

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

COLEGIO DE CIENCIA Y TECNOLOGÍA

MAESTRÍA EN CIENCIAS GENÓMICAS

**Análisis de la localización intracelular de la proteína EhTRF-like II
de *Entamoeba histolytica* y sus interacciones *in silico*
con secuencias de ADN**

TESIS QUE PARA OBTENER EL TÍTULO DE
MAESTRO EN CIENCIAS GENÓMICAS

PRESENTA

Darío Meneses Garmendia

Directora de la Tesis

Dra. Elisa Irene Azuara Liceaga

Codirectora

Dra. Lilia López Cánovas

Ciudad de México, Junio de 2021

SISTEMA BIBLIOTECARIO DE INFORMACIÓN Y DOCUMENTACIÓN



UNIVERSIDAD AUTÓNOMA DE LA CIUDAD DE MÉXICO COORDINACIÓN ACADÉMICA

RESTRICCIONES DE USO PARA LAS TESIS DIGITALES

DERECHOS RESERVADOS ©

La presente obra y cada uno de sus elementos está protegido por la Ley Federal del Derecho de Autor; por la Ley de la Universidad Autónoma de la Ciudad de México, así como lo dispuesto por el Estatuto General Orgánico de la Universidad Autónoma de la Ciudad de México; del mismo modo por lo establecido en el Acuerdo por el cual se aprueba la Norma mediante la que se Modifican, Adicionan y Derogan Diversas Disposiciones del Estatuto Orgánico de la Universidad de la Ciudad de México, aprobado por el Consejo de Gobierno el 29 de enero de 2002, con el objeto de definir las atribuciones de las diferentes unidades que forman la estructura de la Universidad Autónoma de la Ciudad de México como organismo público autónomo y lo establecido en el Reglamento de Titulación de la Universidad Autónoma de la Ciudad de México.

Por lo que el uso de su contenido, así como cada una de las partes que lo integran y que están bajo la tutela de la Ley Federal de Derecho de Autor, obliga a quien haga uso de la presente obra a considerar que solo lo realizará si es para fines educativos, académicos, de investigación o informativos y se compromete a citar esta fuente, así como a su autor ó autores. Por lo tanto, queda prohibida su reproducción total o parcial y cualquier uso diferente a los ya mencionados, los cuales serán reclamados por el titular de los derechos y sancionados conforme a la legislación aplicable.

Integración del jurado:

Presidente: Dra. Abigail Betanzos Fernández

Secretario: Dra. Elisa Irene Azuara Liceaga

1^{er} Vocal: Dra. Lilia López Cánovas

2^{do} Vocal: Dra. Ángeles Concepción Tecalco Cruz

Lugar donde se realizó el posgrado:

POSGRADO EN CIENCIAS GENÓMICAS

PLANTEL DEL VALLE, UACM

Comité tutorial

Dra. Elisa Irene Azuara Liceaga. PCG, UACM

Dra. Lilia López Cánovas. PCG, UACM

Dra. Elizabeth Jacqueline Castañeda Ortiz. PCG, UACM

Dra. Abigail Betanzos Fernández. Cinvestav-IPN

Dr. José de Jesús Olivares Trejo. PCG, UACM

Lectores de tesis

Dra. Ángeles Concepción Tecalco Cruz. PCG, UACM

Dr. José Olivares Trejo. PCG, UACM

Dra. Abigail Betanzos Fernández. Cinvestav-IPN

Este trabajo fue realizado en el Laboratorio 2 del Posgrado en Ciencias Genómicas de la UACM, Plantel del Valle bajo la dirección de las Dras. Elisa Irene Azuara Liceaga y Lilia López Cánovas, y en colaboración con el Dr. Alfonso Olivos García† de La Unidad de Investigación en Medicina Experimental de la UNAM en el Hospital General. El proyecto recibió el apoyo financiero de la UACM mediante la Beca de Posgrado 2020-I y el Proyecto Estudiantil 13, así como de la Beca CONACyT para realizar estudios de Maestría (No. becario: 928090) y el proyecto: CONACyT CF-2019-194163 otorgado a la Dra. Elisa Irene Azuara Liceaga.

Agradecimientos

A la **Dra. Elisa Azuara** por aceptarme en su equipo de trabajo y permitirme desarrollar el proyecto en su laboratorio, por su dedicación y compromiso con el proyecto, por su tiempo, sus valiosos consejos y gran paciencia. Es una persona admirable y por la que tengo un gran respeto.

A la **Dra. Lilia López** por sus grandes aportaciones para la culminación de este proyecto, por su apoyo académico desinteresado y por toda la ayuda durante todo el progreso del mismo.

A la **Dra. Jacqueline Castañeda** por su apoyo en los experimentos, por todas esas horas de trabajo, y sus comentarios críticos y positivos que han enriquecido este trabajo.

A la **Dra. Abigail Betanzos** que me brindó asesoría, por su valiosa ayuda para la revisión y análisis de las microscopías confocales en la Unidad de microscopía del Cinvestav-IPN, y por sus acertados comentarios.

Al **Dr. Alfonso Olivos García†** y **Mario Nequiz Avedaño** por abrirme las puertas de su laboratorio en la Unidad de Investigación en Medicina Experimental (UNAM) en el Hospital General donde adquirí aptitudes y pude poner en marcha los conocimientos y técnicas para poder realizar parte del trabajo experimental.

A mis lectores **Dra. Ángeles Tecalco** y **Dr. José Olivares** por haberme brindado su tiempo para apoyarme enriqueciendo y aportando comentarios constructivos al trabajo de investigación.

Y a todos aquellos que son, fueron y tuve el placer de conocer dentro del Posgrado en Ciencias Genómicas de la UACM, y que son parte importante para mi vida. A mis amigos Roger, Enrique, Jose, Iván, Andrés, Blanca, Flor, Marco, Andy, Jaquie, Olga, Fany, Karen, Esther, Sharon, Vic, Karla, Ame, Eucli, por su compañía y gran apoyo con quienes reí y disfrute cada momento.

Dedicatoria

Quiero dedicar este trabajo a mi madre Martha Garmendía Sánchez y en especial a mi padre Luis Meneses Ortega que desde el cielo me cuida, por brindarme siempre su apoyo y amor incondicional y por siempre darme lo mejor, los amo.

A Natalia por motivarme a seguir adelante y crecer juntos por la vida

Con gran cariño a Erika por su amor, compañía y apoyo.

Índice

Índice de tablas	Página
Tabla 1. Generalidades y características del genoma de <i>E. histolytica</i> .	27
Tabla 2. Características de las unidades de arreglos que codifican ARNt	30
Tabla 3. Secuencias teloméricas en distintos organismos	31
Tabla 4. Secuencias de ADN relacionadas con repetidos teloméricos	50
Tabla 5. Genes que codifican para proteínas TRF ortólogas en Entamoeba, Acanthamoeba y Naegleria	61
Tabla 6. Predicción de la probabilidad de formación de puentes disulfuro en la proteína EhTRF-like II	63
Tabla 7. Motivos identificados y residuos fosforilados para la proteína EhTRF-like II	65

Índice de figuras	Página
Figura 1. Prevalencia y distribución geográfica de las infecciones por Entamoeba spp en humanos detectada por métodos moleculares	18
Figura 2. Ciclo de vida del parásito <i>E. histolytica</i>	19
Figura 3. Factores claves involucrados en la virulencia de <i>E. histolytica</i>	23
Figura 4. Organización del genoma de <i>E. histolytica</i> HM-1: IMSS	28
Figura 5. Análisis de isocoros de 6 diferentes <i>scaffolds</i> de los cromosomas de <i>E. histolytica</i>	29
Figura 6. Telómeros en organismos eucariontes	33
Figura 7. Estructura de los dominios de proteínas TRF1 y TRF2	35
Figura 8. Dominio de unión a ADN Myb de TRF1 y TRF2	37
Figura 9. Modificaciones postraduccionales de las proteínas TRF1 y TRF2	39
Figura 10. Representación esquemática de los dominios, modificaciones postraduccionales y estructura de las proteínas, TRF1 y TRF2 de <i>H. sapiens</i> y las proteínas EhTRF-like I, II y III de <i>E. histolytica</i> , respectivamente	40
Figura 11. Localización de proteínas EhTRF-like en <i>E. histolytica</i>	42

Figura 12. Las proteínas EhTRF-like I y III se unen a las secuencias STR de ADN presentes en los arreglos de ARNt de <i>E. histolytica</i>	43
Figura 13. Modelo de interacciones proteína-proteína en la región telomérica de <i>E. histolytica</i>	44
Figura 14. Obtención del anticuerpo anti-EhTRF-like II	55
Figura 15. Dominio de unión a ADN de las proteínas homólogas a HsTRF en protozoarios de los géneros Entamoeba, Acanthamoeba y Naegleria	58
Figura 16. Árboles filogenéticos de las proteínas TRF-like de los géneros Entamoeba, Acanthamoeba y Naegleria	59
Figura 17. Localización de genes <i>ehtrf-like I, II y III</i> dentro de los <i>contigs</i> en los genomas de amibas del género Entamoeba	60
Figura 18. Motivos conservados en el DBD de proteínas TRF-like en el género Entamoeba	62
Figura 19. Modificaciones postraduccionales predichas en la proteína EhTRF-like II	64
Figura 20. Estructura tridimensional del DBD-EhTRF-like II	66
Figura 21. Modelos moleculares de la interacción ADN-proteína entre el DBD-EhTRF-like II y diferentes secuencias de ADN	68
Figura 22. La interacción del DBD-EhTRF-like II con el ADN muestra una orientación espacial similar a la del DBD-HsTRF1	69
Figura 23. La región del N-terminal <i>linker</i> interacciona con el surco menor del ADN	70
Figura 24. Efecto de mutaciones en el DBD-TRF-like II sobre las interacciones con las secuencias de ADN	73
Figura 25. Estructuras tridimensionales de los complejos formados entre el DBD-EhTRF-like II y HsTRF1 con diferentes secuencias de ADN	74
Figura 26. Estructuras tridimensionales de los complejos formados entre las mutantes del DBD-EhTRF-like II y las diferentes secuencias de ADN	75
Figura 27. Titulación del anticuerpo α -EhTRF-like II	76
Figura 28. Inmunodetección de la proteína EhTRF-like II en los extractos totales y fracciones celulares de <i>E. histolytica</i>	77

- Figura 29.** Localización de la proteína EhTRF-like II en trofozoítos de *E. histolytica*. 79
- Figura 30.** Localización de la proteína EhTRF-like II en los núcleos de los trofozoítos de *E. histolytica*. 80

	Página
1. Resumen	14
2. Abstract	15
3. Introducción	16
3.1 Epidemiología de la amebiasis	17
3.2 Ciclo de vida de <i>Entamoeba histolytica</i>	18
3.3 Factores de virulencia	19
3.4 Resistencia al estrés oxidante	23
3.5 Genoma y organización genómica de <i>E. histolytica</i>	25
3.6 Cromosomas de <i>E. histolytica</i>	28
4. Antecedentes particulares	31
4.1 Telómeros	31
4.2 Características y funciones de las proteínas TRFs	33
4.3 Modificaciones postraduccionales en las proteínas TFR1 y TRF2	37
4.4 Proteínas EhTRF-like en <i>E. histolytica</i>	39
4.5 Localización de las proteínas EhTRF-like I y III	41
4.6 Interacción las proteínas EhTRF-like I y III con secuencias de ADN	42
4.7 Proteína EhTRF-like II	43
5. Justificación	45
6. Hipótesis	46
7. Objetivo General	47
7.1 Objetivos particulares	47
8. Estrategia experimental	48
9. Materiales y métodos	49
9.1 Análisis bioinformático	49
9.2 <i>Docking Molecular</i>	50
9.3 Cultivo de trofozoítos de <i>E. histolytica</i>	51

9.4 Obtención de extractos de proteínas totales, citoplasmáticas y nucleares de amiba	51
9.5 Cuantificación de extractos proteicos	52
9.6 Separación de proteínas por SDS-PAGE y Western blot	53
9.7 Obtención y titulación del anticuerpo α -EhTRF-like II	54
9.8 Localización subcelular de EhTRF-like II por inmunofluorescencia y microscopía confocal	56
10. Resultados	57
10.1 Identificación <i>in silico</i> de los motivos conservados en el DBD Myb de la proteína EhTRF-like II	57
10.2 Predicción de modificaciones postraduccionales de la proteína EhTRF-like II	63
10.3 Determinación de la interacción del DBD de EhTRF-like II con secuencias tipo STR de <i>E. histolytica</i> mediante <i>Docking Molecular</i>	66
10.4 Análisis de la distribución intracelular de la proteína EhTRF-like II en trofozoítos de <i>E. histolytica</i>	75
11. Discusión	80
12. Conclusiones	86
13. Perspectivas	87
14. Bibliografía	88

Abreviaturas

Ala	Alanina
ADN	Ácido desoxirribonucleico
ALT	Vía alternativa de alargamiento telomérico
Arg	Arginina
ATP	Adenosin trifosfato
ARN	Ácido ribonucleico
BSA	Albumina Bovina Sérica
CDK	Cinasa dependiente de ciclina
DBD Myb	Dominio de unión a ADN tipo Myb
dsADN	Ácido desoxirribonucleico de doble cadena
EhTel	Secuencia tipo STR de <i>E. histolytica</i> (TTAGTATT)
EhTRF-like	Factor de unión a repetidos teloméricos de <i>E. histolytica</i>
EMSA	Ensayo de cambio de movilidad electroforética
Gal/GalNAc	Lectina específica de N-acetil D-galactosamina
Glu	Ácido glutámico
h	Horas
HsTel	Secuencia telemérica canónica humana (TTAGGG)
Ile	Isoleucina
kb	kilobases
kDa	kilodaltones
Mb	Megabases
MET	Microscopía electrónica de transmisión
mL	Mililitro
MW	Masa molecular
µg	microgramos
µL	microlitro
NLS	Señal de localización nuclear
Non Rel	Secuencia no relacionada
PBS	Disolución buffer de fosfatos
pI	Punto isoelectrico
POT1	Proteína de protección a telómeros 1
PTM	Modificaciones postraduccionales
Redox	Reacción de óxido-reducción
RMN	Resonancia magnética nuclear
RMSD	Desviación media cuadrática
SDS	Dodecilsulfato sódico
SDS-PAGE	Geles desnaturizantes de poliacrilamida con SDS
STR	Secuencia repetida en tándem corta (2 a 6 pb)
TBS-T	Solución buffer Tris-Borato
TIN2	Factor nuclear 2 de <i>Homo sapiens</i>
TPP1	Proteína de interacción con TIN2 de <i>Homo sapiens</i>

HsTRF
TRFH

Factor de unión a repetidos teloméricos de *Homo sapiens*
Dominio de homodimerización del factor de unión a
repetidos teloméricos

ARNt
RIPA

Ácido ribonucleico de transferencia
Buffer de lisis compuesto por Tris-HCl, EDTA, EGTA,
Triton X-100, ácido desoxicólico, SDS, NaCl, disuelto en
H₂O

1. RESUMEN

Las proteínas TRF1 y TRF2 de humano se unen al ADN telomérico de doble cadena en los extremos terminales de los cromosomas, formando complejos ADN-proteínas que los protegen, regulan su tamaño y evitan las fusiones cromosómicas. En el genoma de *Entamoeba histolytica* no se han identificado secuencias teloméricas, ni está presente un gen que codifique para la enzima telomerasa, y el 10% del mismo corresponde a ARNt asociados a repeticiones en tándem (STR), localizadas en los extremos del 70 % de los cromosomas de este organismo. Por esa razón, se ha propuesto que estos arreglos pueden funcionar como secuencias teloméricas. Adicionalmente, en el genoma de *E. histolytica* se identificaron tres proteínas homólogas a las proteínas TRF1 y TRF2 de *Homo sapiens*, las cuales se denominaron EhTRF-like I, II y III. Los genes *ehtrf-like I* y *II* están contiguos en el genoma, lo cual podría ser el resultado de una duplicación. El porcentaje de identidad del dominio Myb de unión a ADN (DBD Myb) de estas proteínas con sus homólogos humanos es del 93.9%. En este dominio se identificaron los 3 motivos presentes en las proteínas homólogas humanas: N-terminal *linker*, parche ácido y Telebox. En este trabajo se identificaron 18 proteínas TRF-like en los géneros *Entamoeba*, *Acanthamoeba* y *Naegleria*, utilizando la expresión regular [V/L][D/S]L[K/R]DK[W/Y]R, la cual corresponde a la secuencia Telebox. En todos estos géneros, las proteínas ortólogas a TRF contienen en el DBD Myb los tres motivos importantes para el reconocimiento del ADN. Asimismo, se predijo que los dominios TRFH y DBD Myb de la proteína EhTRF-like II pueden modificarse postraduccionalmente, a través de fosforilaciones, SUMOilaciones y la formación de un puente disulfuro, lo que podría modificar su unión al ADN, su localización e interacción con otras proteínas. En esta tesis también se demostró mediante *Docking Molecular* que el complejo formado por el DBD Myb-EhTRF-like II y la secuencia derivada del STR de ARNt de *E. histolytica* (EhTel) tiene una mayor energía de unión y un mayor número de contactos y pseudoenlaces que el formado con la secuencia telemérica canónica (HsTel); sugiriendo que este dominio tiene una mayor afinidad por la secuencia EhTel de *E. histolytica*, que por la secuencia HsTel humana. Por otro lado, la proteína EhTRF-like II se inmunodetectó en la fracción nuclear insoluble de *E. histolytica* HMI-IMSS con un peso molecular de ~50 kDa. Finalmente, se determinó la localización nuclear de EhTRF-like II en la periferia nuclear de los trofozoítos, sugiriendo que ocupan una posición compatible con la protección de las regiones teloméricas en *E. histolytica*.

2. Abstract

Human TRF1 and TRF2 proteins bind to double-stranded telomeric DNA at the terminal ends of chromosomes, forming DNA-protein complexes that protect and regulate their size, and prevent chromosomal fusions. In the genome of *Entamoeba histolytica*, neither telomeric sequences nor a gene coding for the enzyme telomerase have been identified. On the other hand, 10% of the genome corresponds to tRNA associated with short tandem repeats, some located at the ends of the chromosomes of this organism, so it has been proposed that these arrangements could act as telomeric sequences. In the genome of *E. histolytica*, three homologues of TRF1 y 2 proteins were identified, which were called EhTRF-like I, II y III. The genes that encode for EhTRF-like I and II are contiguous in the genome which could be the result of a duplication. These proteins contain a TRFH domain which is involved protein-protein interactions and a MYB DNA Binding Domain (MYBDBD) which recognize the telomeric sequence. The MYB DBD has three important motifs: the N-terminal linker, an acidic patch and the Telebox. In this work we search for TRF-like proteins in all *Entamoeba*, *Acanthamoeba* and *Naegleria* genera using the regular expression [V/L][D/S]L[K/R]DK[W/Y]R which corresponds to the Telebox sequence. We identified 17 Myb-like proteins. In all the genera analyzed the orthologous TRF protein contain in their MYB DBD the three motifs which are important for DNA recognition. Likewise, it was predicted that the TRFH and DBD Myb domains of the EhTRF-like II protein can be post-translationally modified, through phosphorylations, sumoylations and the formation of a disulfide bridge, which could modify its binding to ADN, its location and interaction with other proteins. We analyzed the binding of the DBD-EhTRF-like II with different DNA sequences obtained from de canonical telomeric sequence (HsTel) and the *Entamoeba* STRs (EhTel) using Molecular Docking. Our data showed that the EhTel sequence has higher binding energy and a greater number of contacts and pseudo-bonds than with HsTel sequence, suggesting that this domain has a greater affinity for the sequence derived from the STR of de *E. histolytica* rRNA array, than the canonical telomeric sequence. In this work, we analyzed the localization of EhMybTRF-like II protein in the *E. histolytica* trophozoites though western blot and immunofluorescence assays. The EhTRF-like II was detected in the insoluble nuclear extracts in a 50 kDa band and was located in the nuclear periphery of the trophozoites of *E. histolytica*. This location is as expected for proteins related to the telomeric function.

3. Introducción

Los protozoos pertenecientes al género *Entamoeba* son organismos unicelulares que pueden coexistir como comensales del tracto digestivo humano o como patógenos, por ejemplo, *Entamoeba histolytica*, la cual provoca amibiasis. Esta enfermedad es principalmente entérica, aunque una parte de los individuos infectados presentan síntomas relacionados con la penetración de las amibas al epitelio intestinal, lo cual se conoce como amebiasis invasora (Martínez, 2000). Cuando la enfermedad se vuelve más severa, la diseminación del parásito puede producirse a través del torrente sanguíneo, permitiéndole invadir otros tejidos, en particular el hígado, resultando en abscesos hepáticos (Palomo, 1998). La patogenicidad de *E. histolytica* ha sido atribuida al incremento de la acción concertada de diversos factores que modulan la proteólisis, citotoxicidad, resistencia al complemento e inflamación (Mirelman, 2000; Marie y Petri, 2014). Si se atenúa o inactiva alguno de esos eventos, la virulencia de *E. histolytica* también se modifica. Otros factores que modulan funciones esenciales para la patogenicidad amibiana como la resistencia al estrés oxidante, y el mantenimiento de un sistema eficaz de protección y reparación del daño al material genético causado por el estrés sostenido durante la infección, son de suma importancia para la proliferación y supervivencia de *E. histolytica* (Olivos-García, 2011).

Si bien el control de la amibiasis intestinal humana es relativamente sencillo, en la actualidad no ha sido posible erradicar esta enfermedad. Por eso es importante dar continuidad a los estudios de la biología molecular de este parásito, para entender con profundidad los mecanismos que le permiten proteger y conservar su genoma y de esa forma sobrevivir y proliferar en ambientes adversos.

3.1 Epidemiología de la amebiasis

La amebiasis se define como la infección producida por *E. histolytica*, parásito protozoario cosmopolita incluido tradicionalmente en el phylum *Sarcomastigophora*. Según estimaciones, el 10% de la población mundial presenta la infección, siendo *E. histolytica* la responsable de 100,000 muertes por año, lo que la convierte en la segunda causa de muerte provocada por protozoarios parásitos después de la malaria (Walsh, 1986). A partir del análisis de un total de 110 estudios en 47 países se determinó que la prevalencia global de las infecciones por *E. histolytica*, diagnosticadas por métodos moleculares, era del 3.55% y oscilaba entre 1.72% en Oceanía y Norte América con 21.58% (Figura 1) (Zhaohui Cui, 2019).

En México, se reportaron 191 437 casos nuevos de amebiasis hasta la semana epidemiológica 52 del año 2020 (Boletín Epidemiológico Secretaria de Salud), lo que la coloca como una enfermedad de alto impacto en la salud pública de nuestro país, y pone de manifiesto la insuficiencia en las condiciones de higiene, y el acceso a servicios de salud. Por otro lado, una de las razones por la cual la amebiasis continúa siendo un grave problema de salud en los países pobres es que ningún fármaco, incluyendo al metronidazol, tiene efecto citotóxico sobre los quistes del parásito. Sin embargo, debe reconocerse que desde la implementación masiva del metronidazol en el sistema nacional de salud, la amebiasis invasiva causada por *E. histolytica* han disminuido de forma considerable (Powell, 1966).

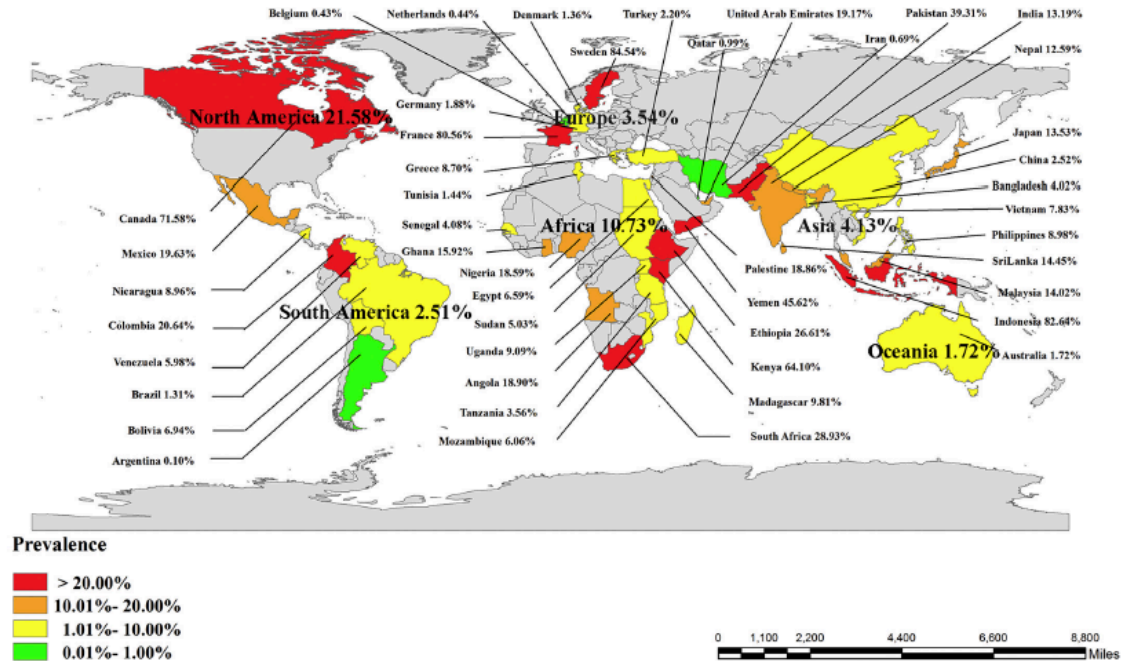


Figura 1. Prevalencia y distribución geográfica de las infecciones por *Entamoeba* spp en humanos detectada por métodos moleculares. Tomado de Zhaohui, 2019.

3. 2 Ciclo de vida de *Entamoeba histolytica*

El ciclo de vida de *E. histolytica* comprende dos estadios bien definidos, el primero dentro del hospedero (humano), dónde se mantiene en forma de trofozoíto, en el lumen del intestino delgado, y es capaz de dividirse para posteriormente invadir el intestino grueso, lugar donde puede permanecer hasta seis meses o más (Palomo A. 1998). Posteriormente, debido a ciertas condiciones del micro-ambiente, el trofozoíto se transforma en quiste, condición en la cual puede ser expulsado del hospedero y sobrevivir fuera de éste, hasta que otro humano lo ingiera, a través de comida o aguas contaminadas con heces de una persona infectada, y así completar el ciclo biológico (Figura 2). Los mecanismos moleculares del proceso de enquistamiento y desenquistamiento del parásito no se conocen en su totalidad todavía. Esto se debe principalmente a que: 1) no hay un modelo animal en el que se pueda reproducir el ciclo biológico del parásito; y 2) hasta la fecha no se ha podido inducir de manera reproducible, en condiciones *in vitro*, el enquistamiento de los trofozoítos de *E. histolytica*, aunque se han identificado algunos genes implicados en el enquistamiento en *E. invadens* (Pumarola, 1991).

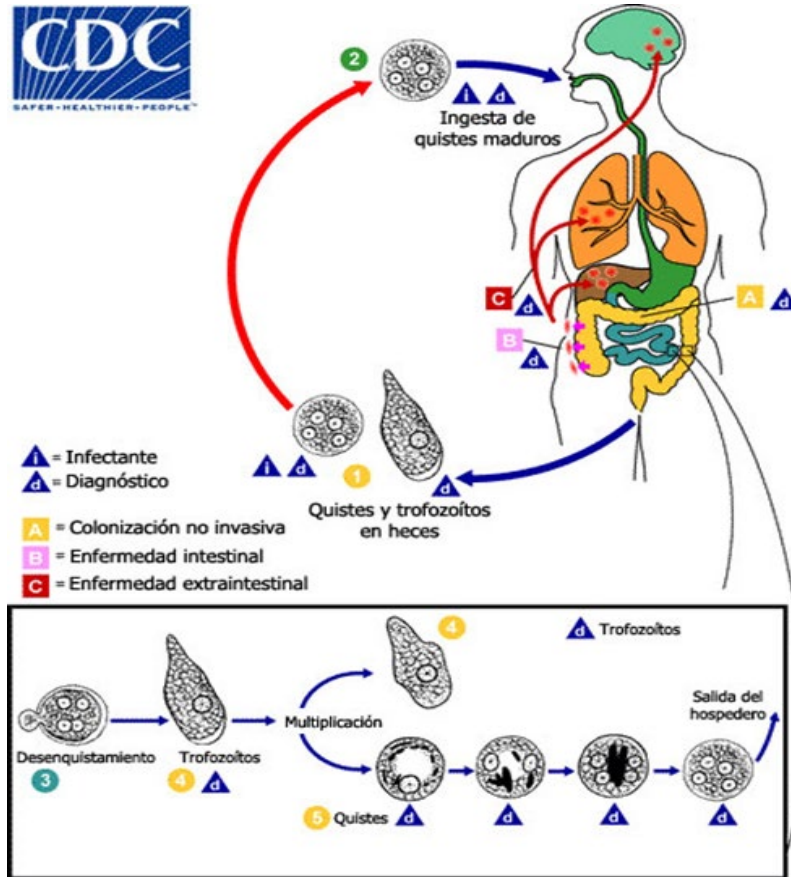


Figura 2. Ciclo de vida del parásito *E. histolytica*. Los quistes de *E. histolytica* provenientes de agua o alimentos son ingeridos y en el íleon, ocurre el desenquistamiento y eclosionan cuatro amibas por quiste. Los trofozoítos liberados alcanzan el lumen del colon, se multiplican, pudiendo invadir y destruir la pared tisular del colon o invaden y alcanzan otros órganos (pulmones, hígado, cerebro, etc.), a través del sistema circulatorio. Los trofozoítos en el lumen pueden enquistarse o mantenerse como trofozoítos y en cualquiera de las dos formas se eliminan en las heces. La mayoría de los trofozoítos pierden su viabilidad al poco tiempo después de excretarse y, los quistes pueden ser ingeridos nuevamente, así el ciclo biológico se completa. Tomado de www.cdc.gov/dpdx/amebiasis.

