

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

COLEGIO DE CIENCIA Y TECNOLOGÍA

DOCTORADO EN CIENCIAS GENÓMICAS

**Identificación del repertorio de proteínas involucradas en la
degradación del RNAm y determinación de su localización celular
en *Entamoeba histolytica***

TESIS PARA OBTENER EL GRADO DE

DOCTORA EN CIENCIAS GENÓMICAS

P R E S E N T A:

ITZEL LÓPEZ ROSAS

Director de Tesis:

Dr. Mario César López Camarillo

México, D.F. Junio 2013.

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.

COMITÉ TUTORIAL

DIRECTOR

Dr. Mario César López Camarillo
Laboratorio de Oncogenómica y Proteómica del Cáncer.
Posgrado en Ciencias Genómicas
Universidad Autónoma de la Ciudad de México

CODIRECTORA

Dra. Laurence A. Marchat
Profesora titular del Programa Institucional en Biomedicina Molecular y Red
en Biotecnología.
Escuela Nacional de Medicina y Homeopatía, Instituto Politécnico Nacional

ASESORES

Dra. Mavil López Casamichana
Profesora Investigadora. Posgrado en Ciencias Genómicas
Universidad Autónoma de la Ciudad de México

Dr. José de Jesús Olivares Trejo
Profesor Investigador. Posgrado en Ciencias Genómicas
Universidad Autónoma de la Ciudad de México

Dr. Máximo Martínez Benítez
Laboratorio de Biología Molecular del CIPRE

Dr. César Sandino Reyes López
Profesor titular del Programa Institucional en Biomedicina Molecular y Red
en Biotecnología.
Escuela Nacional de Medicina y Homeopatía, Instituto Politécnico Nacional

LECTORES DE TESIS

Dra. Laurence A. Marchat

Profesora titular del Programa Institucional en Biomedicina Molecular y Red
en Biotecnología

Escuela Nacional de Medicina y Homeopatía, Instituto Politécnico Nacional

Dra. Mavil López Casamichana

Profesora Investigadora. Posgrado en Ciencias Genómicas

Universidad Autónoma de la Ciudad de México

Dra. María Elizabeth Álvarez Sánchez

Profesora Investigadora. Posgrado en Ciencias Genómicas

Universidad Autónoma de la Ciudad de México

Dr. Israel López Reyes

Centro de Diagnóstico y Vigilancia Epidemiológica del Distrito Federal (CDE-DF)

Universidad Autónoma de la Ciudad de México

AGRADECIMIENTOS

El proyecto de investigación descrito en esta tesis contó con el respaldo académico de la Universidad Autónoma de la Ciudad de México, a la cual agradezco profundamente.

Esta tesis forma parte del Proyecto multilateral M08-S02 “Papel de las proteínas *EhPC4* y *EhCstF64* en la virulencia de *Entamoeba histolytica*”. Apoyado por México-Francia SEP-CONACyT-ANUIES; ECOS-NORD

Agradezco al Instituto de Ciencia y Tecnología del Distrito Federal (ICyT-DF) el apoyo otorgado mediante el proyecto: ICyT 321 “Implementación de una plataforma de análisis proteómico aplicable al estudio de enfermedades humanas, infecciosas y de importancia veterinaria en el Distrito Federal.” De igual manera agradezco por la beca Creación Joven otorgada por el ICyT-DF, México. Para la realización del proyecto: “Identificación del repertorio de proteínas involucradas en la degradación del RNAm y determinación de su localización celular en *Entamoeba histolytica*.”

El trabajo experimental plasmado en esta tesis, fue llevado a cabo en el laboratorio 2 del Posgrado en Ciencias Genómicas, de la Universidad Autónoma de la Ciudad de México, bajo la tutoría del Dr. Mario César López Camarillo. También contamos con el apoyo del laboratorio de la Dra. Esther Orozco del Departamento de Infectómica y Patogénesis Molecular del CINVESTAV-IPN.

"Las ciencias tienen las raíces amargas,

pero muy dulces los frutos. "

Aristóteles

LISTA DE ABREVIATURAS

3' UTR	región 3' no traducida
5' UTR	región 5' no traducida
ActD	actinomicina D
BSA	albúmina sérica de bovino
cDNA	ácido desoxirribonucleico complementario
CHX	cicloheximida
DAPI	4'-6-diamidino-2-fenilindola
DCP1	proteína de <i>decapping</i> 1
DCP2	proteína de <i>decapping</i> 2
DEPC	dietilpirocarbonato
DMSO	dimetilsulfóxido
DNA	ácido desoxirribonucleico
DTT	ditiotreitól
dNTPs	deoxinucleótidos tri-fosfatados
ddTPs	dideoxinucleótidos tri-fosfatados
dT	deoxitimidina
dA	deoxiadenina
dC	deoxicitosina
dG	deoxiguanina
EDTA	ácido etilen diamino tetra-acético
EMSA	ensayo de retardamiento de movilidad electroforética
FISH	hibridación fluorescente <i>in vitro</i>
GST	proteína glutatión S-transferasa

h	horas
HEPES	ácido 4-(2-hydroxietil)-1-piperazinaetanosulfónico
HCl	ácido clorhídrico
IPTG	isopropil- β -D-tiogalactopiranósido
Kb	kilobases
KCl	cloruro de potasio
kDa	kilodaltones
LB	medio de cultivo para bacterias Luria-Broth
M	molar
mM	milimolar
μ M	micromolar
min	minuto
miRNA	microRNA
mg	miligramo
μ g	microgramo
ml	mililitro
μ l	microlitro
NaCl	cloruro de sodio
NaOH	hidróxido de sodio
nm	nanomolar
nt	nucleótidos
ng	nanogramo
ORF	marco de lectura abierto

PABP	proteína de unión al tracto de adeninas
PAGE	electroforesis en geles de poliacrilamida
PCR	reacción en cadena de la polimerasa
pb	pares de bases
PBS	solución salina amortiguadora de fosfatos
pm	picomolar
pre-RNA _m	RNA mensajero precursor
RNA	ácido ribonucleico
RNA _m	ácido ribonucleico mensajero
RNA poli (A)	ácido ribonucleico mensajero poliadenilado
rpm	revoluciones por minuto
RT	transcripción reversa
SDS	dodecil sulfato de sodio
SINAVE	Sistema Nacional de Vigilancia Epidemiológica
siRNA	RNA pequeño de interferencia
seg	segundo
T.A	temperatura ambiente
TAE	solución amortiguadora Tris-Acetato-EDTA
TBE	solución amortiguadora Tris-Borato-EDTA
TBP	proteína de unión a la caja TATA
TEMED	N,N,N',N', tetrametilen-diamina
Tris	tris (hidroximetil)-amino-metano
TYI-S-33	medio de tripticasa, extracto de levadura, hierro y suero bovino para cultivo de <i>E. histolytica</i>

U	unidades
V	voltios

LISTA DE TABLAS

Tabla 1. Número de casos nuevos y tasa de incidencia reportados de amibiasis por entidad federativa para el año 2009.

Tabla 2. Superfamilia de exoribonucleasas identificadas en eubacterias y eucariontes.

Tabla 3. Factores de degradación de RNAm identificados en *P-bodies* en organismos eucariontes

Tabla 4. Oligonucleótidos utilizados en el ensayo de RT-PCR cuantitativa

Tabla 5. Proteínas de degradación predichas en *Entamoeba histolytica*

Tabla 6. Comparación de *EhXRN2* con proteínas relacionadas de diferentes organismos

Tabla 7. Comparación de *EhDCP2* con proteínas relacionadas de diferentes organismos

Tabla 8. Proteínas identificadas que interactúan con la proteína *EhCAF1* de *E. histolytica*

LISTA DE FIGURAS

Figura 1. Número de casos reportados de amibiasis por entidad federativa.

Figura 2. Ciclo de vida de *E. histolytica*.

Figura 3. Vías de degradación del RNAm en eucariontes.

Figura 4. Esquema de la conformación del complejo CCR4-NOT.

Figura 5. Maquinarias de degradación del RNAm con las proteínas predichas para *E. histolytica*.

Figura 6. Perfil de expresión de los genes involucrados en la degradación del RNAm en *E. histolytica*.

Figura 7. Alineamiento Múltiple de la secuencia de aminoácidos de proteínas

Figura 8. Alineamiento Múltiple de la secuencia de aminoácidos de proteínas DCP2.

Figura 9. Relaciones filogenéticas entre *EhXRN2* y miembros de la familia XRN2.

Figura 10. Relaciones filogenéticas entre *EhDCP2* y miembros de la familia DCP2.

Figura 11. Comparación de las proteínas XRN2 y DCP2 de organismos eucariontes y *E. histolytica*.

Figura 12. Expresión e inmunolocalización de las proteínas *EhXRN2*, *EhDCP2* y *EhCAF1*.

Figura 13. Colocalización de los factores de degradación del RNAm con *EhCAF1* en *foci* citoplasmáticos.

Figure 14. Determinación del número de *foci* citoplasmáticos que contienen a *EhCAF1* después de los tratamientos con actinomicina D, cicloheximida y puromicina.

Figura 15. Sustrato RNA Poli(A⁺) RNA y dsRNA colocalizan con *EhCAF1* y *EhAGO2-2* en *foci* citoplasmáticos.

Figura 16. *EhCAF1* forma complejos con proteínas de degradación de RNA, citoesqueleto y de metabolismos celular

Figura 17. *EhCAF1* tiene actividad de desadenilasa.

RESUMEN

En eucariontes superiores la degradación del RNAm y el silenciamiento de genes basado en RNA ocurren en *foci* citoplasmáticos conocidos como cuerpos de procesamiento de RNAm (*P-bodies*). Hasta la fecha la existencia y el potencial papel de los *P-bodies* en la degradación del RNAm siguen siendo poco conocidos en parásitos protozoarios, por lo tanto la identificación de las estructuras citoplasmáticas similares podrían ayudar a comprender los orígenes evolutivos de la degradación del RNAm en eucariontes inferiores. En el presente trabajo, realizamos estudios de inmunofluorescencia y microscopia confocal para investigar la localización de proteínas de degradación de RNAm de *Entamoeba histolytica* y encontrar evidencia acerca de la existencia de *P-bodies* en este parásito intestinal humano. Las proteínas de degradación llamadas exoribonucleasa *EhXRN2* y enzima de *decapping* *EhDCP2* fueron localizadas en *foci* citoplasmáticos en un patrón particular que asemejan la organización de los *P-bodies*. Sin embargo, los *foci* de amiba son más pequeños y menos redondeados que los descritos para eucariontes superiores, por lo tanto los hemos denominado estructuras tipo *P-bodies*. Adicionalmente los *foci* contienen otros factores de degradación de RNAm tales como la desadenilasa *EhCAF1* y la proteína *EhAGO2-2*, que participa en el RNA de interferencia. Análisis bioquímicos revelaron que *EhCAF1* coinmunoprecipita con *EhXRN2*, pero no con *EhDCP2* ni con *EhAGO2-2*, ligando así la desadenilación con la degradación de RNA en dirección 5'-3'. Por otra parte observamos que el número de *foci* que contienen a *EhCAF1* disminuye significativamente después de inhibir la transcripción y la traducción usando actinomicina D y cicloheximida, respectivamente. Por otra

parte, mediante ensayos de hibridación *in situ* (RNA-FISH), encontramos que *EhCAF1* colocaliza con RNAs poli(A)⁺, mientras que durante el silenciamiento del gen *Ehpc4* por RNA de interferencia la proteína *EhAGO2-2* colocaliza con pequeños RNAs de interferencia específicos para PC4 en *foci* citoplasmáticos similares. Nuestros resultados indican que la localización de proteínas que participan en *decapping*, desadenilación y RNA de interferencia en estructuras citoplasmáticas es conservada evolutivamente, sugiriendo que estas estructuras fueron originadas en la evolución temprana del linaje eucarionte. Hasta nuestro conocimiento, este es el primer reporte sobre el estudio de la localización de las proteínas de degradación de RNAm en estructuras tipo *P-bodies* en *E. histolytica*, lo cual abre la oportunidad para entender los mecanismos de degradación del RNAm y el silenciamiento de genes basado en RNA en eucariontes inferiores.

ABSTRACT

In higher eukaryotes mRNA degradation and RNA-based gene silencing occur in cytoplasmic *foci* referred to as mRNA processing bodies (P-bodies). To date the existence and potential role of P-bodies in mRNA decay remain poorly understood in protozoan parasites, thus identification of similar cytoplasmic structures could help to understand the evolutionary origins of mRNA degradation in lower eukaryotes. Here, we performed immunofluorescence and confocal microscopy studies to investigate the localization of *Entamoeba histolytica* mRNA degradation proteins and found evidence about the existence of P-bodies in this human intestinal parasite. Two mRNA decay factors namely the *EhXRN2* exoribonuclease and *EhDCP2* decapping enzyme were both localized in cytoplasmic *foci* in a particular pattern that resemble P-body organization. However, amoeba *foci* are smaller and less rounded to those described in higher eukaryotes, thus we have dubbed them as P-body like structures. The *foci* contained additional mRNA degradation factors including the *EhCAF1* deadenylase and *EhAGO2-2* protein involved in RNA interference. Biochemical analysis revealed that *EhCAF1* coimmunoprecipitated with *EhXRN2*, but not with *EhDCP2* neither *EhAGO2-2*, linking deadenylation to 5'-to-3' mRNA decay. In addition, the number of *EhCAF1*-containing *foci* significantly decreased after transcription and translation inhibition using actinomycin D and cycloheximide. Moreover, by using RNA-FISH assays we found that *EhCAF1* colocalized with poly(A)⁺ RNA, whereas during silencing of *Ehpc4* gene by RNA interference *EhAGO2-2* colocalized with specific small interfering RNAs in similar cytoplasmic *foci*. Our results indicate that location of decapping, deadenylation and RNA interference proteins in P-body like *foci* is

evolutionary conserved suggesting that these structures were originated in the early evolution of eukaryotic lineage. To our knowledge, this is the first report on localization of mRNA decay proteins in P-body like structures in *E. histolytica*, which opens up opportunities for untangling the mechanisms of mRNA degradation and RNA-based gene silencing in deep-branching eukaryotes.

Índice

I. INTRODUCCIÓN	4
1. ANTECEDENTES GENERALES DEL TEMA.....	4
1.1. Amibiasis	4
1.2. Epidemiología de la amibiasis	5
1.3. Morfología de <i>E. histolytica</i>	9
1.4. Ciclo de vida de <i>E. histolytica</i>	10
1.5. Regulación de la expresión génica en eucariontes	15
1.6. Degradación de RNAm en eucariontes	17
1.7. Proteínas que participan en la de degradación del RNAm	20
1.8. Degradación del RNAm y cuerpos de procesamiento (P-bodies)	23
2. ESTADO DEL ARTE	27
2.1. Regulación de la expresión génica en parásitos protozoarios.....	27
2.2. Mecanismos de degradación de RNAm en parásitos protozoarios	34
3. ANTECEDENTES PARTICULARES DEL PROYECTO.....	36
4. IMPORTANCIA DE CONTINUAR EL ESTUDIO DEL TEMA.	37
OBJETIVO GENERAL.....	39
OBJETIVOS ESPECIFICOS	39
III. MATERIALES Y MÉTODOS	40
1. Análisis in silico del repertorio de proteínas involucradas en la degradación del RNAm.....	40
2. Cultivos de trofozoítos de <i>E. histolytica</i>	41
2.1. Choque térmico de los trofozoítos a 42° C	42
2.2. Irradiación de los trofozoítos con luz UV-C	42
2.3. Tratamiento con nitropusato de sodio.....	42

3. Obtención del DNA total de <i>E. histolytica</i> cepa HMI:IMSS	42
4. Diseño, síntesis y cuantificación de los oligonucleótidos.....	43
5. Análisis del RNA.....	46
5.1. Extracción y cuantificación de RNA total de alta calidad	46
5.2. Síntesis de la cadena complementaria de DNA (cDNA).....	47
5.3. Ensayos de RT-PCR cuantitativa	49
6. Clonación de los genes <i>Ehdcp2</i> , <i>Ehxr2</i> y <i>Ehago2-2</i> en el vector de expresión pRSET-A	50
6.1. Amplificación de los genes a partir de DNA genómico de <i>E. histolytica</i> ...	50
6.2. Ligación de los genes en el vector pRSET-A	51
7. Expresión y purificación de las proteínas recombinantes.....	52
7.1. Preparación de las células competentes <i>E. coli</i> BL21 (DE3) pLysS.....	52
7.2. Transformación de células competentes mediante choque térmico	52
7.3. Expresión de las proteínas recombinantes.....	53
7.4. Cuantificación de proteínas por el método de Bradford	53
7.5. Electroforesis de proteínas en geles de poliacrilamida en condiciones desnaturalizantes	54
7.6. Tinción de los geles de poliacrilamida con azul de Coomassie.....	56
7.7. Transferencia de proteínas a membranas de nitrocelulosa	56
7.8. Ensayos de Western blot.....	57
8. Purificación de las proteínas recombinantes y generación de anticuerpos	57
9. Inmunolocalización de las proteínas endógenas en trofozoítos	58
10. Ensayos de inmunoprecipitación	58
11. Expresión de los dsRNA y transfección de trofozoítos	59
12. Hibridación <i>in situ</i> fluorescente de RNA	60

13. Preparación de los sustratos para el ensayo de desadenilación	61
14. Ensayo de desadenilación in vitro	61
IV. RESULTADOS.....	62
1. Análisis in silico del repertorio de proteínas involucradas en la degradación de RNAm.....	62
2. Los genes de degradación del RNAm se expresan diferencialmente en condiciones de estrés.....	66
3. <i>EhXRN2</i> y <i>EhDCP2</i> conservan la arquitectura típica de las proteínas 5'-3' exoribonucleasas y las de decapping.....	68
4. <i>EhXRN2</i> , <i>EhDCP2</i> y <i>EhCAF1</i> son proteínas citoplasmáticas	78
5. <i>EhXRN2</i> , <i>EhDCP2</i> y <i>EhAGO2-2</i> colocalizan en los foci citoplasmáticos que contienen a <i>EhCAF1</i>	82
6. La formación de los <i>foci</i> citoplasmáticos depende de la transcripción activa .	86
7. La inhibición de la traducción afecta la formación de los <i>foci</i> citoplasmáticos	86
8. <i>EhCAF1</i> y <i>EhAGO2-2</i> colocalizan con RNA poli(A) ⁺ y dsRNA.....	90
9. <i>EhCAF1</i> forma complejos con proteínas de degradación de RNA, citoesqueleto y de metabolismo celular.....	94
10. <i>EhCAF1</i> tiene actividad de desadenilasa	98
V. DISCUSIÓN	100
VI. CONCLUSIONES	107
VII. REFERENCIAS	109
VIII. APÉNDICE.....	122

I. INTRODUCCIÓN

1. ANTECEDENTES GENERALES DEL TEMA

1.1. Amibiasis

La amibiasis es la infección del tracto gastrointestinal causada por el protozooario entérico *Entamoeba histolytica*. Amiba es su nombre común, este parásito es capaz de invadir la mucosa intestinal y puede dispersarse a otros organismos, principalmente al hígado (Espinosa-Cantellano *et al.*, 2000). Las amibas presentan dos fases morfológicamente distinguibles durante su ciclo de vida: i) el quiste, infectivo, inmóvil y resistente a cambios medioambientales y ii) el trofozoíto, móvil, sensible a los cambios ambientales y capaz de invadir tejidos y fagocitar células del hospedero. Este microorganismo vive usualmente como comensal en el lumen del intestino grueso del hombre, provocando amibiasis luminal benigna. Sin embargo, bajo ciertas condiciones del hospedero que incluyen la desnutrición, la presencia de un estado inmunológico débil y otras causas no bien entendidas aun, el parásito es capaz de invadir la mucosa intestinal y provocar la disentería amibiana. De manera interesante, estos eventos correlacionan con cambios en la expresión de genes de virulencia de los trofozoítos invasivos. Una vez invadido el epitelio intestinal, el parásito puede diseminarse a través de la sangre y tejidos, dando origen a la amibiasis invasiva. donde los trofozoítos son capaces de causar lesiones extraintestinales, principalmente en el hígado y con menos frecuencia en el pulmón, cerebro, piel, órganos genitales, bazo y riñón (Martínez-Palomo, 1982).

Esta infección prevalece principalmente en los países en vías de desarrollo, donde las condiciones socioeconómicas y de sanidad e higiene son deficientes. Aunado a la desnutrición de la población se propicia un estado inmunológico deficiente en los individuos, los cuales constituyen factores de predisposición y riesgo muy importantes frente a las infecciones, en especial a aquellas producidas por protozoarios parásitos que habitan en el tracto gastrointestinal como lo es *E. histolytica* (Cook, 1980; Martinez-Palomo, 1987).

1.2. Epidemiología de la amibiasis

Las infecciones en humanos producidas por *E. histolytica* presentan una alta incidencia, convirtiendo a la amibiasis humana en uno de los problemas más serios de salud a escala mundial. La infección con este parásito protozoario se estima que ocurre en 50 millones de personas en el mundo, de las cuales mueren al año aproximadamente de 40, 000 a 100, 000 (Walsh, 1986; Gaucher *et al.*, 2002; Stanley, 2003).

La infección inicia por la ingestión de alimentos y agua contaminados con quistes y se considera que los seres humanos son la única fuente directa o indirecta de transmisión del parásito ya sea por medio de las manos, los alimentos y las bebidas contaminadas con heces humanas. La mayor parte de los individuos infectados son portadores asintomáticos. Su alta prevalencia es el resultado de condiciones socioeconómicas precarias, así como estándares de higiene y sanidad deficientes, lo cual es muy frecuente en las comunidades de los países en vías de desarrollo. Aproximadamente, el 10% de la población mundial es

portadora del parásito aunque solo un bajo porcentaje sufre la enfermedad (W.H.O, 1997). En algunos países, la amibiasis se coloca entre las diez primeras causas de muerte, siendo la tercera enfermedad parasitaria de mayor importancia en el ámbito mundial, precedida únicamente por la malaria y la esquistosomiasis (Walsh, 1986). Los estudios serológicos usados para estimar la proporción de la población que ha tenido contacto con el parásito o bien que sufre de amibiasis actualmente, sugieren que uno de cada diez individuos infectados presenta invasión de la mucosa intestinal o hepática (Walsh, 1986).

Las estadísticas de enfermedades gastrointestinales reportadas por el Sistema Nacional de Vigilancia Epidemiológica (SINAVE) son enfermedades estadísticamente relevantes para la salud pública de México entre éstas: absceso hepático amibiano, amibiasis intestinal, fiebre tifoidea, giardiasis, intoxicación alimentaria bacteriana, paratifoidea y otras salmonelosis, shigelosis, infección intestinal debida a virus y otros organismos y otras infecciones intestinales debido a protozoarios. En México, la amibiasis es una de las causas más frecuentes de morbilidad, ocupando el 5° lugar en padecimientos transmisibles su tasa de incidencia general refleja la problemática que México tiene con esta patología (472.2 casos nuevos por cien mil habitantes en 2009) (SINAVE, 2000). Es una enfermedad endémica considerada como un serio problema de salud pública ya que se ha estimado que existe una proporción de un paciente con amibiasis invasiva por cada 4 o 5 portadores asintomáticos (Sepulveda, 1982). En la población mexicana total, el 20% son portadores, el 2% se encuentran enfermos y entre 0.1-0.2% de los enfermos muere (Conde-Bonfil *et al.*, 1992).

Realizando un estudio de hemaglutinación indirecta con anticuerpos específicos dirigidos contra *E. histolytica* se definió que el 8.4% de la población examinada seropositiva, siendo la zona norte de México donde se presento menor incidencia de amibiasis. Los resultados señalan una fuerte relación entre la frecuencia de amibiasis y factores económico-sociales tales como el hacinamiento, la carencia de servicios públicos, los malos hábitos higiénicos y el analfabetismo (Caballero-Salcedo *et al.*, 1994). El Boletín Epidemiológico publicado por la Secretaría de Salud (<http://www.ssa.gob.mx>), reportó los casos acumulados de amibiasis intestinal y abscesos hepáticos amibianos por entidad federativa para el periodo 2000-2009. Se puede observar que la tasa de incidencia general refleja la problemática que México tiene con este padecimiento (472.2 casos nuevos por cien mil habitantes en 2009), con un rango entre las entidades federativas de entre 1 203.9 que obtuvo la entidad con mayor tasa de incidencia, Oaxaca, y 108.6 en la menor, Baja California. En comparación por sexo, fueron las mujeres las más afectadas con una razón de 1:3 (Figura 1). En el último decenio, la incidencia de amibiasis intestinal ha ido decreciendo de presentar una tasa de 1 370.1 casos por cien mil habitantes en el año 2000, a 472.2 en 2009 (Tabla 1).

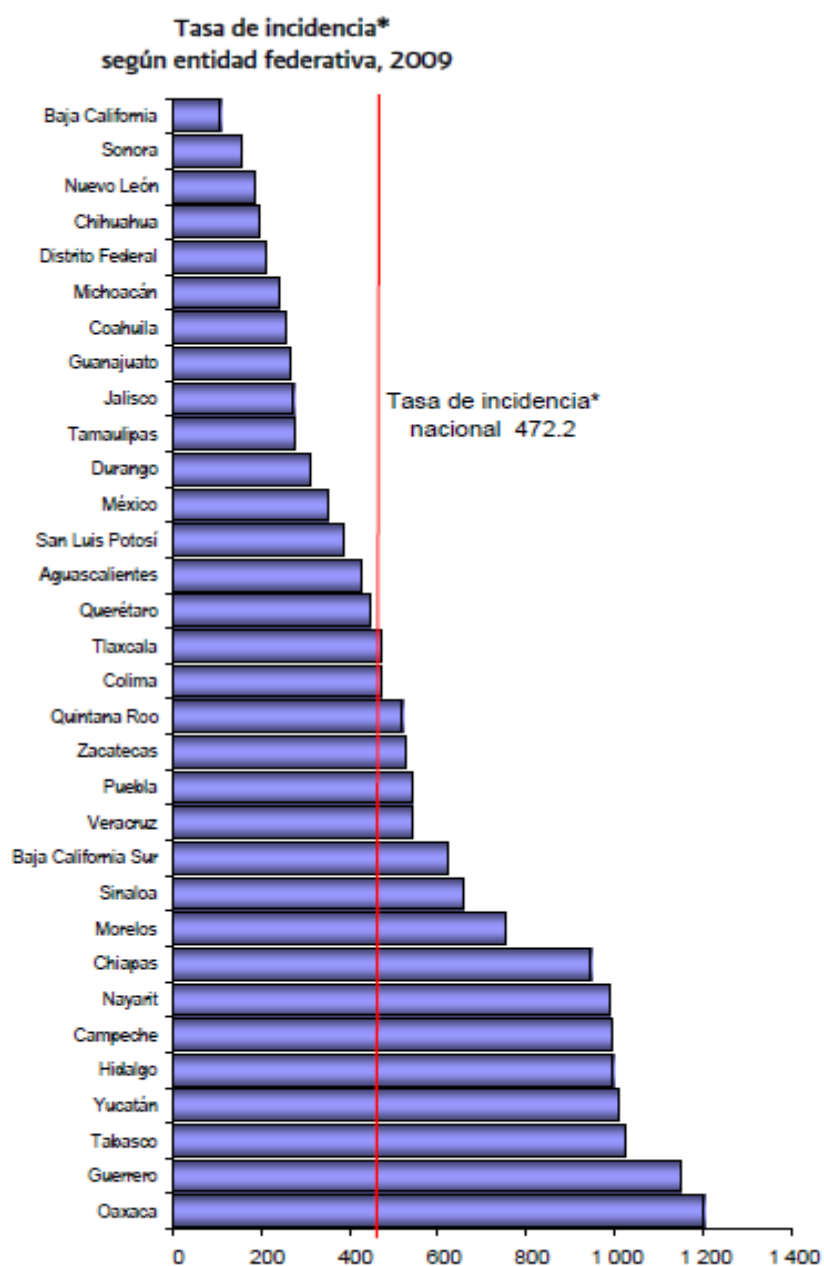


Figura 1. Número de casos reportados de amebiasis por entidad federativa.

Se muestran los casos acumulados de 2000 a 2009, la incidencia de amebiasis intestinal ha ido decreciendo de presentar una tasa de 1370.1 casos por cien mil habitantes en el año 2000, a 472.2 en 2009. Tomado del Boletín Epidemiológico de la Secretaría de Salud.

Casos nuevos y tasas de incidencia* según entidad federativa, 2009

Entidad Federativa	General		Masculino		Femenino	
	Casos	Tasa de incidencia*	Casos	Tasa de incidencia*	Casos	Tasa de incidencia*
Aguascalientes	4 862	425.8	2 150	387.2	2 712	462.2
Baja California	3 437	108.6	1 523	92.2	1 914	126.5
Baja California Sur	3 531	624.5	1 637	552.5	1 894	703.8
Campeche	7 922	995.2	3 723	945.5	4 199	1 043.9
Coahuila	6 757	257.0	2 988	227.5	3 769	286.6
Colima	2 847	473.8	1 293	428.6	1 554	519.3
Chiapas	42 685	947.0	18 250	824.6	24 435	1 065.1
Chihuahua	6 692	197.3	2 927	169.5	3 765	226.1
Distrito Federal	18 840	213.1	7 877	184.3	10 963	240.0
Durango	4 805	309.9	2 200	288.6	2 605	330.5
Guanajuato	13 477	267.1	5 890	245.4	7 587	286.9
Guerrero	36 118	1 150.1	15 509	1 025.7	20 609	1 265.6
Hidalgo	24 152	997.4	10 275	885.1	13 877	1 100.8
Jalisco	19 279	274.8	8 784	254.3	10 495	294.6
México	51 935	350.0	22 606	305.4	29 329	394.5
Michoacán	9 642	243.2	4 316	227.3	5 326	257.9
Morelos	12 633	754.3	5 249	641.6	7 384	861.9
Nayarit	9 576	987.7	4 581	951.3	4 995	1 023.6
Nuevo León	8 188	184.1	3 709	165.8	4 479	202.6
Oaxaca	42 749	1 203.9	17 997	1 064.0	24 752	1 331.3
Puebla	30 655	542.4	13 195	484.7	17 460	596.1
Querétaro	7 711	448.2	3 307	392.9	4 404	501.1
Quintana Roo	6 831	519.8	3 154	476.5	3 677	563.8
San Luis Potosí	9 626	387.4	4 110	341.7	5 516	430.2
Sinaloa	17 463	658.4	7 279	545.2	10 184	773.0
Sonora	3 930	156.5	1 733	136.2	2 197	177.5
Tabasco	21 027	1 025.5	9 787	970.7	11 240	1 078.4
Tamaulipas	8 901	278.8	3 968	248.4	4 933	309.2
Tlaxcala	5 355	471.9	2 262	411.4	3 093	528.8
Veracruz	39 605	544.1	17 112	489.8	22 493	594.3
Yucatán	19 393	1 009.0	8 704	911.5	10 689	1 105.4
Zacatecas	7 284	527.9	3 169	475.4	4 115	577.0
Total	507 908	472.2	221 264	418.6	286 644	524.1

* Tasa de incidencia por 100 000 habitantes.

FUENTE : SINAVE/DGE/Secretaría de Salud/México 2009.

Tabla 1. Número de casos nuevos y tasa de incidencia reportados de amibiasis por entidad federativa para el año 2009. Tomado del Boletín Epidemiológico de la Secretaría de Salud.

1.3. Morfología de *E. histolytica*

La membrana plasmática de los trofozoítos de *E. histolytica* semeja a las membranas celulares de eucariontes superiores (Martínez-Palomo, 1982). En esta

estructura celular se organizan diversas moléculas de superficie que modulan diversos eventos importantes para el establecimiento del parásito en su hospedero tales como: i) la formación del casquete (Vargas *et al.*, 1996) ii) la adhesión a la célula blanco (Arroyo *et al.*, 1987; Petri *et al.*, 1987; Rosales-Encina *et al.*, 1987), ; iii) la fagocitosis (Orozco *et al.*, 1983; Rodriguez *et al.*, 1986); iv) el daño a la membrana de la célula blanco por medio de las proteínas formadoras. de poros o amebaporos (Lynch *et al.*, 1982) y v) la secreción de proteasas (Keene *et al.*, 1990). En conjunto estos eventos le permiten a la amiba colonizar e invadir tejidos en su hospedero, otra característica interesante en *E. histolytica* es su material genético, el cual está confinado por una membrana nuclear muy semejante a la de otros eucariontes. El trofozoíto generalmente tiene un núcleo de 4 a 7 μm y aproximadamente el 20% de las células presentan dos o más núcleos, por lo que la presencia de multiploidia es una característica genética común en este parásito (Das *et al.*, 2002; Mukherjee *et al.*, 2008).

1.4. Ciclo de vida de *E. histolytica*

Durante su ciclo de vida, el parásito presenta dos estadios morfológicos claramente distinguibles: el quiste y el trofozoíto. El quiste se presenta ligeramente ovalado con un diámetro de 8 a 20 μm , cuatro núcleos y una pared celular de aproximadamente 125 a 150 nm de espesor. Los quistes maduros, los cuales son muy resistentes al medio ambiente y a cambios ambientales drásticos, al ser ingeridos por el ser humano se desplazan por el estómago llegando hasta el íleon donde ocurre el desenquistamiento. Los mecanismos moleculares que controlan la reacción de desenquistamiento aun no han sido abordados debido a la falta de un

sistema *in vitro* que reproduzca el fenómeno de manera eficaz. De cada quiste emergen ocho trofozoítos uninucleados los cuales posteriormente se dividen rápidamente por fisión binaria en el lumen del intestino estableciendo la infección (Martínez-Palomo, 1982). Recientemente, se reportó el análisis global de la expresión de genes a escala genómica, donde se analizaron los quistes, de manera interesante se definió que genes tales como diversas sintetisas de quitina se sobreexpresan en los quistes, modulando de manera importante la reacción de enquistamiento (MacFarlane *et al.*, 2005).

Cada vez que los trofozoítos se han adherido a la mucosa intestinal, estos pueden alojarse como comensales alimentándose de bacterias o de *detritus* celulares provenientes del intestino. La fase activa del parásito es el trofozoíto, el cual, en contraste con el quiste, es extremadamente sensible al medio ambiente y muere rápidamente por desecación. El trofozoíto tiene un diámetro de 10 a 60 μm , posee una gran actividad motora y un pleomorfismo característico por la forma ameboide. El trofozoíto es esencialmente anaerobio, aunque puede crecer en presencia de oxígeno. Mediante mecanismos moleculares aún no caracterizados, los trofozoítos atraviesan la mucosa intestinal invadiendo los vasos sanguíneos de los tejidos más próximos, diseminándose hacia los diferentes órganos del cuerpo provocando abscesos, principalmente en el hígado (Figura 2). Este proceso está determinado por la expresión de algunos genes de virulencia los cuales se expresan a bajos niveles en condiciones de no invasión o bien cuando los trofozoítos se crecen en condiciones de laboratorio. Los trofozoítos se pueden desprender de la mucosa intestinal y en la luz del intestino grueso e iniciar

nuevamente el proceso de enquistamiento, durante el cual las células pierden movilidad, toman una forma esférica, se deshidratan y excretan parte de las reservas alimenticias presentes en las vacuolas digestivas, asumiendo la forma de prequiste (Mukherjee *et al.*, 2008).

Posteriormente, secretan una pared quística y forman pilas de ribosomas que forman los cuerpos cromidiales, transformándose en quistes uninucleados inmaduros (Despommier, 1987). El quiste uninucleado presenta sucesivas divisiones nucleares y origina nuevamente al quiste maduro tetranucleado cerrando el ciclo de vida (Figura 2). Finalmente, resulta importante resaltar que un alto porcentaje de los individuos infectados de trofozoítos, pueden permanecer como portadores asintomáticos.

Por otra parte, se conoce que los trofozoítos pueden duplicar su genoma varias veces incluso antes de que ocurra la citocinesis, generando de esta manera poblaciones de trofozoítos multiploides. El fenómeno podría ocurrir a través de un desacoplamiento del ciclo celular con la replicación del DNA (Das *et al.*, 2002). Esto podría explicar también las diferencias en el contenido de DNA observados previamente en algunas cepas de *E. histolytica* (Lopez-Revilla *et al.*, 1978).

