

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

**Universidad Autónoma de la
Ciudad de México**

Colegio de Ciencia y Tecnología

Posgrado en Ciencias Genómicas

**EhTRF-like I: Predicción de su estructura y
análisis de sus interacciones funcionales.**

T E S I S

QUE PARA OPTAR POR EL GRADO DE

MAESTRO EN CIENCIAS GENÓMICAS

P R E S E N T A

Marco Antonio Guemez Peña

DIRECTORA DE TESIS

Dra. Elisa Irene Azuara Liceaga

CO - DIRECTOR

Dr. Jesús Valdés Flores

**POSGRADO
EN CIENCIAS
GENÓMICAS**

CDMX, enero 2020

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. Lilia López Cánovas – UACM

Secretario: Dra. Elisa Irene Azuara Liceaga – UACM

Vocal: Dr. Jesús Valdés Flores – CINVESTAV – IPN

Plantel de Adscripción:

PLANTEL DEL VALLE, UACM

Directora

Dra. Elisa Irene Azuara Liceaga

Universidad Autónoma de la Ciudad de México

Co-Director

Dr. Jesús Valdés Flores

Centro de Investigación y Estudios Avanzados del IPN

TESIS

LECTORES DE TESIS

Dra. Lilia López Cánovas

Profesora - Investigadora de Tiempo Completo. Posgrado en Ciencias
Genómicas.

Universidad Autónoma de la Ciudad de México

Dra. Minerva Camacho Nuez

Profesora - Investigadora de Tiempo Completo. Posgrado en Ciencias
Genómicas.

Universidad Autónoma de la Ciudad de México

Dr. Mauricio Castañón Arreola

Profesor - Investigador de Tiempo Completo. Posgrado en Ciencias Genómicas.

Universidad Autónoma de la Ciudad de México

Este trabajo fue realizado en el laboratorio 2 del Posgrado en Ciencias Genómicas de la UACM bajo la dirección de la Dra. Elisa Irene Azuara Liceaga y en el laboratorio 11 del Departamento de Bioquímica del Cinvestav-IPN bajo la Codirección del Dr. Jesús Valdés Flores.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada con N° de becario 638701 y a la Universidad Autónoma de la Ciudad de México (UACM) por la beca otorgada Para Estudios de Posgrado 2019-I y para Obtención de Grado 2019-II.

Esta tesis fue realizada gracias al financiamiento otorgado por la Universidad Autónoma de la Ciudad de México (UACM) y por el Consejo Nacional de Ciencia y Tecnología (CONACYT) para el proyecto: Factores de unión a Repetidos Teloméricos (TRFs) de *Entamoeba histolytica*: la importancia de mantener la integridad del genoma (CB-222956) otorgado a la Dra. Elisa Irene Azuara Liceaga.

Quiero agradecer de forma especial a la Dra. Elisa Azuara y al Dr. Jesús Valdés por darme la oportunidad de formar parte de sus equipos de trabajo y por el apoyo brindado durante la realización de este proyecto.

TESIS

Para Laura, Marco, Karina, José, Alondra y Virginia.

Gracias por su apoyo.

Nunca olvidaré que cuando tropezaba, tú ahí estabas.

Gracias Andy.

Índice

I. Resumen	15
II. Abstract	16
III. Introducción	17
3.1 Cromosomas.....	17
3.2 Telómeros.....	20
3.3 Complejo Shelterin.....	24
3.4 Dominio de unión a DNA MYB (DBD).....	31
3.5 Proteína TRF1	33
3.6 Estructura de los dominios de la proteína TRF1	36
3.7 Modificaciones postraduccionales de TRF1	41
3.8 Proteínas TRF en diversos organismos.....	45
3.8.1 Tripanosomátidos	45
3.8.2 <i>Plasmodium falciparum</i>	46
3.8.3 <i>Giardia intestinalis</i>	47
3.8.4 <i>Dictyostelium discoideum</i>	48
IV. Antecedentes Particulares.....	50
4.1 <i>Entamoeba histolytica</i>	50
4.2 Genoma de <i>Entamoeba histolytica</i>	51
4.3 Ciclo de vida de <i>Entamoeba histolytica</i>	53
4.4 Secuencias Teloméricas de <i>Entamoeba histolytica</i>	55
4.5 Proteínas TRFs en <i>Entamoeba histolytica</i>	57
V. Justificación	63
VI. Hipótesis.....	64
VII. Objetivos.....	65
7.1 Objetivo General	65

7.2 Objetivos Particulares	65
VIII. Estrategia Experimental.....	66
IX. Metodología.....	67
9.1 Análisis Bioinformático	67
9.2 Obtención y Purificación EhTRF-like I Recombinante	67
9.3 Ensayos de Inmunodetección (Western Blot)	68
9.4 Hibridación de Oligonucleótidos para el EMSA	69
9.5 Ensayos de Retardamiento de Movilidad Electroforética (EMSA)	69
9.6 Cultivo de trofozoítos de <i>E. histolytica</i>	70
9.7 Obtención de extractos totales de trofozoítos de <i>E. histolytica</i>	71
9.8 Obtención de extractos nucleares de trofozoítos de <i>E. histolytica</i>	71
9.9 Cuantificación de proteínas	72
9.10 Ensayo Pull Down.....	72
9.11 Tinción de Plata	73
9.12 Análisis <i>in silico</i> de Interactoma.....	73
X. Resultados.....	74
XI. Discusión	95
XII. Conclusiones	104
XIII. Bibliografía.....	105

Índice de tablas

Tabla 1. Secuencias teloméricas de diferentes organismos	20
Tabla 2. Proteínas no Shelterin relacionadas con los telómeros humanos	29
Tabla 3. Proteínas que interaccionan con las proteínas TRF1 y TRF2	34
Tabla 4. Modificaciones postraduccionales de TRF1	43
Tabla 5. Características de las unidades STR de tRNA	55
Tabla 6. Comparación de las proteínas EhTRF-like de <i>E. histolytica</i> con TRF1 y TRF2	57
Tabla 7. Sondas utilizadas para los ensayos EMSA	69
Tabla 8. Valores de Tm-score y RSMD para la predicción del dominio TRFH	76
Tabla 9. Incorporación del nucleótido radiactivo en los oligonucleótidos que contienen las secuencias teloméricas	81
Tabla 10. Incorporación del nucleótido radiactivo en la sonda EhTel (2)	83
Tabla 11. Proteínas que interaccionan con la proteína rEhTRF-like I identificadas por MS en Extractos Totales	89
Tabla 12. Proteínas que interaccionan con la proteína rEhTRF-like I identificadas por MS en Extractos Nucleares	93

Índice de figuras

Fig. 1. Esquema de un cromosoma mitótico humano	18
Fig. 2. Componentes de los telómeros.....	21
Fig. 3. Estructura de los telómeros.....	22
Fig. 4. Protección de los extremos terminales del cromosoma	23
Fig. 5. Complejo de Shelterin	24
Fig. 6. Organización de los dominios y estructura de las proteínas TRF1 y TRF2	26
Fig. 7. Complejo de Shelterin en el DNA telomérico	26
Fig. 8. Modelo de las interacciones de proteínas de Shelterin con DNA telomérico	27
Fig. 9. Interacción de proteínas en los telómeros humanos	28
Fig. 10. Estructura cristalina del dominio de unión a DNA de c-Myb unido a DNA32	
Fig. 11. Esquema de los dominios de TRF1	33
Fig. 12. Proteínas asociadas a TRF1.....	35
Fig. 13. Arquitectura del dominio TRFH de las proteínas TRF1 y TRF2 de <i>H. sapiens</i>	36
Fig. 14. Estructura de los dominios de dimerización TRFH de las proteínas TRF1 y TRF2	37
Fig. 15. Esquema de las proteínas TRF1 y TRF2 y su secuencia de aminoácidos del MDBD.....	38
Fig. 16. Estructura tridimensional del dominio MDBD de TRF1 unido a DNA bicatenario.....	39
Fig. 17. Estructuras del MDBD de TRF1 humano	40
Fig. 18. Modificaciones postraduccionales de TRF1 (A) y TRF2 (B).....	41
Fig. 19. LaTRF es homóloga a las proteínas TRFs de vertebrados y <i>T. brucei</i>	46
Fig. 20. Factor de unión a repetidos teloméricos de tipo C2H2 de <i>P. falciparum</i> (PfTRZ)	47
Fig. 21. Morfología del parásito <i>E. histolytica</i>	50

Fig. 22. Determinantes de virulencia amebiana	52
Fig. 23. Ciclo de infección de <i>E. histolytica</i>	54
Fig. 24. Representación esquemática de la organización de las unidades de los arreglos de tRNA.....	56
Fig. 25. Proteínas EhTRF-like de <i>E. histolytica</i>	58
Fig. 26. Árbol filogenético de proteínas EhTRF-like	59
Fig. 27. Localización de las proteínas EhTRF-like en los trofozoítos de <i>E. histolytica</i>	60
Fig. 28. Formación de complejos DNA-proteína con la sonda HsTEL y extractos nucleares de <i>E. histolytica</i>	61
Fig. 29. Formación de complejos DNA-proteína con rEhTRF-like I.....	62
Fig. 30. Predicción in silico del dominio MDBD de <i>E. histolytica</i>	75
Fig. 31. Motivo Telebox VxxKDxxR conservado en la tercera α -hélice de EhTRF-like I.....	75
Fig. 32. Predicción in silico del dominio TRFH de <i>E. histolytica</i>	76
Fig. 33. Identificación de los aminoácidos conservados importantes para el dominio TRFH de <i>E. histolytica</i>	77
Fig. 34. Inducción y solubilidad de la proteína rEhTRF-like I	79
Fig. 35. Purificación e Inmunodetección de la proteína rEhTRF-like-I.....	80
Fig. 36. Estandarización de la concentración de oligonucleótidos y proteína rEhTRF-like I en los ensayos EMSA	81
Fig. 37. Determinación de la constante de disociación (Kd) de la proteína rEhTRF-like I a las secuencias EhTel y HsTel.....	82
Fig. 38. Determinación de la concentración de oligonucleótidos y proteína EhTRF-like I.....	84
Fig. 39. Sitios de unión entre EhTRF-like I y EhTel (2)	85
Fig. 40. rEhTRF-like I es afín a diferentes secuencias teloméricas.....	85
Fig. 41. Patrón de bandas de trofozoítos de <i>E. histolytica</i>	86
Fig. 42. Purificación de proteína rEhTRF-like I para los ensayos de Pull Down ...	87

Fig. 43. Ensayo de Pull Down con extractos totales y nucleares de trofozoítos de <i>E. histolytica</i>	88
Fig. 44. Clasificación de proteínas que interaccionan con la rEhTRF-like I identificadas en extractos totales	92
Fig. 45. Clasificación de proteínas que interaccionan con rEhTRF-like I identificadas por MS en extractos nucleares.....	94
Fig. 46. Modelo de interacciones proteína-proteína en la región telomérica del parásito <i>E. histolytica</i>	103

Abreviaturas

Ala	Alanina
ALT	Vía de alargamiento alternativo de los telómeros
Arg	Arginina
Asp	Ácido aspártico
ATM	Proteína Ataxia Telangiectasia Mutada
ATP	Adenosín trifosfato
ATR	Proteína relacionada Ataxia Telangiectasia Mutada y Rad3
BLM	Proteína del Síndrome de Bloom
BSA	Albumina Bovina Sérica
CDK	Cinasa dependiente de ciclina
CDK1	Cinasa 1 dependiente de ciclina
CDK2	Cinasa 2 dependiente de ciclina
DBD	Dominio de unión a DNA
DNA	Ácido desoxirribonucleico
ds-DNA	Ácido desoxirribonucleico de doble cadena
DTT	Ditiotreitol
EDTA	Ácido etilendiaminotetraacético
EGTA	Ácido tetraacético
EhTel	Secuencia tipo STR de <i>E. histolytica</i> 1 (TTAGTATT)
EhTel (2)	Secuencia tipo STR de <i>E. histolytica</i> 2 (TTATATGT)
EhTRF-like I	Factor de unión a repetidos teloméricos I de <i>E. histolytica</i>
EhTRF-like II	Factor de unión a repetidos teloméricos II de <i>E. histolytica</i>
EhTRF-like III	Factor de unión a repetidos teloméricos III de <i>E. histolytica</i>
EMSA	Ensayo de cambio de movilidad electroforética
GalGalNac	Lectina específica de N-acetil D-galactosamina
Glu	Ácido glutámico
HsTel	Secuencia telomérica canónica humana (TTAGGG)
IgA	Inmunoglobulina A
IgG	Inmunoglobulina G
Ile	Isoleucina
K _a	Constante de afinidad
Kb	Kilo bases
K _d	Constante de disociación
kDa	Kilo daltones
LaTBP	Proteína de unión a telómeros de <i>L. amazonensis</i>
LaTRF	Factor de unión a repetidos teloméricos de <i>L. amazonensis</i>
LGT	Transferencia lateral de genes
MA	Análisis mutacional
Mb	Mega bases
mCi	Milicurie
MDBD	Dominio de unión a DNA de tipo Myb
Met	Metionina
mL	Mililitro
MS	Espectrometría de masas
MutTel	Secuencia telomérica humana mutada
MW	Peso molecular

TESIS

µg	Microgramos
µM	Micromolar
µL	Microlitro
NonTel	Secuencia no telomérica
OB	Pliegue de unión oligonucleótido/oligosacárido
P53	Proteína supresora de tumores del gen p53
PARP	Poli ADP ribosa polimerasa
pH	Potencial de hidrógeno
Phe	Fenilalanina
pI	Punto isoeléctrico
PMSF	Fluoruro de fenilmetilsulfonilo
POT1	Proteína de protección a telómeros 1, del inglés Protection of Telomeres protein 1
Poz1	Proteína de protección a telómero 1 en <i>S. pombe</i>
RAP1	Proteína ortóloga de humano de la proteína activadora / represora 1 de levadura del inglés Repressor-activator protein
rDNA	DNA ribosomal
REDOX	Reacción de óxido-reducción
RLIM	Proteína ubiquitina ligasa E3
RMN	Resonancia magnética nuclear
RNA	Ácido ribonucleico
RMSD	Desviación media cuadrática
SDS-PAGE	Geles desnaturalizantes de poliacrilamida
Ser	Serina
ss-DNA	Ácido desoxirribonucleico de cadena sencilla
STR	Secuencia repetida en tándem corta (dos a seis pb)
TbTRF	Factor de unión a repetidos teloméricos de <i>T. brucei</i>
TEM	Microscopia electrónica de transmisión
TERRA	RNA largo no codificante transcrito de los telómeros
TIF	Centros inducidos por la disfunción de los telómeros
TIN2	Factor nuclear 2 que interacciona con el factor 1 de unión a repetidos teloméricos del inglés TRF interacting protein
TPP1	Proteína de interacción con TIN2 acrónimo que proviene de la primera letra de las proteínas TINT1, PTOP, y PIP1.
TRF1	Factor 1 de unión a repetidos teloméricos, del inglés Telomeric repeat-binding factor 1
TRF2	Factor 2 de unión a repetidos teloméricos, del inglés Telomeric repeat-binding factor 2
TRFH	Dominio de homodimerización del factor de unión a repetidos teloméricos
tRNA	Ácido ribonucleico de transferencia
Trp	Triptófano
T4 PNK	T4 polinucleótido cinasa
Val	Valina

I. Resumen

Los telómeros son complejos de protección altamente regulados y dinámicos en los extremos de los cromosomas, que consisten en repeticiones cortas en tándem de DNA y proteínas asociadas. Los factores de unión a repetidos teloméricos 1 y 2 (TRF1 y TRF2), son proteínas esenciales para el mantenimiento telomérico que tienen una arquitectura definida por dos dominios: un dominio TRFH conformado por una estructura de nueve alfa-hélices y un dominio de unión a DNA (MDBD) identificado por su similitud de secuencia con el motivo de unión a DNA de tipo Myb. *Entamoeba histolytica* contiene tres genes que codifican para proteínas similares a TRF1 y TRF2, las cuales fueron denominadas EhTRF-like I, II y III. En este trabajo, mediante análisis bioinformáticos analizamos la estructura tridimensional de los dominios TRFH y MDBD de la proteína EhTRF-like I. También, estudiamos la afinidad de unión de EhTRF-like I a secuencias tipo STR de *Entamoeba histolytica* mediante ensayos EMSA y realizamos una búsqueda de proteínas asociadas a esta proteína que pudieran estar participando en la maquinaria telomérica de este parásito a través de ensayos de tipo *pull down*. Nuestros resultados muestran la presencia de residuos de aminoácidos hidrofóbicos en el dominio TRFH, los cuales podrían ser parte de la superficie de interacción del dominio, posiblemente mediando las interacciones proteína-proteína que se requieren para formar tanto homodímeros como complejos de la maquinaria telomérica. Por otro lado, se confirmó que el dominio Myb de unión a DNA (MDBD) tiene una estructura hélice vuelta hélice altamente conservada. También, se determinó la afinidad de unión de la proteína rEhTRF-like I con la secuencia canónica humana la cual mostró una K_d de 6.8 nM mientras que a la secuencia tipo STR de la unidad NK2 de *E. histolytica* mostró una K_d de 2.9 nM. Finalmente, se identificó como socio de interacción de la proteína rEhTRF-like I a la proteína EhTRF-like II entre otras, lo que podría sugerir que son proteínas necesarias para proteger los extremos de los cromosomas del parásito.

II. Abstract

Telomeres are highly regulated and dynamic complexes at the ends of the chromosomes, which consist of short tandem repeats of DNA and associated proteins. The telomeric repeat binding factors 1 and 2 (TRF1 and TRF2), are essential proteins for telomeric maintenance that have an architecture composed by two domains: the TRFH domain consisting of a structure of nine alpha-helices and a DNA binding domain (MDBD) with a sequence similarity with the MYB DNA binding domain. *Entamoeba histolytica* contains three genes that encode for proteins similar to TRF1 and TRF2, which were dubbed as EhTRF-like I, II and III. In this work, through bioinformatic analysis we obtained the three-dimensional structure of the TRFH and MDBD domains of EhTRF-like I. Also, we determined the binding affinity constant of the rEhTRF-like I to the STRs sequences of *E. histolytica* and to the canonical TRF sequence by EMSA assays. We also perform Pull Down assays to determine proteins associated to rEhTRF-like I that could be participating in the telomeric machinery of this parasite. Our results showed the presence of hydrophobic amino acid residues which could be part of the interaction surface of the TRFH domain, possibly mediating the protein-protein interactions that are required to form both homodimers and complexes of telomeric machinery. On the other hand, it was confirmed that the MDBD domain is highly conserved. Also, the binding affinity of the rEhTRF-like I protein with the human canonical sequence was determined showing a Kd of 6.8 nM while the STR type sequence of the NK2 unit of *E. histolytica* showed a Kd of 2.9. nM. Finally, it was identified EhTRF-like II as an interaction partner of rEhTRF-like I protein, which could suggest that they are necessary to protect the ends of the chromosomes in this parasite.

III. Introducción

3.1 Cromosomas

El genoma es el conjunto de secuencias organizadas de nucleótidos, que especifica las instrucciones para el desarrollo y funcionamiento de un organismo. Todos los seres vivos, desde las bacterias hasta los humanos, almacenan su información genética en estructuras denominadas cromosomas, formadas por ácido desoxirribonucleico (DNA) de doble cadena y proteínas. El DNA es una molécula que forma una doble hélice, en la que cada cadena de la hélice es una molécula lineal formada por subunidades llamadas nucleótidos. Cada nucleótido del DNA tiene una de las cuatro bases nitrogenadas – A (adenina), G (guanina), T (timina) o C (citosa). Estas cuatro bases constituyen el código genético, que en distintas combinaciones específicas determinan la secuencia de aminoácidos de las proteínas. Las proteínas son unos de los productos finales de la expresión génica (Klug, 2006).

Los cromosomas de eucariontes (cuyas células tienen un núcleo delimitado por una membrana) son lineales y se encuentran divididos por una estructura central llamada centrómero, la cual es fundamental en el proceso de división celular. El centrómero divide al cromosoma en dos regiones: la región p (brazos cortos) y la región q (brazos largos) (Fig. 1) (Flores y Mayani, 2013). Cada especie tiene un número característico de cromosomas denominado el número diploide ($2n$). En células diploides los cromosomas se encuentran formando parejas, llamados cromosomas homólogos los cuales son idénticos en tamaño. En la mitosis, los cromosomas se copian y distribuyen de tal manera que cada una de las dos células hijas resultantes recibe una dotación diploide de cromosomas. La meiosis es una forma de división celular asociada con la formación de los gametos en los animales y de esporas en muchos vegetales. Las células producidas por meiosis reciben una sola copia de cada cromosoma, llamado el número haploide (n). La información genética contenida en los cromosomas dirige la función celular, determina en gran medida la apariencia externa de los organismos y sirve de unión entre generaciones en las especies (Klug, 2006).

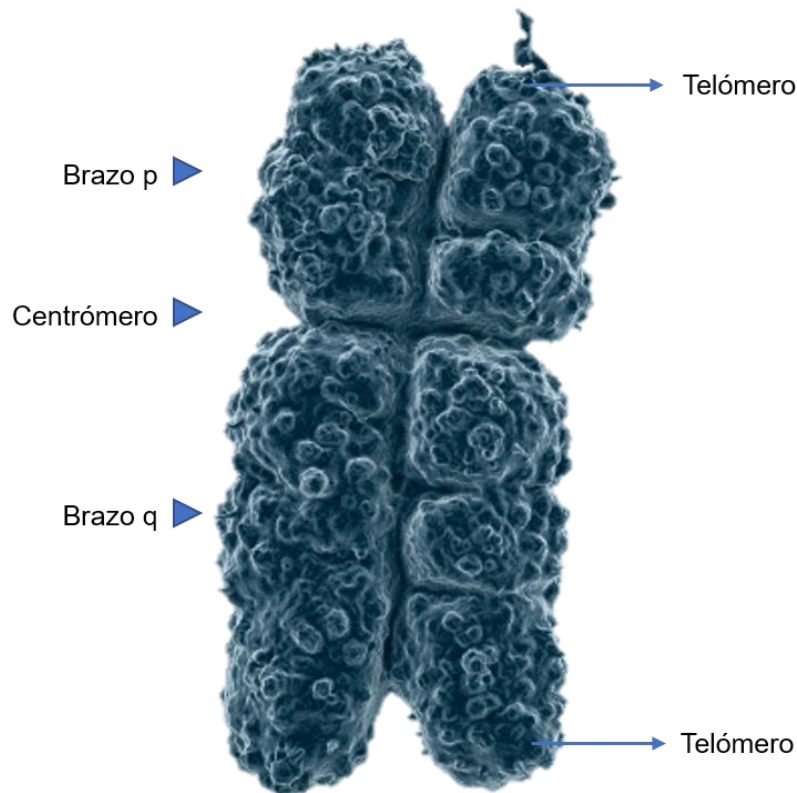


Fig. 1. Esquema de un cromosoma mitótico humano. El cromosoma lineal se subdivide en dos regiones, separadas por una estructura central llamada centrómero. La región más larga del cromosoma corresponde al brazo q y la más corta al brazo p. Puede observarse que los telómeros se localizan en los extremos del cromosoma tanto en su brazo q como p (Tomado de Flores y Mayani, 2013).

Los cromosomas, solo visibles durante la mitosis están formados por fibras de cromatina estrechamente espiralizadas. Tras la división celular, cuando las células entran en el estadio de interfase del ciclo celular, la cromatina se desespiraliza y los cromosomas desaparecen. Durante la interfase, la cromatina se encuentra dispersa en el núcleo, y se replica el DNA de cada cromosoma. Al continuar el ciclo celular, gran parte de las células vuelven a entrar en mitosis, en la que la cromatina se condensa de nuevo y los cromosomas se hacen visibles al microscopio óptico. Esta condensación representa una contracción de unas 10.000 veces para cada fibra de cromatina. En la cromatina eucariótica, hay una cantidad substancial de proteínas asociadas al DNA cromosómico en todas las fases del ciclo celular. Las proteínas asociadas se clasifican en histonas cargadas positivamente (básicas) y en no histonas. De las proteínas asociadas al DNA, las histonas desempeñan claramente la función estructural más esencial. Las

histonas contienen grandes cantidades de los aminoácidos lisina y arginina, que están cargados positivamente, lo que les posibilita unirse electrostáticamente a los grupos fosfatos del esqueleto del DNA, cargados negativamente. El modelo general de la estructura de la cromatina se basa en la suposición de que las fibras de cromatina, compuestas por DNA y proteínas, experimentan diferentes grados de enrollamiento y plegamiento al condensarse dentro del núcleo celular, los cuales se denominan heterocromatina y eucromatina que corresponden a regiones de los cromosomas que permanecen altamente y menos condensadas, respectivamente.

Cada cromosoma eucariótico consiste, en toda su extensión, en una molécula continua de DNA de doble hélice con nucleosomas repetidos, lo que podría llevar a creer que todo el cromosoma es estructuralmente uniforme. Sin embargo, aunque la mayor parte del cromosoma se encuentra presente como cromatina desespiralizada, algunas partes permanecen condensadas y se tiñen fuertemente durante la interfase. Las áreas heterocromáticas son genéticamente inactivas porque no contienen genes o porque los genes que contienen son inactivos. Además, la heterocromatina se replica más tarde que la eucromatina durante la fase S del ciclo celular. El descubrimiento de la heterocromatina proporcionó las primeras pistas de que partes de los cromosomas eucarióticos no siempre codifican proteínas. La heterocromatina se encuentra en diferentes regiones de los cromosomas eucarióticos. Los extremos de los cromosomas, denominados telómeros, también son heterocromáticos (Klug, 2006).

3.2 Telómeros

En la tercera década del siglo XX, Barbara McClintock y Hermann J. Müller encontraron que los cromosomas portaban en sus extremos un componente especial que les confería estabilidad. Muller nombro a estas estructuras telómeros, del griego *telos* (final) y *meros* (parte). Los estudios de McClintock revelaron que, sin estas estructuras terminales, los cromosomas se adherían unos a otros y sufrían cambios estructurales, lo que afectaba la supervivencia y fidelidad en la replicación de los cromosomas (Flores y Mayani, 2013). Posteriormente se encontró que los telómeros del protozooario *Tetrahymena* contenían una secuencia repetida de nucleótidos muy corta, compuesta de dos tipos de nucleótidos: timina (T) y guanina (G) que se repetían muchas veces al final de los cromosomas y que no codificaban para ninguna proteína. A partir de entonces, se han estudiado los telómeros de diversos organismos, y se ha encontrado que su secuencia repetida de nucleótidos es muy similar entre diversas especies (Flores y Mayani, 2013) (Tabla 1).

Tabla 1. Secuencias teloméricas de diferentes organismos

ORGANISMOS	REPETIDOS TELOMÉRICOS
<u>Vertebrados</u>	
<i>Homo sapiens</i>	TTAGGG
<u>Artrópodos</u>	
<i>Bombix mori</i>	TTAGG
<u>Protistas</u>	
<i>Plasmodium vivax</i>	TTTTGGGG
<i>Giardia lamblia</i>	TT(T/C)AGGG
<i>Trypanosoma cruzi</i>	TAGGG
<i>Tetrahymena thermophila</i>	TTAGGG
<i>Oxytricha trifallax</i>	TTGGGG
<i>Physarum polycephalum</i>	TTTTGGGG
<i>Kluyveromyces marxianus</i>	TTAGGG TCGGATTTGATTAGGTATGTGGTGT
<i>Neurospora crassa</i>	TTAGGG
<i>Caenorhabditis elegans</i>	TTAGGG
<u>Algas Verdes</u>	
<i>Chlamydomonas reinhardtii</i>	TTTTAGGG
<u>Plantas</u>	
<i>Arabidopsis thaliana</i>	TTTAGGG
<u>Hongos</u>	
<i>Saccharomyces cerevisiae</i>	TG ₂₋₃ TG ₁₋₆
<i>S. pombe</i>	G ₀₋₆ GGTTACAC ₀₋₁
<i>Kluyveromyces lactis</i>	ACGGATTTGATTAGGTATGTGGTG (TGATTAGGTATGTGGGTACGGATT)

(Tomado de Li, 2005).

Los telómeros son estructuras esenciales que cubren y protegen los extremos de los cromosomas lineales. Los telómeros ocultan los extremos cromosómicos de los mecanismos de detección del daño del DNA y de la maquinaria de reparación del DNA. Cuando los telómeros se vuelven disfuncionales, los extremos se detectan como rupturas de DNA bicatenario (rupturas de ds-DNA). Este reconocimiento da como resultado la activación de puntos de control del daño en el DNA que desencadenan la senescencia o la apoptosis, así como eventos de reparación de los extremos terminales de los cromosomas mediante recombinación no homóloga (NHEJ), un proceso que crea fusiones intercromosómicas. La función protectora de los telómeros está mediada por las actividades de complejos proteicos asociados a los telómeros y por una estructura tridimensional denominada los *T-loops* (de Lange, 2002) (Fig. 2). La longitud de los telómeros es muy heterogénea de cromosoma a cromosoma y de célula a célula dentro de una población.

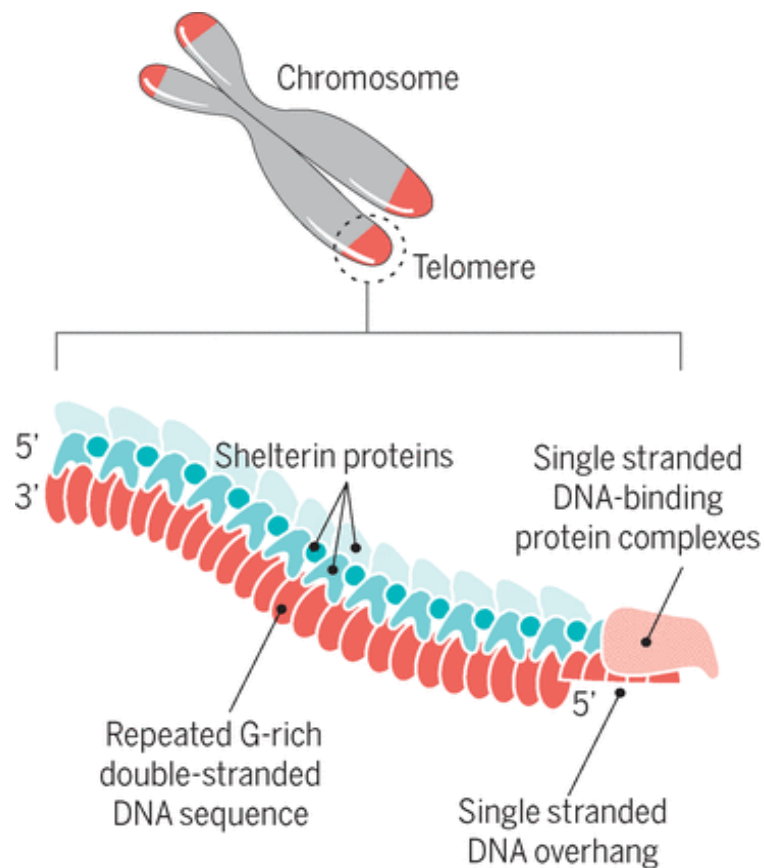


Fig. 2. Componentes de los telómeros. El complejo del telómero humano consiste en una región cromosómica terminal de la secuencia TTAGGG repetida en tándem unida por proteínas del complejo Shelterin (Tomada de Blackburn, 2015).

Los telómeros de organismos como *Homo sapiens* están compuestos de copias en tándem de una repetición simple tipo STR de DNA con la secuencia (TTAGGG) n , que es transcrita por la enzima telomerasa. La telomerasa, es una enzima transcriptasa inversa responsable del mantenimiento telomérico, se reprime transcripcionalmente en la mayoría de las células somáticas humanas diferenciadas, la longitud media de los telómeros está sujeta a un control de acción *cis*, que, en gran parte se ejerce sobre la actividad de la telomerasa en cada telómero. La mayoría de los telómeros están compuestos de DNA telomérico dúplex, con la excepción de las últimas 50-500 bases de la cadena rica en G, que forma una proyección telomérica en el extremo 3' monocatenario (Fig. 3). Esta proyección telomérica de longitud variable es el resultado de un procesamiento por parte de una exonucleasa 5' - 3' aún desconocida activada en fase S tardía y G₂ (Verdun y Karlseder, 2006). La evidencia sugiere que la proyección telomérica 3' está secuestrada en una gran estructura, denominada *t-loop*. La formación de esta estructura implica la inserción de la proyección 3' - telomérica en el DNA telomérico dúplex, su hibridación con la cadena rica en C y la deslocalización de la cadena rica en G en un bucle de desplazamiento (bucle D). Se ha propuesto que los bucles en T están especialmente bien adaptados para proteger los extremos de los cromosomas de la reparación del DNA y los mecanismos de detección del daño del DNA (Choi, 2011).