3.3 Factores de virulencia

La virulencia de *E. histolytica* se caracteriza por su extraordinaria capacidad para penetrar, invadir y destruir tejidos, mediante una serie de eventos en los que participan diferentes factores relacionados tanto con el parásito como con el hospedero, que a continuación se enlistan:

- **Adhesión al epitelio intestinal**

Para que el parásito *E. histolytica* prolifere dentro del lumen intestinal se requiere que los trofozoítos se adhieran a la mucosa intestinal, la colonicen y posteriormente la invadan. La mucosa intestinal está compuesta por distintos tipos de mucina (Muc) unidas mediante puentes disulfuro, siendo Muc2 el principal monómero. Muc2 está compuesta por distintos oligosacáridos que incluyen N—acetil-D-galactosamina (Gal/GalNAc), N-acetil glucosamina (GalNAc), fucosa, galactosa y ácido siálico. Estos compuestos le confieren distintas funciones a la Muc, las cuales incluyen hidratación, formación de gel y resistencia a proteasas. *E. histolytica* ha desarrollado un sistema que le permite degradar con efectividad esta barrera mediante diferentes enzimas como las amilasas, estando mediada su acción por la interacción de la lectina de 260 kDa, a través de la subunidad pesada de 170 kDa específica para residuos de N-acetil-D-galactosamina (Gal/GalNAc). Este parásito también utiliza otras proteínas de anclaje a la mucosa, como las proteínas ricas en lisina y ácido glutámico denominadas KERP1 KERP2; además, proteínas ricas en serina, treonina e isoleucina llamadas STIRP. Tanto KREP1 como STIRP son exclusivas de *E. histolytica* y están implicadas en la adherencia del parásito a las células blancas. Otro complejo importante de proteínas de superficie en *E. histolytica* es el denominado EhCPADH, compuesto por una cisteína proteasa (EhCP112) y una adhesina (EhADH112), las cuales están involucradas en la adhesión, fagocitosis y citólisis de proteínas del hospedero, provocando un aumento en la permeabilidad intercelular y contribuyendo a la desorganización y destrucción de la pared intestinal. Por otra parte, tras un anclaje exitoso a la mucina, favorecido por los factores mencionados, *E. histolytica* libera distintos tipos de moléculas citolíticas (Figura 3), las cuales contribuyen a la virulencia del parásito (Trejo-Suárez., 2008; Olivos-García *et. al.* 2011; Wilson, 2012).

- **Citotoxicidad**

Las proteasas le permiten a la *E. histolytica* degradar la matriz extracelular y producir citólisis de las células epiteliales. Dentro del genoma del parásito se pueden identificar 80 genes que codifican para proteasas, pero 50 de ellos son específicos para cisteín-proteasas (EhCPs), que

pertenecen a la súper familia de la papaína. El 90% de los transcritos de esos genes codifican para EhCP1, EhCP2 y EhCP5. Estas EhCPs se encargan de degradar proteínas de la matriz extracelular como lámina y colágeno, como las regiones C-terminal ricas en cisteína de la Muc, favoreciendo así la unión entre el parásito y las células blancas. Además, las EhCPs contribuyen a suprimir la respuesta inmune del hospedero, degradando moléculas señalizadoras del sistema inmune tales como anticuerpos (Ab) y moléculas involucradas en el sistema del complemento. *E. histolytica* también secreta distintos tipos de glicosidasas como sialidasas, α -glucosidasas y β -N-acetilhexosaminidasa (Que, 2000).

Este parásito también tiene otro complejo de proteínas citolíticas, tipo saponinas (SAPLIP), ubicadas en vesículas denominadas ameboporos (AP) y cuya función está relacionada con la lisis de bacterias ingeridas, y la inducción de la muerte de células epiteliales y leucocitos, durante la invasión del parásito (Figura 3). Los AP, a través de la actividad de las proteínas SAPLIP, son los responsables principales de formar canales iónicos o microporos en las membranas lipídicas de las células del hospedero, por los cuales difunde agua, iones (salida de Na^+ y K^+ y entrada de Ca^{2+}) y otras moléculas pequeñas resultando así en la lisis celular por choque osmótico. *E. histolytica* cuenta con 16 genes que codifican para SAPLIP, pero únicamente AP-A, AP-B y AP-C exhiben actividad formadora de poros. Los AP presentan una estructura y función similares a las N-lisinas, perforinas y granulinas, que se encuentran en las células NK y en los linfocitos T citotóxicos. La secreción de los AP, en conjunto con EhCPs y otros factores de virulencia, trae como resultado la destrucción de los neutrófilos y la consecuente liberación de sus productos tóxicos. Esto amplifica las lesiones causadas por las amibas, produciéndose áreas de necrosis en la mucosa intestinal y focos de ulceración, que condicionan los cuadros clínicos de mala absorción y disentería (Trejo-Suárez, 2008; Olivos-García *et. al.* 2011; Wilson, 2012).

- **Fagocitosis**

Los protozoarios se valen de la fagocitosis como un mecanismo de internalización de nutrientes del medio exterior. El proceso de fagocitosis consta de diversos pasos, de acuerdo con los cambios bioquímicos y morfológicos que ocurren durante la interacción amiba-hospedero, los cuales son: 1) el reconocimiento de partículas o células por la unión de

ligando-receptor; 2) la activación de diversas series de eventos, incluyendo vías de señalización y la reorganización del citoesqueleto, que llevan a la extensión de los pseudópodos; 3) el cierre de los pseudópodos alrededor de la partícula formando el fagosoma, el cual se fusiona secuencialmente con endosomas tempranos o tardíos; 4) y la fusión del fagosoma con el lisosoma, llevando a la digestión de la partícula internalizada. Muchos trabajos realizados sobre este tema indican que la especificidad de las interacciones entre la partícula y los receptores de superficie, activan vías de señalización que guiarán el destino de la partícula fagocitada y, por ende, la respuesta desplegada por la célula fagocítica (Espinosa-Castellano y Martínez-Palomo, 1991; Baxt y Singh, 2008; Boettner *et al.*, 2008). La actividad fagocítica de *E. histolytica*, se considera un indicador de virulencia debido a que los trofozoítos fagocitan y trogocitan células del sistema inmune, reduciendo la respuesta del hospedero. La fagocitosis involucra la reorganización del citoesqueleto, el cual es un proceso en el que están involucradas, además de actina, un gran número de proteínas las cuales son necesarias para que el parásito destruya el tejido del hospedero (Figura 3) (Trejo-Suárez, 2008; Olivos-García *et. al.* 2011; Wilson, 2012). Esta reorganización coordinada del citoesqueleto se encarga de la formación del fagosoma. Esta última estructura madurará principalmente a través de ameboporos y cisteín-proteasas, para la destrucción de la partícula ingerida y que finalmente sea expulsada por exocitosis (Que *et al.*, 2002; Andra *et al.*, 2003).

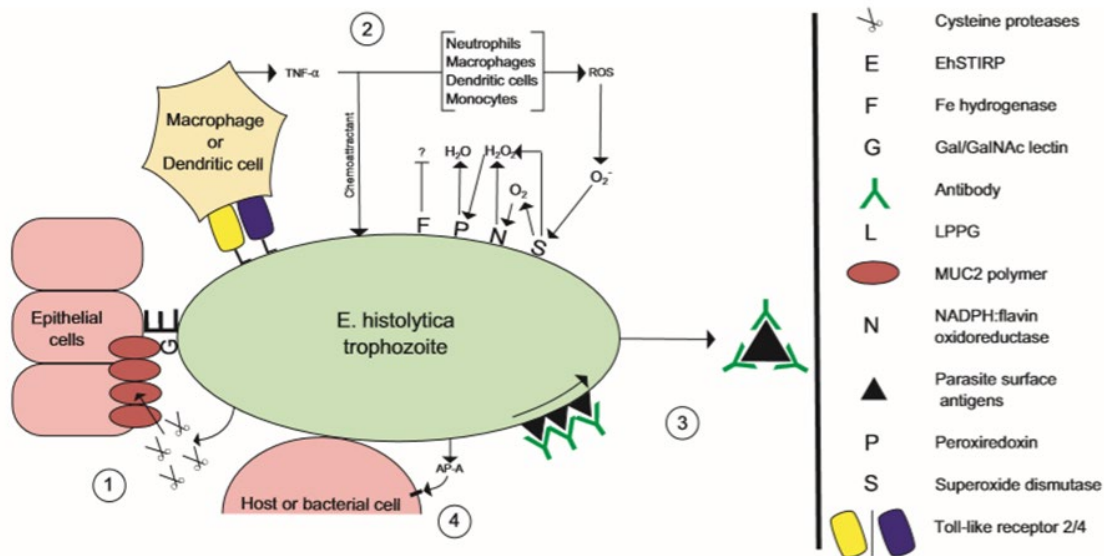


Figura 3. Factores claves involucrados en la virulencia de *E. histolytica*. 1) Unión a la matriz extracelular epitelial a través de lectina Gal/GalNAc y EhSTIRP; y degradación de polímeros MUC2 a través de cisteína proteasas. 2) Resistencia al estrés oxidante provocado por el O_2 y las ROS de la respuesta inmune del hospedero, mediante la vía de reducción de oxígeno. 3) Bloqueo y desprendimiento de antígenos de superficie del trofozoíto por anticuerpos del hospedero y lectinas, que implican el reordenamiento del citoesqueleto y la translocación de complejos antígeno-anticuerpo al uroide de la célula para su eliminación. 4) Contacto directo entre el trofozoíto y la célula huésped o bacteriana, lo que conduce a la secreción de amebaporo-A, que forma poros en la membrana celular diana sin necesidad de receptor. Tomado de Wilson, 2012.

3.4 Resistencia al estrés oxidante

El estrés oxidante se define como un desequilibrio entre los sistemas pro-oxidante y antioxidantes. Cuando predominan las condiciones de estrés oxidante, ocurre daño celular, debido a la oxidación de varias macromoléculas esenciales como el ADN (Guerrat R *et al.* 1981). *E. histolytica* secreta un amplio repertorio de moléculas antioxidantes que le permiten contener con el estrés oxidante y nitrosante, que le provoca el hospedero durante la invasión tisular (Figura 3) (Trejo-Suárez, 2008; Olivos-García *et al.* 2011; Wilson, 2012). A pesar de que este parásito resiste concentraciones bajas de oxígeno en el colon humano (0.1 a 2.3%) (Ladas *et al.* 2007), durante la colonización y la invasión del epitelio intestinal el parásito está expuesto a estrés oxidante, el cual proviene de neutrófilos y macrófagos que secretan una gran variedad de moléculas, tales como especies reactivas de oxígeno (ROS) y nitrógeno (RNS). Las ROS generadas son peróxido de hidrógeno (H_2O_2) y radicales como el anión

superóxido (O_2^-) y el hidroxilo (OH^\cdot), los cuales son productos que se forman durante el denominado estallido respiratorio para la eliminación de patógenos. En general, la mayoría de las ROS se considera dañinas debido a sus efectos sobre el material genético de las células, debido a que oxidan bases en el ADN, preferentemente la base nitrogenada Guanina, generando el aducto 8-oxoguanina (8oxoG). Si esta lesión no es reparada se producen transversiones de las bases del ADN (G:C a T: A), lo cual causa mutaciones y daño al genoma, debido a que también induce roturas en la doble cadena del ADN (David S. *et al.* 2007). Durante el desarrollo del absceso hepático amebiano (AHA), *E. histolytica* está expuesta al oxígeno tisular y a moléculas generadas por el sistema inmune innato (complemento, peróxido de hidrógeno, ácido hipocloroso y óxido nítrico [NO]). Aunque la amiba es un organismo microaerofílico que no tolera más del 5% de oxígeno, la exposición subletal a este gas incrementa su susceptibilidad al complemento, peróxido de hidrógeno y ácido hipocloroso, lo cual ocasiona una destrucción masiva de las amibas (~70%), durante las etapas iniciales de la invasión al hígado (12 h) (Roigothier M, *et al.*, 2002). Es posible que las amibas sobrevivientes induzcan degranulación e inhiban el estallido oxidativo de los polimorfonucleares (PMN), ya que después de este periodo y a pesar de una buena respuesta humoral, la carga parasitaria aumenta de forma continua y termina con la muerte del animal (Olivos-García, *et. al.* 2011).

Las amibas que sobreviven al estrés oxidativo continúan su crecimiento debido a que:

1. El parásito es capaz de reducir el O_2 a radical superóxido (O_2^-). El radical O_2^- , posteriormente es reducido a peróxido de hidrógeno (H_2O_2), mediante la acción de la enzima Fe-superóxido dismutasa (Fe-SOD). El H_2O_2 generado, puede subsecuentemente ser transformado en agua, junto con la participación de otras enzimas como la peroxirredoxinas y ruberitina tioredoxina reductasa (Figura 2) (Hall A, 2009).
2. *E. histolytica* sobre-expresa moléculas antioxidantes (Gonzales M. 2011) y de reparación de los centros de hierro-azufre (Fe-S) de sus proteínas.
3. Los macrófagos del hígado tienen inhibida su capacidad para producir ROS.
4. Los leucocitos polimorfonucleares y macrófagos, que rodean a las amibas durante el desarrollo del AHA, no expresan la enzima iNOS (responsable de la síntesis de NO).

5. *E. histolytica* tiene una familia de proteínas denominadas peroxirredoxinas, las cuales detoxifican el H₂O₂ ayudando a la sobrevivencia del parásito ante el ataque de los macrófagos y neutrófilos (Hall A, 2009).
6. *E. histolytica* tiene un repertorio amplio de proteínas involucradas en la reparación de lesiones producidas por el estrés oxidante del ADN, como las involucradas en la vía de reparación por escisión de bases (vía BER) (Trasviña-Arenas CH, 2019, Azuara-Liceaga E, 2019) y la vía de reparación por escisión de nucleótidos (NER) (López-Camarillo C, 2009; Azuara-Liceaga E, 2019).

E. histolytica expresa también una familia de proteínas relacionadas con el mantenimiento de los extremos de los cromosomas que tienen homología con las proteínas de unión a repetidos teloméricos (TRF). La pérdida de la función de estas proteínas está relacionada con la disfunción telomérica, lo cual conlleva a la fusión cromosómica, senescencia y apoptosis (Hyeung, 2009).

3. 5 Genoma y organización genómica de *E. histolytica*

El genoma de *E. histolytica* es sin duda, una fuente de información muy valiosa que ha permitido realizar un avance significativo en el entendimiento de los mecanismos que le permiten a este organismo parasitar al hospedero, evadir la respuesta inmune, fagocitar, eliminar células blancas y llevar a cabo el proceso de enquistamiento, entre otros. De esta manera, el análisis de la información contenida en el genoma se presenta como uno de los retos actuales más importantes en el estudio de la biología del protozooario parásito *E. histolytica*.

En 2005, Loftus y colaboradores secuenciaron el genoma completo del parásito *E. histolytica*, el cual está compuesto por 23,751,783 pares de bases, de las cuales alrededor del 75% corresponden a adenina (A) y timina (T), distribuidos en 888 segmentos o *contigs*. Sin embargo, posteriormente fue revisado y reanotado, y se predijeron 8,197 genes codificantes para proteínas, pero no se pudieron identificar homólogos para un tercio de las proteínas predichas (31.8%). Actualmente se puede acceder al genoma de *E. histolytica* en el sitio web

<https://eupathdb.org>. Un aspecto importante es que dentro de este genoma no se identificaron secuencias teloméricas canónicas ni el gen que codifica para telomerasa, la cual se encarga de las síntesis de las secuencias teloméricas (Tabla 1). Desafortunadamente, la estructura del genoma predicha por el primer borrador estaba muy fragmentada, debido a la naturaleza repetitiva del genoma, por ejemplo, elementos nucleares intercalados cortos (SINE), elementos nucleares intercalados largos (LINE), arreglos de ARNt que contienen repeticiones en tándem cortas (Loftus *et al.*, 2005; Lorenzi *et al.* 2010), duplicaciones segmentarias y poliploidía, lo que complicaba la caracterización del genoma a nivel cromosómico. Recientemente, con la metodología de secuenciación de nueva generación (NSG) se reanalizó el genoma y se determinó que está compuesto por 38 andamios, con una longitud de 26.879.087 pb (Figura 4). El número de andamios construidos es comparable al sugerido por un estudio anterior, en el cual se resolvieron entre 31 a 35 cromosomas, de más de 300 kb, mediante electroforesis en gel de campos de pulsados (Willhoeft *et al.*, 1999; Riveron *et al.*, 2000). De los 1496 *contigs* del ensamblaje del genoma anterior, se asignaron 1421 al genoma recién ensamblado por BLAST. También se encontró que en los extremos de al menos 23 cromosomas, se ubicaban arreglos de repetidos que codifican para ARNts. Finalmente, en este estudio se identificaron 8734 genes que codifican proteínas (Chang *et al.*, 2020).

Tabla 1. Generalidades y características del genoma de *E. histolytica*

Genoma	Ensamblado
Tamaño (pb)	26.879.087
Contenido GC (%)	24.2
Número de genes	8734
Longitud promedio de genes (pb)	1260.9
Número de genes/10 kb	3.9
Genes Grandes (pb)	15210
Genes pequeños (pb)	147
Porcentaje codificante (%)	49.7
Porcentaje de genes con intrones (%)	24.4
Exones	--
Número	10754
Promedio de número por gen	1.3
Contenido de GC (%)	19.3
Longitud promedio (pb)	962
Longitud total (pb)	10340284
Intrones	--
Número	2553
Contenido de GC (%)	19.3
Longitud promedio (pb)	74.1
Longitud total	189260
Regiones intergénicas	--
Contenido de GC (%)	20.5
Longitud promedio	708.7

Tomado de Lorenzi, 2010; Chang, 2020

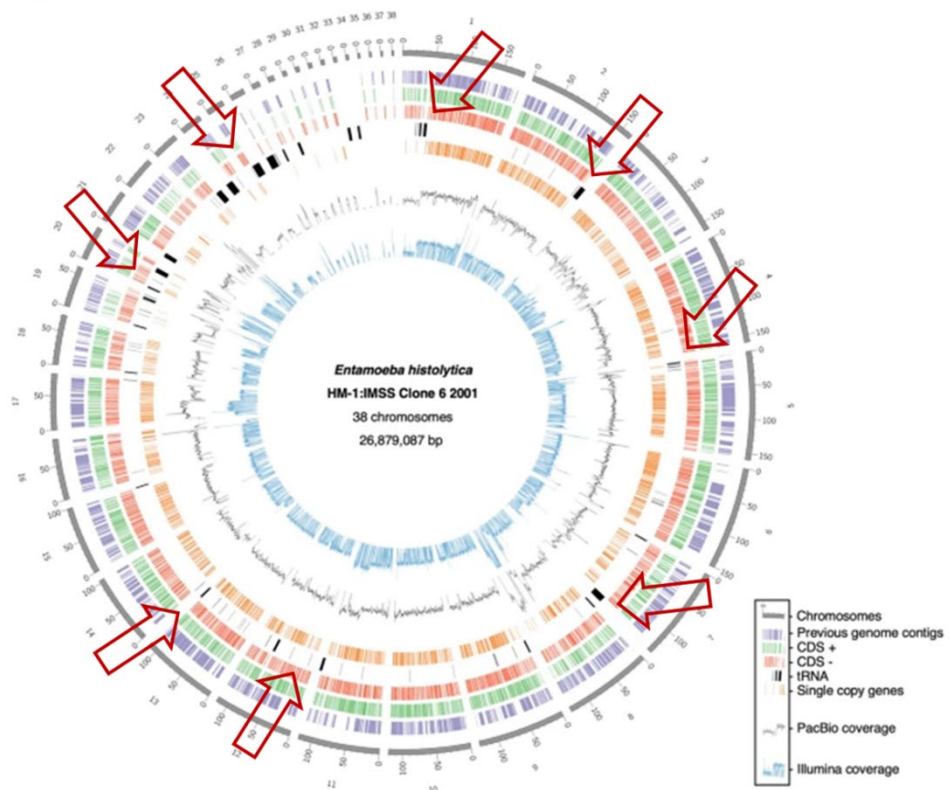


Figura 4. Organización del genoma de *E. histolytica* HM-1: IMSS. En este gráfico circular se muestra la organización y tamaño de los 38 cromosomas (bloques grises; con una escala de 1:10 kb) que componen el genoma de este parásito. Se muestran los *contigs* ensamblados en la versión anterior del genoma (códigos de barras violetas), genes codificantes en la hebra sentido (+, códigos de barras verdes) y anti sentido (-, códigos de barras naranja), arreglos de ARNt (códigos de barras negros), genes de copia única (códigos de barras naranja), y cobertura de lectura mediante PacBio (línea negra) y de Illumina (línea azul). Tomado de Kawano-Sugaya, 2020.

3.6 Cromosomas de *E. histolytica*

En *E. histolytica* se han ensamblado 38 cromosomas y diversos estudios sugieren que el genoma es tetraploide, pero ocasionalmente puede presentar mayor ploidia en los niveles cromosómico y subcromosómico (Kawano-Sugaya, *et al.*, 2020). Estos cromosomas no se condensan en su totalidad por lo que no se pueden observar cromosomas metafásicos como en las células eucariontes superiores (Willhoeft *et al.*, 2000). Además, se ha observado una extensa variabilidad en la longitud de los cromosomas homólogos de diferentes aislados. La variación del tamaño cromosómico observada puede deberse a la expansión y contracción de

repeticiones subteloméricas, como en otros protistas y se especula que en *E. histolytica* estas regiones consisten en arreglos que contienen secuencias que codifican ARNt, lo cual recientemente fue corroborado al identificar su presencia en alrededor del 70% de sus cromosomas. Por lo tanto, se predice que estos conjuntos de genes para ARNt únicos, potencialmente cumplen un papel estructural en el genoma, que podría estar relacionado con su posible función telomérica (Clark C. G., 2006).

Los arreglos que contienen genes para ARNt comprenden casi el 10% del genoma, existen al menos 25 tipos de esos arreglos, y cada uno codifica entre uno y cinco tipos de ARNt por unidad de repetición. Adicionalmente, los arreglos que codifican ARNt tienen repetidos tipo STR, los cuales podrían desempeñar un papel estructural en el genoma, así como brindar protección de las regiones terminales de los cromosomas del parásito (Figura 5).

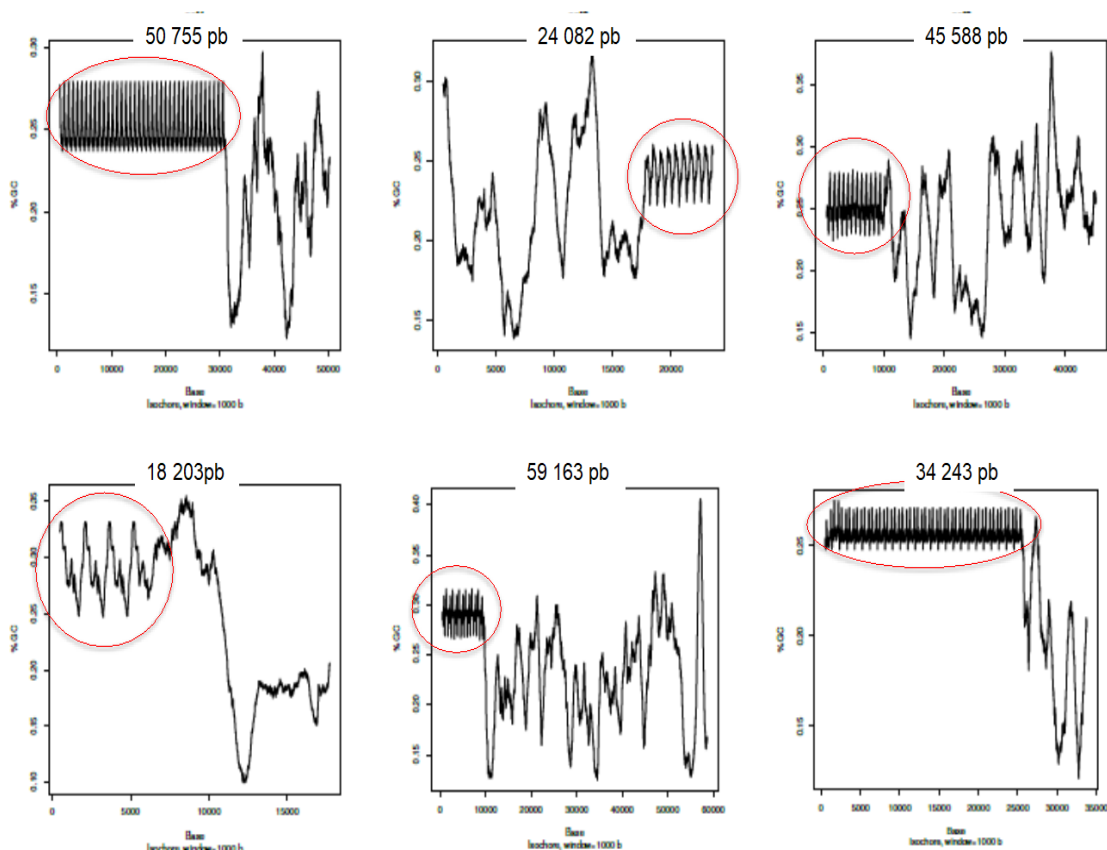


Figura 5. Análisis de isocoros de 6 diferentes scaffolds de los cromosomas de *E. histolytica*. Los scaffolds presentan secuencias repetitivas, las cuales están encerradas en un círculo rojo. El análisis de isocoros fue realizado por el M. en C. Helios Cárdenas Hernández.

Tabla 2. Características de las unidades de arreglos que codifican ARNt				
Nombre del arreglo	Tipo de isoceptor	Unidad longitud (pb)	5S ARN	Núm. Acceso a GenBank
[A ^{AGC}]	Ala ^{AGC}	559	x	BK005648
[ALL]	Ala ^{CGC} , Leu ^{TAA} , LEU ^{CAA}	1154	x	BK005649
[ASD]	Ala ^{TGC} , Ser ^{GCT} , Asp ^{GTC}	1166	x	BK005650
[G ^{GCC}]	Gly ^{GCC}	813	x	BK005662
[G ^{TCC}]	Gly ^{TCC}	490	x	BK005663
[H ^{GTG}]	His ^{GTG}	634	x	BK005664
[LS]	Leu ^{CAG} , Ser ^{CGA}	961	x	BK005667
[LT]	Leu ^{AAG} , Thr ^{AGT}	935	x	BK005666
[MR]	Met ^{CAT} , Arg ^{TCG}	1031	x	BK005653
[NK1]	Asn ^{GTT} , Lys ^{CTT}	1006	x	BK005655
[NK2]	Asn ^{GTT} , Lys ^{CTT}	1251	x	BK005656
[P ^{TGG}]	Pro ^{TGG}	761	x	BK005669
[R ^{TCT}]	Arg ^{TCT}	686	x	BK005654
[R5]	Arg ^{aACG}	846	SI	BK005651
[RT]	Arg ^{CCT} , Thr ^{AGT}	964	x	BK005652
[SD]	Ser ^{TGA} , Asp ^{GTC}	767	x	BK005657
[SSPPCK]	Ser ^{AGA} , Pro ^{AGG} , Pro ^{CGG} , Cys ^{GCA} , Lys ^{TTT}	1775	x	BK005659
[SQCK]	Ser ^{AGA} , Gln ^{CTG} , Cys ^{GCA} , Lys ^{TTT}	1403	x	BK005658
[TQ]	Thr ^{CGT} , Gln ^{TTG}	767	x	BK005660
[TX]	Thr ^{TGT} ,	1107	x	BK005670
[V5]	Val ^{TAC}	812	SI	BK005671
[VF]	Val ^{GAC} , Phe ^{GAA}	965	x	BK005668
[VME5]	Val ^{CAC} , Met ^{CAT} , Gln ^{CTC}	1466	SI	BK005672
[WI]	Trp ^{CCA} , Ile ^{ATT}	1135	x	BK005665
[YE]	Ttr ^{GTA} , Glu ^{TTC}	938	x	BK005661

Se enlistan los tipos de isoceptor de ARNt y su anti-codón. Se proporciona una unidad de longitud, pero a veces están presentes otras variantes de longitud que difieren principalmente en el número de STR. Tomado y modificado de Clark. 2005

4. Antecedentes particulares

4.1 Telómeros

Los telómeros son estructuras esenciales que cubren y protegen los extremos de los cromosomas lineales. Los telómeros protegen los extremos cromosómicos de los mecanismos de detección del daño del ADN y de la maquinaria de reparación. Cuando los telómeros se vuelven disfuncionales, los extremos se detectan como rupturas de ADN bicatenario (rupturas de ADN doble cadena). Este reconocimiento da como resultado la activación de puntos de control del daño en el ADN, que desencadenan la senescencia o la apoptosis, así como eventos de reparación de los extremos terminales de los cromosomas mediante recombinación no homóloga (NHEJ), lo cual genera fusiones inter cromosómicas. La función protectora de los telómeros está mediada por las actividades de los complejos proteicos asociados a los telómeros (de Lange, 2005).