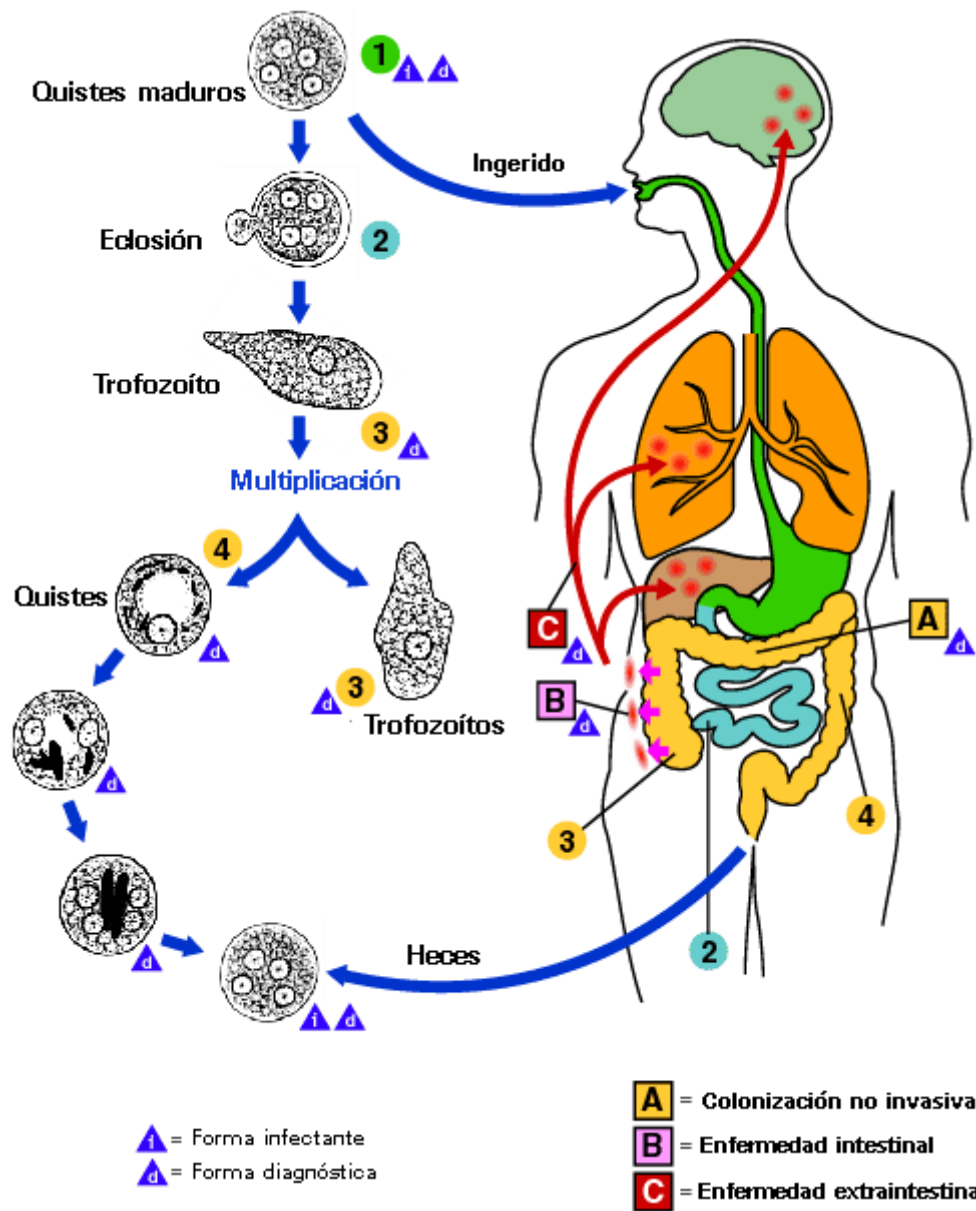


Figura 2. Ciclo de vida de *E. histolytica*. El ser humano se infecta al ingerir quistes maduros (1) en alimentos o agua contaminados. El desenquistamiento (2) ocurre en el intestino delgado y los trofozoítos (3) liberados migran al colón. Los trofozoítos se multiplican por fisión binaria y producen quistes (4), los cuales se eliminan por las heces. Debido a la protección conferida por las paredes de quitina, los quistes pueden sobrevivir días o semanas en el medio externo y transmitir la infección, iniciando nuevamente el ciclo (1). En muchos casos los trofozoítos se mantienen en el lumen intestinal presentándose una infección asintomática (A). En algunos pacientes, los trofozoítos invaden la mucosa intestinal (B), produciendo desde síntomas leves hasta disentería grave, o bien, puede producirse una enfermedad extraintestinal,(C), con manifestaciones patológicas cuando por el torrente sanguíneo, los trofozoítos invaden otros órganos como hígado, cerebro o pulmones. Tomado del URL (<http://www.dpd.cdc.gov/dpdx/HTML/Amebiasis/>)

1.5. Regulación de la expresión génica en eucariontes

En eucariontes los procesos de transcripción y traducción se encuentran compartimentalizados y el destino y modificaciones que sufre el RNA después de ser sintetizado ha sido sujeto de extensa investigación en los últimos años. La complejidad que representan los eventos nucleares requeridos para generar un RNAm exportable y traducible, permite que el control del flujo de la información genética se regule en múltiples puntos de control.

La regulación de la expresión genética modula procesos vitales para la célula tales como el crecimiento, la división, la diferenciación, la replicación del material genético, etc. También regula la expresión de genes en respuesta a estímulos externos. En general, la expresión genética se regula a tres niveles: i) transcripcional, ii) post-transcripcional y iii) traduccional. En el primer nivel, la expresión es regulada por proteínas denominadas factores de transcripción que interaccionan con secuencias específicas presentes en la región 5' no traducida (5' UTR, por sus siglas en inglés), río arriba del sitio de inicio de la transcripción de los genes, también llamadas regiones promotoras. Estas interacciones determinan si un gen se va expresar en un RNA mensajero precursor (pre-RNAm) y en qué cantidad y tiempo. Posteriormente que se sintetiza la molécula de pre-RNAm en el núcleo, esta sufre una serie de modificaciones que forman parte del nivel de regulación post-transcripcional produciendo un RNA mensajero maduro (RNAm). En este nivel de modulación de la expresión genética, los mecanismos celulares operan directamente sobre la molécula de RNA. El correcto procesamiento del RNAm es importante para el desarrollo de los organismos, ya que de él depende

la síntesis de las proteínas (nivel de regulación traduccional) necesarias para el buen funcionamiento de las células.

El pre-RNAm experimenta diversos eventos de procesamiento nucleares tales como la adición de la estructura de *cap* (7-metil-guanosina) mediante la reacción de *capping* en el extremo 5'-UTR, la eliminación de secuencias no codificantes o intrones mediante *splicing*, y el procesamiento del extremo 3' UTR, que incluye las reacciones de corte y poliadenilación. Esta serie de reacciones de procesamiento del pre-RNAm da origen a RNAm maduros que son transportados de manera eficiente hacia el citoplasma para su traducción a proteínas (Fritz *et al.*, 2004).

En particular, el procesamiento del extremo 3'-UTR del pre-RNAm que incluye las reacciones de corte y poliadenilación, es un evento de suma importancia debido a que se ha reportado que alteraciones en este paso pueden modificar la viabilidad, el crecimiento y el desarrollo celular (Zhao *et al.*, 1999). La adición y presencia de tracto de adeninas (proceso conocido como poliadenilación) en el extremo 3' del RNAm regula la expresión génica debido a que favorece el transporte del RNAm del núcleo al citoplasma, provee de estabilidad al RNAm en el citoplasma y promueve de manera considerable la traducción (Colgan *et al.*, 1997). Por otra parte, la eliminación del tracto de poli (A) es el primer paso para iniciar la degradación del RNAm.

1.6. Degradación de RNAm en eucariontes

El RNA es usualmente considerado una molécula muy inestable y es químicamente más lábil que el DNA. De esta manera es sorprendente que el estado estacionario o basal del RNAm en organismos eucariontes es relativamente estable. La principal razón de que haya siempre una cantidad basal de RNAs mensajeros parece debido a que la degradación es catalizada casi exclusivamente por proteínas conocidas como exoribonucleasas que degradan al RNAm desde sus extremos, los cuales están protegidos generalmente por estructuras específicas y proteínas asociadas a las mismas, por lo que la remoción de estas estructuras es considerado el paso limitante para iniciar la degradación o decaimiento del RNAm (Meyer *et al.*, 2004). Ejemplos de esto lo representan la estructura de *cap* en el extremo 5'-UTR; la presencia del tracto de adeninas en el extremo 3'-UTR del RNAm y la presencia de unión de la proteína de unión al tracto de poli(A) (PABP por sus siglas en inglés) en el tracto de adeninas. El control de la degradación de RNAm es un componente importante en la regulación de la expresión de genes (Parker *et al.*, 2004). Los niveles basales del RNAm están determinados por las tasas de síntesis y degradación del mismo, las cuales regulan la expresión final de los genes al proveer de RNA mensajeros a la maquinaria traduccional. Debido a esto, el estudio de la síntesis y degradación del RNAm ha tomado gran importancia en años recientes.

La eliminación del tracto de poli(A) es el primer paso para iniciar la degradación del RNAm. Dos rutas generales de degradación del RNAm han sido descritas en organismos eucariontes. Ambas vías comparten como primer evento,

la remoción exonucleolítica del tracto de poli(A) o desadenilación. Cuando el tracto de poli(A) es reducido por digestión enzimática, producto de la actividad de las desadenilasas, a un tamaño de aproximadamente 10 adeninas, la PABP es liberada, lo que favorece el *decapping* del extremo 5'-UTR, el cual esta mediado por la actividad del complejo de exoribonucleasas *Decapping Protein 1 y 2* (DCP1/DCP2, por sus siglas en inglés). Una vez removida la estructura de *cap* se lleva a cabo una rápida degradación nucleolítica en dirección 5'-3' del resto del mensajero por una enzima llamada XRN1. En la ruta alterna posterior a la desadenilación, el RNAm es degradado en dirección 3'-5' por el complejo enzimático llamado exosoma que comprende al menos 10 diferentes exoribonucleasas del tipo Rrp y algunas RNA helicasas. Por último la estructura de *cap* restante es degradada por la enzima *Decapping Protein Scavenger* (DCPs, por sus siglas en inglés) (Figura 3). En ambas rutas participan proteínas específicas las cuales han sido estudiadas en diversos organismos (Meyer *et al.*, 2004; López-Camarillo, 2012).

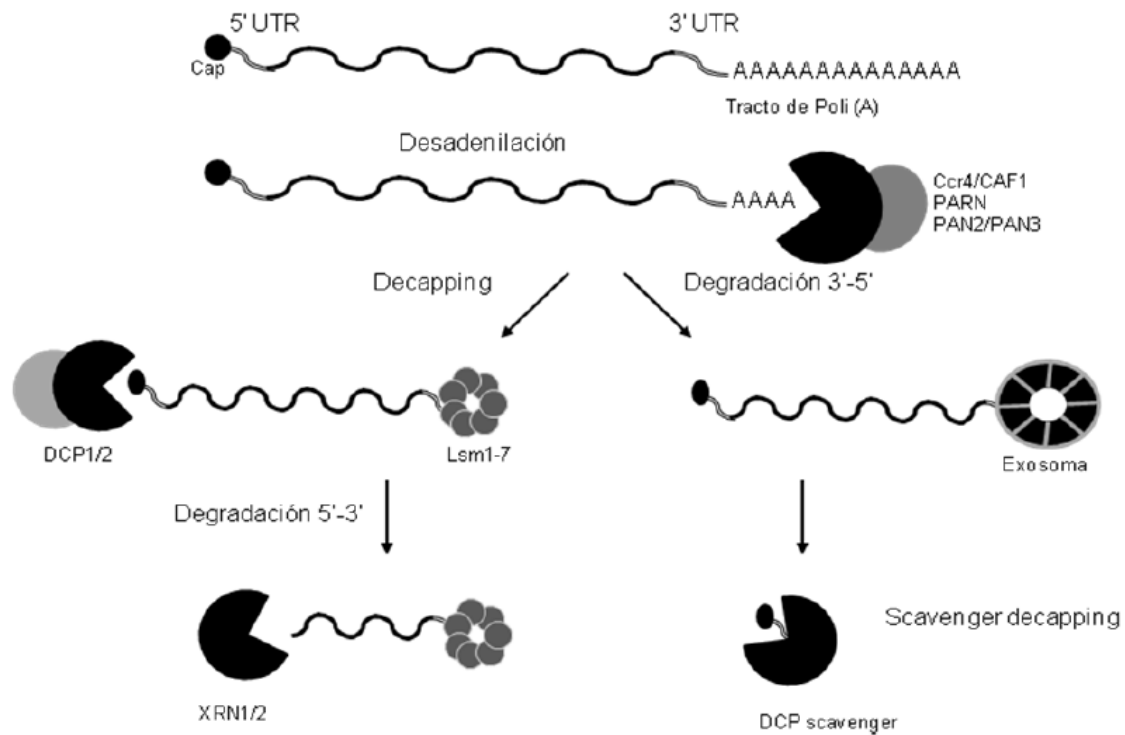


Figura 3. Vías de degradación del RNAm en eucariontes. El inicio de la degradación de RNAm comienza con la desadenilación en la que participan al menos cinco ribonucleasas diferentes conocidas como CCR4, POP2, PARN, PAN2 y PAN3. Una vez degradado el tracto de poli(A) puede seguirse la ruta de degradación en dirección 3'-5 o en dirección 5'-3' por diferentes exoribonucleasas. Tomada de López-Camarillo *et al*, 2012.

1.7. Proteínas que participan en la de degradación del RNAm

A la fecha se han identificado en eubacterias y eucariontes diversas exoribonucleasas que participan en la degradación específica del RNAm (Zuo *et al.*, 2001), las cuales se han clasificado en seis superfamilias en base a sus estructuras proteicas (Tabla 2). De este grupo numeroso de proteínas, requiere especial atención el subgrupo de las RNAsas D, debido a que incluye a las desadenilasas involucradas en el inicio de la degradación del RNAm. En células eucariontes, el principal complejo de poli(A) ribonucleasas se encuentra formando parte del complejo transcripcional CCR4/CAF1/NOT1-5, el cual contiene dos ribonucleasas de trectos de poli(A), además de 5 factores de transcripción denominados NOT1-NOT5 y dos proteínas accesorias CAF 4 y CAF130 (Figura 4) (López-Camarillo, 2012). Dicho complejo también se ha descrito en otros organismos tales como *Homo sapiens*, *Drosophila melanogaster* y *Mus musculus*. Otro complejo enzimático que cataliza la reacción de desadenilación citoplasmática del RNAm está formado por las proteínas PAN2, PAN3 y PARN (Chen *et al.*, 2011).

También participan proteínas para la remoción de la estructura de *cap* DCP1 y DCP2 en un proceso estimulado por las proteínas Dhh1, Edc3, Pat1/Mrt1, Sm-like (Lsm1p-7p) y Hedls (Coller *et al.*, 2004). Adicionalmente participan las exoribonucleasas XRN1 y XRN2 que degradan el RNAm sin estructura de *cap* en dirección 5'-3'. Alternativamente, el RNAm es degradado en dirección 3'-5' por el exosoma que es un complejo proteico de ribonucleasas PH-like (Lykke-Andersen *et al.*, 2011)

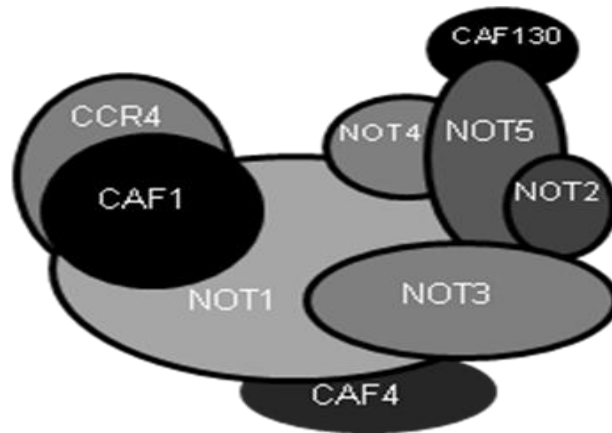


Figura 4. Esquema de la conformación del complejo CCR4-NOT. Poli(A) ribonucleasas: CCR4 y CAF1. Factores transcripcionales: proteínas NOT1-5 y proteínas accesorias: CAF4 y CAF130. Tomada de López-Camarillo *et al*, 2012.

Superfamilia	Miembros en <i>E. coli</i>	Miembros en <i>S. cerevisiae</i>	Distribución en otros organismos	Características catalíticas
RNR	RNasa II RNasa R	-	Muchas eubacterias	Degradación en dirección 3'-5'
		Rrp44 (Dis3) Msu (Dss1) Ssd1	Todos los eucariontes	
DEDD	RNasa D	Rrp6 POP2 CAF1	Todos los eucariontes Ciertas eubacterias	Degradación en dirección 3'-5' Algunos tienen actividad de DNasa
	RNase T	-	γ -proteobacterias	
	Oligoribonucleasa	Ynt20(Rex2)	Todos los eucariontes Ciertas eubacterias	
	-	Pan2 Rex1,3,4	Todos los eucariontes	
RBN	RNase BN	-	Ciertas eubacterias	Degradación en dirección 3'-5'
PDX	PNPasa	-	Muchas eubacterias Ciertos eucariontes	Degradación en dirección 3'-5' Actividad dependiente de fosfatos
	RNasa PH	Rrp41-43, 45, 46 Mtr3	En todos los reinos	Degradación en dirección 3'-5' Actividad dependiente de fosfatos
RRP4	-	Rrp4, 40 Cs14	Todos los eucariontes Muchas archaeas	Degradación en dirección 3'-5'
5XP	-	Xrn1 Rat1	Todos los eucariontes	Degradación en dirección 5'-3'

Tabla 2. Superfamilia de exoribonucleasas identificadas en eubacterias y eucariontes. Tomada y modificada de Zuo *et al*, 2001.

1.8. Degradación de RNA mensajeros y cuerpos de procesamiento (*P-bodies*)

El control de la degradación del RNAm juega un papel clave en la regulación de genes a nivel post-transcripcional (Bevilacqua *et al.*, 2003). A la fecha existe evidencia acumulativa que demuestra que los mecanismos de degradación se llevan a cabo en *foci* citoplasmáticos especializados conocidos como RNPM gránulos, *foci* de degradación de RNAm o cuerpos de procesamiento de RNAm (*P-bodies*/GW182 *bodies*), los cuales están enriquecidos con sustratos de RNA y proteínas de degradación de RNAm. Los *P-bodies* representan centros de procesamiento donde transcritos traducionalmente inactivos son almacenados, reprimidos o degradados (Kulkarni *et al.*, 2010). El inventario real de proteínas que pertenecen a los *P-bodies* no está bien definido pero aproximadamente 25 factores diferentes han sido descritos en dichas estructuras citoplasmáticas (Aizer *et al.*, 2008). Los *P-bodies* existen en insectos, levaduras, nematodos y mamíferos como estructuras citoplasmáticas pequeñas que tienen un papel importante en la degradación, almacenamiento, vigilancia del RNAm y en los mecanismos de silenciamiento de genes basado en RNA (Eulalio *et al.*, 2007). De hecho, se ha demostrado que los transcritos blanco o diana y las proteínas AGO2, DICER y GW182 que participan en la vía RNA de interferencia por microRNAs (miRNAs) y RNAs pequeños de interferencia (siRNAs), se concentran en *P-bodies* (Liu *et al.*, 2005; Sen *et al.*, 2005; Eulalio *et al.*, 2007). Adicionalmente, proteínas UPF y SMG que están involucradas en el mecanismo de degradación de RNAs sin sentido (NMD por sus siglas en inglés), que específicamente degradan RNAm aberrantes que contienen codones de paro prematuros, también se han

encontrado en *P-bodies* (Sheth *et al.*, 2006). En mamíferos, la mayoría de las células tienen de tres a nueve *P-bodies*, aunque el número puede diferir dependiendo de la tasa de proliferación de las células y a condiciones de estrés, tales como privación nutricional, estrés oxidante y choque térmico. Tratamientos con inhibidores de la transcripción y de la traducción como son la actinomicina D y la cicloheximida pueden causar la desintegración de los *P-bodies* (Cougot *et al.*, 2004). Por el contrario, en células tratadas con puromicina se incrementa el número y tamaño de los mismos, lo que refleja la importancia de RNAm activos y la síntesis de proteínas en el ensamblaje de estos agregados citoplasmáticos (Eulalio *et al.*, 2007). A pesar de que los componentes de los *P-bodies* juegan un papel crucial en la degradación del RNAm también hay evidencia que muestra que las proteínas de NMD y de silenciamiento de genes mediado por miRNAs y siRNAs no se ven afectadas en las células que carecen de *P-bodies*. Esto sugiere que la formación de dichas estructuras citoplasmáticas no es la causa sino una consecuencia de la actividad de las maquinarias de degradación del RNAm (Eulalio *et al.*, 2007; Stalder *et al.*, 2009). Además, la represión de la traducción por miRNAs no requiere de las estructuras *P-bodies*, y la ubicación de miRNAs y las proteínas del complejo inducido por RNA (RISC por sus siglas en inglés) en *P-bodies* parecen ser la consecuencia de la represión de traducción (Chu *et al.*, 2006) .

En la tabla 3 (Parker *et al.*, 2007) se enlistan las proteínas identificadas estudiadas a la fecha como componentes de los *P-bodies*.

Componente	Función	Organismo
Proteínas de degradación de RNAs mensajeros:		
DCP2	<i>Decapping</i> (subunidad catalítica)	Levadura Humano Mosca Gusano
DHH1	DEAD <i>box</i> helicasa	Levadura Mosca
Pat1	Activador de <i>decapping</i>	Levadura Humano
Complejo Lsm	Activador de <i>decapping</i>	Levadura Humano
XRN1	5'-3' exoribonucleasa	Levadura Ratón Humano
Edc3	Activador de <i>decapping</i>	Levadura Mosca Humano
CAF1	Principal desadenilasa	Levadura Humano
CCR4	Desadenilación	Levadura Humano
Upf1-3	Componentes de la maquinaria de degradación de RNAs sin sentido (NMD por sus siglas en inglés)	Humano Levadura Mosca
SMG5	Factor requerido para la desfosforilación de Upf1	Humano
SMG7	Recluta la proteína Upf1	Humano

Proteínas de silenciamiento de RNA :		
AGO1/2	Silenciamiento mediado por siRNAs y miRNAs	Gusano Mosca Levadura
GW182	Proteína rica en repetido de GW. Silenciamiento mediado por siRNAs y miRNAs	Gusano Humano
TNRC6b	Parálogo de GW182. Silenciamiento mediado por siRNAs y miRNAs	Humano
MOV10	Helicasa DEAD box, Silenciamiento mediado por siRNAs y miRNAs	Humano
Proteínas de unión a RNA y represores de la traducción:		
CPEB	Proteína de unión al elemento de poliadenilación	Humano
Rbp1	Participa en la regulación de la estabilidad de RNAs mensajeros	Levadura
TTP	Proteína de unión a elementos ricos en AU	Humano
eIF4E	Regulador de la traducción, proteína de unión al CAP	Humano
eIF4E-T	Proteína de unión a eIF4E, represor de la traducción	Humano
Gemina 5	Proteína de unión a RNA.	Humano

Tabla 3. Factores de degradación de RNAm identificados en *P-bodies* en organismos eucariontes. Tomada y modificada de Parker *et al*, 2007.

2. ESTADO DEL ARTE

2.1. Regulación de la expresión génica en parásitos protozoarios

Las infecciones gastrointestinales asociadas a parásitos protozoarios tienen altos índices de morbilidad y mortalidad en el mundo. A pesar de los diversos esfuerzos para el desarrollo de vacunas y medicamentos para el control de la transmisión de estos parásitos se han obtenido pocos avances contribuyendo a que su persistencia se considere como un importante problema de salud pública. Desde hace algunos años diversos grupos de trabajo se han dado a la tarea de realizar estudios acerca de los mecanismos moleculares de los procesos de virulencia y patogenicidad. Una de las principales áreas de estudio se ha dedicado al estudio de los mecanismos en la regulación de la expresión de genes involucrados en patogenicidad, diferenciación, resistencia a fármacos y evasión de la respuesta inmune. La comprensión de los mecanismos implicados en el control de estos eventos puede ayudar para el desarrollo de pruebas diagnósticas, tratamiento médicos más eficientes y blancos de vacunas. Sin embargo, a la fecha los mecanismos de regulación génica son aún poco entendidos en parásitos protozoarios.

En la mayoría de las células, la regulación de la expresión de genes es de suma importancia para el desarrollo, adaptación y homeostasis en el medio ambiente. En organismos eucariontes, la expresión de genes es sometida una regulación llevada a cabo en varios pasos. La regulación ocurre a nivel transcripcional: síntesis del RNA a partir de DNA, adición del tracto de adeninas en

el extremo 3', adición de la estructura de *cap* en el extremo 5' y *splicing*; generando transcritos maduros que pueden ser translocados de núcleo hacia el citoplasma y por último ocurrir la síntesis de proteínas.

Estudios recientes han permitido la identificación de diversos procesos biológicos involucrados en la regulación de la expresión de genes en parásitos protozoarios. Estos procesos requieren de varios componentes, como lo son elementos reguladores en *cis*, factores y cofactores de transcripción, proteínas modificadoras de la cromatina y proteínas responsables de la regulación post-transcripcional.

De los parásitos protozoarios más estudiados con respecto a la regulación de la expresión génica están Tripanosomátidos *Plasmodium sp*, *Trichomonas vaginalis* y *E. histolytica* (Gomez et al., 2010).

2.1.1. Regulación de la expresión génica en *E. histolytica*

E. histolytica, es el agente etiológico de la amibiasis humana. Los mecanismos moleculares que participan en la invasión del parásito no se han entendido completamente. Diversas clonas de *E. histolytica*, derivadas de una misma cepa muestran diferentes fenotipos de virulencia (Bracha et al., 1984; Orozco et al., 1985; Rodriguez et al., 1996), este comportamiento puede ser debido a cambios en la expresión de ciertos genes. Otro de los eventos importantes que se sugiere ocurren debido a la expresión diferencial de sus genes es que lleva un ciclo de vida donde este parásito experimenta la conversión

reversible de las formas infectivas o quistes a las de células invasoras o trofozoítos. El genoma de este parásito consiste 20.8 Mb distribuidas entre 888 scaffolds y contiene aproximadamente 9, 938 genes (Loftus *et al.*, 2005). Ha sido difícil determinar el número de cromosomas en *E. histolytica*, dado que no se condensa, actualmente se ha definido que tiene de 14 a 17 cromosomas (Clark *et al.*, 2007). Las regiones intergénicas son de 400 pb a 2,3 kb, lo que sugiere un embalaje hermético de los genes (Bhattacharya *et al.*, 2000). Aproximadamente el 25% de los genes de este parásito contienen intrones (Loftus *et al.*, 2005). Las secuencias de los intrones son relativamente cortas de 46 a 115 pb, y contiene los dinucleótidos GU y AG en los sitios donantes y aceptores de los sitios de *splicing*, respectivamente. Alrededor del 10% del genoma consiste en genes para RNA de transferencia (tRNA; por sus siglas en inglés) organizados en arreglos en tándem, que varían en longitud de 490 a 1775 pb y que contienen de 1 a 5 genes tRNA (Clark *et al.*, 2006). Los genes para RNA ribosomal se encuentran exclusivamente en moléculas circulares extracromosómicas de DNA con un tamaño aproximado de 26 kb (Bhattacharya *et al.*, 1989). Los genomas de *E. histolytica* y otras especies de *Entamoeba* tienen un alto contenido de A/T (77,6%), con la excepción de *E. moshkovskii*, que tiene aproximadamente 10% menos de contenido A/T (Bhattacharya *et al.*, 2000).

Este parásito constantemente se somete al reversible cambio entre quistes y trofozoítos, a la fecha se han identificado los genes implicados en el desarrollo de este parásito, dichos genes han sido examinados por perfiles de expresión del genoma completo, aproximadamente, el 15% de los genes anotados son

potencialmente regulados y son mayoritariamente de la fase de quiste (672 en total) que incluyen cisteína proteasas y proteínas cinasas transmembranales. Los genes regulados enriquecidos en la fase de trofozoíto (767 en total), genes implicados en la invasión de tejidos, posibles reguladores de diferenciación, donde se encuentran receptores acoplados a proteínas y factores de transcripción (Ehrenkaufer *et al.*, 2007). La mayoría de las personas infectadas con *E. histolytica* permanecen asintomáticos. En algunos casos de infección los trofozoítos invaden la mucosa intestinal produciendo disentería, y en una pequeña fracción de los infectados, existe una difusión de los trofozoítos hacia el hígado, donde ocurre la formación de abscesos. La capacidad de los trofozoítos de *E. histolytica* para sobrevivir dentro el hospedero y para destruir los tejidos del mismo, es realizado por la regulación específica de un gran número de genes y proteínas. Sin embargo a la fecha no se ha logrado definir con exactitud los genes responsables de dicha regulación. Estudios complementarios del análisis transcripcional del genoma de *E. histolytica* en trofozoítos aislados del colon de ratones infectados y de cultivos *in vitro* revelaron 523 transcritos que corresponden al 5,2% de todos los genes predichos en *E. histolytica*, cuya expresión varía significativamente en trofozoítos aislados del intestino. Los genes que modificaron su expresión en los trofozoítos obtenidos a partir del colon de ratones, codifican para proteínas implicadas en metabolismo, virulencia, señalización celular y actividad antibacteriana (Gilchrist *et al.*, 2006). Estudios comparativos de la abundancia de RNA en trofozoítos de *E. histolytica* aislado de abscesos hepáticos de animales y los obtenidos a partir de cultivos *in vitro* demuestran que al menos

siete genes de *E. histolytica* fueron específicamente sobreexpresados y cinco fueron subexpresados en los trofozoítos aislado de los hígados. Los genes regulados en el caso de sobreexpresión genes codifican para proteínas asociadas con choque térmico, algunas proteínas ribosomales, ciclofilina, ferredoxina 2, mientras que dos de los genes subexpresados codifican para miembros de una familia de proteínas que contienen secuencias repetidas ricas en lisina y ácido glutámico (Bruchhaus *et al.*, 2002). Todos estos resultados sugieren que el evento de invasión del hospedero requiere de la regulación de una variedad de genes. La respuesta tan variable por infección con *E. histolytica* es debido a las diferencias en la capacidad de virulencia de los distintos aislados de *E. histolytica*. Otras investigaciones demostraron mediante un estudio comparativo de los transcriptomas de las cepas de alta virulencia HM1: IMSS y de baja virulencia Rahman; la identificación de 353 transcritos que exhiben al menos el doble de diferencia entre las cepas, 152 transcritos fueron sobreexpresado en HM1: IMSS y 202 transcritos sobreexpresaron más en la cepa Rahman. Los genes expresados diferencialmente incluyen cisteína proteasas, calmodulina y cadenas ligeras de lectina (Davis *et al.*, 2007). Estos resultados validan la hipótesis de que las cisteína proteasas tienen un papel importante en la virulencia de *E. histolytica*. Por otra parte, el estudio de genes importantes de sobrevivencia de este parásito en diferentes condiciones ambientales han revelado que el patrón de expresión génica de trofozoítos de *E. histolytica* expuestos a estrés por choque térmico a 42°C mostró una baja regulación masiva de genes. De los 1, 131 genes únicos investigados mediante microarreglos, 471 fueron significativamente reprimidos

durante el tratamiento. Un pequeño número de genes fueron sobreexpresados en respuesta al choque térmico; dentro de este grupo se identificaron genes que codifican para las proteínas de choque térmico HSP90 y HSP70, algunas proteínas hipotéticas y factores reguladores tales como BRF y una posible transcriptasa reversa (Weber *et al.*, 2006). Los factores generales de la transcripción estudiados en *E. histolytica* han sido pocos, identificando algunos los factores de transcripción, algunas proteínas de estos han sido aisladas y caracterizadas no sólo a nivel estructural, sino también a nivel funcional. La transcripción de genes en este parásito es monocistrónica y los RNAm son sintetizados por una inusual RNA polimerasa sensible a α -amanitina (Lioutas *et al.*, 1995). El núcleo de las regiones promotoras de varios genes codificantes contienen motivos conservados, y de acuerdo al estudio de diferentes promotores de genes en *E. histolytica*, el tamaño de la región promotora es entre 200 y 900 pb (Bruchhaus *et al.*, 1993; Singh *et al.*, 1997). La proteína de unión a la caja TATA (TBP; por sus siglas en inglés) es el único miembro de la maquinaria basal de la transcripción de este parásito, caracterizada a la fecha. Esta proteína consta de 234 aminoácidos y su dominio funcional tiene un 55% identidad a nivel de secuencia frente a la proteína TBP de humano (Luna-Arias *et al.*, 1999). La proteína recombinante *EhTBP* forma complejos específicos con la secuencia consenso de la caja TATA de *E. histolytica* y con otras secuencias tipo TATA, lo que indica que esta proteína es más promiscua que las proteínas TBP de levadura y humano (de Dios-Bravo *et al.*, 2005).

Diversos grupos de investigación se han dado a la tarea de estudiar las regiones promotoras de diversos genes, entre ellos los genes que codifican para la subunidad pesada de la proteína Gal/GalNAc lectina; *hgl2* and *hgl5* (Buss *et al.*, 1995; Purdy *et al.*, 1996); los genes de las proteínas de multirresistencia a fármacos; *EhPgp1* y *EhPgp5* (Gomez *et al.*, 1998; Perez *et al.*, 1998); el complejo *EhADHCP* (Azuara-Liceaga *et al.*, 2005; Flores-Soto *et al.*, 2005) y la GTPsa *EhRabB* (Romero-Diaz *et al.*, 2007).

La cromatina de *E. histolytica* es empaquetada en estructuras tipo nucleosoma, no son muy diferentes de la de los metazoos, sin embargo, las regiones de unión entre nucleosomas adyacentes parecen ser irregulares en longitud en comparación con el promedio de 40 pb que se observa en los metazoos (Torres-Guerrero *et al.*, 1991; Gomez-Conde *et al.*, 1998). El genoma de *E. histolytica* posee genes codifican para las cuatro proteínas que componen el *core* de histonas (Fodinger *et al.*, 1993; Binder *et al.*, 1995). Los dominios amino-terminales de las histonas, aunque divergentes de las secuencias de metazoos, son altamente básicos con varios residuos de lisina que son blancos potenciales para la acetilación de las histonas por acetiltransferasas (HAT; por sus siglas en inglés). *E. histolytica* contiene algunos miembros de la familia de acetiltransferasas, como GNAT y MYST que presentan un alto valor de similitud con GCN5, Myst, TafII250, Hat1 y Elp3 de humano. Aunque este parásito contiene genes para la acetilación y la desacetilación de las histonas, hasta ahora no se conoce los mecanismos de estos eventos en este parásito protozoario. Sin

embargo, se considera que estas proteínas realizan una importante actividad implicada en la regulación transcripcional (Ramakrishnan *et al.*, 2004).

La regulación a nivel post transcripcional en *E. histolytica*, aún no está bien definida, dado que la UTR 5' de los RNAm en este parásito son generalmente cortos en longitud (de 5 a 20 nucleótidos) (Bruchhaus *et al.*, 1993). Sólo pocos RNAm tienen una UTR 5' extensa como lo es para los genes: *Ehcp112*, *EhTBP*, *Ehpak*, *Ehmcm3* (Gangopadhyay *et al.*, 1997; Garcia-Rivera *et al.*, 1999; Luna-Arias *et al.*, 1999).

Análisis computacionales para el estudio de la UTR 3' de una colección de secuencias genómicas en este parásito reveló la presencia de elementos conservados, tales como dominios ricos en AU correspondiente a la secuencia consenso de la señal de poliadenilación de los pre-RNAm. Interesantemente, la organización molecular de los elementos en *cis* reguladores de las UTR 3' de los pre RNAm parece ser más o menos conservados a través de la escala evolutiva, mientras que la señal de poliadenilación parece ser más específico en parásitos protozoarios (Zamorano *et al.*, 2008).

2.2. Mecanismos de degradación de RNAm en parásitos protozoarios

Mucho de lo que se conoce acerca de las rutas de degradación en eucariontes ha sido por su estudio en *S. cerevisiae* y en *H. sapiens*. Sin embargo, a pesar de la importancia de los eventos transcripcionales, cotranscripcionales y post-transcripcionales en la regulación de la expresión génica, los mecanismos de

degradación del RNAm en parásitos protozoarios no han sido ampliamente estudiados.