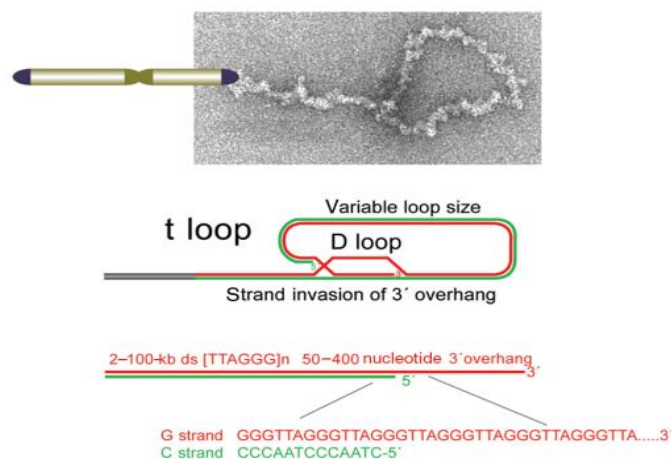


Fig. 3. Estructura de los telómeros. Microscopía electrónica donde se muestra la formación de un t-loop en cromatina aislada. Así mismo, se observa la representación esquemática de un t-loop producto de la hibridación la hebra rica en G en la hebra rica en C. Finalmente se ejemplifica la secuencia de la hebra rica en G que conforma la proyección telomérica 3' de *H. sapiens* (Tomada de de Lange, 2010).

En la mayoría de las células somáticas humanas, los telómeros se acortan en cada ronda de replicación del DNA, en parte, debido a la incapacidad de las DNA polimerasas para completar la secuencia de DNA que queda posterior a la eliminación del último cebador de RNA (Levy et al., 1992). Cuando la longitud del DNA telomérico se vuelve críticamente breve, se activa la respuesta al daño del DNA, lo que desencadena la inducción de la senescencia replicativa (d'Adda di Fagagna et al., 2003). El proceso de senescencia en telómeros debido al desgaste que sufren tras la replicación está íntimamente acoplado a la inducción de una respuesta de daño en el DNA provocada por las vías dependientes de ATM y ATR (Fig. 4), la reparación de telómeros por NHEJ generaría fusiones letales de cromosomas y HDR podría cambiar la longitud de los telómeros. Por lo tanto, los telómeros necesitan bloquear estas cuatro vías distintas. Finalmente, los telómeros actúan como un "reloj" que determina la vida útil a nivel celular, mecanismo denominado "reloj mitótico" o Límite de Hayflick (Diotti y Loayza, 2011).

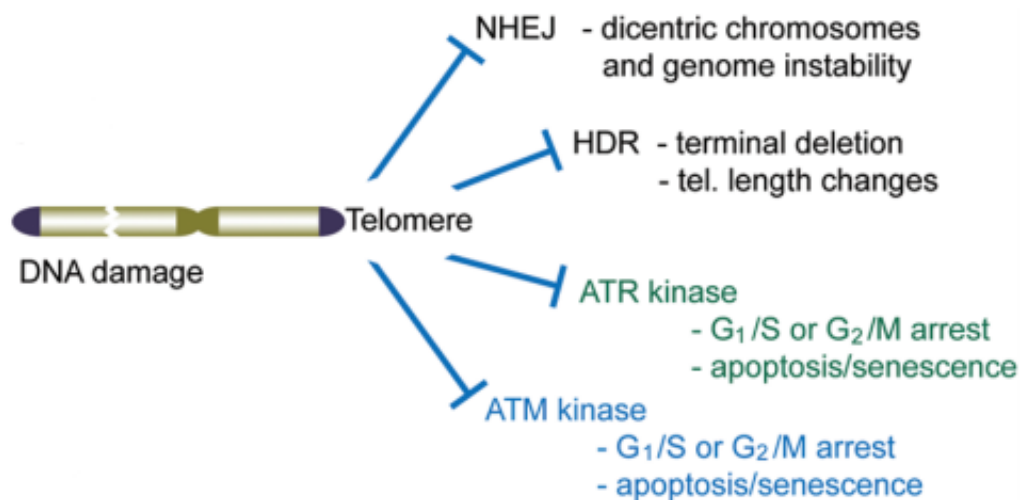


Fig. 4. Protección de los extremos terminales del cromosoma. Los telómeros deben inhibir las cuatro vías que amenazan los extremos finales de los cromosomas. En las roturas de DNA, la señalización de ATM y ATR pueden activarse y las lesiones pueden repararse por NHEJ o HDR (Tomada y modificada de de Lange, 2010).

Tanto la protección como la longitud de los telómeros están mediadas por un complejo proteico estable, llamado complejo de Shelterin. A su vez, este complejo se ha asociado con interacciones transitorias posteriores con factores proteicos que no están asociados cuantitativa o establemente con los telómeros.

3.3 Complejo Shelterin

El complejo Shelterin (también conocido como telosoma) se asocia y juega un papel esencial en la protección de los telómeros y la regulación de la telomerasa. En su forma más abundante, el complejo se compone de seis componentes principales: factor 1 de unión a repetidos teloméricos (TRF1), factor 2 de unión a repetidos teloméricos (TRF2), proteína de protección a telómeros 1 (POT1), factor nuclear 2 que interacciona con el factor 1 de unión a repetidos teloméricos (TIN2), proteína de interacción con TIN2 (TPP1) y proteína ortóloga de humano de la proteína activadora / represora 1 (RAP1), (Choi, 2011) (Fig. 5). TRF1 es una proteína multifuncional que está implicada en el mantenimiento de la longitud de los telómeros, progresión del ciclo celular, la resolución de los telómeros hermanos, así como la reparación de ruptura de doble cadena de ADN. En células que expresan telomerasa, TRF1 actúa como un regulador negativo del mantenimiento de la longitud de los telómeros, mientras que en células con vía ALT, TRF1 está implicado como mediador positivo de recombinación homóloga del mantenimiento de los telómeros (Florence et, al., 2016).

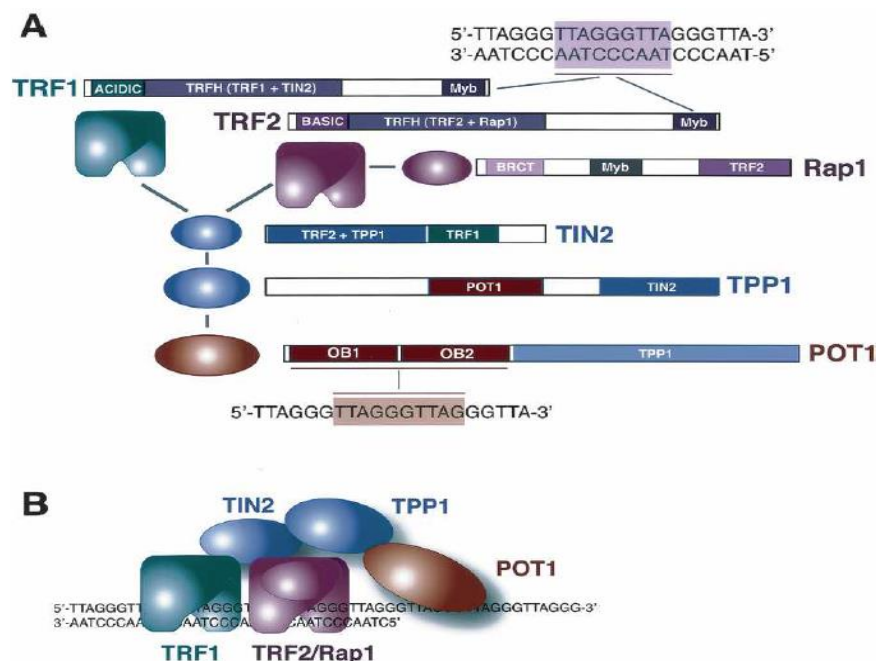


Fig. 5. Complejo de Shelterin. (A) Se muestran las seis subunidades que conforman el complejo, la estructura de sus dominios, sus interacciones proteína-proteína (TRF1-TIN2-TPP1-POT1, TRF2-TIN2-TPP1-POT1, TRF1-TIN2-TRF2) y sus respectivos sitios de unión a DNA. (B) Esquema del complejo Shelterin en el telómero. Las proteínas TRF1 y TRF2 se unen a dsDNA, mientras que POT1 se une a ssDNA en el extremo 3' (Tomada de de Lange, 2005).

La subunidad clave del complejo de Shelterin es TIN2, la cual se une simultáneamente a TRF1 (este enlace contribuye a la estabilización de TRF2 en los telómeros), TRF2 y TPP1, mientras que TPP1 conecta POT1 con TIN2. El sexto componente, RAP1, es un factor asociado a TRF2 (Liu, 2004). Además de este complejo de 6 miembros, el complejo Shelterin también puede estar presente en forma de subcomplejos que carecen de TRF1 o TRF2 / RAP1. Si bien la importancia de estos diferentes subcomplejos sigue sin estar clara, la función particular de cada una de las tres subunidades de unión a DNA (TRF1, TRF2 y POT1) es bien conocida (Choi, 2011). POT1 sirve para ocultar la proyección telomérica 3' de la telomerasa y de los mecanismos de detección de daño a DNA. Por lo tanto, la pérdida de POT1 conduce a la activación de la quinasas ATR, la formación de centros inducidos por la disfunción de los telómeros (TIF) y la inducción de la apoptosis o la detención del ciclo celular (Hockemeyer, 2007). TRF2 sirve para bloquear el reconocimiento de los telómeros a medida que se rompe el ds-DNA. La pérdida de la función de TRF2 conduce a la activación de la quinasas ATM, a la formación de 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 intercromosómicas al igual que en TRF1. TRF2 purificado puede promover el plegamiento de los telómeros artificiales en estructuras de T-loop, una propiedad no mostrada por TRF1 (Celli, 2005).

Entre los complejos asociados a los telómeros, los complejos Shelterin son únicos ya que reúnen las actividades de unión a DNA de POT1, TRF1 y TRF2. Los modelos actuales de las interacciones de estos complejos con los telómeros se extrapolan a partir de las propiedades de unión de DNA bien caracterizadas de TRF1 recombinante, TRF2 y POT1. TRF2 y TRF1 comparten la misma arquitectura básica definida por dos regiones conservadas: un dominio TRFH que media la homodimerización y un dominio de unión a DNA carboxi-terminal de la familia SANT / Myb (Fig. 6). Los dos factores dan lugar a homodímeros que se unen selectivamente al DNA ds-telomérico (Broccoli, 2009).

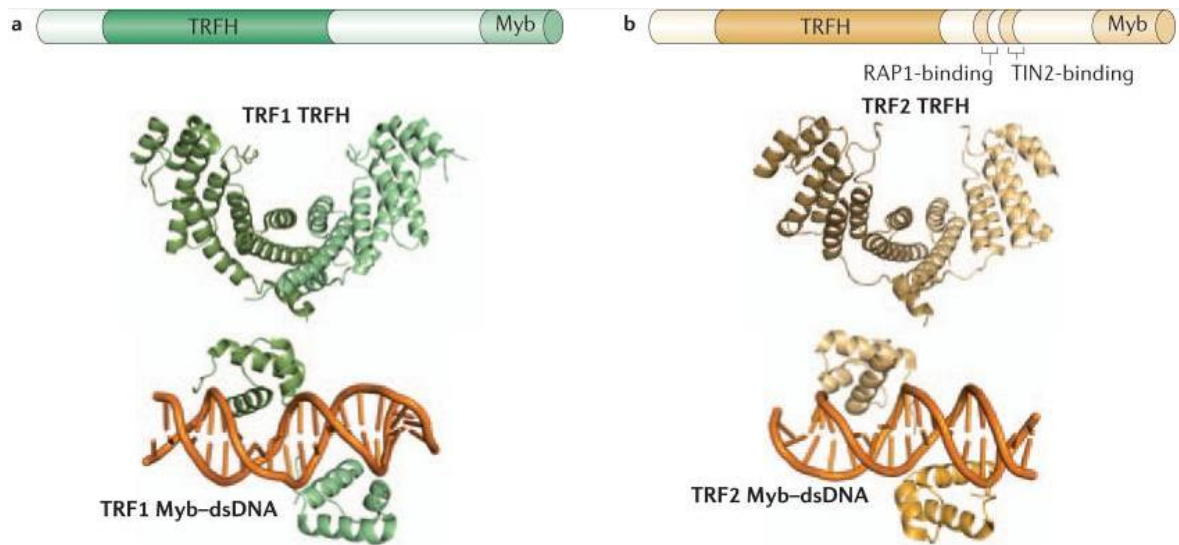


Fig. 6. Organización de los dominios y estructura de las proteínas TRF1 y TRF2. (A) En color verde se muestran los dominios funcionales de la proteína TRF1, así mismo se observa la estructura tridimensional del dominio TRFH (PDB: 1H6O) y el dominio Myb unido a dsDNA. (B) En color anaranjado se muestran los dominios funcionales de la proteína TRF2, también se observa la estructura tridimensional del dominio TRFH (PDB: 1H6P) y el dominio Myb unido a dsDNA (Tomado de Nandakumar, 2013).

A continuación, se muestra un modelo de interacciones simultaneas del complejo Shelterin con ds- y ss-DNA telomérico (Fig. 7), en donde se observa la abundancia relativa de los componentes del Shelterin y como podrían estar ligados a los telómeros en la configuración del t-loop.

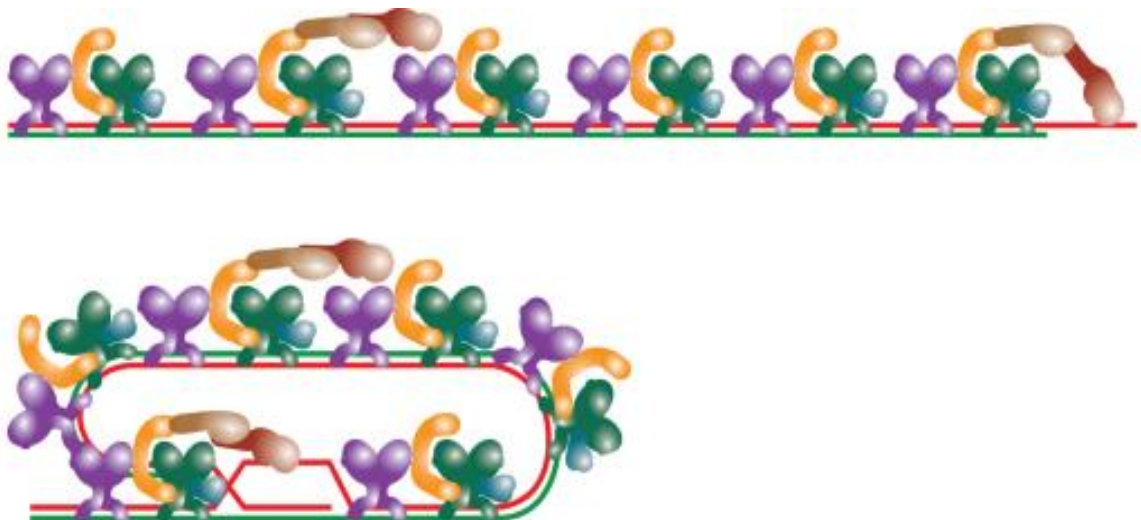


Fig. 7. Complejo de Shelterin en el DNA telomérico. Abundancia de los componentes de Shelterin y su relación funcional en la formación del t-loop (Tomada de de Lange, 2010).

probablemente inhibida por el complejo y se sospecha que TERRA (repetidos teloméricos que contienen RNA) desempeña un papel en esta inhibición (Fig. 9) (Tomado de O`Sullivan, 2010).

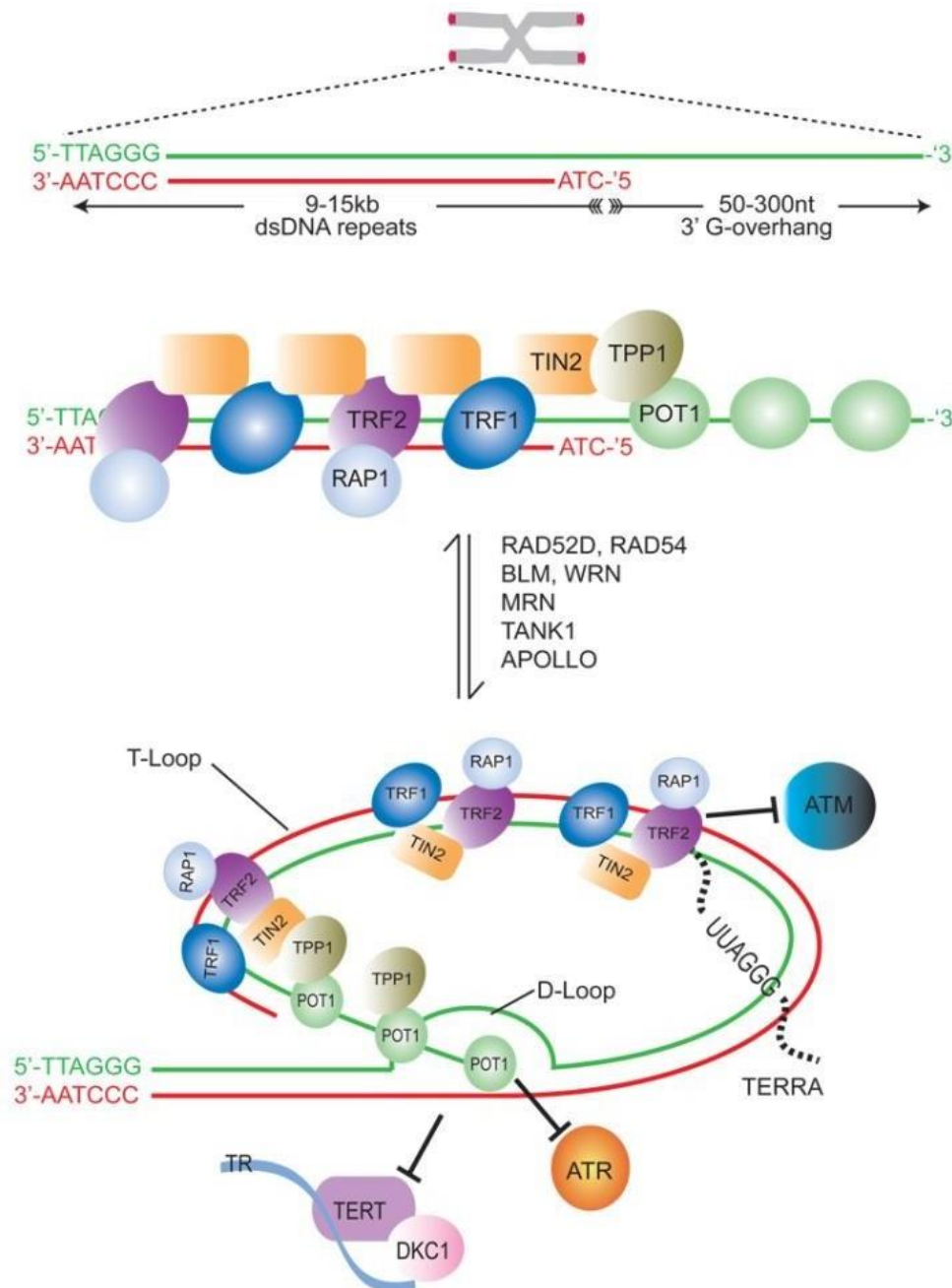


Fig. 9. Interacción de proteínas en los telómeros humanos. Los telómeros consisten en repeticiones TTAGGG, con una cadena principal rica en G y una cadena rezagada rica en C. Los miembros del complejo Shelterin interactúan con una gran cantidad de factores que se localizan transitoriamente, con frecuencia de manera dependiente del ciclo celular. Esos factores ayudan a la formación del t-loop lo que permite distinguir los extremos cromosómicos naturales de las roturas de DNA bicatenario. La cascada de señalización dependiente de ATM es inhibida por TRF2 y la ruta de señalización de ATR por POT1. La telomerasa es probablemente inhibida por el complejo, y se sospecha que TERRA juega un papel en esta inhibición. (Tomada de O`Sullivan, 2010).

En conjunto, aunque el complejo Shelterin consta de seis proteínas solamente, tiene un papel inmensamente complejo en la regulación de la longitud de los telómeros, protección contra ataque enzimático, reclutamiento de actividades enzimáticas requeridas y control de cascadas de señalización en los extremos del cromosoma natural. Si bien, la comprensión de los roles individuales de los componentes de Shelterin crece rápidamente, aún queda mucho por descubrir sobre la regulación transcripcional, traduccional y postraduccional de estos componentes y sobre la importancia de la estequiometría del complejo de Shelterin sobre los cambios del ciclo celular, el daño y la diferenciación del DNA.

No todas las proteínas en los extremos del cromosoma son parte del complejo. Varios criterios distinguen los componentes del complejo Shelterin de las proteínas no asociadas observadas en los telómeros (Tabla 2). El complejo Shelterin es abundante en los extremos del cromosoma y no se acumula en ningún otro lugar; está presente en los telómeros a lo largo del ciclo celular, y su función conocida se limita a ellos. Las proteínas que no son parte del complejo de Shelterin en los extremos de los cromosomas no cumplen con dos o tres de estos criterios, pero pueden desempeñar un papel importante en la función telomérica.

Tabla 2. Proteínas no Shelterin relacionadas con los telómeros humanos

Complejo de proteínas.	Función no telomérica.	Efecto en los telómeros.	Interacción.
MRE11/Rad50/Nbs1	Sensor de daño a DNA, reparación por recombinación.	Formación/Resolución t-loop HR. Mecanismo ALT	Asociado complejo Shelterin.
ERCC1/XPF	NER, reparación crosslink 3' endonucleasa.	Implicado en el proceso de proyección después de la pérdida de TRF2.	Asociado complejo Shelterin.
Helicasa WRN	Migración de hebra, resolución de DNA.	La deficiencia da como resultado la pérdida de latencia en telómeros.	TRF2
Helicasa BLM	Represión de migración de cruce de hebra.	Formación/Resolución t-loops.	TRF2
DNA-PK	NHEJ.	La deficiencia conduce a fusión leve.	Asociado complejo Shelterin.
PARP-2	BER.	No conocida.	TRF2
Tankyrases	Papel en mitosis.	Regulador positivo de los telómeros mediante inhibición de TRF1.	TRF1
Rad50	Desconocido ¿HR?	La deficiencia conduce a fusión leve.	Desconocido.

(Tomada de Lange, 2005).

Los complejos relacionados con el complejo Shelterin también se encuentran en los telómeros de otros eucariotas. Las proteínas tipo POT1 están presentes en casi todos los eucariotas, una proteína similar a TRF1/2 se encuentra la levadura de fisión *Schizosaccharomyces pombe* y en tripanosomátidos; mientras que Rap1 está presente en hongos. En contraste, las subunidades TIN2 y TPP1 hasta ahora solo se han encontrado en vertebrados; su aparición puede haber coincidido con la ganancia de un segundo gen tipo TRF. Por lo tanto, el complejo Shelterin parece estar formado a partir de una proteína dúplex de unión al DNA telomérica, una proteína de unión al DNA monocatenario (ssDNA) y Rap1. La única excepción a esta regla puede ser *Saccharomyces cerevisiae* que carece de una proteína similar a TRF y utiliza en su lugar un ortólogo Rap1 altamente divergente que se une al DNA telomérico de doble cadena. La pérdida de TRF2 activa la ruta de ATM quinasa, lo que conduce a la regulación positiva de p53 y a la detención del ciclo celular en G1/S mediada por p21 (de Lange 2001).

Experimentos recientes en un modelo de knockout de TRF2 de ratón han confirmado que la pérdida de TRF2 da como resultado la activación de ATM y la detención del ciclo celular dependiente de p53. El supresor tumoral p53 también está implicado en la respuesta al acortamiento de los telómeros. Por ejemplo, los ratones deficientes en p53 toleran mejor las consecuencias del acortamiento de los telómeros en las generaciones posteriores de ratones deficientes en telomerasa (Celli, 2015).

Varias características del complejo Shelterin son altamente conservadas. Una proteína de unión al dúplex de DNA telomérico puede ser reconocida en *S. pombe* y *Trypanosoma brucei* y en cada caso el ortólogo se une a Rap1 (Li, 2005). En vertebrados el dímero TPP1/POT1 es conectado a TRF/Rap1 y en *S. pombe* esta conexión es mediada por Poz1, la cual se une a Rap1 y el ortólogo de TPP1 es Tpz1, y la proteína de unión a DNA de doble cadena es Taz1, mientras que, en vertebrados, TIN2 se une a TPP1 y ambas a TRFs. La arquitectura conservada de esta estructura general del complejo de Shelterin en eucariontes sugiere que el complejo ya existió en su ancestro en común (Palm, 2008).

3.4 Dominio de unión a DNA MYB (DBD)

Dentro del genoma de los vertebrados se reportan pocos genes *myb*, sus productos se clasifican en tres formas generales, A-myb, B-myb y C-myb que comparten los dominios de unión al DNA con secuencias de aminoácidos similares y que se unen a la misma secuencia de reconocimiento. Una característica común de las proteínas Myb es la presencia de un dominio de unión a DNA que se encuentra conservado entre los animales, plantas, protozoarios y hongos, típicamente cerca de la región N-terminal. Un dominio Myb mínimo consiste en dos regiones de aminoácidos que son capaces de formar tres hélices alfa (H1, H2, H3) cada una y así unirse al DNA. Gran parte de las proteínas Myb presentan tres secuencias repetidas imperfectas en esta región, aunque también se han identificado proteínas con uno, dos, cuatro o más segmentos repetidos. Debido a la poca similitud de secuencia, las proteínas Myb se han clasificado de acuerdo con el número de secuencias repetidas adyacentes: Myb R1R2R3 (Myb3R), MybR2R3, MybR o MybR1R2R2R1/2 (Myb 4R). Las proteínas Myb con sólo una secuencia repetida no se pueden unir al ADN como tal, sino que deben formar heterodímeros u homodímeros, lo que permite a estas proteínas reconocer al DNA con alta afinidad y especificidad (Arratia y Aguirre, 2013).

El DBD media la actividad de unión de una manera específica de secuencia. Este dominio contiene tres repeticiones imperfectas conservadas de 52 aminoácidos (R1, R2, R3) (Fig. 10). Dentro de cada repetición hay tres residuos de triptófanos que están separados por 18 o 19 aminoácidos. Las proteínas que contienen DBD MYB son factores de transcripción que contienen un dominio relacionado con el DBD del c-MYB humano. El producto del protooncogén c-Myb humano es la proteína Myb prototípica y su DBD se repite (R1, R2, R3) donde, cada repetido adopta una conformación de hélice-giro-hélice que se intercala con un surco principal de la secuencia diana de DNA (C/TAACG/TG), estructura estabilizada por un núcleo hidrofóbico (Ogata, et. al; 1992).

Las estructuras resueltas de varios dominios relacionados con Myb han revelado que la longitud, orientación y la organización espacial de las tres hélices son muy similares a la estructura de las repeticiones de c-Myb. Los residuos altamente

conservados entre las proteínas relacionadas con Myb son los que participan en la formación del núcleo hidrofóbico, lo que indica que la identidad estructural del dominio es esencial para las funciones de repetición de Myb que se unen al DNA y que no se unen al DNA (Ko et. al, 2008).

Además de los residuos esenciales para la integridad estructural, otros residuos conservados entre todos los dominios Myb son un segmento de residuos de aminoácidos ácidos cerca del comienzo de la primera hélice. Sorprendentemente, estos aminoácidos están mas altamente conservados que los residuos involucrados en la unión a DNA. Los residuos de ácido glutámico y ácido aspártico a menudo están presentes en la posición seis a diez. Los dominios de unión a DNA típicamente exhiben un carácter fuertemente básico para formar enlaces iónicos con el esqueleto de fosfato de la doble hélice del DNA, lo que hace que la presencia de un parche de residuos ácidos altamente conservados en el MDBD sea bastante inusual. Debido a que estos residuos ácidos también están presentes en repeticiones relacionadas con Myb que no se unen al DNA, se ha planteado que estos residuos podrían ser importantes para las interacciones proteína-proteína en lugar de proteínas-DNA (Ko et. al, 2008).

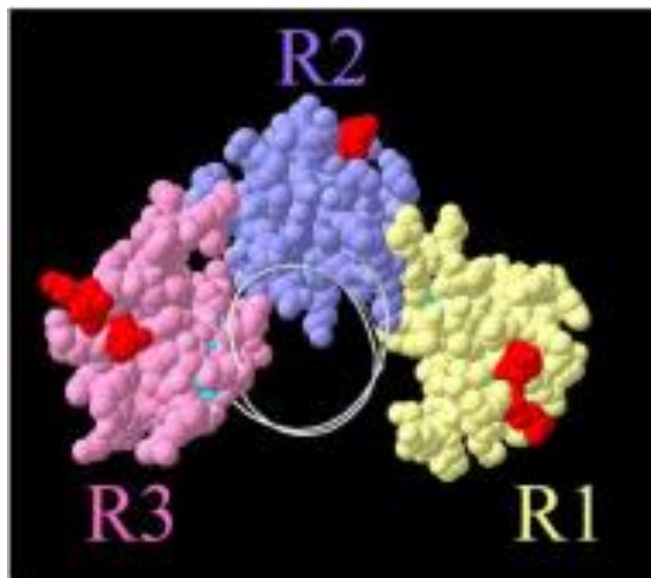


Fig. 10. Estructura cristalina del dominio de unión a DNA de c-Myb unido a DNA. La estructura fue resuelta por cristalografía a partir del cristal del dominio. La primera repetición (R1) se representa en amarillo, la segunda repetición (R2) en púrpura y la tercera repetición (R3) en rosa. Identificación de un grupo de residuos ácidos altamente conservados se representa en rojo. (Tomado de Ko, et. al. 2008).

3.5 Proteína TRF1

TRF1 es una proteína telomérica que se une a la matriz dúplex de repeticiones TTAGGG en los extremos de cromosomas de mamíferos. La proteína tiene un tamaño de 439 aa y un peso molecular de 50.246 kDa. TRF1 tiene homología con el dominio de unión a DNA de la familia Myb de factores de transcripción, pero, a diferencia de la mayoría de las proteínas relacionadas con Myb, TRF1 lleva uno en lugar de múltiples motivos de unión a DNA de tipo Myb. TRF1 se une al DNA como un homodímero usando un dominio MYB conservado cerca del extremo N-terminal de la proteína. TRF1 requiere ambas repeticiones de Myb para la formación de un complejo estable con DNA, lo que indica una similitud entre el modo de unión a DNA de TRF1 y otras proteínas relacionadas con Myb. TRF1 de humano y de ratón son proteínas con tres dominios reconocibles (Fig. 11), un dominio ácido en el dominio N-terminal, un dominio específico TRF conservado en el dominio C-terminal con una fuerte homología con el dominio de unión a DNA con las oncoproteínas Myb (Bianchi, 1997) (Fig. 12).

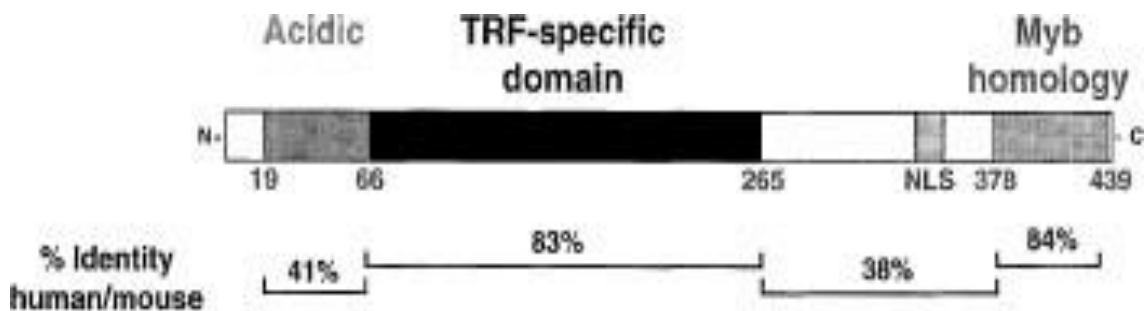


Fig. 11. Esquema de los dominios de TRF1. TRF1 de humano (hTRF1) y su identidad con la proteína TRF1 de ratón (mTRF1), se muestra el dominio ácido, el TRFH, la señal de localización nuclear (NLS) y el dominio MDBD (Tomada de Bianchi, 1997).

Se cree que el grado de ocupación de TRF1 en el DNA telomérico se correlaciona positivamente con la longitud de los telómeros, lo que implica que los telómeros más largos soportan más TRF1. Esta información se retransmite a POT1 a través de TPP1, lo que resulta en una mejor asociación de POT1, haciéndolo inaccesible para la telomerasa (Xin, et. al., 2008).

TRF1 también funciona como centro de interacción proteína-proteína dentro de la red de señalización de los telómeros (Tabla 3), interactuando directamente con otros miembros del complejo Shelterin (TIN2, TPP1 y POT1) y con un conjunto

diverso de proteínas y complejos de proteínas involucrados en el ciclo celular, reparación de DNA y recombinación, para mantener la estructura y longitud de los telómeros (Fig. 12) (Xin, et. al., 2008).

Tabla 3. Proteínas que interaccionan con las proteínas TRF1 y TRF2

Proteína accesoria.	Miembro Shelterin.	Sitio de interacción.	Significado de interacción.
Ck2	TRF1	N/D	Fosforila TRF1, promueve su unión al DNA telomérico y lo protege de la degradación.
Fbx4	TRF1	TRFH	Se dirige a TRF1 para la degradación mediada por ubiquitina.
Pin1	TRF1	TRFH	Regula la estabilidad de TRF1, orientación de TRF1 a degradación y longitud del telómero.
PinX1	TRF1	TRFH	Control de localización subcelular, mejora la unión de TRF1 a los telómeros, inhibe la telomerasa.
RLIM	TRF1	C-terminal aa 265-378	Se dirige a TRF1 para la degradación mediada por ubiquitina.
RNA pol II	TRF1	N/D	Puede regular la transcripción de TERRA.
Tankirasa 1, 2	TRF1	N-terminal ácido	Elimina TRF1 de los telómeros, regulación de su longitud.
ATM	TRF1, TRF2	N/D	Fosforila TRF1, dificulta su unión a DNA telomérico, la autofosforilación y regulación de ATM pueden ser inhibidas por TRF2.
Ku	TRF1, TRF2	TRFH	Protección final.