Tabla 3. Secuencias teloméricas en distintos organismos

Organismo	Secuencia Telomérica
Vertebrados	
<i>Homo sapiens</i>	TTAGGG
Artrópodos	
<i>Bombix mori</i>	TTAGG
Hongos	
<i>Saccharomyces cerevisiae</i>	TG(2-3) TG(1-6)
Plantas	
<i>Arabidopsis thaliana</i>	TTTAGGG
Protistas	
<i>Plasmodium vivax</i>	TTTTGGGG
<i>Giardia lamblia</i>	TT(T/C) AGGG
<i>Trypanosoma cruzi</i>	TAGGG
<i>Dictyostelium discoideum</i>	AG(1-8)

Tomado y modificado de Li, 2005 Rodríguez-Centeno, 2019.

En organismos eucariontes, las proteínas que se unen al telómero desempeñan un papel crítico en la regulación de la longitud y protección del cromosoma. Esta maquinaria es conocida como Telosoma o complejo Shelterin, el cual está conformado por un conjunto de 6 proteínas: TRF1, TRF2, RAP1, TIN2, TPP1 y POT1 (Figura 6) (Yang Q, *et al.*, 2007). Las

proteínas TRF1 y TRF2 reconocen la secuencia telomérica, mientras que POT1 reconoce el ADN de cadena sencilla del extremo 3' *overhang*. La subunidad clave del complejo Shelterin es TIN2, ya que se une simultáneamente a TRF1 y TRF2, lo cual contribuye a la estabilización del complejo; mientras que TPP1 interacciona con POT1. El sexto componente RAP1, es un factor asociado a TRF2 (Diotti y Loayza, 2011). Entre los complejos asociados a los telómeros, los complejos Shelterin son únicos ya que reúnen las actividades de unión a ADN de POT1, TRF1 y TRF2. El modelo actual se basa en las interacciones de estos complejos con los telómeros y se extrapolan a partir de las propiedades de unión del ADN con TRF1 recombinante, TRF2 y POT1, las cuales están bien caracterizadas. El complejo Shelterin tiene una alta especificidad para las repeticiones teloméricas de TTAGGG, debido a la presencia de múltiples pliegues de reconocimiento de TTAGGG en el complejo. Los dominios de unión a ADN SANT/ tipo Myb de TRF1 y TRF2 se unen a la secuencia 5'-YTAGGGTTR-3' en el ds-ADN, mostrando muy baja tolerancia a los cambios de una sola base (Figura 6) (Giraud-Panis, 2013).

En *Homo sapiens* existen dos genes que codifican para TRF1 y TRF2. TRF1 regula de manera negativa la longitud de los repetidos teloméricos y promueve la replicación de ellos, y TRF2 está involucrada en el ensamblaje del *T-loop* terminal, cubre y protege del extremo cromosomal, y además regula de manera positiva la longitud de los telómeros (Palm and de Lange, 2008). TRF1 es una proteína con un peso molecular aproximado de 50 kDa, y forma homodímeros para reconocer la secuencia telomérica de doble cadena, a través de su dominio de unión a ADN Myb (DBD Myb) (Figura 7). La dimerización está mediada a través de un dominio llamado TRFH (Figura 7) (Broccoli *et al.*, 1997, Bianchi *et al.*, 1997). Esta proteína está implicada en el mantenimiento de la longitud de los telómeros, la progresión del ciclo celular, así como la reparación de las rupturas de la doble cadena del ADN. En las células que expresan la enzima telomerasa, TRF1 actúa como un regulador negativo del mantenimiento de la longitud de los telómeros, mientras que en células con vías alternativas (vía ALT), TRF1 media positivamente la recombinación homóloga en los extremos de los cromosomas, para el mantenimiento de los telómeros (van Steensel y de Lange 1997; Mogorzewska, 2000).

TRF2 tiene un peso molecular aproximado de 65–69 kDa, también reconoce la secuencia telomérica de doble cadena mediante su DBD Myb, formando homodímeros por el dominio TRFH (Figura 7). La función de esta proteína es bloquear el reconocimiento de los telómeros, a medida que se rompe la doble cadena de ADN. La pérdida de la función de TRF2 conduce a la activación de la quinasa ATM, a la formación de centros inducidos por la disfunción del telómero (TIF) y a la inducción de senescencia o apoptosis. En células que carecen de puntos de control funcionales, los telómeros desprovistos de TRF2 sirven como sustratos para la activación de la vía NHEJ y dan lugar a fusiones inter cromosómicas al igual que las desprovistas de TRF1. La TRF2 purificada *in vitro* puede promover el plegamiento de los telómeros artificiales en estructuras de *T-loop*, una propiedad que no posee TRF1 (Mogorzewska, 2000; Nandakumar y Cech, 2013).

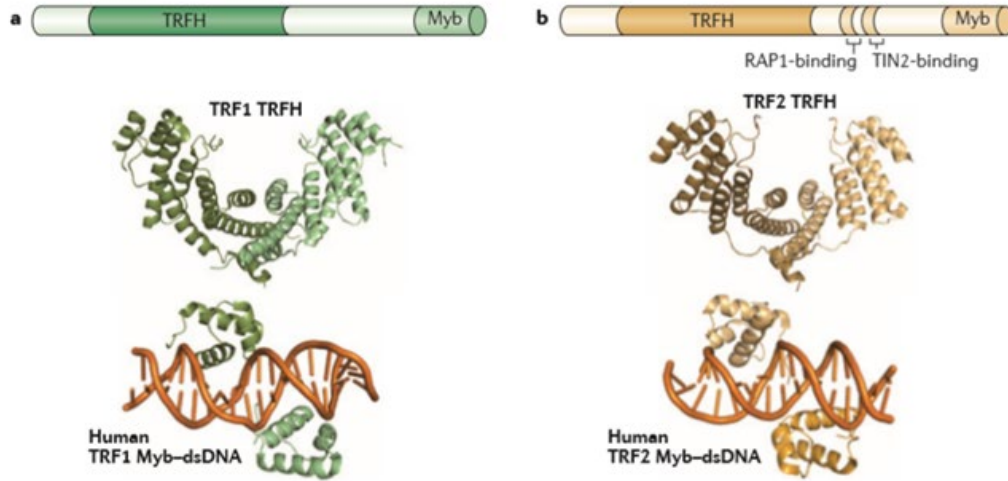


Figura 7. Estructura de los dominios de las proteínas TRF1 y TRF2. Esquema de las proteínas TRF1 (A) y TRF2 (B), donde se muestra el dominio de dimerización TRFH y el dominio de unión a ADN Myb con el cual se unen a ADN de doble cadena. Tomado de Nandakumar, 2013.

TRF1 y TRF2 comparten la misma arquitectura básica definida por dos regiones conservadas: un dominio TRFH que media la homodimerización y un dominio de unión a ADN carboxilo-terminal de la familia SANT / Myb (DBD Myb) (Figura 7) (Broccoli *et al.*, 1997; Biaud *et al.*, 1996; Fairall *et al.*, 2001). El dominio TRFH está en la región N-terminal de la proteína y además de ser responsable de la homodimerización, permite la unión de las TRFs a otras proteínas del Shelterin. En esta misma región se encuentra el dominio de unión al ADN, el cual está involucrado en el reconocimiento de la secuencia telomérica. Los dominios TRFH y de unión a ADN están conectados a través de una región enlazadora larga, presumiblemente no estructurada (Figura 7). Los dominios N-terminales de TRF1 y TRF2 difieren significativamente entre sí, ya que en TRF1 posee características ácidas, mientras que en TRF2 tiene características básicas y es rico en residuos de glicina y arginina (dominio GAR) (Broccoli *et al.*, 1997). Curiosamente, a pesar de la alta similitud estructural, se ha demostrado que las proteínas TRF median distintas funciones biológicas. De hecho, su sobreexpresión o deficiencia desencadenan diferentes respuestas celulares (Walker, 2012)

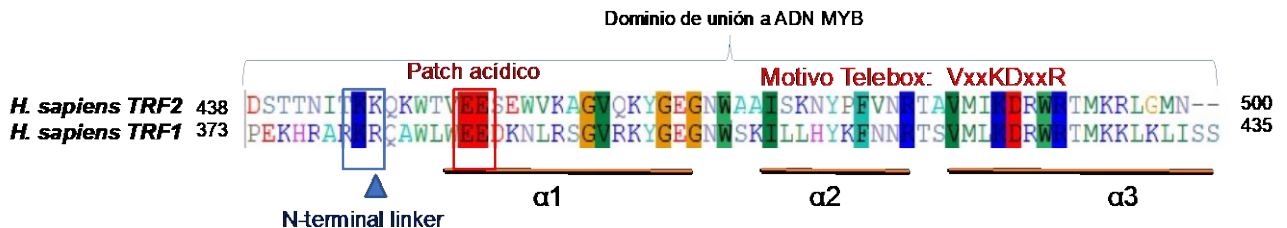
Las proteínas TRF1 y TRF2 reconocen la secuencia telomérica mediante su DBD-Myb. El DBD-Myb está conformado por 3 α -hélices, en una organización hélice-vuelta-hélice, en

donde la tercer α -hélice es la que media el reconocimiento del ADN (Figura 8A). En el DBD-Myb se pueden identificar tres motivos relevantes para el reconocimiento del ADN: N-terminal linker, *patch* ácido y Telebox (Court *et al.*, 2005; Wieczór, 2014). El N-terminal linker se encuentra en una región no estructurada, cercana a la primer alfa-hélice, está enriquecido en aminoácidos básicos como arginina y lisina, y se une al surco menor del ADN (Figura 8A). La región del *patch* está formado por aminoácidos ácidos como ácido aspártico y glutámico, los cuales se unen a las colas del N-terminal de las histonas. Finalmente, el motivo Telebox contiene la secuencia conservada VDLKDKWRT (VxxKDXXR) en la región de la tercer α -hélice y aminoácidos que median la interacción con el ADN (Figura 8A) (Bilaud *et al.*, 1996; Wieczór, 2014). Los complejos TRF-ADN se forman a través de las interacciones entre los residuos de aminoácidos y las bases nucleicas, mediadas por la atracción electrostática entre las proteínas cargadas positivamente y los nucleótidos cargados negativamente (puente salino) y por puentes de hidrógeno, con una contribución menor de la asociación hidrófoba (Figura 8B). En el caso de TRF1, tan solo cuatro residuos de aminoácidos son responsables de la formación de prácticamente todos los enlaces de hidrógeno específicos de secuencia. Tres de estos sitios de reconocimiento (Lys-421, Asp-422, Arg-425) están ubicados en la tercera α -hélice C-terminal, la cual se une al surco mayor del ADN; mientras que la presencia de un solo residuo de Arg (R380), en la región N-terminal del DBD Myb, explicaría la menor especificidad de esta región por el surco menor del ADN (Wieczór, 2014) (Figura 8C).

Los cuatro residuos mencionados son homólogos para TRF1 y TRF2, lo que respalda el supuesto de que el mecanismo de reconocimiento del ADN es muy similar para las dos proteínas. La secuencia VxxKDXXR en la tercera hélice C-terminal está altamente conservada, incluso entre especies relacionadas lejanamente, y sugieren que este motivo es específico para la secuencia T[A/T]AGGG. Por lo tanto, el patrón de conservación observado indica que Val418, Lys-421, Asp-422 y Arg-425 están involucrados en el reconocimiento de secuencias teloméricas. La Arg-380, como residuo de unión al surco menor, parece ser insensible a la sustitución de Arg/Lys, ya que ambos residuos cargados positivamente parecen ser intercambiables en esta posición, en la secuencia de la familia TRF. Sin embargo, estos residuos muestran diferentes patrones de enlaces de hidrógeno, en el surco menor del

ADN. La cadena lateral de arginina está involucrada en la interacción con T10 y A10', mientras que el grupo amino de lisina interactúa casi exclusivamente con T10 (Wieczór, 2014).

A



B

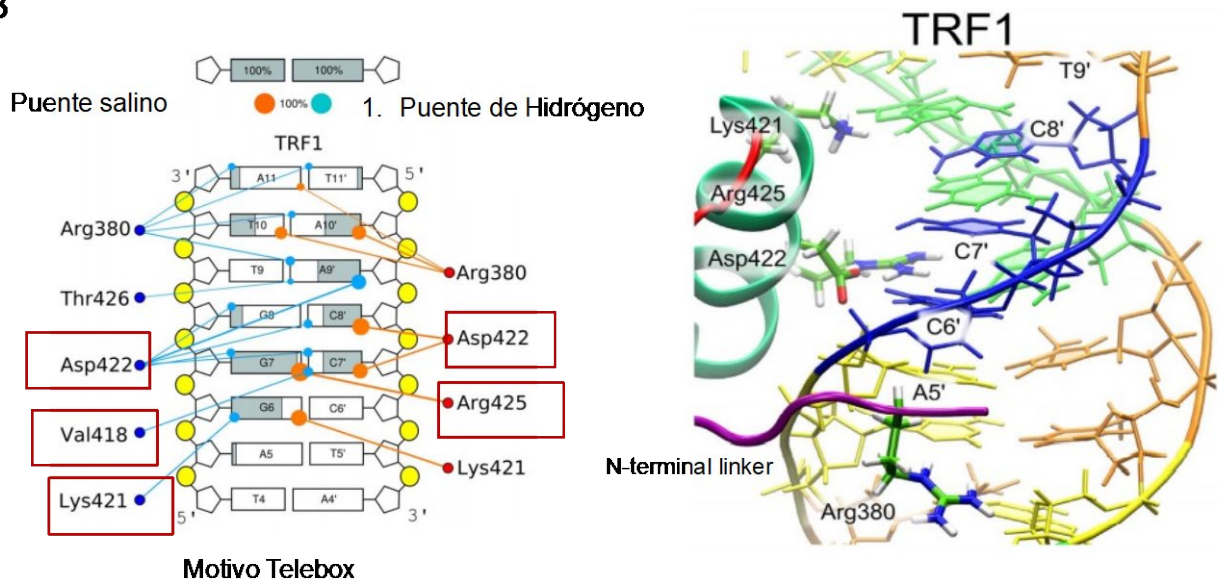


Figura 8. Dominio de unión a ADN Myb de TRF1 y TRF2. (A) Alineamiento de la secuencia de aminoácidos del DBD Myb de TRF1 y TRF2, donde se señalan las 3 alfa hélices y los motivos: N-terminal *linker*, parche ácido y Telebox. (B) Esquema y estructura tridimensional (C) de las interacciones del DBD Myb de TRF1 con la secuencia telomérica, donde se muestran los residuos de aminoácidos que forman interacciones con la secuencia telomérica, puentes salinos (anaranjado) y puentes de hidrogeno (azul). Tomado y adaptado de Wieczór, 2013

4.3 Modificaciones postraduccionales en las proteínas TRF1 y TRF2

Las modificaciones postraduccionales (PTM) de las proteínas inducen una gran variedad de cambios estructurales y funcionales y constituyen un mecanismo esencial de regulación de su actividad biológica (Audagnotto, 2017). En el caso de las proteínas TRF1 y TRF2, su

función está controlada por diferentes PTM, las cuales regulan el mantenimiento telomérico afectando funciones como las interacciones con otras proteínas, la unión a ADN, la dimerización, y median la localización celular, así como su degradación. Entre estas se encuentran la fosforilaciones, SUMOilaciones, metilaciones, parsilaciones y ubiquitinaciones, las cuales se encuentran conservadas en los vertebrados (Broccoli *et al.* 1997) (Figura 9). La fosforilación mediada por proteínas cinasas es una PTM con gran importancia reguladora. En el caso de TRF1, la fosforilación afecta la localización y estabilidad de la interacción con el ADN y está implicada en la respuesta al daño en la vía de reparación del ADN de doble cadena. Además, se encuentra implicada en la regulación de la unión de TRF2 con el ADN telomérico. Por otra parte, TRF1 experimenta poliubiquitinación, lo cual es una señal de que puede ser degradada mediante la vía del proteosoma. También se ha evidenciado que la SUMOilación de TRF1 y TRF2 es importante para el mantenimiento de la longitud y estabilidad telomérica, ya que se reportó que la pérdida de esta modificación resulta en el acortamiento e inestabilidad de los telómeros e inducción de senescencia celular. Además, las proteínas TRF1 y TRF2 SUMOiladas participan en la recombinación homóloga de la vía alternativa de alargamiento telomérico (ALT), en células carentes de telomerasa, lo cual da lugar a los cuerpos APB. A pesar de que la poli-ADP-ribosilación (parsilación) de TRF2 reduce su unión con ADN telomérico, no afecta su participación en el mantenimiento de la integridad telomérica. Sin embargo, la parsilación de TRF1 causa su liberación de los telómeros y la vuelve susceptible a degradación por el proteosoma, por lo que se ha propuesto que este mecanismo permitiría el acceso de la telomerasa a los extremos terminales de los cromosomas (Walker, J. *et al.* 2012).

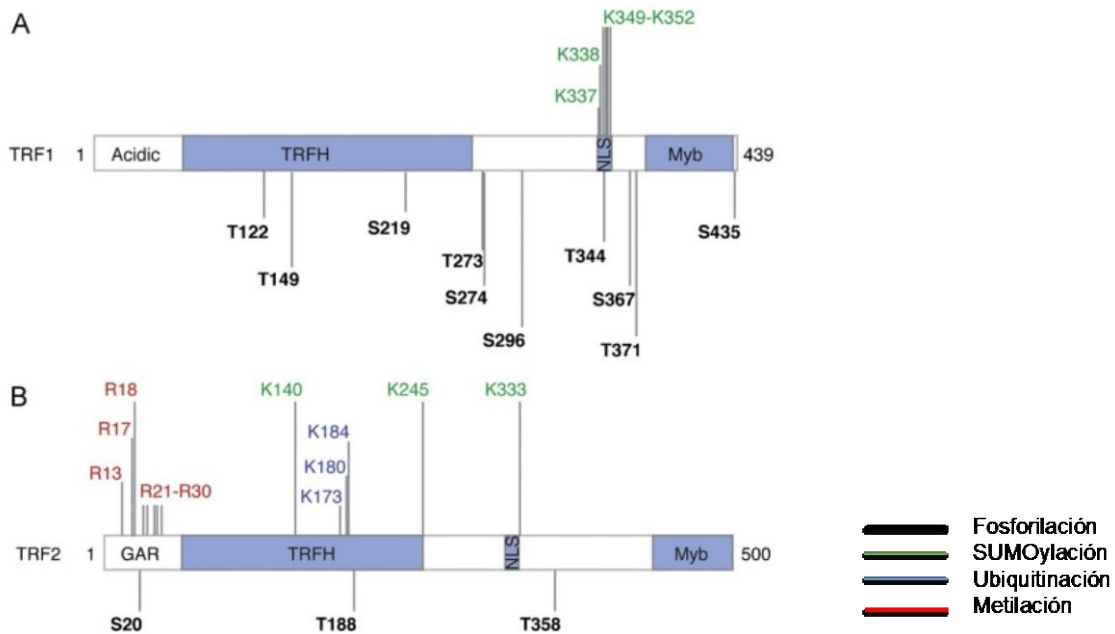


Figura 9. Modificaciones postraduccionales de las proteínas TRF1 y TRF2. En el esquema se muestran los dominios TRFH, Myb y la señal de localización nuclear (NLS) de las proteínas TRF1 y TRF2. Los sitios de fosforilación se indican en negro, ubiquitinación en azul, SUMOilación en verde y los residuos de arginina que se metilan en rojo

4.4 Proteínas EhTRF-like de *E. histolytica*

En el genoma de *E. histolytica* se identificaron 3 genes que codifican para proteínas que contienen un dominio de unión Myb a ADN similar al de las proteínas TRF y que se denominaron proteínas de unión a repetidos teloméricos, EhTRF-like I, II y III. Las proteínas EhTRF-like tienen arquitecturas similares a la de las proteínas TRF humanas. En primer lugar, tienen un dominio de unión al ADN Myb en el extremo carboxilo terminal, el cual tiene conservado el motivo Telebox (VDLKDKWRT). En segundo lugar, tienen un dominio localizado en el extremo N-terminal, de aproximadamente 200 aminoácidos denominado dominio de homología (TRFH), que está relacionado con la homodimerización, pero también está implicado en la interacción con otras proteínas teloméricas. En cuanto a las proteínas EhTRF-like I y III presentan similitud a la HsTRF1, debido a que en el extremo amino terminal se encuentran los residuos acídicos, tal como ocurre en la proteína TRF1; de igual manera tienen similitudes en la masa molecular y en sus puntos isoeléctricos. La proteína EhTRF-like III presenta similitudes con la proteína HsTRF2, ya que presenta en el extremo

amino terminal, residuos básicos y la presencia del motivo de unión a Rap (RBM) (Figura 10).

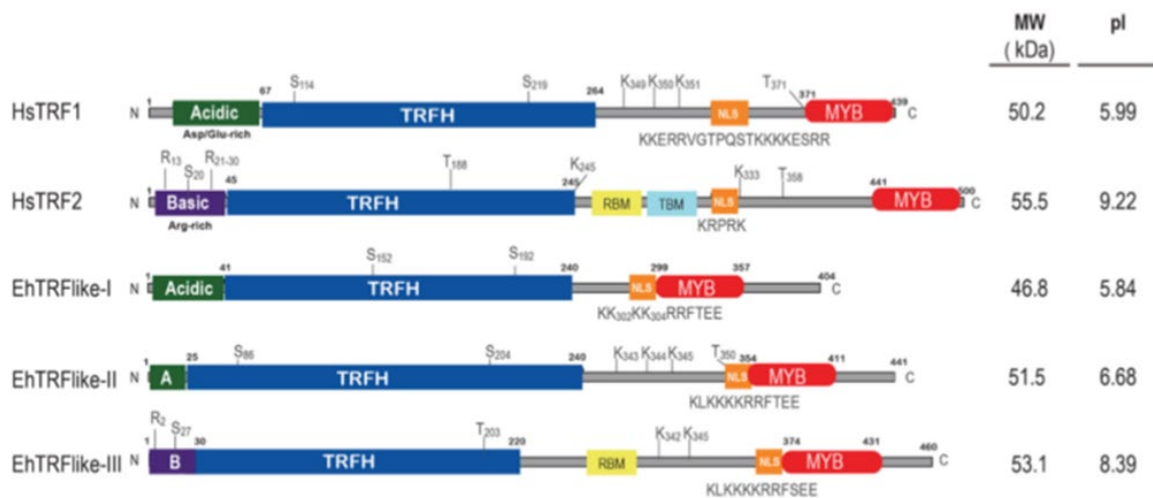


Figura 10. Representación esquemática de los dominios, modificaciones postraduccionales y estructura de las proteínas, TRF1 y TRF2 de *H. sapiens* y las proteínas EhTRF-like I, II y III de *E. histolytica*. TRFH, dominio de dimerización; NLS, señal de localización nuclear; DBD Myb, dominio de unión a ADN; A, amino terminal ácido; B, amino terminal básico; MW, masa molecular; y pI, punto isoeléctrico. Tomado de Rendón-Gandarilla, *et. al.* 2018.

En *E. histolytica*, los genes que codifican para los factores de unión a repetidos teloméricos pudieran haberse originado debido a duplicaciones en el genoma, que le permitieron adaptarse al estrés oxidante. En particular, los genes que codifican para *ehtrf-like I* y *II* están contiguos en el genoma y sus proteínas codificadas presentan un porcentaje de identidad del 93.9%, lo cual sugiere que podrían ser el resultado de una duplicación. La duplicación es un tipo de mutación que implica la generación de una o más copias de un gen o región de un cromosoma. Las duplicaciones génicas y cromosómicas se producen en todos los organismos, aunque son especialmente prominentes en las plantas. La duplicación génica es un mecanismo importante en el proceso de la evolución y pueden estar implicadas en la adaptación y colonización de nuevos nichos. Por ejemplo, durante la evolución, una copia de un gen puede continuar realizando la función original y la otra copia dar lugar a una nueva función (Lawrence C, 2008). Por esta razón, se cree que los genes que codifican para EhTRF-like son duplicaciones, que posteriormente sufrieron neofuncionalización. Así se ha propuesto que en el genoma de *E. histolytica*, pudieron originarse duplicaciones de los genes

ehtrf-like en respuesta evolutiva al mecanismo de respuesta al daño a los telómeros, por estrés oxidativo.

4.5 Localización de las proteínas de EhTRF-like I y III

Las proteínas EhTRF-Like I y III se colocaron con Lamin B1 y H4K20me3, las cuales son marcadores de periferia nuclear y heterocromatina telomérica, respectivamente. EhTRF-like I y III se ubican cercanas a Lamin B1 y H4K20me3, en la periferia nuclear o distribuidas en *focos* (Figura 11A y B), por lo que se propuso que participan en la protección de los extremos terminales del cromosoma de *E. histolytica*. Por otro lado, por medio de microscopía electrónica de transmisión (MET), en trofozoítos que sobreexpresaban a la proteína EhTRF-like III, se localizó a esta proteína enriquecida en las regiones de heterocromatina o cromatina altamente condensada, cerca de la periferia nuclear (Figura 11C). Estos datos muestran que EhTRF-like III se encuentra en regiones de cromatina con alto grado de compactación, compatible con regiones teloméricas (Rendón-Gandarilla *et al.*, 2018).

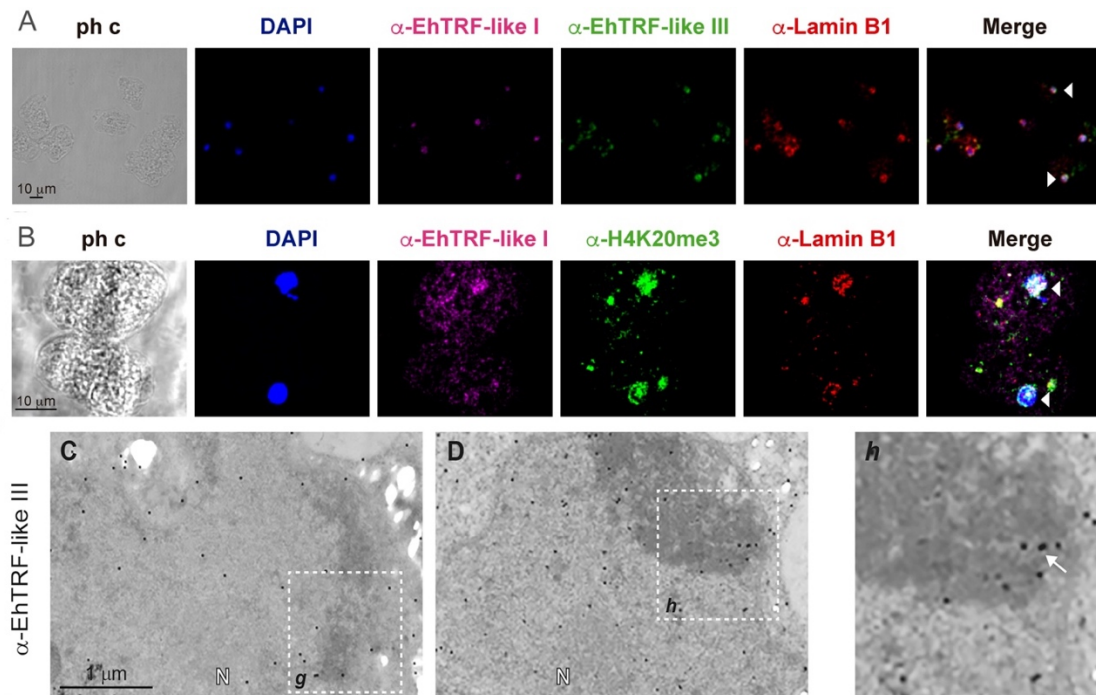


Figura 11. Localización de las proteínas EhTRF-like I y III en *E. histolytica*. A y B) Detección por microscopia confocal de las proteínas EhTRF-like I y III con los marcadores de función telomérica Lamin B1 y H4K20me3. Se muestra a EhTRF-like I (magenta), EhTRF-like III (verde), Lamin B1 (rojo), H4K20me3 (verde) y a los núcleos contrateñidos con DAPI (azul). Las puntas de flechas blancas indican colocalización de las proteínas en la periferia nuclear. ph c, contraste de fase. C) Inmunodetección con oro coloidal por MET de EhTRF-like III, en núcleos de trofozoítos que sobreexpresan a la proteína EhTRF-like III. Panel h: ampliación del recuadro (D) N, núcleos y C, citoplasma. Flechas: partículas de oro. Tomado y modificado de Rendón- Gandarilla *et. al.*, 2018.