A la fecha existen pocos trabajos realizados en parásitos sobre la descripción de los mecanismos de degradación de transcritos y la existencia de *P-bodies*. En el parásito protozoario *Trypanosoma brucei* se identificaron y caracterizaron cuatro diferentes proteínas exoribonucleasas denominadas XRNA, XRNB, XRNC y XRND responsables de degradar el RNAm en dirección 5'-3', dichas proteínas presentan homología a las exoribonucleasas Xrn1p y Xrn2p de levadura. La proteína XRND se localiza en el núcleo, mientras que las proteínas XRNB y XRNC se encuentran en citoplasma y la proteína XRNA en ambos compartimentos. También se determinó que la actividad de XRND y XRNA es esencial para el crecimiento del parásito (Li *et al.*, 2006).

En parásitos protozoarios, las estructuras citoplasmáticas que asemejan *P-bodies* se describieron inicialmente en tripanosomas. En el año 2007 se reportó que el RNAm y proteínas relacionadas con el metabolismo del RNAm se concentran en gránulos citoplasmáticos en respuesta a condiciones de estrés por inanición como una estrategia para la protección de transcripción transitoria durante períodos de estrés en *T. cruzi* y *T. brucei*. Sin embargo, parece que estas estructuras no parecen representar la estructura típica de los *P-bodies*. (Cassola *et al.*, 2007).

En *T. brucei* se ha descrito un complejo proteico conformado las proteínas CAF1, NOT1, NOT2 y NOT5, DHH1 y un posible homólogo a Caf130. La proteína

CAF1 de este parásito tiene actividad de desadenilación y tiene una distribución en citoplasma, aunque no se logró determinar una distribución clara que indique la existencia de estructuras citoplasmáticas de almacenamiento o degradación de RNAm (Schwede *et al.*, 2008). En *T. cruzi* la helicasa TcDhh1 se encuentra en complejos proteicos y en *foci* citoplasmáticos discretos que parecen *P-bodies*. Dichas estructuras varían en cantidad en respuesta a condiciones de estrés nutricional y a tratamientos con cicloheximida y puromicina (Holetz *et al.*, 2007).

Sin embargo, actualmente no se han caracterizado otras proteínas que participen en la degradación del RNAm como son las diversas RNAsas, enzimas de *decapping* ni posibles desadenilasas en otros parásitos protozoarios de importancia médica.

3. ANTECEDENTES PARTICULARES DEL PROYECTO.

Previamente, en nuestro grupo de trabajo durante la maestría se identificó una pequeña familia de poli(A) ribonucleasas, específicamente nos dimos a la tarea de caracterizar a la desadenilasa *EhCAF1*. En dicho trabajo se describió la clonación, expresión y purificación de la proteína recombinante con la que se obtuvieron anticuerpos policlonales anti-*EhCAF1*. Mediante ensayos de *Western Blot* utilizando extractos proteicos de trofozoítos utilizando los anticuerpos específicos anti-CAF1 se identificó a la proteína endógena correspondiente a una banda de 37kDa en la fracción citoplasmática. Por otra parte la realización de inmunofluorescencias nos confirmó la localización de *EhCAF1* en pequeños

puntos en citoplasma (López-Rosas, 2007). En consecuencia a esta observación, nos dimos a la tarea de una revisión bibliográfica, encontrando datos que reportan que proteínas que participan en mecanismos de degradación de RNAm se localizan en pequeñas estructuras citoplasmáticas denominadas *P-bodies*. En el presente trabajo se hizo la propuesta de la existencia de estructuras tipo *P-bodies* en el parásito protozoario *E. histolytica*. A la fecha, la descripción de dichas estructuras no se ha realizado en este organismo, por lo que nos interesamos en definir si la proteína *EhCAF1* se localiza en estos agregados citoplasmáticos y describir la presencia de cuerpos de procesamiento y almacenamiento de RNAs mensajeros (*P-bodies*).

4. IMPORTANCIA DE CONTINUAR EL ESTUDIO DEL TEMA.

En *E. histolytica* se conoce poco acerca de la degradación de los RNAm por lo que nos interesamos en continuar con el estudio de proteínas que pertenecen a las maquinarias de degradación de mensajeros, como lo describiremos a continuación, en donde hemos continuado con lo planteado en el proyecto iniciado durante la maestría en relación al estudio de proteínas de degradación de RNAm, específicamente desadenilasas. En el presente trabajo proponemos el estudio de la localización de la proteína *EhCAF1* con una distribución citoplasmática en pequeños *foci* enriquecidos, similares a los *P-bodies*. La colocalización en citoplasma de *EhCAF1* con proteínas de distintas maquinarias de degradación en amiba nos permitirá sugerir la existencia de estructuras tipo *P-bodies*. que abre la

oportunidad para entender los mecanismos de degradación del RNAm y el silenciamiento de genes basado en RNA este parásito protozoario.

II. OBJETIVOS

OBJETIVO GENERAL

Identificar el repertorio de proteínas involucradas en la degradación del RNAm y determinar si algunas de ellas se localizan en estructuras tipo *P-bodies* en *E. histolytica*.

OBJETIVOS ESPECIFICOS

1. Identificar los posibles genes que codifican para proteínas de degradación del RNAm en *E. histolytica*.
2. Clonar los genes correspondientes a las maquinarias de: *decapping*: *EhDCP2*, *EhXRN2* y silenciamiento del RNAm (RISC): *EhAGO2-2*.
3. Expresar las proteínas recombinantes y generar anticuerpos específicos contra las mismas.
4. Co-inmunolocalizar las proteínas *EhCAF1*, *EhDCP2*, *EhXRN2* y *EhAGO2-2* en trofozoítos de *E. histolytica*.
5. Co-inmunolocalizar *EhCAF1* con RNAm en trofozoítos.
6. Analizar el efecto de la inhibición de la transcripción y la traducción en el tamaño y número de *P-bodies*.

III. MATERIALES Y MÉTODOS

1. Análisis *in silico* del repertorio de proteínas involucradas en la degradación de RNAm

Para la identificación de los genes que codifican para proteínas de las maquinarias de degradación de RNAm se llevó a cabo una búsqueda de secuencias homólogas en base de datos del genoma de *E. histolytica*, AmoebaDB database (<http://amoebadb.org/amoeba/>). Las secuencias de proteínas de degradación de RNAm para humano fueron utilizadas como templatado. Las proteínas ortólogas en *E. histolytica* fueron confrontadas mediante búsquedas heurísticas contra todas las bases de datos y alineadas con secuencias de aminoácidos de proteínas de degradación de RNAm de diferentes organismos obtenidas de UniProtKB database utilizando el programa ClustalW (<http://www.ebi.ac.uk/clustalw/>). A partir del alineamiento obtenido, se realizó un análisis filogenético de las secuencias por el método de *Neighborn joining* mediante el programa *Molecular Evolutionary Genetics Analysis* (MEGA por su siglas inglés) versión 5.1. La predicción de dominios estructurales y funcionales se realizó con el programa *ScanProsite* (<http://prosite.expasy.org/scanprosite/>). La predicción de la estructuras terciarias para las proteínas *EhCAF1*, *EhXRN2* and *EhDCP2* fueron predichas con el programa *Swiss model* (<http://www.expasy.org/swissmod/SWISS-MODEL.html>) utilizando como templatado la estructura del cristal para RAT1 de *S. pombe* (PDB 3FQD) y la estructura del cristal para DCP2 de *S. cerevisiae* (PDB 2KQMB). El modelaje y visualización de

las estructuras terciarias de las proteínas se realizó con el visualizador RCSB PDB *protein workshop 3.9 viewer*.

2. Cultivos de trofozoítos de *E. histolytica*

Trofozoítos cepa HMI:IMSS de *E. histolytica* se cultivaron axénicamente a 37° C en medio TYI-S-33 (biosate, glucosa anhídrida, NaCl, KH₂PO₄, K₂HPO₄, L(+) ácido ascórbico, L-cisteína, citrato férrico, agua desionizada, pH 6.8), complementado con 15% de suero bovino (GIBCO), 3% de mezcla de vitaminas *Diamond* (SIGMA) (Diamond *et al.*, 1978) y antibióticos 0.25 U/ml de bencilpenicilina (Laboratorios Pisa) y 35 µg/ml de estreptomina (Lakeside). Para realizar los experimentos, se cosecharon trofozoítos en fase logarítmica de crecimiento mediante incubación en hielo durante 15 min. y se agitaron suavemente hasta despegarlos por completo de las paredes del tubo o caja de cultivo. Las células se empastillaron mediante centrifugación a 3000 rpm /10 min. en una centrifuga Sorvall Biofuge *primo R* a 4° C en tubos Falcon (Nalgene) de 50 ml. El sobrenadante se eliminó y el paquete celular se lavó por centrifugación con solución salina amortiguadora de fosfatos (PBS por sus siglas en inglés) (NaCl 140 mM, KCl 2.7 mM, Na₂PO₄ 8 mM, KH₂PO₄ 8 mM, pH 7.2) durante 5 min. a 4° C. Durante el último lavado, se tomó una alícuota de la suspensión y se contaron las amibas en una cámara de Neubauer. La viabilidad celular se controló mediante microscopía utilizando azul de tripano que es un colorante de exclusión. Los trofozoítos se utilizaron de manera inmediata para los siguientes experimentos.

2.1. Choque térmico de los trofozoítos a 42° C

Una caja mediana (75cm²) de cultivo confluyente de trofozoítos (10 x10⁶ células) de la cepa HMI:IMSS se incubó a 42° C/ 4 h. (Weber *et al.*, 2006), posteriormente se realizó la extracción de RNA total con la técnica de Trizol (Invitrogen).

2.2. Irradiación de los trofozoítos con luz UV-C

Trofozoítos cepa HMI:IMSS se incubaron durante 15 min., en cajas Petri de vidrio, posteriormente se les retiró el medio y las células no adheridas. Las células pegadas al fondo de las cajas fueron sometidas a irradiación con una dosis de 150 J/m² de luz UV-C durante 8 seg., en un Stratalinker 1800 (Stratagene). A continuación, se les añadió medio TYI-S33 completo previamente calentado y se incubaron durante 3 h/37°C (Lopez-Casamichana *et al.*, 2008), para continuar con la extracción del RNA total.

2.3. Tratamiento con nitropusiato de sodio

Trofozoítos cepa HMI:IMSS de *E. histolytica* se cultivaron axénicamente a 37° C en tubos de 15 ml. Las células fueron tratadas con 1mM de nitropusiato de sodio (SIGMA) durante 1h (Ramos *et al.*, 2007). Posteriormente se les retiró el medio y se les cambio por nuevo. La obtención de RNA total se realizó a las 24 h después de retirar el compuesto químico.

3. Obtención del DNA total de *E. histolytica* cepa HMI:IMSS

Para la obtención del DNA total a partir de trofozoítos, se utilizó el protocolo de extracción de DNA mediante el uso del kit Wizard Genomic DNA purification

(Promega). Se cosechó una caja mediana (10×10^6 células) por centrifugación a 3000 rpm/ 10 min. en la centrifuga Sorvall Biofuge *primo R* a 4° C en tubos Falcon (Nalgene) de 50 ml. El sobrenadante se eliminó y las células se resuspendieron y lisaron con la adición de 1 200 μ l de amortiguador de lisis (Nuclei Lysis Solution), la suspensión se distribuyó en tubos *Eppendorf* nuevos de 1.5 ml (800 μ l a cada tubo), se les adicionaron 200 μ l de solución de precipitación de proteínas (Protein Precipitation Solution) y se incubaron por 5 min., en hielo. Los tubos se centrifugaron a 13 000 rpm/4 min., se recuperó el sobrenadante y se transfirió a tubos *Eppendorf* de 1.5 ml y se adicionaron 600 μ l de isopropanol, mezclando cuidadosamente por inversión. Se centrifugaron los tubos a 13 000 rpm/5 min. a temperatura ambiente (T.A), el sobrenadante se decantó y la pastilla se secó durante 15 min. Por último, se adicionaron 30 μ l de agua grado biología molecular (SIGMA) para rehidratar el DNA. Por último el DNA se cuantificó por espectrofotometria a 260 nm. Las alícuotas se almacenaron a -20° C para su posterior uso.

4. Diseño, síntesis y cuantificación de los oligonucleótidos

Después de realizar los alineamientos y determinar cuáles fueron las regiones más conservadas entre las proteínas, con el programa Oligo se diseñaron oligonucleótidos específicos para amplificar los genes de interés y realizar los clonaciones y ensayos de RT-PCR cuantitativa (Tabla 4). Todos los oligonucleótidos utilizados, se generaron por síntesis química (IDT) y fueron cuantificados a través de visualización por UV en geles de agarosa teñidos con

bromuro de etidio y espectrofotométricamente mediante la absorción a una longitud de onda de 260 nm (Biomate 3 Thermo Spectronic). Para determinar la concentración de los oligonucleótidos se utilizó la siguiente fórmula: $(\mu\text{g/ml}) = \text{Abs}_{260 \text{ nm}} \times 33 \times \text{factor de dilución}$.

Gen	Primer sentido	Primer Antisentido	Producto (bp)	Tm (°C)
<i>Ehago2-2</i>	5'- AAGAATCTAGCCGACGTGAC-3'	5'-TCTCTTCTTGGTGAACGGTC-3'	100	51.5
<i>Ehdcp2</i>	5'-TGACTTGTGTGCACGATTTG-3'	5'-GCTCCTACTACTGATATTCTTGACT-3'	267	50.8
<i>Ehxrnl</i>	5'-TGACAGAAGTGATAATGATATTTGC-3'	5'-TGTTTCAGGTGATTGTCGGT-3'	290	50
<i>Ehedc3</i>	5'- GGTTAGCTGTTGCTCAAGTTG--3'	5'-ACAAACCAACCCATCTCCTC-3'	110	51.7
<i>Ehlsm1</i>	5'-ACGTACATTTTCTAAAGAAGGAGTT-3'	5'-GCTTGAGAAAAGTGGATTGGAAA-3'	150	51
<i>Ehcaf1</i>	5'-CATATCCAGGTATGGAATACCCAT-3'	5'-ATGGTGGAATTGTCTATTGGG-3'	120	51
<i>Ehnot1</i>	5'-TGCTCTGAACATTTTCAAATGG-3'	5'-GACCATATCATCCATCAACGC-3'	195	50.5
<i>Ehupf1</i>	5'-TGAGGAAAATGGCATGGGAG-3'	5'-TGTCTATCAGCTTCAGCCTC-3'	163	51.5
<i>Ehrnasa III</i>	5'-AGCTCAACTACATTACACAATGC-3'	5'-TGCTCGTTCTTCAAGATAGGA-3'	183	51
<i>EhL31</i>	5'-AGAGAGCTCCAAAAGCCATC-3'	5'-CTACGGTTTTAGTTACGAGACC-3'	244	52

Tabla 4. Oligonucleótidos utilizados en el ensayo de RT-PCR cuantitativa

5. Análisis del RNA

5.1. Extracción y cuantificación de RNA total de alta calidad

La extracción del RNA total se realizó utilizando el reactivo comercial TRIzol (mezcla de fenol-isotiocianato de guanidina), siguiendo las indicaciones del proveedor (Invitrogen). Se cosechó una caja mediana (10×10^6 células) por centrifugación a 3000 rpm/10 min, en la centrifuga Sorvall Biofuge *primo R* a 4° C en tubos Falcon (Nalgene) de 50 ml. El sobrenadante se eliminó y las células se resuspendieron y lisaron con 500 µL de TRIzol. La mezcla se incubó a T.A durante 5 min. y se agregaron 300 µl de cloroformo. Inmediatamente de mezclar por inversión suave para homogenizar, se incubó durante 3 min a T.A., se centrifugó a 8000 rpm/ 5 min. para obtener la fase acuosa. El RNA se obtuvo de la fase acuosa mediante precipitación con 800 µl de isopropanol incubando durante 10 min a T.A., se centrifugó a 8000 rpm/15 min para precipitar el RNA. El sobrenadante se eliminó por decantación y la pastilla de RNA se lavó con etanol al 75%, se centrifugó nuevamente y el sobrenadante se eliminó. Finalmente, la pastilla de RNA se secó durante 10 min a T.A y se resuspendió en agua libre de ribonucleasas tratada con DEPC al 0.1%. El RNA se cuantificó en un espectrofotómetro (Biomate 3 Thermo Spectronic) y se visualizó en un gel de agarosa al 1% en TBE 1X. El RNA en forma pura exhibe una relación lineal entre la concentración y la absorbancia a 260 nm. Dado que una unidad de absorbancia

a 260 nm corresponde a 44 µg/ml de RNA de cadena sencilla, la concentración de RNA pudo estimarse de acuerdo a la siguiente ecuación:

$$[\text{RNA } \mu\text{g/ml}] = A_{260 \text{ nm}} \times 44.19 \times \text{factor de dilución.}$$

5.2. Síntesis de la cadena complementaria de DNA (cDNA)

Con la finalidad de realizar la cuantificación de los transcritos de los genes que codifican para proteínas involucradas en la degradación de RNAs mensajeros y el silenciamiento de genes por RNA de interferencia en diversas condiciones, implementamos la técnica de transcripción reversa. Para la síntesis de las cadenas de cDNA se realizó el protocolo experimental que se describe a continuación.

5.2.1. Tratamiento del RNA total con Dnasa I

Se procedió a mezclar 1 µg de RNA, 1U de la enzima DNasa I (NEB) y el volumen necesario de agua libre de RNAsas para completar a 7 µl la reacción. Se incubo la muestra por 15 min. a 37° C. Posteriormente se adicionó 1 µl de EDTA 25 mM y se incubó 10 min. a 65° C para desactivar la enzima. Se prosiguió con el protocolo para realizar el proceso de síntesis del cDNA por retrotranscripción.

5.2.2. Retrotranscripción

Primeramente determinamos la cantidad de RNA total inicial óptima para lograr detectar diferencias en las concentraciones de los transcritos de interés, de manera tal que no se saturara la reacción de PCR. Para esto se procedió a mezclar diferentes concentraciones de RNA en el rango de 50 a 500 ng

previamente tratados con la DNasa I, con 100 ng de oligo (dT₁₈) en un volumen de 9 µl el cual se completó mediante la adición de agua tratada con DEPC al 0.1%. Las muestras se desnaturalizaron a 64°C/10 min. Posteriormente, a T.A se les adicionó el amortiguador de reacción 5X (Invitrogen), dNTPs 10 mM, DTT 100 mM y 40 U de inhibidor de RNasas (Promega) en un volumen final de 20 µl. Las mezclas de reacción se incubaron a 25°C/5 min. Finalmente, se agregaron 200 U de la enzima reversa transcriptasa Superscript II (Invitrogen) y se incubaron a 42°C/90 min. Transcurrido este tiempo se le adicionó 1 µl de RNasa H con la finalidad de obtener un mayor rendimiento en la reacción de PCR debido a que se elimina el RNA templado y se impide la formación de moléculas dúplex DNA-RNA, que disminuyen el número de hebras cDNA libre que sirven de templado en la reacción de PCR. Para algunos templados, la incubación del RNA y el iniciador a 65°C en ausencia de amortiguador y sales, seguida de enfriamiento en hielo, elimina la mayoría de las estructuras secundarias y permite el perfecto alineamiento del iniciador. Para eliminar la posibilidad de que durante la reacción de síntesis de cDNA existieran o se formaran nuevas estructuras secundarias, realizamos una curva de calibración utilizando las temperaturas 37°C, 42°C y 45°C para determinar la temperatura óptima de síntesis de cDNA. La retrotranscripción a 42°C fue la óptima para lograr una adecuada amplificación por PCR de los transcritos, por lo que los ensayos los realizamos a esta temperatura. La enzima Superscript II (Invitrogen) que utilizamos es una transcriptasa reversa dependiente de RNA la cual es capaz de generar con mayor eficiencia la síntesis de moléculas de cDNA de hasta 12.3 kb y no presenta actividad residual de RNasa H.

5.3. Ensayos de RT-PCR cuantitativa

El RNA total obtenido de trofozoítos cepa HMI:IMSS crecidos en condiciones basales de cultivo y células estresadas por choque térmico, nitropusiato de sodio y células irradiadas fue utilizado para sintetizar el cDNA, como anteriormente se explicó. La reacción en cadena de la polimerasa (PCR por sus siglas en inglés), en tiempo real (RT por su siglas en inglés); (RT-PCR cuantitativa) se realizó con los oligonucleótidos que se presentan en la Tabla III. Tres muestras biológicas independientes fueron analizadas. La reacción de RT-PCR cuantitativa se realizó utilizando un equipo *ABI PRISM 7000 Sequence Detection System* (Applied Biosystems). Las reacciones se realizaron con 500 nM de cada oligonucleótido, 1X *PCR SYBRGreen Master Mix* (Qiagen) y 1 ul de cDNA como templado. Después de la desnaturalización inicial de 15/95°C, continuó el ciclo de amplificación (repetido 40 veces) como sigue: 15 seg/95°C, 15 seg/52°C y 15 seg/72°C. Realizamos una curva de desnaturalización se realizó (55°C/95°C, 2°C/1 min) para comprobar que los amplicones obtenidos fueron específicos para cada gen. Para cada triplicado, la media de las concentraciones relativas obtenidas para el RNAm problema se dividió por la media de los valores correspondientes obtenidos para la amplificación del gen normalizador *EhL31*, correspondiente a la proteína ribosomal L31. Esta relación representa la variación de la abundancia del transcrito en la condición problema en comparación con el control. Cada experimento de PCR se repitió tres veces y tres muestras biológicas independientes fueron analizadas. La expresión relativa de los genes de

degradación de RNAm en los diferentes tratamientos se calculó por el método $2^{-\Delta\Delta Ct}$ (Livak *et al.*, 2001).

6. Clonación de los genes *Ehdcp2*, *Ehxr2* y *Ehago2-2* en el vector de expresión pRSET-A

6.1. Amplificación de los genes a partir de DNA genómico de *E. histolytica*

El DNA genómico se obtuvo a partir de trofozoítos de *E. histolytica* cepa HMI:IMSS, empleando el *kit Wizard Genomic DNA purification* (Promega) siguiendo el protocolo recomendado por el fabricante para aislamiento de DNA genómico de cultivos de tejidos y tejidos animales. Se diseñaron oligonucleótidos sentido y antisentido para cada uno de los genes seleccionados, con los respectivos sitios de restricción para las enzimas *Bam*HI (NEB) y *Hind*III (NEB). Mediante la reacción PCR, se amplificó el fragmento deseado utilizando como molde 50 ng de DNA genómico, una mezcla de amortiguador de reacción 10X, dNTPs 10mM, $MgCl_2$ 2.5 mM, 100 ng de cada oligonucleótido y 1 U de la enzima *Taq* DNA polimerasa (Invitrogen). Se utilizó un programa de 30 ciclos correspondientes a la desnaturalización, alineamiento y extensión. Los productos de amplificación por PCR de los genes se analizaron mediante electroforesis en geles de agarosa al 1% en amortiguador TAE 1X. Los geles se tiñeron con bromuro de etidio y los productos resueltos se visualizaron mediante su exposición a luz UV en un transiluminador (BioRad). Las bandas del gel de agarosa correspondientes a los amplificados de los genes se cortaron y se purificó el DNA

del gel utilizando el kit *Geneclean II* (BIO 101), siguiendo las instrucciones del fabricante.

6.2. Ligación de los genes en el vector pRSET-A

Los productos de PCR previamente purificados, correspondientes a los genes, *Ehdcp2*, *Ehxr2* y *Ehago2-2* y el vector pRSET-A se cortaron paralelamente con las enzimas *Bam*HI (NEB) y *Hind*III (NEB) utilizando 5 U de cada enzima en el amortiguador de reacción 2 (NEB) durante 4 h/37° C. Posteriormente, se analizaron los productos de digestión en un gel de agarosa al 1% en TAE 1X. Los insertos y el vector digeridos se purificaron con el kit *Geneclean II* (BIO 101) y nuevamente se analizaron en un gel de agarosa al 1% para su cuantificación. Los genes se clonaron en el vector de expresión pRSET-A por separado. La reacción de ligación se realizó utilizando una proporción molar vector:inserto de 1:5, 5µl de amortiguador con ATP y 1 U de enzima T4 DNA ligasa (Invitrogen) en un volumen de reacción de 20 µl. La mezcla se incubó durante toda la noche (O.N por sus siglas en inglés) a 16°C. Como control se realizó una mezcla de ligación similar que carecía del inserto. Las reacciones de ligación se utilizaron para transformar células *E. coli* DH5 α competentes, las cuales se sembraron e incubaron O.N a 37°C para posteriormente picar colonias y crecerlas O.N a 37°C en medio líquido Luria Broth (LB por sus siglas en inglés) con ampicilina para la obtención de DNA plasmídico por lisis alcalina y por último se realizó un análisis por digestión enzimática de las clonas positivas. Las clonas positivas se seleccionaron, algunas de ellas se expandieron para la obtención y

purificación de los plásmidos correspondientes para utilizarlos en la expresión de las proteínas recombinantes.

7. Expresión y purificación de las proteínas recombinantes

7.1. Preparación de las células competentes *E. coli* BL21 (DE3) pLysS

Para la obtención de células competentes, bacterias de la cepa BL21(DE3)pLysS fueron sembradas en una placa de medio LB y se incubaron O.N/ 37° C. Al día siguiente, una colonia se sembró en 5 ml de medio LB líquido y se incubó a 37°C hasta que alcanzó una D.O a 600 nm de 0.5. Entonces se realizó una dilución 1:20 de este cultivo en 50 ml de medio LB líquido fresco y se permitió que creciera a 37°C hasta una DO de 0.5 a 600 nm. Posteriormente, las células se incubaron en hielo durante 20 min y se centrifugaron a 7000 rpm en la centrifuga Sorvall Biofuge *primo R* por 5 min/ 4°C; la pastilla se resuspendió en 20 ml de solución de CaCl₂ 1 mM, se incubó por 5 min/ 4°C y se centrifugó a 7000 rpm durante 5 min/4°C. Finalmente, la pastilla se resuspendió en 4 ml de solución de CaCl₂ 1 mM. Las células se utilizaron inmediatamente.

7.2. Transformación de células competentes mediante choque térmico

100 µl de bacterias *E. coli* BL21 (DE3) pLysS competentes se transformaron con 1 µg de los plásmidos recombinantes, pRSET-*Ehxn2*, pRSET-*Ehdcp2* y pRSET-*Ehago2-2*, se incubaron por 20 min/4°C y después por 90 seg/42°C en un termobloque. Las células se pasaron inmediatamente a 4°C/3 min. Después, se agregaron 200 µl de medio LB y las bacterias se incubaron por 1 h/37° C en agitación a 200 rpm. Por último, las bacterias se espatularon sobre cajas Petri con

medio LB sólido con 100 µg/ml de ampicilina y 34 µg/ml de cloranfenicol. Las cajas se incubaron a 37° C/O.N.

7.3. Expresión de las proteínas recombinantes

Las bacterias *E. coli* BL21 (DE3) pLysS previamente transformadas con los plásmidos pRSET-*Ehxrn2*, pRSET-*Ehdcp2* y pRSET-*Ehago2-2* se inocularon en 100 ml de medio LB con ampicilina 100 µg/ml y cloranfenicol 34 µg/ml a 37° C en agitación a 200 rpm hasta alcanzar una D.O. a 600 nm de 0.6. En este momento se tomó una muestra de bacterias correspondiente al tiempo 0 o tiempo sin inducción. Posteriormente, se agregó isopropil-β-D-tiogalactósido (IPTG) a una concentración final de 1 mM y se dejó el cultivo crecer durante 3 h adicionales a 37° C. Se cosecharon las bacterias a 7000 rpm/5 min y la pastilla se guardó a – 20°C hasta su uso.

7.4. Cuantificación de proteínas por el método de Bradford

Para determinar la cantidad de proteínas obtenidas se siguió el método de cuantificación descrito por Bradford (Bradford, 1976). Para esto, se realizó una curva patrón, brevemente en tubos *Eppendorf* de 1.5 ml de volumen se colocaron 0, 5, 10, 15 y 20 µl correspondiente a 0, 2.5, 5, 7.5 y 10 µg, respectivamente de albúmina sérica de bovino (BSA) (NEB). A cada tubo se le agregó la cantidad necesaria de una concentración inicial de NaCl 150 mM para tener un volumen final de 100 µl. Luego se agregó 1 ml del reactivo de Bradford (100 mg de azul de Comassie G-250, 50 ml de etanol al 95%, 100 ml de ácido fosfórico al 85%). Los tubos se agitaron vigorosamente y se incubaron 3 min/T.A. Después, se determinó

la absorbancia de las muestras al medirlas a una longitud de onda de 595 nm, en la región visible del campo electromagnético, utilizando un espectrofotómetro automático (Biomate 3 Thermo Spectronic). Los datos obtenidos se graficaron realizando una regresión lineal, obteniendo así los valores de la pendiente (m) y la ordenada al origen (b). Para medir la concentración de las proteínas de interés, se tomaron alícuotas de 5 μ l de cada muestra y se siguió el mismo procedimiento. Los datos obtenidos se interpolaron empleando los valores de m y b en la curva patrón, determinando así las cantidades exactas de proteínas en cada muestra.

7.5. Electroforesis de proteínas en geles de poliacrilamida en condiciones desnaturalizantes

La electroforesis en geles de poliacrilamida es uno de los métodos más utilizados para la purificación, análisis y caracterización de proteínas. La técnica permite separar moléculas cargadas y se basa en diferencias en movilidad cuando se les somete a la acción de un campo eléctrico. La electroforesis se lleva a cabo sobre un soporte inerte, generalmente sobre geles de poliacrilamida (PAGE por sus siglas en inglés). En la electroforesis desnaturalizante (SDS-PAGE), como su propio nombre indica, se utilizan agentes desnaturalizantes de proteínas, como pueden ser detergentes (SDS), agentes caotrópicos (urea) y agentes reductores (2-mercaptoetanol, DTT). Para la visualización de las proteínas se utilizan diferentes métodos de tinción, siendo el más habitual el azul de Coomassie. En la electroforesis SDS-PAGE, la separación es en función del tamaño molecular, lo que permite determinar el peso molecular de las proteínas.

El volumen de muestra correspondiente a 50 µg de proteínas se colocó en un tubo *Eppendorf* de 1.5 ml y se les adicionó un volumen igual de solución amortiguadora de muestra Laemmli 2X (125 mM de Tris-HCl pH 6.8, SDS al 4% y glicerol al 20%), posteriormente se le agregó 14 µl/ml del agente reductor, β-mercaptoetanol al 5%. Las mezclas se hirvieron/5 min en baño María y se colocaron en hielo hasta el momento de cargarse en el gel.

Las proteínas se resolvieron electroforéticamente mediante SDS-PAGE al 12%. Para preparar los geles se empleó el sistema de Bio-Rad (Mini-Protean II System). Todas las soluciones se prepararon en agua destilada estéril y filtrada a través de una membrana de 0.45 micras (Millipore). Las soluciones utilizadas para la preparación de los geles fueron las siguientes: solución acrilamida 29:1, que contiene 29.2% de acrilamida (Bio-Rad) y 0.8% de bis-N, N' metilen-bis-acrilamida (Bio-Rad), solución amortiguadora de separación que contiene Tris (Sigma) 1.5 M, pH 8.8 ajustado con HCl, solución amortiguadora concentradora, contiene Tris 0.5 M, pH 6.8 ajustado con HCl; SDS al 10%; persulfato de amonio (Bio-Rad) al 10% y como catalizador Tetrametiletilenediamina (TEMED por sus siglas en inglés) (Bio-Rad). Los geles se prepararon en molde en vidrio de 10 x 8.2 cm, con separadores de 1.5 mm. de espesor. Una vez armado el molde y sellado, se vaciaron 10 ml de la solución del gel separador; después se adicionaron 2 ml de agua para aplanar el gel. Cuando el gel polimerizó (≈ 20 min.) se retiró el agua y se agregó en el espacio restante del molde la solución del gel concentrador e inmediatamente se colocó el peine, formando los pozos en los cuales se inyectaron las muestras y se dejó polimerizar por 20 min adicionales. Después, el gel se montó en la cámara de

electroforesis (Bio-Rad) y se le agregó solución amortiguadora de corrida 1X preparada con 25 mM Tris, glicina 192 mM y SDS al 0.1%.. En cada pozo se colocaron las alícuotas correspondientes y los geles se corrieron durante 2 h 30 min. a 100 V.

7.6. Tinción de los geles de poliacrilamida con azul de Coomasie

Las bandas de proteínas, una vez terminada la electroforesis, se visualizan tiñendo los geles con una solución de azul de Coomasie R-250. La tinción se realiza con agitación suave durante, al menos, 30 min. Posterior del tiempo de tinción y para poder observar claramente las bandas de proteínas, el gel se contrastó con una solución desteñidora (metanol al 30% y ácido acético al 10%) durante toda la noche en agitación ligera. Como marcadores de peso molecular se utilizaron los marcadores preteñidos *Precision Plus Protein Standards Dual color* (BioRad) del rango entre 10-250 kDa.

7.7. Transferencia de proteínas a membranas de nitrocelulosa

Las proteínas separadas por SDS-PAGE al 10%, se transfirieron a membranas de nitrocelulosa (Bio-Rad). Los geles, las membranas de nitrocelulosa y el papel *Whatman* se preincubaron por 5 min en solución de transferencia de proteínas (20% metanol, 10% Tris-Glicina). La electrotransferencia se llevo a cabo a 400 mA por 45 min/4°C. Por último, la presencia de las proteínas en la membrana se confirmó por tinción con rojo de Ponceau (100 mg/ml). Después las membranas se lavaron con agua bidestilada para retirar el colorante y se realizaron los ensayos de *Western blot*.

7.8. Ensayos de *Western blot*

Las proteínas transferidas a las membranas de nitrocelulosa se incubaron durante 2 h. en solución bloqueadora con 5% de leche descremada en PBS 1X. Posteriormente, se agregó el anticuerpo primario anti-His6x (Santa Cruz) en una dilución 1:1000 en PBS 1X y 2% de leche descremada por 2 h. a temperatura ambiente en agitación constante. Después las membranas se lavaron tres veces con PBS-Tween 20 al 0.05% y se les agregó el anticuerpo secundario anti-IgG de ratón acoplado a peroxidasa (Zymed) en una dilución 1:5000 en PBS 1X y leche descremada al 2% por 1 h. a temperatura ambiente en agitación. Por último, las membranas se lavaron tres veces con PBS-Tween 20 al 0.05% y se revelaron por quimioluminiscencia con el reactivo ECL plus (Amersham Pharmacia Biotech) utilizando placas de auto-radiografía (Kodak).

8. Purificación de las proteínas recombinantes y generación de anticuerpos

Una vez evaluada la expresión de las proteínas recombinantes, se analizó la solubilidad de cada una de las proteínas para la elección del protocolo de purificación de la misma. La purificación se realizó con columnas de afinidad Ni-NTA (Qiagen). Con las proteínas purificadas y cuantificadas se iniciaron las inmunizaciones siguiendo el protocolo diseñado para la obtención de anticuerpos en ratones y en conejo. Suero pre-inmune e inmune se obtuvieron y se verificó su especificidad mediante *Western blot* utilizando las proteínas recombinantes.

9. Inmunolocalización de las proteínas endógenas en trofozoítos

Cultivos de trofozoítos *E. histolytica*, se obtuvieron cosechando aproximadamente 10×10^6 de células por ensayo y se utilizó el volumen adecuado en cada pozo de placas de 6 pozos (Corning) previamente preparadas con un cubreobjetos por pozo esterilizados por flameo y se incubaron a 37°C/ 60 min para la adhesión de las amibas a la superficie de vidrio. Una vez pegadas las amibas se eliminó cuidadosamente el medio y las amibas sueltas con una pipeta Pasteur. Se realizaron lavados con PBS1X calentado a 37°C. Se fijaron las células con paraformaldehído al 4% por 40 min/37°C, y se lavaron con PBS1X. A continuación se permeabilizaron las amibas con Tritón al 0.5% durante 40 min/37°C y se lavaron con PBS1X. Se bloquearán las células con PBS1X/BSA 1% durante 60 min/37°C. A continuación las células se incubaron con los anticuerpos primarios con anti-*EhCAF1* y anti-*EhAGO2-2* generados en conejo, o con los anticuerpos generados en ratón anti-*EhXRN2*, and anti-*EhDCP2* (1:200) 1 h/37°C Las células fueron lavadas tres veces e incubadas con los anticuerpos secundarios fluoresceinados (1:200) 1 h/37°C. Finalmente se realizó la tinción de los núcleos con 4'-6-diamidino-2-fenilindola (DAPI por sus siglas en inglés) en una concentración de 5mg/ml 1 min/37°C. Las muestras fueron analizadas a través de microscopia confocal con el sistema Leica TCS SP2 (Leica, CO).