(Modificado de Oganessian, 2009)

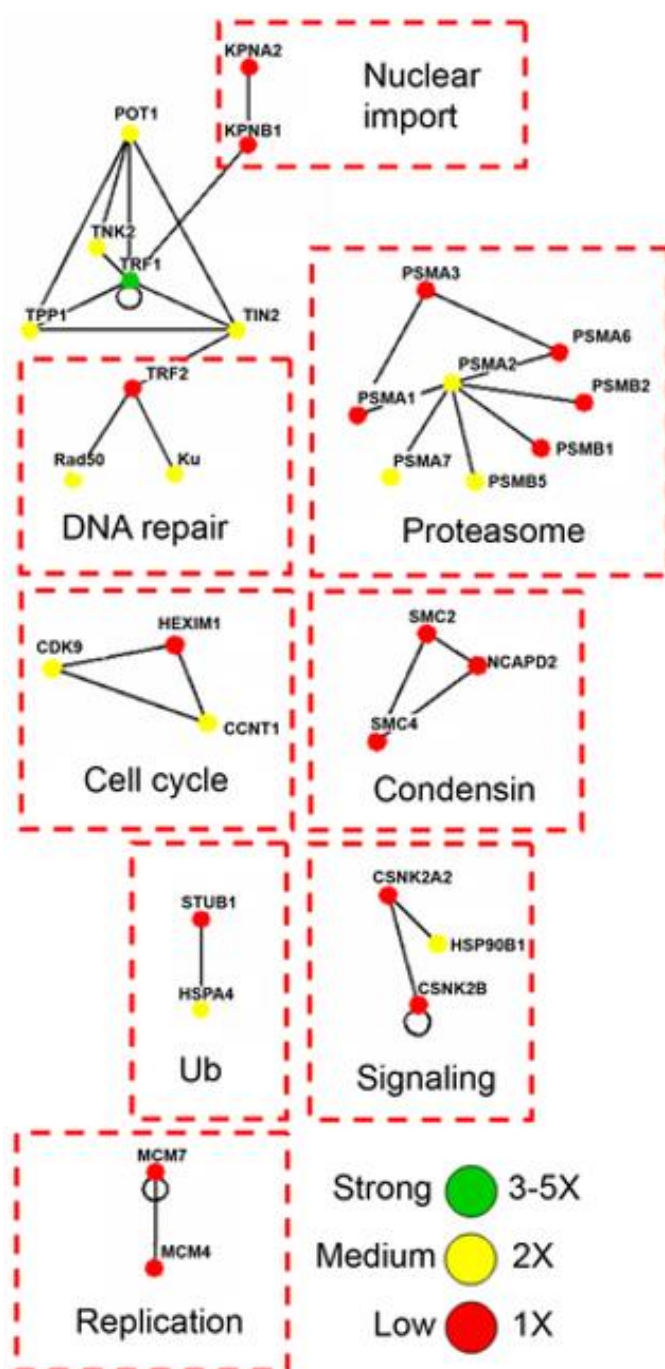


Fig. 12. Proteínas asociadas a TRF1. Diagrama de proteínas asociadas a TRF1 agrupadas en varias categorías según su función. Color indicativo de la prevalencia/frecuencia de la asociación: Verde (alto, 3-5 veces), amarillo (medio, 2 veces), rojo (bajo, 1 vez) (Tomado de Giannone, et. al., 2010).

3.6 Estructura de los dominios de la proteína TRF1

La familia de las proteínas TRF tiene una arquitectura similar (Fig. 13), definida por dos dominios característicos. En primer lugar, tiene el dominio MYB de unión al DNA en el extremo C-terminal. En segundo lugar, tienen un motivo de secuencia situado en el extremo amino terminal de aproximadamente 200 aminoácidos, denominado dominio de homología de TRF (TRFH) que es exclusivo de esta familia de genes (Broccoli, 1997). Los motivos TRFH son más divergentes en secuencia que sus motivos adyacentes Myb. TRF1 se une al DNA telomérico bicatenario como homodímeros. La función del dominio de TRFH está relacionada con la homodimerización pero también está implicado en la interacción con otras proteínas teloméricas a través de interacciones homeotípicas fuertes. La estructura terciaria del dominio de homodimerización de TRF1 en humanos se determinó con una resolución de 2,9 Å (Fig. 14). El dominio tiene una arquitectura enteramente α -helicoidal, y la dimerización se produce a través de nueve α -hélices. Las interfaces de dimerización presentan interacciones únicas que evitan la heterodimerización.

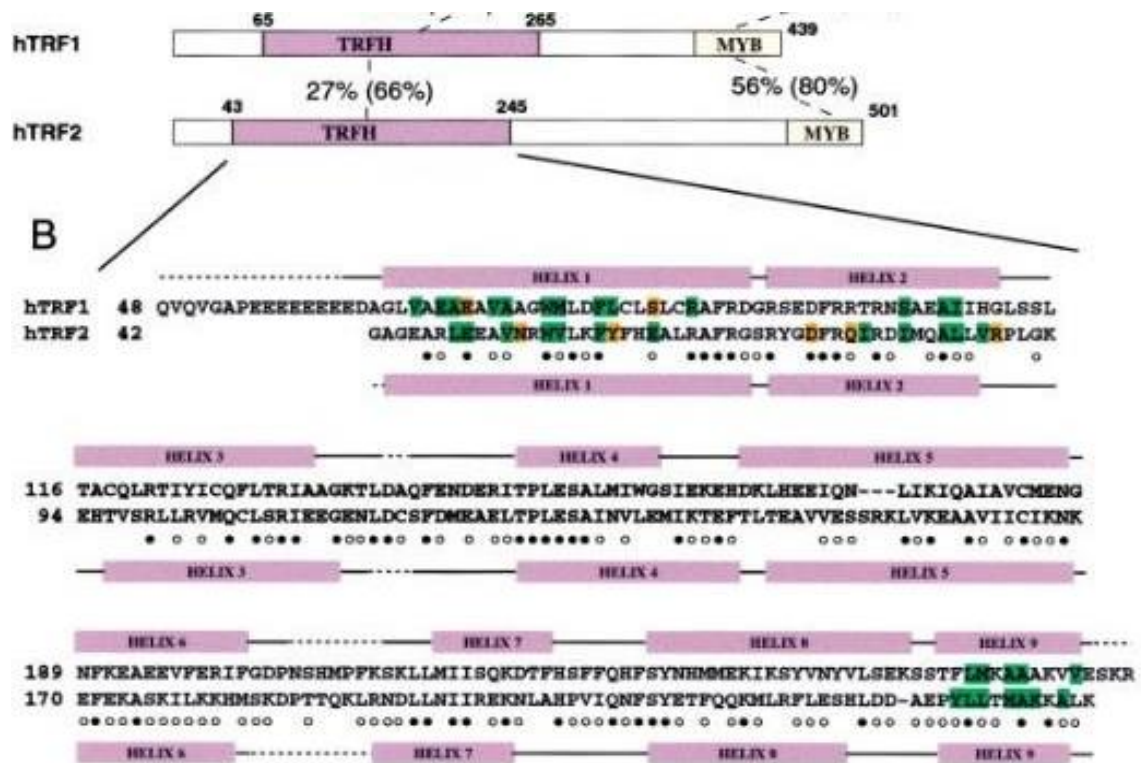


Fig. 13. Arquitectura del dominio TRFH de las proteínas TRF1 y TRF2 de *H. sapiens*. A) Representación esquemática que indican la ubicación de los dominios de secuencia. B) Alineamiento de las secuencias del dominio TRFH de TRF1 y TRF2 (Tomada de Fairall, 2001).

La secuencia de aminoácidos del dominio TRFH de las proteínas TRF1 y TRF2 está conformado por residuos de aminoácidos conservados como Ala, Trp, Leu, Phe y Arg, los cuales están involucrados en interacciones hidrofóbicas (Fairall, 2001). La secuencia de aminoácidos de los dominios TRFH tiene un 27% de identidad, sin embargo, la superposición de la estructura del dominio TRFH de TRF1 con la de TRF2 muestra que los dos dominios TRFH tienen estructuras 3D casi idénticas. Cada dímero está formado por dos monómeros que interactúan en una disposición antiparalela formando un dímero simétrico cuya estructura general se asemeja a una herradura retorcida (Fig. 14).

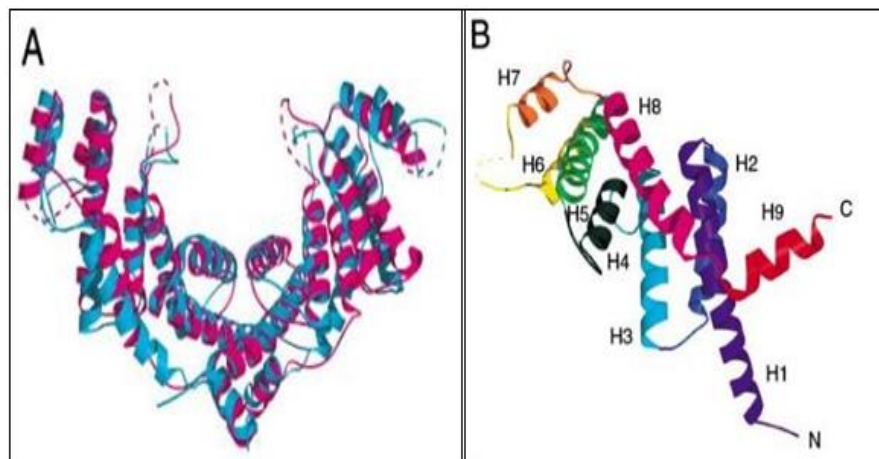


Fig. 14. Estructura de los dominios de dimerización TRFH de las proteínas TRF1 y TRF2. (A) Superposición de los dominios de dimerización de TRF1 (cyan) y TRF2 (magenta) que muestra la forma de herradura de la estructura. (B) Estructura de un monómero TRFH en el dominio de dimerización de TRF1. Las α -hélices son de color púrpura a rojo y van numeradas del extremo N al extremo C-terminal (Tomada de Fairall, 2001).

El dominio de unión a DNA carboxi-terminal (DBD) fue identificado originalmente a través de su similitud de secuencia con el motivo de unión a DNA tipo Myb. Los dominios DBD se unen a sus dominios de dimerización por medio de enlaces muy largos (aproximadamente 100 aminoácidos en TRF1 y 150 aminoácidos en TRF2) (Fig. 15), que parecen no tener una estructura inherente. Los residuos de aminoácidos para la estructura del dominio son Trp450-Arg496. Las estructuras contienen tres alfa hélices, la hélice I está formada por Val452-Tyr465, la hélice dos formada por Trp470-Asn476 y la hélice tres formada por Ala484-Arg496. Estas tres hélices mantienen su conformación debido a la presencia de Trp450, Val458, Val642, Trp470, Ile473, Phe479, Trp491, Met494 los cuales mantienen un

reconocimiento de DNA, que median interacciones de enlace de hidrógeno adicionales entre la proteína y el DNA lo que aumenta la especificidad y afinidad de reconocimiento. La conservación de la estructura del agua sugiere que las proteínas TRF reconocen la estructura del DNA hidratando el lugar de la secuencia de DNA *per se*. Estas estructuras confirman que las proteínas teloméricas utilizan homeodominios para el reconocimiento de DNA de doble cadena (Koning et. al; 1996).

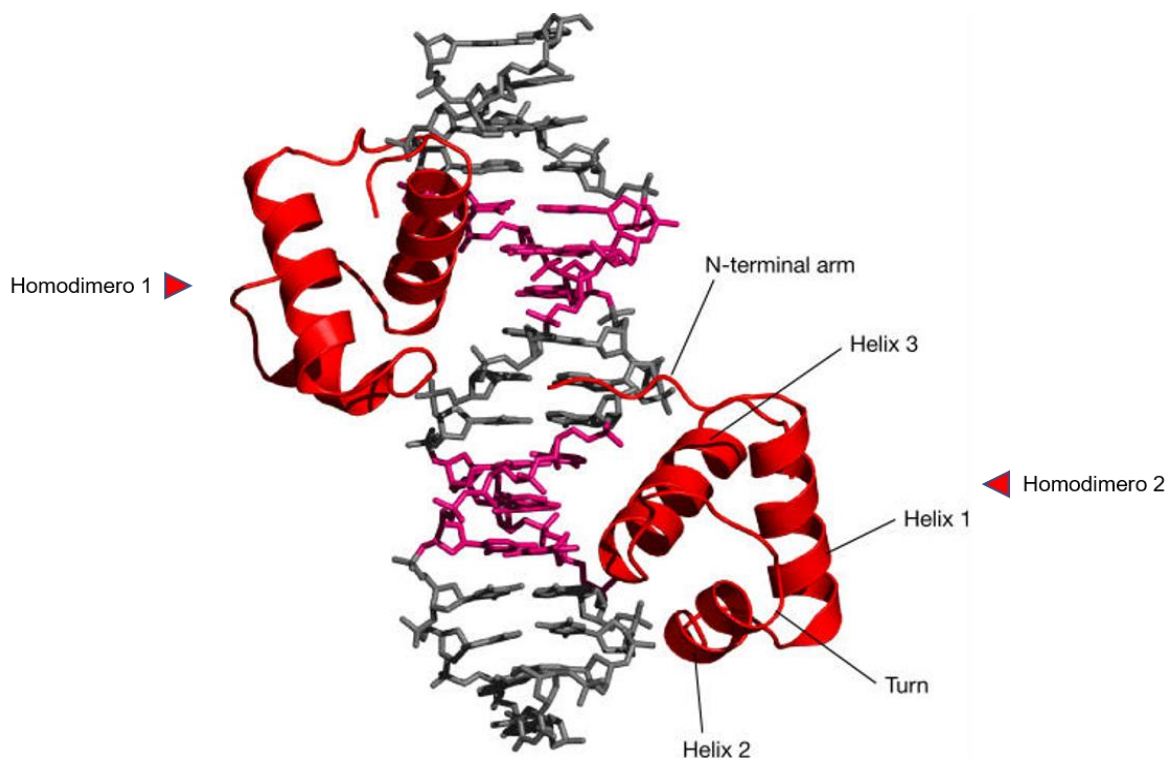


Fig. 16. Estructura tridimensional del dominio MDBD de TRF1 unido a DNA bicatenario. Las moléculas de la proteína en el complejo se muestran en color rojo. Se señalan las tres α -hélices características y su conformación hélice-giro-hélice. Los conjuntos de pares de bases GC presentes en el sitio de unión a DNA están resaltados en magenta (Tomado de Court, 2005).

La estructura tridimensional del dominio de unión a DNA Myb (DBD) de la proteína TRF1 fue obtenida a través de la superposición de 20 estructuras obtenidas por medio de MRN del dominio de unión a DNA. A continuación (Fig. 17) se muestra una vista estereoscópica de los átomos (N, C α , C') de DNA y los aminoácidos Arg378-Ser435 involucrados en una conformación bastante rígida que se une al DNA.

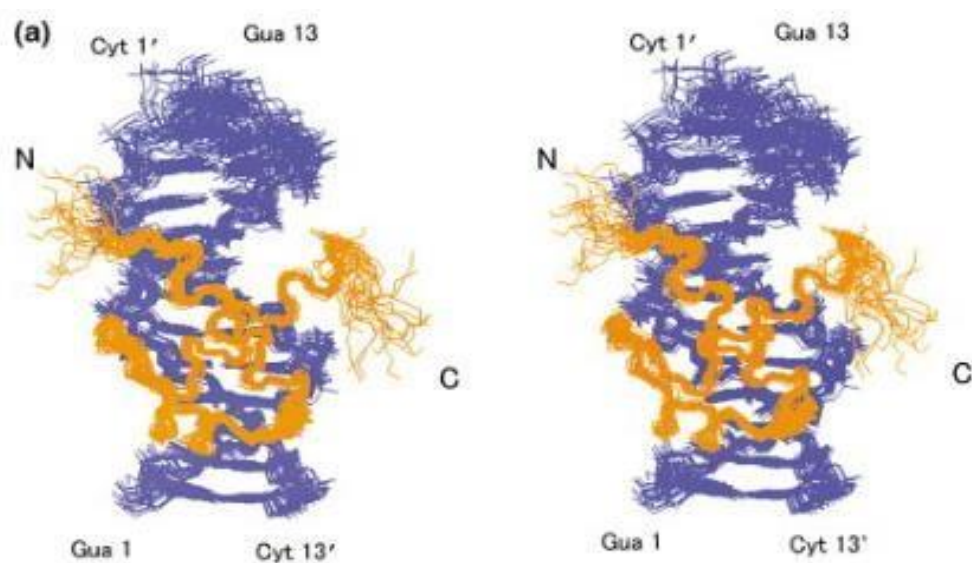


Fig. 17. Estructuras del MDBD de TRF1 humano. Vista estereoscópica de la estructura MRN del dominio DBD Myb de la proteína TRF1 unida a un fragmento de DNA de 13 pb. Los residuos de aminoácidos Arg378–Ser435 se muestran en anaranjado y los átomos de DNA se muestran en azul (G1-G11). (Tomado de Nishikawa y cols., 2001).

3.7 Modificaciones postraduccionales de TRF1

Tanto TRF1 como TRF2 desempeñan múltiples funciones en los telómeros, incluida la protección, replicación y mantenimiento de la longitud telomérica. La regulación de TRF1 y TRF2 en estos procesos se controla mediante modificaciones postraduccionales (Fig. 18), a veces de forma dependiente del ciclo celular, que afecta funciones claves como la unión a DNA, la dimerización, la localización, la degradación y las interacciones con otras proteínas (Walker, 2012).

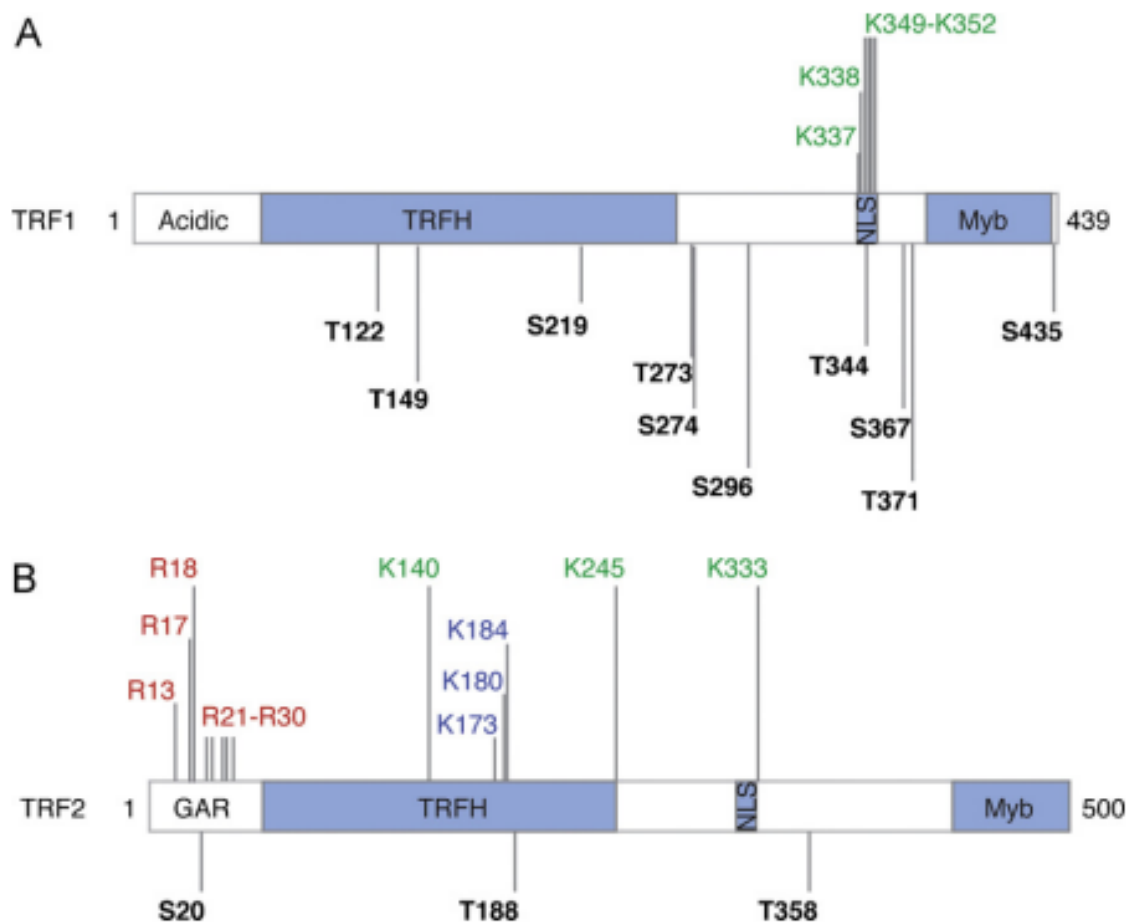


Fig. 18. Modificaciones postraduccionales de TRF1 (A) y TRF2 (B). Se muestran los dominios de TRF1 y TRF2. Los sitios de fosforilación se muestran en color negro, los sitios de ubiquitinación en azul, los sitios de SUMOilación en verde y los sitios de metilación de arginina en rojo. NLS: señal de localización nuclear (Tomado de Walker, 2012).

Varias líneas de evidencia sugieren que el papel de TRF1 como un modulador de la longitud de los telómeros se rige por sus modificaciones postraduccionales que determinan su abundancia en los telómeros; estas modificaciones y su efecto consecuente sobre el destino de TRF1 y los telómeros se resumen a continuación (Tabla 4) (Oganesian, 2009).

La fosforilación, proceso de adición de un grupo fosfato, tiene un papel principal en la unión, estabilidad y localización de TRF1. La ubiquitinación es un proceso usado para la degradación de proteínas en el proteosoma, así como, eventos de señalización (Al-Hakim, 2010). La SUMOilación (small ubiquitin-like modifier) requiere una cascada de enzimas que implican una enzima activadora, una enzima conjugadora y una proteína ligasa. En general, las SUMOilaciones controlan la ubicación celular, la interacción con otras proteínas y otras funciones reguladoras como la respuesta al daño de DNA y la reparación de DNA (Bekker-Jensen, 2011). Las poli (ADP-ribosila) polimerasas (PARP) son una familia de enzimas que usan NAD⁺ como sustrato para generar polímeros de ADP-ribosa sobre residuos de ácido glutámico de los receptores de proteína, implicadas en un gran número de procesos celulares que implican principalmente reparación de DNA y apoptosis. La tankirasa 1 y su homólogo tankirasa 2 se identificaron como proteínas que interactúan con TRF1, uniéndose a una secuencia corta localizada en el dominio ácido N-terminal de TRF1 (Guettler, 2011).

Tabla 4. Modificaciones postraduccionales de TRF1

Modificación	Aminoácidos	ID de los sitios	Enzimas Modificadoras	Función
Fosforilación	T122	MS, MA	CK2	La inhibición de CDK2 deteriora la unión de TRF1 a DNA telomérico y su estabilidad.
	T149	MA, Ab	CDK	La fosforilación de T149 es requerida para la interacción de TRF1 con PIN2.
	S219	MA	ATM	S219 está implicada en la respuesta a daño al DNA.
	T273	MA	AKT	La sobreexpresión de AKT aumenta los niveles de TRF1 y los promotores de acortamiento telomérico.
	S274	MS, MA	-	La mutación de S274 afecta la unión de TRF1 al DNA telomérico.
	S296	MA	Aurora A	S296 está implicada en los defectos mitóticos inducidos por Aurora A.
	T344	MA	Cdkl	T344 está implicado en la fosforilación de TRF1 por Plk1.
	S367	MS, MA, Ab	ATM	La fosforilación de S367 inhibe la unión de TRF1 a DNA telomérico.
	T371	MS, MA, Ab	Cdkl	La fosforilación de T371 mantiene libre a TRF1 y se requiere para la resolución de telómeros hermanos.

	S435	MS, MA	Pik1	La fosforilación de S435 está implicada en la unión de TRF1 a DNA telomérico.
Ubiquitinación	-	-	Fbx4	La eliminación de Fbx4, aumenta los niveles de TRF1, lo que produce un acortamiento de los telómeros y un deterioro del crecimiento celular.
	-	-	RLIM	La sobreexpresión de RLIM, una ligasa E3, disminuye los niveles de TRF1, lo que conlleva a un acortamiento de los telómeros.
SUMOylación	K333, K338, K349-K352	MA	MMS21	La inhibición de MMS21 promueve el acortamiento y la inestabilidad de los telómeros, lo que resulta en la inducción de senescencia celular en células ALT.
PARsylación	-	-	Tankirasa 1 y 2	La poli ADP-ribosylación de TRF1 por tankirasa elimina TRF1 de los telómeros, lo que conduce a un mayor acceso de la telomerasa en los extremos de los telómeros.

ID: Identificado por método (MS: espectrometría de masas, MA: análisis mutacional, Ab: anticuerpo específico de sitio). (Modificado de Walker, 2012).

3.8 Proteínas TRF en diversos organismos

La clonación y secuenciación de las secuencias teloméricas en otros eucariotas reveló que la naturaleza de las repeticiones teloméricas y la presencia de una cadena rica en G son características comunes de casi todos los telómeros eucariotas, además, se estableció que la hebra que contiene el extremo 3' sobresale en ambos extremos del cromosoma plegándose en las estructuras llamadas T-loop (Pfeiffer, 2013). La estructura de los telómeros en los extremos de los cromosomas eucariotas lineales y las proteínas presentes en el complejo telomérico como TRF1, se ha encontrado no solo en mamíferos, sino también en plantas y protozoos lo que sugiere un papel conservado en la protección de los extremos cromosómicos para que no sean reconocidos como daño a DNA (Griffith, 1999).

Además de los organismos modelo eucariotas, como la levadura y los vertebrados, se han identificado telómeros o telomerasas en eucariotas filogenéticamente diversos, incluidos los ciliados, plantas, algas verdes, dinoflagelados, tripanosomátidos y apicomplexas, en donde se ha encontrado extensa heterogeneidad de secuencia en las repeticiones teloméricas, sin embargo, el mecanismo de mantenimiento telomérico cromosómico parece estar conservado evolutivamente (Uzlíková, et. al., 2017).

3.8.1 Tripanosomátidos

Homólogos de TRF humano se han encontrado codificados en los genomas de *T. brucei*, *T. cruzi* y *L. major* basados en similitudes de secuencia con el dominio de unión a DNA tipo Myb C-terminal. En *L. amazonensis* se han identificado dos proteínas de protección a los telómeros: LaTBP1 y LaTRF. La proteína LaTBP1 se ha propuesto que comparte alta identidad y similitud con las proteínas TRFs putativas de *L. major*, *L. infantum*, *L. braziliensis*. Además, TbTRF tiene similitudes de secuencia con LaTRF con el dominio MYB y el dominio TRFH de TRF1 y TRF2 humana. En conjunto, se cree que laTRF comparte alta similitud de secuencia con TbTRF debido a que las especies están relacionadas filogenéticamente (Fig. 19) (da Silva, 2010).

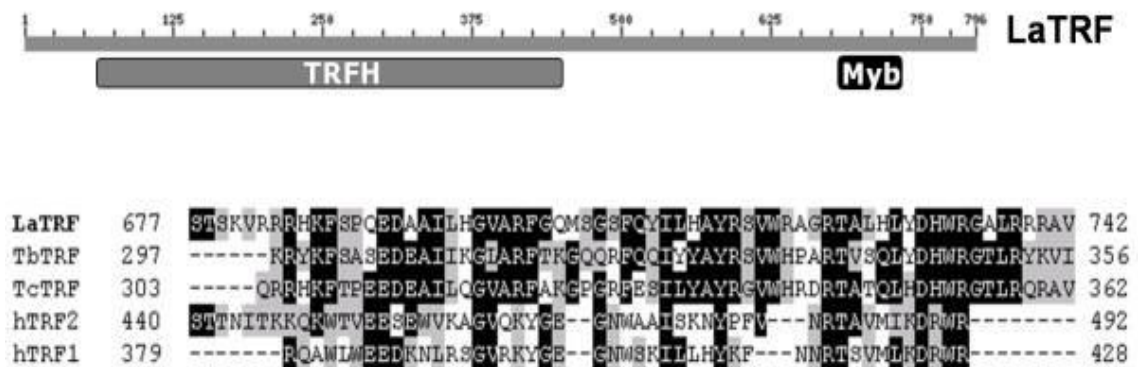


Fig. 19. LaTRF es homóloga a las proteínas TRFs de vertebrados y *T. brucei*. A) Representación de los dominios de la proteína LaTRF B) Alineamiento múltiple del dominio Myb-like de las proteínas TRFs de humano (hTRF1 y hTRF2), *L. amazonensis* (LaTRF), *T. brucei* (TbTRF) y *T. cruzi* (TcTRF). (Tomado de da Silva, 2010).

3.8.2 *Plasmodium falciparum*

En *P. falciparum*, el agente causal de malaria en humanos, las regiones teloméricas y subteloméricas juegan un papel fundamental en la supervivencia, virulencia y transmisión del parásito. Como organismos unicelulares, dependen de mecanismos eficientes de mantenimiento de los telómeros para asegurar la capacidad de proliferación a lo largo de su complejo ciclo de vida. A pesar del papel esencial y evolutivamente conservado de las proteínas TRF de tipo MYB en estos procesos, dicho factor no se ha descrito en parásitos de *P. falciparum* y parece faltar también en el ciliado *Tetrahymena thermophila*, un modelo tradicional para estudiar la biología de los telómeros. Esto es sorprendente, dado que la estructura general del tracto telomérico de *P. falciparum* es idéntico al observado en otros eucariotas. Los telómeros de *P. falciparum* consisten en repeticiones AGGGTT(T/C) degeneradas, tienen una longitud media de 1-1.2 Kb y tienen una proyección telomérica en el extremo 3' de la hebra rica en G, que parece formar estructuras *t-loop* (Bertschi, et. al., 2017).

Una proteína con una posible actividad similar a TRF se ha identificado en *P. falciparum* denominada PfTRZ o proteína de unión a repeticiones teloméricas, carece de un dominio MYB pero se une a las repeticiones de telómeros a través de un dominio de dedo de zinc tipo C2H2 (Fig. 20). *In vivo*, PfTRZ se une con alta especificidad al tracto telomérico y a las repeticiones de telómeros intersticiales río arriba de los genes de virulencia ubicados en las regiones subteloméricas. Se

sabe que PfTRZ regula la homeostasis de la longitud de los telómeros y es necesaria para la progresión eficiente del ciclo celular. Curiosamente, PfTRZ también se une y regula la expresión de los genes 5S rDNA. En un análisis filogenético, PfTRZ se ha identificado como un homólogo funcional del factor de transcripción básico TFIIIA, el cual pudo haber adquirido una nueva función en el mantenimiento de los telómeros al inicio del linaje apicomplexa (Bertschi, et. al., 2017).

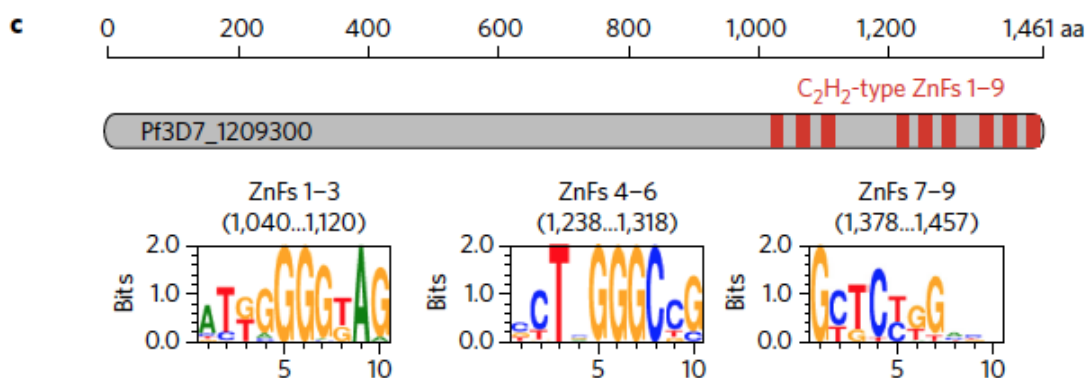


Fig. 20. Factor de unión a repetidos teloméricos de tipo C₂H₂ de *P. falciparum* (PfTRZ). Representación esquemática de la proteína con dominio de dedos de Zinc de tipo C₂H₂ (PF3D7_1209300). También se representan las secuencias de unión a DNA predicho para secuencias teloméricas en este parásito (Tomado de Bertschi, et. al., 2017).

3.8.3 *Giardia intestinalis*

El parásito *G. intestinalis* condensa sus cromosomas lineales extremadamente cortos, hecho que también se ha descrito en otros organismos unicelulares y que valida la cuestión de que la longitud y la reorganización dentro de regiones subteloméricas dinámicas constituyen una ventaja evolutiva para la adaptación del organismo. En el parásito *G. intestinalis* se ha encontrado la repetición del pentanucleótido TAGGG de 5pb rico en G que forma una estructura G-cuadruplex. Diversos experimentos han demostrado la presencia de una enzima telomerasa activa que sintetiza las repeticiones teloméricas. Además de la repetición telomérica de tipo TAGGG se demostró que la telomerasa de *G. intestinalis* sintetiza *in vitro* también las variantes TAAGG y TAAGGG. Se ha propuesto que la longitud del telómero de *G. intestinalis* no excede 1 kb. Además, el genoma de *G. intestinalis* codifica también la subunidad catalítica de la

telomerasa TERT, que se ha demostrado es bastante divergente de secuencias y carece del motivo N-terminal que se requiere para la unión de la subunidad de RNA de la telomerasa. Cabe señalar que *G. intestinalis* carece de varios genes, como Ku70, Ku80 y proteínas láminas canónicas, que generalmente son responsables de la unión de los telómeros a la envoltura nuclear. En resumen, a pesar de sus características inusuales, que incluyen una telomerasa estructuralmente divergente pero activa, repetidos teloméricos únicos y telómeros relativamente cortos, los datos actuales respaldan la opinión de que los extremos cromosómicos en *G. intestinalis* se mantienen de una manera conservadora que es similar a otros eucariotas (Uzlíková, et. al., 2017).

3.8.4 *Dictyostelium discoideum*

El genoma de la ameba *D. discoideum* se terminó de secuenciar en 2005 y no se encontró ninguna secuencia telomérica canónica en los extremos de ninguno de sus seis cromosomas. Este organismo presenta elementos palindrómicos extracromosómicos que contienen dos copias de rDNA, presentes al final de los cromosomas. La estructura de la región terminal del rDNA consiste en repeticiones de secuencia A(G)_n donde el número de Gs es variable con una longitud de 29 pb. Estas repeticiones se extienden como una proyección telomérica en el extremo 3' rico en G, similar a la de vertebrados, que está precedido por cuatro repeticiones en tándem de dos motivos de DNA diferentes (Rodríguez, et. al., 2019). La secuencia de nucleótidos de los extremos de la secuencia palindrómica y de los cromosomas no pudo determinarse completamente debido a la presencia de secuencias repetidas. La secuenciación del genoma completo de *D. discoideum* indica que las partes terminales del palíndromo de rDNA están presentes en los extremos de los cromosomas por lo que se propone representan la región telomérica. Así mismo, contiene un gen que codifica para un homólogo de la transcriptasa inversa (TERTHP), que tiene actividad de supuesta telomerasa y es necesaria para mantener la estructura de los telómeros. Sin embargo, no ha sido posible encontrar proteínas homologas a las del complejo Shelterin en mamíferos, sin embargo, se han encontrado varias proteínas que se unen a secuencias teloméricas como TRF1 y TRF2 de humano que tiene en común un dominio de unión a DNA de tipo MYB y que podrían

desempeñar un papel en la unión y protección de los telómeros como son MybAA y MybL quienes tienen la mayor similitud con las proteínas TRF1 y TRF2. La proteína MybAA y MybL presenta un porcentaje de identidad del 35 y 39%, respectivamente de identidad con respecto al motivo Telebox. También, *D. discoideum* tiene una proteína que se une al DNA telomérico monocatenario denominada DDB_GO286375 que presenta 26% de identidad con el pliegue OB (pliegue de unión oligonucleótido/oligosacárido) de POT1 humana. Se ha propuesto que los telómeros de este organismo podrían mantenerse mediante mecanismos de recombinación homóloga como la vía de alargamiento alternativo de los telómeros (ALT) (Rodríguez, *et. al.*, 2019).

IV. Antecedentes Particulares

4.1 *Entamoeba histolytica*

E. histolytica es un protozooario comensal del intestino grueso, que en ocasiones invade la mucosa intestinal, puede diseminarse por vía hemática y es el agente responsable de la amebiasis, afección cosmopolita. Este protozooario pertenece al subphylum *Sarcodina*, que incluye a los protozoarios móviles dotados de pseudópodos (Pumarola, 1991). Las características claves para la supervivencia de *E. histolytica* son su capacidad de adherirse, lisar y fagocitar el epitelio intestinal y otros tejidos del huésped. Varios genes que codifican adhesinas, proteasas y ameboporos han sido implicados en la patogénesis y virulencia.

El parásito presenta dos fases de desarrollo bien establecidas: el trofozoíto y el quiste, que constituyen la forma invasiva e infectante, respectivamente (Fig. 21). El trofozoíto, o forma móvil, es extraordinariamente pleomórfico, ya que su aspecto y movilidad están influenciados por los cambios de pH, potencial REDOX y osmolaridad. *E. histolytica* se multiplica por fisión binaria y es sensible al jugo gástrico y a los agentes externos. Su hábitat comprende la luz y pared del colon. Su tamaño variable oscila entre 10 y 60 nm (Pumarola, 1991).

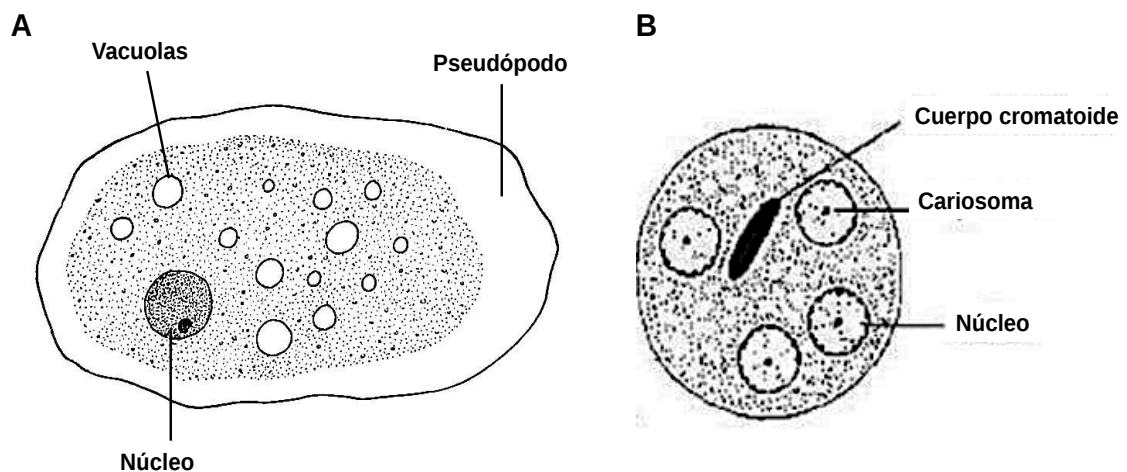


Fig. 21. Morfología del parásito *E. histolytica*. (A) Trofozoíto activo en el cual se muestra el núcleo, vacuola con un eritrocito y pseudópodos. (B) Quiste tetra nucleado el cual muestra sus núcleos y su cariosoma (tomado de *E. histolytica*: Morphology, life cycle, Pathogenesis, clinical manifestation, lab diagnosis and Treatment - Online Biology Notes).

4.2 Genoma de *Entamoeba histolytica*

El genoma de *E. histolytica* se encuentra en un núcleo esférico con un cúmulo de cromatina pequeño y puntiforme en el centro, encerrado en una cápsula llamado cariosoma o endosoma (Petri, 1993). También presenta cromatina adherida a la cara interna de la membrana nuclear, distribuida en forma más o menos homogénea (Romero, 1993).

El genoma de *E. histolytica* consta de 23.7 megabases distribuidas entre 888 andamios. Los 9,938 genes predichos tienen un tamaño promedio de 1.17 kilobases (kb) y comprenden el 49% del genoma. Los cromosomas de *E. histolytica* no se condensan en su totalidad y la extensa variabilidad de longitud observada entre los cromosomas homólogos de diferentes aislamientos hace que el número exacto de cromosomas sea difícil de determinar. El metabolismo de *E. histolytica* parece haber sido conformado por la pérdida de genes secundarios y la transferencia lateral de genes (LGT), principalmente de linajes bacterianos. La glucosa es la principal fuente de energía, sin embargo, varias glucosidasas y azúcares quinasas parecen haber sido adquiridas a través de LGT, que probablemente permiten a *E. histolytica* usar azúcares que no sean glucosa, por ejemplo, fructosa y galactosa. La mayoría de las vías para la biosíntesis de aminoácidos se ha eliminado, excepto las de serina y cisteína, componentes principales de la resistencia al estrés oxidativo. El citoesqueleto tiene varias funciones importantes en la motilidad del parásito, la muerte dependiente de contacto y la fagocitosis de las células epiteliales intestinales del huésped. *E. histolytica* contiene más de 100 fosfatasas, que desfosforilan proteínas, algunas con una característica inusual que es la presencia de un número variable de dominios de repetidos ricos en leucina involucrados en las interacciones proteína – proteína.

E. histolytica se alimenta de bacterias en la luz del colon y lisan las células epiteliales del huésped después de la invasión de la pared intestinal. Se han caracterizado varios determinantes de virulencia amebiana (Fig. 22), incluida una lectina GalGalNAc de múltiples subunidades implicada en la adhesión a células huésped, cisteína proteasas que degradan la matriz extracelular del huésped y

péptidos formadores de poros (ameboporos) capaces de lisar células diana. El análisis del genoma revela redundancia en los genes que codifican estos factores de virulencia (Loftus, 2005).

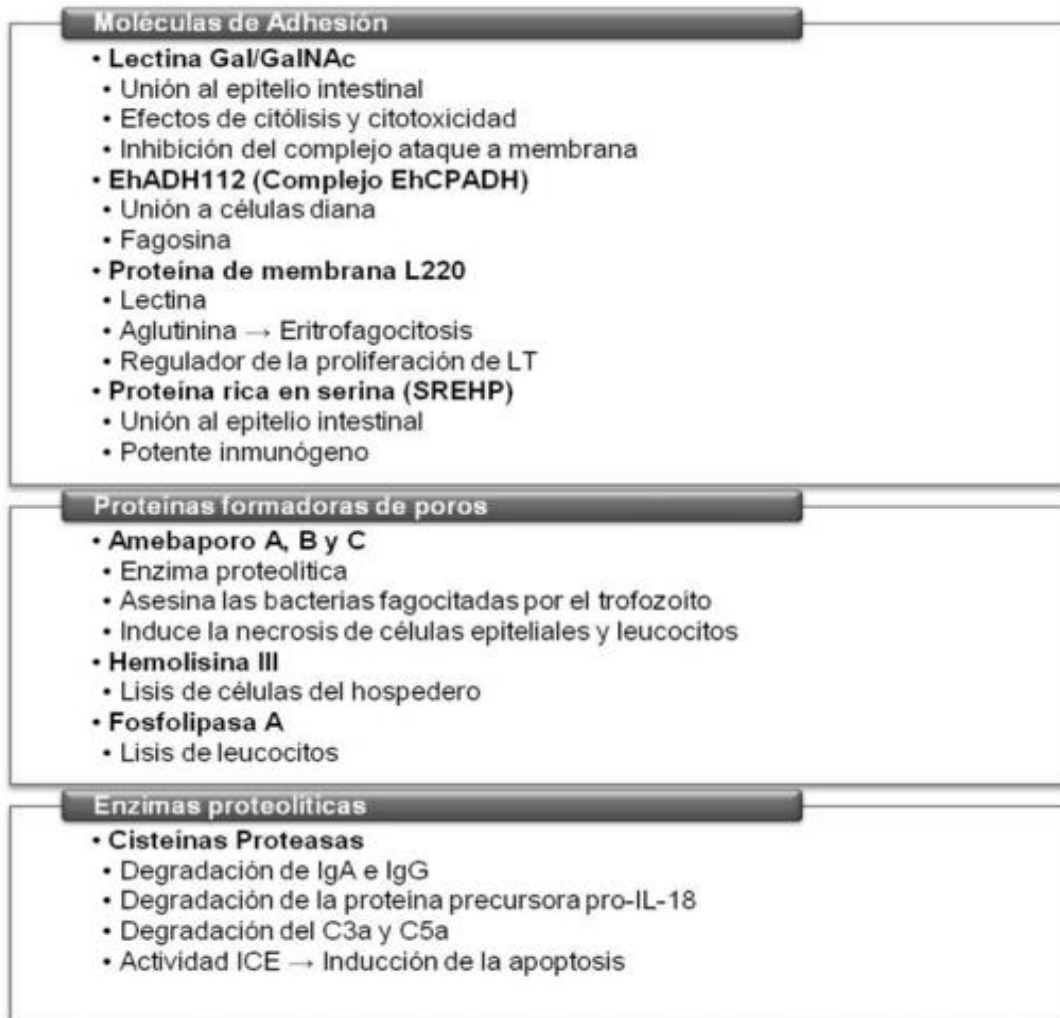


Fig. 22. Determinantes de virulencia amebiana. Principales funciones de las proteínas secretadas por trofozoítos de *Entamoeba histolytica* (Tomado de Trejo y Castañón, 2009).

4.3 Ciclo de vida de *Entamoeba histolytica*

E. histolytica es monogenética o monoxénica, lo que significa que su ciclo de vida se completa sólo en un huésped (Fig. 23), que generalmente es el humano. La ingestión de quistes maduros infecciosos (que poseen 4 núcleos), en comida, agua o manos contaminadas, va seguida de su excitación sólo cuando alcanzan la parte terminal del intestino delgado (íleo). El jugo gástrico no tiene efecto y no puede disolver la pared del quiste.

La exquistación es el proceso de transformación de los quistes en trofozoítos mediante la disrupción de la pared del quiste, a través de la acción de la enzima tripsina, lo que permite que emerja la ameba cuadri-nucleada e inicie la proliferación. La ameba sufre una serie de divisiones nucleares seguidas de 3 rondas sucesivas de citocinesis (división celular) para producir 8 pequeños trofozoítos uninucleados, llamados amebulas. Bajo condiciones hostiles, también podrían producir quistes; una vez que se agregan a la capa de mucina del intestino acumulan una cantidad considerable de alimentos en forma de glucógeno, gránulos de cromátides en forma de barras negras y forman una pared de quiste delgada, redondeada y resistente. Por lo tanto, la enquistación ocurre en la luz del intestino, y la formación del quiste se completa cuando cuatro núcleos están presentes. Las señales que conducen a enquistación o exquistación son poco conocidas, pero podrían ser desencadenadas por una lectina de unión a una galactosa en la superficie del parásito. Después de eso, ambas etapas pasan en las heces: los trofozoítos generalmente están presentes en las heces blandas, mientras que los quistes se encuentran en las heces firmes. Debido a la protección conferida por sus paredes, los quistes pueden resistir las condiciones ambientales durante días o semanas y pueden ser responsables de la transmisión de la enfermedad. Sin embargo, una vez fuera del cuerpo, los trofozoítos que pasan en las heces se dañan rápidamente y, si se ingieren, no sobreviven a la exposición al entorno gástrico. Lo anterior describe la colonización típica asintomática que representa casi el 80 al 90% de todas las infecciones (Petri, 2017).

4.4 Secuencias Teloméricas de *Entamoeba histolytica*

En *E. histolytica*, no ha sido posible identificar las secuencias de los extremos terminales de los cromosomas ni se han identificado secuencias teloméricas canónicas ni ortólogos de genes de la telomerasa (Clark et. al. 2007). Sin embargo, el 10% del genoma de *E. histolytica* corresponde a genes de tRNA que están asociados con repeticiones en tándem (Loftus et. al. 2010). Hay 25 tipos diferentes de arreglos en tándem largos que contienen entre 1 y 5 tipos de tRNA por unidad de repetición y STR que se asemejan a microsatélites (Tabla 5). Se ha propuesto que estas matrices pudieran localizarse en los extremos del cromosoma, actuando como regiones teloméricas que cumplen un papel estructural en el genoma. Además, los cromosomas de *E. histolytica* no se condensan completamente y existe una variación considerable en el tamaño de sus cromosomas debido a la extracción y contracción de las repeticiones teloméricas, como en otros protistas como en *D. discoideum* (Clark et. al 2006).

Tabla 5. Características de las unidades STR de tRNA

Nombre del arreglo	Tipo de isoceptor	Longitud (pb)	5S RNA?	Numero de acceso 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]	eMet ^{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 ^{ACG}	846	Si	BK005651
[RT]	Arg ^{CCT} , Thr ^{AGT}	964	X	BK005652
[SD]	Ser ^{TGA} , Asp ^{GTC}	767	X	BK005657
[SPPCK]	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} , iMet ^{CAT} , Glu ^{CTC}	1466	Si	BK005672

[W1]	Trp ^{CCA} , Ile ^{AAT}	1135	X	BK005665
[YE]	Tyr ^{GTA} , Glu ^{TTC}	938	X	BK005661

Los tipos de isoceptor de tRNA se administran con su anticodón. 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).

Casi todos los genes que codifican para tRNA son adyacentes a STRs, lo que indica que su asociación con STR es un fenómeno general, y a su vez las unidades están organizadas en matrices en tándem que comprenden múltiples copias idénticas o casi idénticas. Fuera de las regiones de codificación de tRNA, las secuencias son extremadamente ricas en A+T en contraste con las secuencias tRNA ricas en G+C. Tres de las secuencias ricas en G+C son idénticas y se identificaron como genes que codifican para RNA 5S previamente desconocidos en *E. histolytica* (Fig. 24). La cuarta secuencia rica en G+C en la matriz, no se ha identificado, pero se sospecha que puede ser un pequeño gen que codifica RNA nuclear. Las longitudes de las unidades de matriz varían entre 490 - 1775 pb, y entre los genes adyacentes de tRNA o 5S RNA hay una región STR en casi todos los casos las cuales tienen cierta homología con la secuencia canónica humana (TAAGGG) (Clark et. al., 2006).

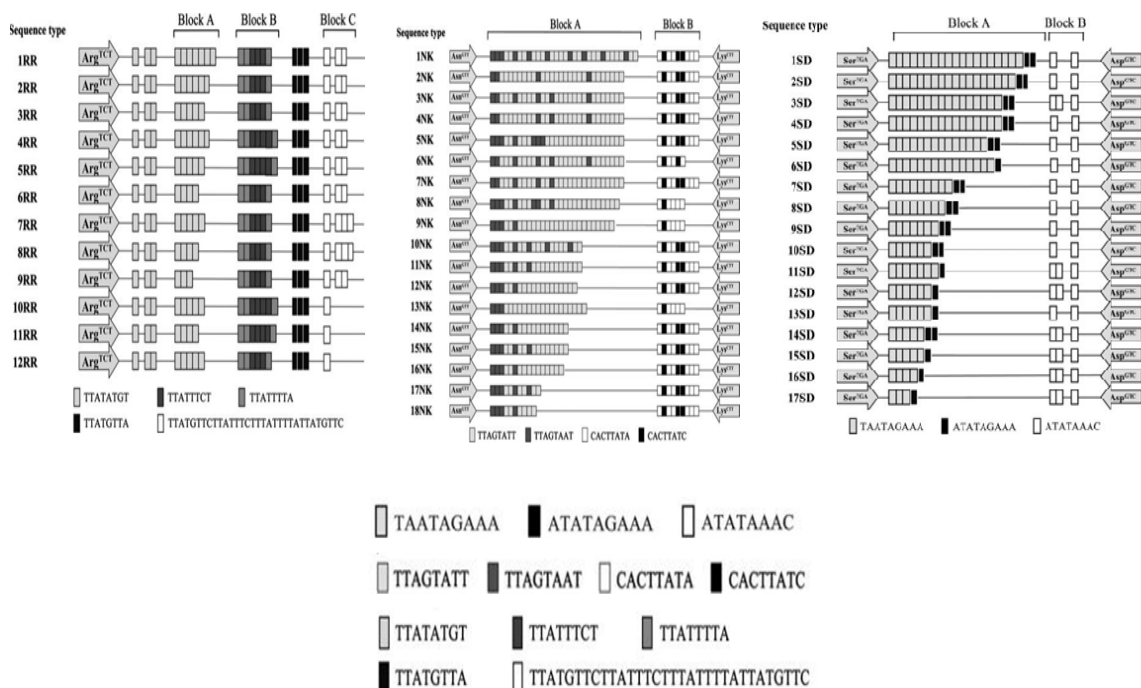


Fig. 24. Representación esquemática de la organización de las unidades de los arreglos de tRNA. Se muestran unidades representativas. La orientación del gen está indicada por la dirección de la flecha. Los STR se indican mediante cuadros de diferente color, cada color en cada matriz indica una secuencia STR distinta (Tomado y modificado de Clark, 2006).

4.5 Proteínas TRFs en *Entamoeba histolytica*

Hasta ahora, no se han descrito secuencias teloméricas o complejos de proteína implicados en la función de los telómeros de este parásito. El estudio de los homólogos de TRF ayudará a comprender mejor la biología de los telómeros de *E. histolytica*. Por lo tanto, en nuestro grupo de trabajo se han estudiado proteínas de *E. histolytica* que tienen el dominio Myb de unión a DNA (Myb-DNA), estas proteínas se agrupan en las familias I, II y III y están relacionadas a factores de transcripción. Además, se identificó, que los miembros que componen la familia II denominados EhTRF-like tienen alta identidad de secuencia con proteínas TRF de *H. sapiens*.

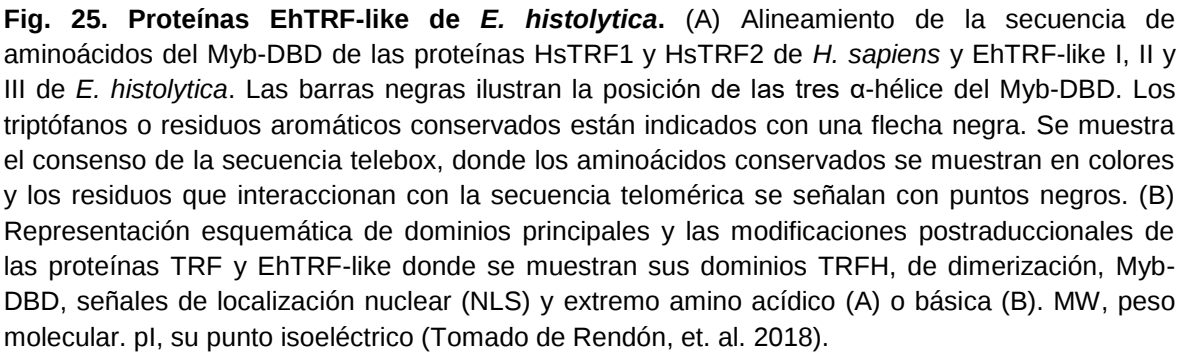
A diferencia de *H. sapiens*, el cual tiene un solo gen que codifica TRF1 y otro a TRF2, *E. histolytica* tiene tres genes codificantes para proteínas TRF, los cuales están ubicados en los *loci*: EH1_001110, EHI_001090 y EHI_148140. En *E. histolytica* los genes EHI_001110 y EHI_001090 están localizados dentro del mismo contig, lo cual sugiere que ocurrió una duplicación génica en el genoma de este parásito. A las proteínas codificadas por estos genes se le denominó EhTRF-like I, II, III y aunque en la base de datos aparecen como proteínas hipotéticas, debido a su alta identidad con proteínas de unión a telómeros, su posible función podría estar relacionada al reconocimiento de las secuencias teloméricas.

Tabla 6. Comparación de las proteínas EhTRF-like de *E. histolytica* con TRF1 y TRF2

Locus	Anotación	Proteína H. sapiens	Acceso	E-value	Pid
EHI_148140	HP	TRF1	P54274	4.8e-7	29.7
EHI_001090	HP	TRF1	P54274	4.8e-7	29.7
EHI_001110	HP	TRF2	P54274	9.9e-7	25.8

(Tomado y modificado de Rendón, et. al. 2018).

Las proteínas EhTRF-like además, tienen conservado el motivo Telebox que contiene los aminoácidos VDLKDKWRT en la tercera hélice del dominio DBD MYB, involucrados en los contactos proteína-DNA de los factores de unión a repetidos teloméricos de eucariontes superiores. EhTRF-like I y II tienen un extremo amino terminal ácido, similar al extremo rico en ácido aspártico y ácido glutámico de TRF1 con un pI de 5.8 y 7.14 respectivamente. Contrastando con el



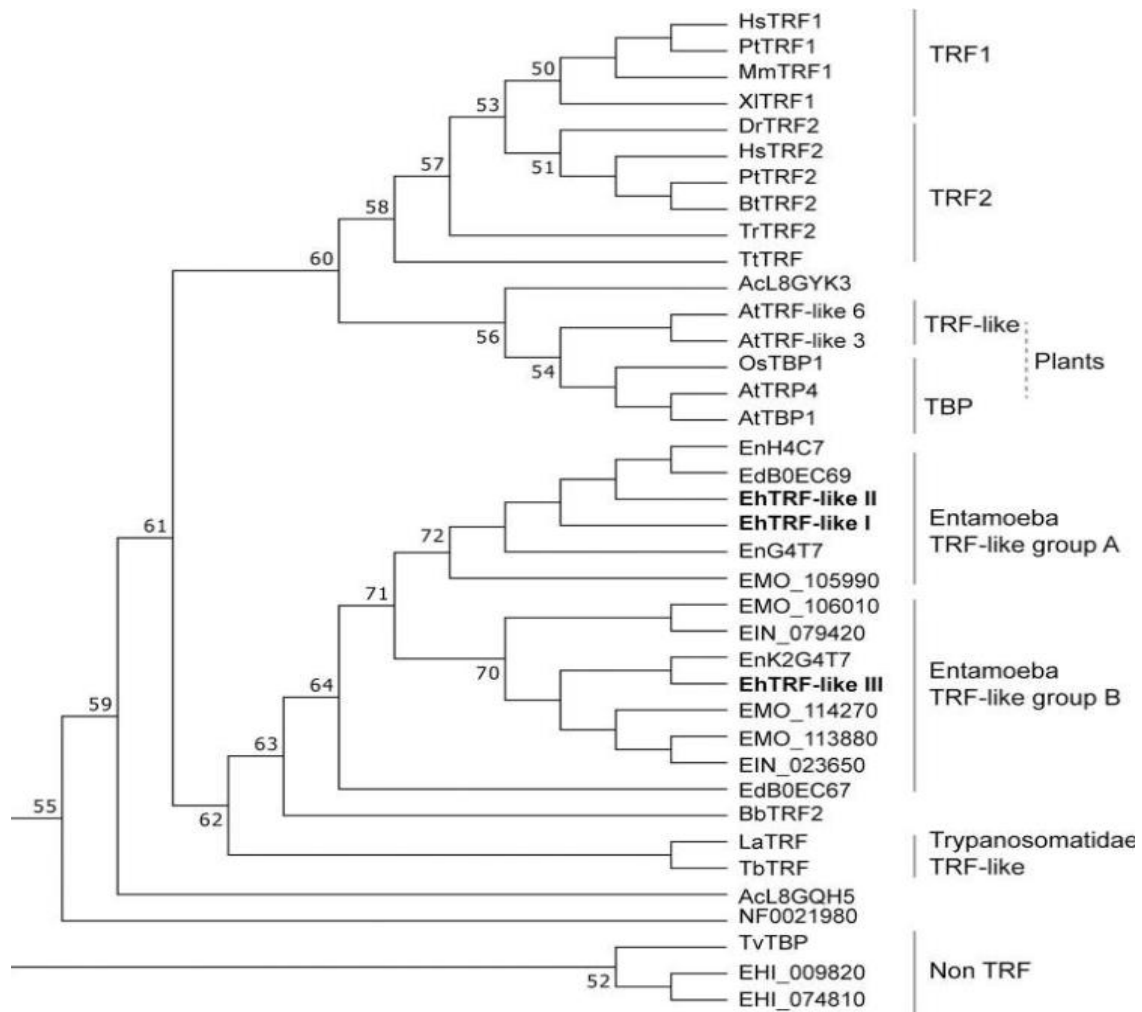


Fig. 26. Árbol filogenético de proteínas EhTRF-like. Las secuencias de aminoácidos del DBD MYB fueron alineadas usando ClustalW y posteriormente se infirió el análisis filogenético usando el método de Neighbor-Joining. Las distancias evolutivas se calcularon mediante el método de corrección de Poisson. Los valores de bootstrap >50% (de 1,000 repeticiones) se muestran cerca de las ramas individuales. Los análisis evolutivos se realizaron con MEGA 7. HsTRF1 (*Homo sapiens*, NP_059523.2), MmTRF1 (*Mus musculus*, NP_033378.1), PtTRF1 (*Pan troglodytes*, XP_016815057.1), XITRF1 (*Xenopus laevis*, DQ146827.1), HsTRF2 (*H. sapiens*, NP_005643.2), PtTRF2 (*P. troglodytes*, XP_016785565.1), BtTRF2 (*Bos taurus*, NP_001341392.1), BbTRF2 (*Branchiostoma belcheri*, XP_019634930.1), DrTRF2 (*Danio rerio*, BC096857.1), TrTRF2 (*Thecamonas trahens*, XP_013759355.1), TtTRF (*Takifugu rubripes*, XP_011616883.1), AtTRF-like 3 (*A. thaliana*, NP_001185023.1), AtTRF-like 6 (*A. thaliana*, NP_974133.1), AtTBP1 (*A. thaliana*, XP_015625433.1), AtTRP4 (*A. thaliana*, NP_196886.1), OsTBP1 (*Oryza sativa*, XP_015625433.1), EhTRF-like I (*E. histolytica*, XP_649878.1), EhTRF-like II (*E. histolytica*, XP_649880.1), EhTRF-like III (*E. histolytica*, XP_657528.1), EdBOEC67 (*Entamoeba dispar*, XP_001735901.1), EdB0EC69 (*E. dispar*, XP_001735903.1), Ein079420 (*E. invadens*, XP_004184366.1), Ein023650 (*E. invadens*, XP_004257447.1), EnH4C7 (*E. nuttalli*, XP_008857362.1), EnK2H4C7 (*E. nuttalli*, XP_008860337.1), EnK2G4T7 (*E. nuttalli*, XP_008860339.1), NF0021980 (*Naegleria fowleri*), LaTRF2 (*Leishmania amazonensis*, ABU53006.1), TbTRF (*Trypanosoma brucei*, AAX86992.1), TvTBP (*Trichomonas vaginalis*, XP_001314728.1), EHI_009820 (*E. histolytica*, XP_001913661.1), and EHI_074810 (*E. histolytica*, XP_649576.1) (Tomado de Rendón, et. al. 2018).

Mediante microscopía confocal también se ha determinado que las proteínas EhTRF-like I y III co-localizan en el núcleo cercanas a la periferia nuclear y en algunos *foci*. De manera interesante, estas proteínas ocupan posiciones similares a las de la proteína lámina B1 y las de la histona H4K20me3 las cuales son marcadores de periferia nuclear y de regiones teloméricas respectivamente. También, por medio de microscopía electrónica de transmisión (TEM) han encontrado a EhTRF-like III en regiones de cromatina condensada, lo que podría corresponder a regiones de heterocromatina (Fig. 27). Estos hallazgos han sugerido que las proteínas EhTRF-like de *E. histolytica* pueden desempeñar funciones similares a las proteínas TRF de *H. sapiens*.

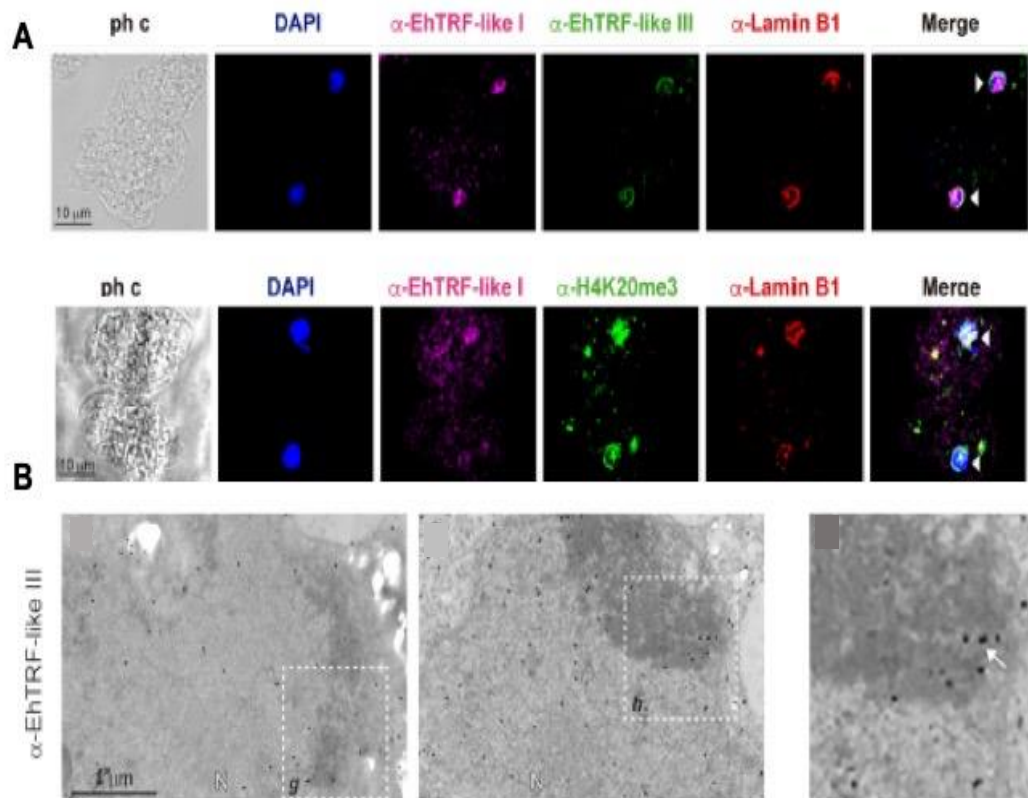


Fig. 27. Localización de las proteínas EhTRF-like en los trofozoítos de *E. histolytica*. (A) Localización de las proteínas EhTRF-like I y III. Los trofozoítos se procesaron para inmunofluorescencia directa y se incubaron con los anticuerpos acoplados a Alexa α-EhTRF-like I (Alexa-647, magenta), α-EhTRF-like III (Alexa-488, verde), α-Lamin B1 (Alexa-555, rojo) y α-H4K20me3 (Alexa-488, verde). Los núcleos se tiñeron con DAPI y las preparaciones se analizaron por microscopía confocal. Las puntas de flecha blancas indican co-localización de las proteínas. (B) Trofozoítos que sobre expresan a la proteína EhTRF-like III se incubaron en presencia del anticuerpo α-EhTRF-like III seguido de un anticuerpo secundario acoplado a una partícula de oro y visualizados por TEM (Tomado de Rendón, et. al. 2018).