4.6 Interacción las proteínas EhTRF-like I y III con secuencias de ADN

Una de las preguntas planteadas en el estudio de las proteínas EhTRF-like I, II y III es si reconocen secuencias teloméricas, debido a que tienen una identidad del 31% en su DBD Myb con las proteínas HsTRF; y si además pueden reconocer los repetidos tipo STR, presentes en los arreglos de ARNts que se proponen como secuencias teloméricas en este organismo. Es por ello, que mediante ensayos EMSA se estudió si la proteína EhTRF-like I recombinante (rEhTRF-like I) podría reconocer la secuencia derivada del arreglo de ARNt [NK1]. En la Figura 12A se observa que la proteína rEhTRF-like I puede formar un complejo ADN-proteína con esta secuencia. De manera interesante, el complejo ADN-proteína compitió específicamente con la secuencia telomérica canónica humana y no así con una secuencia no relacionada (Figura 12A). Este hallazgo demostró que las secuencias derivadas

de los STR pueden ser reconocidas por las proteínas EhTRF y que están relacionadas a la secuencia telomérica canónica (Rendón-Gandarilla, 2018).

La unión de las proteínas EhTRF-like se corroboró empleando un ensayo de ELISA-DPI (*DNA Protein Interaction*) para evaluar la interacción ADN-proteína EhTRF. Para ello, se utilizaron dos secuencias derivadas de los STR de los arreglos de ARNt: secuencia derivada del arreglo [NK1] (EhTR) y secuencia derivada del arreglo [R^{TCT}] (EhTR2); y también se usaron diferentes fragmentos del arreglo de ARNt de [Gly^{GCC}] (Figura 12B). Las proteínas EhTRF-like I y III forman complejos ADN-proteína y tiene especificidad con secuencias STR y diferentes fragmentos de secuencias de los arreglos de ARNt derivados del genoma de *E. histolytica*, lo cual indica que estas secuencias podrían estar relacionadas con secuencias teloméricas (Jordan Alejandro, 2020).

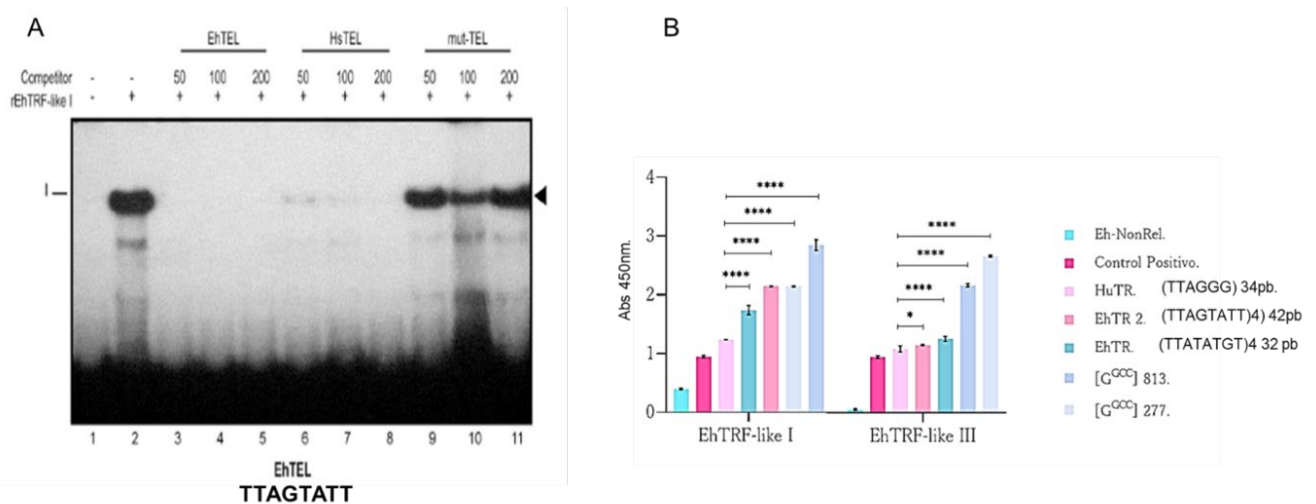


Figura 12. Las proteínas EhTRF-like I y III se unen a las secuencias STR de ADN presentes en los arreglos de ARNt de *E. histolytica*. A) Ensayo de EMSA empleando la proteína rEhTRFlIke I y como sonda marcada radioactivamente la secuencia obtenida de un STR del arreglo de ARNt EhTel (TTAGTATT). Se realizaron competencias empleando concentraciones crecientes de las secuencias EhTel, HsTel y mut-Tel (Rendón- Gandarilla *et. al.* 2018). B) Ensayo de ELISA-DPI para determinar la unión de las proteínas EhTRFlIke I y III a diferentes fragmentos del arreglo de ARNt de [Gly^{GCC}], secuencia derivada del arreglo [NK1] (EhTR) y secuencia derivada del arreglo [R^{TCT}] (EhTR2). Tesis de Jordan-Alejandre, 2020.

4.7 La proteína EhTRF-like II

Las proteínas EhTRF-like I y II presentan una identidad del 93.94% en su DBD Myb y se encuentran ubicadas de manera contigua en el mismo *contig* del genoma de este organismo, por lo que se ha propuesto que pudieron haberse duplicado (<https://amoebadb.org/>). La

5. Justificación

La amibiasis constituye un grave problema de salud en México y en el mundo, por lo que es importante conocer los mecanismos a nivel molecular que permiten la supervivencia y patogenicidad de *E. histolytica* durante la invasión a los tejidos. Por otro lado, se conoce que las proteínas TRFs, en humanos, protegen las regiones teloméricas durante la proliferación y la exposición al estrés oxidante proveniente de la respuesta inmune. Así que el estudio de las proteínas TRF-like de *E. histolytica*, contribuirá a entender los mecanismos de protección de los extremos de los cromosomas de este parásito, en condiciones basales de crecimiento y durante la proliferación celular.

6 Hipótesis

La proteína EhTRF-like II se localiza en el núcleo de los trofozoítos de *E. histolytica* y su DBD presenta motivos conservados que le permiten unirse a secuencias relacionadas con las regiones terminales de los cromosomas.

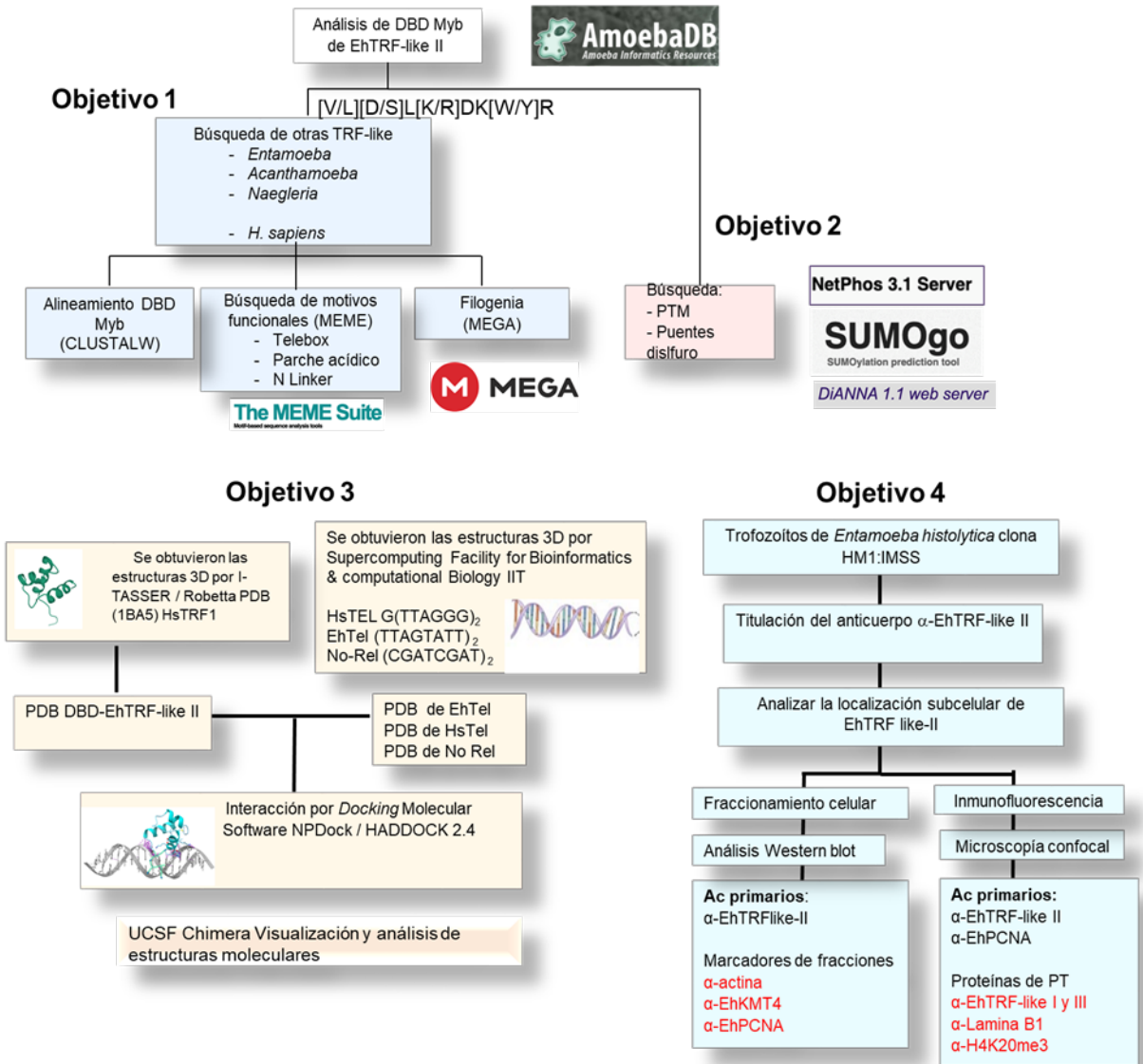
7. Objetivo General

Estudiar la estructura y localización de la proteína EhTRF-like II de *E. histolytica*.

7.1 Objetivos particulares

1. Identificar *in silico* los motivos conservados en el DBD de la proteína EhTRF-like II.
2. Predecir las modificaciones postraduccionales de la proteína EhTRF-like II que modifiquen su interacción con el ADN y su localización subcelular.
3. Determinar mediante *docking molecular* la interacción del DBD de EhTRF-like II con secuencias tipo STR de *E. histolytica*.
4. Determinar la distribución subcelular de la proteína EhTRF-like II en trofozoítos y su colocalización con proteínas relacionadas con la protección telomérica.

8. Estrategia Experimental



9. Materiales y Metodología

9.1 Análisis bioinformático

Para identificar a las proteínas TRF-like codificadas en el genoma de los miembros del género *Entamoeba* y de las amibas *Naegleria fowleri* y *Acanthamoeba castellani*, se utilizó la *expresión* regular [V/L][D/S]L[K/R]DK[W/Y]R, correspondiente al motivo Telebox, como sonda para realizar una búsqueda en la base de datos AmoebaDB de los posibles ortólogos de TRF-like. Las secuencias de aminoácidos del DBD de las TRF-like se alinearon con las secuencias de TRF1 (UniProtKB: P54274) y TRF2 (UniProtKB: Q15554) de *Homo sapiens* empleando el programa ClustalW2.

Para determinar la relación evolutiva entre las proteínas ortólogas a TRF en los géneros de *Entamoeba*, *Acanthamoeba* y *Naegleria*, se realizó un análisis filogenético de las secuencias del DBD Myb mediante el programa MEGA X (<https://www.megasoftware.net/>) y el método de sustitución Neighbor-Joining para 1000 replicados y se construyó el árbol filogenético correspondiente. Este análisis se realizó también con la secuencia completa de aminoácidos de las proteínas TRF-like del género *Entamoeba*, utilizando el programa phylogeny.fr y como método de sustitución Máxima Similitud para 1000 replicados. Las secuencias de las proteínas TRF-like también se analizaron con el programa MEME (meme-suite.org/meme/tolos), para identificar firmas y motivos conservados.

Por otra parte, para predecir sitios susceptibles a modificaciones postraduccionales se utilizó el servidor NethPhos 3.1 (<http://www.cbs.dtu.dk/services/NetPhos/>), que predice sitios de fosforilación de serina, treonina o tirosina en proteínas eucariotas, utilizando conjuntos de redes neuronales, y realiza predicciones tanto genéricas como específicas de cinasas (ATM, CKI, CKII, INSR, PKA, PKB, PKC, entre otras). Para la predicción de SUMOilaciones se utilizó la herramienta SUMOgo (<http://predictor.nchu.edu.tw/SUMOgo/>), la cual utiliza algoritmos que toman en cuenta las propiedades fisicoquímicas y bioquímicas de proteínas o motivos consenso. Para la predicción de puentes disulfuro se utilizó el servidor web DiANNA 1.1 (<http://clavius.bc.edu/~clotelab/DiANNA/>), el cual usa una red neuronal que incluye información de enlaces disulfuro extraídos de proteínas reales, estructura secundaria, regiones flanqueantes y algoritmos para reconocer pares de cisteínas que puedan formar puentes disulfuro. Para realizar los diagramas de la proteína y ubicar las modificaciones

postraduccionales se empleó el software IBS Illustrator for Biological Sequences (<http://ibs.biocuckoo.org/>).

9.2 Docking Molecular

La predicción y modelado de la estructura terciaria del DBD de la proteína EhTRF-like II (DBD-EhTRF-like II) se realizó empleando los servidores Robetta (<https://robeta.bakerlab.org/index.html>), que utiliza un método de predicción de Novo (*ab initio*) e I-Tasser (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>), el cual utiliza modelos comparativos y datos de plantilla, que permiten alineamientos de secuencias para modelado por homología. Como control se empleó la estructura tridimensional del DBD de HsTRF1 (PDB:1ba5). Los modelos tridimensionales de las secuencias de ADN de doble cadena, empleadas en este trabajo, se obtuvieron mediante el servidor web Supercomputing Facility for Bioinformatics & computational Biology IIT (<http://www.scfbio-iitd.res.in/software/drugdesign/bdna.jsp>), el cual predice la estructura tridimensional basado en parámetros conformacionales extraídos de estudios experimentales de difracción de fibras de ADN. Las secuencias empleadas fueron la secuencia canónica telomérica humana (HsTel), la secuencia derivada del arreglo de ARNt [NK1] y una secuencia no relacionada (Tabla 4). La interacción de las diferentes secuencias de ADN y el DBD-EhTRF-like II se realizó mediante *docking molecular* (DM), empleando los servidores NPDock (<http://genesilico.pl/NPDock/>) y HADDOCK 2.4 (<https://wenmr.science.uu.nl/haddock2.4/>). Para visualizar y analizar los modelos de las diferentes estructuras de proteínas, ADN y de los complejos ADN-proteína, se utilizó el programa UCFS Chimera (<https://www.cgl.ucsf.edu/chimera/>).

Tabla 4. Secuencias de ADN relacionadas con repetidos teloméricos

Nombre de la secuencia	Secuencia	Origen
HsTel	GTTAGGGTTAGGG	Secuencia telomérica canónica humana
EhTel	TTAGTATTTTAGTATT	Secuencia derivada del arreglo de ARNt [NK1]
NonRel	CGATCGATCGATCGAT	Secuencia no relacionada

En negritas se muestra el elemento repetido dentro de cada secuencia.

9.3 Cultivo de trofozoítos de *E. histolytica*

Los trofozoítos de *E. histolytica* (cepa HM1: IMSS) se cultivaron a 37 °C en medio TYI-S-33, suplementado con 15% de suero bovino (BIOWEST), y una mezcla de vitaminas Diamond (Sigma-Aldrich) al 1% (Diamond *et al.*, 1978). Los trofozoítos se cosecharon durante la fase de crecimiento logarítmico, enfriando las placas en hielo durante 15 min. Posteriormente, se transvasaron a tubos falcón de 50 mL y se centrifugaron a 1,500 rpm durante 7 min a 4° C, para preparar los extractos de proteínas totales o para realizar estudios de microscopía (Azuara-Liceaga *et al.*, 2018).

9.4 Obtención de extractos de proteínas totales, citoplasmáticas y nucleares de amiba

Para obtener los extractos de proteínas totales después de cosechar los trofozoítos, se lavaron en PBS 1X pH 6.8 (NaCl 137 mM, KCl 2.7 mM, Na₂HPO₄ 8 mM, y KH₂PO₄ 2 mM), a 1,500 rpm por 5 min. Posteriormente, se realizó el conteo celular en una cámara de Neubauer y se colectaron 3x 10⁶ células en tubos Eppendorf, a las cuales se les adicionó Tris-pHMB 100 mM (Sigma-Aldrich), 1 µL de E64 10 Mm (Sigma-Aldrich), 1 µL de PMSF 200 mM (Sigma-Aldrich) y 5 µL de Complete™ Protease inhibidor cocktail (Sigma-Aldrich). A continuación, la suspensión celular se sometió a 3 ciclos de congelamiento y descongelamiento, se centrifugó a 1,400 rpm durante 5 min a 4 °C y finalmente se recuperaron los sobrenadantes. Los extractos obtenidos se cuantificaron por el método de Bradford y se analizó la integridad de las proteínas por SDS-PAGE.

La obtención de proteínas citoplasmáticas y nucleares, así como de los extractos nucleares solubles (Ns) de los trofozoítos, se llevó a cabo empleando el método descrito por Schreiber *et al.* (1989). Para ello, 3x10⁶ trofozoítos que se encontraban en fase de crecimiento logarítmico, se cosecharon por centrifugación a 2,000 rpm, durante 5 min a 4°C. La pastilla se lavó dos veces con PBS 1X pH 6.8. Las pastillas de células se re suspendieron en 400 µL de buffer frío A (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, y 0.5 mM PMSF) y los trofozoítos se incubaron en hielo durante 15 min, para favorecer la turgencia de los mismos, después se añadió 25 µL de solución NP-40 al 10%

(Sigma-Aldrich) y se agitó vigorosamente durante 30 seg en vórtex. El homogeneizado se centrifugó durante 10 min a 14,000 rpm y, se recuperó el sobrenadante. El sobrenadante, contiene los extractos citoplasmáticos y la pastilla que contiene los núcleos de los trofozoítos. Los núcleos se re suspendieron en 1 mL de *buffer A* conteniendo sacarosa 0.34 M, luego se mezclaron y centrifugaron durante 10 min a 14, 000 rpm. El sedimento nuclear se resuspendió en 50 μ L de *buffer C* frío (HEPES 20 mM, pH 7,9; NaCl 0.4 M; EDTA 1 mM; EGTA 1 mM; DTT 1 mM y PMSF 1 mM) a 4°C durante 15 min. A continuación, la suspensión nuclear se centrifugó durante 10 min a 13,000 rpm a 4 °C, y se recuperó el sobrenadante, el cual contiene los extractos nucleares solubles (Ns), y se conservó en alícuotas a -70 °C hasta su uso, mientras que el sedimento nuclear se lisó en *buffer* RIPA (Tris-Cl 50 mM, pH 8, NaCl 150 mM, EDTA 2 mM, Tritón X100 al 1%, y SDS al 0,1%). Posteriormente, para romper la cromatina, el sedimento nuclear lisado en *buffer* RIPA frío se sonicó aplicando tres pulsos de 30 s (30% de potencia) y un período refractario entre ellos de 2 min. Por último, la muestra se centrifugó a 14,000 rpm a 4 °C durante 10 min y el sobrenadante se rotuló como fracción nuclear insoluble (Ni).

9.5 Cuantificación de los extractos proteicos

La concentración de las proteínas extraídas se determinó mediante el método de Bradford (Bio-Rad), para lo cual se construyó una curva estándar que relaciona la absorbancia a 595 nm contra la concentración de BSA. Primero, se diluyó el reactivo de Bradford (Bio-Rad) en una relación 1:4 y se filtró, con ayuda de una pirinola, a través de una membrana de papel Whatman #1 y protegido de la luz. Después, en una placa de 96 pozos, se adicionó a cada pozo 200 μ L de la solución recién diluida de Bradford y 10 μ L de BSA disuelta en agua desionizada estéril a diferentes concentraciones (150, 450, 600, 750 y 1005 ng/ μ L) por triplicado; y se incubó a temperatura ambiente por 5 min. A continuación, se midió la absorbancia de cada pozo en un espectrofotómetro a 595 nm, se graficó la absorbancia contra la concentración de BSA y se ajustó una ecuación de regresión lineal. Para estimar la concentración de proteínas de las muestras se realizó el mismo procedimiento experimental descrito para construir la curva y los valores de absorbancia obtenidos se interpolaron en la ecuación de regresión obtenida. Cada muestra se analizó por triplicado.

9.6 Separación de proteínas por SDS-PAGE y Western blot

Las proteínas de los extractos totales, nucleares solubles e insolubles y citoplasmáticos obtenidos previamente, se separaron mediante electroforesis en geles de poliacrilamida con SDS (SDS-PAGE). Para ello, el gel separador se preparó con acrilamida al 12%, bis-acrilamida al 0.9%, Tris-HCl 0.375 M pH 8.8 y SDS al 0.1% se polimerizó agregando 0.05% de persulfato de amonio (PSA) y 0.03% de TEMED. El gel concentrador se preparó con acrilamida al 5%, bis-acrilamida 0.36%, Tris-HCl 0.125M pH 6.8, SDS al 0.1%, PSA 0.055 y TEMED al 0.005%. Para correr el gel, se cargaron 40 µg/mL de cada extracto de proteínas, resuspendidas en *buffer* de carga 10X (Tris-HCl 1M; 50% glicerol; 1% SDS; 1% azul de bromofenol y 10% β-mercaptoetanol), y calentado a 100°C durante 5 min. La electroforesis se realizó en *buffer* de corrida 1X (Tris base 0.125M; glicina 0.96 M y 0.5% SDS) durante 20 min a 70 V. A continuación, se elevó el voltaje a 100 V durante 10min. Al concluir la electroforesis se tiñó el gel durante 2 h con azul de Coomassie 0.5 g/L (Bio-Rad); posteriormente se destiñó con solución desteñidora (metanol 45% y ácido acético 10%), hasta que quedaran solamente evidenciadas las bandas de proteínas. El gel duplicado sin tinción fue transferido a membranas de nitrocelulosa de 0.45 µm (Bio-Rad), en una cámara semiseca de transferencia Trans-Blot turbo (BioRad) de acuerdo a las instrucciones del fabricante. Para ello se montó el sistema de transferencia en el siguiente orden: ánodo, papel filtro, membrana de nitrocelulosa (Bio-Rad), gel de poliacrilamida, papel filtro y cátodo. Todos los materiales utilizados para el ensayo fueron incubados en *buffer* de transferencia (Tris base 25 mM, glicina 190 mM y metanol 20%, pH 8.3) frío durante 5 min y se mantuvieron ligeramente húmedos durante el montaje. Para llevar a cabo la transferencia de las proteínas se aplicó 20 V durante 20 min con amperaje constante. Una vez concluido este tiempo y siguiendo el protocolo se desmontó el sistema, se colocó cuidadosamente la membrana de nitrocelulosa en un recipiente de fondo plano y se incubó 5 min con rojo de Ponceau S (Sigma-Aldrich). Al terminar la tinción se decantó el colorante y se lavó la membrana con agua desionizada estéril, hasta visualizar las bandas en la membrana.

Para los ensayos de Western Blot (WB), las membranas se incubaron 1 h a 37°C con leche descremada (Carnation) al 1.5%. Al concluir se lavó rápidamente con *buffer* de lavado TBST

(NaCl 137mM, Tris 20mM, Tween 20 0.025%) y se incubaron durante toda la noche a 4°C con *buffer* de bloqueo nuevo, en un rotador serológico (LW Scientific) a 100 rpm. Al finalizar el bloqueo, la membrana se lavó vigorosamente durante 1 h con 6 cambios de *buffer* de lavado. Al concluir este proceso de lavado, la membrana se incubó con el anticuerpo primario α -EhTRF-like II (1: 1,000), toda la noche a 4°C, en un rotador serológico (LW Scientific), a velocidad de 100 rpm. Posteriormente, se lavó la membrana con el mismo programa de lavado descrito anteriormente con *buffer* TBST, y luego se incubaron durante 2 h con los correspondientes anticuerpos secundarios marcados con peroxidasa de rábano (HRP) (1: 8,000, Santa Cruz Biotechnology) en un rotador serológico a 100 rpm. Al concluir la incubación se lavó nuevamente la membrana y se adicionó 500 μ L de sustrato Luminata Forte (Merck-Millipore) y se incubó 20 seg. Las bandas de proteínas se revelaron por quimioluminiscencia en el equipo Chemidoc (Bio-Rad), con periodos de 10 y hasta 100 seg de captura. También se revelaron las imágenes por autoradiografía, utilizando Luminata Forte (Merck-Millipore) y en un cuarto oscuro.

9.7 Obtención y titulación del anticuerpo anti- EhTRF-like II

Para generar el anticuerpo anti EhTRF-like II se diseñó un péptido, a partir del análisis de la secuencia de aminoácidos de la proteína, que permitiera identificar una región diferencial entre las tres proteínas EhTRF-like. Para ello se realizó el alineamiento de la secuencia de aminoácidos de las proteínas EhTRF-like I, II y III, mediante el programa Bioedit (<https://bioedit.software.informer.com/>), para determinar regiones divergentes. Una vez seleccionada la región, se hizo una predicción de epítomos B para finalmente seleccionar el péptido diferencial CIEDKCVAHHIIDQ, el cual se mandó a sintetizar acoplado a KLH (GL Biochem). En la Figura 14A se muestra el alineamiento de las tres proteínas y la secuencia del péptido sintetizada. Antes de la inmunización, se obtuvo el suero pre inmune de un conejo macho Nueva Zelanda de 2 kg. Después, dicho conejo se inmunizó con 500 μ g del péptido por vía subcutánea, con el adyuvante Titermax (Sigma-Aldrich) (durante 5 ocasiones con intervalos de 15 días) (Figura 14B). Para la titulación del anticuerpo α -EhTRF-like II, se empleó el suero proveniente del 4to sangrado y los extractos de proteínas totales de los trofozoítos como blanco. Para ello, estas proteínas se separaron en un gel preparativo SDS-

PAGE al 12%, y se transfirieron a una membrana de nitrocelulosa la cual se bloqueó con 5% de leche en polvo (Carnation) en TBS-T. Posteriormente, diferentes réplicas de la membrana se incubaron con diferentes diluciones (1:100, 1:250, 1:500 y 1:1000) de los sueros pre inmune e inmune durante toda la noche a 4°C. A continuación, las membranas se lavaron tres veces con TBS-T y se les adicionó el anticuerpo secundario (HRP-anti-IgG de conejo) diluido 1: 25,000 en TBS-T por 2 h, a temperatura ambiente y con agitación. Finalmente, las membranas se lavaron tres veces con TBS-T, la señal se reveló con HRP Western Luminata Forte (Merck Millipore) y se visualizó con Chemidoc MP de Bio-Rad.