10. Ensayos de inmunoprecipitación

Utilizamos el protocolo previamente descrito por Tovy *et al*, con mínimas modificaciones. Brevemente, proteínas citoplasmáticas (50 µg) se diluyeron en 300 µl de amortiguador HNTG (20 mM Hepes pH 7.5, 150 mM NaCl, 0.1% Tritón,

10% glicerol), y se incubaron 10 µl de perlas de proteína G (Sigma) por 30 min/ 4°C. Las interacciones de inespecíficas de proteínas se excluyeron por centrifugación 3000 rpm a 4°C/ 5 min. El sobrenadante se incubo por separado con los anticuerpos anti-*EhCAF1* (1:100) o anti-*EhXRN2* (1:100) por 2 h/4°C. Posterior a la incubación a las perlas de proteína G se les adicionaron las muestras e incubadas a 16 h/4°C. Las proteínas inmunoprecipitadas se colectaron por centrifugación, se lavaron tres veces con el amortiguador HNTG, y se separaron por geles SDS-PAGE al 12% (Tovy *et al.*, 2010). Las proteínas se transfirieron a membranas de nitrocelulosa, y detectadas mediante ensayos de *Western blot* utilizando los anticuerpos anti-*EhCAF1* de conejo o anti-*EhXRN2* de ratón.

11. Expresión de los dsRNA y transfección de trofozoítos

El dsRNA se obtuvo con la técnica previamente descrita por Solis *et al*, la secuencia completa del gen *Ehpc4* (EHI_192520) se clono con los sitios de restricción *SmaI* y *XhoI* en el plásmido pL4440, el cuál es flanqueado bidireccionalmente por promotores T7 (Timmons *et al.*, 2001). Células competentes *E. coli* HT115 se transformaron con el vector pL4440-*Ehpc4* y espatuladas en cajas LB-agar ampicilina/tetraciclina. Colonias aisladas se obtuvieron después de incubarlas a 37°C/O.N. La síntesis de los dsRNA se indujo por la adición de IPTG 2 mM durante 4 h/37°C. El marcaje con el fluoroforo Cy3 de los dsRNA purificados se realizo de acuerdo las recomendaciones del fabricante para el uso del *Silencer siRNA labeling kit* (Ambion). La mezcla de reacción, 5 µg de dsRNA, agua libre de nucleasas, amortiguador de marcaje y el reactivo de

marcaje Cy3 se incubó a 37°C/ 1h protegida de la luz. Los dsRNAs marcados se precipitaron con etanol y resuspendieron en agua DEPC e inmediatamente se utilizaron para los ensayos de inmunolocalización por microscopia confocal.

12. Hibridación *in situ* fluorescente de RNA

Para la detección de los RNA poli(A)⁺ por hibridación *in situ* fluorescente (FISH), parásitos se cosecharon y se adhirieron a cubreobjetos. Las células se fijaron con paraformaldehído al 4% en PBS, y desnaturalizados en formamida al 70% a 70°C, seguidos de 10 min de incubación con NH₄Cl 25 mM. Los parásitos se permeabilizaron y bloquearon durante 30 min en Triton X-100 al 0.5% (Sigma), BSA 2%, seguido de por una pre-hibridación de 2 h a T.A. en solución de hibridación (2% BSA, 4X SSC, 5% dextran, 35% formamida desionizada y 10 U de inhibidor de RNAsas). La hibridación se realizó durante toda la noche a T.A. en una cámara húmeda en presencia de 300 ng/ml del oligo-(dT)₃₀ conjugado a FITC en solución de hibridación. Las laminillas se lavaron una vez en solución 4X SSC más formamida desionizada al 35%, una vez con 4X SSC y las células se contratiñeron con DAPI (1 mg/ml). Por último las laminillas se montaron con el medio de montaje *Vectashield* (Vector). El análisis de localización subcelular se realizó en un microscopio Leica TCS SP2. La combinación de las imágenes se obtuvo por sobreposición de las imágenes indicadas utilizando el programa Leica LCS lite (Leica Microsystem).

13. Preparación de los sustratos para el ensayo de desadenilación

Para estudiar la desadenilación, se sintetizaron sustratos de RNA *in vitro*. La preparación de los sustratos se realizó utilizando una sonda de RNA de 256 pb correspondiente al gen *EhPgp5* (Lopez-Camarillo *et al.*, 2003) con un tracto de poli (A₅₀). La sonda de RNA se obtuvo por transcripción *in vitro* utilizando la RNA polimerasa T7.

14. Ensayo de desadenilación *in vitro*

El ensayo de desadenilación *in vitro* se realizó con 20 mM Tris/Cl (pH 7.0), 150 mM NaCl, 2 mM MgCl₂, 5 U RNasin (Promega), 5 µg de RNA y 5 µg proteína recombinante *EhCAF1* purificada. La mezcla de reacción de 20 µl de volumen se incubó a 30 °C/1 h. Las reacciones se detuvieron con la adición de EDTA y se resolvieron en geles de poliacrilamida (19:1) al 18% con urea 7 M que fueron teñidos con una solución de azul de toluidina al 1%.

IV. RESULTADOS

1. Análisis *in silico* del repertorio de proteínas involucradas en la degradación de RNAm

Con el fin de identificar genes homólogos potencialmente implicados en la degradación del RNAm en *E. histolytica* inspeccionamos el genoma del parásito en la base de datos *AmoebaDB* (<http://amoebadb.org/amoeba/>) utilizando como templatado las secuencias de aminoácidos de las proteínas implicadas en la degradación del RNAm en levadura y humano. Identificamos posibles genes codificantes para proteínas que participan en los mecanismos de decaimiento del RNAm, tales como desadenilación, *decapping*, NMD y RNA de interferencia (Figura 5, Tabla 5). Nuestros datos indican que los mecanismos del decaimiento del RNAm están bien conservados en *E. histolytica*, aunque carece de varios genes importantes, detectamos genes putativos que codifican para proteínas de *decapping* como los factores *Ehdcp2*, *Ehlsm1-6*, y *Ehedc3*, *Ehdhh1*, así como el gen homólogo para la 5'-3' exoribonucleasa *Ehxr2*. Por el contrario, no encontramos homólogos de los genes DCP1 o XRN1. También identificamos los genes que codifican para las desadenilasas *EhCAF1* y *EhCAF1-like*, así como las subunidades NOT1-5, mientras que el gen homólogo para la desadenilasa CCR4 no se encontró, lo que sugiere la existencia de un complejo CAF1/NOT en donde faltan una de las principales desadenilasas. Curiosamente, no se detectaron los genes *pan2*, *pan3* y *parn*, lo que indica que la maquinaria de desadenilación en *E. histolytica* es más pequeña que las descritas para levadura y humano. El genoma

de la amiba también contiene genes para la maquinaria NMD representado por tres genes *Ehupf* (Tabla 5). También detectamos componentes de las vías del RNA de interferencia, que incluyen a las proteínas *EhRdRP*, *EhRNAsaIII* y *EhAGO2-2*, que previamente habían sido reportadas (Abed *et al.*, 2005) (Zhang *et al.*, 2008). Al igual que en otros trabajos, no detectamos ningún homólogo para DICER y la proteína GW182, de hecho se ha sugerido que en el parásito *E. histolytica* el silenciamiento del RNAm puede ocurrir por pequeños RNA de interferencia generados por mecanismos independientes a DICER. (Zhang *et al.*, 2011).

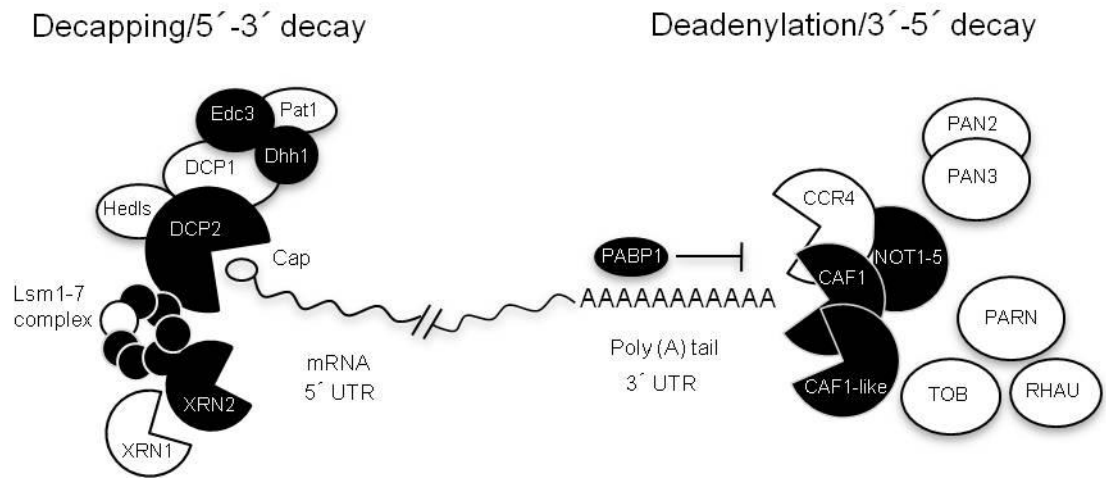


Figura 5. Maquinarias de degradación del RNAm con las proteínas predichas para *E. histolytica*. El inicio de la degradación de RNAm comienza con la desadenilación, posteriormente se puede continuar la ruta de degradación en dirección 3'-5' o en dirección 5'-3' por diferentes exoribonucleasas.

<i>Entamoeba histolytica</i>			<i>Homo sapiens</i>						
Proteína predicha	Tamaño (aa)	Locus ^b	Proteína	No. de acceso ^a	Tamaño (aa)	e-value	I (%)	H (%)	Función reportada
Proteínas involucradas en <i>decapping</i>									
-	-	-	DCP1	Q9NPI6	582	-	-	-	Enzima de <i>decapping</i>
EhDCP2	233	EH1_058810	DCP2	Q8IU60	420	8.4e-25	28	50	Subunidad catalítica de la enzima de <i>decapping</i>
-	-	-	Hedls/Ge-1	Q6P2E9	1401	-	-	-	Componente del complejo de <i>decapping</i>
-	-	-	RAP55	Q8ND56	463	-	-	-	Proteína con dominios Sm-like y FDF; Involucrada en represión de la traducción
-	-	-	XRN1	Q81ZH2	1706	-	-	-	
EhXRN2	938	EH1_133330	XRN2 (Rat1p)	Q9H0D6	950	4.9e-139	46	64	5' to 3' exonucleasa
EhLsm1	110	EH1_188020	Lsm1	O15116	133	8.4e-19	31	59	Proteínas Sm-like involucradas en <i>decapping</i> , <i>splicing</i> del mRNA
EhEdc3	576	EH1_198940	Edc3	Q96F86	508	8.9e-10	29	45	Activador de <i>decapping</i>
EhDhh1	392	EH1_093900	Dhh1	P26196	483	1.7e-72	38	61	RNA helicasa requerida para la represión de la traducción. Activador de <i>decapping</i>
Proteínas involucradas en desadenilación									
EhCAF1	312	EH1_048150	CAF1 (Pop2p)	Q9UIV1	285	3.6e-56	46	63	Deadenilasa (complejo CCR4-NOT)
-	-	-	PAN2	Q504Q3	1202	-	-	-	Deadenilasa
-	-	-	PAN3	Q58A45	741	-	-	-	Estimula la actividad de PAN2
-	-	-	PARN	O95453	639	-	-	-	3'-exoribonucleasa
EhCCR4	509	EH1_026640	CCR4	Q9ULM6	557	7.8e-8	28	51	Deadenilasa (complejo CCR4-NOT)
Proteínas involucradas en NMD									
EhUpf1	938	EH1_035550	UPF1	Q92900	1129	3.5e-36	48	62	RNA helicasa necesaria para NMD
-	-	-	SMG5	Q92900	1129	-	-	-	Factor requerido para la desfosforilación de UPF1
-	-	-	SMG7	Q92540	1137	-	-	-	Factor requerido para la desfosforilación de UPF1
Factores del RNA interferente									
EhAGO2a	938	EH1_125650	AGO2	Q9UKV8	859	1.9e-23	24	42	Silenciamiento de genes mediado por RNA
EhRNAse III	257	EH1_068740	DICER	Q9UPY3	483	7.6e-8	22	51	Silenciamiento de genes mediado por RNA
-	-	-	GW118	Q8NDV7	1962	-	-	-	Silenciamiento de genes mediado por RNA
-	-	-	TNR6B	Q9UPQ9	1833	-	-	-	Silenciamiento por miRNA/ siRNA
-	-	-	MOV10	Q9HCE1	1003	-	-	-	Dead box helicasa, silenciamiento por miRNA/ siRNA
Proteínas de unión a RNA y represores de la traducción									
-	-	-	CPEB	Q9NPI6	566	-	-	-	Poliadenilación citoplasmática; regulador traduccional
EhTTP	213	EH1_008770	TTP	P26651	326	9.6e-24	64	71	Proteína de unión a elementos ricos en AU,
Ehelf4E	201	EH1_180000	eIF4E	P06730	217	1.2e-27	35	57	Componente del complejo de <i>decapping</i>
-	-	-	eIF4E-T	Q8ND56	463	-	-	-	
EhGemin 5	464	EH1_130870	Gemin5	Q8TEQ6	1508	4.9e-12	22	38	Proteína de unión a eIF4E, <i>splicing</i>
^a swiss-Prot/Tr-EMBL databases ^b AmoebaDB database I, Identidad; H, homología NMD, Nonsense-mediated mRNA decay Componentes básicos de los P-bodies son marcados en negritas. Proteínas de levadura homologas a las proteínas de humano en los P- bodies están en paréntesis. Para las familias de las proteínas (XRN2, Lsm 1-8, Upf1-3, eIF4E) sólo un miembro representativo se muestra									

Tabla 5. Proteínas de degradación predichas en *Entamoeba histolytica*

2. Los genes de degradación del RNAm se expresan diferencialmente en condiciones de estrés

Analizamos si los genes de *E. histolytica* involucrados en la degradación de RNAm que se transcriben en trofozoítos cultivados en condiciones basales son modulados en respuesta al estrés celular inducido por choque térmico, daño al DNA inducido por la radiación UV-C y en respuesta al tratamiento con nitropusiato de sodio. Con los oligonucleótidos específicos diseñados (Tabla 4) se realizaron los ensayos de PCR cuantitativa en tiempo real para los genes representativos de los mecanismos de degradación del RNAm (Figura 6). Después del choque térmico la expresión de los genes *Ehxr2*, *Ehls1*, *Ehupf1*, y *Ehrnasa III* aumento de 1,5 hasta 12 veces genes, mientras que los genes *Ehcaf1*, *Ehnot1*, *Ehdcp2* y *Ehago2-2* fueron reprimidos 2 veces en comparación con las células no tratadas (Figura 6). Después de la radiación UV-C el gen *Ehupf1* se reguló positivamente hasta 6 veces, mientras que la expresión de *Ehcaf1*, *Ehnot1*, *Ehdcp2*, *Ehls1*, *Ehedc3* y *Ehrnasa III* fue reprimida de manera significativa. Los niveles de RNAm de *Ehxr2* y *Ehago2-2* se mantuvieron sin cambios. Después del tratamiento con nitropusiato de sodio, los genes *Ehnot1* y *Ehdcp2* se reprimieron de manera significativa, mientras que *Ehcaf1*, *Ehls1*, *Ehupf1*, *Ehago2-2* y *Ehrnasa III* se indujeron de 2 a 16 veces en comparación con trofozoítos no tratados. Como normalizador endógeno se cuantificó la expresión del RNAm de la proteína ribosomal L31 cuya expresión se mantuvo sin cambios en las condiciones estudiadas.

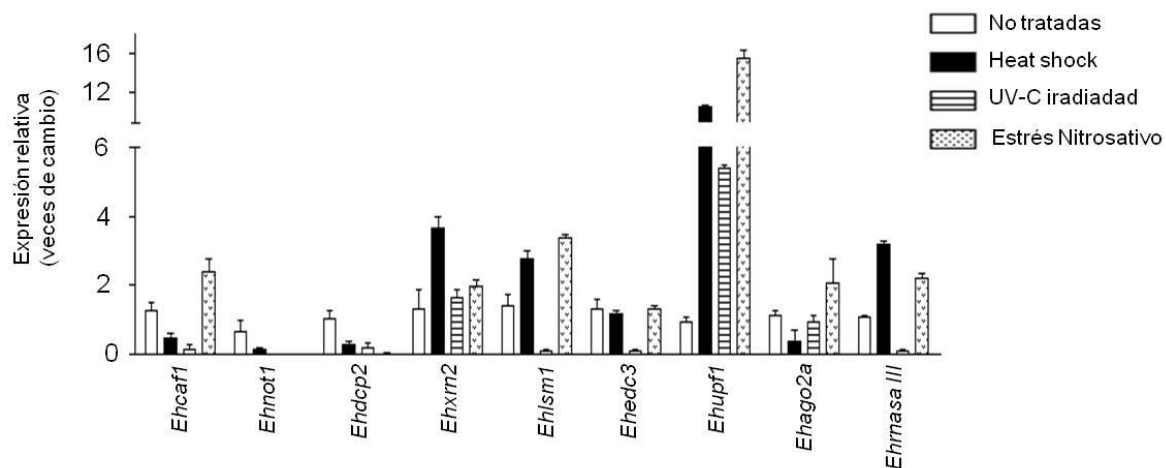


Figura 6. Perfil de expresión de los genes involucrados en la degradación del RNAm en *E. histolytica*. Ensayos de PCR cuantitativa en tiempo real se realizaron para analizar la expresión relativa de los genes representativos de las maquinarias degradación del RNAm después de choque térmico, radiación UV-C y el tratamiento con nitroprusiato de sodio. Para cada triplicado, la media de las concentraciones relativas obtenidos para el RNAm de los genes estudiados se dividió por la media de los valores correspondientes obtenidos para la amplificación del normalizador endógeno correspondiente al transcrito de la proteína ribosomal L31. Cada experimento de PCR se repitió tres veces y tres muestras biológicas independientes fueron analizadas. La expresión relativa (veces de cambio) de los genes de degradación del RNAm en los diferentes tratamientos se calculó por el método doble $\Delta\Delta Ct$. Las barras indican el valor promedio $\pm$ desviación estándar que se muestra.

3. *EhXRN2* y *EhDCP2* conservan la arquitectura típica de las proteínas 5'-3' exoribonucleasas y las de *decapping*

Seleccionamos las proteínas *EhXRN2* y *EhDCP2* debido a que sus homólogos en humano y levadura han sido definidos como marcadores de *P-bodies* (Parker *et al.*, 2007). Proteínas homólogas a XRN2 y DCP2 de diversos organismos fueron seleccionadas por análisis de BLAST y alineamientos múltiples en ClustalW de las secuencias de aminoácidos predichas (Figura 7 y 8) (Tabla 6 y 7). *EhXRN2* es un gen sin intrones de 2813 nt, que predice un polipéptido de 938 aa. Búsquedas por similitud de secuencia mostró que *EhXRN2* exhibe 46% ($e=4.9e-139$) y 45% ($e=1.2e-108$) de identidad con la proteína XRN2 de *Homo sapiens* y RAT1 de *Schizosaccharomyces pombe*, respectivamente. La inferencia filogenética con las secuencias completas de proteínas XRN2 reveló que *EhXRN2* está estrechamente relacionada con proteínas homólogas de otros protozoos (Figura 9). La comparación de las secuencias de aminoácidos de *EhXRN2* con *HsXRN2* y *SpRAT1* reveló la presencia de motivos conservados funcionales y estructurales representadas por los dominios nucleasa XRN_N y el *tower domain* (Figura 11A), que ha sido reportado como responsable de la actividad catalítica en *SpRAT1* (Xiang *et al.*, 2009). La identidad entre *EhXRN2* y *SpRAT1* nos permitió construir un modelo estructural de *EhXRN2*, utilizando como templado los datos del cristal para *SpRAT1*. La estructura terciaria predicha para *EhXRN2* mostró que el dominio N-terminal está constituido por seis α -hélices y dos cadenas β , la estructura muestra una orientación y longitud similar a la descrita para el modelo estructural de *SpRAT1* (Figura 11B-C).

El gen *Ehdcp2* es de 699 nt contiene dos intrones y predice un marco de lectura abierto (ORF por sus siglas en inglés) de 233 aminoácidos. *EhDCP2* exhibe 28% ($e = 8.4e-25$) y 29% ($e = 1.6e-19$) de identidad con las proteínas DCP2 de humano y levadura, respectivamente. La inferencia filogenética reveló que *EhDCP2* está estrechamente relacionada con proteínas homólogas de otros protozoos (Figura 10). *EhDCP2* contiene el dominio conservado DCP2 y el motivo NUDIX (Figura 11D). En particular, la secuencia consenso para el motivo NUDIX, $GX_5EX_7REUXEEXGU$, está presente en *EhDCP2*, incluyendo los tres residuos de glutamato (E), residuos implicados en la función catalítica. Ambos dominios son los responsables de la eliminación de la estructura de *cap* en células de mamífero (She *et al.*, 2008). El modelaje de la estructura terciaria para *EhDCP2* utilizando como templatado la estructura de *ScDCP2* mostró la conformación y longitud de ambas moléculas son similares (Figura 11E-F).

Proteínas XRN2

```

tr|C4MB40|C4MB40_ENTHI 70 -----VEVDNFIYDMNGLIHPCSHFQD-RQSPETIEDMMVMGRYLDLIIIVRPRKLLYLAV
tr|B0E6C7|B0E6C7_ENTDS 40 -----VEVDNFIYDMNGLIHPCSHFQD-RQSPETIEDMMVMGRYLDLIIIVRPRKLLYLAV
tr|Q8RV84|Q8RV84_ORYSJ 46 -----IEFDNLYLDMNGIIHPCSHPED-KPSPTTFTEVFQCMFDYIDRLFVMVRPRKLLYMAI
sp|Q9FQ02|XRN2_ARATH 48 -----VEVDNLYLDMNGIIHPCSHPED-KPSPTTFTEVFQCMFDYIDRLFVMVRPRKLLYMAI
sp|Q9DBR1|XRN2_MOUSE 46 -----VEFDNLYLDMNGIIHPCSHPED-KPAPKNEDEMMVAIFEYIDRLFIVRPRKLLYMAI
sp|Q9H0D6|XRN2_HUMAN 46 -----VEFDNLYLDMNGIIHPCSHPED-KPAPKNEDEMMVAIFEYIDRLFIVRPRKLLYMAI
tr|A3KNE2|A3KNE2_XENLA 46 -----VEFDNLYLDMNGIIHPCSHPED-KPAPKNEDEMMVAIFEYIDRLFIVRPRKLLYMAI
sp|Q9VM71|XRN2_DROME 47 -----IEFDNLYLDMNGIIHPCSHPED-KPAPKNEDEMMVAIFEYIDRLFIVRPRKLLYMAI
tr|P40848|XRN2_SCHPO 47 -----VEVDNLYLDMNGIIHPCSHPED-KPAPKNEDEMMVAIFEYIDRLFIVRPRKLLYMAI
tr|A2DLG3|A2DLG3_TRIVA 40 -----SQEFDNLYLDMNGLIHPCSHFQD-RECPKNIHEVYENIKLYLRIFIVRPRKLLYLAI
tr|Q8I358|Q8I358_PLAF7 44 KVVHDLNDYVNNDELHCSDNINGYFDNLYLDMNGIIHPCSHFQD-SKRAKSNBEIFINVFYVRLFIIEPRKLLYMAI
tr|Q4Q1P5|Q4Q1P5_LEIMA 42 -----AYSVDNFIYDMNGLIHPCCHDTAPLPEFENEEMFERMFTCHDILVRVVRPRKLLYMAI
tr|E1F363|E1F363_GIAIA 32 -----LYIDFNALIHESIRKISYIIFTFTEAMFGELADYIQVIVSTVTPRRRIETAT
consensus 81 .....*.*.*.*.*.....

tr|C4MB40|C4MB40_ENTHI 127 DGVAPRAKMNQORSRRFSAKDVENAELKH-AEENRNKVRNG---LEVPYIKPLDSN-----CITPGTEFMSVW
tr|B0E6C7|B0E6C7_ENTDS 97 DGVAPRAKMNQORSRRFSAKEVDENRERKQ-AEENRNKVRNG---LEIPYVRPLDSN-----CITPGTEFMSAV
tr|Q8RV84|Q8RV84_ORYSJ 103 DGVAPRAKMNQORSRRFRAAKDAADAAAE-ERLEEFERERGRKIFPK-QQSQCDSN-----VITPGTEFMAVL
sp|Q9FQ02|XRN2_ARATH 105 DGVAPRAKMNQORSRRFRAAKDAADAAAE-ERLEEFERERGRKIFPK-VDSQVFDN-----VITPGTEFMAVL
sp|Q9DBR1|XRN2_MOUSE 103 DGVAPRAKMNQORSRRFRASKEGMAAAVEK-QRVREELIARGGFIFPE-EIKERFDSN-----CITPGTEFMDNL
sp|Q9H0D6|XRN2_HUMAN 103 DGVAPRAKMNQORSRRFRASKEGMAAAVEK-QRVREELIARGGFIFPE-EIKERFDSN-----CITPGTEFMDNL
tr|A3KNE2|A3KNE2_XENLA 103 DGVAPRAKMNQORSRRFRASKEGVSVEBK-NRIRREVLIRGGYIPKE-EMKERFDSN-----CITPGTEFMDNL
sp|Q9VM71|XRN2_DROME 104 DGVAPRAKMNQORSRRFRAAKETTKRLEI-ARIREELISRGCKIFPEKEKEGEHFDN-----CITPGTEFMDNL
tr|P40848|XRN2_SCHPO 104 DGVAPRAKMNQORSRRFRASKEGMAAAVEK-QRVREELIARGGFIFPE-EIKERFDSN-----CITPGTEFMDNL
tr|A2DLG3|A2DLG3_TRIVA 98 DGVAPRAKMNQORSRRFRAAKTSEYNRNKNFOEITLKGQKEAEKIMKEQYIQDHDN-----IITPGTEFMCCL
tr|Q8I358|Q8I358_PLAF7 123 DGVAPRAKMNQORSRRFSAISCSIEIKR---AYILKERFIAENKMVPEETTYWDSN-----VITPGTEFMSHL
tr|Q4Q1P5|Q4Q1P5_LEIMA 101 DGVAPRAKMNQORSRRFRAAKETTKRLEI-ARIREELISRGCKIFPEKEKEGEHFDN-----CITPGTEFMDNL
tr|E1F363|E1F363_GIAIA 85 DGVAPRAKAKOORTRRMGAAEISGVGLYRNNIATATLSEVEFLTSKHIANGNNDQOTLPPQNLKAIIPGTLFMRL
consensus 161 *****.*.*.*.*.....

tr|C4MB40|C4MB40_ENTHI 193 SKDLRYIYIRKMS-QSSYWNKVVILSDSCVPGEGEHKIMDEIRKQESFTYNNP-----IHHCLYGLD
tr|B0E6C7|B0E6C7_ENTDS 163 SKDLRYIYIRKMS-QSSYWNKVVILSDSCVPGEGEHKIMDEIRKQESFTYNNP-----IHHCLYGLD
tr|Q8RV84|Q8RV84_ORYSJ 171 SIALQYYIYIRLN-NDPGWKNIRVILSDASVPGEGEHKIMSYIRKQENLGEHNP-----THHCLYGLD
sp|Q9FQ02|XRN2_ARATH 173 SFALRYIYIRLN-SDPGWKNIRVILSDASVPGEGEHKIMSYIRKQENLGEHNP-----THHCLYGLD
sp|Q9DBR1|XRN2_MOUSE 171 AKCLRYIYIRLN-NDPGWKNIRVILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
sp|Q9H0D6|XRN2_HUMAN 171 AKCLRYIYIRLN-NDPGWKNIRVILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
tr|A3KNE2|A3KNE2_XENLA 171 AKCLRYIYIRLN-NDPGWKNIRVILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
sp|Q9VM71|XRN2_DROME 173 SKCLHYEVHRLN-LNPAAWGRHVILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
tr|P40848|XRN2_SCHPO 173 AKSLRYIYIRLN-SDPGWKNIRVILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
tr|A2DLG3|A2DLG3_TRIVA 168 SKCLRSYISMCN-TSPAWSNRIILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
tr|Q8I358|Q8I358_PLAF7 189 SIALQYYIYIRLN-NDPGWKNIRVILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
tr|Q4Q1P5|Q4Q1P5_LEIMA 168 ALALRYIYIRKLN-BDAWHLTVVILSDASVPGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
tr|E1F363|E1F363_GIAIA 165 HSFILIAVAKNVDSEAPGWVCPAVTISGWLKCGEGEHKIMDYIRKQACQNHDPN-----THHCLYGLD
consensus 241 ..*.....*.*.*.*.*.....

```

Figura 7. Alineamiento Múltiple de la secuencia de aminoácidos de proteínas XRN2.

Proteínas DCP2

[illegible]

Figura 8. Alineamiento Múltiple de la secuencia de aminoácidos de proteínas DCP2.

Proteína	Organismo	No. de acceso ^a	E-value	I (%)	S (%)
5'-3' exoribonucleasa	<i>Entamoeba dispar</i>	B0E6C7	3.4e-138	44	62
XRN 5'-3' exoribonucleasa	<i>Trichomonas vaginalis</i>	A2DLG3	7.2e-104	52	70
exoribonucleasa 2 putativa	<i>Leishmania major</i>	Q4Q1P5	5.6e-60	27	46
5'-3' exoribonucleasa-2	<i>Giardia intestinalis</i>	E1F363	1.6e-96	44	59
Exoribonucleasa, putativa	<i>Plasmodium falciparum</i>	Q8I358	7.0e-101	49	68
XRN2, 5'-3' exoribonucleasa 2	<i>Mus musculus</i>	Q9DBR1	1.5e-139	46	65
XRN2, 5'-3' exoribonucleasa 2	<i>Homo sapiens</i>	Q9H0D6	4.9e-139	46	64
LOC100049102 proteína	<i>Xenopus laevis</i>	A3KNB2	4.9e-100	47	66
XRN2, 5'-3' exoribonucleasa 2	<i>Drosophila melanogaster</i>	Q9VM71	8.2e-98	33	53
XRN2, 5'-3' exoribonucleasa 2	<i>Schizosaccharomyces pombe</i>	P40848	2.1e-138	47	63
XRN2, 5'-3' exoribonucleasa 2	<i>Arabidopsis thaliana</i>	Q9FQ02	4.5e-134	45	62
Putative 5'-3' exoribonucleasa	<i>Oryza sativa</i>	Q8RV84	3.4e-138	44	62

^aUniProtKB database. I, identidad. S, similitud

Tabla 6. Comparación de *EhXRN2* con proteínas relacionadas de diferentes organismos

Proteína	Organismo	No. de acceso ^a	E-value	I (%)	S (%)
Enzima de <i>decapping</i> de RNAm	<i>Entamoeba dispar</i>	B0EP63	2.3e-122	98	99
Hidrolasa, proteína de la familia NUDIX	<i>Trichomonas vaginalis</i>	A2DDL9	3.3e-14	37	48
Proteína putativa sin caracterizar	<i>Giardia intestinalis</i>	A8B9U6	1.3e-10	30	52
Proteína putativa sin caracterizar	<i>Leishmania major</i>	Q4Q239	1.1-03	26	31
hidrolasa NUDIX, putativa	<i>Plasmodium falciparum</i>	Q8IEM5	4.2e-19	29	49
DCP2,enzima 2 de <i>decapping</i> de RNAm	<i>Mus musculus</i>	Q9CYC6	5.2e-25	28	50
DCP2,enzima 2 de <i>decapping</i> de RNAm	<i>Homo sapiens</i>	Q8IU60	8.5e-25	30	50
Enzima de <i>decapping</i> de RNAm	<i>Xenopus laevis</i>	Q45NB9	9.1e-28	29	54
Decapping protein 2	<i>Drosophila melanogaster</i>	Q9VUU4	8.8e-25	32	52
Enzima de <i>decapping</i> de RNAm, subunidad 2	<i>Arabidopsis thaliana</i>	F4K3Z9	4.4e-20	27	46
Proteína de <i>decapping</i> de RNAm -like	<i>Oryza sativa</i>	Q6K837	8.7e-23	32	51

^aUniProtKB database. I, identidad. S, similitud

Tabla 7. Comparación de *EhDCP2* con proteínas relacionadas de diferentes organismos

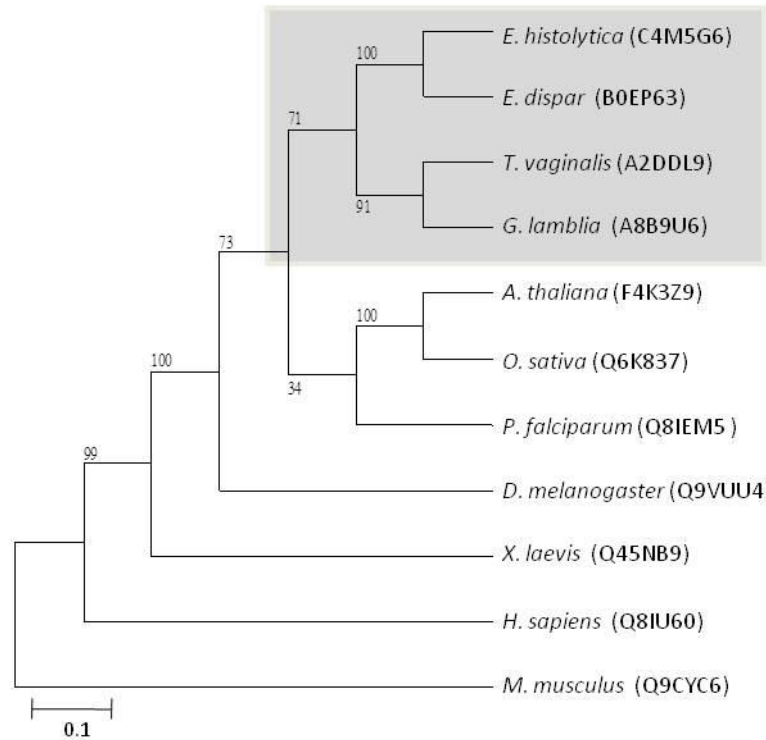


Figura 10. Relaciones filogenéticas entre *EhDCP2* y miembros de la familia DCP2. El árbol enraizado fue generado con el programa MEGA 5.1 por el método *Neighbor Joining*, algoritmo basado en los alineamientos en ClustalW de la secuencia de aminoácidos completa. Los números de acceso de cada proteína están indicados entre paréntesis.

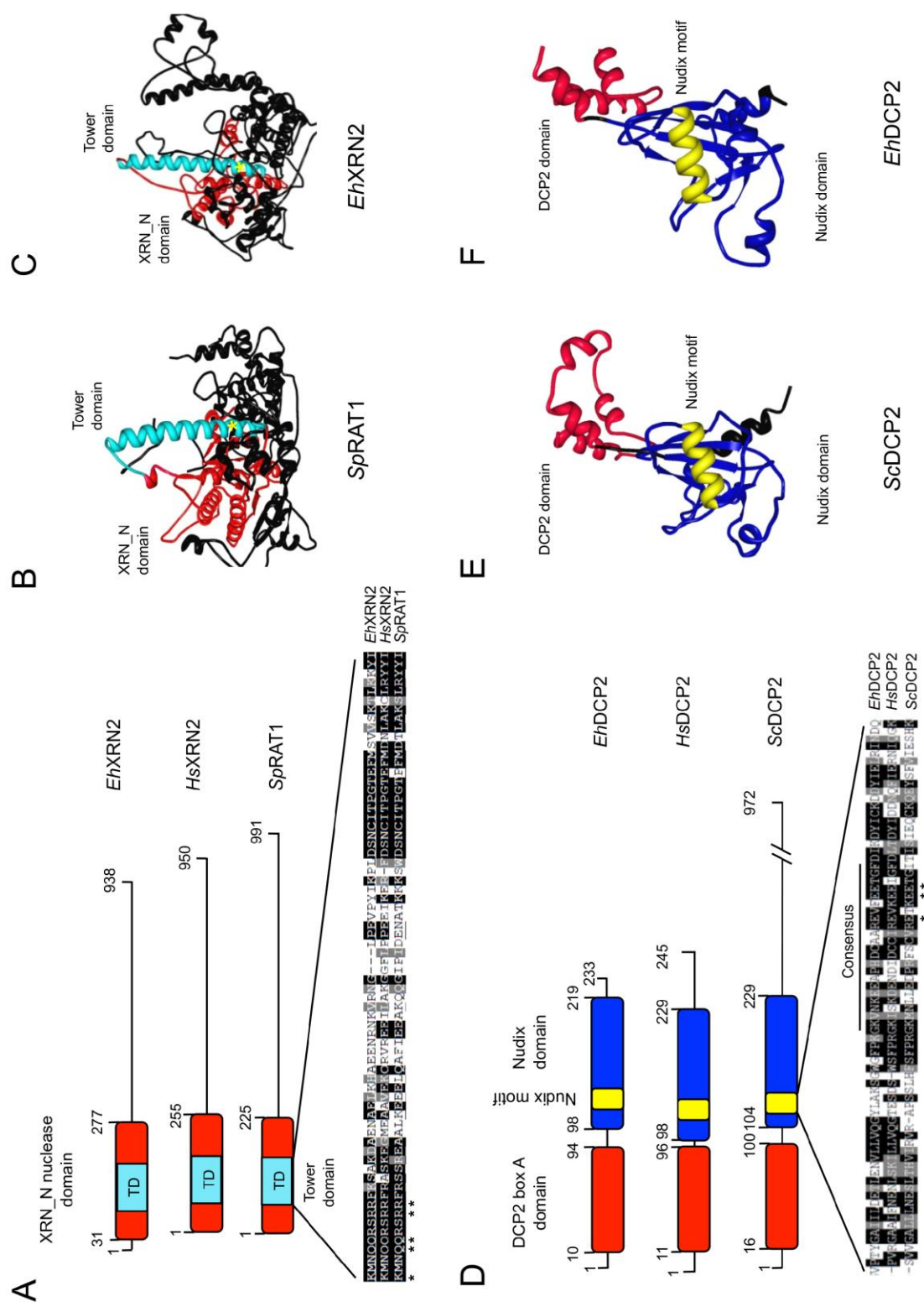


Figura 11. Comparación de las proteínas XRN2 y DCP2 de organismos eucariontes y *E. histolytica*. (A) Representación esquemática de las proteínas XRN2 de *E. histolytica*, *H. sapiens* y *S. pombe*. Caja Roja, dominio XRN_N; caja azul, *tower domain*. Los números son relativos a la metionina inicial en cada uno. Panel inferior, alineamiento múltiple de residuos de aminoácidos del *tower domain* de proteínas homólogas. Cuadro negro, residuos conservados, caja gris, sustitución por aminoácidos conservados. Los asteriscos marcan los residuos conservados en el sitio activo. (B, C) los diagramas de la estructura terciaria para *SpRAT1* (B) y el predicho para la proteína *EhXRN2* (C). Rojo, dominio N-terminal XRN_N; azul, *tower domain* (TD). (D) Representación esquemática de las proteínas DCP2 de *E. histolytica*, *H. sapiens* y *S. cerevisiae*. Caja roja, dominio DCP2; caja azul, dominio NUDIX. Los números son relativos a la metionina inicial en cada proteína. Panel inferior, alineamiento múltiple de residuos de aminoácidos del motivo NUDIX (PF00293) en proteínas homólogas. Los asteriscos marcan los residuos conservados que participan en la actividad catalítica. Cuadro negro, residuo conservado; cuadro gris, sustitución por aminoácidos conservados. (E, F) Diagramas de la estructuras terciaria de *ScDCP2* (E) y el predicho para *EhDCP2* (F). Rojo, dominio DCP2; azul, dominio NUDIX; amarillo, motivo NUDIX. Los modelos 3D se visualizaron y refinaron con el programa RCSB.

4. *EhXRN2*, *EhDCP2* y *EhCAF1* son proteínas citoplasmáticas

Con los anticuerpos específicos obtenidos para las proteínas *EhXRN2* y *EhDCP2* determinamos la localización celular de ambas en trofozoítos de *E. histolytica*. Las proteínas recombinantes se sobreexpresaron en bacterias (Figura 12A y B; carril 2), se purificaron mediante cromatografía de afinidad y se utilizaron como inmunógenos en ratón para obtener anticuerpos específicos. En ensayos de *Western blot*, utilizando los anticuerpos específicos para *EhXRN2* y *EhDCP2* detectamos las correspondientes proteínas recombinantes purificadas (Figura 12A y B, carril 5). La localización subcelular de *EhXRN2* y *EhDCP2* se investigó por ensayos de *Western blot* utilizando extractos proteicos citoplasmáticos y nucleares obtenidos de trofozoítos (Figura 12C). En la fracción citoplasmática, los anticuerpos específicos reconocieron las bandas de 103 kDa y 27.5 kDa, que corresponden al peso molecular predicho para las proteínas endógenas *EhXRN2* y *EhDCP2*, respectivamente (Figura 12C, carriles 2 y 4). En contraste, los anticuerpos no reconocieron ninguna proteína en la fracción nuclear (carriles 1 y 3). Como control, se utilizó suero pre-inmune, que no dio señal, y anticuerpos que reconocen a la proteína *EhPC4* principalmente en la fracción nuclear (carril 5), evidenciando la ausencia de contaminación cruzada entre ambas fracciones proteicas.

Posteriormente se estudió la distribución celular de las proteínas *EhXRN2* y *EhDCP2* en trofozoítos a través de ensayos de inmunofluorescencia y microscopía confocal. Como se muestra en la figura 12, tanto *EhXRN2* y *EhDCP2* se inmunodetectaron difusamente en el citoplasma (D-H y I-M), se observó que las

proteínas principalmente se concentran en *foci* brillantes en el citoplasma con un diámetro que oscila entre 50 a 300 nm y asemejan la organización de los cuerpos de procesamiento o *P-bodies* descritos en organismos eucariontes superiores. Para investigar más a fondo la localización de las proteínas adicionales que participan en la degradación del RNAm en trofozoítos de *E. histolytica*, decidimos estudiar la distribución celular de la desadenilasa EhCAF1 utilizando anticuerpos específicos generados en conejo. En particular, la proteína EhCAF1 se encontró distribuida en el citoplasma en *foci* detectables con una tinción difusa más general (Figura 12N-R), en un patrón similar al de las proteínas EhXRN2 y EhDCP2.

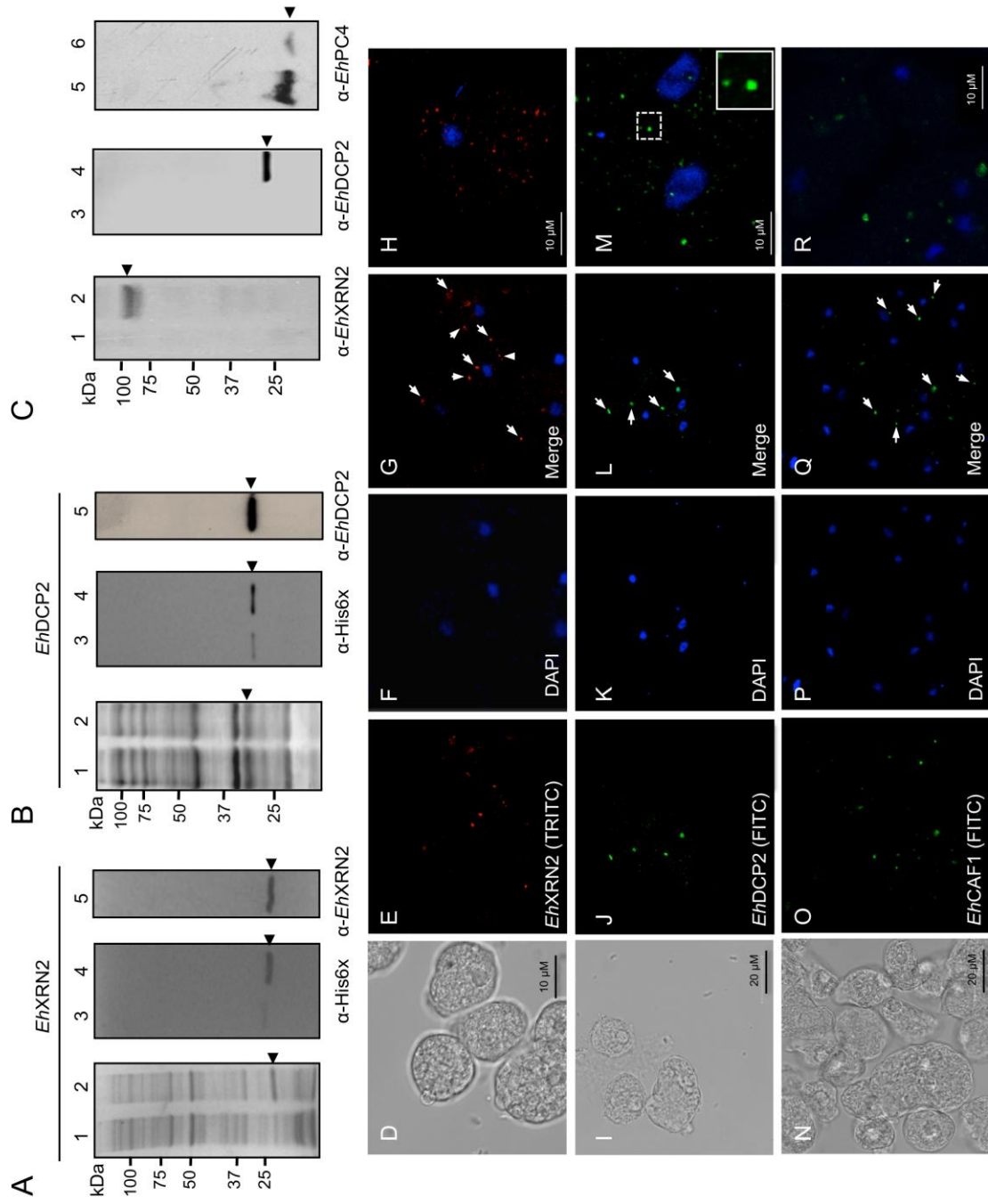


Figura 12. Expresión e inmunolocalización de las proteínas *EhXRN2*, *EhDCP2* y *EhCAF1*. (AB) Expresión de las proteínas *EhXRN2* (A) y *EhDCP2* (B). Las proteínas bacterianas se separaron a través de SDS-PAGE al 12% y los geles se tiñeron con azul de Coomassie. Carril 1, extractos bacterianos no inducidos; carril 2, extractos bacterianos inducidos con IPTG. El *Western blot* se realizó utilizando anticuerpo anti-His6x (Santa Cruz) para detectar la etiqueta de histidinas; extractos bacterianos no inducidos (carril 3) o extractos bacterianos inducidos (carril 4). Las proteínas recombinantes *EhXRN2* y *EhDCP2* también se inmunodetectaron con sus correspondientes anticuerpos específicos (carril 5). (C) Inmunodetección de *EhXRN2* (carriles 1 y 2), *EhDCP2* (carriles 3 y 4) y *EhPC4* (carriles 5 y 6). Los carriles 1, 3 y 5, extractos nucleares; carriles 2, 4 y 6, extractos citoplasmáticos. (D-R). Ensayos de microscopía confocal mediante inmunofluorescencias para la localización de *EhXRN2* (E), *EhDCP2* (J), y *EhCAF1* (O) en trofozoítos fijados con PFA al 4% utilizando anticuerpos secundarios acoplados a TRITC y FITC. Paneles de en medio (F, K, P) muestran la tinción nuclear con DAPI, y los paneles de la derecha (G, L, Q) muestran la sobreposición de las dos señales. Diferentes aumentos de las células individuales se utilizaron para mejorar la visibilidad de las estructuras citoplasmáticas (H, M, R). Algunos *foci* se indican con flechas.

5. *EhXRN2*, *EhDCP2* y *EhAGO2-2* colocalizan en los *foci* citoplasmáticos que contienen a *EhCAF1*

Las observaciones indican que *EhXRN2*, *EhDCP2* y *EhCAF1* se encuentran en *foci* similares, lo que plantea la posibilidad de que puedan ser localizados en estructuras citoplasmáticas comunes. Para probar esta hipótesis, se realizaron los ensayos de colocalización mediante microscopía confocal en trofozoítos, utilizando a *EhCAF1* como un marcador de *foci* citoplasmáticos (Figura 13). Los resultados mostraron que *EhXRN2* y *EhDCP2* coimmunolocalizan en la mayoría de las estructuras citoplasmáticas que contienen a *EhCAF1* (Figura 13, A-D y E-H). Entonces, nos preguntamos si otros factores relacionados con la degradación del RNAm también podrían acumularse en los mismos sitios que la proteína *EhCAF1*. Para examinar esta hipótesis, experimentos de inmunofluorescencia indirecta se realizaron a través de fabricación casera de anticuerpos contra la proteína Argonauta *EhAGO2-2*, proteína involucrada en RNA de interferencia (Zhang *et al.*, 2011). Curiosamente, los resultados demostraron que la señal para la proteína *EhAGO2-2* también colocalizó en los mismos sitios que la señal para *EhCAF1* (I-L). En conjunto, nuestros datos indican que los factores de degradación del RNAm se concentran en la estructuras tipo *P-bodies*. El hecho de que las proteínas *EhXRN2*, *EhDCP2*, *EhCAF1* y *EhAGO2-2* colocalizaron en estructuras citoplasmáticas similares planteó la posibilidad de que podría existir una interacción física. Por lo tanto, realizamos ensayos de coinmunoprecipitación con los anticuerpos para *EhCAF1* y *EhXRN2* como anzuelos y las proteínas citoplasmáticas de los trofozoítos como captura (Figura 13M). Los resultados

evidenciaron que *EhCAF1* coinmunoprecipitó con *EhXRN2* y viceversa, lo cual coincide con los resultados obtenidos en la coinmunolocalización y sugiere un vínculo funcional entre los eventos de desadenilación y la degradación en dirección 5'-3' (Figura 13M, carriles 2 y 4 del panel superior). Por el contrario, no se detectó coimmunoprecipitación de *EhCAF1* con *EhDCP2* o *EhAGO2-2*, lo que sugiere que a pesar de su localización en los mismos *foci* en el citoplasma, estas proteínas no interaccionan físicamente en los trofozoítos. De manera interesante, se observó coinmunoprecipitación de la proteína actina con ambos anzuelos (Figura 13M, carriles 2 y 4 del panel inferior). Estos resultados concuerdan con lo observado en el ensayo de *Pull-down* donde utilizamos a la proteína recombinante de fusión GST-*EhCAF1* como anzuelo y proteínas citoplasmáticas como captura o presa, se identificó a la proteína actina como una proteína que interactúa con *EhCAF1*. Adicionalmente, se ha reportado en mamíferos que los *P-bodies* están anclados al citoesqueleto y asociados a tubulina y a las haces de actina, lo que sugiere que los *P-bodies* requieren una interacción dinámica con el citoesqueleto (Aizer *et al.*, 2008).

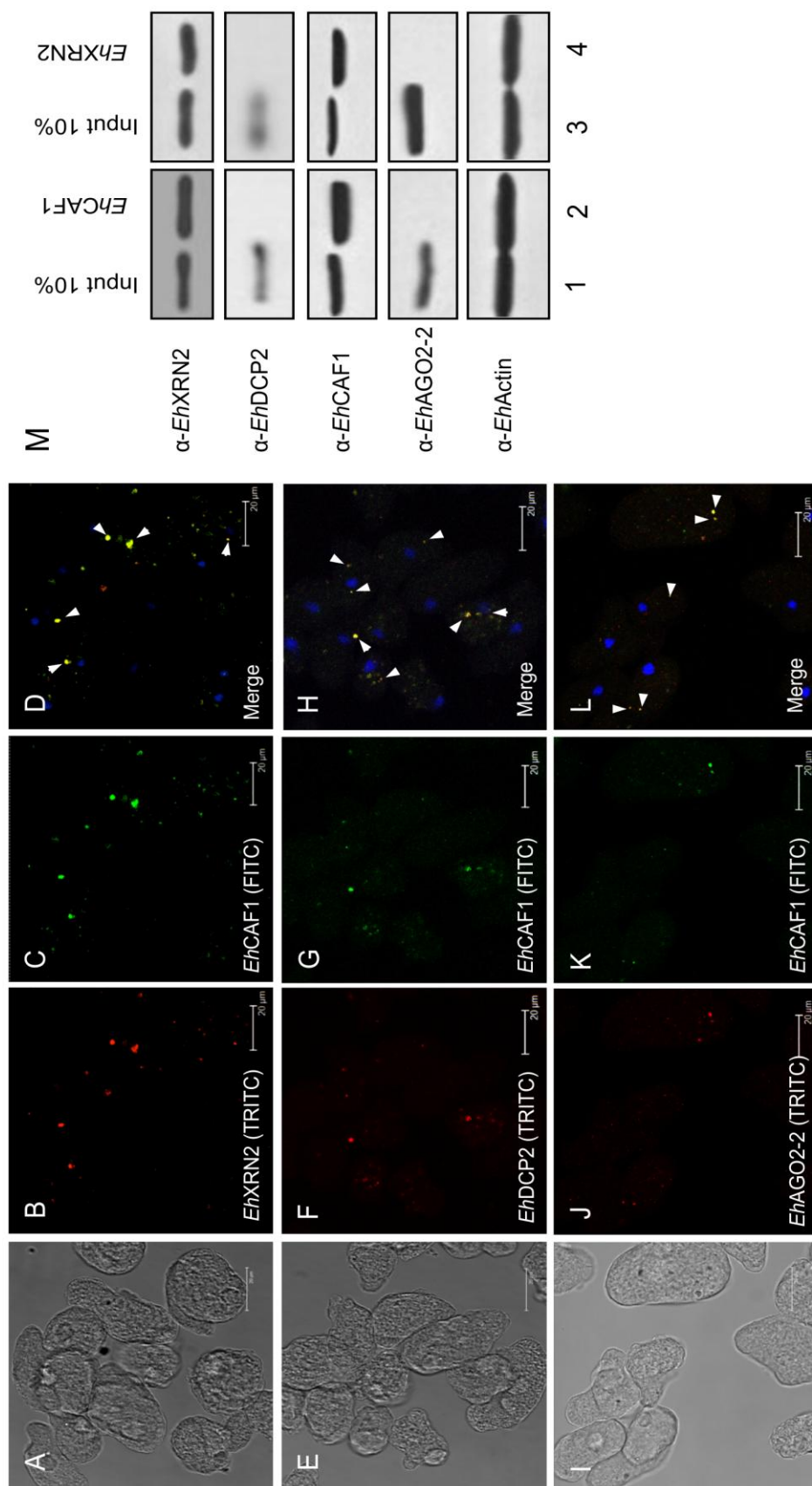


Figura 13. Colocalización de los factores de degradación del RNAm con *EhCAF1* en *foci* citoplasmáticos. (A-L) Ensayos de inmunofluorescencia a través de microscopia confocal. Trofozoítos se incubaron con anticuerpos primarios anti-*EhXrn2* (B), anti-*EhDCP2* (F), anti-*EhAGO2-2* (J), y anti-*EhCAF1*-(C, G, K) (1:100). Los anticuerpos secundarios incluyeron anticuerpos anti-ratón IgG conjugado a TRITC (1:200) y anti-IgG de conejo conjugado a FITC (1:200). (A, E e I), microscopía de luz, (B, F y J), canal rojo; (C, G y K), canal verde; (D, H y L), superposición de las imágenes. Las puntas de flecha señalan las proteínas colocalizadas (M), ensayos de inmunoprecipitación con anticuerpos anti-*EhCAF1* y anti-*EhXRN2* conjugados a perlas de proteína G con lisados citoplasmáticos de trofozoítos de *E. histolytica*. Los inmunoprecipitados se separaron por SDS-PAGE y se analizaron por *Western blot*, utilizando los anticuerpos indicados a la izquierda de cada panel. Los carriles 1 y 3, 10% del lisado total utilizada para inmunoprecipitación; carriles 2 y 4; proteínas inmunoprecipitadas utilizando los anticuerpos indicados del lado izquierdo en cada panel.

6. La formación de los *foci* citoplasmáticos depende de la transcripción activa

Para evaluar la importancia de la transcripción el tamaño y número de los *P-bodies*, se realizaron experimentos y medición del número de *foci* que contienen a *EhCAF1* en respuesta al tratamiento con el inhibidor de la transcripción actinomicina D como se describió previamente (Lopez-Camarillo *et al.*, 2003). Los resultados muestran que el bloqueo de la transcripción altera significativamente el número de *foci* citoplasmáticos (Figure 14A). Se observó que el número de *foci* citoplasmáticos por células disminuyó después de 4 h de incubación con actinomicina D. Después de 12 h de tratamiento, el número promedio de *foci* por célula disminuyó hasta el 50%. Este efecto no se observó en las células tratadas con DMSO solo, utilizado como control. Ensayos de *Western blot* evidenciaron que el nivel de expresión de *EhCAF1* no se vio afectado significativamente por el tratamiento actinomicina D, lo que indica que la reducción del número de *foci* no fue resultado de la inhibición de la proteína *EhCAF1*. Estos datos indican que la formación de estos *foci* citoplasmáticos está vinculada a la transcripción activa, que es una característica conservada en los *P-bodies* de organismos eucariontes superiores.

7. La inhibición de la traducción afecta la formación de los *foci* citoplasmáticos

Una característica común de los *P-bodies* de mamíferos es que pueden ser dispersados por fármacos como la cicloheximida que estabiliza los RNAm en polisomas; mientras que la puomicina libera los RNAm de los ribosomas, se

observa un aumento en el número y tamaño de dichas estructuras (Eulalio *et al.*, 2007), (Cougot *et al.*, 2004). Para evaluar el efecto de la inhibición de la síntesis de proteínas sobre el número y tamaño de los *foci* citoplasmáticos que contienen a *EhCAF1*, trofozoítos de *E. histolytica* se trataron con los inhibidores de la traducción, cicloheximida y puromicina. El resultado mostró que el número de *foci* por células disminuyó rápidamente después de 15 min de incubación con cicloheximida y casi desaparecieron a los 120 min de tratamiento (Figura 14B). En contraste, el tratamiento con puromicina sólo induce un aumento moderado en el número de *foci* citoplasmáticos después de 120 min de tratamiento (Figura 14C). Los ensayos control mostraron que el disolvente DMSO solo, no causó un efecto significativo en la distribución de *foci* que contiene a *EhCAF1*.

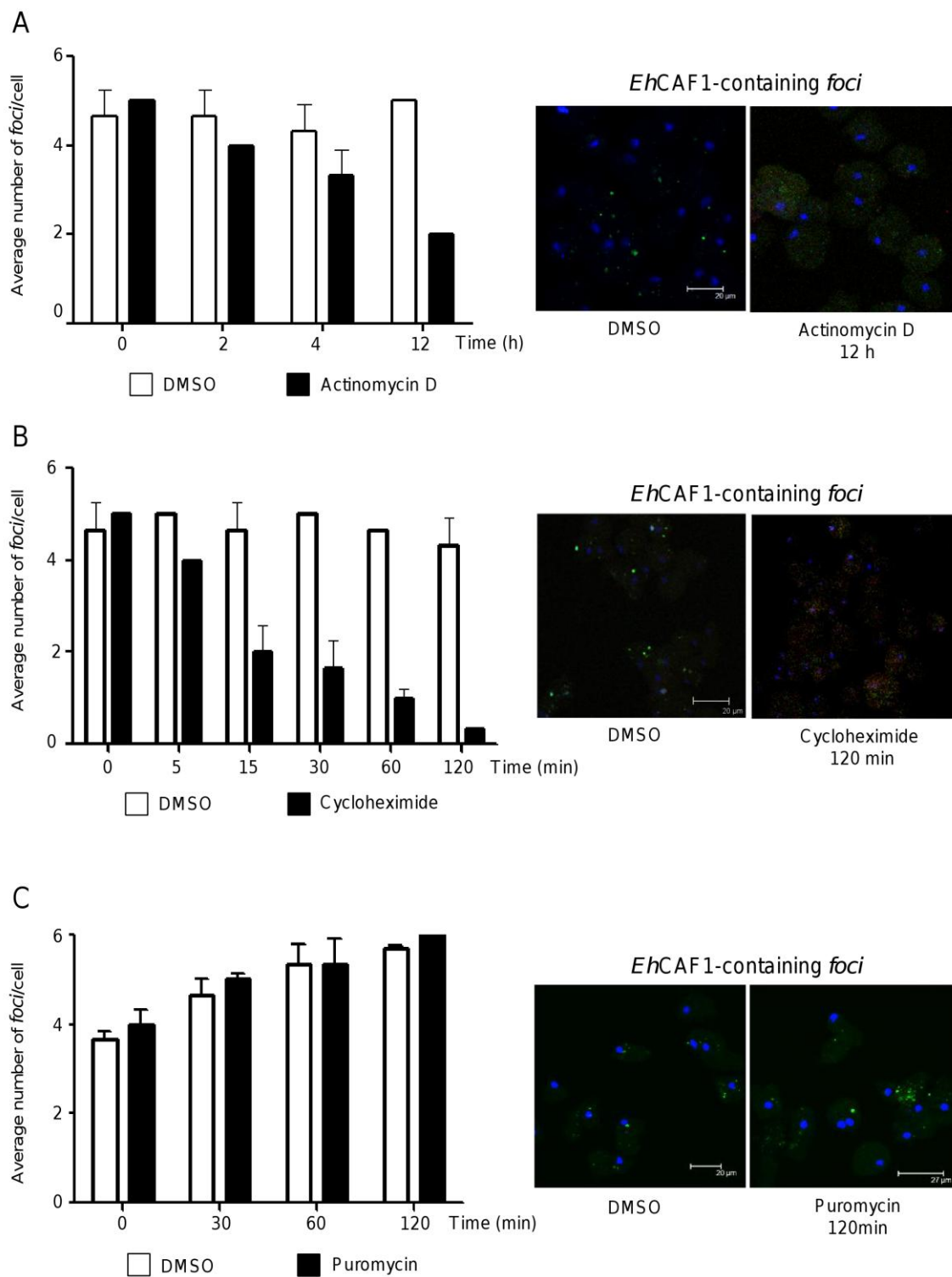


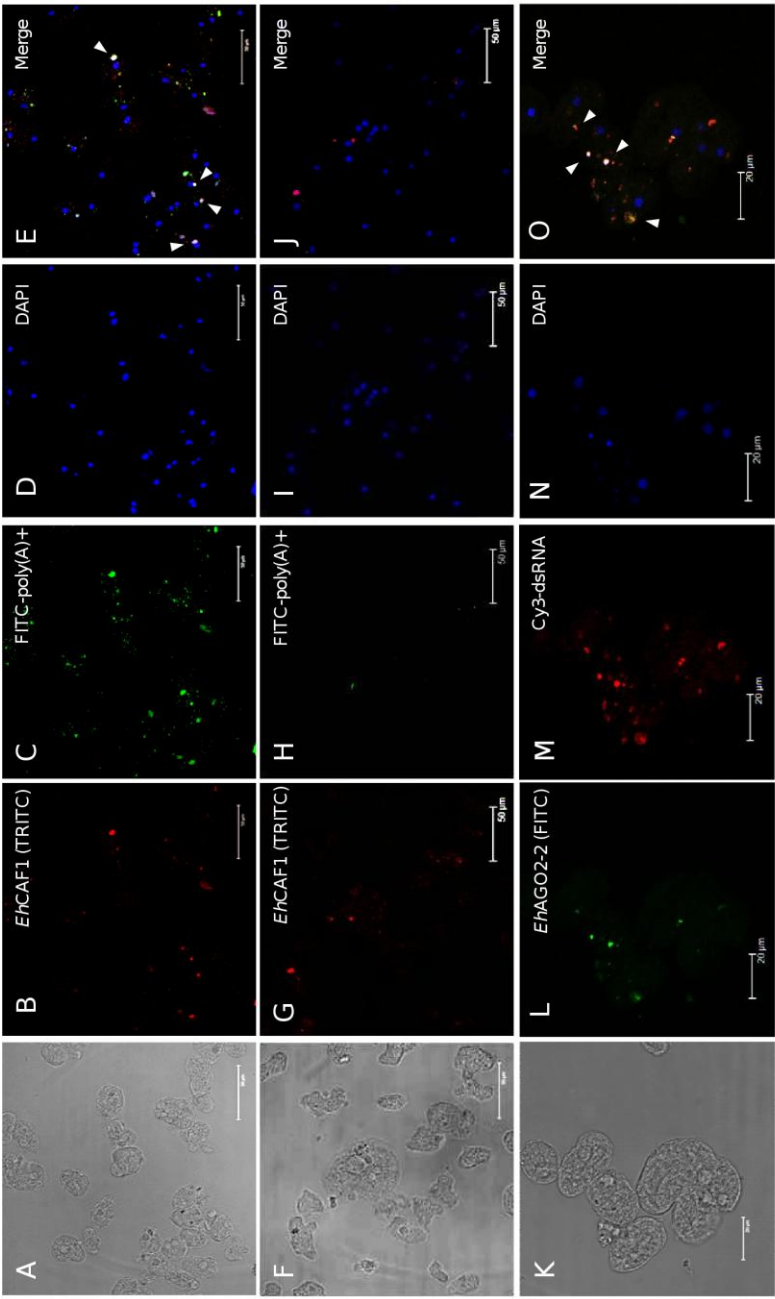
Figure 14. Determinación del número de *foci* citoplasmáticos que contienen a *EhCAF1* después de los tratamientos con actinomicina D (A), cicloheximida (B) y puromicina (C). Después del tratamiento con actinomicina D (5µg/ml), cicloheximida (10µg/ml) o puromicina (200µg/ml), los trofozoítos se fijaron, se incubaron con el anticuerpo anti-*EhCAF1* y anticuerpos secundarios marcados con FITC, los trofozoítos se analizaron a través microscopía de confocal. Izquierda, evolución temporal de la cantidad promedio de *foci* por trofozoítos en respuesta a la inhibición transcripcional (A) o traduccional (B-C). Los datos corresponden a seis campos elegidos al azar. Derecha, imágenes representativas del control (DMSO) y las células tratadas con actinomicina D, cicloheximida o puromicina se observan en el canal verde.

8. *EhCAF1* y *EhAGO2-2* colocalizan con RNA poli(A)⁺ y dsRNA

Para investigar la naturaleza de *foci* citoplasmáticos, se investigó si sustratos de RNA poliadenilados se acumulan en estas estructuras. Mediante hibridación fluorescente *in situ* (RNA-FISH) se detectaron RNAs poli(A)⁺ en trofozoítos, utilizando como sonda un oligo(dT)₃₀ marcado con FITC y anticuerpos específicos *EhCAF1* para evidenciar los *foci* que contienen a *EhCAF1*. Los resultados mostraron que la señal del oligo (dT)₃₀-FITC apareció dispersa en el citoplasma y una proporción menor de la sonda fluorescente colocalizó con *EhCAF1* en condiciones basales de cultivo (Figura 15E). Los experimentos control mostraron la señal para oligo(dT)₃₀ -FITC fue específica para el RNA, ya que desapareció después de tratamiento RNasa A (Figura 15F-J).

Para obtener ideas sobre el papel potencial de los *P-bodies* como estructuras de RNA de interferencia, se investigó si RNAs de doble cadena (dsRNA) se acumulan en *foci* citoplasmáticos y colocalizan con *EhAGO2-2* durante el silenciamiento de un gen endógeno. Se diseñó y validó un RNA largo de doble cadena para el silenciamiento del gen *Ehpc4* utilizando la técnica de *soaking* previamente descrito (Solis *et al.*, 2009). Brevemente, dsRNA se generaron en bacterias, se purificaron, marcaron con Cy3 y se transfectaron en trofozoítos para producir siRNAs fluorescentes. Los resultados mostraron que siete días después de la transfección la señal para los RNAs-Cy3 se encontró en *foci* citoplasmáticos discretos dentro de los trofozoítos (Figura 15M). Por otra parte, ensayos de FISH acoplados a inmunofluorescencia, utilizando la sonda dsRNA-Cy3 y anticuerpos para *EhAGO2-2* mostraron que ambas señales

colocalizaron en grandes *foci* citoplasmáticos (Figura 15 7K-O). Sorprendentemente, se observó un aumento en el tamaño de los *foci* y la intensidad de la señal para *EhAGO2-2* en comparación con las células control no transfectadas, lo que sugiere que la expresión de la proteína *EhAGO2-2* fue inducida por la transfección de dsRNA. Ensayos de *Western blot* confirmaron que a los tres y siete días después de la transfección de RNAds-Cy3, los niveles de la proteína *EhPC4* se redujeron hasta un 50 y 80%, respectivamente, lo que indica que la interferencia de RNA fue exitosa (Figura 15P-Q). Descartamos que estos *foci* citoplasmáticos que contienen siRNAs dirigidos contra el gen *Ehpc4* correspondan a vesículas o endosomas que podría seguir permaneciendo por la transfección ya que utilizamos el método de *soaking* para liberar los dsRNA-Cy3 en los trofozoítos. Sin embargo, no podemos asegurar que sólo la cadena guía del siRNA se encuentra en los *foci* citoplasmáticos, ya que se ha descrito que los siRNAs también pueden localizarse en *P-bodies* en forma de dsRNA (Jagannath *et al.*, 2009).



RNAse A



Figura 15. RNA Poli(A)⁺ y dsRNA colocalizan con *EhCAF1* y *EhAGO2-2* en *foci* citoplasmáticos

(A-E) Ensayos de colocalización de *EhCAF1* y RNAs Poli (A)⁺. Los trofozoítos se inmunotñieron con anticuerpos *EhCAF1* (B), RNAs poli (A)⁺ se detectaron mediante hibridación con oligo(dT)₃₀ conjugado a FITC(C), las células se contratiñeron con DAPI (D) y analizaron mediante microscopía confocal .La combinación de imágenes mostro la sobreposición de las señales. Las flechas blancas indican los sitios de RNAs poli (A)⁺ y la señal de *EhCAF1*. Las flechas amarillas indican la señal del tracto poli (A)⁺ que no se sobrepone con la de los *foci* que contiene a *EhCAF1*. (F-J) Control, células tratadas con RNAsa A. Inmunodetección de *EhCAF1* (G), detección of poli(A)⁺ (H), DNA se contratiño con DAPI (I). Combinación de imágenes (J). (K-O) Ensayos de colocalización de *EhAGO2-2* y dsRNA-Cy3. Trofozoítos se transfectaron con dsRNA-Cy3 específicos para *Ehpc4* (K-O) e inmunotñieron con anticuerpos anti-*EhAGO2-2* a los siete días post-transfección (L). Las células se contratiñeron con DAPI (N) y analizaron a través de microscopia confocal. Combinación de imágenes (O) muestra ambas señales. Las flechas señalan las señales de colocalización en *foci* citoplasmáticos. (P) Western blot para *EhPC4* y *EhActina* de proteínas obtenidas a los siete días post-transfección de los dsRNAs-Cy3. (Q) Análisis densitométrico de las bandas inmunodetectadas en P.

9. *EhCAF1* forma complejos con proteínas de degradación de RNA, citoesqueleto y de metabolismo celular

Dado que en el ensayo de inmunoprecipitación con el anticuerpo anti-*EhCAF1* encontramos que precipita la proteína actina como previamente se ha reportado nos dimos a la tarea de corroborar el resultado de la posible interacción física entre la ribonucleasas *EhCAF1*, actina y otras proteínas citoplasmáticas de *E. histolytica*.

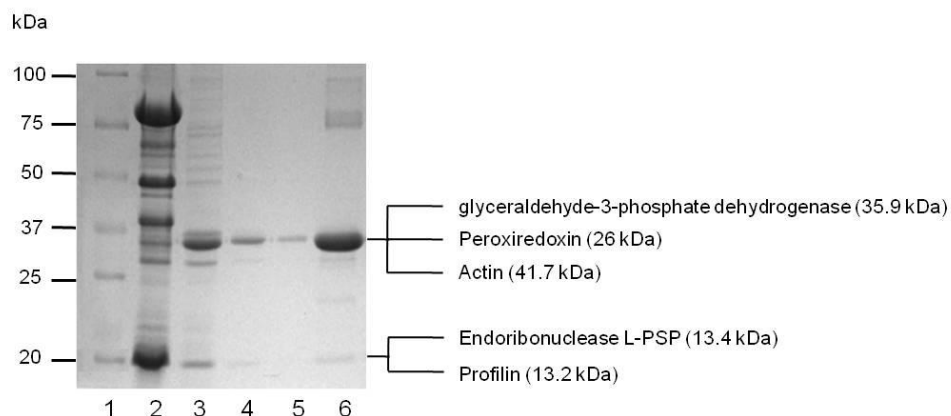
Mediante ensayos de *Pull-down* con la proteína de fusión GST-*EhCAF1* utilizada como anzuelo y proteínas citoplasmáticas como captura, se identificaron diferentes complejos proteicos (Figura 16A). Las bandas correspondientes a los complejos identificados como fueron escindidos del gel de manera manual utilizando puntas estériles de pipetas de 20 a 250 μ l. Las proteínas se identificaron por espectrometría de masas ESI/MS-MS. El análisis se realizó mediante la digestión de los péptidos utilizando tripsina para su posterior conversión a iones en fase gaseosa mediante una ionización-desorción laser y separación según su relación masa-carga (m/z). El análisis de masas de los péptidos se realizó utilizando la base de datos MASCOT. Los péptidos se identificaron comparando con las secuencias de amiba en las bases de datos del NCBI y *AmoebaDB*. Finalmente, se realizó un análisis bioinformático estándar de las secuencias predichas de aminoácidos utilizando los programas BLAST, Clustal W, Pfam y *Prosite*, entre otros. De las proteínas que interactúan con *EhCAF1* interesantemente se identificaron la endoribonucleasa L-PSP, entre otras como, actina, profilina, proteína heat shock 70 y la enzima málica (Tabla 8). No es de

sorprendernos que hayamos identificado actina como una proteína de interacción con *EhCAF1* debido a que ha sido reportado que los *P-bodies* son estructuras citoplasmáticas que se anclan a citoesqueleto y se asocian a tubulina y actina humana (Aizer *et al.*, 2008).

Protein	Molecular mass/pl	Gene symbol ^a	Accession number ^a	Mascot score	No. of matched peptides	Sequence coverage (%)	MS/MS peptide sequence (ion scores)	Function
glyceraldehyde-3-phosphate dehydrogenase	35991/7.04	EH1_008200	C4LVR9	1003	20	66	²⁶ KDFELVAINDPFMDPK ⁴⁰ (75) ⁵⁴ QFEGTVEAGENAIIVNGHK ⁷² (79) ⁸⁰ DPAQIGWGALGVYVVESTGVFTTIK ¹⁰⁶ (75)	Glucose metabolic process
enolase	47297/5.87	EH1_130700	C4LXE8	405	9	24	⁶⁵ AVENVNTIIGPALLGK ⁸⁰ (84) ²²⁰ EALDLLVEAIA ²³¹ (86) ³¹⁰ FTVEHGNFQIVGDDLLVTNPAR ³³¹ (83)	glycolysis
Elongation factor 1-alpha	47234/8.77			107	14	26	⁸⁵ YYFTIIDAPGHR ⁹⁶ (73) ²¹¹ GPTLIGALDSVTPPERPVDKPLR ²³³ (54) ²⁵³ VETGILKPGTIVQFAPSGVSSECK ²⁷⁶ (32)	
peroxiredoxin	26237/7.36	EH1_145840	C4M3Q0	182	17	33	¹³⁵ LTFFPLVSDIKR ¹⁴⁵ (62) ¹⁵¹ YGMNLVEAGIAR ¹⁶² (92) ¹⁷⁵ YQMNDGIGR ¹⁸⁶ (62)	antioxidant and peroxiredoxin activity
actin	41700/5.26			123	13	26	¹⁰⁵ GYAFTTTAER ²⁰⁴ (61) ³¹⁴ EMIQLAPPTMK ³²⁵ (67) ³⁵⁸ EEYDESGPAIVHR ³⁷⁰ (69)	
malic enzyme	53286/5.92	EH1_044970	Q9NH04	1137	46	50	⁶⁰ GNFVGVVSDSTR ¹⁰¹ (72) ²⁴⁰ EAGAYIVATGR ²⁵⁰ (69) ⁴³³ VTDLTWQQVYDIAEHDK ⁴⁵⁰ (75)	malate metabolic process
pyruvate:ferredoxin oxidoreductase	127554/6.35	EH1_051060	C4LTX6	877	24	19	⁵²⁰ ISVLDR ⁵²⁵ (56) ⁷¹¹ APAEFVQIDGK ⁷²¹ (64) ⁹¹⁸ YAAVKPLLAEEK ⁹³⁰ (65)	iron ion binding
pyruvate phosphate dikinase	97836/5.89	EH1_009530	Q24801	439	15	17	⁷⁵ VFGGEENPLLVSR ⁸¹ (57) ³⁵¹ TGVDMEVEGLTK ³⁶³ (61) ³⁹⁵ GLPASPGAATGAVVFDADDAVEQAK ⁴¹⁹ (43)	Kinase and ligase activity
acetyl-CoA synthetase	77518/6.14			425	13	14	²²⁷ EFAADKPIILLK ²³⁸ (58) ⁴²⁵ IPNFETPERIPNFETPER ⁴³⁴ (57) ⁴⁷¹ LIADVADGR ⁴⁸⁰ (65)	
heat shock protein 70	71482/5.32			354	12	18	²⁴⁰ LVNHFAIEFKR ²⁵⁰ (55) ³⁴⁸ VVQLQDFFNGKEPNK ³⁶⁴ (62) ³⁹⁵ SINPDEAVAYGAAVQAAILTGTGGK ³⁸⁶ (67)	
elongation factor 1 alpha	19690/7.19			142	5	22	¹⁸ STTTGHLIYK ²⁷ (25) ⁶² YYFTIIDAPGHR ⁹³ (59) ¹⁴⁴ YLVVGVVAK ¹⁵¹ (57)	
profilin	3228/7.63	EH1_176140	C4MB21	360	26	56	⁷⁴ VDADEGTAMGK ⁸⁴ (62) ⁸⁵ KGAEGISYK ⁹⁴ (75) ⁹⁷ QAVIIGYFSDASVSAGQNSDATYK ¹²⁰ (30)	actin cytoskeleton organization
endonuclease L-PSP	13374/5.78	EH1_087570	C4LXT9	211	10	38	⁶² YVLEEAGSSMDK ⁷³ (74) ⁸⁶ DFGVFNIGYAEAFGNHKPAR ¹⁰⁵ (51) ¹⁰⁸ ACFAAAALPK ¹¹⁵ (66)	
putative fructose-1,6-bisphosphate aldolase	35806/5.96			200	7	23	²⁴⁵ LESAIGIPEDQIR ²⁵⁷ (56) ²⁷⁵ LAMTGSIRR ²⁸³ (57) ²⁹⁶ QYLGPAR ³⁰² (48)	
eukaryotic translation elongation factor	22699/6.23			126	4	17	¹¹⁸ IISEEQRPGLFNVNR ¹³³ (54) ¹⁷⁹ VGSIVAIR ¹⁸⁷ (59) ¹⁹¹ GLPEGVPGLDK ²⁰¹ (22)	
enhancer binding protein-2	26300/9.82	EH1_182670	C4LV48	120	4	20	¹²⁹ EVPTFEFSDNLIFVK ¹⁴² (30) ¹⁴³ NLAYAIKDEDLK ¹⁵⁴ (50) ¹⁸⁰ GYGFVEVK ¹⁸⁷ (26)	

Tabla 8. Proteínas identificadas que interactúan con la proteína *EhCAF1* de *E. histolytica*

A



B

EHI_087570 endoribonuclease L-PSP
 MSKLTVVASP LAPEAVGAYS QAIICNGMVY CSGQIGLDRK TGDFAGKTIE
 EQSKQVMTNL KYVLEEAGSS MDKVVKTCL LADIKDFGVF NGIYAEAFGN
 HKPARACFAA AALPKGALVE VECIATL

C

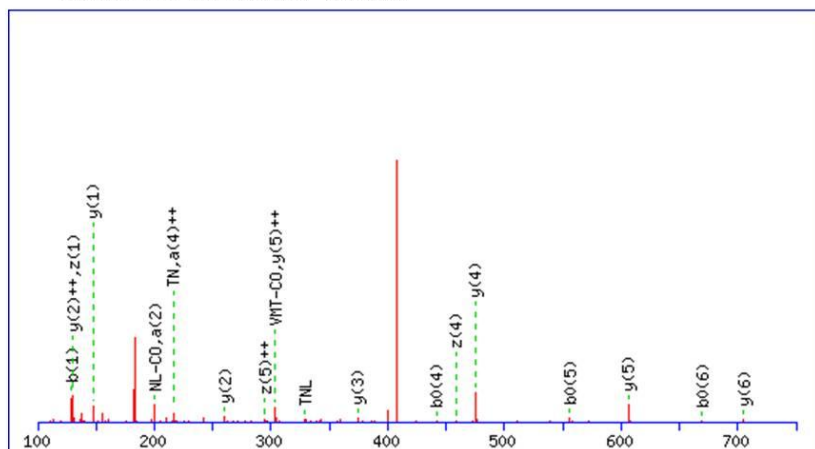


Figura 16. *EhCAF1* forma complejos con proteínas de degradación de RNA, citoesqueleto y de metabolismos celular. (A) Ensayo de *Pull-down* donde se señala el nombre de las proteínas identificadas en los complejos observados mediante PAGE. (B) Secuencia de la proteína endoribonucleasa L-PSP identificada como resultado de la interacción con la proteína *EhCAF1*. (C) Espectro de masas correspondiente a la proteína endoribonucleasa L-PSP.

10. *EhCAF1* tiene actividad de desadenilasa

La unión a los sustratos de RNA es el primer paso requerido para la actividad de desadenilación. Sin embargo la proteína *EhCAF1* aparentemente no tiene dominios de unión a RNA, por lo tanto con el fin de definir y continuar con la caracterización de la *EhCAF1* que se inicio durante el trabajo de maestría, nos dimos a la tarea de realizar un ensayo funcional. Los resultados anteriores evidenciaron que la proteína recombinante *EhCAF1* purificada es capaz de unirse a una sonda de RNA de manera específica. El complejo proteína-RNA se mantuvo sin cambios en presencia de un competidor inespecífico y desapareció cuando se utilizo un exceso molar de 350 veces de un competidor específico (López-Rosas, 2007).

Para investigar la participación de la proteína *EhCAF1* en la degradación de los tractos de adeninas del extremo 3' de los RNAm, realizamos un ensayo de actividad de desadenilasa usando como sustrato un RNA poliadenilado (A_{50}) generado por transcripción *in vitro*. Los resultados muestran una desadenilación progresiva en el sustrato de RNA, lo que resulta en un producto completamente desadenilado después de 60 min, el transcrito no adenilado usado como control no se afecto durante la reacción. (Figura 17). Estos resultados muestran que la proteína *EhCAF1* es una desadenilasa funcional.

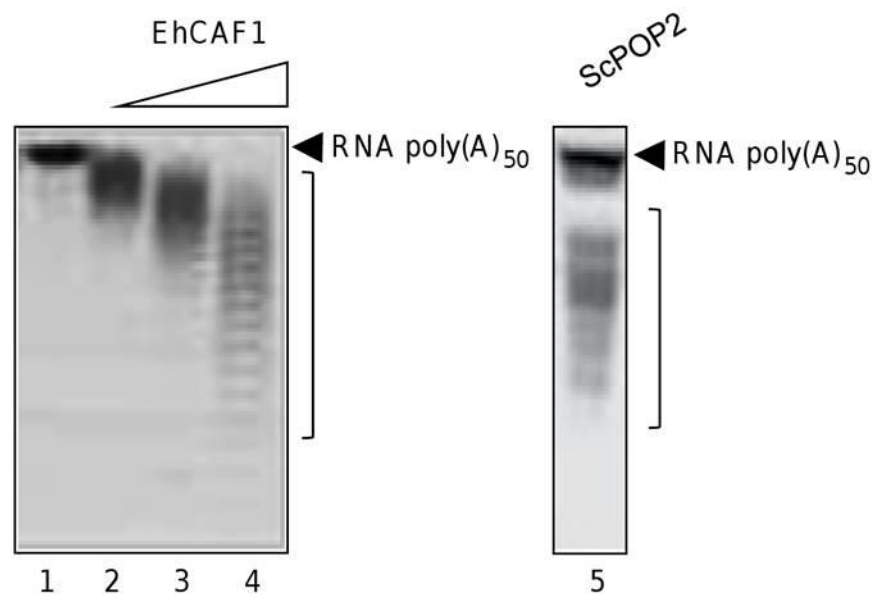


Figura 17. *EhCAF1* tiene actividad de desadenilasa. Ensayo de desadenilación *in vitro*. Carriles 1-5, se muestran los sustratos de RNA a los 0, 15, 30 y 60 min. Las muestras se analizaron en un gel desnaturante de poliacrilamida al 18% con urea 7M. Carril 5, control positivo ensayo de desadenilación utilizando la enzima POP2 de *S. cerevisiae*.

V. DISCUSIÓN

En el presente trabajo reportamos la identificación de proteínas que participan en la degradación del RNAm. Mediante análisis *in silico* identificamos genes codificantes para proteínas representativas de los mecanismos de degradación de RNAm, tales como de *decapping*, desadenilación, NMD y RNA de interferencia en *E. histolytica*, lo cual sugiere que los componentes de la degradación del RNAm y los mecanismos se conservan en este parásito. Sin embargo, la composición de los mecanismos es variable e incompleta en comparación con células eucariontes superiores, una característica común en los genomas reducidos de parásitos divergentes. En particular, encontramos que la maquinaria de desadenilación pareciera estar incompleta en ameba, ya que no encontramos CCR4, PAN2, PAN3, ni la desadenilasa PARN, lo que sugiere que *EhCAF1* representa la principal poli(A) ribonucleasa en este parásito. Además, la maquinaria de *decapping* aparentemente carece de las proteínas DCP1 y XRN1, pero si identificamos homólogos a DCP2 y XRN2, que se conocen como componentes de los *P-bodies*. Curiosamente, *EhXRN2* y *EhDCP2* proteínas implicadas en la degradación del RNAm se acumulan por lo menos en cierta cantidad de *foci* citoplasmáticos discretos en un patrón similar que asemeja la organización de los *P-bodies* de organismos eucariontes superiores. Este patrón de localización se observó también para la desadenilasa *EhCAF1* y la proteína *EhAGO2-2*. Dado que las proteínas DCP2, XRN2, CAF1 y Ago2 son componentes de los *P-bodies* células de humano y levadura, nuestros datos sugieren que los *foci* citoplasmáticos identificados podrían representar los *P-bodies* en *E. histolytica*. Sin embargo, los

foci detectados en amiba parecen ser diferentes en forma y tamaño en comparación con los descritos en células eucariontes superiores, siendo estos de forma más pequeña y menos redondeada, por lo que dichas estructuras que hemos detectado las hemos denominado estructuras tipo *P-bodies*.

Dada la similitud entre las proteínas que intervienen en la degradación del RNAm en levadura y humano, y *E. histolytica*, y la colocalización de *EhCAF1* con *EhXRN2* y *EhAGO2-2* en estructuras citoplasmáticas, nuestros resultados también sugieren que éstas pueden representar sitios de la degradación del RNAm. Sin embargo, esto debe ser confirmado a través de experimentos adicionales diseñados para demostrar la degradación de sustratos de RNA y la localización de las proteínas adicionales de decaimiento del RNAm en estos cuerpos citoplasmáticos.

Los *foci* citoplasmáticos que se asemejan a los *P-bodies*, previamente fueron descritos en el parásito protozoario *Trypanosoma* (Cassola *et al.*, 2007). En *T. cruzi* y *T. brucei* se reportó que los RNAm y proteínas asociadas al metabolismo del RNA se concentran en gránulos citoplasmáticos de RNA en respuesta a condiciones de estrés, principalmente el nutricional como una estrategia de protección para la transcripción transitoria durante periodos Sin embargo, parece que dichas estructuras no representan a los *P-bodies* típicos. Un complejo conformado por CAF1, NOT1, NOT2 y NOT5, DHH1 y un posible homólogo a Caf130 se encontraron en *T. brucei* y la actividad de desadenilación de CAF1 mostró ser esencial para la supervivencia celular (Schwede *et al.*, 2008).

Notablemente, la proteína CAF1 parece estar predominantemente distribuida en el citoplasma, aunque un claro patrón de tinción punteada indicativo de su asociación con gránulos de almacenamiento de RNA o degradación no es evidente. Por otra parte, la helicasa TcDhh1 se encontró presente en complejos independientes de polisomas y en *foci* citoplasmáticos que asemejan *P-bodies*, que varían en número en respuesta a condiciones de estrés nutricional y a tratamientos con cicloheximida y puromicina (Holetz *et al.*, 2007). Sin embargo, a pesar de los esfuerzos iniciales para estudiar los factores de degradación del RNAm en parásitos, a la fecha la existencia y el papel potencial de los *P-bodies* en los eventos de degradación de RNAm no se han abordado plenamente en parásitos protozoarios.

No excluimos la existencia de otros *foci* que contengan RNA en *E. histolytica*, tales como los gránulos de estrés descritos en células eucariontes. Los gránulos de estrés representan agregados citoplasmáticos compuestos de complejos 48S de pre-inicio estancados, que incluyen subunidades pequeñas ribosomales, los factores de inicio de la traducción (eIF2, eIF3, eIF4, eIF4G), PABP, p54/Rck y TTP; y CPEB, proteínas de unión a RNA que regulan la traducción del RNAm y la degradación (Kedersha *et al.*, 2002; Wilczynska *et al.*, 2005). Los *P-bodies* y los gránulos de estrés comparten algunas proteínas como son TTP, p54/Rck, Lsm1, y Xrn1, pero difieren uno del otro porque los *P-bodies* contienen componentes de la maquinaria de degradación del RNAm, mientras que los gránulos de estrés contienen componentes de la maquinaria de inicio de la traducción y marcadores específicos que incluyen a la proteína TIA1 (Gilks *et al.*,

2004; Kedersha *et al.*, 2005). La existencia de gránulos de estrés en *E. histolytica* permanece sin ser abordada con marcadores específicos, con el fin de obtener una imagen completa de las estructuras citoplasmáticas que proteínas de unión a RNA en este parásito protozoario.

Tratamientos con actinomicina D en células de mamíferos induce la pérdida de los *P-bodies* (Cougot *et al.*, 2004), sugiriendo que los transcritos son degradados en los *P-bodies* hasta que no hay más RNAm para sostener la estructura. En tripanosomas, no hay pérdida de los gránulos de RNAm después de ser tratados con actinomicina D, lo que sugiere que a diferencia de los *P-bodies*, estas estructuras no son los sitios donde se lleva a cabo la degradación (Cassola *et al.*, 2007). La formación de los *P-bodies* en humano y levadura depende del incremento de los RNAm no traducibles, lo que indica que los RNAm deben salir del ciclo de traducción para entrar en los (Brenques *et al.*, 2005; Teixeira *et al.*, 2005). El uso de inhibidores de la transcripción y de la traducción que se sabe inhiben la degradación del RNAm sugiere que los *foci* citoplasmáticos representan sitios de decaimiento del RNAm. Nuestros resultados mostraron un comportamiento similar en trofozoítos de *E. histolytica*, ya que se observó un número reducido de *P-bodies* en células tratadas con actinomicina D y cicloheximida. Se ha reportado que el inhibidor de la traducción puromicina induce un incremento en el tamaño y número de los *P-bodies* por la liberación de RNAm de los ribosomas. En trofozoítos de *E. histolytica*, observamos un ligero incremento de las estructuras tipo *P-bodies* después del tratamiento con puromicina. Nuestros datos indican que las estructuras tipo *P-bodies* no son

estáticas sino que son dinámicas y están vinculadas a la síntesis de RNA en curso y a la traducción en *E. histolytica*.

La coinmunolocalización de la desadenilasa *EhCAF1* y la exoribonucleasa *EhXRN2* en los mismos *foci* citoplasmáticos apoya la idea de que estas estructuras tienen un papel importante en la degradación del RNAm. Además, estos resultados proporcionan por primera vez, una relación funcional entre la desadenilación y la degradación en dirección 5'-3' del RNAm, que representa datos nuevos con respecto a las redes de interacción de los factores de desintegración de RNAm.

Una fracción pequeña de RNAs poliadenilados se concentraron en *foci* citoplasmáticos con la proteína *EhCAF1*. Adicionalmente, observamos que siRNAs fluorescentes localizaron con la proteína *EhAGO2-2* en *foci* citoplasmáticos durante el silenciamiento del gen *Ehpc4* por RNA de interferencia, lo que implica que las estructuras tipo *P-bodies* podrían tener un papel importante en el silenciamiento de genes mediado por siRNA, como centros que faciliten la interacción entre los siRNAs y las proteínas Argonauta. Recientemente se reportó que la proteína *EhAGO2-2* localiza en el núcleo con siRNAs que tienen un 5' polifosfato terminal durante el silenciamiento transcripcional en la cepa G3 de *E. histolytica* G3 (Zhang *et al.*, 2011). Sorprendentemente, en este trabajo no detectamos alguna señal nuclear para la proteína *EhAGO2-2*. Esto podría ser debido al hecho de las proteínas Argonauta pueden transportar clases específicas de RNAs pequeños reguladores en distintos compartimentos celulares para una

regulación negativa de la expresión de genes, dependiendo de los siRNAs generados durante la interferencia del RNA. Durante el silenciamiento transcripcional, *EhAGO2-2* transporta pequeños RNAs reguladores específicamente hacia el núcleo para mediar el silenciamiento de genes vía modificaciones de histonas (Zhang *et al.*, 2011). Nuestros experimentos indican que durante la interferencia del RNAm tanto la proteína *EhAGO2-2* y siRNAs específicos se encuentran preferentemente en el citoplasma.

La interacción de la proteína actina con *EhCAF1* y *EhXRN2* también sugiere que las estructuras tipo *P-bodies* en amibas podría estar asociadas a citoesqueleto como se ha descrito en humanos (Aizer *et al.*, 2008). Experimentos adicionales actualmente en curso ayudará a entender las implicaciones de estas interacciones en los mecanismos de degradación de RNAm.

En conjunto, nuestros resultados proporciona evidencia de que las proteínas de degradación del RNAm involucradas en las rutas de desadenilación y *decapping* están organizadas en estructuras citoplasmáticas microscópicas llamadas en este trabajo estructuras tipo *P-body* (*P-body like structures*).

En un trabajo futuro será necesario comprender la composición de estas estructuras y definir la localización de otras proteínas de degradación de RNAm de las maquinarias de NMD y RNA de interferencia, lo que contribuirá a una mejor comprensión de la degradación del RNA en este protista. Esto también nos dará la oportunidad de responder las preguntas básicas intrigantes acerca de la

regulación post transcripcional de la expresión de genes en este parásito protozoario.

VI. CONCLUSIONES

Los resultados del presente trabajo permitieron la identificación del repertorio de proteínas involucradas en la degradación del RNAm. identificamos genes codificantes para proteínas representativas de los mecanismos de degradación de RNAm, tales como de *decapping*, desadenilación, NMD y RNA de interferencia en *E. histolytica*, lo cual sugiere que los componentes de la degradación del RNAm y los mecanismos se conservan en este parásito.

Las proteínas endógenas *EhXRN2*, *EhDCP2* y *EhCAF1* tienen una distribución citoplasmática en agregados discretos, parecidos a las estructuras P-bodies de organismos eucariontes como humano y levadura. Estas proteínas colocalizan en las mismas estructuras citoplasmáticas. También observamos en los trofozoítos una disminución en el número y tamaño de los *foci* citoplásmicos tipo P-bodies en respuesta a los tratamientos con inhibidores de la transcripción y traducción, actinomicina D y cicloheximida. indicando que la presencia de los transcritos y las proteínas en estas estructuras se requiere para el mantenimiento de los mismos.

Adicionalmente observamos que las proteínas *EhCAF1* y *EhAGO2-2* colocalizan con RNAm y dsRNAs en las mismas estructuras citoplasmáticas, respectivamente, lo que implica que las estructuras tipo P-bodies podrían tener un papel importante en el silenciamiento de genes mediado por siRNA

En conjunto, nuestros resultados proporciona evidencia de que las proteínas de degradación del RNAm involucradas en las rutas de desadenilación y

decapping están organizadas en estructuras citoplasmáticas microscópicas llamadas en este trabajo estructuras tipo *P-bodies*.

En un trabajo futuro será necesario comprender la composición de estas estructuras y definir la localización de otras proteínas de degradación de RNAM de las maquinarias de NMD y RNA de interferencia, lo que contribuirá a una mejor comprensión de la degradación del RNA en este protista. Esto también nos dará la oportunidad de responder las preguntas básicas intrigantes acerca de la regulación post transcripcional de la expresión de genes en este parásito protozoario.

VII. REFERENCIAS

- Abed, M. and S. Ankri** (2005). "Molecular characterization of *Entamoeba histolytica* RNase III and AGO2, two RNA interference hallmark proteins." *Exp Parasitol* **110**(3): 265-269.
- Aizer, A. and Y. Shav-Tal** (2008). "Intracellular trafficking and dynamics of P bodies." *Prion* **2**(4): 131-134.
- Arroyo, R. and E. Orozco** (1987). "Localization and identification of an *Entamoeba histolytica* adhesin." *Mol Biochem Parasitol* **23**(2): 151-158.
- Azuara-Liceaga, E., E. Flores-Soto, et al.** (2005). "*Entamoeba histolytica*: structural and functional analysis of the Ehadh112 gene promoter." *Exp Parasitol* **110**(3): 280-285.
- Bevilacqua, A., M. C. Ceriani, et al.** (2003). "Post-transcriptional regulation of gene expression by degradation of messenger RNAs." *J Cell Physiol* **195**(3): 356-372.
- Bhattacharya, A., S. Satish, et al.** (2000). "The genome of *Entamoeba histolytica*." *Int J Parasitol* **30**(4): 401-410.
- Bhattacharya, S., A. Bhattacharya, et al.** (1989). "Circular DNA of *Entamoeba histolytica* encodes ribosomal RNA." *J Protozool* **36**(5): 455-458.
- Binder, M., S. Ortner, et al.** (1995). "Sequence and organization of an unusual histone H4 gene in the human parasite *Entamoeba histolytica*." *Mol Biochem Parasitol* **71**(2): 243-247.

- Bracha, R. and D. Mirelman** (1984). "Virulence of *Entamoeba histolytica* trophozoites. Effects of bacteria, microaerobic conditions, and metronidazole." *J Exp Med* **160**(2): 353-368.
- Bradford, M. M.** (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." *Anal Biochem* **72**: 248-254.
- Brengues, M., D. Teixeira, et al.** (2005). "Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies." *Science* **310**(5747): 486-489.
- Bruchhaus, I., M. Leippe, et al.** (1993). "Unusual gene organization in the protozoan parasite *Entamoeba histolytica*." *DNA Cell Biol* **12**(10): 925-933.
- Bruchhaus, I., T. Roeder, et al.** (2002). "Differential gene expression in *Entamoeba histolytica* isolated from amoebic liver abscess." *Mol Microbiol* **44**(4): 1063-1072.
- Buss, H., C. Lioutas, et al.** (1995). "Analysis of the 170-kDa lectin gene promoter of *Entamoeba histolytica*." *Mol Biochem Parasitol* **72**(1-2): 1-10.
- Caballero-Salcedo, A., M. Viveros-Rogel, et al.** (1994). "Seroepidemiology of amebiasis in Mexico." *Am J Trop Med Hyg* **50**(4): 412-419.
- Cassola, A., J. G. De Gaudenzi, et al.** (2007). "Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes." *Mol Microbiol* **65**(3): 655-670.
- Clark, C. G., I. K. Ali, et al.** (2006). "Unique organisation of tRNA genes in *Entamoeba histolytica*." *Mol Biochem Parasitol* **146**(1): 24-29.

- Clark, C. G., U. C. Alsmark, et al.** (2007). "Structure and content of the *Entamoeba histolytica* genome." *Adv Parasitol* **65**: 51-190.
- Colgan, D. F. and J. L. Manley** (1997). "Mechanism and regulation of mRNA polyadenylation." *Genes Dev* **11**(21): 2755-2766.
- Coller, J. and R. Parker** (2004). "Eukaryotic mRNA decapping." *Annu Rev Biochem* **73**: 861-890.
- Conde-Bonfil, M. C. and C. de la Mora-Zerpa** (1992). "[*Entamoeba histolytica*: a standing threat]." *Salud Publica Mex* **34**(3): 335-341.
- Cook, G. C.** (1980). "Hepatic function in the presence of systemic infection." *P N G Med J* **23**(4): 153-154.
- Cougot, N., S. Babajko, et al.** (2004). "Cytoplasmic foci are sites of mRNA decay in human cells." *J Cell Biol* **165**(1): 31-40.
- Chen, C. Y. and A. B. Shyu** (2011). "Mechanisms of deadenylation-dependent decay." *Wiley Interdiscip Rev RNA* **2**(2): 167-183.
- Chu, C. Y. and T. M. Rana** (2006). "Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54." *PLoS Biol* **4**(7): e210.
- Das, S. and A. Lohia** (2002). "Delinking of S phase and cytokinesis in the protozoan parasite *Entamoeba histolytica*." *Cell Microbiol* **4**(1): 55-60.
- Davis, P. H., J. Schulze, et al.** (2007). "Transcriptomic comparison of two *Entamoeba histolytica* strains with defined virulence phenotypes identifies new virulence factor candidates and key differences in the expression patterns of cysteine proteases, lectin light chains, and calmodulin." *Mol Biochem Parasitol* **151**(1): 118-128.

- de Dios-Bravo, G., J. P. Luna-Arias, et al.** (2005). "Entamoeba histolytica TATA-box binding protein binds to different TATA variants in vitro." *FEBS J* **272**(6): 1354-1366.
- Despommier, D. D. K., J. W.** (1987). " Parasite life cycles " *Parasite life cycles*: 127.
- Diamond, L. S., D. R. Harlow, et al.** (1978). "A new medium for the axenic cultivation of *Entamoeba histolytica* and other Entamoeba." *Trans R Soc Trop Med Hyg* **72**(4): 431-432.
- Ehrenkaufer, G. M., R. Haque, et al.** (2007). "Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite." *Cell Microbiol* **9**(6): 1426-1444.
- Espinosa-Cantellano, M. and A. Martinez-Palomo** (2000). "Pathogenesis of intestinal amebiasis: from molecules to disease." *Clin Microbiol Rev* **13**(2): 318-331.
- Eulalio, A., I. Behm-Ansmant, et al.** (2007). "P bodies: at the crossroads of post-transcriptional pathways." *Nat Rev Mol Cell Biol* **8**(1): 9-22.
- Eulalio, A., I. Behm-Ansmant, et al.** (2007). "P-body formation is a consequence, not the cause, of RNA-mediated gene silencing." *Mol Cell Biol* **27**(11): 3970-3981.
- Flores-Soto, E., E. Azuara-Liceaga, et al.** (2005). "The *Entamoeba histolytica* Ehcp112 gene has a distal and weak promoter." *Exp Parasitol* **110**(3): 286-291.
- Fodinger, M., S. Ortner, et al.** (1993). "Pathogenic *Entamoeba histolytica*: cDNA cloning of a histone H3 with a divergent primary structure." *Mol Biochem Parasitol* **59**(2): 315-322.

- Fritz, D. T., N. Bergman, et al.** (2004). "Messenger RNA decay in mammalian cells: the exonuclease perspective." *Cell Biochem Biophys* **41**(2): 265-278.
- Gangopadhyay, S. S., S. S. Ray, et al.** (1997). "Unusual genome organisation in *Entamoeba histolytica* leads to two overlapping transcripts." *Mol Biochem Parasitol* **89**(1): 73-83.
- Garcia-Rivera, G., M. A. Rodriguez, et al.** (1999). "*Entamoeba histolytica* : a novel cysteine protease and an adhesin form the 112 kDa surface protein." *Mol Microbiol* **33**(3): 556-568.
- Gaucher, D. and K. Chadee** (2002). "Construction and immunogenicity of a codon-optimized *Entamoeba histolytica* Gal-lectin-based DNA vaccine." *Vaccine* **20**(27-28): 3244-3253.
- Gilchrist, C. A., E. Houpt, et al.** (2006). "Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome." *Mol Biochem Parasitol* **147**(2): 163-176.
- Gilks, N., N. Kedersha, et al.** (2004). "Stress granule assembly is mediated by prion-like aggregation of TIA-1." *Mol Biol Cell* **15**(12): 5383-5398.
- Gomez-Conde, E., P. Hernandez-Jauregui, et al.** (1998). "Chromatin organization during the nuclear division stages of live *Entamoeba histolytica* trophozoites." *Exp Parasitol* **89**(1): 122-124.
- Gomez, C., M. Esther Ramirez, et al.** (2010). "Regulation of gene expression in protozoa parasites." *J Biomed Biotechnol* **2010**: 726045.

- Gomez, C., D. G. Perez, et al.** (1998). "Transcriptional analysis of the EhPgp1 promoter of *Entamoeba histolytica* multidrug-resistant mutant." *J Biol Chem* **273**(13): 7277-7284.
- Holetz, F. B., A. Correa, et al.** (2007). "Evidence of P-body-like structures in *Trypanosoma cruzi*." *Biochem Biophys Res Commun* **356**(4): 1062-1067.
- Jagannath, A. and M. J. Wood** (2009). "Localization of double-stranded small interfering RNA to cytoplasmic processing bodies is Ago2 dependent and results in up-regulation of GW182 and Argonaute-2." *Mol Biol Cell* **20**(1): 521-529.
- Kedersha, N. and P. Anderson** (2002). "Stress granules: sites of mRNA triage that regulate mRNA stability and translatability." *Biochem Soc Trans* **30**(Pt 6): 963-969.
- Kedersha, N., G. Stoecklin, et al.** (2005). "Stress granules and processing bodies are dynamically linked sites of mRNP remodeling." *J Cell Biol* **169**(6): 871-884.
- Keene, W. E., M. E. Hidalgo, et al.** (1990). "*Entamoeba histolytica*: correlation of the cytopathic effect of virulent trophozoites with secretion of a cysteine proteinase." *Exp Parasitol* **71**(2): 199-206.
- Kulkarni, M., S. Ozgur, et al.** (2010). "On track with P-bodies." *Biochem Soc Trans* **38**(Pt 1): 242-251.
- Li, C. H., H. Irmer, et al.** (2006). "Roles of a *Trypanosoma brucei* 5'->3' exoribonuclease homolog in mRNA degradation." *RNA* **12**(12): 2171-2186.
- Lioutas, C. and E. Tannich** (1995). "Transcription of protein-coding genes in *Entamoeba histolytica* is insensitive to high concentrations of alpha-amanitin." *Mol Biochem Parasitol* **73**(1-2): 259-261.

- Liu, J., F. V. Rivas, et al.** (2005). "A role for the P-body component GW182 in microRNA function." *Nat Cell Biol* **7**(12): 1261-1266.
- Livak, K. J. and T. D. Schmittgen** (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." *Methods* **25**(4): 402-408.
- Loftus, B., I. Anderson, et al.** (2005). "The genome of the protist parasite *Entamoeba histolytica*." *Nature* **433**(7028): 865-868.
- López-Camarillo, C., López-Rosas, I** (2012). "Degradación del RNAm en eucariontes el caso del parásito *Entamoeba histolytica*." *Editorial Académica Española*: 84.
- Lopez-Camarillo, C., J. P. Luna-Arias, et al.** (2003). "EhPgp5 mRNA stability is a regulatory event in the *Entamoeba histolytica* multidrug resistance phenotype." *J Biol Chem* **278**(13): 11273-11280.
- Lopez-Casamichana, M., E. Orozco, et al.** (2008). "Transcriptional profile of the homologous recombination machinery and characterization of the EhRAD51 recombinase in response to DNA damage in *Entamoeba histolytica*." *BMC Mol Biol* **9**: 35.
- Lopez-Revilla, R. and R. Gomez** (1978). "*Entamoeba histolytica*, *E. invadens*, and *E. moshkovskii*: fluctuations of the DNA content of axenic trophozoites." *Exp Parasitol* **44**(2): 243-248.
- López-Rosas, I.** (2007). "Identificación de las poli(A) ribonucleasas de *Entamoeba histolytica* y estudio funcional de la poli(A) ribonucleasa EhCAF1." *Tesis de Maestría, UACM*.

- Luna-Arias, J. P., R. Hernandez-Rivas, et al.** (1999). "The TATA-box binding protein of *Entamoeba histolytica*: cloning of the gene and location of the protein by immunofluorescence and confocal microscopy." *Microbiology* **145** (Pt 1): 33-40.
- Lykke-Andersen, S., R. Tomecki, et al.** (2011). "The eukaryotic RNA exosome: same scaffold but variable catalytic subunits." *RNA Biol* **8**(1): 61-66.
- Lynch, E. C., I. M. Rosenberg, et al.** (1982). "An ion-channel forming protein produced by *Entamoeba histolytica*." *EMBO J* **1**(7): 801-804.
- MacFarlane, R. C., P. H. Shah, et al.** (2005). "Transcriptional profiling of *Entamoeba histolytica* trophozoites." *Int J Parasitol* **35**(5): 533-542.
- Martinez-Palomo, A.** (1987). "The pathogenesis of amoebiasis." *Parasitol Today* **3**(4): 111-118.
- Martínez-Palomo, A.** (1982). "The biology of *Entamoeba histolytica*." *Tropical Medicine Research Studies Series. Research Studies Press, John Wiley and Sons Ltd*: 161.
- Meyer, S., C. Temme, et al.** (2004). "Messenger RNA turnover in eukaryotes: pathways and enzymes." *Crit Rev Biochem Mol Biol* **39**(4): 197-216.
- Mukherjee, C., C. G. Clark, et al.** (2008). "Entamoeba shows reversible variation in ploidy under different growth conditions and between life cycle phases." *PLoS Negl Trop Dis* **2**(8): e281.
- Orozco, E., F. de la Cruz Hernandez, et al.** (1985). "Isolation and characterization of *Entamoeba histolytica* mutants resistant to emetine." *Mol Biochem Parasitol* **15**(1): 49-59.

- Orozco, E., G. Guarneros, et al.** (1983). "Entamoeba histolytica. Phagocytosis as a virulence factor." *J Exp Med* **158**(5): 1511-1521.
- Parker, R. and U. Sheth** (2007). "P bodies and the control of mRNA translation and degradation." *Mol Cell* **25**(5): 635-646.
- Parker, R. and H. Song** (2004). "The enzymes and control of eukaryotic mRNA turnover." *Nat Struct Mol Biol* **11**(2): 121-127.
- Perez, D. G., C. Gomez, et al.** (1998). "Transcriptional analysis of the EhPgp5 promoter of *Entamoeba histolytica* multidrug-resistant mutant." *J Biol Chem* **273**(13): 7285-7292.
- Petri, W. A., Jr. and J. I. Ravdin** (1987). "Cytopathogenicity of *Entamoeba histolytica*: the role of amebic adherence and contact-dependent cytolysis in pathogenesis." *Eur J Epidemiol* **3**(2): 123-136.
- Purdy, J. E., L. T. Pho, et al.** (1996). "Upstream regulatory elements controlling expression of the *Entamoeba histolytica* lectin." *Mol Biochem Parasitol* **78**(1-2): 91-103.
- Ramakrishnan, G., C. A. Gilchrist, et al.** (2004). "Histone acetyltransferases and deacetylase in *Entamoeba histolytica*." *Mol Biochem Parasitol* **138**(2): 205-216.
- Ramos, E., A. Olivos-Garcia, et al.** (2007). "*Entamoeba histolytica*: apoptosis induced in vitro by nitric oxide species." *Exp Parasitol* **116**(3): 257-265.
- Rodriguez, M. A. and E. Orozco** (1986). "[Characterization of 3 mutants of *Entamoeba histolytica* deficient in virulence with different phenotypes for adhesion and toxin activity]." *Arch Invest Med (Mex)* **17 Suppl 1**: 161-166.

- Rodriguez, M. A., M. A. Vargas, et al.** (1996). "Phenotypic variation in clones of *Entamoeba histolytica*." *Acta Parasitologica Turcica* **20**: 203-210.
- Romero-Diaz, M., C. Gomez, et al.** (2007). "Structural and functional analysis of the *Entamoeba histolytica* EhrabB gene promoter." *BMC Mol Biol* **8**: 82.
- Rosales-Encina, J. L., I. Meza, et al.** (1987). "Isolation of a 220-kilodalton protein with lectin properties from a virulent strain of *Entamoeba histolytica*." *J Infect Dis* **156**(5): 790-797.
- Schwede, A., L. Ellis, et al.** (2008). "A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells." *Nucleic Acids Res* **36**(10): 3374-3388.
- Sen, G. L. and H. M. Blau** (2005). "Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies." *Nat Cell Biol* **7**(6): 633-636.
- Sepulveda, B.** (1982). "Amebiasis: host-pathogen biology." *Rev Infect Dis* **4**(6): 1247-1253.
- She, M., C. J. Decker, et al.** (2008). "Structural basis of dcp2 recognition and activation by dcp1." *Mol Cell* **29**(3): 337-349.
- Sheth, U. and R. Parker** (2006). "Targeting of aberrant mRNAs to cytoplasmic processing bodies." *Cell* **125**(6): 1095-1109.
- Singh, U., J. B. Rogers, et al.** (1997). "Transcription initiation is controlled by three core promoter elements in the hgl5 gene of the protozoan parasite *Entamoeba histolytica*." *Proc Natl Acad Sci U S A* **94**(16): 8812-8817.
- Solis, C. F., J. Santi-Rocca, et al.** (2009). "Use of bacterially expressed dsRNA to downregulate *Entamoeba histolytica* gene expression." *PLoS One* **4**(12): e8424.

- Stalder, L. and O. Muhlemann** (2009). "Processing bodies are not required for mammalian nonsense-mediated mRNA decay." *RNA* **15**(7): 1265-1273.
- Stanley, S. L., Jr.** (2003). "Amoebiasis." *Lancet* **361**(9362): 1025-1034.
- Teixeira, D., U. Sheth, et al.** (2005). "Processing bodies require RNA for assembly and contain nontranslating mRNAs." *RNA* **11**(4): 371-382.
- Timmons, L., D. L. Court, et al.** (2001). "Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*." *Gene* **263**(1-2): 103-112.
- Torres-Guerrero, H., D. A. Peattie, et al.** (1991). "Chromatin organization in *Entamoeba histolytica*." *Mol Biochem Parasitol* **45**(1): 121-130.
- Tovy, A., R. Siman Tov, et al.** (2010). "A new nuclear function of the *Entamoeba histolytica* glycolytic enzyme enolase: the metabolic regulation of cytosine-5 methyltransferase 2 (Dnmt2) activity." *PLoS Pathog* **6**(2): e1000775.
- Vargas, M., P. Sansonetti, et al.** (1996). "Identification and cellular localization of the actin-binding protein ABP-120 from *Entamoeba histolytica*." *Mol Microbiol* **22**(5): 849-857.
- Walsh, J. A.** (1986). "Amebiasis in the world." *Arch Invest Med (Mex)* **17 Suppl 1**: 385-389.
- Walsh, J. A.** (1986). "Problems in recognition and diagnosis of amebiasis: estimation of the global magnitude of morbidity and mortality." *Rev Infect Dis* **8**(2): 228-238.

- Weber, C., G. Guigon, et al.** (2006). "Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in *Entamoeba histolytica*." *Eukaryot Cell* **5**(5): 871-875.
- Wilczynska, A., C. Aigueperse, et al.** (2005). "The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules." *J Cell Sci* **118**(Pt 5): 981-992.
- Xiang, S., A. Cooper-Morgan, et al.** (2009). "Structure and function of the 5'→3' exoribonuclease Rat1 and its activating partner Rai1." *Nature* **458**(7239): 784-788.
- Zamorano, A., C. Lopez-Camarillo, et al.** (2008). "In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for *Entamoeba histolytica* mRNA polyadenylation." *Comput Biol Chem* **32**(4): 256-263.
- Zhang, H., H. Alramini, et al.** (2011). "Nucleus-localized antisense small RNAs with 5'-polyphosphate termini regulate long term transcriptional gene silencing in *Entamoeba histolytica* G3 strain." *J Biol Chem* **286**(52): 44467-44479.
- Zhang, H., G. M. Ehrenkaufer, et al.** (2008). "Small RNAs with 5'-polyphosphate termini associate with a Piwi-related protein and regulate gene expression in the single-celled eukaryote *Entamoeba histolytica*." *PLoS Pathog* **4**(11): e1000219.
- Zhang, H., J. M. Pompey, et al.** (2011). "RNA interference in *Entamoeba histolytica*: implications for parasite biology and gene silencing." *Future Microbiol* **6**(1): 103-117.
- Zhao, J., L. Hyman, et al.** (1999). "Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis." *Microbiol Mol Biol Rev* **63**(2): 405-445.

Zuo, Y. and M. P. Deutscher (2001). "Exoribonuclease superfamilies: structural analysis and phylogenetic distribution." *Nucleic Acids Res* **29**(5): 1017-1026.

VIII. APÉNDICE

Medios de cultivo:

LB:

10g de triptona, 5g de extracto de levadura, 10g de NaCl, Agua c.b.p. 1L. Se ajusta el pH a 7.5. Se esteriliza por autoclave.

TYI-S33

34.48g de Biosate, 10.5g de glucosa anhidra, 2.3g de NaCl, 0.7g de KH_2PO_4 , 1.1g de K_2HPO_4 , 1.44g de L-cisteína, 0.3g de ácido ascorbico, 36mg de citrato ferrico, Agua c.b.p 1L. Se ajusta el pH a 6.8. Se esteriliza por autoclave

Soluciones:

Amortiguador de carga para proteínas (2X):

2.5ml de Tris-HCl 0.5 M, 0.4mg de SDS, 2ml de glicerol, 200 μg de β -mercaptoetanol y 0.1mg de azul de bromofenol en un volumen final de 10ml. pH 6.8.

Amortiguador de corrida para electroforesis de proteínas:

Tris-base 0.025 M, glicina 0.192 M y SDS 1%.

TAE (50X):

242g de Tris-base, 57.1ml de ácido acético glacial, 37.2g de $\text{Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O}$. Agua c.b.p. 1 L.

TBE (10X):

108g de Tris-base, 55g de ácido bórico, 40ml 0.5 M de EDTA pH 8. Agua c.b.p.
1L.

Solución bloqueadora para *Western blot*:

5% de leche descremada en PBS 1x pH 7.4

Amortiguador de transferencia (1X):

Tris (0.025 M), glicina (0.192 M), 20 % de metanol.

PBS pH 7.4:

NaCl 137 mM, KH₂PO₄ 1.47 mM, NaH₂PO₄.12H₂O 6.47 mM y KCl 2.68 mM.

Solución de Azul de Coomasie (1 L):

Coomasie R-250 al 0.25%, 500 ml de metanol y 70 ml de ácido acético. Agua
c.b.p.1L.

Solución desteñidora:

Etanol 30 %, ácido acético 10 % y Agua bidestilada 60%.

Reactivo de Bradford:

50mg de Azul de Coomasie G-250, 25ml de etanol al 100%, 50ml de ácido
fosfórico al 85 % y 500ml de agua bidestilada. La solución se filtra con papel
Whatman No. 1.

mRNA Decay Proteins Are Targeted to poly(A)⁺ RNA and dsRNA-Containing Cytoplasmic Foci That Resemble P-Bodies in *Entamoeba histolytica*

Itzel López-Rosas¹, Esther Orozco^{1,2}, Laurence A. Marchat³, Guillermina García-Rivera², Nancy Guillen^{4,5}, Christian Weber^{4,5}, Eduardo Carrillo-Tapia¹, Olga Hernández de la Cruz¹, Carlos Pérez-Plasencia^{6,7}, César López-Camarillo^{1*}

1 Programa en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, México City, México, **2** Departamento de Infectómica y Patogénesis Molecular, Centro de Investigación y de Estudios Avanzados, Instituto Politécnico Nacional, México City, México, **3** Programa Institucional de Biomedicina Molecular y Red en Biotecnología, Escuela Nacional de Medicina y Homeopatía, Instituto Politécnico Nacional, México City, México, **4** Unité Biologie Cellulaire du Parasitisme, Institut Pasteur, Paris, France, **5** INSERM U786, Paris, France, **6** Unidad de Genómica y Secuenciación Masiva, Instituto Nacional de Cancerología, México City, México, **7** Unidad de Biomedicina, Facultad de Estudios Superiores-Iztacala, Universidad Nacional Autónoma de México, México City, México

Abstract

In higher eukaryotes, mRNA degradation and RNA-based gene silencing occur in cytoplasmic foci referred to as processing bodies (P-bodies). In protozoan parasites, the presence of P-bodies and their putative role in mRNA decay have yet to be comprehensively addressed. Identification of P-bodies might provide information on how mRNA degradation machineries evolved in lower eukaryotes. Here, we used immunofluorescence and confocal microscopy assays to investigate the cellular localization of mRNA degradation proteins in the human intestinal parasite *Entamoeba histolytica* and found evidence of the existence of P-bodies. Two mRNA decay factors, namely the *Eh*XRN2 exoribonuclease and the *Eh*DCP2 decapping enzyme, were localized in cytoplasmic foci in a pattern resembling P-body organization. Given that amoebic foci appear to be smaller and less rounded than those described in higher eukaryotes, we have named them “P-body-like structures”. These foci contain additional mRNA degradation factors, including the *Eh*CAF1 deadenylase and the *Eh*AGO2-2 protein involved in RNA interference. Biochemical analysis revealed that *Eh*CAF1 co-immunoprecipitated with *Eh*XRN2 but not with *Eh*DCP2 or *Eh*AGO2-2, thus linking deadenylation to 5'-to-3' mRNA decay. The number of *Eh*CAF1-containing foci significantly decreased after inhibition of transcription and translation with actinomycin D and cycloheximide, respectively. Furthermore, results of RNA-FISH assays showed that (i) *Eh*CAF1 colocalized with poly(A)⁺ RNA and (ii) during silencing of the *Ehpc4* gene by RNA interference, *Eh*AGO2-2 colocalized with small interfering RNAs in cytoplasmic foci. Our observation of decapping, deadenylation and RNA interference proteins within P-body-like foci suggests that these structures have been conserved after originating in the early evolution of eukaryotic lineages. To the best of our knowledge, this is the first study to report on the localization of mRNA decay proteins within P-body-like structures in *E. histolytica*. Our findings should open up opportunities for deciphering the mechanisms of mRNA degradation and RNA-based gene silencing in this deep-branching eukaryote.

Citation: López-Rosas I, Orozco E, Marchat LA, García-Rivera G, Guillen N, et al. (2012) mRNA Decay Proteins Are Targeted to poly(A)⁺ RNA and dsRNA-Containing Cytoplasmic Foci That Resemble P-Bodies in *Entamoeba histolytica*. PLoS ONE 7(9): e45966. doi:10.1371/journal.pone.0045966

Editor: Georg Stoecklin, German Cancer Research Center, Germany

Received: April 4, 2012; **Accepted:** August 23, 2012; **Published:** September 24, 2012

Copyright: © 2012 López-Rosas et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by Mexico-France (SEP-CONACYT-ANUIES; ECOS NORD) grant M08-S02 and European community grants. The authors acknowledge Autonomous University of Mexico City for support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: cesar.lopez@uacm.edu.mx

Introduction

The control of messenger RNA (mRNA) degradation plays a key role in the posttranscriptional regulation of gene expression [1]. In eukaryotic cells, mRNA degradation is initiated by shortening of the 3' end-poly(A)⁺ tail by the CCR4/CAF1/NOT1–5 complex and the PAN2, PAN3 and PARN deadenylases [2]. Next, the DCP1 and DCP2 decapping enzymes remove the cap from the 5' end of the deadenylated mRNA in a process stimulated by Dhh1, Ede3, Pat1/Mrt1, Sm-like (Lsm1p–7p) and Hedls proteins [3]. Lastly, the uncapped mRNA body is degraded by XRN1 and XRN2 exoribonucleases in the 5'-to-3' direction.

Alternatively, mRNA is degraded in the 3'-to-5' direction by a complex of PH-like ribonucleases known as the exosome [4].

There is growing body of evidence to show that mRNA degradation occurs in specialized cytoplasmic foci variously referred to as mRNP granules, mRNA-decay foci, GW182-bodies and mRNA processing bodies (P-bodies). These structures are enriched in RNA substrates and mRNA turnover proteins. P-bodies are mRNA processing centers within which non-translating transcripts are sorted and either silenced or degraded [5]. Although the protein inventory of P-bodies has not been defined in detail, around 25 different factors have been detected within these small cytoplasmic domains [6]. The P-bodies observed in

yeast, insect, nematode and mammalian cells have critical roles in mRNA degradation, mRNA storage, mRNA surveillance and RNA-based gene silencing mechanisms [7]. In fact, it has been demonstrated that P-bodies have high concentration of target transcripts and the AGO2, DICER, and GW182 proteins involved in RNA interference by microRNAs (miRNAs) and small interfering RNAs (siRNAs) [8,9,10]. In addition, P-bodies also contain the UPF and SMG families of proteins involved in the nonsense mediated decay (NMD) mechanism that degrades aberrant mRNAs containing premature translation termination codons. [11]. In mammals, most cells contain from three to nine P-bodies, although this number appears to depend on the cell's proliferation ratio and stress conditions, such as nutrient deprivation, oxidative stress and heat shock. Cell treatments with inhibitors of transcription or translation (such as actinomycin D and cycloheximide, respectively) lead to the disassembly of P-bodies [12]. In contrast, treatment with puromycin increases the number and size of P-bodies; this observation emphasizes the relevance of active mRNA and protein synthesis in the assembly of these structures [10]. Even though P-body components clearly play crucial roles in mRNA degradation, there is also evidence to show that NMD and silencing mediated by miRNAs and siRNAs occur in cells lacking detectable P-bodies. This suggests that P-body formation is a consequence, rather than a cause, of mRNA degradation activity [10,13]. Furthermore, translational repression by miRNAs does not require P-body structures; the localization of miRNAs and RNA-induced silencing complex (RISC) proteins within P-bodies appears to be a consequence of translational repression [14].

Human amebiasis is caused by infection with the intestinal protozoan parasite *Entamoeba histolytica*. The disease is endemic in many developing countries and affects around 50 million people worldwide [15]. The symptoms of amebiasis include colitis, dysentery and extra-intestinal abscesses. Despite the importance of post-transcriptional events in the regulation of gene expression, the parasite's mRNA degradation and RNA-based gene silencing mechanisms are poorly understood. Some proteins involved in the RNA interference pathway, including *Eh*AGO2-2, *Eh*RNAseIII and *Eh*RdRP, have been identified and characterized in *E. histolytica* [16,17]. However, proteins involved in mRNA decay pathways have not previously been studied and the potential role of P-bodies in mRNA decay and RNA interference has not been clearly established in protozoan parasites. Here, we sought to localize mRNA degradation factors within *E. histolytica* and found evidence of the existence of P-body-like structures. Our findings open up opportunities for deciphering the mechanisms of mRNA decay and RNA gene silencing in this parasitic, deep-branching eukaryote.

Results

mRNA degradation machineries in *Entamoeba histolytica*

In order to identify homologous genes potentially involved in mRNA decay in *E. histolytica*, we screened the parasite genome at AmoebaDB (<http://amoebadb.org/amoeba/>) by using the amino acid sequences of yeast and human proteins involved in mRNA turnover as probes. Heuristic searches identified several putative protein-encoding genes for mRNA degradation machineries, including deadenylation, decapping, NMD, and RNA interference activities (Figure 1, Table S1). Our data indicated that although several important genes are lacking, mRNA decay machineries are generally well conserved in *E. histolytica*. We detected putative *Eh*dcp2, *Eh*lsm1–6, *Eh*edc3 and *Eh*dhh1 genes (coding for mRNA decapping factors) and an *Eh*xrn2 gene (coding for a putative 5'-to-

3' exoribonuclease). In contrast, we did not find any *dcp1* or *xm1* homologs. We also identified genes for NOT1–5 subunits and for *Eh*CAF1 and *Eh*CAF1-like deadenylases. In contrast, the lack of a *ccr4* deadenylase homolog suggested the existence of a CAF1/NOT complex lacking one of the major deadenylases. Intriguingly, we did not detect *pan2*, *pan3* and *pan* genes; this indicates that *E. histolytica*'s deadenylation machinery is smaller than those described in yeast and in the human. Amoeba genome also contains genes for the NMD machinery, as represented by three *Eh*upf genes (Table S1). RNA interference pathway members, including *Eh*RdRP, *Eh*RNAseIII and *Eh*AGO2-2 proteins, have already been reported in this parasite [16,17]. We (like others) did not detect DICER and GW182 homologs, suggesting that RNA interference may occur by small interfering RNAs generated by DICER-independent mechanisms in *E. histolytica* [18].

mRNA degradation genes are differentially expressed under diverse stress conditions

We next sought to establish whether mRNA degradation genes are transcribed in *E. histolytica* trophozoites grown in basal culture conditions and how these genes are modulated in response to the stress induced by heat shock, UV-C-induced DNA damage and nitric oxide treatments, as described elsewhere [19,20,21]. We designed specific primers (Table S2) and performed quantitative real time-PCR assays for representative genes from the various mRNA degradation machineries (Figure 2). After heat shock, *Eh*xrn2, *Eh*lsm1, *Eh*upf1, and *Eh*mase III gene expression was found to be upregulated 1.5- to 12-fold relative to non-treated cells, whereas the *Eh*caf1, *Eh*not1, *Eh*dcp2 and *Eh*ago2-2 genes were repressed 2-fold (Figure 2). After UV-C-induced DNA damage, *Eh*upf1 was upregulated by a factor of 6, whereas the expression of *Eh*caf1, *Eh*not1, *Eh*dcp2, *Eh*lsm1, *Eh*edc3 and *Eh*mase III genes was significantly repressed. *Eh*xrn2 and *Eh*ago2-2 mRNA levels were similar in treated and non-treated cells. *Eh*not1 and *Eh*dcp2 genes were significantly repressed after exposure to nitrosative stress in comparison with untreated trophozoites, whereas *Eh*caf1, *Eh*lsm1, *Eh*upf1, *Eh*ago2-2 and *Eh*mase III were upregulated 2- to 16-fold. We quantified the mRNA expression of the L31 ribosomal protein-encoding gene as an endogenous control; its expression did not change significantly under any of the tested stress conditions.

*Eh*XRN2 and *Eh*DCP2 have the typical architecture of a 5'-3' exoribonuclease and a decapping protein, respectively

To initiate the study of mRNA degradation proteins in *E. histolytica*, we focused on the putative *Eh*XRN2 exoribonuclease and the *Eh*DCP2 decapping enzyme. We selected these proteins because their homologs have been described as P-body markers in the human and in yeast [22]. Homologous XRN2 and DCP2 proteins from a variety of organisms were assigned by BLAST analysis and ClustalW multi-alignments of predicted amino acid sequences (Figure S1A, S2A, and Table S3, S4). *Eh*xrn2 is an intronless gene (2813 nt) that predicts a 938 amino acids polypeptide. Sequence similarity searches revealed that *Eh*XRN2 exhibits 46% identity ($e = 4.9\text{e-}139$) and 45% identity ($e = 1.2\text{e-}108$) with *Homo sapiens* XRN2 and *Schizosaccharomyces pombe* RAT1, respectively. Phylogenetic inference of full-length XRN2 sequences revealed that *Eh*XRN2 is closely related to homologous proteins from other protozoa (Figure S1B). Comparison of the amino acids sequence of *Eh*XRN2 with those of *Hs*XRN2 and *Sp*RAT1 homologs indicate the presence of conserved functional and structural motifs as represented by the XRN_N nuclease domain and the internal tower domain (Figure 3A), which is responsible for catalytic activity in *Sp*RAT1 [23]. Importantly, the active site

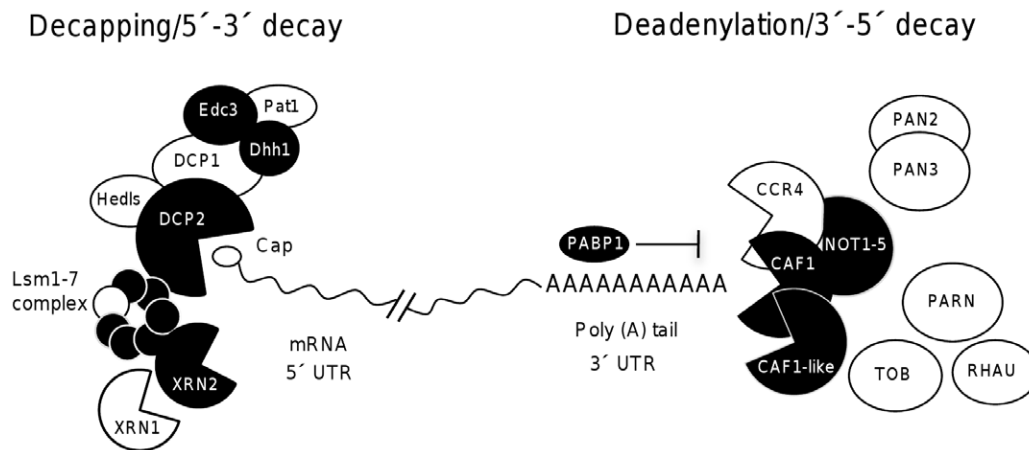


Figure 1. A comparative representation of mRNA decay machineries in *E. histolytica* (black shapes) and *H. sapiens* (white and black shapes). Homologous proteins were assigned via BLAST and ClustalW multiple alignments. The undulated line denotes mRNA. doi:10.1371/journal.pone.0045966.g001

motif KX_2QQX_2RR , which is critical for *SpRAT1* ribonuclease function, is well conserved in *EhXRN2*. The high degree of homology between *EhXRN2* and *SpRAT1* allowed us to construct a structural model of *EhXRN2* by using *SpRAT1* crystal data as template. The predicted tertiary structure suggested that the *EhXRN2* N-terminal domain is constituted by a six α -helix bundle and a two-stranded β -sheet, the orientation and length of which is similar to that found in the *SpRAT1* structural model (Figure 3B–C).

The *Ehdcp2* gene (699 nt) contains two introns and predicts an open reading frame of 233 amino acids. *EhDCP2* exhibits 28% ($e = 8.4e-25$) and 29% ($e = 1.6e-19$) identity with human DCP2 and yeast DCP2, respectively. Phylogenetic inference revealed that

EhDCP2 is closely related to homologous proteins from other protozoa (Figure S2B). *EhDCP2* contains the conserved DCP2 box A domain (Figure 3D) and the consensus sequence for the nudix motif $GX_5EX_7REUXEEXGU$ (including the three catalytic glutamate (E) residues). In mammalian cells, both these domains are responsible for cap structure removal [24]. Modeling of *EhDCP2*'s tertiary structure using the *ScDCP2* structure as a template indicated that the two proteins are similar in terms of conformation and length (Figure 3E–F).

EhXRN2, *EhDCP2* and *EhCAF1* are cytoplasmic proteins

We next investigated the cellular localization of *EhXRN2* and *EhDCP2* factors in *E. histolytica*. An *Ehxm2* gene fragment (351 nt)

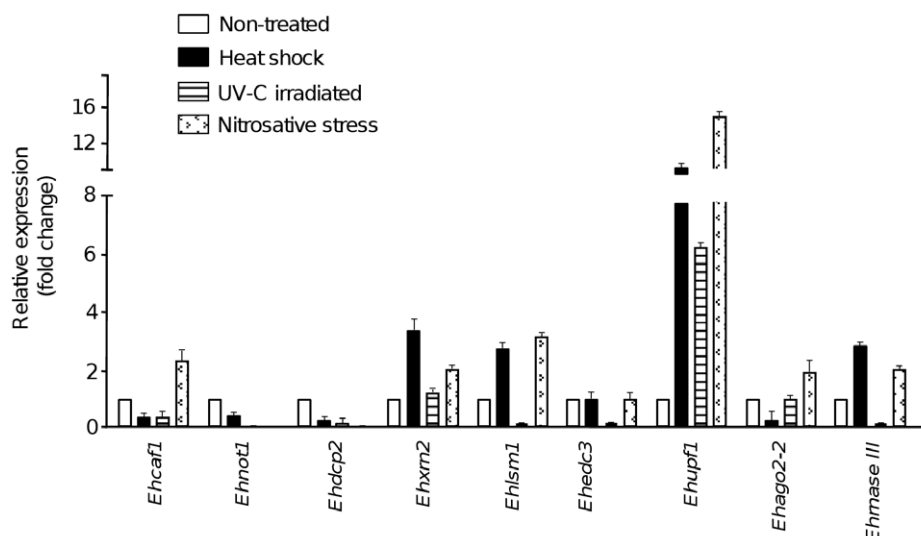


Figure 2. mRNA expression profiles of genes for *E. histolytica* mRNA degradation proteins. Quantitative real-time PCR assays were performed to analyze the relative expression of representative mRNA degradation genes after heat shock, UV-C induced DNA damage and sodium nitroprusside treatments. For each triplicate experiment, the mean of the relative concentrations obtained for the tested mRNA were divided by the mean of the corresponding values obtained for endogenous ribosomal *L31* amplification. Each PCR experiment was carried out three times and three independent biological samples were analyzed. The relative expression (fold change) of mRNA degradation genes in the different treatments was calculated by the $2(-\Delta\Delta Ct)$ method) using as reference the Ct data for the untreated condition. Bars indicate the mean $\pm$ SD. doi:10.1371/journal.pone.0045966.g002

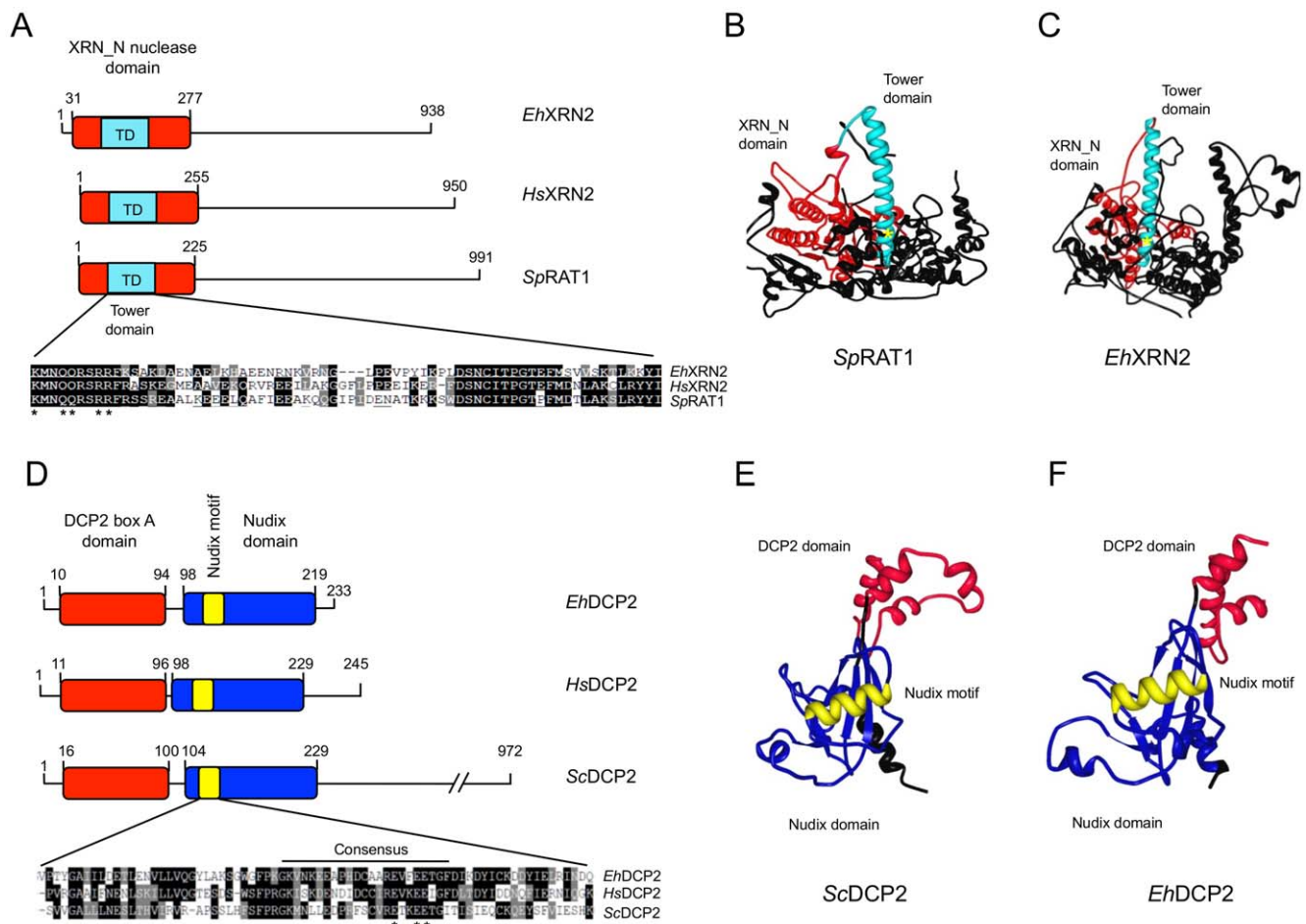


Figure 3. Comparison of XRN2 and DCP2 from eukaryotic organisms and *E. histolytica*. (A) A schematic representation of XRN2 protein from *E. histolytica*, *H. sapiens* and *S. pombe*. Red box: the XRN domain; cyan box: the tower domain (TD). The numbers are relative to the initial methionine in each protein. Lower panel: multiple alignments of amino acid residues from the TD of homologous proteins. Black box: a conserved residue; grey box: a conserved substitution. Asterisks mark the conserved residues in the active site. (B, C) Ribbon diagrams of ScRAT1's tertiary structure (B) and the predicted model of EhXRN2 (C). Red: the N-terminal XRN_N domain; cyan: the TD. (D) A schematic representation of the DCP2 protein from *E. histolytica*, *H. sapiens*, and *S. cerevisiae*. Red box: the DCP2 domain; blue box: the nudix domain. The numbers are relative to the initial methionine in each protein. Lower panel: multiple alignments of amino acid residues from the nudix motif (PF00293) of homologous proteins. Asterisks indicate the conserved residues involved in catalytic activity. Black box: a conserved residue; grey box: a conserved substitution. (E, F) Ribbon diagrams of ScDCP2's tertiary structure (E) and the predicted model of EhDCP2 (F). Red: the DCP2 domain; blue: the nudix domain; yellow: the nudix motif. Three-dimensional models were displayed and refined using the RCSB PDB Protein Workshop 3.9 viewer. doi:10.1371/journal.pone.0045966.g003

and the full-length *Ehdcp2* gene were PCR-amplified from genomic DNA and cloned in-frame into the pRSET-A expression vector. Recombinant 6xHis-tagged *EhXRN2* and *EhDCP2* polypeptides were expressed in bacteria (Figure 4A and B; lane 2), purified by affinity chromatography and used as immunogens in the mouse to obtain specific antibodies. In Western blot assays based on anti-*EhXRN2* and *EhDCP2* antibodies, we detected the corresponding purified recombinant proteins (Figure 4A and B, lane 5). The subcellular localization of *EhXRN2* and *EhDCP2* were investigated by Western blotting cytoplasmic and nuclear protein extracts from *E. histolytica* trophozoites (Figure 4C). In the cytoplasmic fraction, specific antibodies recognized bands at 103 kDa and 27.5 kDa, which correspond to the predicted molecular weights for endogenous *EhXRN2* and *EhDCP2* proteins, respectively (Figure 4C, lanes 2 and 4). In contrast, no proteins in the nuclear fraction (lanes 1 and 3) were recognized. The results of control assays with used pre-immune serum (which did not give any signal;

data not shown) and antibodies that recognized the *EhPC4* transcription factor (found mainly in the nuclear fraction (lane 5), as reported elsewhere [25]) confirmed the absence of cross-contamination between cytoplasmic and nuclear fractions.

We then used immunofluorescence and confocal microscopy assays to study the distribution of mRNA degradation proteins within trophozoites. As shown in Figure 4, *EhXRN2* and *EhDCP2* proteins were diffusely immunodetected throughout the cell (D-H and I-M) but were mainly concentrated in bright cytoplasmic foci with a diameter ranging from 50 to 300 nm; this resembled the organization into P-bodies seen in higher eukaryotes. To further investigate the localization of additional mRNA degradation proteins within trophozoites, we studied the cellular distribution of *EhCAF1* deadenylase using specific rabbit antibodies generated in-house (manuscript in preparation). In particular, *EhCAF1* had a similar pattern to *EhXRN2* and *EhDCP2* within cytoplasmic foci and a more diffuse general staining (Figure 4N-R).

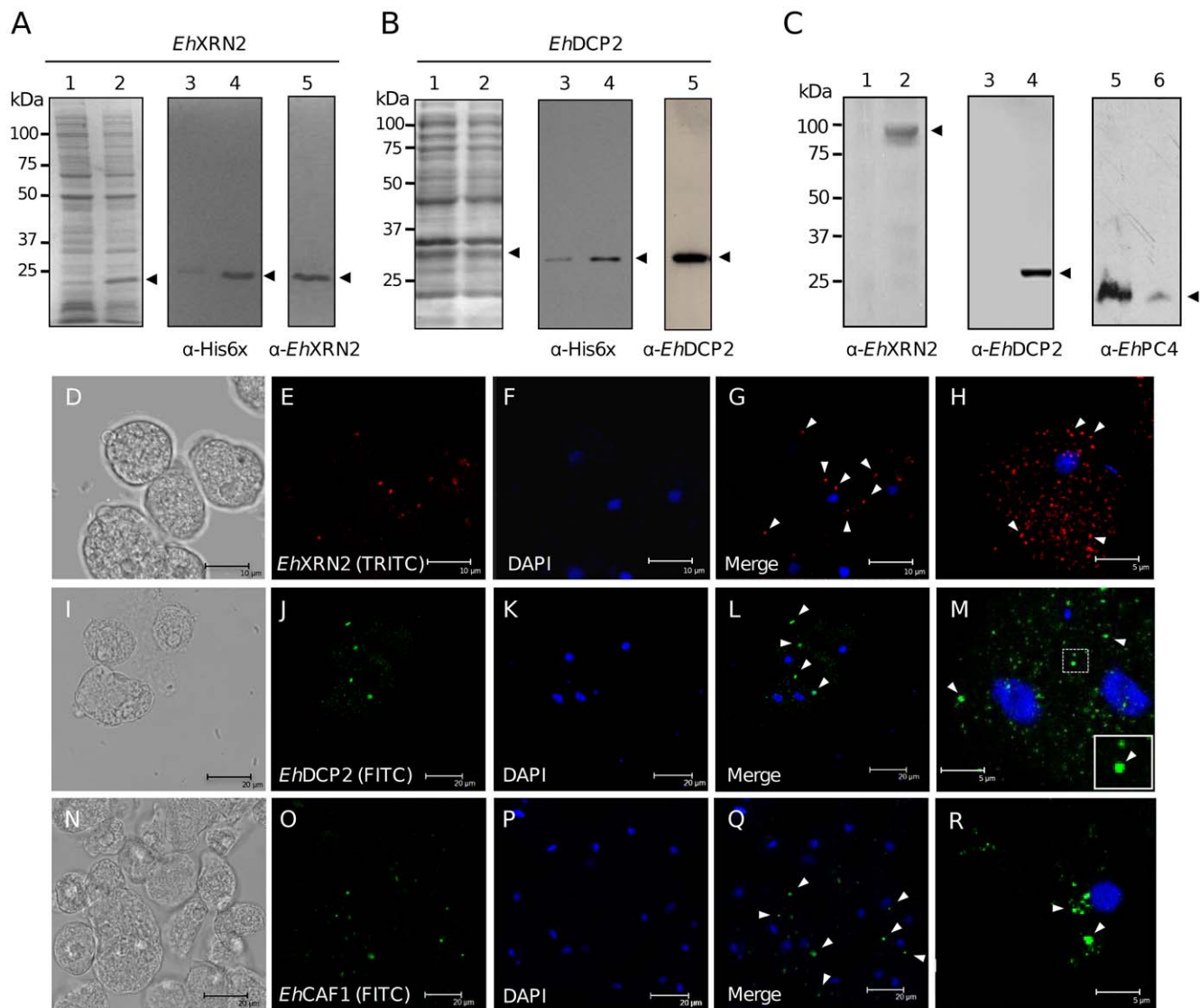


Figure 4. Expression and immunolocalization of *EhXRN2*, *EhDCP2* and *EhCAF1*. (A–B) Expression of 6xHis-tagged *EhXRN2* (A) and *EhDCP2* (B) proteins. Bacterial proteins were separated by 12% SDS-PAGE and the gels were stained with Coomassie blue. Lane 1: non-induced bacterial extracts; lane 2: IPTG-induced bacterial extracts. Western blot assays were performed using anti-6xHis tag antibodies and non-induced bacterial extracts (lane 3) or IPTG-induced bacterial extracts (lane 4). Recombinant *EhXRN2* and *EhDCP2* proteins were also immunodetected with specific antibodies (lane 5). (C) Immunodetection of *EhXRN2* (lanes 1 and 2), *EhDCP2* (lanes 3 and 4) and *EhPC4* (lanes 5 and 6) proteins in trophozoites, using specific antibodies. Lanes 1, 3 and 5: nuclear extracts; lanes 2, 4 and 6: cytoplasmic extracts. (D–R) Confocal immunofluorescence microscopy assays for the localization of *EhXRN2* (E), *EhDCP2* (J) and *EhCAF1* (O) in fixed trophozoites using FITC-antibodies. Middle panels (F, K, P) show nuclear staining with DAPI and right-hand panels (G, L, Q) show overlays of the two signals. Different magnifications of single cells were used to enhance the visibility of cytoplasmic structures (H, M, R). Arrowheads indicated some typical foci.
doi:10.1371/journal.pone.0045966.g004

EhXRN2, *EhDCP2* and *EhAGO2-2* colocalize within *EhCAF1*-containing cytoplasmic foci

The observed colocalization of *EhXRN2*, *EhDCP2* and *EhCAF1* prompted us to look for common focal, cytoplasmic structures. We performed co-immunolocalization and confocal microscopy assays in trophozoites by using *EhCAF1* as a marker for the P-body-like structures (Figure 5). Our data showed that both *EhXRN2* and *EhDCP2* co-immunolocalized within the majority of *EhCAF1*-containing-cytoplasmic foci (Figure 5, A–D and E–H). We next wondered whether additional factors related to mRNA degradation might also accumulate with *EhXRN2* and *EhDCP2* in a unique set of *EhCAF1*-containing foci. Hence, we performed immunofluorescence experiments using antibodies

raised against *EhAGO2-2* an Argonaut protein involved in RNA interference [18]. Indeed, the *EhAGO2-2* signal also colocalized with *EhCAF1* in trophozoites (Figure 5, I–L). None of these proteins gave a significant nuclear signal. Taken as a whole, our data indicate that these mRNA degradation factors are concentrated within cytoplasmic P-body-like structures.

We next sought to establish whether the colocalized *EhXRN2*, *EhDCP2*, *EhCAF1*, and *EhAGO2-2* proteins could interact physically by performing co-immunoprecipitation assays with anti-*EhCAF1* or -*EhXRN2* antibodies as bait and trophozoite cytoplasmic proteins as prey (Figure 5M). Our results showed that *EhCAF1* co-immunoprecipitated with *EhXRN2* and *vice versa*; these data fit with the co-immunolocalization studies and suggest a

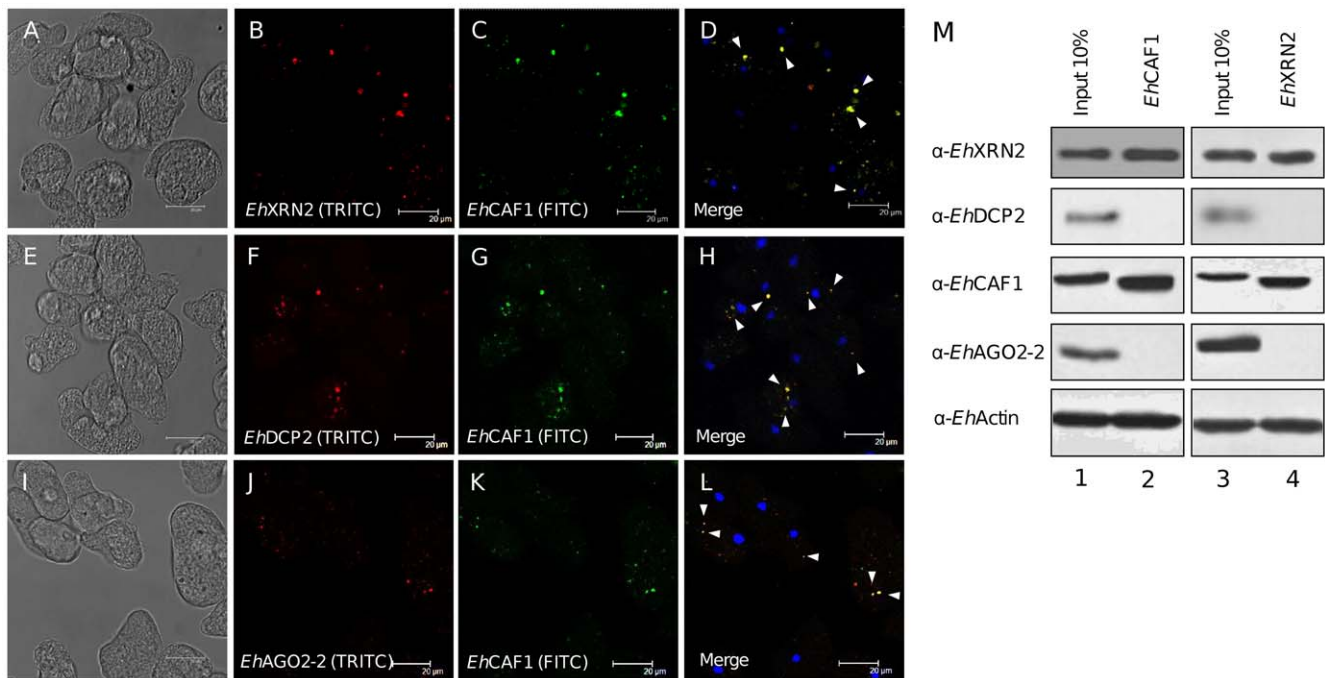


Figure 5. Colocalization of mRNA decay factors with *EhCAF1* in cytoplasmic foci. (A–L) Confocal immunofluorescence microscopy assays. Trophozoites were incubated with anti-*EhXRN2* (B), anti-*EhDCP2* (F), anti-*EhAGO2-2* (J) and anti-*EhCAF1* (C, G, K) primary antibodies (1:100). Secondary antibodies included TRITC-conjugated anti-mouse IgG (1:200) and FITC-conjugated anti-rabbit IgG (1:200). (A, E and I): light microscopy; (B, F and J): the red channel; (C, G and K): the green channel; (D, H and L): merged images. Arrowheads mark typical, colocalized signals in cytoplasmic foci. (M) Immunoprecipitation assays using anti-*EhCAF1* and anti-*EhXRN2* antibodies conjugated to protein G beads with cytoplasmic lysates of *E. histolytica* trophozoites. Immunoprecipitates were separated with SDS-PAGE and analyzed by Western blotting with the antibodies indicated at the left side of each panel. Lanes 1 and 3: input (10% of the total lysate used for immunoprecipitation); lanes 2 and 4: immunoprecipitated proteins. doi:10.1371/journal.pone.0045966.g005

novel functional link between deadenylation and 5'-to-3' mRNA degradation (Figure 5M, upper panel, lanes 2 and 4). In contrast, we did not detect co-immunoprecipitation of *EhCAF1* with either *EhDCP2* or *EhAGO2-2* - suggesting that despite their colocalization in cytoplasmic foci in trophozoites, these proteins do not interact. Surprisingly, we also found co-immunoprecipitation of actin protein with both baits (Figure 5M, bottom panel, lanes 2 and 4). In agreement with these results, we also identified actin as an *EhCAF1*-interacting protein in pull-down assays using GST-*EhCAF1* recombinant fusion protein as bait and cytoplasmic proteins as prey (manuscript in preparation). It has been reported that mammalian P-body proteins are anchored to the cytoskeleton and associate with tubulin and actin-bundles in human cells, suggesting that P-bodies require a dynamic interaction with the cytoskeleton [26].

The formation of cytoplasmic foci depends on the presence of active transcription

To evaluate the impact of active ongoing transcription on P-body size and number, we measured the change over time in the number of *EhCAF1*-containing foci after transcription inhibition with actinomycin D, as described elsewhere [27]. Our results showed that abrogation of transcription considerably altered the number of cytoplasmic foci (Figure 6A). We observed that number of cytoplasmic foci per cell diminished significantly after 4 h of incubation with actinomycin D. After 12 h treatment, the number of foci per cell had fallen by 50%. This effect was not observed in cells treated with the vehicle DMSO alone. The results of Western blot assays showed that the expression level of *EhCAF1* was not significantly affected by actinomycin D treatment (data not

shown). Hence, the fall in the number of foci was not due to the inhibition of *EhCAF1* expression. Our data indicate that cytoplasmic foci formation is linked to active transcription – a feature that is conserved in P-bodies in higher eukaryotes.

Inhibition of translation affects cytoplasmic foci formation

A common feature of mammalian P-bodies is that they are disrupted by drugs such as cycloheximide that stabilize mRNAs in polysomes, whereas their number and size are increased by drugs such as puromycin that release ribosomes from mRNA [10,12]. To evaluate the effect of inhibiting protein synthesis on the number and size of *EhCAF1*-containing cytoplasmic foci, we treated *E. histolytica* trophozoites with cycloheximide or puromycin. The results showed that the number of foci per cell rapidly fell after 15 min of incubation with cycloheximide and almost disappear after 120 min of treatment (Figure 6B). In contrast, puromycin treatment led to a slight increase in the number of cytoplasmic foci after 120 min (Figure 6C). Control assays showed that DMSO alone did not have a significant effect on the number of *EhCAF1*-containing foci.

EhCAF1 and *EhAGO2-2* colocalize with poly(A)⁺ RNA and dsRNA substrates

We next looked at whether polyadenylated RNA substrates accumulate in P-body-like foci under basal culture conditions. Poly(A)⁺ RNAs were detected in trophozoites by fluorescence *in situ* hybridization (FISH) with a FITC-labeled oligo(dT)₃₀ probe. *EhCAF1* foci were evidenced by immunofluorescence with specific antibodies. Interestingly, our results showed that although

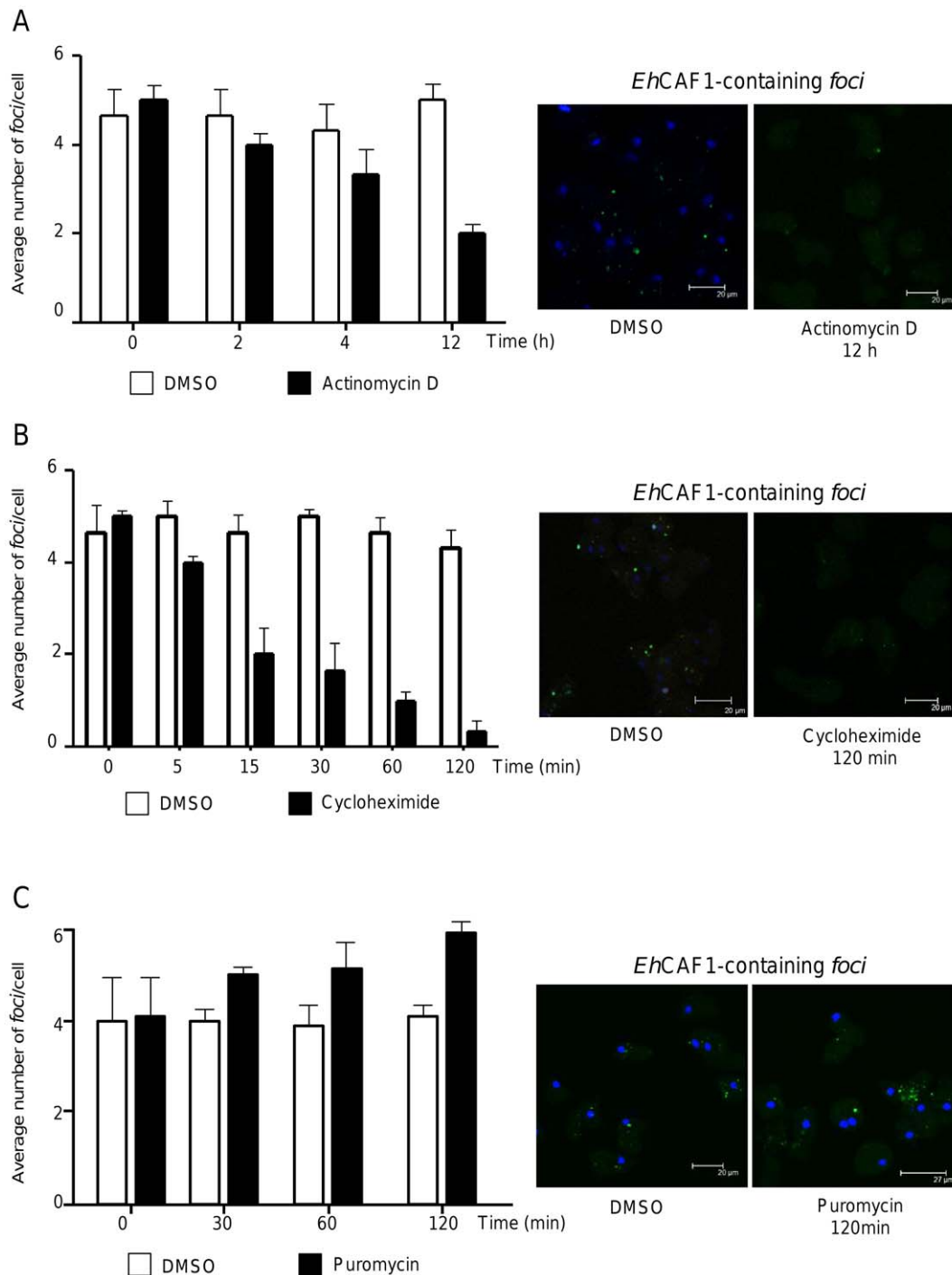


Figure 6. Determination of the number of *EhCAF1*-containing cytoplasmic foci after treatment with actinomycin D, cycloheximide or puromycin. After treatment with actinomycin D (5 μ g/ml, panel A), cycloheximide (10 μ g/ml, panel B) or puromycin (200 μ g/ml, panel C), trophozoites were fixed, incubated with anti-*EhCAF1* and FITC-labeled secondary antibodies, counterstained with DAPI and analyzed under an immunofluorescence confocal microscope. Left panel: change over time in the mean number of foci per trophozoite, in response to inhibition of transcription (A) or translation (B–C). The data correspond to six randomly chosen fields. Right panel: representative images of cells treated with DMSO vehicle (control) and actinomycin D, cycloheximide or puromycin (observed in the green channel). doi:10.1371/journal.pone.0045966.g006

the FITC-oligo(dT)₃₀ signal was dispersed throughout the cytoplasm, a small proportion of probe colocalized with immunostained *EhCAF1* in cytoplasmic foci (Figure 7E). Control experiments showed that the FITC-oligo(dT)₃₀ signal was specific

for RNA, since it disappeared after RNase A treatment (Figure 7F–J).

To gain insights into the potential role of P-body-like structures in RNA interference, we sought to determine whether a specific double-stranded RNA (dsRNA) accumulated in cytoplasmic foci

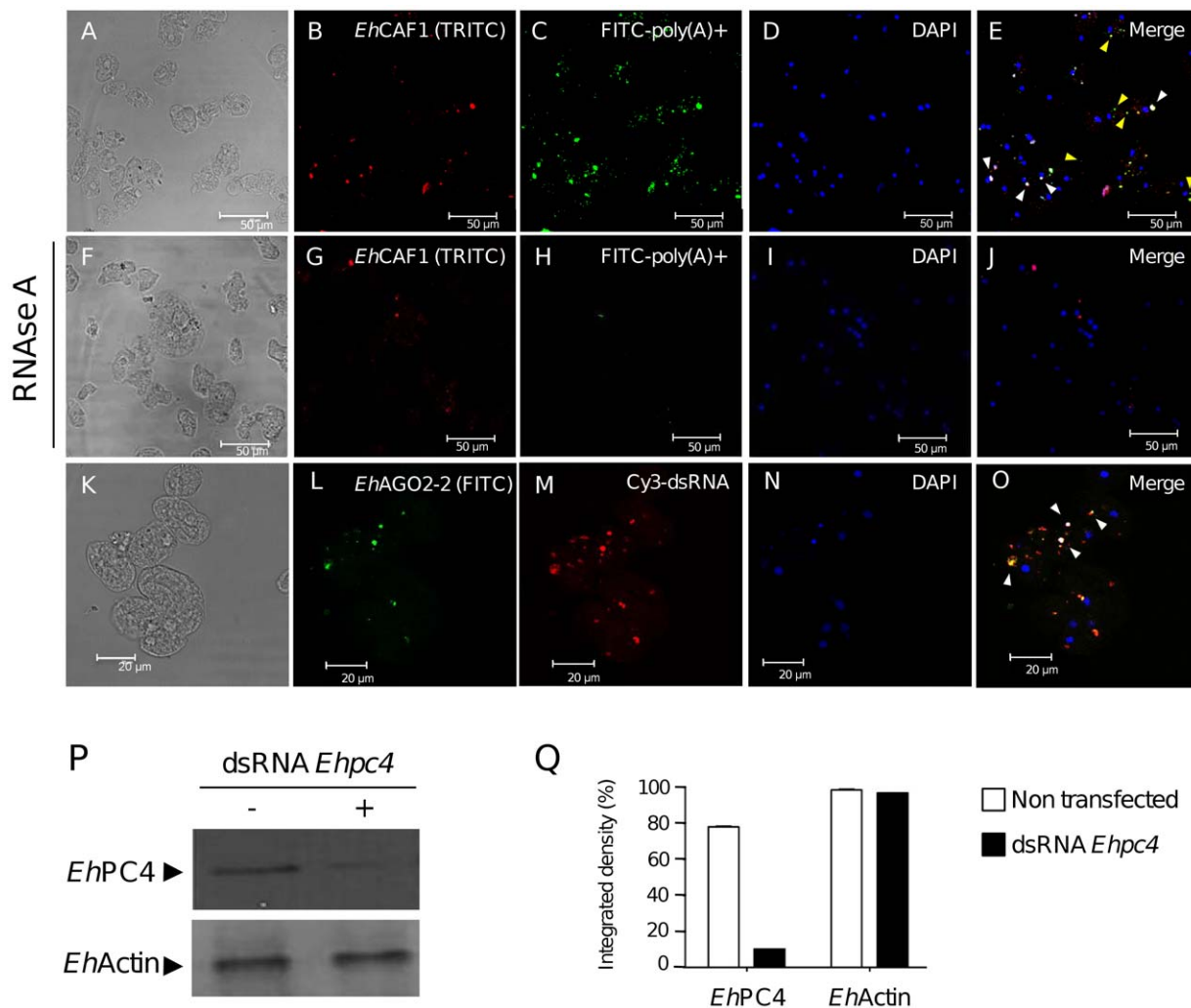


Figure 7. Colocalization of poly(A)⁺ RNA and dsRNA substrates with *EhCAF1* and *EhAGO2-2* in cytoplasmic foci. (A–E) Poly(A)⁺ RNA and *EhCAF1* colocalization assays. Trophozoites were immunostained with *EhCAF1* (B) antibodies. Poly(A)⁺ RNAs were detected by hybridization with FITC-conjugated oligo-(dT)₃₀ (C). The cells were counterstained with DAPI (D) and analyzed with immunofluorescence confocal microscopy. The merged image (E) shows the overlapping signals. White arrowheads indicate the accumulation of poly(A)⁺ RNA and *EhCAF1* signals. Yellow arrowheads indicate FITC-poly(A)⁺ signal that does not overlap with *EhCAF1*-containing foci. (F–J) RNase A treatment, performed as a control experiment. Immunodetection of *EhCAF1* (G), detection of poly(A)⁺ (H) and DNA counterstaining with DAPI (I). Merged image (J). (K–O) *EhAGO2-2* and Cy3-dsRNA colocalization assays. Trophozoites were transfected with Cy3-dsRNA targeting *Ehpc4* (K–O) and then immunostained with *EhAGO2-2* antibodies on day seven after transfection (L). Cells were counterstained with DAPI (N) and analyzed with immunofluorescence confocal microscopy. The merged image (O) shows both signals. Arrowheads mark colocalized signals in cytoplasmic foci. (P) Western blot analysis for *EhPC4* and *EhActin* from proteins extracts obtained on day seven after Cy3-dsRNA transfection. (Q) A densitometric analysis of the bands in P. doi:10.1371/journal.pone.0045966.g007

and colocalized with *EhAGO2-2* during silencing of an endogenous gene. The long dsRNA was designed to specifically knock down the *Ehpc4* gene expression and its efficacy was validated by using the soaking approach described for amoeba [28]. Briefly, dsRNA was generated in bacteria, purified, Cy3-labeled and transfected into trophozoites to produce fluorescent siRNA. Our data showed that the dsRNA-Cy3 signal was found in discrete, cytoplasmic foci seven days after transfection of the trophozoites (Figure 7M). Moreover, coupled FISH and immunofluorescence assays using the dsRNA-Cy3 probe and *EhAGO2-2* antibodies showed that the two signals colocalized in large, cytoplasmic foci (Figure 7K–O). Remarkably, the focus size and signal intensity for *EhAGO2-2* were greater in cells than in non-transfected, control cells - suggesting that *EhAGO2-2* protein expression was induced by siRNA transfection. Western blot assays confirmed that *EhPC4*

protein levels were reduced by 50% and 80% three and seven days after dsRNA-Cy3 transfection, respectively. This indicated that RNA interference had been successful (Figure 7P–Q). We can rule out the possibility that the cytoplasmic foci containing siRNA against the *Ehpc4* gene were non-transfected vesicles or endosomes, since we used the soaking method to deliver the dsRNA-Cy3 probe to trophozoites. However, we cannot be sure that only the siRNA guide strand was located in the cytoplasmic foci, since it has been reported that siRNAs can localize in P-bodies as dsRNA [29].

Discussion

mRNA decay pathways and machineries have been extensively studied in *S. cerevisiae* and *H. sapiens*. In contrast, little is known about mRNA degradation factors in protozoan parasites. Here, we

identified protein-encoding genes for mRNA degradation machineries in *E. histolytica*. Our identification of decapping, deadenylation, NMD and RNA interference activities suggests that mRNA degradation components and mechanisms are conserved in this parasite. However, the composition of these machineries is simpler than that seen in higher eukaryotic cells - a frequent feature in the small genomes of divergent parasites. Importantly, we found that the deadenylation machinery is incomplete in *E. histolytica* because we did not find *ccr4*, *pan2*, *pan3*, or *pam* deadenylases; this result suggests that *EhCAF1* is the parasite's main poly(A) ribonuclease. Furthermore, the decapping machinery lacks *dcp1* and *xrn1* genes but contains DCP2 and XRN2 homologs, which are typical P-body markers. Interestingly, the *EhXRN2* and *EhDCP2* proteins involved in 5'-to-3' mRNA decay accumulated in discrete cytoplasmic foci in a manner similar to that seen in P-bodies in higher eukaryotes. This localization pattern was also observed for *EhCAF1* deadenylase and *EhAGO2-2*. Given that DCP2, XRN2, CAF1 and AGO2 are resident P-body proteins in human cells, our data suggest that the foci observed in *E. histolytica* correspond to P-bodies. However, the amoebic foci are smaller and less rounded than P-bodies and so we prefer to refer to them as P-body-like structures. Given the similarity between proteins involved in mRNA decay in the human, in yeast and in *E. histolytica* and the colocalization of *EhCAF1* with *EhXRN2* and *EhAGO2-2* in cytoplasmic structures, our results also suggest that mRNA degradation occurs at these foci. However, this hypothesis requires formal confirmation through additional experiments designed to evidence the degradation of RNA substrates and the localization of other mRNA decay proteins in these P-body-like structures.

Cytoplasmic foci resembling P-bodies have previously been described in the protozoan parasite *Trypanosoma*. Cassola *et al.* [30] reported that mRNA and proteins related to mRNA metabolism were concentrated in cytoplasmic mRNA granules in starved *T. cruzi* and *T. brucei* parasites and suggested that the said granules provided transient protection of transcription during periods of stress. However, these structures may not represent typical P-bodies. A complex containing CAF1, NOT1, NOT2, NOT5, DHH1 and a possible homolog of Caf130 was found in *T. brucei*, in which CAF1 deadenylation activity was shown to be essential for cell survival [31]. In particular, CAF1 protein appeared to be predominantly distributed in the cytoplasm and the characteristic, punctate staining indicative of association with RNA storage or degradation granules was not found. Holetz *et al.* [32] showed that the TcDhh1 helicase protein was present in polysome-independent complexes and in discrete cytoplasmic foci resembling P-bodies. The latter varied in number as a function of the cell's nutritional status and following cycloheximide and puromycin treatments. Despite these initial efforts, the existence and potential roles of P-bodies in mRNA degradation events in protozoan parasites have not been comprehensively addressed.

We cannot rule out the existence of other RNA-containing foci in *E. histolytica*, such as the stress granules described in eukaryotic cells. Stress granules are cytoplasmic aggregates of stalled 48S pre-initiation complexes. They include small ribosomal subunits, translation initiation factors (eIF2, eIF3, eIF4, eIF4G), PABP, p54/Rck and the TTP and CPEB RNA binding proteins that regulate mRNA translation and decay [33,34]. Although P-bodies and stress granules share some proteins (e.g. TTP, p54/Rck, Lsm1, and Xrn1), there are some notable differences; P-bodies contain components of the mRNA decay pathway, whereas stress granules contain components of the translation initiation machinery and other specific markers (including TIA1) [35,36]. The presence of stress granules in *E. histolytica* remains to be established;

studies with specific markers would be required to gain a comprehensive view of cytoplasmic structures containing RNA binding proteins in this parasite.

In mammalian cells, treatment with actinomycin D induces the loss of P bodies [12], suggesting that the latter disaggregate once mRNA transcripts are no longer available for degradation. In contrast, the fact that trypanosome mRNA granules persist after actinomycin D treatment suggests that transcript degradation does not take place in these structures [30]. P-body formation in human and yeast also depends on the presence of elevated levels of non-translating mRNAs, which indicates that mRNAs exit the translation cycle before they enter P-bodies [37,38]. The observed effects of transcriptional and translational inhibitors that are known to inhibit mRNA decay suggest that cytoplasmic foci represent sites in which mRNA decay occur. Our data revealed a similar situation in *E. histolytica* trophozoites, since the number of P-body-like structures number fell after treatment with actinomycin D and cycloheximide. It has been reported that the translation inhibitor puromycin induces an increase in the number and size of P-bodies by releasing ribosomes from mRNA. In *E. histolytica* trophozoites, the number of P-body-like structures increase slightly after puromycin treatment. Hence, our data indicate that P-body-like structures are dynamic and are associated with ongoing mRNA synthesis and translation in *E. histolytica*.

The co-immunolocalization of the deadenylase *EhCAF1* and the exoribonuclease *EhXRN2* in cytoplasmic foci in *E. histolytica* supports the hypothesis whereby P-body-like structures have an important role in mRNA decay. Furthermore, these results constitute the first evidence of a functional interaction between deadenylation and 5'-to-3' mRNA degradation. The interaction between actin and both *EhCAF1* and *EhXRN2* further suggests that the proteins in amoebic P-body-like structures are associated with the cytoskeleton, as is the case in the human [26]. Additional, ongoing experiments should help us to understand the implications of these interactions within mRNA degradation machineries.

A small proportion of polyadenylated RNAs was concentrated in cytoplasmic foci, along with *EhCAF1*. In addition, fluorescent siRNAs were rapidly localized with *EhAGO2-2* protein in cytoplasmic foci during RNA interference of *Ehpc4* gene; this observation implies that P-body-like structures may have an important role in siRNA-mediated gene silencing by facilitating the interaction between siRNAs and Argonaut proteins. It was recently reported that *EhAGO2-2* protein was localized in nuclei with siRNAs containing 5'-polyphosphate termini during transcriptional gene silencing in *E. histolytica* strain G3 [39]. Surprisingly, we did not detect a nuclear signal for *EhAGO2-2*. This may be due to the fact that depending on the siRNAs generated during RNA interference, Argonaut proteins can downregulate gene expression by transporting specific classes of small regulatory RNAs to distinct cellular compartments. During transcriptional silencing, *EhAGO2-2* specifically transports small regulatory RNAs to nuclei in which gene silencing is mediated by histone modifications [39]. Our results indicate that both *EhAGO2-2* and specific siRNAs are preferentially targeted to the cytoplasm during RNA interference.

Taken as a whole, our present results provide evidence to suggest that mRNA degradation proteins involved in deadenylation and decapping pathways are organized within cytoplasmic, P-body-like structures that can be observed under the microscope. Future work is needed to understand the composition of these structures and define the localization of other mRNA decay proteins from NMD and RNA interference machineries and thus provide a better understanding of mRNA decay in this protist. This research should also provide us with an opportunity

to answer intriguing, fundamental questions about the posttranscriptional regulation of gene expression in this deep-branching eukaryotic parasite.

Materials and Methods

Ethics statement

This study was carried out in strict accordance with the recommendations of the Guide for the Use of Laboratory Animals of the Investigation Center and Advanced Studies of the National Polytechnic Institute. The protocols and experiments were approved by the Institutional Animal Care of the Investigation Center and Advanced Studies of the National Polytechnic Institute. Animals were kept in environmentally controlled animal facilities at the Investigation Center and Advanced Studies of the National Polytechnic Institute. All surgery was performed under sodium pentobarbital anesthesia and efforts were always made to minimize suffering.

Bacterial strains

Escherichia coli strain BL21pLysS was used for generation of *EhDCP2*, *EhXRN2* and *EhAGO2-2* recombinant proteins. We used the RNase-III deficient *E. coli* strain HT115 (rnc14::VTn10) to generate long dsRNAs targeting the *Ehpc4* gene [40].

E. histolytica cultures

Trophozoites of HM1-IMSS strain were axenically cultured in TYI-S-33 medium [41]. Cells in the logarithmic growth phase were used in all experiments. Cell viability was monitored by microscopy using the trypan blue dye exclusion method.

In silico identification of mRNA degradation genes

Identification of mRNA decay machinery genes was performed by screening the *E. histolytica* genome sequence at AmoebaDB (<http://amoebadb.org/amoeba/>). Human mRNA degradation protein sequences were used as queries. The amino acids sequences of orthologous *E. histolytica* proteins were aligned with those of mRNA degradation proteins from a range of organisms in the UniProt Knowledgebase, using ClustalW software with gap penalties of 10. Structural domains and sequence patterns were predicted with Scan Prosite. Three-dimensional structures for *EhXRN2* and *EhDCP2* were predicted with the SWISS-MODEL program by using the *S. pombe* RAT1 crystal structure (PDB 3FQD) and the *S. cerevisiae* DCP2 crystal structure (PDB 2KQMB) as a template, respectively. Three-dimensional models were displayed and refined using the RCSB Protein Data Bank Protein Workshop 3.9 viewer.

Quantitative real-time PCR assays

Total RNA was obtained using Trizol (Invitrogen) from HM1-IMSS strain trophozoites grown in basal culture conditions and after the induction of cell stress by heat shock, DNA damage and nitric oxide treatment. cDNAs were synthesized using 1 µg total RNA, 100 ng oligo (dT)₁₈, 100 mM DTT, 10 mM dNTPs, 40 U RNase inhibitor (Promega) and 200 U of Superscript II reverse transcriptase (Invitrogen) in first-strand buffer at 42°C for 90 min. After primer-specific reverse transcription, quantitative real-time PCR was performed with the primers listed in Table S2. Three independent biological samples were analyzed.

Quantitative real-time RT-PCR was carried out using an ABI PRISM 7000 Sequence Detection System (Applied Biosystems). Reactions contained 500 nM of each primer, 1× PCR SYBR-Green Master Mix (Qiagen) and 1 µL of template cDNA. Primers used for amplification are listed in Table S2. After an initial

denaturation for 15 min at 95°C, the following amplification cycle (repeated 40 times) was applied: 15 s at 95°C, 15 s at 52°C and 15 s at 72°C. Next, a denaturation curve was established (55°C to 95°C, 2°C min⁻¹) to check that a unique amplicon was produced. For each triplicate, the mean of the relative concentrations obtained for the tested mRNA was divided by the mean of the corresponding values obtained for amplification of the endogenous normalizer ribosomal *L31* gene. This ratio represents the variation of the mRNA abundance in the sample condition, compared with control. Each PCR experiment was repeated three times and three independent biological samples were analyzed. The relative expression of the genes coding for mRNA degradation enzymes was calculated using the 2^(-ΔΔCt) method [42].

Expression and purification of recombinant proteins

A 351-base-pair internal region of the *EhXm2* gene (EHI_133330) was PCR-amplified from HM1-IMSS genomic DNA using *Ehxm2-S* (5'-CCGGATCCAGGGAACAACCT-CCTGTA-3') and *Ehxm2-AS* (5'-CCAAGCTTTAACAG-TAGTTGAAAAAACCAC-3') primers. The full-length 699 nt *Ehdcp2* gene (EHI_058810) was PCR-amplified from cDNA using *Ehdcp2-S* (5'-CCCGGATCCATGGATACTCATGTTCCA-3') and *Ehdcp2-AS* (5'-CCAAGCTTTCATCTGGACTTCTT-TATT-3') primers. *Bam*HI and *Hind*III restriction sites are underlined. Amplified PCR products were purified and cloned in-frame into the pRSETA expression vector (Invitrogen). