1073/pnas.93.10.4862>
- Ong S-J, Huang S-C, Liu H-W, Tai J-H (2004) Involvement of multiple DNA elements in iron-inducible transcription of the ap65-1 gene in the protozoan parasite *Trichomonas vaginalis*. *Mol Microbiol* 52:1721–1730. <https://doi.org/10.1111/j.1365-2958.2004.04088.x>
- Petterson EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF chimera a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612. <https://doi.org/10.1002/jcc.20084>
- Quon DV, Delgadillo MG, Khachi A, Smale ST, Johnson PJ (1994) Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. *Proc Natl Acad Sci* 91:4579–4583. <https://doi.org/10.1073/pnas.91.10.4579>
- Rigby PWJ (1993) Three in one and one in three: it all depends on TBP. *Cell* 72:7–10. [https://doi.org/10.1016/0092-8674\(93\)90042-O](https://doi.org/10.1016/0092-8674(93)90042-O)
- Ruvalcaba-Salazar OK, Ramírez-Estudillo MDC, Montiel-Condado D et al (2005) Recombinant and native *Plasmodium falciparum* TATA-binding-protein binds to a specific TATA box element in promoter regions. *Mol Biochem Parasitol* 140:183–196. <https://doi.org/10.1016/j.molbiopara.2005.01.002>
- Schumacher MA, Lau AOT, Johnson PJ (2003) Structural basis of core promoter recognition in a primitive eukaryote. *Cell* 115:413–424. [https://doi.org/10.1016/S0092-8674\(03\)00887-0](https://doi.org/10.1016/S0092-8674(03)00887-0)
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2014) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539–539. <https://doi.org/10.1038/msb.2011.75>
- Simoës-Barbosa A, Meloni D, Wohlschlegel JA, Konarska MM, Johnson PJ (2008) Spliceosomal snRNAs in the unicellular eukaryote *Trichomonas vaginalis* are structurally conserved but lack a 5'-cap structure. *RNA* 14:1617–1631. <https://doi.org/10.1261/ma.1045408>
- Smith A, Johnson P (2011) Gene expression in the unicellular eukaryote *Trichomonas vaginalis*. *Res Microbiol* 162:646–654. <https://doi.org/10.1016/j.resmic.2011.04.007>
- Smith AJ, Chudnovsky L, Simoes-Barbosa A, Delgadillo-Correa MG, Jonsson ZO, Wohlschlegel JA, Johnson PJ (2011) Novel core promoter elements and a cognate transcription factor in the divergent unicellular eukaryote *Trichomonas vaginalis*. *Mol Cell Biol* 31:1444–1458. <https://doi.org/10.1128/MCB.00745-10>
- Thomas MC, Chiang C-M (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* 41:105–178. <https://doi.org/10.1080/10409230600648736>
- Torres-Machorro AL, Hernández R, Sánchez J, López-Villaseñor I (2006) The 5S ribosomal RNA gene from the early diverging protozoa *Trichomonas vaginalis*. *Mol Biochem Parasitol* 145:269–273. <https://doi.org/10.1016/j.molbiopara.2005.10.009>
- Vanacova S, Liston DR, Tachezy J, Johnson PJ (2003) Molecular biology of the amitochondriate parasites, *Giardia intestinalis*, *Entamoeba histolytica* and *Trichomonas vaginalis*. *Int J Parasitol* 33:235–255. [https://doi.org/10.1016/S0020-7519\(02\)00267-9](https://doi.org/10.1016/S0020-7519(02)00267-9)
- Vannini A (2013) A structural perspective on RNA polymerase I and RNA polymerase III transcription machineries. *Biochim Biophys Acta Gene Regul Mech* 1829:258–264. <https://doi.org/10.1016/j.bbagr.2012.09.009>
- Vannini A, Cramer P (2012) Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol Cell* 45:439–446. <https://doi.org/10.1016/j.molcel.2012.01.023>
- Webb B, Sali A (2016) Comparative protein structure modeling using MODELLER. In: *Current protocols in bioinformatics*. Wiley, Hoboken, pp 5.6.1–5.6.37
- White RJ, Jackson SP (1992) The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. *Trends Genet* 8:284–288. [https://doi.org/10.1016/0168-9525\(92\)90255-3](https://doi.org/10.1016/0168-9525(92)90255-3)
- World Health Organization (2011) Prevalence and incidence of selected sexually transmitted infections, *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, syphilis and *Trichomonas vaginalis*: methods and results used by WHO to generate 2005 estimates. WHO 1–38
- Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its Core. *J Mol Biol* 430:2237–2243. <https://doi.org/10.1016/j.jmb.2017.12.007>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.