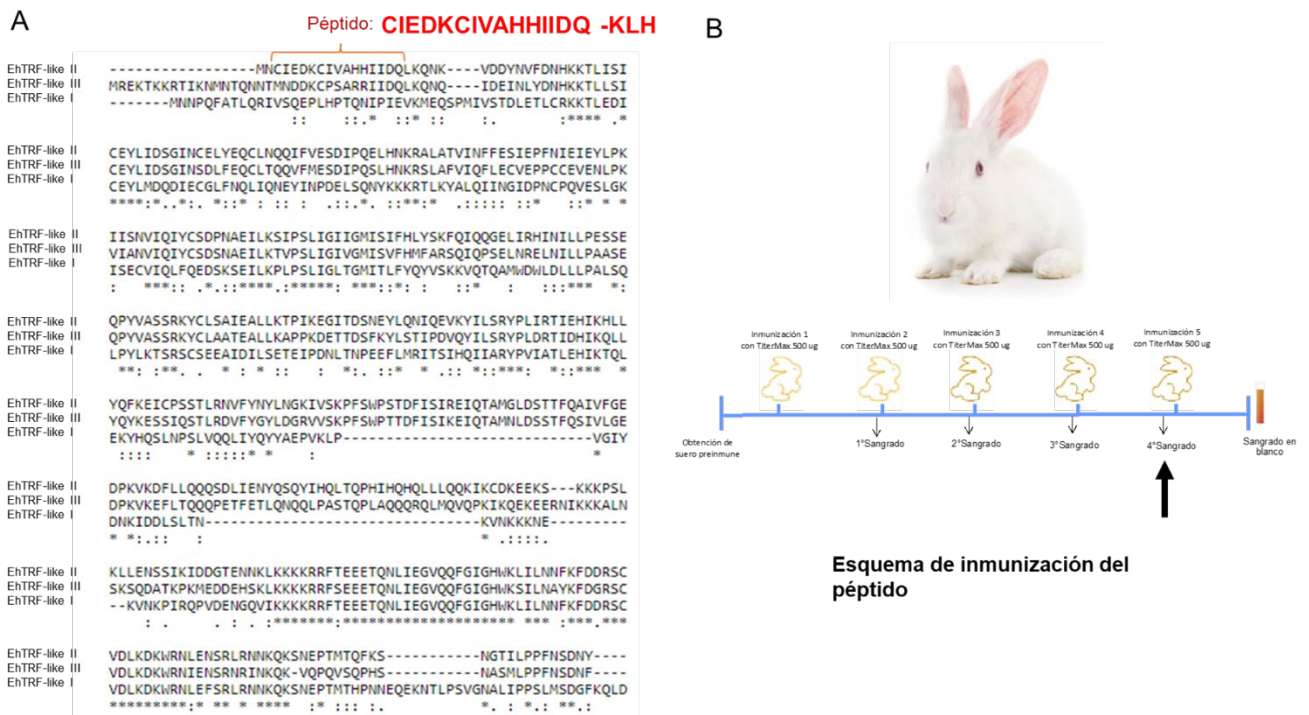


Figura 14. Obtención del anticuerpo anti-EhTRF-like II. A) Alineamiento de las secuencias de aminoácidos de las proteínas EhTRF-like I, II y III con Bioedit y la secuencia del péptido. B) Esquema de inmunización del conejo macho Nueva Zelanda. La flecha señala el sangrado empleado en la titulación del anticuerpo.

9.8 Localización subcelular de EhTRF-like II por inmunofluorescencia y microscopía confocal

Para determinar la localización de la proteína EhTRF-like II se cultivaron los trofozoítos de *E. histolytica* en placas de 6 pozos, conteniendo un cubreobjeto estéril dentro de cada pozo. En cada pozo se inocularon 8×10^5 trofozoítos en 200 μ L de medio de cultivo TYI-S-33 y para que las células se pegaran se incubó durante 30 min a 37°C. A continuación, sin retirar el medio de cultivo, se adicionó 1 mL de etanol absoluto a -20°C por las paredes de cada pozo y se incubó durante 20 min a temperatura ambiente para fijar y permeabilizar las células adheridas a los cubreobjetos. Posteriormente, se eliminó el contenido de cada pozo y los cubreobjetos se lavaron tres veces con PBS pH 6.8, a 37°C por 30 min. Luego, los cubreobjetos se incubaron con 1mL de NH₄Cl 50 Mm, durante 30 min a 37°C, se retiró el NH₄Cl y se adicionó 1mL de BSA/PBS pH 6.8 al 1% y se incubó 30 min a 37°C. A continuación, los cubreobjetos se incubaron toda la noche a 4°C con dos diluciones diferentes (1:50 y 1:100) del anticuerpo α -EhTRF-like II. Posteriormente, las células adheridas a los cubreobjetos se lavaron 4 veces con 1mL de PBS pH 6.8 y se incubaron con el anticuerpo secundario anti-conejo FITC a una dilución 1:250 a 37°C durante 1 h y en la oscuridad. Los portaobjetos se montaron con Vectashield, el cual contiene 10 μ L de DAPI (Vector Lab), sobre la cara del cubreobjeto donde se encuentra el cultivo. Finalmente se selló el porta-cubreobjeto con esmalte transparente, las imágenes se visualizaron a través de un sistema de escaneo confocal de microscopio invertido Nikon (Leica TCS SP2) y se analizaron mediante el software Confocal Assistant, con la asesoría de la Dra. Abigail Betanzos en la Unidad de Microscopía del Cinvestav-IPN.

10 Resultados

10.1 Identificación *in silico* de los motivos conservados en el DBD Myb de la proteína EhTRF-like II.

Las proteínas TRF tienen un dominio de unión DBD Myb de un solo repetido (R2/R3). Este repetido está conformado por aproximadamente 50 aminoácidos que forman tres hélices alfa, siendo la tercera hélice la que está involucrada en la interacción con el ADN (Bilaud *et al.* 1996). En las proteínas ortólogas a TRF en humanos se han identificado tres motivos relevantes para la unión a ADN, entre los que se encuentran: el parche ácido, importante para estabilizar el ADN cuando interacciona con el extremo N-terminal de proteínas histonas; el N-terminal *linker* conformado por lisinas, las cuales se unen al surco menor de la doble cadena de ADN telomérico; y el motivo Telebox (VxxDKxxR) presente en la tercera α -hélice que es quien establece las interacciones con el surco mayor del ADN. Con esta información, en este trabajo nos propusimos identificar a los ortólogos de las proteínas TRF en protozoarios del género *Entamoeba* y en las amibas *Naegleria folwleri* y *Acanthamoeba castellani*. Para ello, se realizó una búsqueda de todas las proteínas TRF-like en la base de datos AmoebaDB, utilizando como sonda la expresión regular [V/L][D/S]L[K/R]DK[W/Y]R, la cual corresponde al motivo Telebox, presente en el DBD Myb. Este análisis se reveló la existencia de 18 secuencias de proteínas TRF-like, codificadas en los genomas de *E. histolytica*, *E. dispar*, *E. moshkovskii*, *E. invadens*, *E. nuttalli*, *N. folwleri* y *A. castellani*. Las secuencias de aminoácidos del DBD de las 18 proteínas TRF-like se alinearon con las secuencias de TRF1 y TRF2 de *Homo sapiens*, empleando ClustalW. El alineamiento mostró que todas las secuencias del DBD Myb analizadas conservan el motivo Telebox (VxxKDxxR) (Figura 15); y que los residuos de valina, lisina, asparagina y arginina, implicados en el reconocimiento específico de la secuencia telomérica en *H. sapiens*, también están presentes en este motivo del dominio DBD Myb de las proteínas TRF-like de estos protozoarios. Sin embargo, cabe destacar que las proteínas de *Acanthamoeba* y *Naegleria* fueron las que presentaron mayor divergencia en la secuencia del DBD Myb. Así, en *Naegleria* se observaron sustituciones conservadas de leucina por valina y de arginina por lisina.

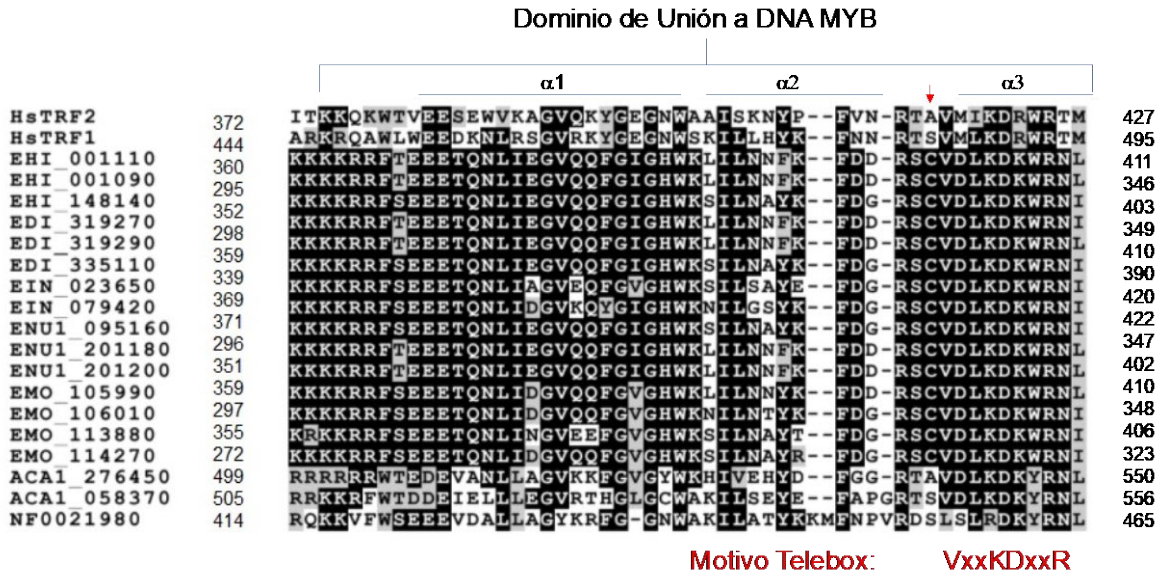


Figura 15. Dominio de unión a ADN de las proteínas homólogas a HsTRF, en protozoarios de los géneros Entamoeba, Acanthamoeba y Naegleria. Alineamiento múltiple de los 51 aminoácidos del DBD Myb, presentes en las proteínas de Entamoeba, Naegleria, Acanthamoeba y *H. sapiens*. En la figura se muestra la posición de las 3 α -hélices, el motivo Telebox presente en la tercer α -hélice y las cisteínas conservadas en el género Entamoeba (flecha roja).

De manera adicional, se determinó la relación evolutiva entre la proteína EhTRF-like II y las proteínas ortólogas a TRF en otras especies de Entamoeba, Acanthamoeba y Naegleria mediante la realización de un análisis filogenético. Empleando las secuencias de aminoácidos del DBD Myb de todas las proteínas TRF-like se construyó un árbol filogenético que agrupó a las proteínas en dos clados. En el clado I se encontraron las proteínas con características similares a HsTRF1 en donde están agrupadas EhTRF-like I y II. En tanto que en el clado II se agruparon las proteínas similares a HsTRF2, donde se ubicó EhTRF-like III (Figura 16A). Las proteínas TRF-like de Acanthamoeba y Naegleria no se agruparon en ninguno de estos dos clados (Figura 16A). Debido a que no se resolvieron las ramas relacionadas con las proteínas similares a EhTRF-like I y II, se realizó un análisis adicional en el que se emplearon las secuencias de aminoácidos completas de todas las proteínas TRF de las proteínas del género Entamoeba (Figura 16B). En este análisis si se pudieron resolver tres ramas, las cuales separaron a las proteínas en 3 clados los de las proteínas tipo TRF-like I, II y III (Figura 16B). El clado de las proteínas TRF-like II es en donde se agrupan un menor número de ortólogos. Las proteínas de *E. invadens* no se agruparon con ninguno de esos clados.

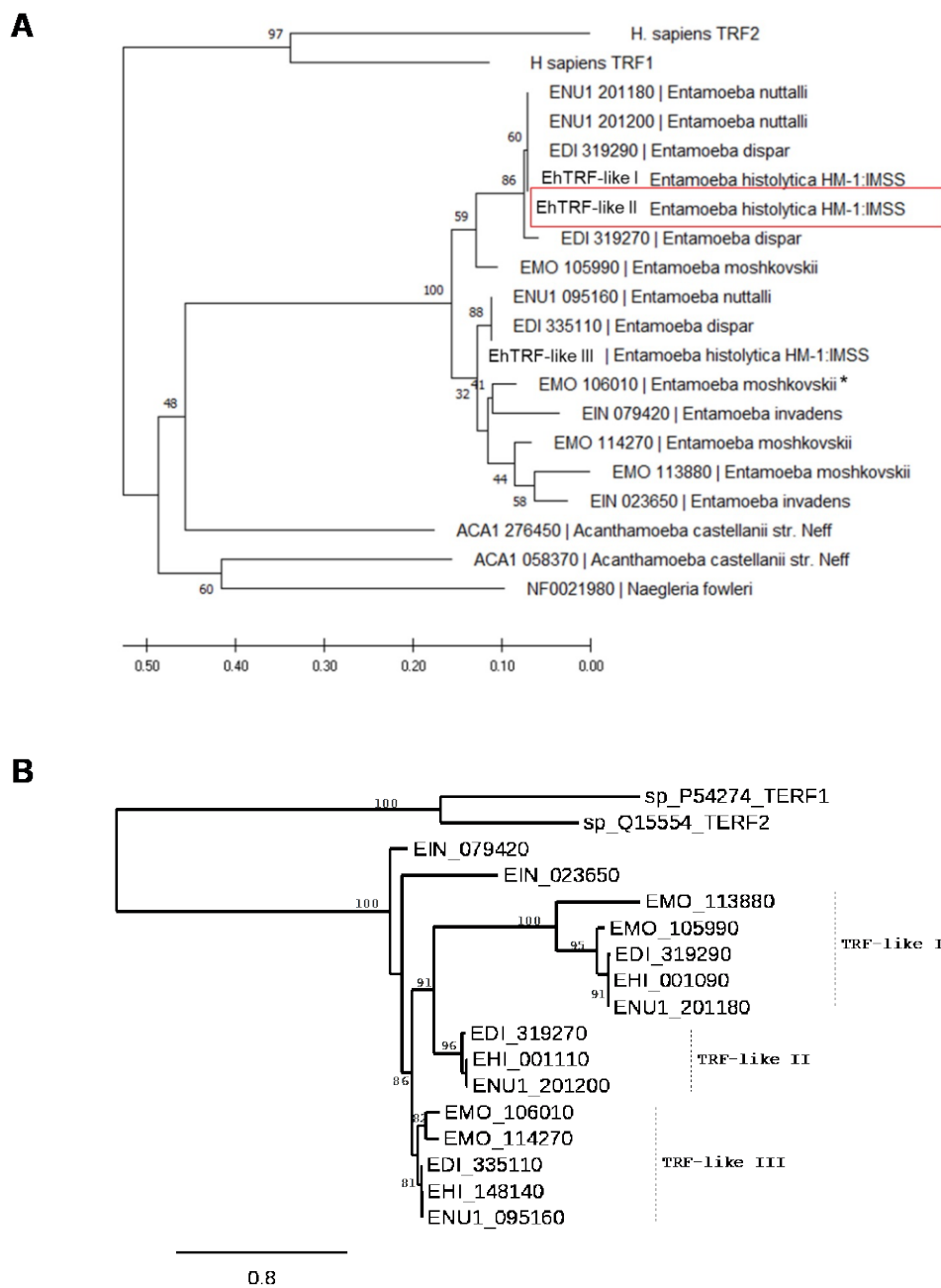


Figura 16. Árboles filogenéticos de las proteínas TRF-like de los géneros Entamoeba, Acanthamoeba y Naegleria. A) Árbol filogenético construido a partir del alineamiento múltiple de las secuencias de 51 aminoácidos del DBD Myb de las 18 proteínas TRF-like. El análisis filogenético se llevó a cabo usando el programa MEGA X y como método de sustitución Neighbor-Joining para 1,000 replicados. Los valores de *Bootstrap* se muestran en los nodos. B) Árbol filogenético construido con la secuencia completa de aminoácidos de las proteínas TRF-like del género Entamoeba. El análisis filogenético se llevó a cabo usando el programa phylogeny.fr y como método de sustitución Máxima Similitud para 1,000 replicados. Los valores de *Bootstrap* se muestran en los nodos.

Se estudió también la ubicación de los genes que codifican para las proteínas TRF-like I, TRF-like II y TRF-like III en los diferentes genomas de los organismos estudiados. Se encontró que el mapa físico de estos genes es semejante en todos los organismos estudiados (Figura 17). En particular, los genes *trf-like I* y *trf-like II* se encuentran contiguos en los genomas del género *Entamoeba*, excepto en *E. invadens* (Figura 17A y Tabla 5); mientras que *ehtrf-like III* está mapeado en otro *contig*, para todos los organismos en estudio (Figura 17B y Tabla 5). En el caso del genoma de *E. invadens*, *Acanthamoeba castellanii* y *Naegleria fowleri* sus genes *trf-like* no se organizan de manera sinténica con respecto a *E. histolytica*.

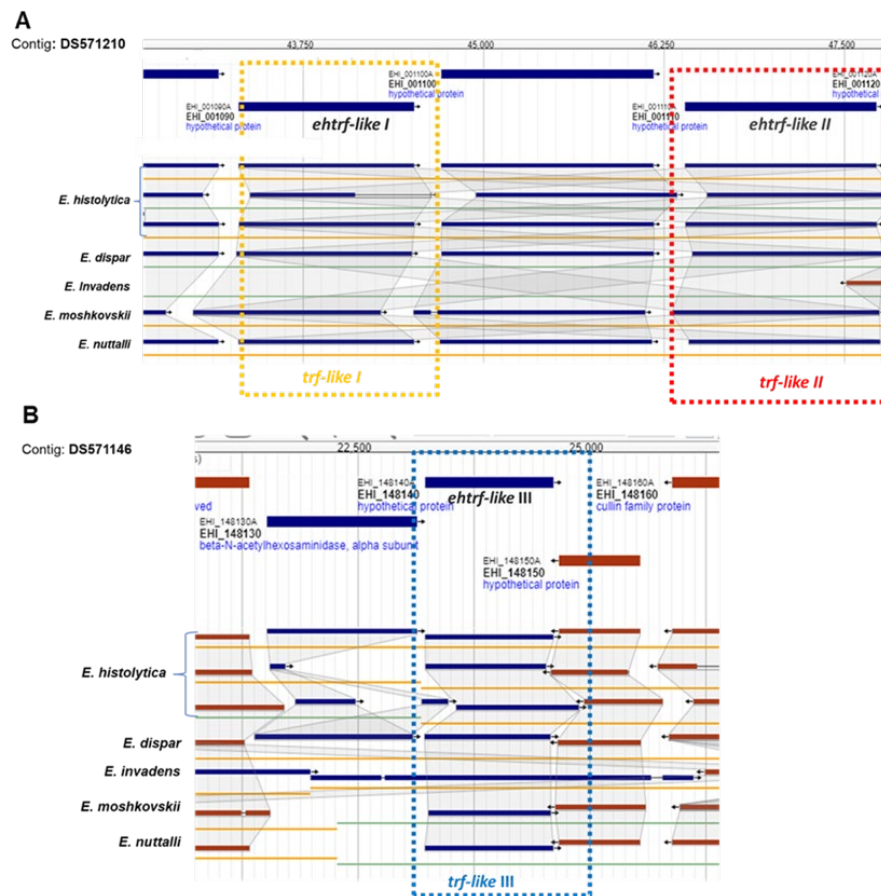


Figura 17. Localización de genes *ehtrf-like I, II* y *III* dentro de los *contigs* en los genomas de amibas del género *Entamoeba*. A) Se muestra los genes *ehtrf-like I* y *II* para *E. histolytica*, *E. dispar*, *E. invadens*, *E. moshkovskii* y *E. nuttalli*. B) En otro *contig* se muestra el gen *ehtrf-like III* para *E. histolytica*, *E. dispar*, *E. invadens*, *E. moshkovskii* y *E. nuttalli*.

En conjunto, estos resultados aportaron elementos adicionales para proponer que las proteínas EhTRF-like II de *E. histolytica*, *E. dispar* (EDI_319270), *E. moshkovskii* (EMO_106010) y *E. nuttalli* (ENU1_201200), son ortólogas (Tabla 5).

Tabla 5. Genes que codifican para proteínas TRF ortólogas en Entamoeba, Acanthamoeba y Naegleria

Organismo	<i>trf-like I</i>	<i>trf-like II</i>	<i>trf-like III</i>	<i>trf-like</i>
<i>E. histolytica</i>	EHI_001090*#	EHI_001110*#	EHI_148140	
<i>E. dispar</i>	EDI_319290*#	EDI_319270*#	EDI_335110	
<i>E. nuttalli</i>	ENU1_201180*#	ENU1_201200*#	ENU1_095160	
<i>E. moshkovskii</i>	EMO_1133880 EMO_105990*#	EMO_106010*#	EMO_114270	
<i>E. invadens</i>				EIN_023650 EIN_079420
<i>A. castellanii</i>				ACA1_058370^ ACA1_276450^
<i>N. fowleri</i>				NF0021980^

*Genes contiguos en el genoma, # Genes sinténicos, ^ De acuerdo a la base de datos AmoebaDB

La búsqueda de los motivos conservados en las secuencias de las proteínas TRF-like de las distintas especies de Entamoeba se realizó empleando la herramienta MEME. Además, se emplearon las secuencias de las proteínas de humano como referencia. De manera interesante, en el DBD Myb de las principales especies de Entamoeba se identificaron los tres motivos conservados en las secuencias homólogas de humano: 1) el motivo Telebox, importante en la afinidad de unión y el reconocimiento específico de secuencia; 2) el parche ácido, enriquecido en ácido glutámico e importante para estabilizar la unión al ADN al interactuar con el extremo N-terminal de histonas; y 3) el 'linker' N-terminal, conformado por lisinas que se unen al surco menor de la doble cadena de ADN telomérico, el cual probablemente media la unión independiente del reconocimiento secuencia específico (Figura 18). Hasta el momento, estos hallazgos sugieren que EhTRF-like II posee en su DBD, otros dominios, además del Telebox, importantes para su unión con el ADN.

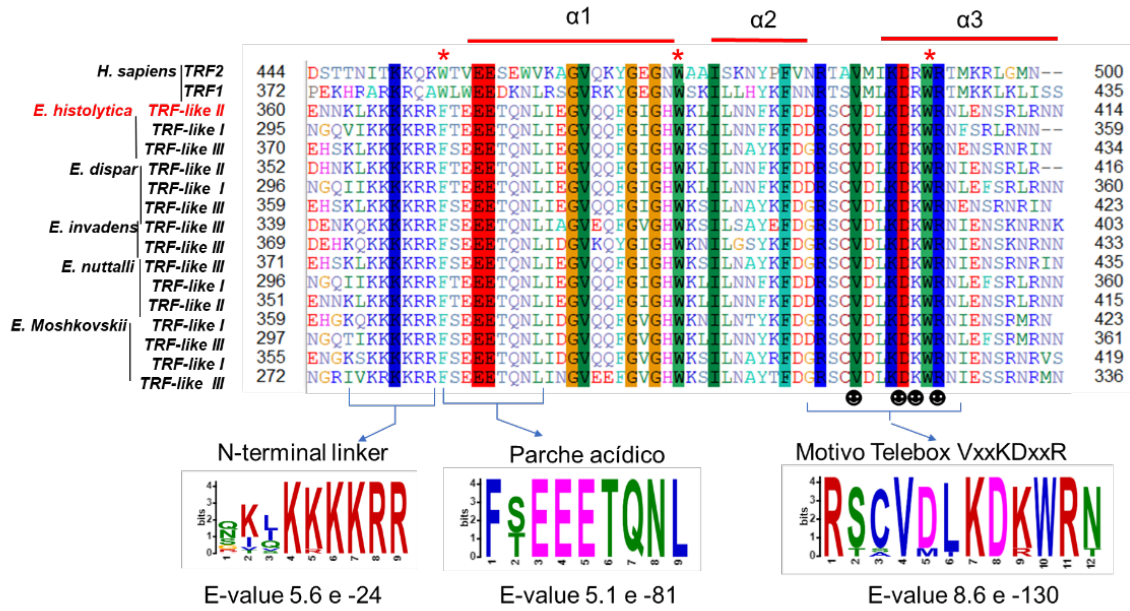


Figura 18. Motivos conservados en el DBD de proteínas TRF-like, en el género Entamoeba. Las tres α -hélices características del DBD y las posiciones de los residuos de triptófano conservados, responsables de la conformación hélice-vuelta-hélice, se muestran en la parte superior, marcadas con barras rojas. En la parte inferior, se muestran las firmas propuestas para las regiones del N-linker terminal, el parche ácido y el motivo Telebox, donde se marcan los aminoácidos involucrados en el reconocimiento de la secuencia. Se muestra también el valor E de cada una de las firmas propuestas.

Finalmente, también, se analizó el dominio de homodimerización (TRFH) de EhTRF-like II. En este dominio se identificaron 6 motivos conservados, los cuales se muestran en la Tabla 6. De estos motivos se identificó el motivo C[L/T][N/T]QQ[IVA]F[M/V/L/A][S/C]D que presenta el residuo Phe104 conservado, este residuo es relevante en la interacción proteína-proteína durante la homodimerización de las proteínas HsTRF (Tabla 6). Adicionalmente, se identificaron otros 7 motivos conservados en este dominio, los cuales se incluyeron para el análisis sobre su susceptibilidad a ser modificados postraduccionalmente.

10.2 Predicción de modificaciones postraduccionales de la proteína EhTRF-like II.

El dominio de unión a ADN (DBD Myb) de las proteínas TRFs puede ser modificado a nivel postraducciona por fosforilación, ubiquitinación, acetilación, SUMOilación y por el estado *redox*. Estas modificaciones pueden afectar la actividad de las proteínas TRF, su presencia en el núcleo, su unión a la secuencia telomérica y su capacidad de interactuar con otras proteínas. En particular, la cisteína, un aminoácido conservado durante la evolución de las proteínas con DBD, está implicado en la unión al ADN, en las interacciones proteína-proteína y en la regulación *redox*. Al respecto, las proteínas TRF-like del género Entamoeba tienen una cisteína conservada en el dominio DBD Myb, que no está presente en las proteínas de los géneros *Naegleria*, *Acanthamoeba* o en *H. sapiens* (Figura 15). Esta cisteína podría estar implicada en la formación de puentes disulfuro en condiciones oxidantes, regulando así la unión de las proteínas TRF-like del género Entamoeba al ADN. Por lo tanto, en este trabajo analizamos la factibilidad de la formación de puentes disulfuro, que tuvieron como resultado la generación de estructuras con diferentes estabilidades de la proteína EhTRF-like II en ambientes *redox*, debido a que las cisteínas conservadas formaran enlaces entre ellas (Tabla 6). Los resultados obtenidos sugieren que al menos tres estructuras, entre las Cys3-Cys226, Cys56-Cys160, Cys109-Cys396 que poseen una alta probabilidad de formarse en la proteína EhTRF-like II. Llama la atención que puente disulfuro formado entre las Cys109-Cys396, pues la Cys396 se encuentra en el DBD Myb y solo está presente en las proteínas TRF del género Entamoeba (Figura 19, Tabla 6). De manera que, la predicción del puente disulfuro entre la cisteína localizada en el DBD Myb solo se predice para la proteína EhTRF-like II y no así para los parálogos EhMyb-like I y III (datos no mostrados).

Tabla 6. Predicción de la probabilidad de formación de puentes disulfuro en la proteína EhTRF-like II, de acuerdo a DiANNA 1.1

Posición de la cisteína	Distancia (aa)	Score	Puente Disulfuro
3-226	223	0.997	XXXMNCIEDKC - QFKEICPSSTL
8-324	316	--	CIEDKCIVAAH - QQKIKCDKEEK
40-50	10	--	TLISICEYLID - SSRKYCLSAIE
56-170	114	0.984	ELYEQCLNQQI - SSRKYCLSAIE
109- 396	257	0.843	VIQIYCSDPNA - FDDRSCVDLKD

En la proteína EhTRF-like II también se predijo que los residuos de serina y treonina son susceptibles a otras modificaciones postraduccionales. Mediante el software NethPhos 3.1, se identificaron 8 motivos conservados, y 6 de ellos contienen residuos de serina y treonina, que pueden ser fosforilados (Figura 19, Tabla 7). En este análisis, se determinó específicamente que los residuos Ser122, Ser218, Thr263 y Thr350 están conservados en las proteínas EhTRF-like II y HsTRF, y en esta última proteína se ha demostrado que se fosforilan (Walker y Zhu, 2012). La fosforilación de estos residuos en HsTRF, afecta la unión de la proteína al ADN y además está implicada en respuesta a daño al ADN (DDR) (Walker y Zhu, 2012), lo cual podría indicar que la fosforilación en EhTRF-like II podría tener consecuencias similares.

Adicionalmente, utilizando el software SUMOgo identificamos residuos de lisina susceptibles de SUMOilación en la proteína EhTRF-like II (Figura 19). La SUMOilación se puede encontrar en organismos eucariotas y tiene un papel importante en la expresión génica, la reparación al ADN la recombinación cromosómica y la señalización celular (Chi-Chang, 2018; Geiss-Friedlander, 2007).

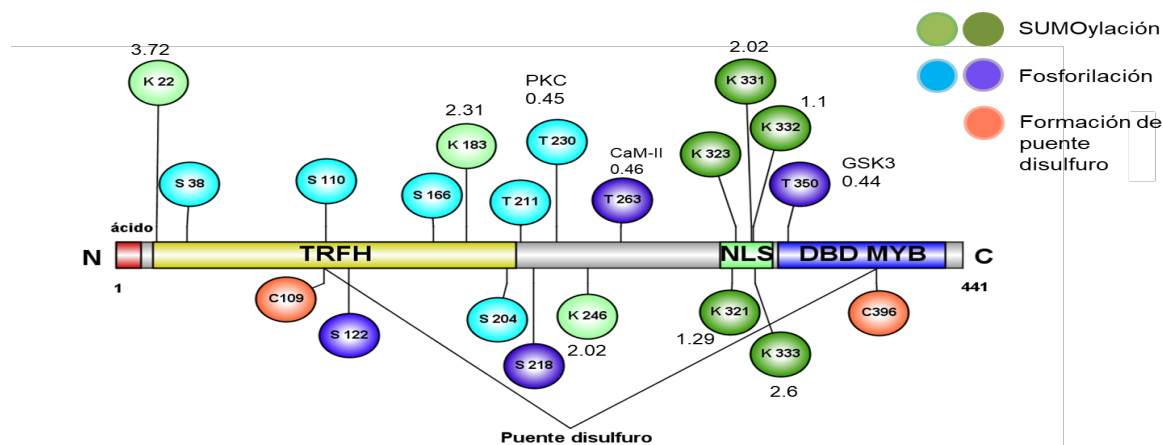










Figura 19. Modificaciones postraduccionales predichas en la proteína EhTRF-like II. Esquema que muestra los dominios TRFH, DBD Myb, el extremo N-terminal ácido y la señal de localización nuclear (NLS) de la proteína EhTRF-like II. Los residuos de serina o treonina fosforilados identificados con el programa NetPhos se señalan en cian y en morado los residuos conservados con HsTRF1. Los residuos de lisina SUMOilados identificados con el programa SUMOgo, se muestran en verde. También se indica el puente disulfuro formado entre Cys109-Cys396 y los *scores* con valores superiores a 0.4. El esquema se realizó en el programa IBS Ilustrator for Biological Sequences.

Tabla 7. Motivos identificados y residuos fosforilados para la proteína EhTRF-like II

Dominio	Motivo predicho*	Valor E	Residuo Fosforilado [^]	Enzima modificadora (Score) [^]
N-Terminal		1.5 e -81	S 38	PKA 0.68
TRFH		3.9 e -37	-----	----
TRFH		1.1 e -81	S 110	CK II 0.56
TRFH		5.4 e -83	-----	----
TRFH		3.1 e -70	S 166	PKC 0.83
TRFH		9.4 e -76	S 204 T 211	PKA 0.55 CDC2 0.48
TRFH		3.7 e -46	S 218	PKA 0.58
DBD		4.7e-99	T350	PKA 0.581

*De acuerdo al programa MEME; [^]de acuerdo a NethPhos 3.1

Con los resultados anteriores podemos concluir que la proteína EhTRF-like II conserva residuos de serina, treonina y lisina los cuales podrían modificarse postraduccionalmente y participar en funciones relevantes de esta proteína como la unión a ADN, su dimerización, localización, degradación e interacción con otras proteínas.

10.3 Determinación de la interacción del DBD de EhTRF-like II con secuencias tipo STR de *E. histolytica*, mediante *docking molecular*

El *docking molecular* (DM) automatizado tiene como objetivo proponer un modelo de unión entre dos moléculas, por medio del entendimiento de las fuerzas de interacción involucradas en el reconocimiento molecular. Nuestro trabajo puso en evidencia la aplicabilidad de los métodos computacionales basados en el DM para estudiar las interacciones proteína-ADN. En primer lugar, se obtuvo el modelo tridimensional del DBD de la proteína EhTRF-like II, mediante los programas I-Tasser y Robetta, y tomando como modelo el PDB (1ba5A), correspondiente al DBD de la proteína HsTRF1. Sin embargo, el mejor modelo se obtuvo con el programa Robetta con un valor de confianza de 0.94 (Figura 20). Esta estructura se superlapó con la del DBD de HsTRF1, encontrándose que las principales diferencias entre ambas estructuras se observan en la zona del N-terminal *linker* (Figura 20).

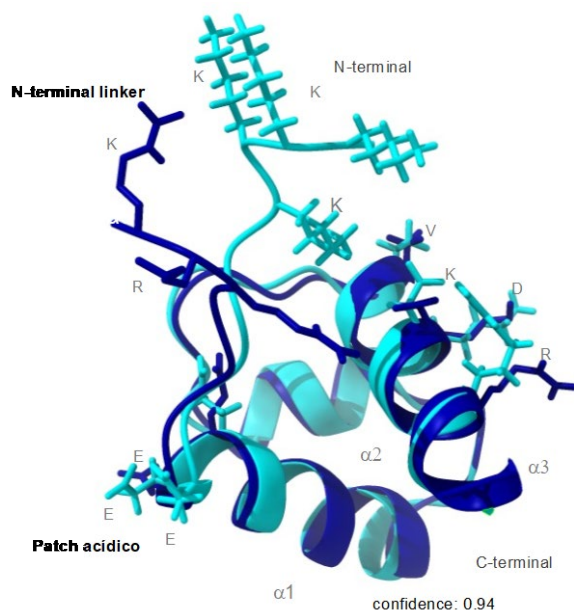


Figura 20. Estructura tridimensional del DBD-EhTRF-like II. Estructuras superlapadas del DBD-EhTRF-like II (cian) y DBD-HsTRF1 (azul) empleando el programa ChimeraX. La estructura 3D del DBD-EhTRF-like II se obtuvo con el programa Robetta, tomando como molde al DBD-HsTRF1 (PDB 1ba5), y la estructura tridimensional se obtuvo con un valor de confianza de 0.94. Se muestran los residuos de los motivos N-terminal *linker*, parche ácido y Telebox.

Para realizar la interacción mediante DM se emplearon diferentes secuencias de ADN, la secuencia telomérica canónica humana (HsTel-TTAGGG), la secuencia NK del arreglo de ARNt [Asn^{GTT}-Lys^{CTT}] de *E. histolytica* (EhTEL-TTAGTATT) y una secuencia no relacionada (Non Rel-CGAT) (Tabla 4). Los modelos tridimensionales de estas secuencias de ADN se obtuvieron mediante el programa Supercomputing Facility for Bioinformatics & computational Biology IIT.

Empleando el programa NPDock se obtuvo la interacción DBD-EhTRF-like II con las tres secuencias HsTel, EhTel y Non-Rel presentando un RMSD de 2 Å. Para la interacción del DBD-EhTRF-like II con HsTel se establecieron 179 contactos, entre las bases nitrogenadas GTTAG y los aminoácidos de la tercer α -hélice: Val397, Arg394, Lys400 y Asp405 (Figura 21A). De manera interesante, se observó que la Lys359 del N-terminal *linker* hace contacto con AT en el surco menor del ADN, con un score (NPDock) de -2.84. La interacción DBD-EhTRF-like II y EhTel mostró aproximadamente 276 contactos, entre las bases nitrogenadas TTTAGTA y los aminoácidos Arg394, Cis396, Val397 Lys402 y Asp401 (Figura 21B). También se observaron enlaces entre las Lys368 y Lys369 del N-terminal *linker*, que hacen contacto con el surco menor del ADN y presentaron un score (NPDock) de -2.82. En la interacción DBD-EhTRF-like II con la secuencia no relacionada, solo se establecieron 48 enlaces y un *score* (NPDock) de -2.64.

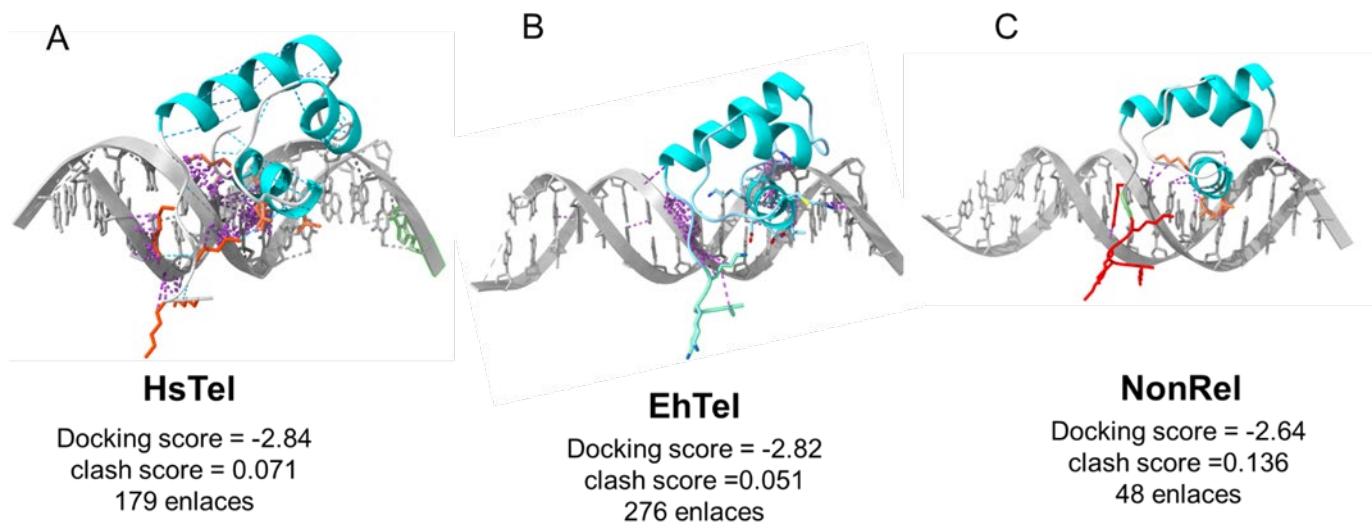


Figura 21. Modelos moleculares de la interacción ADN-proteína entre el DBD-EhTRF-like II y diferentes secuencias de ADN. Interacción del DBD-EhTRF-like II con la secuencia telomérica canónica de A) *H. sapiens* (HsTel, TTAGGG), B) con la secuencia NK del arreglo de ARNt [Asn^{GTT}] de *E. histolytica* (EhTel, TTAGTATT) y C) con la secuencia no relacionada (Non-Rel, CGATCGAT). En líneas punteadas moradas se señalizan los pseudoenlaces. Los acoplamientos moleculares mostrados en esta figura se realizaron con el servidor NPDock

Dado que las secuencias específicas tuvieron más asociaciones que la no relacionada, estos resultados sugirieron que las secuencias derivadas de los STR podrían interactuar con la proteína EhTRF-like II, tal y como se ha demostrado que sucede con las proteínas parálogas EhTRF-like I y EhTRF-like III (Rendón-Gandarilla *et al*, 2018; Jordan-Alejandre, tesis 2020). Finalmente, ya que en la base de datos PDB (Protein Data Bank) ya se encuentra una estructura tridimensional del DBD-TRF1 (1IV6) y la secuencia telomérica canónica (HsTel), se hizo la comparación con las estructuras tridimensionales del DBD-EhTRF-like II y las secuencias HsTel y EhTel. En la Figura 22 se observa que los aminoácidos Val397, Arg394, Lys400 y Asp405 se orientan de manera similar en los DBD de las proteínas HsTRF1 y EhTRF-like II, para ambas secuencias. También se analizó la orientación del motivo del terminal *linker* para la estructura tridimensional del DBD-HsTRF y la secuencia canónica, y

la estructura del DBD-EhTRF-like II y la secuencia EhTel (Figura 23). Los resultados obtenidos son congruentes con los reportados para la proteína HsTRF1 (Figura 23A), en la cual se ha reportado que el N- terminal *linker* del DBD Myb interactúa con el surco menor del ADN (Wieczór, 2014, Nishikawa *et al.* 2001) Cabe destacar que el extremo terminal *linker* del DBD-EhTRF-like II ocupa una posición similar en el surco menor como se ha determinado para el DBD-HsTRF1 (Figura 23B). En el caso de la proteína EhTRF-like II, ésta región tiene una mayor longitud, ya que está formada por los aminoácidos Lys365, Lys366, Lys367, Lys368, Arg369 y Arg370.

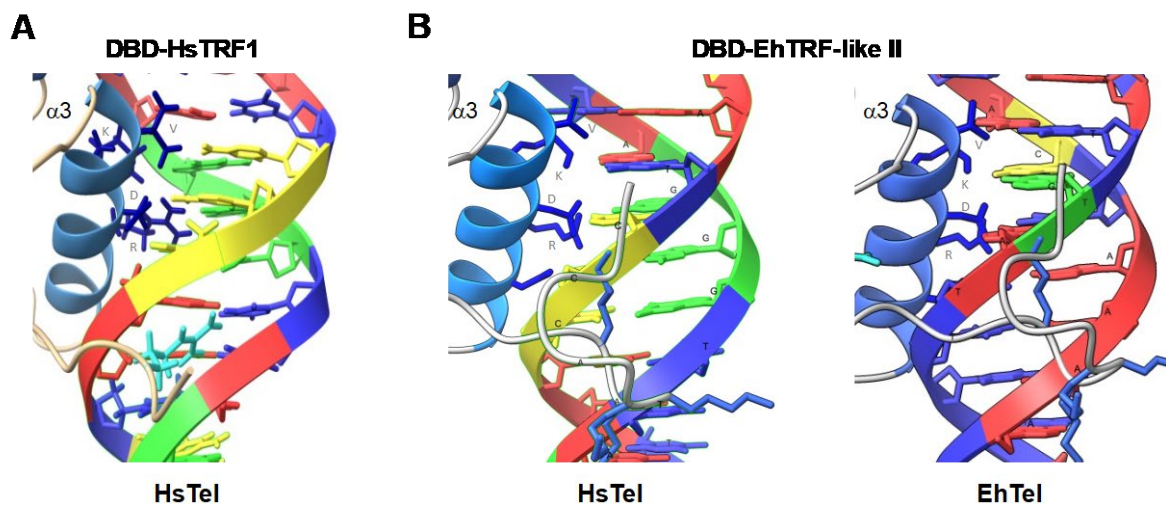


Figura 22. La interacción del DBD-EhTRF-like II con el ADN muestra una orientación espacial, similar a la del DBD-HsTRF1. A) Modelo tridimensional de DBD-HsTRF1 y la secuencia telomérica canónica HsTel (PDB:1IV6), y B) DBD-EHTRF-like II con las secuencias HsTel y EhTel. Se observa la tercera α -hélice y los aminoácidos valina, lisina, asparagina y arginina orientados hacia el ADN. Se utilizó como visualizador el programa Chimera-13.1.

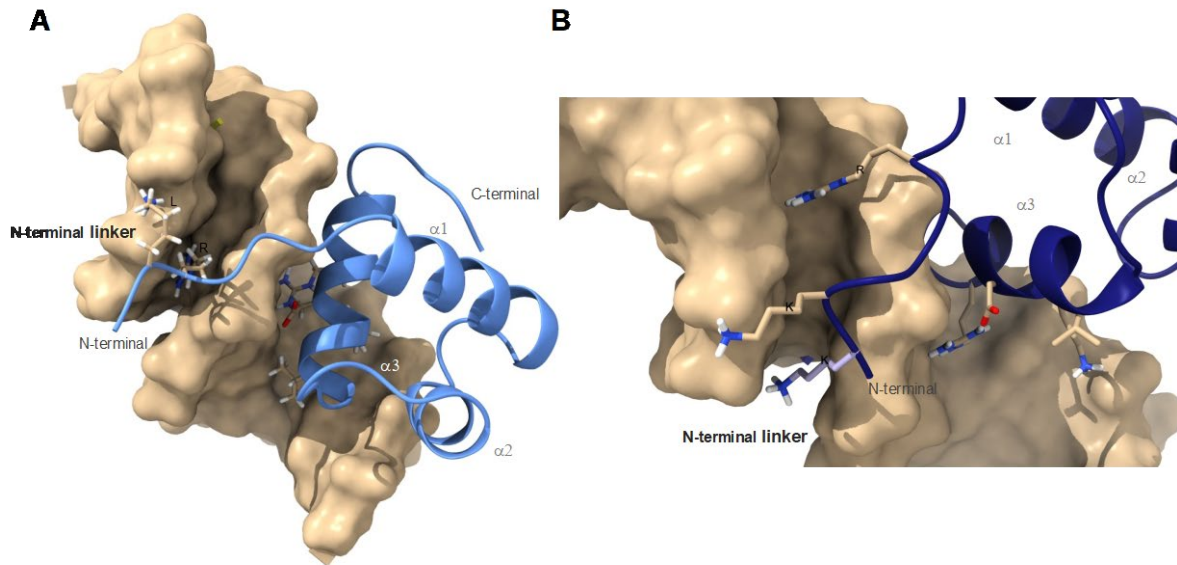


Figura 23. La región del N-terminal *linker* interactúa con el surco menor del ADN. A) Modelo tridimensional de DBD-HsTRF1 y la secuencia telómerica canónica HsTel (PDB:1IV6), y B) DBD-EHTRF-like II y la secuencia EhTel. Se muestra la orientación del motivo N-terminal *linker* en el surco menor de ambas proteínas. Se utilizó como visualizador el programa Chimera-1.1.1.

Una vez obtenidos los modelos tridimensionales del DBD-EhTRF-like II y de las diferentes secuencias, se observó que los motivos N-terminal *linker* y Telebox estaban orientados de manera similar a lo que ocurre en el DBD-HsTRF1. De tal manera que se analizó el efecto de mutaciones, en la formación de contactos y pseudoenlaces, en los motivos N-terminal *linker* y Telebox con las secuencias HsTel, EhTel y nonRel. Para ello se introdujeron mutaciones en el DBD-EhTRF-like II y los efectos de éstas se analizaron mediante DM utilizando el programa HADDOCK 2.4 (<https://wenmr.science.uu.nl/haddock2.4/>). Las mutantes diseñadas fueron tres: 1) N-*linker*-K, en la que se sustituyeron cuatro Lys (K365-369A) de la región del N-*linker*; 2) N-*linker*-R, donde se sustituyó una Arg (R370A) también en la región del N-*linker*; 3) Telebox-3, en la que se sustituyeron los residuos de Val, Lys y Asp por Ala (V402A, K405A y D406A) en el Telebox (Figura 24A). Cabe mencionar que las mutaciones introducidas, no afectaron la estructura tridimensional del DBD (Figura 24B). Una vez obtenidas las estructuras 3D de las proteínas mutantes, se realizaron los *docking moleculares* para todas ellas y las tres secuencias de ADN (HsTel, EhTel y non-Rel). Se analizaron los valores obtenidos del *Haddock score* y se graficaron. La puntuación de *Haddock score* fue diferente para el DBD de la proteína *wild type* y las mutantes con las tres

secuencias de ADN obteniéndose el menor valor para la asociación con la secuencia canónica (-117 ± 1.6) y el mayor para la no relacionada (-94.2 ± 5.9) (Figura 24C). Estos resultados sugirieron que el DBD-EhTRF-like II es capaz de interactuar con la secuencia HsTel, así como con los repetidos STR de *E. histolytica* contenidos en la secuencia EhTel (-104.2 ± 3.9), como se muestra en la Figura 24C.

Por el contrario, para la interacción de la proteína HsTRF1 con las secuencias de ADN se observó que no hay diferencias en las puntuaciones de *Haddock* (Figura 24C). Sin embargo, el RMSD obtenido de la interacción entre el DBD-HsTRF1 con la secuencia HsTel, difirió de los modelos con la secuencia EhTel y la no relacionada, sugiriendo un mejor ajuste del modelo y una mayor cantidad de interacciones entre la proteína humana y la secuencia canónica ($2.7 \text{ \AA} \pm 1.6$). En el caso de la mutante del DBD-EhTRF-like II en el N-linker-K asociada a la HsTel, se observó un aumento en la puntuación del *Haddock* (-102 ± 5.6) (Figura 24C), con respecto a las secuencias EhTel y la non-Rel, donde no hubo diferencias significativas y, además, se observó un aumento del RMSD ($12.1 \text{ \AA} \pm 0.2$). Esto sugirió que los cambios en la secuencia de aminoácidos modifican el ajuste del modelo. La mutante N-linker-R fue la que mayor efecto tuvo en la puntuación del *Haddock* (-86.4 ± 12.1 y -86.4 ± 12.1) para las secuencias HsTel y EhTel, respectivamente, y una diferencia significativa con la secuencia no relacionada (Figura 24C). En el caso del RMSD la mutación de la arginina afectó considerablemente el valor de RMSD entre las tres secuencias (Figura 24C). Finalmente, para la mutante en el Telebox se observó una diferencia significativa en el valor de *Haddock* para la secuencia EhTel (-96.5 ± 14.2) (Figura 24C). Con este análisis se pudo corroborar que para el DBD de la proteína EhTRF-like II, específicamente las regiones N-linker terminal y Telebox son relevantes para la interacción con las secuencias teloméricas.

Finalmente, estos análisis permitieron identificar las bases nitrogenadas que interactuaban con los residuos de aminoácidos del N-terminal linker y el Telebox de los DBD de HsTRF y EhTRF-like II. Por ejemplo, la región del N-terminal linker de ambas proteínas reconoció *in silico* la secuencia TTA, en el surco menor de la secuencia canónica humana. Por otra parte, los aminoácidos del Telebox del DBD de HsTRF1 y EhTRF-like II reconocieron los nucleótidos AG y GG, respectivamente, en el surco mayor de la secuencia canónica humana

(Figura 25). Para la secuencia EhTel los DBD de HsTRF1 y EhTRF-like II reconocieron los nucleótidos TT y TA, respectivamente en el surco menor. También se observaron diferencias en los nucleótidos AG y CTA que reconocen los DBD de HsTRF1 y EhTRF-like II en el Telebox, respectivamente (Figura 25). De manera interesante ambos DBD reconocieron los mismos nucleótidos de la secuencia nonRel (TCGA para el surco menor y ATC para el mayor).

En el caso de las dos mutantes del N- terminal *linker*, se modificó el reconocimiento a la secuencia, reconociendo guaninas en lugar de timinas para la secuencia HsTel (Figura 26). De manera interesante en estas mutantes no se observaron cambios en los nucleótidos que reconocen el N-terminal *linker* con la secuencia EhTel. En el caso de la mutante Telebox-3, se observaron cambios en los nucleótidos que son reconocidos en ambas secuencias en el surco mayor (GTTAG→CCTA) para la secuencia HsTel y para la secuencia EhTel (AG→TAAA). Todas las mutantes reconocieron de igual manera la secuencia no relacionada (TCGA para el surco menor y ATC para el surco mayor) (Figura 26). Estos resultados resaltaron la importancia de los aminoácidos Val, Lys y Asp en el motivo Telebox, por su interacción con la secuencia canónica humana y los STR de *E. histolytica* y que, además, la secuencia no relacionada que se diseñó fue adecuada para evaluar la posible unión inespecífica de la proteína EhTRF-like II. Aunque definitivamente, estos hallazgos predictivos deberán comprobarse experimentalmente.

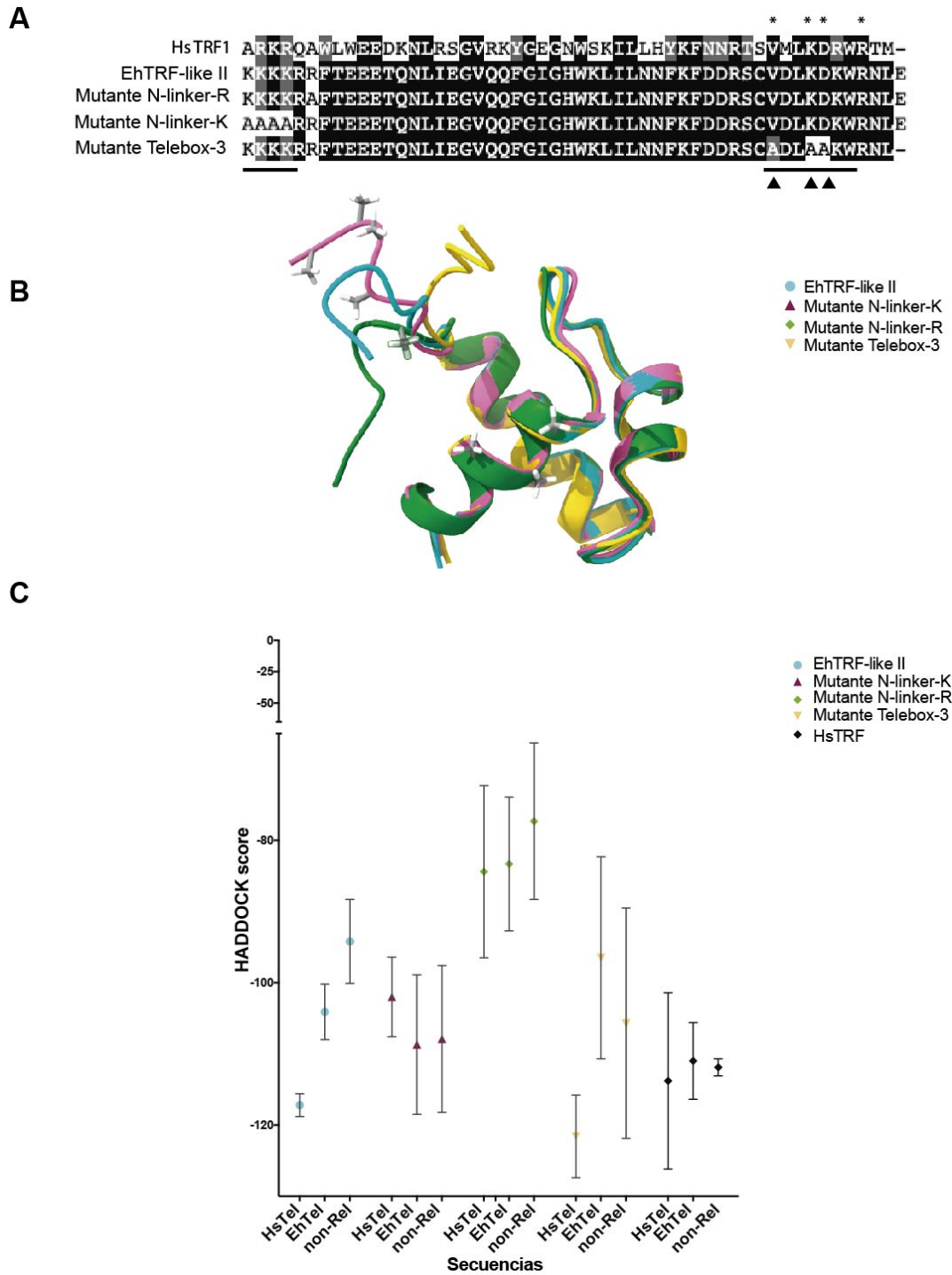


Figura 24. Efecto de mutaciones en el DBD-TRF-like II sobre las interacciones con las secuencias de ADN. A) Alineamiento múltiple del DBD de las proteínas HsTRF1 y EhTRF-like II wt y mutantes (*N-linker-K*, *N-linker-R* y Telebox-3). La barra señala el *N-linker* y los aminoácidos modificados. Los aminoácidos relevantes para el Telebox se indican con un * y con una punta de flecha se muestran las mutaciones introducidas. B) Imagen sobrelapada de las estructuras 3D de los DBD de la proteína EhTRF-like II wt (cian) y las mutantes *N-linker-K* (rosa), *N-linker-R* (verde) y Telebox-3 (amarillo). Gráfico de los valores de *Haddock Score* (C), obtenido de las interacciones entre el DBD de las proteínas mencionadas anteriormente con las secuencias de ADN.

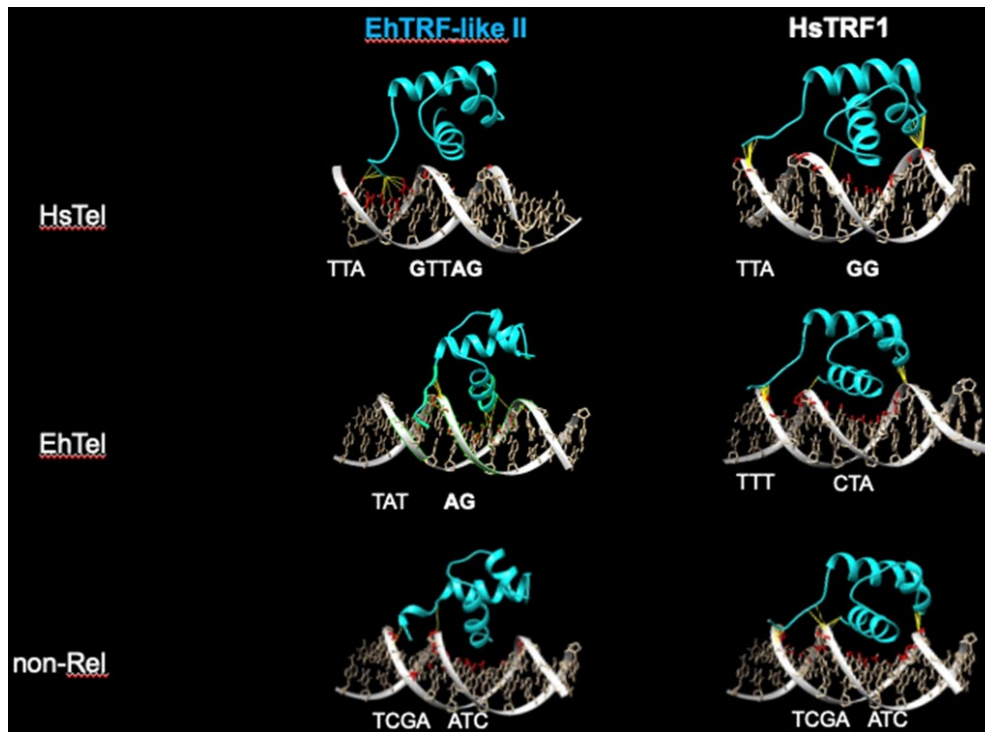


Figura 25. Estructuras tridimensionales de los complejos formados entre el DBD-EhTRF-like II y HsTRF1, con diferentes secuencias de ADN. Se emplearon las secuencias HsTel (TTAGGG), EhTel (TTAGTATT) y una no relacionada (CGATCGAT). En amarillo se muestran los pseudoenlaces y en rojo las bases que interactúan. Se muestran las bases de ADN que interactúan con el N-terminal *linker* y la tercera α -hélice del dominio del DBD Myb. Los complejos ADN-proteína se modelaron empleando el programa Haddock 2.4 y se visualizaron y analizaron con el programa ChimeraX.1.13.1

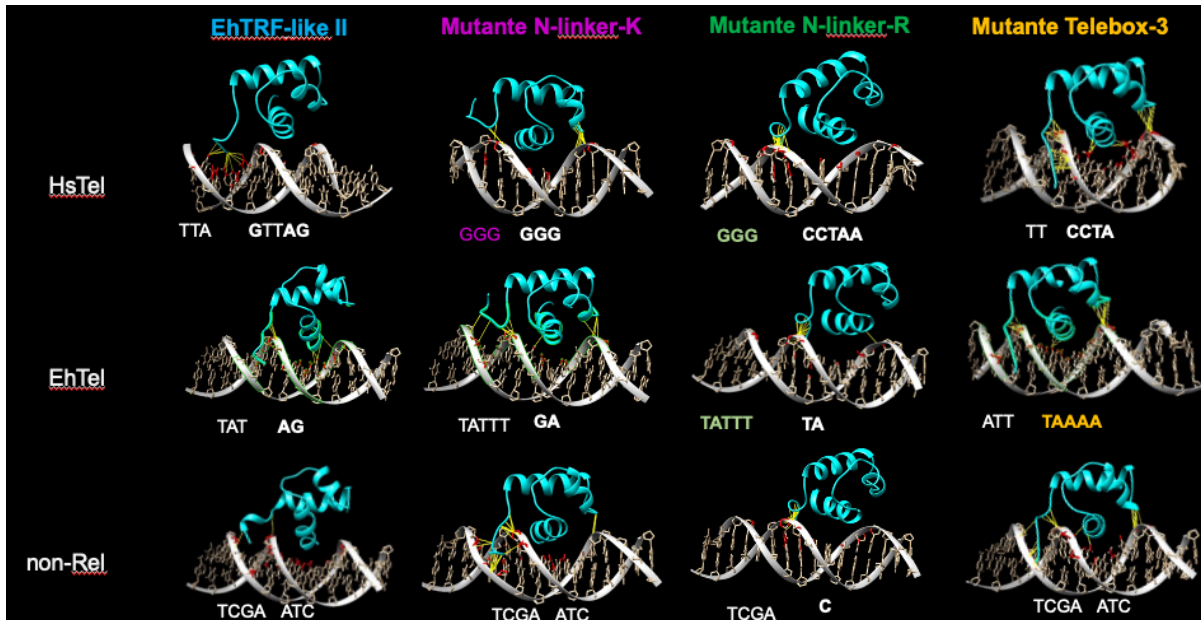


Figura 26. Estructuras tridimensionales de los complejos formados entre las mutantes del DBD-EhTRF-like II y las diferentes secuencias de ADN. Mutantes N-linker K (K361-364A), N-linker R (R365A) y Telebox-3 (V402A, K405A y D406A). Se muestran las bases de ADN que interaccionan con el N-terminal *linker* y la tercer α -hélice del dominio del DBD Myb. Se emplearon las secuencias HsTel (TTAGGG), EhTel (TTAGTATT) y una no relacionada (CGATCGAT). En amarillo se muestran los pseudoenlaces y en rojo las bases que interactúan. Los complejos ADN-proteína se modelaron empleando el programa Haddock 2.4 y se visualizaron en el programa ChimeraX.1.13.1.

10.4 Análisis de la distribución intracelular de la proteína EhTRF-like II en trofozoítos de *E. histolytica*

Con el objetivo de conocer la distribución intracelular de la proteína EhTRF-like II en trofozoítos de *E. histolytica*, se empleó un anticuerpo policlonal de conejo, previamente obtenido en el laboratorio mediante inmunización con el péptido-CIEDKCIVAHHIIDQ acoplado a la proteína KLH, pero que no había sido titulado; por lo que la titulación del anticuerpo se realizó con extractos totales de trofozoítos. Diferentes concentraciones de extractos totales (5, 20 y 30 μ g) se separaron en un gel de SDS-PAGE al 12% y se transfirieron a una membrana de nitrocelulosa, la cual fue bloqueada con leche al 5%. Posteriormente, se prepararon diferentes diluciones (1:1000, 1:500, 1:250 y 1:100) del anticuerpo anti-EhTRF-like II, para realizar su titulación. Los resultados revelaron que la

mejor dilución del anticuerpo del suero inmune fue de 1:100, empleando 5 µg del extracto de proteínas totales (Figura 27). El anticuerpo detectó una banda que migró hasta los 50 kDa, el cual es el peso molecular predicho de la proteína EhTRF-like II. En tanto que el suero pre-inmune a la misma dilución, no detectó proteína alguna, lo cual demostró la especificidad del suero inmune.

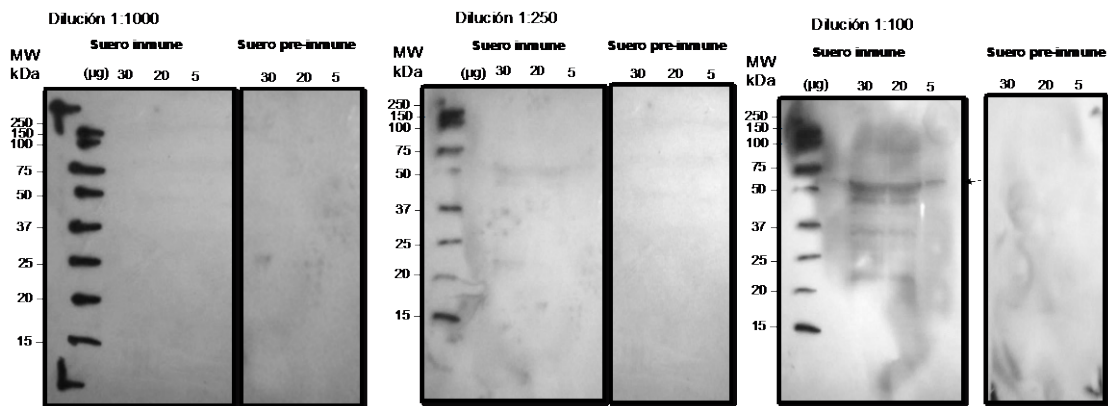


Figura 27. Titulación del anticuerpo α -EhTRF-like II. Diferentes concentraciones de extractos totales de trofozoítos se separaron en un gel de SDS-PAGE al 12% y se transfirieron a membranas de nitrocelulosa, las cuales se incubaron con diferentes diluciones del suero inmune y pre inmune 1:1000, 1:250 y 1:100. La flecha indica la banda reconocida por el anticuerpo en un peso molecular de aproximadamente 50 kDa. Los números a la izquierda indican el peso de los estándares moleculares.

Posteriormente, para analizar la localización subcelular de la proteína EhTRF-like II, se inmunodetectó a la proteína en extractos proteicos obtenidos de la fracción citoplasmática, nuclear soluble y nuclear insoluble. En estas muestras se identificó la presencia de la proteína EhTRF-like II, mediante ensayos de Western blot empleando el anticuerpo α -EhTRF-like II a la dilución determinada durante la titulación del anticuerpo (Figura 28A). Los resultados indicaron que la proteína EhTRF-like II sólo se encontraba en la fracción nuclear insoluble, ya que la banda detectada migraba hasta la misma posición que migró la banda detectada en los extractos totales (Figura 28B). Estos ensayos demostraron que la proteína EnTRF-like II se expresa en condiciones basales de cultivo y que posiblemente se encuentra adosada a la membrana nuclear. Sin embargo, la pureza de las diferentes fracciones debe comprobarse, incluyendo como controles: actina como control de carga; EhPCNA, como control de

proteína nuclear soluble; y Lamina B1, como proteína adosada a la membrana nuclear de la fracción insoluble.

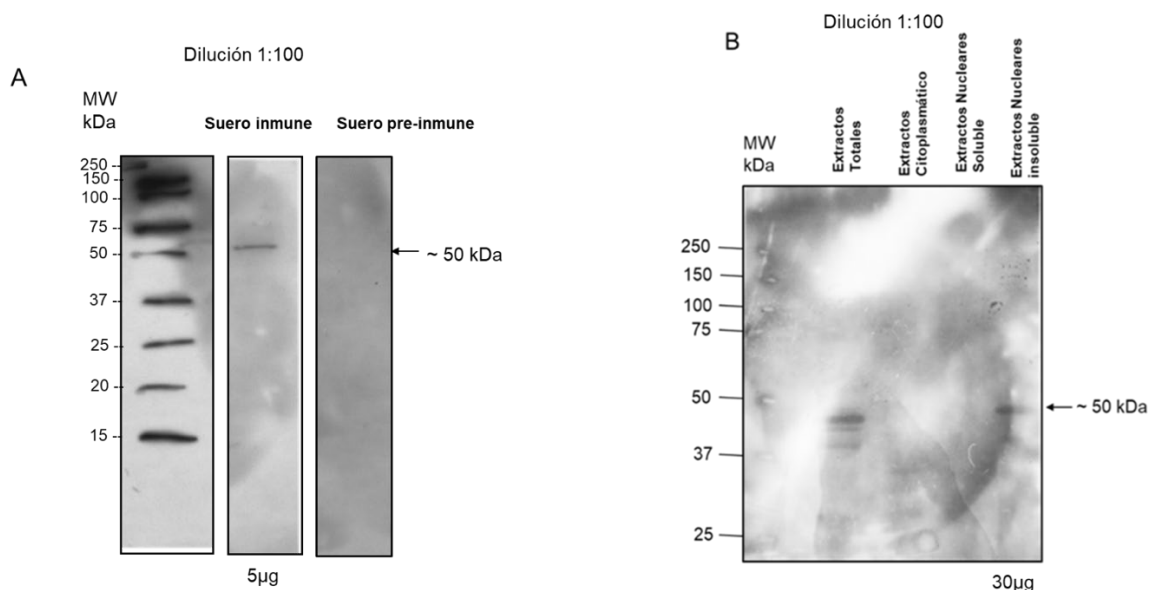


Figura 28. Inmunodetección de la proteína EhTRF-like II en los extractos totales y fracciones celulares de *E. histolytica*. A) Expresión de la proteína EhTRF-like II en trofozoítos en condiciones basales empleando proteínas totales (5µg) separadas en un gel SDS-PAGE al 10% y transferidas a una membrana de nitrocelulosa, la cual fue bloqueada con 5% de leche. La membrana se incubó con suero inmune y pre-inmune a una dilución de 1:100, se reveló empleando sustrato Luminata y se expuso a una placa fotográfica. Carril 1: marcadores de peso molecular; carril 2: suero inmune; y carril 3: suero pre-inmune. La flecha señala una banda de 50 kDa, que corresponde a la proteína EhTRF-like II. B) Fraccionamiento celular de trofozoítos. Los extractos totales (control), citoplasmáticos y nucleares solubles e insolubles, se cuantificaron, separaron en un gel SDS-PAGE al 10% y se transfirieron a una membrana de nitrocelulosa. La inmunodetección se realizó empleando el anticuerpo anti-EhTRF-like II, a una dilución 1:100. La flecha señala una banda a 50 kDa, que corresponde a la proteína EhTRF-like II.

Para corroborar la localización subcelular de la proteína EhTRF-like II, se realizaron ensayos de inmunofluorescencia empleando al anticuerpo anti-EhTRF-like II y trofozoítos fijados en cubreobjetos y permeabilizados con etanol absoluto a -20 °C. Las imágenes analizadas por microscopía confocal muestran que la proteína EhTRF-like II se localiza en el núcleo de los trofozoítos, predominantemente en la periferia nuclear (Figura 29A). Este hallazgo correlaciona con la localización subcelular de las proteínas HsTRF 1 y HsTRF 2 previamente descrita (Palm y de Lange, 2008). En el ensayo realizado en este trabajo se incluyó como control positivo, a la proteína nuclear EhPCNA, la cual se observó también en la membrana nuclear; y como control negativo, únicamente se empleó el anticuerpo secundario, el cual no

brindó ningún patrón de tinción (Figura 29B). La presencia de EhTRF-like II en la periferia nuclear se observó más claramente en las imágenes con un acercamiento del núcleo (Figura 30). Sin embargo, este patrón se deberá corroborar realizando ensayos de colocalización, empleando el anticuerpo anti-Lamin B1 (marcador de periferia nuclear). Además, para saber si EhTRF-like II se asocia con proteínas de protección telomérica, se deben realizar ensayos de cotinción con el anticuerpo anti-H4K20me3. La localización de la proteína EhTRF-like II en la periferia nuclear es similar a la ya reportada para las proteínas EhTRF-like I y III, las cuales se han relacionado con la protección telomérica en *E. histolytica*. De manera que, la ubicación subcelular de EhTRF-like II sugiere que también podría participar en esta función.

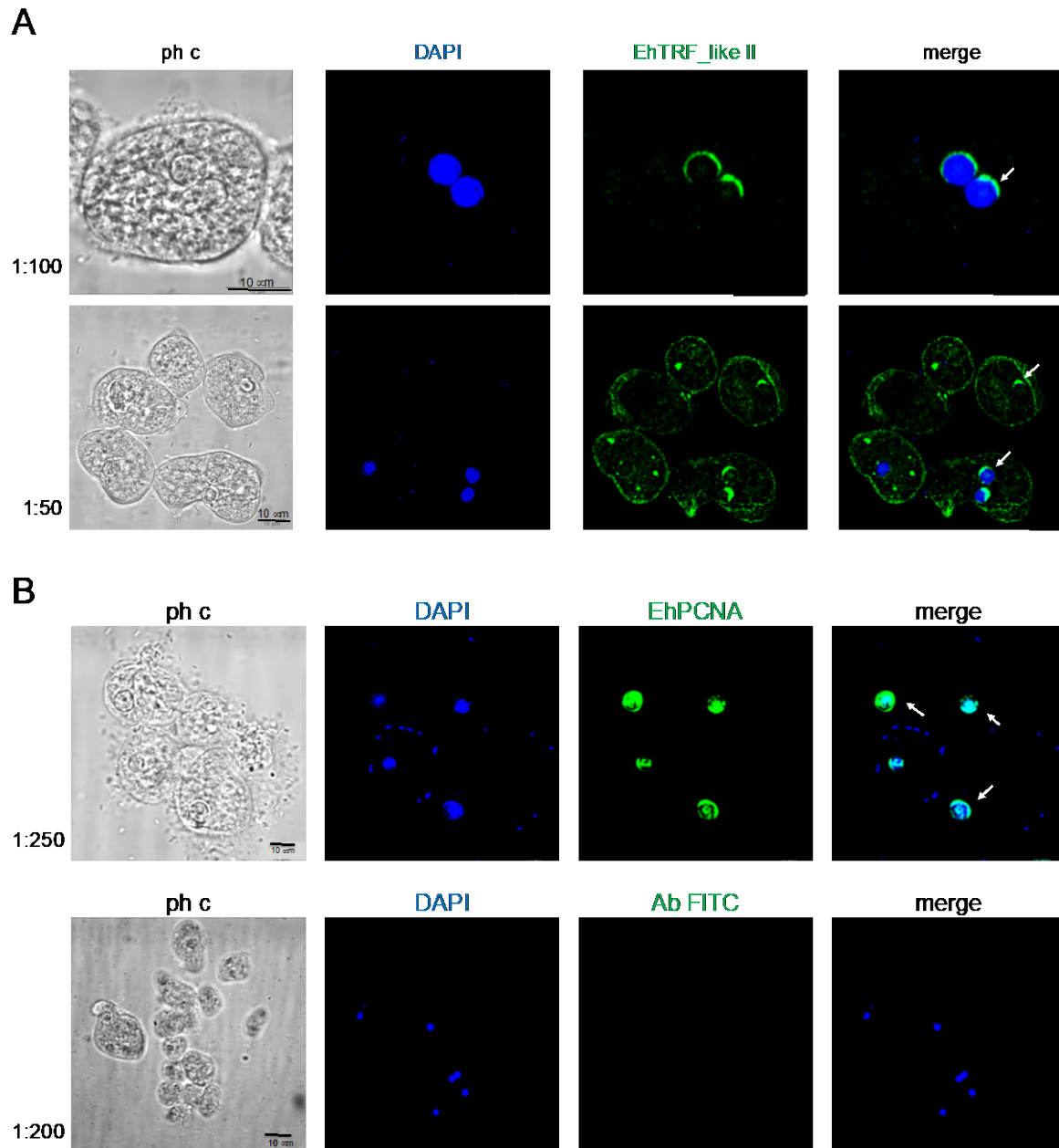


Figura 29. Localización de la proteína EhTRF-like II en trofozoítos de *E. histolytica*. A). Los trofozoítos se adhirieron a cubreobjetos, se fijaron y permeabilizaron con etanol absoluto a $-20\text{ }^{\circ}\text{C}$. Las preparaciones se bloquearon con albúmina al 1%, se incubaron con dos diluciones del anticuerpo α -EhTRF-like II (1:50 y 1:100) y con el anticuerpo anti-conejo acoplado a FITC (verde). Los núcleos se contra tiñeron con DAPI (azul). B) Control positivo, empleando el anticuerpo α -EhPCNA; y control negativo, usando sólo el anticuerpo secundario. Todas las muestras se analizaron por microscopía confocal.

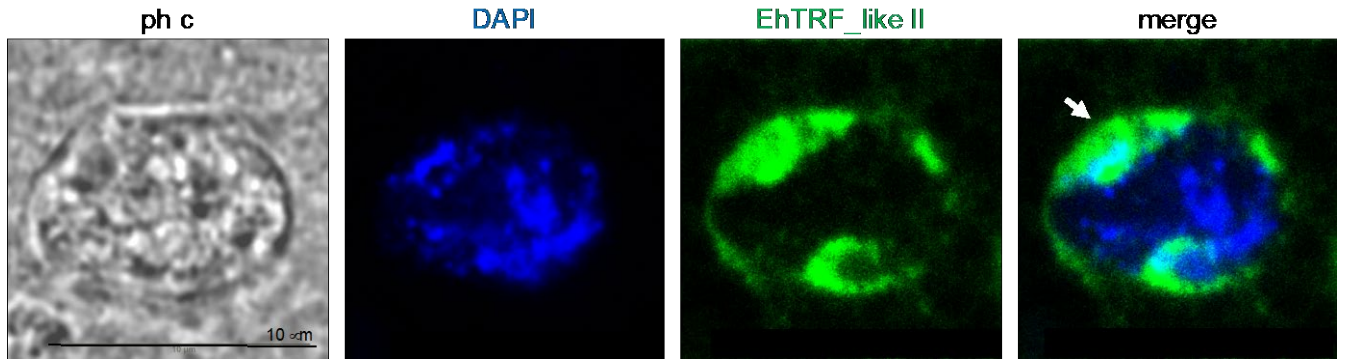


Figura 30. Localización de la proteína EhTRF-like II en los núcleos de los trofozoítos de *E. histolytica*. Aumento de los núcleos, donde la flecha muestra la localización perinuclear de la proteína EhTRF-like II.

11. Discusión

En células eucariontes los extremos terminales de los cromosomas interactúan con un gran número de proteínas, dentro de las más importantes están las proteínas TRFs. En humanos se han identificado dos proteínas: TRF1 y TRF2, las cuales se unen específicamente a las repeticiones de TTAGGG teloméricas bicatenarias y tienen una estructura similar compuesta por los dominios TRFH y de unión a ADN Myb. El acceso de las proteínas TRF al ADN está modulado por diferentes factores, entre ellos, su localización subcelular, las modificaciones postraduccionales que puedan sufrir, sus interacciones con otras proteínas y la estructura de la cromatina (Prochazkova, 2016). De manera interesante, en el parásito *E. histolytica* se identificaron tres proteínas similares a la TRF1 y TRF2 de *Homo sapiens* (HsTRF), las cuales se denominaron EhTRF-like I, II y III. Las proteínas EhTRF tienen una organización de dominios similar entre ellas, que comparten con las de humano (Rendón-Gandarilla *et al.*, 2018). Aunque estas proteínas ya se han identificado en otros protozoarios como *Trypanosoma brucei*, *Trypanosoma cruzi*, y *Leishmania donovani* (Li B *et al.*, 2005; Da Silva *et al.*, 2010), en los organismos protozoarios del género Entamoeba, Acanthamoeba y Naegleria se conoce poco acerca de estas proteínas y su participación en la organización de los extremos terminales de los cromosomas. Es por ello que en este trabajo para identificar las proteínas ortólogas a las proteínas TRF en los géneros Entamoeba, Acanthamoeba y Naegleria se realizó una nueva búsqueda en la base de datos AmoebaDB, y se utilizó como

sonda la secuencia correspondiente al Telebox, pues es el motivo que se une específicamente al ADN. Nuestro análisis identificó a 18 proteínas codificadas en los genomas de los organismos protozoarios parásitos *N. fowleri*, *A. castellanii*, *E. dispar*, *E. moshkovskii*, *E. invadens*, *E. nuttalli*, similares a las HsTRF (Figura 15). Los resultados mostraron que todas estas proteínas contienen en su DBD-Myb los tres dominios relevantes para su función: el N-terminal *linker*, el parche ácido y el motivo Telebox, distribuidos en las tres α -hélices (Figura 18). Estos motivos están implicados en la interacción con las bases nucleotídicas del dsADN telomérico. En los organismos estudiados se observó que el N-terminal *linker* del DBD Myb de las proteínas ortólogas a TRF, está formado por 6 residuos básicos (Lys y Arg), en comparación con HsTRF1 que solo contiene la Lys379 y Arg380 (Figuras 15 y 18). En el caso del motivo Telebox (VxxKDxxR), se observó que está bien conservado en todas las proteínas de estos organismos protozoarios (Figura 18). Sin embargo, en *Acanthamoeba* y *Naegleria* el motivo Telebox mostró mayor divergencia (Figura 18). Aquí, es importante enfatizar que, en ninguno de los organismos protozoarios parásitos de estos géneros, se han identificado secuencias teloméricas canónicas, ni el gen que codifica para el complejo telomerasa (Wang *et. al.* 2003; Loftus *et. al.* 2005; Clarke *et. al.* 2013; Wang *et. al.* 2003). En consecuencia, la identificación de proteínas ortólogas a las proteínas HsTRF, proporciona evidencias de cómo podría llevarse a cabo la organización telomérica en estos organismos. Es por ello, que se realizó un análisis filogenético de las secuencias provenientes de esos organismos para determinar si se organizaban en clados debido a que compartieran algunas características de las proteínas EhTRF-like I, II y III identificadas previamente en *E. histolytica*. Inicialmente se realizó un análisis filogenético empleando solo la secuencia de aminoácidos del DBD-Myb de las 18 proteínas TRF-like identificadas. El árbol filogenético resultante mostró que las proteínas TRF-like de *Acanthamoeba* y *Naegleria* no se agrupaban con las proteínas del género Entamoeba y tampoco fue posible discriminar las proteínas tipo TRF-like I y II, las cuales se agruparon en el mismo clado (Figura 16A). Por esa razón, se decidió emplear las secuencias completas de las proteínas TRF-like de todos los organismos del género Entamoeba en un nuevo análisis, el cual permitió obtener tres clados. El clado que contiene el menor número de proteínas es el clado II y es donde se agrupan EhTRF-like II de *E. histolytica* y sus ortólogos EDI_319270 (*E. dispar*), EMO_106010 (*E. moshkovskii*) y ENU1_201200 (*E. nuttalli*) (Figura 16B).

El análisis de las secuencias del DBD-Myb de todas las proteínas TRF-like del género *Entamoeba* también permitió identificar la presencia de residuos de cisteína conservados (Figura 18, Tabla 6). Se sabe que estos residuos, en el caso del DBD de la proteína c-Myb, son críticos para la regulación *redox* durante su unión con el ADN (Ness, 2003), así como en los factores de transcripción de la familia Myb R2R3 en plantas (Heine *et al.*, 2004). Es por ello que en este trabajo se analizó *in silico*, empleando el programa DIANNA 1.1, si el residuo Cys396 del DBD-TRF-like II podría estar involucrado en la formación de un puente disulfuro, revelando que las cisteínas 109 y 396 podrían formar un enlace de ese tipo (Tabla 6). La formación de enlaces disulfuro en los DBD-Myb puede afectar la unión de la proteína al ADN, cuando esta molécula se somete a un ambiente con agentes oxidantes; ya que al formarse un puente disulfuro se produce un cambio conformacional en la proteína que limita el reconocimiento de la secuencia de ADN. Este mecanismo al ser reversible puede regular mediante reacciones *redox*, la función de la proteína cuando se modifica el microambiente. En el caso de *E. histolytica*, esto ocurre cuando el parásito se encuentra con un ambiente oxidante, producto de la respuesta inmunológica de su hospedero o cuando cambia la tensión de oxígeno, durante el absceso hepático, lo cual podría explicar la redundancia de tres genes codificantes para las proteínas TRF.

En este trabajo también se predijeron otras modificaciones postraduccionales en las proteínas TRF-like, relacionadas con la presencia de motivos conservados (Figura 19, Tabla 7). Este resultado correlaciona con reportes previos donde se demuestra que las proteínas HsTRF pueden ser modificadas postraduccionalmente y que estas modificaciones regulan su interacción con otras proteínas, como la formación de dímeros, sirven como plataforma de superficie para la formación de complejos, regulan su localización celular y participan en su unión al ADN (Walker y Zhu., 2012), por lo que sugiere que las proteínas TRF-like podrían tener funciones semejantes a las HsTRF.

El análisis de la secuencia de las proteínas TRF-like realizado en este trabajo reveló la presencia de 10 motivos conservados (Figura 18, Tabla 7). Tres de ellos están localizados en el DBD Myb (Figura 18), los cuales corresponden a los tres motivos relevantes en el

reconocimiento de la secuencia telomérica (N-terminal *linker*, *patch* ácido y Telebox). Además, se identificaron 6 motivos en las regiones amino terminal, el dominio TRFH y el DBD Myb de la proteína EhTRF-like II, los cuales presentaron una alta probabilidad de fosforilarse por cinasas implicadas en daño y reparación al ADN (Tabla 7). Por ejemplo, en el dominio TRFH la Ser110 podría ser fosforilada por la CK II (Tabla 7), lo cual modularía la formación de dímeros, así como la capacidad para unirse al ADN, como ocurre con su ortólogo HsTRF1 (Walker y Zhu., 2012); o bien los residuos Ser204 y Thr222 del dominio TRFH, podrían ser sustrato de las enzimas PDK y cinasa dependientes de ciclina B2 (CDC2), respectivamente (Tabla 7), lo que modificaría sus interacciones con miembros del complejo de Shelterin y su unión al ADN (Walker y Zhu., 2012). Finalmente, también se predijeron 8 lisinas que podrían SUMOilarse las cuales se encuentran localizadas en la señal de localización nuclear y el extremo N-terminal de la proteína EhTRF-like II (Figura 19). La SUMOilación de TRF1 está relacionada con la relocalización de esta proteína en cuerpos APB, en donde se puede llevar a cabo el mantenimiento de los telómeros por la vía alternativa (Vía ALT) (Walker y Zhu., 2012; Yalçin et al., 2017). La SUMOilación ha sido recientemente caracterizada en *E. histolytica* y las enzimas involucradas también han sido descritas (Díaz-Hernández, 2021), lo cual permite proponer que modificación que podría estar involucrada en el mantenimiento de los extremos de los cromosomas en este organismo (Karpivovich, y Artavanis-Tsakonas, 2020).

En este trabajo nos propusimos analizar si la proteína EhTRF-like II podría reconocer las secuencias derivadas de los STRs contenidas en los arreglos de ARNt, como lo hacen las proteínas parálogas EhTRF-like I y EhTRF-like III. Para ello estudiamos *in silico*, mediante *docking molecular* la interacción del DBD-EhTRF-like II con la secuencia derivada del arreglo [NK1] y se incluyó la secuencia canónica humana y una secuencia no relacionada. Este análisis requirió del modelado de DBD-EhTRF-like II, mediante homología con el DBD-HsTRF1 (PDB:1ba5). El DBD-EhTRF-like II presentó una estructura tridimensional, la cual conserva las tres α -hélices y una región N-terminal *linker* más grande que el DBD-HsTRF1, por la presencia de los 6 residuos básicos ricos en lisina (Figura 20). Los resultados sugirieron que el DBD de la proteína EhTRF-like II es capaz de interactuar con la secuencia canónica humana, pero también con la secuencia derivadas de STR del arreglo [NK1] (Figura

21). También se mostró que el DBD se orienta espacialmente de la misma manera que lo hace HsTRF1, ya que para el caso de HsTRF1 con la secuencia HsTel tiene un RSMD de 1.76 Å y para EhTRF-like II con EhTel presenta un RSMD de 2 Å (Figuras 21-22 y 25), lo que apoya la idea de que estas secuencias podrían funcionar como secuencias teloméricas en amiba. Un aspecto relevante de estos análisis mostró que el motivo N-terminal *linker* se orienta hacia el surco menor de la secuencia telomérica (Figura 23). Los análisis estructurales por RMN mostraron que el motivo N-terminal *linker* de HsTRF1 interacciona con los nucleótidos TAA del surco menor de la secuencia telomérica humana (Nishikawa *et al.* 2001). El N-terminal *linker* de la proteína EhTRF-like II es de un tamaño mayor que el de la HsTRF1, y está formado por 4 lisinas y dos argininas lo cual es similar para las proteínas TRF-like de todos organismos de los géneros *Entamoeba*, *Acanthamoeba* y *Naegleria* (Figuras 18 y 20), sugiriendo que pudiera participar en el reconocimiento de secuencias teloméricas que tienen un mayor grado de variación en su secuencia y un mayor enriquecimiento en pares TA. Esto podría ser válido, al menos para el parásito *E. histolytica*, ya que para este organismo se ha reportado la presencia en los extremos terminales de los cromosomas, arreglos de ARNt flanqueados por secuencias repetidas tipo STR ricas en pares TA, las cuales se han propuesto como secuencias teloméricas (Clark *et al.*, 2007; Tawari *et al.* 2008; Clark *et al.*, 2016).

Empleando las metodologías del *docking molecular* también analizamos el efecto de mutaciones en la región del motivo Telebox del DBD-EhTRF-like II, donde se encuentran los aminoácidos implicados en la interacción directa con el ADN; y en la región el N-terminal *linker* (Figura 24). Las mutaciones introducidas ocasionaron cambios en el reconocimiento de las bases tanto en el N-terminal *linker* como en el motivo Telebox (Figuras 24-26). Los resultados mostraron que las mutantes tuvieron una menor formación de enlaces y pseudoenlaces, así como diferencias en los nucleótidos que reconocen (Figura 26). Estos datos proveen evidencias *in silico* de cómo podría ocurrir el reconocimiento de las secuencias STR en *E. histolytica*, sin embargo, requieren de ser validados experimentalmente para verificar que la interacción y el reconocimiento efectivamente suceda de esa manera.

Por medio de ensayos de *Western Blot* se determinó la localización subcelular de la proteína EhTRF-like II, detectándola en la fracción nuclear insoluble con un peso molecular de 50 kDa, el cual es similar al peso predicho (Figura 28). La presencia de esta proteína en la fracción nuclear insoluble sugirió que podrían estar interaccionando con proteínas adosadas a la membrana nuclear. Esto es diferente a lo descrito para las proteínas EhTRF-like I y II, las cuales se inmunodetectaron en la fracción nuclear soluble (Rendón-Gandarilla *et al.*, 2018). La localización en la membrana nuclear se corroboró mediante ensayos de inmunofluorescencia empleando microscopía confocal. La proteína EhTRF-like II se concentró en la periferia nuclear de los trofozoítos de *E. histolytica*, particularmente en regiones compatibles con la función telomérica (Figuras 29 y 30). Esta localización es diferente a la descrita para la proteína EhTRF-like I, la cual se observó localizada difusamente en la totalidad del núcleo de este parásito, sin embargo, coincide con la localización en la membrana nuclear de EhTRF-like III (Rendón-Gandarilla, 2018; Tesis de Vanesa Medina, 2018). Para corroborar la localización de esta proteína se requeriría colocalizarla con marcadores de membrana nuclear como la Lamin B1 y la marca telomérica H4K20me3.

En resumen, estos resultados apuntan a que en *E. histolytica* la proteína EhTRF-like II por sus características estructurales, la unión a secuencias STR *in silico*, su localización subcelular, y la semejanza con otras TRF de amibas y humanas, forma parte de la maquinaria del Telosoma y participa en las funciones teloméricas.

12. Conclusiones

1. En el DBD de la proteína EhTRF-like II se identificaron tres motivos: *N-linker* terminal, parche ácido y Telebox, los cuales están conservados en las proteínas homólogas a TRF en los géneros Entamoeba, Acanthamoeba y Naegleria.
2. Se predijo que el DBD de la proteína EhTRF-like II puede modificarse postraduccionalmente, a través de fosforilaciones, SUMOilaciones y la formación de puentes disulfuro, lo que podría alterar su localización y unión al ADN.
3. La secuencia EhTel tiene una energía de unión *in silico* similar a la de la secuencia HsTel por el DBD-EhTRF-like II, sugiriendo que las secuencias STR funcionan como secuencias teloméricas.
4. Los motivos conservados en *N-linker* Terminal y el Telebox son relevantes para la interacción *in silico* del DBD-EhTRF-like II con las secuencias teloméricas de humano y STR de amiba.
5. La proteína EhTRF-like II se inmunodetectó en la fracción nuclear insoluble con un peso molecular de ~50 kDa y se inmunolocalizó en la periferia nuclear de los trofozoítos de *E. histolytica*, sugiriendo que ocupan posiciones compatibles con la función telomérica.

13. Perspectivas

- Validar los resultados obtenidos experimentalmente mediante la realización de ensayos EMSA con la proteína EhTRF-like II recombinante y las secuencias de ADN utilizadas en este trabajo.
- Predecir mediante *docking molecular* las interacciones de las proteínas EhTRF-like I y III con las secuencias de ADN utilizadas en este trabajo, para observar diferencias y similitudes con las interacciones predichas para EhTRF-like II.
- Comprobar la pureza de las fracciones celulares mediante la inmunodetección de las proteínas Lamin B1 y EhPCNA.
- Co-localizar a la proteína EhTRF-like I con otro componente del complejo Shelterin, como EhTRF-like III, que comparte la misma localización y con marcadores de membrana nuclear (Lamin B1) y de heterocromatina telomérica (H4K20me3).
- Demostrar la interacción entre las proteínas EhTRF-like I y II mediante ensayos de *pull-down*.

14. Referencias

1. Akbar MA, Chatterjee NS, Sen P, *et al.* (2004) Genes induced by a high-oxygen environment in *Entamoeba histolytica*. *Mol. Biochem. Parasitol*;133(2):187-96.
2. Álvarez Hernández V. (2018) Tesis “Caracterización de la proteína EhTRF-like I en *Entamoeba histolytica*”. 29 de enero 2018. Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México.
3. Andra J., Herbst R., and Leippe M. (2003). Amoebapores, archaic effector peptides of protozoan origin, are discharged into phagosomes and kill bacteria by permeabilizing their membranes. *Dev. Comp. Immunol.* 27:291–304.
4. Audagnotto M, and Dal Peraro, M. (2017) Protein post-translational modifications: *in silico* prediction tools and molecular modeling. *Compu. Struct. Biotechnol. J.* 15, 307-319
5. Baxt L.A. and Singh U. (2008) New insights into *Entamoeba histolytica* pathogenesis. *Curr. Opin. Infect. Dis.* 21:489-494.
6. Betanzos A. *et al.* (2013) The EhCPADH112 complex of *Entamoeba histolytica* interacts with tight junction proteins occludin and claudin-1 to produce epithelial damage. *PLoS one* 8. E65100
7. Bianchi, A., Smith, S., Chong, L., Elias, P., and de Lange, T. (1997) TRF1 is a dimer and bends telomeric DNA. *EMBO J.* 16 1785–1794
8. Biaud, T., Koerig, C.E., Binet-Brasselet, E., Ancelin, K., Pollice, A., Gasser, S.M., and Gilson, E. (1996) The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucleic Acids Res.* 24 1294–1303.
9. Blackburn EH, (1990) Telomeres: Structure and synthesis. *J. Biol Chem.* Abril, 1999; 265 (11); 5919-21
10. Boettner D.R., Huston C.D., Linford A.S., Buss S.N., Houpt E., Sherman N.E., Petri W.A., Jr. (2008) *Entamoeba histolytica* phagocytosis of human erythrocytes involves PATMK, a member of the transmembrane kinase family. *PLoS Pathog.*4:e8.
11. Boletín epidemiológico Sistema Nacional de Vigilancia Epidemiológica (Secretaría de Salud) semana 52 del 2020.
12. Broccoli, D., Smogorzewska, A., Chong, L. *et al.* (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat Genet* 17, 231–235
13. Burla R, La Torre M, Saggio I. (2016) Mammalian telomeres and their partnership with lamins. *Nucleus.* Apr 25;7(2):187-202. PMID: 27116558; PMCID: PMC4916877.
14. Chi-Cheng *et al.* (2018) SUMOgo: Predicción de sitios de SUMOylación en lisinas mediante modelos de cribado de motivos y efectos de diversas modificaciones postraduccionales. *Sci Rep* 8 15512
15. Choi, K. H. *et al.* (2011). Characterization of the DNA binding specificity of Shelterin complexes. *Nucleic Acids Research*, 39(21), 9206–9223.

16. Clark, C. G, Ali I, *et al.* (2006) Unique organisation of tRNA genes in *Entamoeba histolytica*. *Molecular & Biochemical Parasitology*; 146 24-29
17. Clark, C. G, Alsamark, U. Tazreiter M. *et al.* (2007) Structure and content of the *Entamoeba histolytica* genome. *ADV. Parasitol.* 65, 51-190
18. Clarke M, Lohan AJ, Liu B, Lagkouvardos I, Roy S, Zafar N, Bertelli C, Schilde C, Kianianmomeni A, Bürglin TR, Frech C, Turcotte B, Kopec KO, Synnott JM, Choo C, Paponov I, Finkler A, Heng Tan CS, Hutchins AP, Weinmeier T, Rattei T, Chu JS, Gimenez G, Irimia M, Rigden DJ, Fitzpatrick DA, Lorenzo-Morales J, Bateman A, Chiu CH, Tang P, Hegemann P, Fromm H, Raoult D, Greub G, Miranda-Saavedra D, Chen N, Nash P, Ginger ML, Horn M, Schaap P, Caler L, Loftus BJ.(2013) Genome of *Acanthamoeba castellanii* highlights extensive lateral gene transfer and early evolution of tyrosine kinase signaling. *Genome Biol.* 2013 Feb 1;14(2): R11. PMID: 23375108; PMCID: PMC4053784.
19. Court R, Chapman L, Fairall L, Rhodes D. (2005) How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures [published correction appears in *EMBO Rep.* 2005 Feb;6(2):191]. *EMBO Rep.*;6(1):39-45.
20. da Silva MS, Perez AM, da Silveira Rde C, de Moraes CE, Siqueira-Neto JL, Freitas Lde H Jr, Cano MI. (2012) The *Leishmania amazonensis* TRF (TTAGGG repeat-binding factor) homologue binds and co-localizes with telomeres. *BMC Microbiol.* 2010 May 7;10:136. PMID: 20459667; PMCID: PMC2896790.
21. de Lange T. (2005) Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes Dev.* 19(18):2100-10
22. Diamond, L. Harlow, D., Cunnick C. (1978) Un nuevo medio para el cultivo axénico de *Entamoeba histolytica* y otras Entamoeba. *TRSTMH* Vol. 72 No 4 pp 431-432
23. Díaz-Hernández M *et al*; (2021) Protein SUMOylation is crucial for Phagocytosis in *Entamoeba histolytica* trophozoites. *Int J. of Mol. Scien.* 2021, 22(11); 5709
24. Díaz Medina V. (2018) Tesis: Análisis funcional de una proteína TRF-like en *E. histolytica*. Posgrado en Ciencias Genómicas-Universidad Autónoma de la Ciudad de México. Fecha de examen: 27 de febrero 2018 Cotutoria: Dr. Francisco Rendón Gandarilla
25. Diotti R. & Loayza D. (2011) Shelterin complex and associated factors at human telomeres, *Nucleus*, 2:2, 119-135
26. Espinosa-Cantellano M., and Martinez-Palomo A. (1991) The plasma membrane of *Entamoeba histolytica*: structure and dynamics. *Biol.Cell* 72:189-200.
27. Palomo-Martinez A. and Espinosa-Cantellano M., (2000) Pathogenesis of intestinal amebiasis: from molecules to disease. *Clin.Microbiol.Rev.* 13:318-331.
28. Fairall L, Chapman L, Moss H, de Lange T, Rhodes D. (2009) Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. *Mol Cell.* 2001 Aug;8(2):351-61. PMID: 11545737.
29. Faust, D. M. & Guillen, N. (2012) Virulence and virulence factors in *Entamoeba histolytica* the agent of human amoebiasis. *Microbes and infection* 14. 1428-1441

30. Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu. Rev. Plant Biol.* 60, 433–453.
31. Gaspar, *et al.* (2018) Telomere maintenance mechanisms in cáncer. *Genes (Basel)*. 2018 May; 9(5): 241
32. Geiss- Friedlander, R. y Melchor F. (2005) Conceptos de SUMOylación: una década después. *Reseñas de la naturaleza. Biología celular molecular* 8, 947-956
33. Giraud- Panis, M. J. Pisano, *et al.* (2013) One identity or more for telomeres? *Front. Oncol.* 3:48
34. Gónzales M. C. *et al.* (2011) La batalla contra el estrés oxidativo. *Avance y perspectiva Vol. 3 No. 2; Nueva época.*
35. Guerrant RL, Brush J, Ravdin JI, *et al.* (1981) Interaction between *Entamoeba histolytica* and human polymorphonuclear neutrophils. *J Infect Dis.*;143(1):83-93.
36. Guemez Peña M. (2020) “EhTRF-like I: Análisis estructural e interacciones proteína-proteína”. Enero de 2020. Cotutoria: Dra. Abigail Beatzos Fernandez. Tesis de Maestria; Posgrado en Ciencias Genómicas, UACM
37. Hall, A., Karplus, P.A. and Poole, L.B. (2009) Typical 2-Cys peroxiredoxins – structures, mechanisms and functions. *FEBS J.*, 276: 2469-2477
38. Hay R. (2005) SUMO: una historia de modificaciones. *Molecular cell* 18, 1-12
39. He, X. L., and Zhang, J. Z. (2005) Gene complexity and gene duplicability. *Curr. Biol.* 15, 1016–1021.
40. Heine G, Hernández M, & Grotewold E, (2004) Two cysteines in plant R2R3 MYB domains participate in REDOX- dependent DNA- binding. *J of Biol. Chem.* Vol 279 36; 37878-37885 pp
41. Hyeung Kim, *et al*; (2009) TRF2 funtions as a protein hub and regulates telomere maintenance by recognizing specific peptide motifs. *Nat. Struct Mol. Biol.* 2009 Apr; 16(4); 372-9
42. Jordan Alejandro E. (2020) Cambios en la expresión de las proteínas EhTRF-like I y III en respuesta del daño al DNA en *Entamoeba histolytica*. Tesis de Licenciatura. Septiembre, 2020.
43. Kato, K. *et al.* (2013) Sialic acid-dependent attachment of mucins from three mouse strains to *Entamoeba histolytica*. *Biochemical and biophysical research communications* **436**. 252-258
44. Karpiyevich, M., & Artavanis-Tsakonas, K. (2020). Ubiquitin-Like Modifiers: Emerging Regulators of Protozoan Parasites. *Biomolecules*, 10(10), 1403.
45. Kawano-Sugaya Y, *et al.* (2020) Near-chromosome level genome assembly reveals ploidy diversity and plasticity in the intestinal protozoan parasite *Entamoeba histolytica*. *MBC Genomics* **21**, 813
46. Linger BR, Price M (2009). Conservation of telomere protein complexes: shuffling through evolution. *Crit Rev Biochem mol Biol*; 44:434-46.

47. Lidell, M. E., Moncada, D. M., Chadee, K. & Hansson, G. C. (2006) *Entamoeba histolytica* cysteine proteases cleave the MUC2 mucin in its C-terminal domain and dissolve the protective colonic mucus gel. *Proceedings of the National Academy of Science of the United States of America* **103**. 9298-9303
48. Li B, Espinal A, Cross GA. (2005) *Trypanosome* telomeres are protected by a homologue of mammalian TRF2. *Mol Cell Biol*. 2005 Jun;25(12):5011-21. PMID: 15923618; PMCID: PMC1140576.
49. López-Camarillo C, Lopez-Casamichana M, Weber C, Guillén N, Orozco E, Marchat L, (2009) DNA repair mechanisms in eukaryotes: Special focus in *Entamoeba histolytica* and related protozoan parasites. *Infectar Genet Evol*. Dic, 2009; 9 (6); 1051-6
50. Lorenzi HA *et. al.* (2010) New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. *PLoS Negl Trop Dis*. 2010 Jun 15;4(6):e716.
51. Loftus B. *et. al.* (2005) The genome of the protist parasite *Entamoeba histolytica* PubMed 15729342 10.1038/nature03291
52. Lustig AJ (2016) Hypothesis: Paralog Formation from Progenitor Proteins and Paralog Mutagenesis Spur the Rapid Evolution of Telomere Binding Proteins. *Front. Genet*. 7:10.
53. Mandakova, T., and Lysak, M. A. (2008). Chromosomal phylogeny and karyotype evolution in $x = 7$ crucifer species (Brassicaceae). *Plant Cell* 20, 2559–2570.
54. Martinez A. (2000) Recent developments in amoebiasis research. *Curr. Opin. Infect. Dis*. 13:451-456.
55. Meneses E. *et. al.* (2010) The R2R3 Myb protein family in *Entamoeba histolytica*. *Elsevier Gene* 455 (2010) 32–42
56. Mogorzewska A, van Steensel B, Bianchi A, Oelmann S, Schaefer MR, Schnapp G, de Lange T. (2000) Control of human telomere length by TRF1 and TRF2. *Mol Cell Biol*. 2000 Mar;20(5):1659-68. PMID: 10669743; PMCID: PMC85349.
57. Moncada D., Keller K. & Chadee K. (2003). *Entamoeba histolytica* cysteine proteinase disrupt the polymeric structure of colonic mucin and alter its protective function. *Infection and Immunity* 71:838-844
58. Morf L. *et. al.* (2013) Robust gene silencing mediated by antisense small RNAs in the pathogenic protist *Entamoeba histolytica*. *Nucleic Acids Research*, 2013, 9424–9437 Vol. 41, No. 20
59. Moyzis, R. K, Buckingham *et. al.* (1988) A highly conserved repetitive DNA sequence (TTAGGG)_n present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. U.S.A* 1988 85, 6622-6626.
60. Nandakumar, J., & Cech, T. R. (2013). *Finding the end: recruitment of telomerase to telomeres*. *Nature Reviews Molecular Cell Biology*, 14(2), 69–82.
61. Nishikawa T, Okamura H, Nagadoi A, Koig P, Rhodes D, Nishimura Y. (2001) Structure of the DNA-binding domain of human telomeric protein, TRF1 and its

- interaction with telomeric DNA. *Nucleic Acids Res Suppl.* 2001;(1):273-4. PMID: 12836370.
62. Ness SA. (2003) Myb protein specificity: evidence of a context-specific transcription factor code. *Blood Cells Mol Dis.* 2003 Sep-Oct;31(2):192-200. PMID: 12972026
 63. Olivos-García, A; Saavedra, E; Nequiz Avendano, M y Pérez-Tamayo, R. (2011) Amibiasis: mecanismos moleculares de la patogenicidad de *Entamoeba histolytica*. *Rev. Fac. Med. (Méx.)*. 2011, vol.54, n.2 [citado 2020-01-31], pp.10-20.
 64. Palm W., and de Lange, T. (2008) How shelterin protects mammalian telomeres *AnnuRev. Genet.* 42, 301-334
 65. Palomo A. and Espinosa M. (1998) Amoebiasis: New understanding and new goals. *Parasitol Today (Regul. Ed.)* 14:1-3.
 66. Pearson J, Morf L, Singh U. (2013) Regulation of H₂O₂ stress-responsive genes through a novel transcription factor in the protozoan pathogen *Entamoeba histolytica*. *J. Biol Chem.* Feb 2013, 8;288(6):4462-74
 67. Powell SJ, MacLeod I, Wilmot AJ, *et al.* (1966) Metronidazole in amoebic dysentery and amoebic liver abscess. *Lancet.* 1966;2(7477):1329-31.
 68. Procházková Schrumpfová P, (2016) Telomere and Telomerase-Associated Proteins and Their Functions in the Plant Cell. *Front. Plant Sci.* 7:851.
 69. Pumarola A. *et al.* (1991) *Microbiología y Parasitología Médica*. Ediciones Científicas y técnicas. 2da Edición. Barcelona, España
 70. Que Xuchu and Reed S. (2000) Cysteine proteinases and the pathogenesis of Amebiasis. *Clin Microbiol. Rev.* 2000 Apr;13(2): 196-206
 71. Que X., Brinen L.S., Perkins P., Herdman S., Hirata K., Torian B.E., Rubin H., McKerrow J.H., and Reed S.L. (2002). Cysteine proteinases from distinct cellular compartments are recruited to phagocytic vesicles by *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* 119:23–32.
 72. Rendón-Gandarilla FJ, *et al.* (2018) Telomeric Repeat-Binding Factor Homologs in *Entamoeba histolytica*: New Clues for Telomeric Research. *Front. Cell. Infect. Microbiol.* 8:341.
 73. Rhodes *et al.* (2002) Telomere architecture. *EMBO report.* (12) pp1139-1145.
 74. Riveron AM, Lopez-Canovas L, Baez-Camargo M, Flores E, Perez-Perez G, Luna-Arias JP, Orozco E. (2000) Circular and linear DNA molecules in the *Entamoeba histolytica* complex molecular karyotype. *Eur Biophys J.* 2000;29(1):48-56. PMID: 10826778.
 75. Rodriguez- Centeno *et al.* (2019) Structure of *Dictyostelium discoideum* telomeres. Analysis of possible replication mechanisms. *PLoS One.* 14(9): e0222909
 76. Schreiber, E, *et al.* (1989). Rapid detection of octamer binding proteins with “mini extracts”, prepared from a small number of cells. *Nucleic Acids Res.* 17:6419

77. Stanley Kim, H., (2004). Transcriptional divergence of the duplicated oxidative stress-responsive genes in the *Arabidopsis* genome. *The Plant Journal*, 41(2), 212–220.
78. Tawari, B. *et al.* (2008) Patterns of evolution in the unique tRNA gene arrays of the genus *Entamoeba*. *Mol. Biol. Evol.* 25, 187-198
79. Trasviña-Arenas CH, Diaz-Quezada C, Orozco E, Brieba LG. (2018) The Sole DNA Ligase in *Entamoeba histolytica* is a High-Fidelity DNA Ligase Involved in DNA Damage Repair. *Front Cell Infect Microbiol.* 2018 Jul; 12;8:214. PubMed PMID: 30050869; PubMed Central PMCID: PMC6052137.
80. Trasviña-Arenas CH, David SS, Delaye L, Azuara-Liceaga E, Brieba LG. (2019) Evolution of Base Excision Repair in *Entamoeba histolytica* is shaped by gene loss, gene duplication, and lateral gene transfer. *DNA Repair (Amst)*. 2019 Apr;76:76-88. Epub 2019 Feb 22. PubMed PMID: 30822689.
81. Trejos J, Castaño JC. (2009) Factores de virulencia del patógeno intestinal *Entamoeba histolytica*. *Infectio*, 2009;13-2
82. Von Zglinicki, T (2002) Oxidative stress shortens telomeres. *Trends in Biochemical Science.* 27(7); 339-344
83. Van Steensel B, de Lange T. (1997) Control of telomere length by the human telomeric protein TRF1. *Nature.* 1997 Feb 20;385(6618):740-3 PMID: 9034193.
84. Xin, H. Liu, D. & Songyang, Z. (2008) The telosome/Sleterin complex ad its funtions. *Genome Biology*, 9(9), 232
85. Yalçın, Z., Selenz, C., & Jacobs, J. (2017). Ubiquitination and SUMOylation in Telomere Maintenance and Dysfunction. *Frontiers in genetics*, 8, 67.
86. Yang Q, Zhang R, Horikawa I, Fujita K, Afshar Y, Kokko A, Laiho P, Aaltonen LA, Harris CC (2007). «Functional diversity of human protection of telomeres 1 isoforms in telomere protection and cellular senescence». *Cancer Res* 67 (24). PMID 18089797
87. Walsh JA. (1986) Problems in recognition and diagnosis of amebiasis: estimation of the global magnitude of the morbidity and mortality. *Rev. Infect. Dis.* 1986;8(2):228–238
88. Wang Z, Samuelson J, Clark CG, Eichinger D, Paul J, Van Dellen K, Hall N, Anderson I, Loftus B. (2003) Gene discovery in the *Entamoeba invadens* genome. *Mol Biochem Parasitol.* 2003 Jun;129(1):23-31. PMID: 12798503.
89. Walker JR, Zhu XD. (2012) Post-translational modifications of TRF1 and TRF2 and their roles in telomere maintenance. *Mech Ageing Dev.* 2012 Jun;133(6):421-34. Epub 2012 May 23. PMID: 22634377.
90. Welter, B. H., Goldston, A. & Temesvari, L. A. (2011) Localisation to lipid rafts correlates with increased funtion of the Gal/GalNAc lectin in the human protozoan parasite. *Entamoeba histolytica. Internacional journal for parasitology* 41. 1409-1419
91. Willhoeft U, Tannich E. (1999) The electrophoretic karyotype of *Entamoeba histolytica*. *Mol Biochem Parasitol.* 1999 Mar 15;99(1):41-53. PMID: 10215023.

92. Willhoeft U, Tannich E. (2000) Fluorescence microscopy and fluorescence in situ hybridization of *Entamoeba histolytica* nuclei to analyse mitosis and the localization of repetitive DNA. *Mol Biochem Parasitol.* 2000 Feb 5;105(2):291-6.
93. Wilson I. W. *et al.* (2012) Host–Parasite interactions in *Entamoeba histolytica* and *Entamoeba dispar*: what have we learned from their genomes? *Parasite Immunology*, 2012, 34, 90–99
94. Wilson I. W, *et al.* (2019) Genetic diversity and Gene family expansion in members of the genus *Entamoeba*. *Genome Biol. Evol.* Mar 2019; 11(3):688-705
95. World Health Organization 2020. WHO/PAHO/UNESCO report, a consultation with experts on amoebiasis. *Epidemiological Bulletin*
96. Zhaohui Cui *et al.* (2019) Molecular Epidemiology, evolution and phylogeny of *Entamoeba* spp. *Elvesiere. Infection, Genetics and Evolution* Vol. 75 Nov. 2019
97. Zhang X, *et al.* (1997) Identification of possible reactive oxygen species involved in ultraviolet radiation- induced oxidative DNA damage. *Free Radic. Biol Med.* 1997; 23(7):980-5