Así mismo, se analizó si los extractos de proteínas nucleares de *E. histolytica* contenían proteínas que pudieran reconocer secuencias canónicas teloméricas. Por lo que se realizaron ensayos tipo EMSA empleando extractos nucleares obtenidos de trofozoítos de *E. histolytica* y una secuencia telomérica canónica humana (HsTEL-TTAGGG) obteniéndose tres complejos DNA-proteína (Fig. 28). Presumiblemente, la formación de estos complejos se debe a la unión de las tres proteínas similares a TRF con la sonda HsTEL.

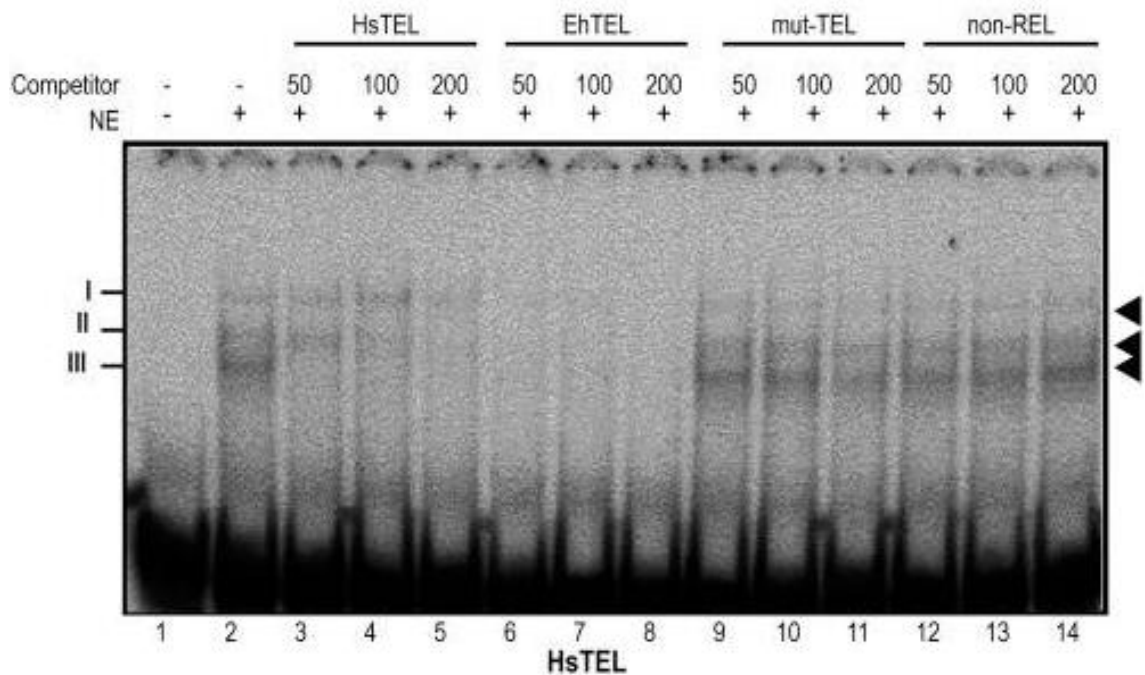


Fig. 28. Formación de complejos DNA-proteína con la sonda HsTEL y extractos nucleares de *E. histolytica*. Se realizó el ensayo EMSA utilizando DNA telomérico canónico humano de doble cadena radiomarcado (HsTEL) como sonda y extractos nucleares obtenidos de trofozoítos de *E. histolytica*. Los ensayos de competencia se realizaron en presencia de un exceso de 50, 100 y 200 veces de secuencias no marcadas: competidor específico de HsTEL, secuencia de *E. histolytica* (EhTEL), secuencia telomérica mutada (mut-TEL) o secuencia no relacionada (Non-Rel) (Tomado de Rendón, et. al. 2018).

Para investigar si las proteínas EhTRF podían unirse a DNA telomérico bicatenario, se purificó la proteína recombinante EhTRF-like I y se utilizó una secuencia (EhTEL-TTAGTATT) derivada de los STR de la subunidad NK2 de los genes de tRNA elegida por ser la que tiene mayor identidad de secuencia con el repetido canónico humano (TTAGGG). Como se observa en la Fig. 29 rEhTRF-like I se unió específicamente a las sondas HsTEL y EhTEL. Los ensayos de

competencia mostraron que el complejo formado por la proteína recombinante desapareció en presencia de sondas en excesos molares de HsTEL y EhTEL sin marcar. No hubo competencia por la unión cuando la sonda mut-TEL se usó con el mismo exceso molar; lo cual corrobora que la proteína EhTRF-like I es capaz de unirse a secuencias teloméricas canónicas.

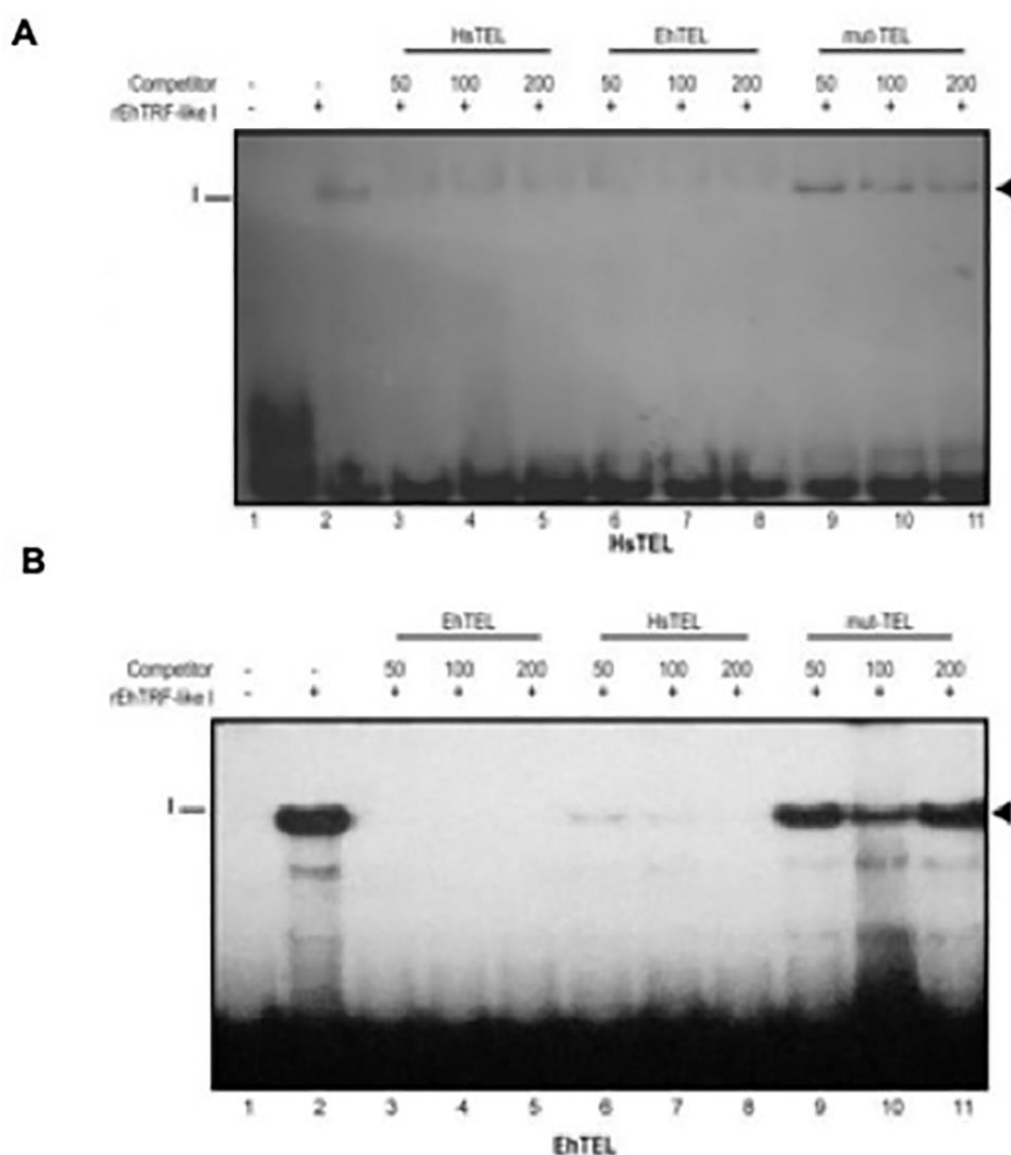


Fig. 29. Formación de complejos DNA-proteína con rEhTRF-like I. Ensayos EMSA empleando rEhTRF-like I y las sondas HsTel (A) y EhTel (B) y ensayo de competencia en presencia de excesos molares de 50, 100 y 200 veces de secuencias no marcadas (HsTel, EhTel y Mut-Tel).

V. Justificación

La regulación y el mantenimiento de los telómeros mediado por el complejo Shelterin parece ser compleja y los detalles moleculares de este interesante aspecto de la biología de los telómeros no han sido completamente elucidados en el parásito *E. histolytica*.

En el genoma de *Entamoeba histolytica* no hay registros de secuencias teloméricas clásicas por lo cual es importante determinar la afinidad por las secuencias STR y obtener un posible interactoma de proteínas asociadas a la proteína EhTRF-like I que nos proporcione más información sobre cómo está organizada la maquinaria telomérica de este parásito.

VI. Hipótesis

Los dominios TRFH y de unión a DNA MYB de las proteínas EhTRF-like de *E. histolytica* presentan una estructura tridimensional e interacciones similares a las de sus homólogos humanos.

VII. Objetivos

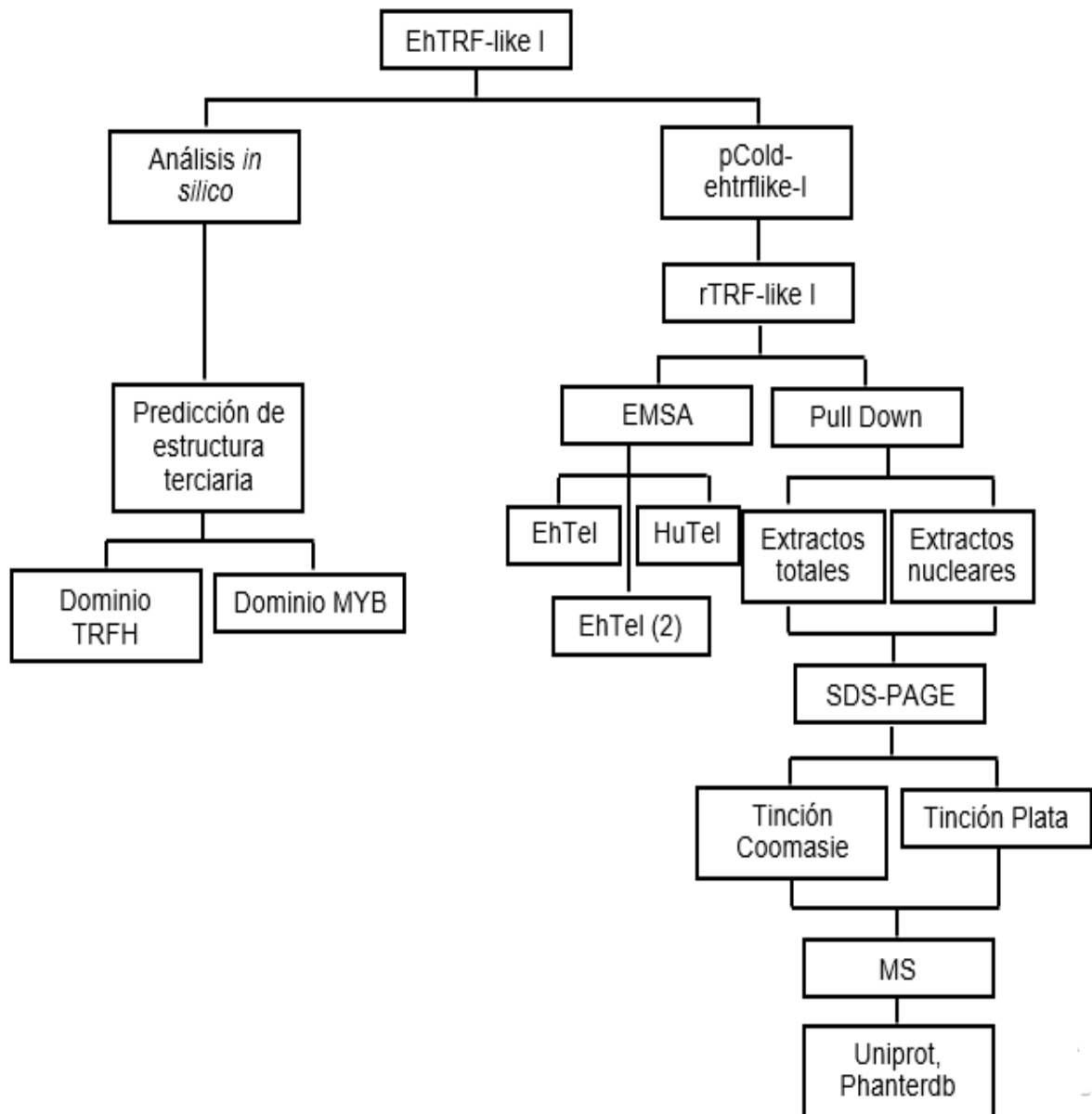
7.1 Objetivo General

Caracterizar la estructura tridimensional de la proteína EhTRF-like I de *E. histolytica* y sus interacciones con DNA y proteína.

7.2 Objetivos Particulares

1. Determinar la estructura *in silico* de los dominios TRFH y el dominio MYB de unión a DNA de la proteína EhTRF-like I.
2. Determinar la K_d de la proteína rEhTRF-like I con secuencias teloméricas canónicas y no canónicas.
3. Identificar proteínas que interaccionen con la proteína rEhTRF-like I.

VIII. Estrategia Experimental



IX. Metodología

9.1 Análisis Bioinformático

La predicción de la estructura secundaria para cada uno de los dominios de dimerización TRFH y de unión a DNA MYB se realizó empleando los servidores PFAM (<https://pfam.xfam.org/>) y PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>). La predicción de estructura terciaria se realizó mediante modelado libre (*ab initio*) empleando el servidor ROBETTA (<http://rosetta.bakerlab.org/index.html>). Robetta es un servidor de predicción de estructura de proteínas de cadena completa, analiza las cadenas de proteínas en dominios putativos con el protocolo de Ginzu y modela esos dominios por el método de *ab initio*. Para visualizar y analizar el modelado de las estructuras se utilizó el programa UCFS Chimera (<https://www.cgl.ucsf.edu/chimera/>).

9.2 Obtención y Purificación EhTRF-like I Recombinante

El gen de 1210 pb que codifica para la proteína EhTRF-like I se clonó en el vector pCold I (Takara) utilizando los oligonucleótidos sentido 5'-CGCGGATCCTTATTGAGAAAGATCCAATTG-3' y el antisentido 5'-GGAATTCCATATGAATAACCCTCAGTTTGC-3' los cuales contienen los sitios de restricción *Bam*HI y *Nde*I respectivamente para generar la construcción *pCold-ehtrf-like I* de acuerdo a lo publicado en Rendón-Gandarilla et al., 2018

Para la producción de la proteína rEhTRF-like I se empleó la cepa de *Escherichia coli* Rosetta (DE3) (Novagen), las cuales se transformaron con el plásmido *pCold-ehtrf-like I*. Posteriormente, una colonia se usó para inocular 25 mL de medio LB suplementado con 100 µg/mL de ampicilina y se dejó crecer durante toda la noche (ON). Con este cultivo se inoculó un matraz de 1L y la expresión de la proteína rEhTRF-like I se indujo con 1mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) a 16° C por 16 h. Posteriormente, las células se cosecharon y lisaron en un buffer de lisis (20 mM Na₂HPO₄, 500 mM NaCl, 8 mg lisozima y 4 µL de 100 mM PMSF) y se incubaron por 30 minutos, seguido de un ciclo de sonicación (30 segundos, 30 pulsos, 80% amplitud). El lisado se centrifugó a 14,000 rpm durante 30 minutos a 4°C. La fracción soluble se filtró con un filtro estéril de 0.45 µm y se procedió a

purificar por el método de *batch* incubándose con la agarosa HIS-Select® Nickel Affinity Gel (Sigma-Aldrich) por 1 hora a 4° C. Posteriormente, se centrifugó a 5000 rpm por 1 minuto. El primer lavado consistió en 25 ml de tampón de lisis y el segundo lavado se realizó con 25 mL de tampón de lisis suplementado con 20 mM imidazol. La proteína se eluyó con 2 ml de 50 mM Tris-HCl, 100 mM NaCl, 500 mM imidazol y se dializó durante la noche en 1 L de un tampón que contenía 25 mM Tris-HCl pH 7,5 y 300 mM NaCl. También la proteína se purificó por medio de cromatografía líquida de proteínas de alta velocidad (FPLC) en el equipo ÄKTAprime plus (GE Healthcare Life Sciences) y una columna Ni Sepharose High Performance (GE Healthcare) empleando el programa His Tag Purification HISTRAP. Para evaluar la inducción y purificación de la proteína rEhTRF-like I, los extractos y la proteína dializada se analizaron por medio de SDS-PAGE e inmunodetección empleando un anticuerpo anti-histidina (α -His) (Santa Cruz Biotechnology, Inc) en una dilución de 1:1000 y un anticuerpo secundario anti-ratón (Invitrogen-Gibco) diluido 1:3000. Para la purificación de la proteína recombinante, se obtuvo la fracción soluble utilizando cromatografía de afinidad con His-Select Nickel Affinity Gel, siguiendo las indicaciones del fabricante. La proteína rEhTRF-like I se dializó en un buffer con 25mM Tris-HCL pH 7.4 y 300 mM NaCl. Finalmente, se cuantificó por el método de Bradford y se empleó para los ensayos EMSA y Pull down.

9.3 Ensayos de Inmunodetección (Western Blot)

La proteína inducida y purificada rEhTRF-like I fue analizada por electroforesis en geles SDS-PAGE al 10% teñidas con Azul de Coomassie R-250. A continuación, las proteínas fueron transferidas a membranas de nitrocelulosa en una cámara de transferencia Trans-Blot turbo (BIO-RAD) por 30 min a voltaje constante 25V. Posteriormente, fueron bloqueadas con leche descremada al 5% en PBS 1X durante toda la noche. Para realizar la inmunodetección se agregó el anticuerpo primario α -His en una dilución 1:5000 al 1% en leche descremada en PBS 1X y se incubaron en agitación ON a 4°C. Al terminar la incubación, se lavó 4 veces la membrana con PBS-Tween 20 1X cada 15 minutos por 1 hora a temperatura ambiente. Enseguida, se agregó el anticuerpo secundario (anti-IgG de ratón) en una dilución 1:3000 en leche descremada al 1.5% en PBS 1X y se incubó en

agitación por 1 hora a 37°C. Finalmente, se realizaron 4 lavados cada 15 minutos por 1 hora con PBS-Tween 20 1X y se visualizó con reactivo Merck™ Sustrato HRP Western Luminata Forte (Merck-Millipore) en el equipo ChimeDoc MP (BIO-RAD).

9.4 Hibridación de Oligonucleótidos para el EMSA

Para la realización de la hibridación se tomó 1 µg de cada uno de los oligonucleótidos (HsTel, EhTel, EhTel (2), NonTel) sentido y antisentido (Tabla 7) y se colocaron en un tubo eppendorf con buffer TM (50 mM Tris-HCL, 2.5 mM MgCl₂ pH 7.0). Los tubos se calentaron y mantuvieron a 85 °C durante 5 minutos y se dejaron enfriar hasta alcanzar la temperatura ambiente. Posteriormente, fueron guardados a -20°C hasta su uso.

Tabla 7. Sondas utilizadas para los ensayos EMSA

Sonda	Secuencia.
HsTel	S: 5´ GCTCTAGATG TTAGGG TTAGGGTTAGGGTTAGGG 3´ AS: 5´ AATCCCAATCCCAATCCCAATCCCGAGATCTAC 3´
EhTel	S: 5´ GCTCTAGATG TTAGTATTT AGTATTTTAGTATTTTAGTATT 3´ AS: 5´ AATCATAAAATCATAAAATCATAAAATCATAACGAGATCTAC 3´
EhTel (2)	S: 5´ GCTCTAGAT TTATATG TTTATATGTTTATATGTTTATATG 3´ AS: 5´ GCAGATCTACATATAACCATATAAACATATAAACATATAA 3´
NonTel	S: 5´ GCTCTAGATGTGGGCTAGCTAGCTCTG 3´ AS: 5´ CATCTAGAGCCCTGAGTCAGAGCTAG 3´

9.5 Ensayos de Retardamiento de Movilidad Electroforética (EMSA)

Los ensayos de unión a DNA se realizaron empleando los siguientes oligonucleótidos: HsTEL (TTAGGGx4), el cual contiene la secuencia canónica humana, EhTEL (secuencia derivada del STR (TTAGTATTx4) de la unidad NK2 del arreglo de tRNA arrays de *E. histolytica*), EhTel (2) (secuencia derivada del STR (TTATATGTx4) de la unidad [RTCT] del arreglo de tRNA arrays de *E. histolytica*) y MutTel (secuencia (TGGGTAGx4) no relacionada a una secuencia telomérica). Para ello los oligonucleótidos se hibridaron y se marcaron con 0.042 mCi [γ -³²P] dATP (Perkin Elmer), usando la enzima T4 Polynucleotide Kinase (Thermo Fisher Scientific) para posteriormente purificarlos con el kit QIAquick nucleotide removal kit (Qiagen) de acuerdo con las instrucciones del fabricante y finalmente fueron cuantificados con un equipo espectrómetro de centelleo liquido.

Concentraciones crecientes de la proteína recombinante rEhTRF-like I (5.02, 10.04, 15.06, 20.08, 25.1, 30.12, 35.14, 40.16, 45.18 x10³ nM) se incubaron por 25 min a 4°C con 0.05 ng/μL de los oligonucleótidos marcados en buffer de binding Bdg 2X (12 mM HEPES pH 7.8, 60 mM KCL, 5mM MgCl, 20% Glicerol y un coctel de inhibidores de proteasas), solución de SpMg (4mM espermidina, 4mM MgCl₂). Las muestras se depositaron en un gel nativo no desnaturalizante de poliacrilamida al 6% y de baja fuerza iónica con 0.5% TBE. La electroforesis se llevó a cabo a 85 V y el gel fue secado al vacío a 80 °C por 30 minutos; posteriormente, el gel se expuso en un cassette para autoradiografía Storage phosphor screen (GE-Healthcare) toda la noche y los complejos proteína-DNA se evidenciaron a través de un detector Storm (GE-Healthcare) o por autoradiografía utilizando películas Kodak X-omat film (Sigma) y exponiendo a - 70°C para su posterior revelado. La constante de disociación (K_d) se estimó a partir de los ensayos, cuantificando la cantidad de oligonucleótido libre y del complejo DNA-proteína con el programa Quantity One 1-D Analysis Software (BIO-RAD), determinando la concentración en la que el 50% de la sonda está unida a la proteína a través del análisis de datos empleando el programa Prism - GraphPad y la herramienta Equation: One site - Specific binding.

9.6 Cultivo de trofozoítos de *E. histolytica*

Los trofozoítos de la Cepa HM1: IMSS clona A se cultivaron axénicamente a 37° C en tubos de medio TYI-S-33 (Diamond et al. 1978) suplementados con 15% de suero bovino (Microlab) al 15% (inactivado por calor), suplementado con 3% de vitaminas de Diamond (Sigma- Aldrich), mezcla de penicilina y estreptomicina 100x (Sigma-Aldrich). Para realizar los experimentos, se usaron los trofozoítos en fase de crecimiento logarítmico, los cuales fueron incubados en un baño de hielo durante 10 minutos y se cosecharon por centrifugación a 1500 rpm durante 5 minutos a 4° C.

9.7 Obtención de extractos totales de trofozoítos de *E. histolytica*

Para la obtención de proteínas totales después de cosechar los trofozoítos de *E. histolytica*, se lavaron en PBS 1x pH 6.8 a 1500 rpm por 5 minutos. Posteriormente, se realizó conteo celular en una cámara Neubauer y se hicieron alícuotas de 3×10^6 células en tubos eppendorf a los cuales se les adicionarán 100 μ L de TRIS-HCL 100 mM (Sigma- Aldrich), 1 μ L de PMSF 200 mM (Sigma- Aldrich) y 5 μ L de 50x CompleteTM Protease inhibidor Cocktail (Sigma-Aldrich) y posteriormente se realizaron tres ciclos de congelamiento–descongelamiento para luego ser centrifugados a 1500 rpm durante 5 minutos a 4° C y recuperar los sobrenadantes. Los extractos obtenidos de esta manera se cuantificaron por el método de Bradford empleando el reactivo Bio-Rad protein Assay (Bio-Rad).

9.8 Obtención de extractos nucleares de trofozoítos de *E. histolytica*

Para obtener proteínas nucleares de *E. histolytica*, 1×10^6 trofozoítos se cosecharon por centrifugación a 1500 rpm durante 5 minutos a 4° C. La pastilla se resuspendió en 1ml de PBS 1X pH 6.8 para lavarla. Posteriormente, las células se resuspendieron en 400 μ L de buffer A (Hepes 1M pH 7.9 KCL 1M, EDTA 0.5M, EGTA 100mM, DTT 1M, PMSF 200mM Complete 50x) por cada 1×10^6 trofozoítos mezclando suavemente. Luego de 5 a 10 min, se añadió 25 μ L de NP40 al 10% por cada 1×10^6 trofozoítos. Éstas se agitaron vigorosamente en el vórtex durante 30 segundos y se centrifugaron a 14000 rpm durante 10 minutos. El sobrenadante se transfirió a nuevos tubos. Se resuspendió la pastilla en el mismo volumen de buffer A y se colocó sobre un colchón de sacarosa 0.3M (1mL) y se centrifugo por 20 minutos a 9,000 rpm. La pastilla de núcleos obtenida se rompió realizando 3 ciclos de congelamiento – descongelamiento. Posteriormente, fueron centrifugados durante 15 minutos a 14,000 rpm y se hicieron alícuotas con el sobrenadante en tubos eppendorf, se cuantificaron por el método de Bradford y se almacenaron a -70° C hasta su uso.

9.9 Cuantificación de proteínas

La cuantificación de proteínas se realizó utilizando el kit Protein Assay II (Bio-Rad), el cual está basado en el método de Bradford. El reactivo se preparó una dilución 1:4 con agua destilada y enseguida se filtró en papel Whatman #1 para remover partículas precipitadas. Se preparó también una solución concentrada de BSA a una concentración de 10 mg/ μ L en agua desionizada a partir de la cual se preparó una curva patrón donde se adicionaron 200 μ L de la solución preparada de Bradford a cada pozo considerado en la curva. Se adicionaron también 10 μ L de muestra según corresponda y se incubó la placa en oscuridad a temperatura ambiente por 5 min. Posteriormente, se leyó la absorbancia a 595 nm en un lector de placas ELISA.

9.10 Ensayo Pull Down

El ensayo de pull down es un método *in vitro* utilizado para determinar una interacción física entre dos o más proteínas. En un ensayo de pull down una proteína "cebo" se marca y captura con un ligando de afinidad inmovilizado específico para la etiqueta, generando de ese modo un "soporte de afinidad secundario" para purificar otras proteínas que interactúan con la proteína cebo. Se utilizó la proteína rEhTRF-like I que contiene la etiqueta de fusión de hexahistidina (6xHis) previamente purificada. Los extractos nucleares se preclarificaron mediante incubación con la resina de Ni^{2+} en buffer de unión (25mM Tris-HCl, 500 mM NaCl y 20 mM Imidazol, pH 8.0) durante una hora para eliminar las proteínas que interaccionan de manera inespecífica. La resina de Ni^{2+} (50 μ L) se lavó con 1-2 volúmenes de agua para eliminar el etanol del medio, enseguida fue equilibrada con buffer de equilibrio (50mM fosfato de sodio, 300mM NaCl y 20mM Imidazol, pH 8.0) y a continuación se añadió e inmovilizó a la proteína rEhTRF-like I (6.2 μ g/ μ L) mediante incubación por una hora en agitación suave. Una vez transcurrido este tiempo se hicieron lavados con 1-5 volúmenes de buffer de lavado (50mM fosfato de sodio, 300mM NaCl y 50mM Imidazol, pH 8.0) y se agregaron los extractos nucleares, posteriormente se incubó por una hora. Una vez transcurrido el tiempo se realizaron lavados con buffer de lavado (50mM fosfato de sodio, 300mM NaCl y 50mM Imidazol, pH 8.0) y se llevó a cabo la

elución de la interacción proteica con buffer de elución (50mM fosfato de sodio, 300mM NaCl y 300 mM Imidazol, pH 8.0). Los complejos de proteínas resultantes fueron recuperados y analizados mediante SDS-PAGE, teñidos con Azul de Coomassie G-250 o mediante tinción con plata compatible con espectrometría de masas. Para la identificación de las proteínas las proteínas eluidas se separaron en un gel de SDS-PAGE 10% y se tiñeron mediante tinción con plata compatible con espectrometría de masas. Los geles se enviaron a la Unidad de Genómica, Proteómica y Metabolómica del CINVESTAV-Zacatenco y el análisis de los péptidos se realizó por medio MALDI-TOF/TOF

9.11 Tinción de Plata

Para la tinción con plata se utilizó el Kit Silver Stain Plus (Bio-Rad). El gel se incubó por 20 minutos en solución fijadora (50% metanol, 10% ácido acético, 10% fixative enhancer concentrado y 30% agua Milli-Q) y enseguida se realizaron dos lavados de 10 min con agua Milli-Q. Posteriormente, se sometió a tinción con la solución de tinte por 20 min en agitación suave, (5 mL silver complex solution, 5 mL reduction moderator solution, 5 mL image development y 50 mL development accelerator solution y 35 mL de agua Milli-Q). Una vez que se observó el patrón de bandas se adicionaron 250 mL de solución de paro (5% de ácido acético) y se incubó por 15 minutos para detener la reacción.

9.12 Análisis *in silico* del Interactoma

Una vez obtenido el resultado del análisis por espectrometría de masas, se realizó una búsqueda *in silico* a partir de las huellas digitales peptídicas obtenidas para identificar proteínas en la base de datos MASCOT y se compararon con la base de datos del genoma de *Entamoeba histolytica*. Posteriormente se analizó la función celular en la base de datos GO Pantherdb (<http://www.pantherdb.org>).

X. Resultados

Simulación *in silico* de la estructura tridimensional de los dominios TRFH y DBD MYB de la proteína EhTRF-like I

La proteína EhTRF-like I con número de acceso EHI_001090 (de acuerdo con AmoebaDB), tiene una alta homología con HsTRF1 de *H. sapiens*. Esta proteína conserva los dominios de unión a DNA MYB (MDBD) y el dominio de dimerización TRFH; sin embargo, hasta ahora no se conoce su estructura tridimensional. Es por ello que primero para el análisis de dominios MDBD y TRFH, se obtuvo una predicción de su estructura secundaria con el fin de localizar y delimitar los aminoácidos pertenecientes a cada uno de los dominios. Para el dominio TRFH se empleó servidor PSIPRED, mientras que para el dominio MDBD el servidor PFAM. Posteriormente, se predijo la estructura tridimensional de los dominios TRFH y MDBD con la herramienta bioinformática Robetta, para obtener un modelo por el método de *ab initio* y homología respectivamente.

La predicción de la estructura terciaria por Robetta, arrojó cinco modelos tridimensionales diferentes para cada uno de los dominios predichos (MDBD y TRFH) con un determinado valor de confianza. Para seleccionar la predicción más confiable en el caso del dominio TRFH debido a que éste es muy divergente se determinó el valor de Tm-score y RSMD empleando el servidor <https://zhanglab.ccmb.med.umich.edu/TM-score>. Y finalmente se visualizaron las estructuras tridimensionales de los modelos seleccionados empleando el programa UCFS Chimera (<https://www.cgl.ucsf.edu/chimera/>).

a) Dominio DBD MYB

El análisis de la estructura terciaria del dominio MDBD de la proteína EhTRF-like I mostró un alto grado de identidad estructural y una confianza de 0.97 con el dominio MDBD de HsTRF1. Posteriormente, se realizó la superposición de los cinco modelos obtenidos para el dominio MDBD de EhTRF-like I y una vez elegido el mejor, se empalmó con la plantilla PDB del dominio MDBD de hTRF1 (datos no mostrados). El MDBD de EhTRF-like I conservó las características tres

α -hélice con una conformación hélice-giro-hélice tal y como su homólogo HsTRF1 (Fig. 30).

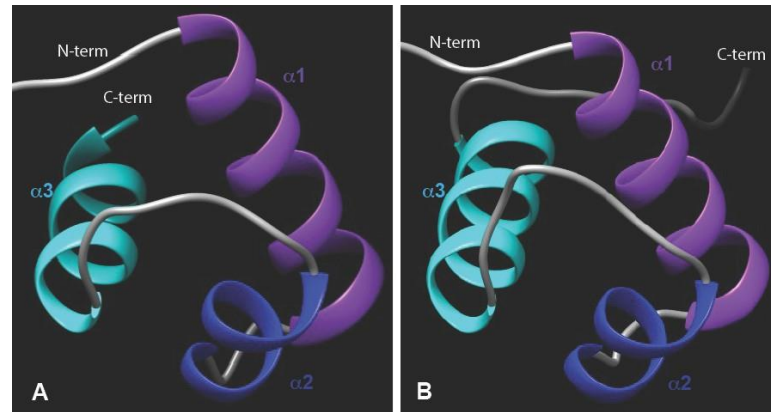


Fig. 30. Predicción in silico del dominio MBD de *E. histolytica*. A) Modelo de estructura tridimensional del dominio MBD de EhTRF-like I. B) Modelo de estructura 3D del dominio MBD de HsTRF1 con base en su plantilla PDB 1ITY.

Una vez confirmada la similitud de las estructuras tridimensionales, se analizó la secuencia de aminoácidos que conforman al dominio MBD, donde se corroboró la presencia de los residuos de aminoácidos importantes para la función biológica del dominio, como es la conservación del motivo telebox VxxKDxxR en la tercera α -hélice (Fig. 31), indispensable para el reconocimiento con el surco mayor de DNA de doble cadena a través de homeodominios y por especificidad mediante contactos directos con el grupo de guaninas (G), típico de las repeticiones teloméricas (Curt, 2005).

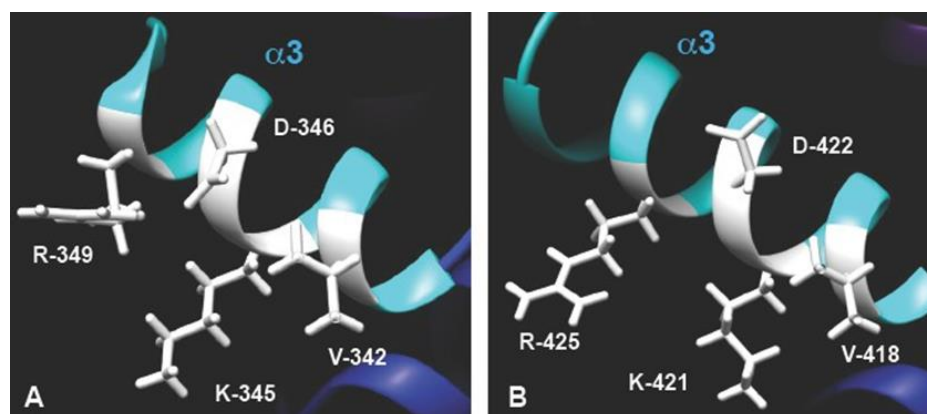


Fig. 31. Motivo Telebox VxxKDxxR conservado en la tercera α -hélice de EhTRF-like I. A) Dominio MBD de EhTRF-like I, donde se muestra la presencia de la secuencia Telebox (α -hélice 3 – Val342, Lys345, Asp346, Arg349) en posiciones similares a HsTRF1. B) Dominio MBD de HsTRF1, Donde se muestra la presencia de la secuencia Telebox (α -hélice 3–Val418, Lys421, Asp422, Arg425) en posiciones similares a HuTRF.

b) Dominio TRFH

La estructura terciaria del dominio de homodimerización TRFH fue difícil de predecir debido a que presenta una identidad de estructura secundaria del 12.5% con respecto al dominio TRFH de *H. sapiens*, por lo cual los modelos del dominio TRFH de la proteína EhTRF-like I predichos presentaron un valor de confianza de 0.21 con respecto a la estructura tridimensional del dominio TRFH de HsTRF1, lo cual denota un bajo grado de identidad estructural. Sin embargo, se seleccionó el modelo 1 que obtuvo los mejores valores de TM-score y RMSD (Tabla 8). Con este modelo se realizó una superposición de los dominios TRFH de EhTRF-like I y HsTRF1. La superposición entre ambas estructuras permitió determinar que las regiones conformadas por la α -hélice 2 (azul), la α -hélice 3 (cyan) y la α -hélice 8 (magenta), presentan orientaciones similares (Fig. 32).

Tabla 8. Valores de Tm-score y RMSD para la predicción del dominio TRFH		
Dominio TRFH	Tm-score	RMSD
Modelo 1	0.1684	2.63
Modelo 2	0.1473	3.24
Modelo 3	0.1636	3.26
Modelo 4	0.1734	3.49
Modelo 5	0.1512	2.52

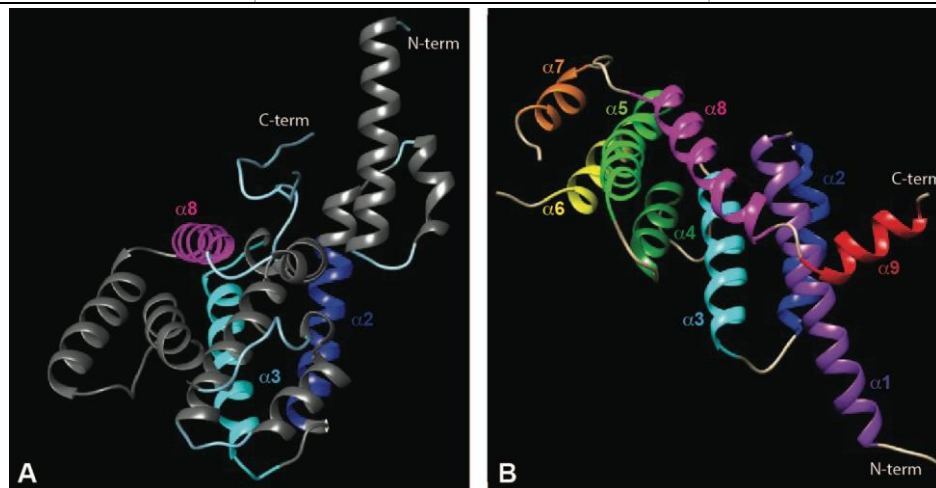


Fig. 32. Predicción in silico del dominio TRFH de *E. histolytica*. A) Modelo estructural tridimensional del dominio TRFH de EhTRF-like I. Se muestra la presencia de α -hélices conservadas (α -hélice 2–azul, α -hélice 3–cyan y α -hélice 8–magenta) y la arquitectura enteramente helicoidal característica. B) Estructura tridimensional del dominio TRFH de HsTRF1 con base en su plantilla PDB 1H6O. Se muestra la presencia de nueve α -hélices (N-terminal α -hélice 1 – morado, α -hélice 2– azul, α -hélice 3 – cyan, α -hélice 4 – verde, α -hélice 5 – verde claro, α -hélice 6 – amarillo, α -hélice 7 – anaranjado, α -hélice 8 – magenta, α -hélice 9 – rojo C-terminal).

Una vez identificadas las α -hélice conservadas en el modelo de TRFH de la proteína EhTRF-like I, se analizó la secuencia de aminoácidos que poseen estas hélices para la búsqueda de residuos relacionados con la función del dominio. Dicho análisis permitió la localización de varios residuos hidrofóbicos en el dominio TRFH de la proteína EhTRF-like I necesarios en la interface para formación del homodímero. Entre estos residuos están leucinas, isoleucinas y fenilalaninas, las cuales ocupan posiciones similares a los presentes en el dominio TRFH de su homólogo HsTRF1 (Fig. 33). De manera interesante se identificó al residuo del aminoácido fenilalanina 142, responsable de la unión de HsTRF1 con TIN2 a través del dominio TRFH (Chen, 2013).

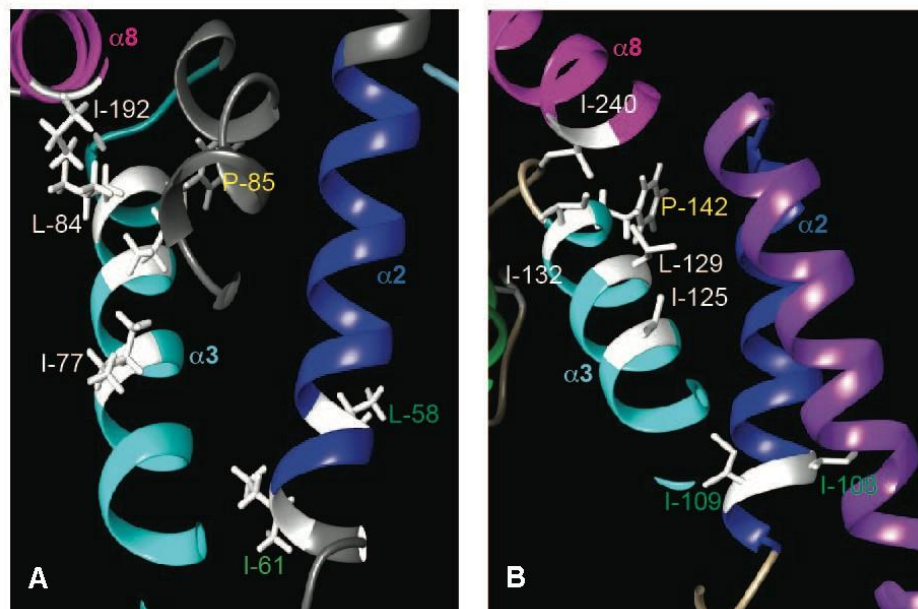


Fig. 33. Identificación de los aminoácidos conservados importantes para el dominio TRFH de *E. histolytica*. A) Estructura tridimensional del dominio TRFH de EhTRF-like I, donde se muestra la presencia de aminoácidos conservados (α -hélice 2 – Leu58 y Ile61, α -hélice 3 – Ile77, Val81, Leu84, Phe85, α -hélice 8 –Ile192) en posiciones similares a HsTRF1. B) Estructura tridimensional del dominio TRFH de HsTRF1, donde se muestra la presencia de aminoácidos conservado en posiciones similares a EhTRF-like I (α -hélice 2 – Ile108, Ile109, α -hélice 3 – Ile125, Leu129, Ile132, Phe142, α -hélice 8 –Ile240).

Nuestros resultados sugirieron la presencia de un dominio TRFH en la proteína EhTRF-like I a pesar de la baja identidad de secuencia y su divergencia con el dominio TRFH de *H. sapiens*, debido a que se encontraron arquitecturas similares y la presencia de residuos de aminoácidos con características hidrofóbicas necesarios para la interacción proteína-proteína. Por otro lado, el dominio DBD MYB se encuentra altamente conservado en secuencia y estructura.

Determinar la K_d de la proteína rEhTRF-like I con una secuencia telomérica canónica humana (HsTEL) y con una secuencia obtenida de los repetidos tipo STR de los tRNA de *E. histolytica* (EhTEL)

a) Purificación de la proteína EhTRF-like I

Con el fin de obtener a la proteína EhTRF-like I recombinante con una pureza adecuada se expresó la proteína en la cepa de *E. coli* BL21 Rosetta™ (DE3), la cual expresa codones de eucariotas superiores en el plásmido *pCold-ehtrf-like I*. Al analizar el extracto de proteínas de nuestro cultivo mediante SDS-PAGE al 10% teñido con Azul brillante de Coomassie, observamos el enriquecimiento de una banda de aproximadamente 47 kDa, correspondiente al peso teórico de la proteína recombinante, lo cual sugirió que la proteína rEhTRF-like I se expresó en las condiciones de trabajo (Fig. 34A). Para determinar si rEhTRF-like I se encontraba en fracción soluble se realizó un proceso de lisis celular donde las bacterias *E. coli* BL21 Rosetta™ (DE3) que sobre-expresan a rEhTRF-like I fueron empastilladas y posteriormente incubadas en buffer de lisis con lisozima seguido de 3 ciclos de sonicación (30 pulsos, 30 segundos, 80% amplitud) con la finalidad de extraer proteínas solubles y eliminar ácidos nucleicos del cultivo bacteriano. El fraccionamiento se analizó mediante SDS-PAGE al 10% teñido con Azul brillante de Coomassie observándose a la proteína rEhTRF-like I predominantemente en la fracción soluble (Fig. 34B). Para corroborar la expresión y solubilidad de la proteína rEhTRF-like I se realizó un ensayo de inmunodetección utilizando un anticuerpo primario comercial α -His dirigido contra el tag de histidinas y un anticuerpo secundario α -ratón HRP. Mediante este ensayo observamos el reconocimiento del anticuerpo α -his hacia el tag de histidinas de la proteína rEhTRF-like I y la abundancia de esta en la fracción soluble en una banda aproximada a los 47 kDa (Fig. 34C).

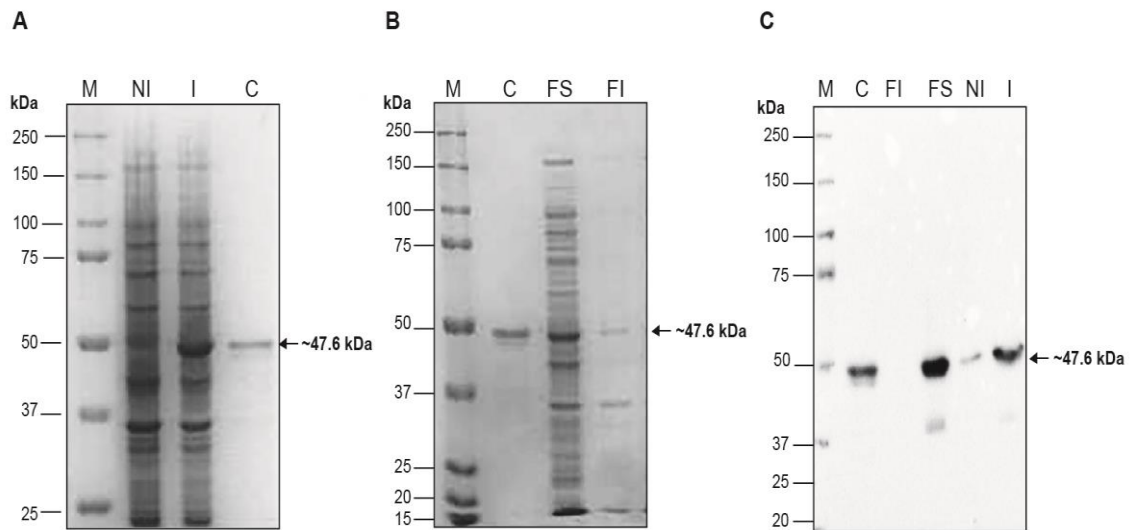


Fig. 34. Inducción y solubilidad de la proteína rEhTRF-like I. A) Extractos totales de bacterias Rosetta BL21. No inducido (NI), inducido (I). Control (C) de proteína rEhTRF-like I. Se indica su peso molecular B) Se muestra (C) control de proteína rEhTRF-like I, Fracción soluble (FS), Fracción insoluble (FI). C) Inmunodetección de la proteína rEhTRF-like I inducida. Control (C) Control, Fracción insoluble (FI), Fracción soluble (FS), No inducido (NI), Inducido (I). La flecha indica a la proteína rEhTRF-like I en su peso molecular teórico de 47.6 kDa.

Para la realización de los ensayos EMSA fue necesario purificar a la proteína rEhTRF-like I con suficiente pureza. La purificación se llevó a cabo mediante cromatografía de afinidad, la cual se basa en la unión específica y reversible de una proteína a un ligando unido a una matriz. El ligando se puede unir directamente a la proteína de interés o a una etiqueta que se une covalentemente a la proteína. Utilizamos el His-Select Nickel Affinity Gel y se realizó una purificación en modo discontinuo o *batch*. La proteína rEhTRF-like I etiquetada con un *tag* de histidinas se adsorbió a través del ligando de Ni-NTA unido covalentemente a la matriz de agarosa. Utilizamos Imidazol para eluir a la proteína debido a su capacidad para coordinarse con el ion Ni^{+2} y desplazar a la proteína rEhTRF-like I unida. La fracción obtenida se dializó con buffer de diálisis empleando una membrana con un tamaño de poro de 12 kDa. La proteína purificada se analizó mediante SDS-PAGE al 10% y se tiñó con Azul brillante de Coomassie en donde se observó una banda que correspondía al peso molecular predicho de la proteína EhTRF-like I de 47 kDa (Figura 35A). Para verificar el grado de purificación de la proteína rEhTRF-like I se realizó un ensayo de inmunodetección con el anticuerpo comercial primario α -His, donde se observó en

las fracciones eluidas la presencia de la proteína rEhTRF-like I purificada en el peso molecular esperado de 47 kDa (Figura 35B). Como se observa en el carril E1 y E2, la proteína rEhTRF-like I purificada tiene la cantidad, concentración y el grado de pureza adecuada para los ensayos EMSA.

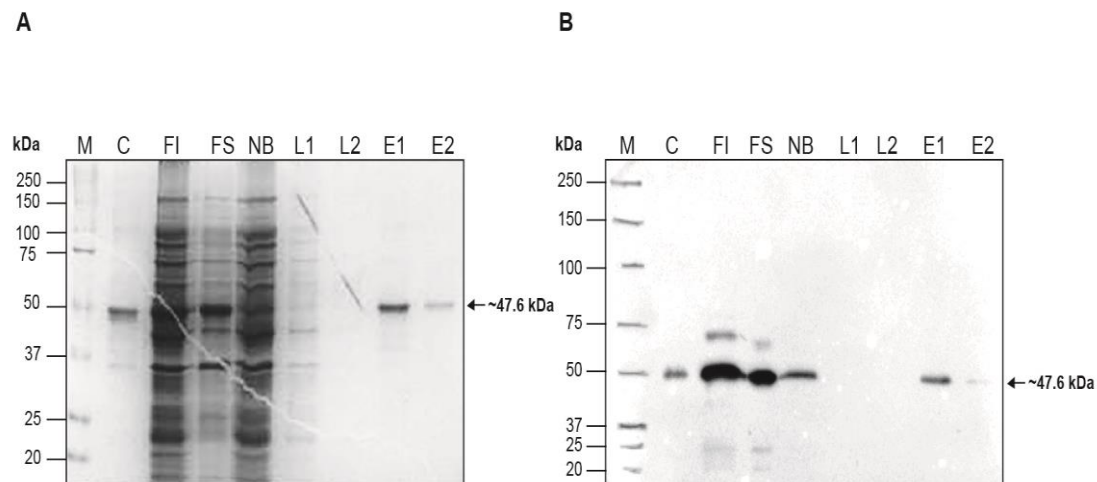


Fig. 35. Purificación e Inmunodetección de la proteína rEhTRF-like I. A) Purificación en batch de la proteína. Control (C), Fracción soluble (FS), Fracción insoluble (FI), Fracción no unida (NB), Lavados (L1 y L2), Elución de rEhTRF-like I (E1 y E2). B) Inmunodetección de la proteína rEhTRF-like I purificada. Control (C), Fracción soluble (FS), Fracción insoluble (FI), Fracción no unida (NB), Lavados (L1 y L2), Elución de rEhTRF-like I (E1 y E2). La flecha indica a la proteína rEhTRF-like I en su peso molecular teórico de 47.6 kDa.

b) Ensayos de Retardamiento de Movilidad Electroforética (EMSA)

El dominio de unión a DNA MDBD de HsTRF1 reconoce de forma específica la secuencia canónica humana TTAGGG. En este contexto y considerando que la predicción de estructura terciaria indica que la proteína EhTRF-like I tiene un dominio MDBD similar en estructura al de los humanos y que los aminoácidos relevantes para el reconocimiento de la secuencia telomérica están conservados, en este proyecto planteamos determinar la afinidad que tiene la proteína EhTRF-like I por la secuencia canónica humana (HsTel) y por una secuencia obtenida de los STR de los arreglos NK2 de tRNA de *E. histolytica* (EhTel). Para ello, se sintetizaron los oligonucleótidos HsTel y EhTel, se verificó su integridad en un gel de agarosa al 2% por tinción con Bromuro de Etidio, se procedió a hibridarlos y se prepararon a una concentración final de 100 ng/μL. Posteriormente, se marcaron con ATP [γ - 32 P] y la incorporación del nucleótido radiactivo se cuantificó en un contador de centelleo (Tabla 9).

Tabla 9. Incorporación del nucleótido radiactivo en los oligonucleótidos que contienen las secuencias teloméricas

Oligonucleótido	CPM
HsTel	219.298
EhTel	215.778

CPM: Cuentas por minuto

Primeramente, se estandarizaron los ensayos EMSA empleando diferentes concentraciones de los oligonucleótidos HsTel y EhTel, diluyendose en 5000, 10000 y 20000 CPM. Todos los geles realizados durante los ensayos EMSA fueron analizados en geles nativos no desnaturalizantes de poliacrilamida al 6% de baja fuerza iónica (0.5X TBE). Se determinó que la concentración de 5,000 CPM, la cual corresponde a 0.012 ng, sería la utilizada para la interacción con la proteína rEhTRF-like I (Fig. 36A). Para determinar las concentraciones de proteína rEhTRF-like I a utilizar se partió de un stock a una concentración de 5.02 nM/ μ L y se utilizaron diferentes concentraciones molares. En la Fig. 36B se observa que a partir de 60.24×10^3 nMoles de proteína rEhTRF-like I se presenta una retención completa de la sonda EhTel a causa de su interacción y formación del complejo DNA-proteína. Por lo que se procedió a emplear un rango de proteínas a partir de 5.02 nM/ μ L hasta 45.18 nM/ μ L.

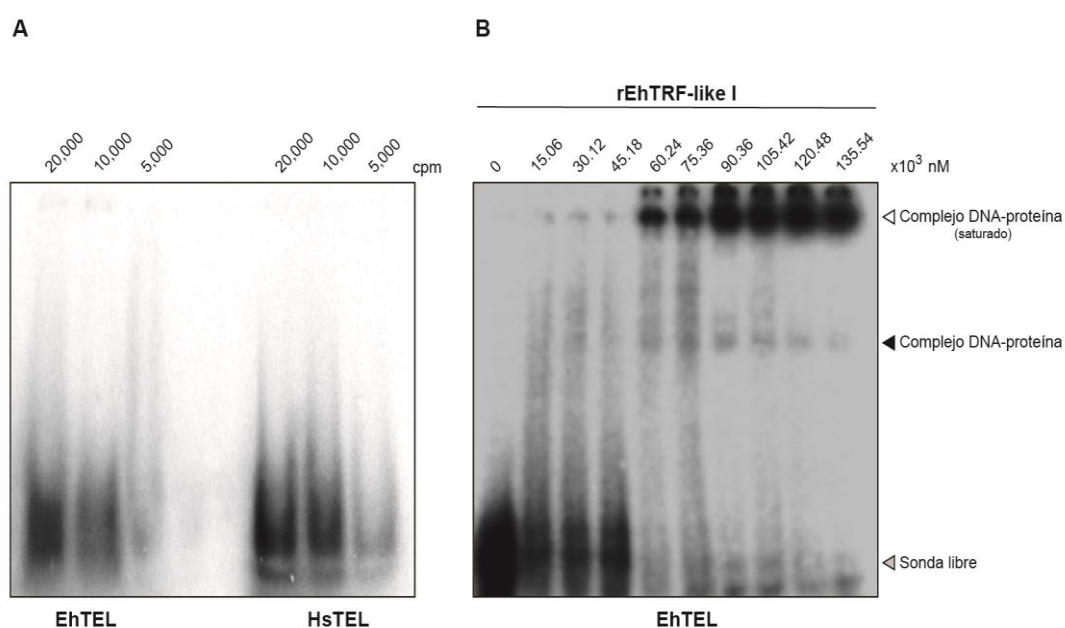


Fig. 36. Estandarización de la concentración de oligonucleótidos y proteína rEhTRF-like I en los ensayos EMSA. A) Sonda EhTel y HsTel en 20000, 10000 y 5000 CPM. B) Sonda EhTel a una dilución de 5000 CPM en interacción con la proteína rEhTRF-like I en concentraciones molares crecientes.

Para calcular la constante de disociación (K_d) se hicieron al menos tres ensayos con cada una de las sondas HsTel y EhTel y se cuantificaron los complejos DNA-proteína y el DNA libre mediante el programa Quantity One 1-D Analysis Software (BIO-RAD). La K_d se calculó empleando el programa Prism - GraphPad y la herramienta Equation: One site -- Specific binding. Con este análisis pudimos determinar que la proteína EhTRF-like I tiene una K_d de 6.8 nM para la secuencia telomérica EhTel y una K_d de 2.8 para la HuTel (Fig. 37). Este resultado sugiere que la proteína EhTRF-like I presenta una mayor afinidad por la secuencia canónica humana, lo cual era de esperarse ya que el MDBD de la proteína EhTRF-like I conserva los aminoácidos que interaccionan con esta secuencia como lo es el motivo VxxKDxxR.

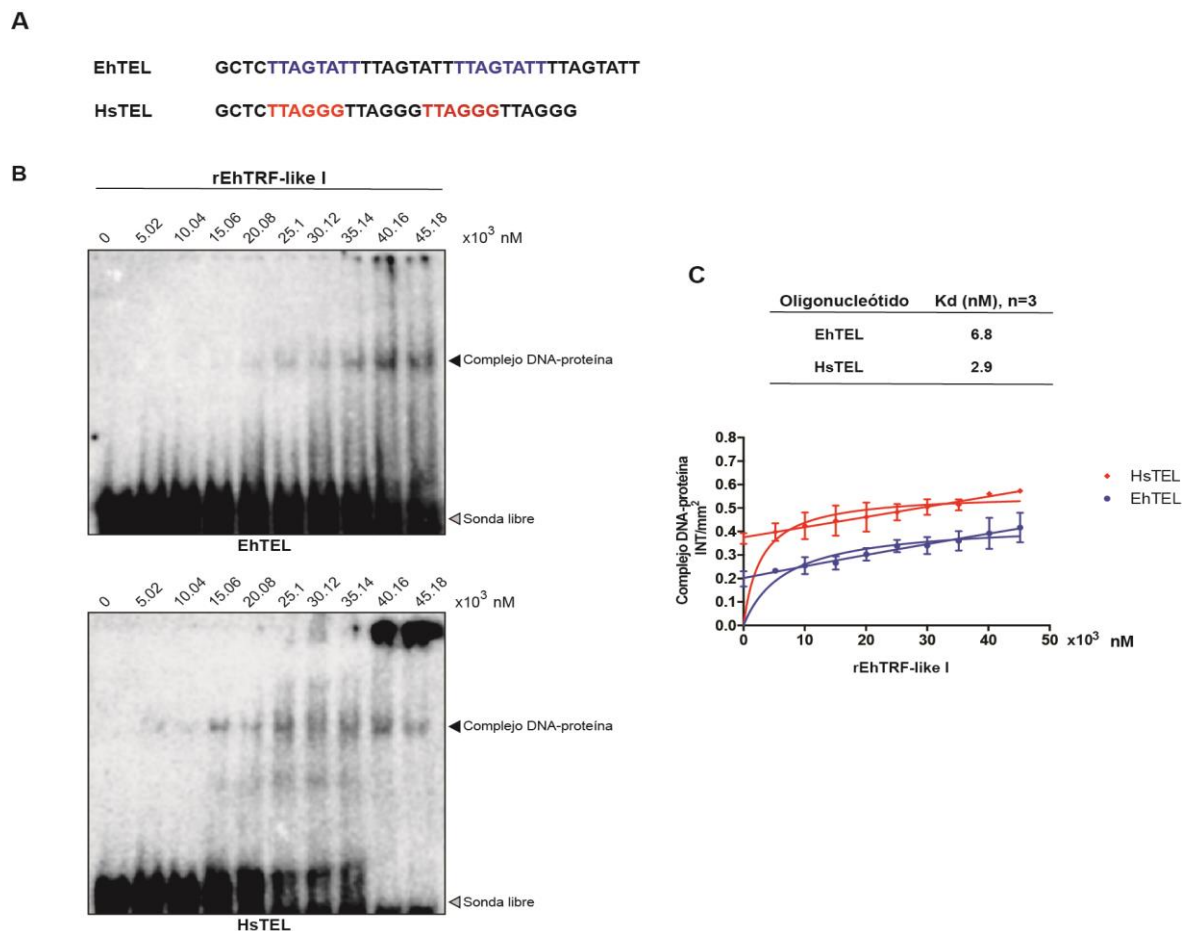


Fig. 37. Determinación de la constante de disociación (K_d) de la proteína rEhTRF-like I a las secuencias EhTel y HsTel. A) Secuencias de los oligonucleótidos marcados y utilizados para los ensayos. B) Ensayos EMSA representativos donde se muestran los complejos DNA-proteína utilizando las sondas EhTel y HsTel, se observa el complejo DNA-proteína y el DNA no unido. C) Constante de disociación de la proteína EhTRF-like I y las sondas EhTel y HsTel que fueron determinadas con una curva de regresión no lineal. Las barras de error representan la desviación estándar de 3 ensayos independientes (n=3).

c) Otros STR de *Entamoeba histolytica*

Con la finalidad de ampliar nuestro conocimiento en la búsqueda de posibles secuencias con funciones teloméricas de *E. histolytica*, decidimos probar una nueva sonda denominada EhTel(2) derivada del arreglo de tRNA [R^{TCT}]. La secuencia de EhTel(2) es TTATATGT la cual fue también marcada y cuantificada siguiendo el protocolo descrito anteriormente (Tabla 10).

Tabla 10. Incorporación del nucleótido radiactivo en la sonda EhTel (2)

Oligonucleótido	CPM
EhTel (2)	273,318

De igual manera se corrió un gel con diferentes diluciones de la sonda EhTel(2) y se determinó que la concentración de 5,000 CPM era la que sería la utilizada para la interacción con la proteína rEhTRF-like I (Fig. 38A). Por otro lado, también se determinaron las concentraciones molares de la proteína rEhTRF-like. Interesantemente, la proteína EhTRF-like I formó complejos DNA-proteína con la sonda EhTel(2), pero a diferencia de la sonda EhTel se obtuvieron 4 complejos DNA-proteína, lo cual podría evidenciar la presencia de varios sitios de unión para la proteína en la secuencia de DNA o que la proteína rEhTRF-like I se unía de manera inespecífica (Fig. 38B). En el siguiente ensayo empleamos diferentes concentraciones molares de rEhTRF-like I (de 2.51 a 40.16 x10³ nM) en presencia de 0.005 µM/µL de heparina para eliminar el efecto de uniones DNA-proteína inespecíficas. En estos ensayos se obtuvieron dos complejos DNA-proteína los cuales se definieron mejor en presencia de heparina y por esta razón se continuó utilizando en los ensayos posteriores (Fig. 39A). Posteriormente se hizo un primer ensayo con diferentes concentraciones de proteína rEhTRF-like I a partir de 5.02 nM/µL y hasta 45.18 nM/µL. En estos ensayos se obtuvieron dos complejos DNA-proteína por lo cual no fue posible determinar la *K_d* mediante el análisis anteriormente utilizado (Fig 39B). Sin embargo, realizamos un ensayo de competencias el cual nos permitió determinar la especificidad de los complejos obtenidos, para lo cual empleamos como competidores concentraciones molares crecientes (50, 100, 200 nM) de las sondas EhTel(2), EhTel, HsTel y NonTel. En

la Fig. 40 se observa que los complejos DNA-proteína obtenidos con la sonda EhTel (2) compitieron con concentraciones molares crecientes (50, 100, 200 nM) de las sondas no marcadas de EhTel (2) y EhTel, no así con las mismas concentraciones de las sondas HsTel y NonTel. Estos resultados sugirieron que la proteína rEhTRF-like I tiene capacidad de unirse a otra secuencia derivada de los STR de *E. histolytica*, sin embargo, los complejos obtenidos no compitieron con el exceso molar de la sonda HsTel, lo cual pudiera deberse al mayor número de sitios de interacción requiriendo mayor concentración de esta sonda para poder competir.

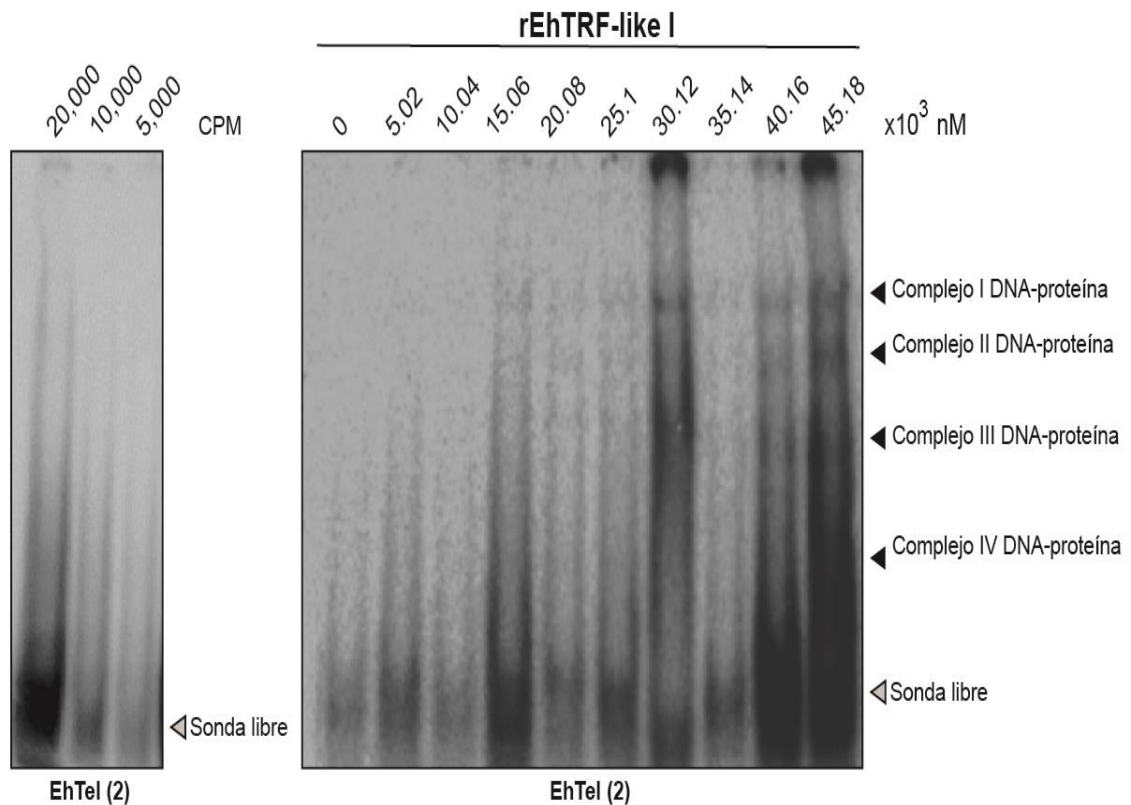


Fig. 38. Determinación de la concentración de oligonucleótidos y proteína EhTRF-like I. A) Sonda EhTel (2) en 20,000, 10,000 y 5,000 CPM. B) Sonda EhTel (2) en interacción con la proteína EhTRF-like I en concentraciones molares crecientes.

Identificar proteínas que interaccionen con la proteína rEhTRF-like I

a) Obtención de Extractos Totales y Nucleares de *E. histolytica*

Para la realización del ensayo de interacción proteína–proteína mediante Pull Down se requiere de extractos totales y nucleares. Para ello, los extractos fueron obtenidos de un cultivo axénico de trofozoítos de la cepa HM1: IMSS de *E. histolytica* en fase de crecimiento logarítmico. El análisis de extractos proteicos totales y nucleares obtenidos se hizo mediante SDS-PAGE al 10% teñido con Azul brillante de Coomasie (Fig. 41). En la cual se puede observar que las proteínas obtenidas cuentan con la integridad y la concentración suficiente para realizar los ensayos de Pull Down.

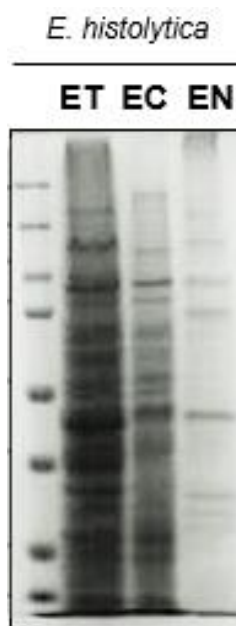


Fig. 41. Patrón de bandas de trofozoítos de *E. histolytica*. Las proteínas totales y nucleares de *E. histolytica* se obtuvieron de 1×10^6 trofozoítos las proteínas obtenidas se separaron en un gel SDS-PAGE al 10%. Extractos totales (ET), extractos citoplasmáticos (EC) y extractos nucleares (EN).

b) Purificación de la proteína rEhTRF-like I

Previamente habíamos purificado a la proteína EhTRF-like I recombinante mediante cromatografía de afinidad mediante batch empleando columnas de sepharosa His-Select Nickel Afinity Gel. Sin embargo, mediante esta metodología la proteína estaba acompañada de algunas proteínas contaminantes, lo cual para EhTRF-like I de *Entamoeba histolytica*.

los ensayos de EMSA no interfiere con los resultados. Para el caso de las interacciones proteína – proteína, es necesario el aislamiento de la proteína rEhTRF-like I con un mayor grado de pureza con el fin de garantizar que la proteína “cebo” sea la que capture a sus posibles aliados de interacción y no los contaminantes. Por lo tanto, la proteína rEhTRF-like I se purificó por afinidad a través del equipo AKTAprime plus con una columna HisTRAP™ HP de 1 mL siguiendo el método HisTag Purification His Trap. Para verificar la pureza de la proteína purificada se tiñieron los geles con el Kit Silver Stain Plus (Bio-Rad) de manera que se pudieran visualizar todos los contaminantes. En la Fig. 42 se muestran las diferentes fracciones eluidas en donde podemos observar que a partir de la fracción 17 se eluye la proteína rEhTRF-like I observándose un grado de pureza mayor.

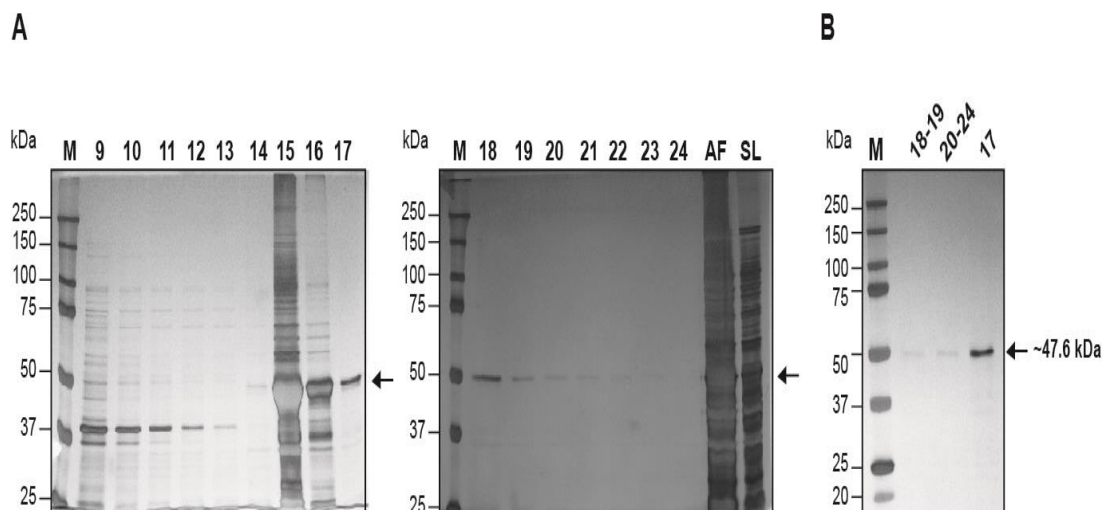


Fig. 42. Purificación de proteína rEhTRF-like I para los ensayos de Pull Down. Se muestra la obtención de la proteína rEhTRF-like I a partir de la fracción 17 y hasta la 24. B) Fracciones concentradas para usarse en el ensayo de Pull Down.

c) Ensayo de Pull Down

Posteriormente, se realizó el ensayo de Pull Down empleando los extractos totales y nucleares de los trofozoítos de *E. histolytica*. Para ello la proteína rEhTRF-like I se unió a la matriz de sepharosa His-Select Nickel Affinity Gel. Por otro lado, los extractos nucleares y totales de *E. histolytica* se preclarificaron en la resina para eliminar las proteínas que se unían a la resina de manera inespecífica, la fracción no unida correspondió a los extractos totales y nucleares

preclarificacos. Posteriormente, estos extractos preclarificados se incubaron con la proteína rEhTRF-like I para finalmente eluirse y separarse en un gel de acrilamida al 10% en el cual se incluyeron diferentes controles y se tiñeron con el Kit Silver Stain Plus (Bio-Rad). Estos ensayos nos permitieron obtener bandas diferenciales en los extractos nucleares y totales en las cuales se obtuvieron proteínas de pesos moleculares comprendidos entre 20 y 200 kDa las cuales podrían ser el resultado de las interacciones proteína-proteína. Se obtuvieron alrededor de 26 y 18 bandas diferenciales en los extractos totales y nucleares, respectivamente (Fig. 43). Los pesos moleculares de las proteínas obtenidas podrían corresponder a proteínas que previamente se ha descrito mediante análisis bioinformáticos relacionados con la función de la proteína EhTRF-like I, por ejemplo, proteínas involucradas en la maquinaria de Shelterin como EhRap (17.85kDa), proteínas involucradas en la vía ALT como: EhATM (287 kDa), EhRad50 (146.4 kDa) EhBLM (137.5 kDa), EhWRN (58.5 kDa), EhNBSI (51.5 kDa), EhRad51 (40.3 kDa) y EhRad52 (26.9kDa), EhMRE11 (68 kDa), EhMRE11-like (25.2kDa) y EhRPA (16.7kDa) o proteínas estructurales de los cromosomas como la proteína EhSMC (143.9 kDa).

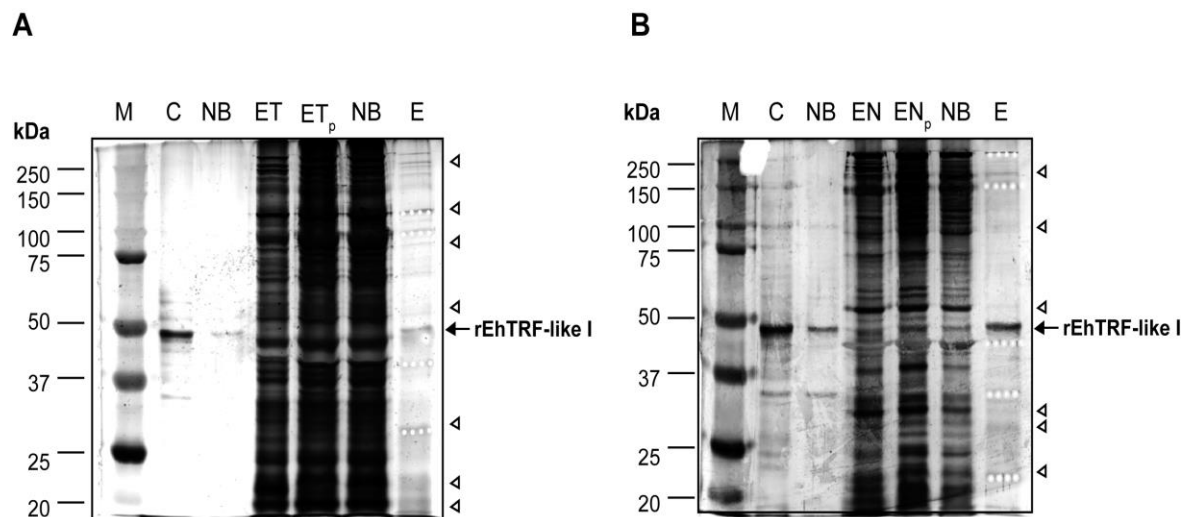


Fig. 43. Ensayo de Pull Down con extractos totales y nucleares de trofozoítos de *E. histolytica*. A) Se observa la presencia de proteínas diferenciales que interaccionan con la proteína rEhTRF-like I en extractos totales I. (C) Control, (NB) Fracción no unida de rEhTRF-like I, (ET) Extractos totales, (ET_p) Extractos totales preclarificados, (NB_p) Fracción no unida de los extractos totales preclarificados (E) Elución. Las muestras que se identificaron por MS son C, ET_p, E. B) Se observa la presencia de proteínas diferenciales que interaccionan con la proteína rEhTRF-like I en extractos nucleares. (C) Control, (NB) Fracción no unida de rEhTRF-like I (EN) Extractos nucleares (EN_p) Extractos nucleares preclarificados (NB_p)

Fracción no unida de los extractos nucleares preclarificados (E) Elución. Las muestras que se identificaron por MS son C, ENp, N.

Para conocer la identidad de estas proteínas, las fracciones eluidas de los extractos totales y nucleares se enviaron a identificar por medio de espectrometría de masas a la Unidad de Proteómica del CINVESTAV. En estos ensayos también se incluyeron las fracciones de las proteínas que se unieron de manera inespecífica a la matriz de His-Select Nickel Affinity Gel para excluir de los resultados a estas proteínas. En los extractos totales se identificaron 665 proteínas de las cuales solo se consideraron 89 proteínas capaces de interactuar con la proteína rEhTRF-like I (Tabla 11), estas se agruparon en funciones relacionadas con el citoesqueleto de actina, morfogénesis celular, procesos de oxidación y reducción, glicolisis, detoxificación celular y traducción (Fig. 44).

Tabla 11. Proteínas que interactúan con la proteína rEhTRF-like I identificadas por MS en Extractos Totales

Nº Acceso	Anotación	Score	Protein Matched Products	Protein Matched Peptides	Protein Seq Cover (%)
B1N2P0	Actin	12.683.14	137	20	46.01
B1N621	Actin putative	11.918.23	148	18	43.67
B1N5D1	Actin 2 protein putative	3.810.285	21	5	36.84
C4LYJ9	Actin-related protein 2/3 complex subunit	782.232	19	5	16.94
C4LTB5	Actin-related protein 2/3 complex subunit	782.232	21	5	16.94
C4LSV0	Actin-related protein 2/3 complex subunit	669.9	11	4	20.94
C4LXD7	Fructose 1 6-bisphosphate aldolase putative	2.834.370	32	4	11.52
C4M295	Adenylyl cyclase-associated protein	3.649.737	18	7	18.83
B1N3T9	Adenylylcyclase-associated protein putative	3.048.575	11	3	23.46
C4M2R2	Aldehyde-alcohol dehydrogenase	2.068.295	74	22	24.14
C4M230	Aldehyde-alcohol dehydrogenase	2.039.070	66	21	23.91
B1N387	Aldehyde-alcohol dehydrogenase 2 putative	1.816.677	44	15	18.09
C4LWS2	Alcohol dehydrogenase 3 putative	3.624.280	49	11	24.74
B1N3M8	Alcohol dehydrogenase 3 putative	2.920.628	26	5	32.65
C4M1G7	Alcohol dehydrogenase putative	2.792.999	13	2	8.12
C4M7F5	Alcohol dehydrogenase putative	615.538	15	5	11.69
C4LUV9	Acetyl-CoA synthetase putative	563.323	40	11	14.45
C4LXE8	Enolase putative	11.459.92	97	12	34.4
C4LVR9	Glyceraldehyde-3-phosphate dehydrogenase	2.451.102	44	7	19.46

TESIS

Q24801	Pyruvate phosphate dikinase	6.447.500	130	23	26.44
C4LTX6	Pyruvate ferredoxin oxidoreductase	1.543.691	93	15	14.54
C4LVG4	Actophorin putative	2.720.281	13	2	19.57
B1N2J0	Malic enzyme putative	1.546.707	39	9	24.02
Q9NH04	Malic enzyme	1.546.707	39	9	24.02
C4M6Q4	Peroxiredoxin	5.018.954	45	10	35.04
C4MAR7	Peroxiredoxin	4.970.804	43	9	35.37
B1N623	Peroxiredoxin putative	4.970.804	41	8	38.02
B1N5Y9	Peroxiredoxin putative	4.970.804	41	8	54.89
C4MBP3	Peroxiredoxin	4.970.804	43	9	34.18
C4MA54	Peroxiredoxin	4.970.804	43	9	34.76
C4M3Q0	Peroxiredoxin	4.970.804	43	9	34.76
C4M8W7	Peroxiredoxin	4.970.804	43	9	34.76
B1N5A8	Peroxiredoxin	3.335.361	35	8	29.54
B1N5L9	Peroxiredoxin putative	1.635.444	12	3	15.6
C4M3S5	Heat shock protein 70 putative	4.283.253	27	7	14.33
B1N4N6	Heat shock protein70 hsp70A2 putative	4.048.687	17	4	12.12
C4LZ45	Heat shock protein70 hsp70A2 putative	4.048.687	17	4	9.35
B1N486	Heat shock protein 70 putative	4.048.687	17	4	8.79
C4M9V9	Heat shock protein 70 putative	4.048.687	17	4	8.79
C4M1G4	Heat shock protein70 hsp70a2 putative	4.048.687	17	4	8.42
B1N5W1	Heat shock protein 70 putative	2.880.747	7	2	11.9
C4M8D4	Heat shock protein 70 putative	2.880.747	7	2	10.26
B1N5L3	Heat shock protein70 hsp70a2 putative	2.880.747	7	2	19.75
B1N5L2	Heat shock protein70 hsp70a2 putative	1.262.630	12	3	23.4
B1N398	Heat shock protein70 hsp70A2 putative	1.167.940	10	2	10.91
B1N5M8	Heat shock protein70 hsp70A2 putative	1.167.940	10	2	21.24
Q9BLF3	Rab family GTPase	1.895.938	20	4	24.39
Q963D1	Rab family GTPase	1.651.178	15	4	14.57
Q5NT43	Rab family GTPase	1.614.545	12	2	10.63
Q5NT38	Rab family GTPase	1.239.903	8	1	5.67
B1N3J0	Rab family GTPase	1.233.234	8	1	1.47
Q5NSZ0	Rab family GTPase	1.233.234	8	1	5.47
Q5NT40	Rab family GTPase	1.233.234	8	1	5.42
Q5NT36	Rab family GTPase	1.233.234	8	1	5.64
C4M1Y3	Rab family GTPase	1.233.234	8	1	4.17
C4MA07	Rab family GTPase	1.233.234	8	1	5.53
Q6AW59	EhRabC2 protein	1.239.903	8	1	5.34
Q9NJV1	Grainin 2	1.309.470	19	5	15.02
B1N3T3	Grainin 2	1.038.491	12	3	16.67
B1N2I7	60S ribosomal protein L10a 2 putative	1.506.578	17	5	26.64
C4M5A0	40S ribosomal protein	1.101.468	15	4	18.5

TESIS

B1N336	40S ribosomal protein S4 putative	725.876	11	3	6.6
C4LYF4	40S ribosomal protein S4 putative	725.876	11	3	6.44
C4LT02	40S ribosomal protein S3 putative	682.205	6	2	10.25
C4M7N8	40S ribosomal protein S17 putative	619.187	6	1	9.4
C4M0Z2	40S ribosomal protein S17 putative	619.187	7	1	9.4
C4M609	40S ribosomal protein S3a	487.308	17	5	13.28
C4LTQ5	40S ribosomal protein S3a	487.308	17	5	13.28
B1N2F9	Type A flavoprotein putative	728.765	19	6	11.58
C4M7D4	Elongation factor 1-alpha	970.721	39	8	16.52
C4MAC9	Elongation factor 1-alpha	927.097	30	7	14.03
C4M7U2	Ribosomal protein L15	607.755	9	3	6.83
C4M6Z7	Triosephosphate isomerase	683.417	6	2	10.34
C4M192	Phosphoglycerate kinase	751.615	9	2	5.54
C4M6N9	Protein kinase putative	603.015	9	7	7.85
C4M1I9	Uncharacterized protein	7.281.259	162	19	33.17
B1N2H6	Uncharacterized protein	4.173.372	31	6	21.14
B1N5L0	Uncharacterized protein	2.880.747	7	2	7.16
C4M779	Uncharacterized protein	2.080.943	8	2	16.58
C4LWQ4	Uncharacterized protein	864.25	14	4	26.73
C4M1J1	Uncharacterized protein	760.059	7	1	1.59
B1N2Q8	Uncharacterized protein	466.203	5	2	11.35

Protein.score: Es un valor asignado por el equipo, no tiene límite específico, pero entre más grande hay más confiabilidad que la proteína que se detectó en la interacción sea la que dice ser. En este caso, entre más grande mejor. **Protein.MatchedProducts:** Indica cuantas veces se encontró esa proteína, habla de la abundancia de la muestra. **Protein.MatchedPeps:** Como la columna anterior, habla de la abundancia de péptidos de esa proteína que fueron detectados por el equipo. **Protein.DigestPeptides:** Indica cuantos péptidos se encontraron de la proteína digeridos con una enzima, en este caso tripsina. **Protein.seqCover(%):** El porcentaje de cobertura de la secuencia.

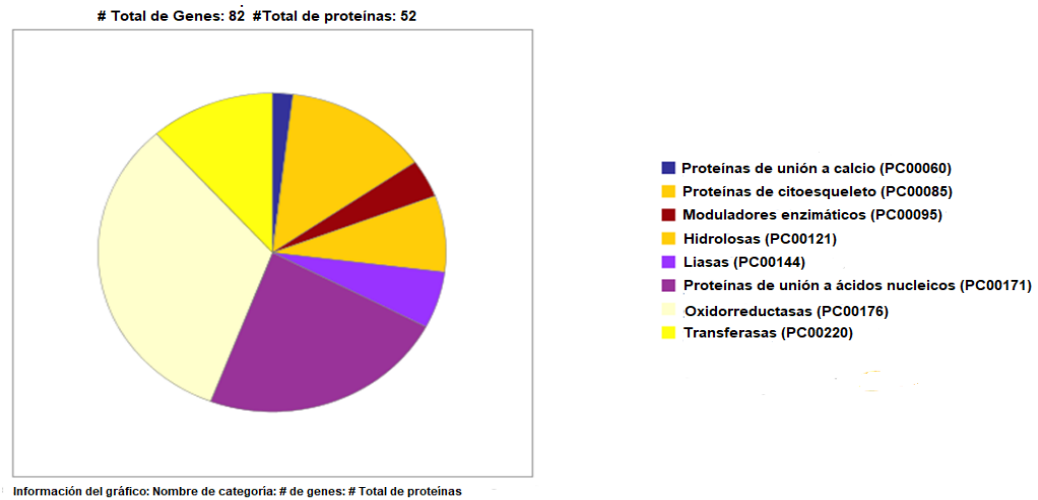


Fig. 44. Clasificación de proteínas que interactúan con la rEhTRF-like I identificadas en extractos totales. Clasificación funcional de las proteínas de extractos totales de trofozoítos de *E. histolytica* que interactúan con rEhTRF-like I identificadas por espectrometría de masas y analizadas mediante GO Panther (<http://www.pantherdb.org>).

En el caso de las proteínas nucleares que interactuaron con rEhTRF-like I se identificaron 544 proteínas, de igual manera se excluyeron las proteínas que se unieron a la sepharosa de manera inespecífica y solo se obtuvieron 29 proteínas que se consideraron que interactúan específicamente con la proteína rEhTRF-like I (Tabla 12). De manera interesante se identificó a la proteína EhTRF-like II como una de las proteínas que interactuaron con la proteína rEhTRF-like I (C4M1J1). Este hallazgo es relevante ya que EhTRF-like I y II comparten características similares a la proteína TRF1, lo cual podría significar que ambas proteínas podrían formar dímeros EhTRF-like I-EhTRF-like II. Además, se identificaron otras proteínas relacionadas con funciones como procesos de glicosilación, polimerización de actina, traducción y transducción de señales mediadas por GTPasas (Fig. 45). Adicionalmente, se identificaron 2 proteínas las cuales están anotadas como hipotéticas en la base datos AmoebaDB, sin embargo, al ser analizadas mediante la herramienta BLASTP las proteínas con los números de acceso C4M2R6 y C4LSC8 presentan homología con la Histona H1.4 (31.2%) y proteínas de mantenimiento estructural de los cromosomas SMC (27.4%) respectivamente.

Tabla 12. Proteínas que interaccionan con la proteína rEhTRF-like I identificadas por MS en Extractos Nucleares

Nº Acceso	Anotación	Score	Protein Matched Products	Protein Matched Peptides	Protein Seq Cover (%)
B1N2P0	Actin	10.664.0	126	18	38.83
B1N621	Actin putative	10.664.0	147	18	43.98
C4M4P4	Actin-binding protein cofilin tropomyosin family	1.234.75	6	2	25.68
C4M782	Arp2/3 complex 34 kDa subunit	481.636	12	4	10.92
C4LXD7	Fructose 1 6-bisphosphate aldolase putative	4.024.43	29	5	30
C4LVR9	Glyceraldehyde-3-phosphate dehydrogenase	16.765.6	99	11	39.22
C4LXE8	Enolase putative	2.872.84	55	8	22.94
C4LVY0	Alcohol dehydrogenase putative	452.557	9	3	7.83
C4M230	Aldehyde-alcohol dehydrogenase	545.408	29	12	8.62
C4M2R2	Aldehyde-alcohol dehydrogenase	579.992	32	13	8.62
B1N387	Aldehyde-alcohol dehydrogenase 2	483.823	28	9	6.39
C4M295	Adenylyl cyclase-associated protein	3.005.96	20	5	10.04
B1N3T9	Adenylylcyclase-associated protein_ putative	2.989.62	12	2	14.81
C4LXG4	Purine nucleoside phosphorylase putative	1.115.13	16	3	20.59
C4M296	Calreticulin putative	896.828	12	2	4.88
C4LTX6	Pyruvate ferredoxin oxidoreductase	824.863	45	13	10.67
Q24801	Pyruvate phosphate dikinase	779.692	26	10	12.43
C4LZC2	Pyrophosphate fructose 6-phosphate 1-phosphotransferase	769.159	12	4	9.71
C4M1I9	Uncharacterized protein (EhTRF-like I)	6.127.05	139	15	31.44
C4M1J1	Uncharacterized protein (EhTRF-like II)	561.473	8	1	1.59
C4M2R6	Uncharacterized protein	1.074.56	20	5	13.82
C4LSC8	Uncharacterized protein EHI_153370	516.329	11	9	8.91
B1N2H6	Uncharacterized protein EHI_045090	475.034	7	2	10.98
C4M1Y2	Rab GDP dissociation inhibitor EHI_164890	468.785	9	4	11.62
C4MAC9	Elongation factor 1-alpha	441.747	19	5	13.35
C4M7D4	Elongation factor 1-alpha	441.747	20	5	13.35
C4LZ42	40S ribosomal protein S18 putative	565.357	7	3	14.84
C4M367	40S ribosomal protein S18 putative	710.387	9	4	24.48
B1N309	Rab GDP dissociation inhibitor	435.519	8	3	11.62

Protein.score: Es un valor asignado por el equipo, no tiene límite específico, pero entre más grande hay más confiabilidad que la proteína que se detectó en la interacción sea la que dice ser. En este caso, entre más grande mejor. **Protein.MatchedProducts:** Indica cuantas veces se encontró esa proteína, habla de la abundancia de la muestra. **Protein.MatchedPeps:** Como la columna anterior, habla de la abundancia de péptidos de esa proteína que fueron detectados por el equipo. **Protein.DigestPeptides:** Indica cuantos péptidos se encontraron de la proteína digeridos con una enzima, en este caso tripsina. **Protein.seqCover(%):** El porcentaje de cobertura de la secuencia.



Fig. 45. Clasificación de proteínas que interaccionan con rEhTRF-like I identificadas por MS en extractos nucleares. Clasificación funcional de las proteínas de extractos totales de trofozoítos de *E. histolytica* que interaccionan con rEhTRF-like I identificadas por espectrometría de masas y analizadas mediante GO Panther (<http://www.pantherdb.org>).

XI. Discusión

Los telómeros son los extremos terminales de los cromosomas y en ellos interaccionan un gran número de proteínas entre ellas las proteínas TRF. El acceso de las proteínas TRF es modulado por diferentes factores, entre ellos, la interacción con diversas proteínas, la estructura de la cromatina, su localización nuclear y diferentes modificaciones post-traduccionales (Prochazkova, 2016). Sin embargo, poco se conoce la organización de los extremos terminales de los cromosomas en organismos protozoarios. Un aspecto inicial para el estudio de los telómeros en estos organismos es la identificación de proteínas homólogas a las proteínas TRF. Dicha identificación se ha logrado a través del uso de sondas de la región más conservada de dichas proteínas, esto es el dominio de unión a DNA MYB, para buscar proteínas homólogas en estos organismos. De esta manera se han podido identificar proteínas TRF en organismos como *T. cruzi*, *T. brucei*, *L. amazonensis*, *D. discoideum* y, recientemente, en el parásito *E. histolytica*. En este protozoario se han identificado tres proteínas similares a las proteínas TRF, las cuales se denominaron EhTRF-like I, II y III.

El dominio de unión a DNA MYB de las proteínas TRF tiene un tamaño de alrededor de 100 aminoácidos en donde están conservados los aminoácidos LKDKWRT [telebox motif (VxKDxxR)], los cuales median la formación de puentes de hidrógeno entre las proteínas y el DNA (Bilaud T., 1996; Brocoli y col. 1997). Otro dominio relevante para la función de las proteínas TRF es el dominio de dimerización TRFH; sin embargo, éste presenta una mayor divergencia en su secuencia de aminoácidos que el DBD Myb (Bianchi y col. 1996, Brocoli y col. 1997). La estructura tridimensional de los dominios TRFH y MDBD de las proteínas TRF ya ha sido determinada (Nishikawa y col, 1998; Fairall et., 2001).

En este trabajo se propuso predecir la estructura tridimensional de los dominios funcionales MDBD y TRFH de la proteína EhTRF-like I con la finalidad de profundizar en el conocimiento de esta proteína en este parásito. La estructura tridimensional del dominio MDBD de la proteína EhTRF-like I se predijo por medio de un método de modelaje por homología empleando el Programa Robetta. Los

resultados mostraron la presencia de tres α -hélices de las cuales, la segunda y tercera α -hélices forman el característico motivo hélice-giro-hélice presente en las proteínas que tienen el DBD MYB. Además, se corroboró la presencia del motivo Telebox VxxKDxxR en la tercera α -hélice el cual está enriquecido con aminoácidos hidrofóbicos (V342, K345, D346, R349) en posiciones similares a los de la proteína TRF1. Estos aminoácidos son relevantes para mantener el plegamiento del dominio de unión a DNA de esta proteína y para el reconocimiento del surco mayor del DNA. La proteína EhTRF-like I conserva los aminoácidos lisina, arginina y ácido aspártico los cuales están involucrados en la formación de enlaces por puentes de hidrógeno con la G4, G5 y la C de la cadena complementaria de la secuencia telomérica canónica humana. Este hecho es interesante debido a que el genoma del parásito *E. histolytica* no contiene secuencias teloméricas canónicas por lo que las proteínas EhTRF-like podrían ser una herramienta para la búsqueda de las secuencias teloméricas en este organismo.

El dominio TRFH de las proteínas TRF1 y TRF2 tiene un tamaño de 200 aminoácidos. Este dominio, media la homodimerización e interacción con un conjunto de proteínas necesarias para la protección de los telómeros (Fairall L, 2001). También, este dominio interactúa con el DNA, formando un complejo donde el DNA se envuelve alrededor del dominio para controlar la topología del DNA telomérico (Benarroch-Popivker et al., 2016). Se ha reportado que este dominio posee una mayor divergencia de secuencia de aminoácidos entre las proteínas TRF de diferentes organismos, sin embargo, los dominios exhiben arquitecturas tridimensionales similares conformadas de 9 a 10 α -hélices (Brocoli y col. 1997; Fairall et al., 2001, Giraud-Panis et al., 2017). El dominio TRFH de las proteínas TRF1 y TRF2 tiene un porcentaje de identidad del 27%, sin embargo, la superposición de las estructuras tridimensionales de ambos dominios presenta que tienen estructuras 3D casi idénticas (Fairall L, 2001). Cada dímero está formado por dos monómeros que interactúan en una disposición antiparalela formando un dímero simétrico cuya estructura general se asemeja a una herradura (Fairall L, 2001). En el análisis de predicción de estructura secundaria empleando PSI-PRED se determinó que las proteínas EhTRF-like de *E. histolytica*

conservan al menos nueve α -hélices en esta región lo cual permitió delimitar este dominio (Rendón-Gandarilla, 2018). Sin embargo, debido a la divergencia de la secuencia, predecir la estructura tridimensional del dominio TRFH de la proteína EhTRF-like I representó un mayor reto debido a que tiene una identidad del 12.12% con respecto a la secuencia de aminoácidos del dominio de TRF1. Es por ello que se realizó la predicción de su estructura tridimensional por el método de *ab initio* empleando el software *Robetta*. Los modelos tridimensionales de este dominio TRFH que se construyeron presentaron un valor de confianza de 0.21% por lo que para seleccionar el mejor modelo se utilizaron criterios adicionales de selección como Tm-score [medida entre la similitud de dos estructuras de proteínas, (Zhang y Skolnick, 2005)] y el RMSD [medida de la distancia promedio entre los átomos, (Kedem et al, 1999)]. A pesar de la baja identidad de la secuencia del dominio TRFH de *H. sapiens*, encontramos que la proteína EhTRF-like I conserva una estructura tridimensional conformada por una arquitectura enteramente helicoidal que incluye un conjunto de diez α -hélices del mismo modo que ocurre en la proteína Taz1 de organismos como *S. pombe* y *S. cerevisiae* (Giraud-Pani, 2013). Por otro lado, encontramos que las regiones conformadas por las α -hélices 2, 3 y 8 de *E. histolytica* son las más conservadas en relación con la proteína TRF1; estas tres α -hélices contienen la presencia de residuos de aminoácidos como isoleucinas, leucinas y valinas con características hidrofóbicas necesarios para la interacción proteína-proteína. De manera interesante, encontramos la presencia del residuo de F80 de EhTRF-like I en la misma posición estructural que la F120 de TRF1 el cual se denomina motivo de unión TRFH (TBM). Este motivo TBM interacciona con el motivo FxLxP presente en las proteínas que interaccionan con TRF1, como la proteína TIN2 (Diotti y Loayza, 2011; Boutou, et. al. 2013) y son cruciales para la interacción específica con otros factores asociados a telómeros como la proteína PINX1, la cual es un inhibidor de la telomerasa y media la unión de TRF1 a los telómeros (Oganesian et. al., 2009). Como consecuencia de su arquitectura, los dominios TRFH, crean un sitio de acoplamiento específico para las secuencias de péptidos, lo cual explica su función como un centro de reclutamiento para varios socios del complejo de Shelterin.

Se ha sugerido un origen evolutivo común para estos dominios TRFH a pesar de su bajo nivel de identidad de secuencia, lo que dificulta su identificación a través de algoritmos de búsqueda por homología. Varias plantas, e incluso Trypanosomátidos, expresan proteínas que contienen un MDBD con motivo Telebox en su extremo C-terminal y un dominio N-terminal que probablemente se pliega en una estructura TRFH. De confirmarse, esto sugeriría que los dominios MDBD y TRFH son dos motivos teloméricos ancestrales que se propone surgieron en el momento de la divergencia entre organismos procariotas-eucariotas (Giraud-Panis, 2017).

La identificación de firmas teloméricas ha sido una cuestión desafiante en el parásito *E. histolytica*, al no encontrarse secuencias teloméricas canónicas ni ortólogos de genes de la telomerasa, en contraste con parásitos como *T. cruzi* y *G. lamblia* en los cuales si ha sido posible identificar repetidos teloméricos localizados en los extremos de sus cromosomas y expresan el gen de la telomerasa (da Silva et. al., 2010; Uzlikova et. al 2017). En *E. histolytica* se han propuesto los arreglos de repetidos de tRNA que incluyen repetidos tipo STR como secuencias teloméricas. La presencia de las proteínas EhTRF-like ha ayudado a explorar el reconocimiento de estas proteínas a dichas secuencias tipo STR. Una de estas secuencias es la EhTel (proveniente del arreglo NK2), la cual se seleccionó debido a que presenta una mayor similitud a la secuencia canónica humana y se determinó que la proteína EhTRF-like I podía interaccionar con ella y formar complejos DNA-proteína (Rendón-Gandarilla, 2018). Sin embargo, a la fecha no se había determinado la afinidad de EhTRF-like I por secuencias de DNA canónicas (EhTel) o no canónicas (HsTel) por lo que en este trabajo realizamos ensayos para conocer la afinidad de la proteína EhTRF-like I con la secuencia EhTel en comparación con la secuencia canónica humana (HsTel). Las K_d de la proteína EhTRF-like I con las secuencias EhTel y HsTel determinadas en este trabajo fueron de 6.8 y 2.9 nM, respectivamente. Estos valores indican que la proteína rEhTRF-like I tiene una mayor afinidad por la sonda telomérica humana HsTel, sin embargo, en el genoma de este organismo no se han identificado secuencias teloméricas canónicas. Las K_d de las proteínas TRF1 y TRF2 con la secuencia telomérica (5'-GTTAGGGTTAGGG-3') son de

2.0×10^{-7} nM y 7.5×10^{-7} , respectivamente, lo cual indica que TRF1 se une con mayor afinidad a la secuencia telomérica que TRF2. Al parecer la unión de TRF2 presenta menor afinidad debido a que requiere mayor flexibilidad en su unión debido a que su función fundamental está en la formación del t-loop (Hanaoka et al., 2004). Las diferencias entre la afinidad encontradas entre la proteína EhTRF-like I y las proteínas TRF1 y 2 podrían estar relacionadas con varios factores como el buffer de unión empleando, en este trabajo empleamos un buffer que contenía 12 mM HEPES, 60 mM KCl, 5 mM MgCl y 4 mM Spermidina-Magnesio en contraste con el buffer empleado por Hanaoka y colaboradores que usaron 10 mM HEPES-KOH, 3 mM EDTA, 180 mM KCl, y 0.003% Triton X-100 (v/v) (pH 6.8). Otra razón podría estar relacionada a que las proteínas EhTRF-like de *E. histolytica* se unen al DNA con menor afinidad debido a la gran variabilidad de las secuencias tipo STR que se encuentran en su genoma (42 arreglos de tRNA diferentes los cuales existen al menos 12 STRs diferentes) (Clark et al., 2005; Clark et al., 2006). Una pregunta que surge alrededor de esta gran variabilidad de las secuencias STRs de este parásito está en relación a su posible origen. En este trabajo se propone que la secuencia canónica TTAGGG, la cual está enriquecida en residuos de Guaninas, pudo haber estado presente en el genoma de un organismo primitivo que dio origen a este parásito. Sin embargo, al estar expuesto al estrés oxidativo y la acumulación de especies reactivas de oxígeno (ROS) pudo haber tenido un efecto deletéreo en las secuencias teloméricas, debido a que el nucleótido G es más proclive a la oxidación, generando aductos 8-oxoG (Zhang et al., 1997), induciendo mutaciones tipo transversión G: CT: A que resultó en secuencias ricas en AT. A nivel telomérico, 8-oxoG representa la lesión oxidativa de DNA más frecuente debido al alto nivel de guanina en la secuencia. Además, debido al estado de heterocromatina, el daño telomérico oxidativo se repara de manera menos eficiente en comparación con el daño oxidativo generalizado en el genoma (Coluzzi, 2019). Todas estas características hacen que los telómeros sean las secuencias que más frecuentemente se dañan por el estrés oxidativo. (Walker et al., 2012). *E. histolytica* está continuamente expuesta a estrés oxidativo proveniente de neutrófilos y macrófagos, los cuales producen especies reactivas de oxígeno (ROS) y nitrógeno (ERN) por lo que este parásito posee una serie de mecanismos para evadir la actividad de estas

especies reactivas (Nandakumar, 2013). De esta manera, al oxidarse los nucleótidos de G y ocasionarse transversiones de bases nucleotídicas podría explicarse la mayor afinidad de la proteína rEhTRF-like I por la secuencia HsTel y de igual manera podríamos asociarlo a la ausencia de secuencias teloméricas clásicas, debido al estado oxidativo en el que se encuentra viviendo comúnmente *E. histolytica*.

Por otro lado, probamos una segunda secuencia tipo STR de *E. histolytica* obtenida del microarreglo de tRNA [R^{TCT}] debido a que es la secuencia que se repite con mayor frecuencia, de la cual no pudimos obtener la K_d por el mismo método que con la sonda EhTel debido a la presencia de múltiples sitios de interacción DNA-proteína, pero determinamos su especificidad por la proteína rEhTRF-like I a través de ensayos de competencia. En este trabajo pudimos probar que la proteína EhTRF-like I también es capaz de formar complejos DNA-proteína con esa secuencia. Finalmente, el origen de la gran variabilidad de las secuencias tipo STR de *E. histolytica* podría estar relacionado a que el mecanismo de mantenimiento telomérico del parásito podría estar mediado por la vía Alternativa (vía ALT). Esta vía ALT tiene entre sus características una gran heterogeneidad en las secuencias teloméricas debido a que estas se mantienen a través de mecanismos de recombinación (Boutou et al., 2013; Cesare et al., 2010).

Otro aspecto a considerar en la función de la proteína EhTRF-like I son las asociaciones proteína-proteína las cuales participan en diferentes procesos biológicos en los telómeros. El proteoma de los telómeros consta de alrededor de 200 proteínas que se han asociado con diferentes aspectos de la biología de los telómeros, incluida la protección telomérica, la síntesis de DNA telomérico y el alargamiento de los telómeros (Déjardin, 2009). En el caso de TRF1, hay diferentes proteínas que interaccionan con esta proteína las cuales están involucradas en la regulación de la longitud de los telómeros o en el acceso a la secuencia telomérica. Giannone y colaboradores, realizaron ensayos tipo pull down empleando como “proteína cebo” a TRF1 identificando 5 de los 6 componentes del complejo de Shelterin, además de proteínas asociadas al telómero como la Tankyrase 1, PARP1, BLM, WRN, RAD50, DNA-pkc,

Ku70/Ku86, y ATM (Giannone et al., 2010). También estos cambios identificaron proteínas que se clasificaron en categorías funcionales tales como reparación del daño del DNA, ubiquitinación, cohesión cromosómica, modificación y remodelación de la cromatina, replicación del DNA, regulación del ciclo celular, transcripción, metabolismo de nucleótidos, procesamiento de RNA (Giannone et al., 2010).

Los ensayos de pull down empleando la proteína rEhTRF-like I nos permitieron identificar 89 proteínas provenientes de extractos totales y 29 proteínas obtenidas de extractos nucleares, las cuales mediante la herramienta GO Panther y el análisis de clasificación funcional se asociaron, en el caso de los extractos totales, a funciones relacionadas con el citoesqueleto de actina, morfogénesis celular, procesos de oxidación y reducción, glicolisis, detoxificación celular y traducción; y en el caso de los extractos nucleares, con funciones de glicosilación, polimerización de actina, traducción y transducción de señales mediadas por GTPasas. Posteriormente, las secuencias de aminoácidos de cada una de las proteínas identificadas se analizaron mediante Blast contra la base de datos de NCBI donde interesantemente logramos identificar a la proteína EhTRF-like II (locus EHI_001110) con un porcentaje de cobertura de 1.59. Esta proteína tiene una longitud de 441 aminoácidos, comparte características similares a EhTRF-like I debido a que también tiene el extremo amino terminal ácido, y un porcentaje de identidad del 93.44 con el MDBD de EhTRF-like I. Coincidentemente, estos dos genes se localizan de manera contigua en el genoma de este parásito, lo cual permite proponer que los genes *ehtrf-like I* y *II* podrían ser el resultado de una duplicación y pueden estar sujetos a co-regulación transcripcional. El hecho de identificar a EhTRF-like II interaccionando con la proteína rEhTRF-like I revela que estas dos proteínas podrían formar heterodímeros, como ocurren en las proteínas AtTRB2 y AtTRB3 de *Arabidopsis thaliana* (Procházková et al., 2016). Los ensayos de pull down también permitieron la identificación de una proteína hipotética similar a la Histona H1.4 (locus EHI_160980) con una longitud de secuencia de 304 aminoácidos y un porcentaje de cobertura de 13.82. Los miembros de la familia de histonas H1 son un componente clave de la cromatina y pueden estabilizar tanto la estructura de los nucleosomas como la arquitectura de

la cromatina de orden superior. Existen múltiples subtipos de H1 con una gran variedad de modificaciones postraduccionales (Hergeth, 2015). Las histonas H1 se diferencian de las otras cuatro familias (H2A, H2B, H3 y H4) por su interacción con el exterior del nucleosoma, donde se ubican en el sitio de entrada y salida de DNA de la core central y se unen otros 19-20 pb de DNA que unen un nucleosoma al siguiente, por lo que se denominan histonas "enlazadoras". Dichas proteínas son esenciales para regular el espaciamiento de los nucleosomas y la condensación de la cromatina. Se cree que la H1 brinda estabilidad genómica porque en células de levadura con carencia de H1 se ha demostrado inestabilidad genómica debido al aumento de recombinación homóloga (HR) (Parseghian, 2015). El subtipo de histona H1.4 es una histona que se ha encontrado relacionada a heterocromatina que se ubica en regiones subteloméricas y teloméricas (Parseghian, 2015). Finalmente, en algunos organismos como *T. brucei* se ha predicho que las proteínas H1 están codificadas en una familia multigen de al menos cinco genes dispuestos en tándem. Dichas proteínas son esenciales para regular el espaciamiento de los nucleosomas y la condensación de la heterocromatina (Povelones, 2012).

También identificamos una proteína hipotética de mantenimiento estructural de los cromosomas (SMC) esencial para la transmisión exitosa de los cromosomas durante la replicación. Estas proteínas, además, funcionan junto con otras proteínas en la condensación cromosómica, la cohesión de las cromátidas hermanas, la recombinación, la reparación del DNA y el silenciamiento epigenético de la expresión génica debido a su interacción con la proteína CTCF (Harvey, 2002). Además, es un componente central que coordina las vías de recombinación homóloga (HR) para la reparación de roturas de doble cadena del DNA.

Adicionalmente, identificamos a una proteína del complejo Arp2/3 con el número de acceso C4M782, una longitud de secuencia de 293 aminoácidos y un porcentaje de cobertura de 10.92. Las proteínas Arp2/3 son proteínas de unión a actina nuclear 2 y 3 implicadas en procesos relacionados con el DNA como son la remodelación de la cromatina, la transcripción, la replicación y la reparación del DNA (Schrank, et. al., 2018). Es probable que todo el complejo esté involucrado

en un proceso que ayuda a regular la longitud telomérica (Ungar, 2009). Esta proteína se recluta en la cromatina dañada y promueve la reparación por recombinación homóloga dirigida (HDR) mediante la organización de la cromatina al generar dominios de reparación esenciales lo que conduce a mayor movilidad en sitios específicos de DSB. La polimerización de actina nuclear es necesaria para la migración de un subconjunto de DSBs en grupos discretos subnucleares durante HDR (Schrack, et. al., 2018).

Finalmente, nosotros identificamos menos proteínas en nuestros ensayos (89 y 29 a partir de extractos totales y nucleares, respectivamente) en comparación con el proteoma asociado a TRF1 realizado por el grupo de Giannone y colaboradores en el cual identificaron 328 y 88 proteínas en extractos totales y nucleares, respectivamente (Giannone et al., 2010). Es posible que esta diferencia sea un reflejo de las condiciones empleadas en el ensayo de pull down, que en nuestro caso fueron con alta astringencia (300mM de NaCl y 50 mN de Imidazol) obteniendo de esta manera solamente interacciones proteína-proteína fuertes.

Con los resultados obtenidos se plantea el siguiente modelo para la organización telomérica de *E. histolytica*.

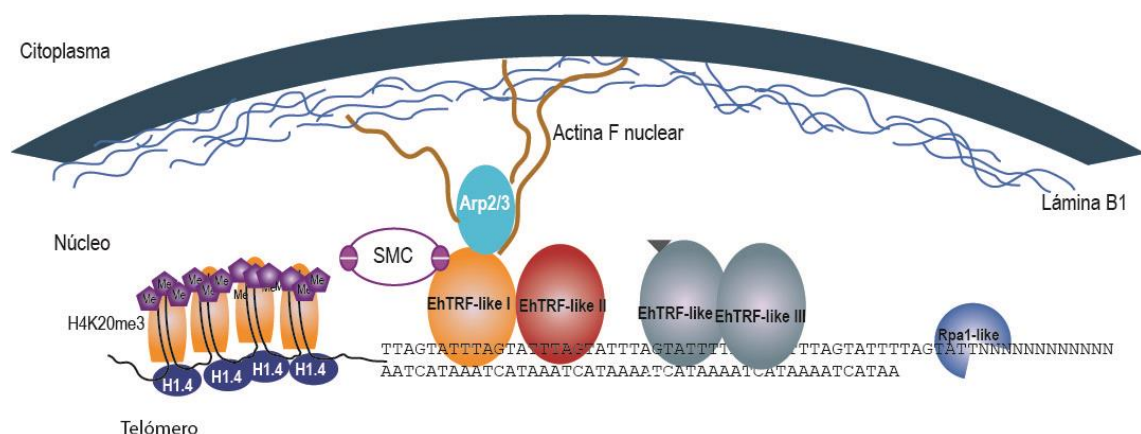


Fig. 46. Modelo de interacciones proteína-proteína en la región telomérica del parásito *E. histolytica*. En el modelo se observan a la proteína EhTRF-like I, interactuando mediante la formación de un heterodímero con EhTRF-like II, las cuales están unidas al DNA de los repetidos tipo STR. También se observa la interacción de la proteína Arp2/3 y la proteína SMC con EhTRF-like I. Por otro lado, podemos observar la presencia de la histona H1.4 involucrada en la formación de heterocromatina.

XII. Conclusiones

Se predijo la estructura terciaria de los dominios TRFH Y MDBD de la proteína EhTRF-like I de *E. histolytica*. El dominio de unión a DNA MYB presenta una estructura tridimensional con una conformación hélice-vuelta-hélice y los residuos de aminoácidos del motivo Telebox conservados. Para el caso del dominio TRFH se determinó una arquitectura enteramente helicoidal y la presencia de aminoácidos con características hidrofóbicas relevantes para la interacción proteína-proteína.

Se determinó que la proteína EhTRF-like I tiene una Kd de 6.8 nM para la secuencia telomérica EhTel y una Kd de 2.8 nM para la HsTel. Lo cual quiere decir la proteína EhTRF-like I presenta una mayor afinidad por la secuencia canónica humana.

Se identificaron proteínas asociadas a la maquinaria telomérica de *E. histolytica* como son la proteína Eh-TRF-like II, la proteína histona H1.4, la proteína de mantenimiento estructural de los cromosomas (SMC) y la proteína Arp2/3.

XIII. Bibliografía

1. Al-Hakim, A., Escribano-Diaz, C., Landry, M.C., O'Donnell, L., Panier, S., Szilard, R.K., Durocher, D. (2010). **The ubiquitous role of ubiquitin in the DNA damage response.** DNA Repair 9, 1229–1240.
2. Álvarez HV. (2018). **Interacción de la proteína EhTRF-like I con proteínas relacionadas con la función telomérica en *Entamoeba histolytica*.** Tesis de Maestría. UACM.
3. Arratia J., Aguirre J. (2013). **Los Factores de Transcripción tipo Myb, una Familia de Reguladores de la Diferenciación Celular Conservada en los Organismos Eucariontes.** TIP Revista Especializada en Ciencias Químico-Biológicas, 16(2):98-108
4. Bekker-Jensen, S., Mailand, N. (2011). **The ubiquitin- and SUMO-dependent signaling response to DNA double-strand breaks.** FEBS letters 585, 2914–2919.
5. Bertschi N. L., Toenhake C. G., Niederwieser I. *et. al.*, (2017). **Malaria parasites possess a telomere repeat-binding protein that shares ancestry with transcription factor IIIA.** Nat Microbiol. 13;2:17033. doi: 10.1038/nmicrobiol.2017.33.
6. Bianchi, A., Smith, S., Chong, L., Elias, P., & de Lange, T. (1997). **TRF1 is a dimer and bends telomeric DNA.** *The EMBO Journal*, 16(7), 1785–1794. <http://doi.org/10.1093/emboj/16.7.1785>
7. Bilaud T, Koering CE, Binet-Brasselet E, Ancelin K, Pollice A, Gasser SM, Gilson E. (1996). **The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human.** *Nucleic Acids Res. Apr 1; 24(7):1294-303*
8. Benarroch-Popivker D, Pisano S, Mendez-Bermudez A. *et. al.* (2016). **TRF2-Mediated Control of Telomere DNA Topology as a Mechanism for Chromosome-End Protection.** Mol Cell. Jan 21; 61(2):274-86.
9. Blackburn E. H., Epel E. S. Lin J. (2015). **Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection.** Science. Vol. 350, Issue 6265, pp. 1193-1198 DOI: 10.1126/science.aab3389

10. Boutou E., Vlachodimitropoulos D., Pappa V., et. al. (2013). **DNA Repair and Telomeres – An Intriguing Relationship**. New Research Directions in DNA Repair. <https://dx.doi.org/10.5772/56115341>
11. Broccoli D. de Lange T. et al. **Comparison of the human and mouse gene encoding the telomeric protein, TRF1: chromosomal localization, expression and conserved protein domains**. Human Mol. Genet. 1997; 6: 69-76
12. Broccoli D, Smogorzewska A, Chong L, de Lange T (1997). **Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2**. Nat. Genet. 17:231–235
13. Celli GB, de Lange T (2005). **DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion**. Nat. Cell Biol. 7:712–718
14. Cesare AJ, Reddel RR. **Alternative lengthening of telomeres: models, mechanisms and implications**. Nat Rev Genet. 2010;11(5):319-30. doi: 10.1038/nrg2763.
15. Choi, K. H., Farrell, A. S., Lakamp, A. S., & Ouellette, M. M. (2011). **Characterization of the DNA binding specificity of Shelterin complexes**. *Nucleic Acids Research*, 39(21), 9206–9223. <http://doi.org/10.1093/nar/gkr665>
16. Clark C. G., Ali I. K., Zaki M., Loftus B. J., Hall N. (2006). **Unique organisation of tRNA genes in *Entamoeba histolytica***. Mol Biochem Parasitol. 146, 24–29. 10.1016/j.molbiopara.2005.10.013
17. Clark C. G., Alsmark U. C., Tazreiter M., Saito-Nakano Y., Ali V., Marion S., et al. (2007). **Structure and content of the *Entamoeba histolytica* genome**. Adv. Parasitol. 65, 51–190. 10.1016/S0065-308X(07)65002-7
18. Coluzzi E., Leone S., Sgura A. (2019). **Oxidative Stress Induces Telomere Dysfunction and Senescence by Replication Fork Arrest**. Cells. doi: 10.3390/cells8010019.
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 cristal structures**. *EMBOReports*, 6(1), 39–45. <http://doi.org/10.1038/sj.embor.7400314>

20. da Silva, M. S., Perez, A. M., da Silveira, R., de Moraes, C. E., Siqueira-Neto, J. L., Freitas, L., & Cano, M. I. (2010). **The *Leishmania amazonensis* TRF (TTAGGG repeat-binding factor) homologue binds and co-localizes with telomeres.** *BMC microbiology*, 10, 136. doi:10.1186/1471-2180-10-136
21. d'Adda di Fagagna, F., Reaper, P. M., Clay-Farrace, L., Fiegler, H., Carr, P., Von Zglinicki, T., Saretzki, G., Carter, N. P. and Jackson, S. P. (2003). **A DNA damage checkpoint response in telomere-initiated senescence.** *Nature* 426, 194- 198. doi:10.1038/nature02118
22. de Lange T (2001). **Telomere capping—one strand fits all.** *Science* 292: 1075-1076.
23. de Lange T (2002). **Protection of mammalian telomeres.** *Oncogene*. 21:532–540
24. de Lange T (2005). **Shelterin: the protein complex that shapes and safeguards human telomeres.** *Genes & Dev* 19:2100-2110 doi:10.1101/gad.1346005
25. de Lange T (2010). **How shelterin solves the telomere end-protection problem.** *Cold Spring Harb Symp Quant Biol.* 2010;75:167-77. doi: 10.1101/sqb.2010.75.017.
26. Déjardin, J., & Kingston, R. E. (2009). **Purification of proteins associated with specific genomic Loci.** *Cell*, 136(1), 175–186. doi:10.1016/j.cell.2008.11.045
27. Diotti, R., & Loayza, D. (2011). **Shelterin complex and associated factors at human telomeres.** *Nucleus* (Austin, Tex.), 2(2), 119–135. doi:10.4161/nucl.2.2.15135
28. Fairall L., Chapman L., Moss H., et. al. (2001). **Structure of the TRFH Dimerization Domain of the Human Telomeric Proteins TRF1 and TRF2.** *Molecular Cell.* VOLUME 8, ISSUE 2, P351-361 DOI: [https://doi.org/10.1016/S1097-2765\(01\)00321-5](https://doi.org/10.1016/S1097-2765(01)00321-5)
29. Florence R. Wilson, Angus Ho, John R. Walker, Xu-Dong Zhu (2016). **Cdk-dependent phosphorylation regulates TRF1 recruitment to PML bodies and promotes C- circle production in ALT cells.** *J Cell Sci* 2016 129: 2559- 2572; doi: 10.1242/jcs.186098

30. Flores F. Eugenia, Mayani Hector (2003). **Cromosomas, control celular y cáncer: una cuestión de telomerasas**. Revista Ciencia. Academia Mexicana de Ciencias
31. Giannone, R. J., McDonald, H. W., Hurst, G. B., Shen, R. F., Wang, Y., & Liu, Y. (2010). **The protein network surrounding the human telomere repeat binding factors TRF1, TRF2, and POT1**. PloS one, 5(8), e12407. doi:10.1371/journal.pone.0012407
32. Giraud-Panis M. J., Pisano S., Benarroch-Popivker D., Pei B., Le Du M. H., Gilson E. (2013). **One identity or more for telomeres?** Front. Oncol. 3:48 10.3389/fonc.2013.00048
33. Giraud-Panis, M. J., Ye, J., & Gilson, E. (2018). **TRFH domain: at the root of telomere protein evolution?**. Cell research, 28(1), 7–8. doi:10.1038/cr.2017.152
34. Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., de Lange, T. (1999). **Mammalian telomeres end in a large duplex loop**. Cell 97, 503–514.
35. Guettler, S., LaRose, J., Petsalaki, E., Gish, G., Scotter, A., Pawson, T., Rottapel, R., Sicheri, F. (2011). **Structural basis and sequence rules for substrate recognition by Tankyrase explain the basis for cherubism disease**. Cell 147, 1340– 1354.
36. Hanaoka, S., Nagadoi, A., & Nishimura, Y. (2005). **Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA-binding activities**. *Protein science : a publication of the Protein Society*, 14(1), 119–130. doi:10.1110/ps.04983705
37. Harvey, S. H., Krien, M. J., & O'Connell, M. J. (2002). **Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases**. Genome biology, 3(2), REVIEWS3003. doi:10.1186/gb-2002-3-2-reviews3003
38. Helgeson, L. A., & Nolen, B. J. (2013). **Mechanism of synergistic activation of Arp2/3 complex by cortactin and N-WASP**. eLife, 2, e00884. doi:10.7554/eLife.00884
39. Hergeth, S. P., & Schneider, R. (2015). **The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle**. EMBO reports, 16(11), 1439–1453. doi:10.15252/embr.201540749

40. Hockemeyer D, Palm W, Else T, Daniels JP, Takai KK, Ye JZ, Keegan CE, de Lange T, Hammer GD (2007). **Telomere protection by mammalian Pot1 requires interaction with Tpp1**. *Nat. Struct. Mol. Biol.* 14:754–761.
41. Kedem, K., Chew, L. P., and Elber, R. (1999). **Unit-vector RMS (URMS) as a tool to analyze molecular dynamics trajectories**. *Proteins*, 37, 554-64.
42. Ko, E. R., Ko, D., Chen, C., & Lipsick, J. S. (2008). **A conserved acidic patch in the Myb domain is required for activation of an endogenous target gene and for chromatin binding**. *Molecular cancer*, 7, 77. doi:10.1186/1476-4598-7-77
43. König P, Giraldo R, Chapman L, Rhodes D (1996) **The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA**. *Cell* 85: 125–136
44. Klug, W. S., Cummings M. R., Spencer C. A. (2006). **Conceptos de Genética**. 8va edición. Editorial Pearson Educación S. A. Madrid, España.
45. Levy, M. Z., Allsopp, R. C., Futcher, A. B., Greider, C. W. and Harley, C. B. (1992). **Telomere end-replication problem and cell aging**. *J. Mol. Biol.* 225, 951-960. doi:10.1016/0022-2836(92)90096-3
46. Li, B; Espinal, A; & Croos, G. A. (2005). **Trypanosoma telomeres are protected by a homologue of mammalian TRF2**. *Molecular and Cellular Biology*, 25(12), 5011-5021
47. Liu D., O'Connor, M.S., Qin, J., and Songyang, Z. (2004). **Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins**. *J. Biol. Chem.* 279: 51338-51342.
48. Loftus B., et. al. (2005). **The genome of the protist parasite *Entamoeba histolytica***. *Nature*. 24;433(7028):865-8.
49. Meneses E., et. al. (2010). **The R2R3 Myb protein family in *Entamoeba histolytica***. *Gene* 1;455(1-2):32-42. doi: 10.1016/j.gene.2010.02.004.
50. Nandakumar, J., & Cech, T. R. (2013). **Finding the end: recruitment of telomerase to telomeres**. *Nature Reviews Molecular Cell Biology*, 14(2), 69–82.
51. Nishikawa T, Nagadoi A, Yoshimura S., et. al. (1998). **Solution structure of the DNA-binding domain of human telomeric protein, hTRF1**. VOLUME 6, ISSUE 8, P1057-1065 DOI:[https://doi.org/10.1016/S0969-2126\(98\)00106-3](https://doi.org/10.1016/S0969-2126(98)00106-3)

52. Nishikawa T., Okamura, H., Nagadol, A. (2001). **Solution structure of a telomeric DNA complex of human TRF1.** *CellPres.* Volume 9, ISSUE 12, P1237-1251 DOI: [https://doi.org/10.1016/S0969-2126\(01\)00688-8](https://doi.org/10.1016/S0969-2126(01)00688-8)
53. Oganessian, L., & Karlseder, J. (2009). **Telomeric armor: the layers of end protection.** *Journal of Cell Science*, 122(22), 4013–4025. <http://doi.org/10.1242/jcs.050567>
54. Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Nishimura, Y. (1992). **Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core.** *Proceedings of the National Academy of Sciences of the United States of America*, 89(14), 6428–6432. doi:10.1073/pnas.89.14.6428
55. Ogata K., Nishimura Y., et al. **Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices.** *Cell*. 1994; 79 (95042761): 639-648
56. O'Sullivan, R. J., & Karlseder, J. (2010). **Telomeres: protecting chromosomes against genome instability.** *Nature Reviews. Molecular Cell Biology*, 11(3), 171–181. <http://doi.org/10.1038/nrm2848>
57. Palm, W; & de Lange, T. (2008). **How shelterin protects mammalian telomeres.** *Annual review of genetics*, 42, 301-334
58. Parseghian M. H. (2015). **What is the role of histone H1 heterogeneity? A functional model emerges from a 50 year mystery.** *AIMS biophysics*, 2(4), 724–772. doi:10.3934/biophy.2015.4.724
59. Petri Jr WA, Haque R. (2017). **Entamoeba Species, Including Amebic Colitis and Liver Abscess, Infectious Disease.** <https://oncohemakey.com/entamoeba-species-including-amebic-colitis-and-liver-abscess/>
60. Peska V., Prochazkova S. P., Fajkus J. (2011). **Using the Telebox to Search for Plant Telomere Binding Proteins.** *Current Protein and Peptides Science*, 12, 75-83
61. Pfeiffer, V., & Lingner, J. (2013). **Replication of telomeres and the regulation of telomerase.** *Cold Spring Harbor perspectives in biology*, 5(5), a010405. doi:10.1101/cshperspecta.010405

62. Povelones, M. L., Gluenz, E., Dembek, M., Gull, K., & Rudenko, G. (2012). **Histone H1 plays a role in heterochromatin formation and VSG expression site silencing in *Trypanosoma brucei***. *PLoS pathogens*, 8(11), e1003010. doi:10.1371/journal.ppat.1003010
63. Procházková Schruppfová, P., Schořová, Š., & Fajkus, J. (2016). **Telomere- and Telomerase-Associated Proteins and Their Functions in the Plant Cell**. *Frontiers in plant science*, 7, 851. doi:10.3389/fpls.2016.00851
64. Pumarola A., et. al. (1991). **Microbiología y Parasitología Médica**. Ediciones Científicas y Técnicas. 2da Edición, Barcelona, España.
65. Rendón-Gandarilla, F. J., Álvarez-Hernández, V., Castañeda-Ortiz, et. al., (2018). **Telomeric Repeat-Binding Factor Homologs in *Entamoeba histolytica*: New Clues for Telomeric Research**. *Frontiers in cellular and infection microbiology*, 8, 341. doi:10.3389/fcimb.2018.00341
66. Rodríguez Centeno, J., Manguán-García, C., Perona, R., & Sastre, L. (2019). **Structure of Dictyostelium discoideum telomeres. Analysis of possible replication mechanisms**. *PloS one*, 14(9), e0222909. doi:10.1371/journal.pone.0222909
67. Romero R. (1993). **Microbiología y Parasitología Humana**. Editorial Médica Panamericana, D. F. México.
68. Schrank, B. R., Aparicio, T., Li, Y., et. al., (2018). **Nuclear ARP2/3 drives DNA break clustering for homology-directed repair**. *Nature*, 559(7712), 61–66. doi:10.1038/s41586-018-0237-5
69. Ungar, L., Yosef, N., Sela, Y., et. al., (2009). **A genome-wide screen for essential yeast genes that affect telomere length maintenance**. *Nucleic acids research*, 37(12), 3840–3849. doi:10.1093/nar/gkp259
70. Uzlíková M., Fulnečková J., Weisz F., et. al., (2017). **Characterization of telomeres and telomerase from the single-celled eukaryote *Giardia intestinalis***. *Mol Biochem Parasitol*. 211:31-38. doi: 10.1016/j.molbiopara.2016.09.003.
71. Verdun R, Karlseder J. (2006). **The DNA Damage Machinery and Homologous Recombination Pathway Act Consecutively to Protect Human Telomeres**. *CellPress Nov 17*;127(4):709-20. DOI: 10.1016/j.cell.2006.09.034

72. Walker J. R; Zhu Xu-Dong (2012). **Post-translational modifications of TRF1 and TRF2 and their roles in telomere maintenance.** Mechanism of Agein and Development. Elsevier Lieland Ltd. 133 (2012) 421 – 434
73. Xin, H., Liu, D., & Songyang, Z. (2008). **The telosome/shelterin complex and its functions.** Genome biology, 9(9), 232. doi:10.1186/gb-2008-9-9-232
74. Zhang X, Rosenstein BS, Wang Y, Lebowitz M, Wei H. (1997) **Identification of possible reactive oxygen species involved in ultraviolet radiation-induced oxidative DNA damage.** Free Radic Biol Med. 23(7):980-5.
75. Zhang Y, Skolnick J (2005). **TM-align: a protein structure alignment algorithm based on the TM-score.** Nucleic Acids Res.;33(7):2302–2309. Published 2005 Apr 22. doi:10.1093/nar/gki524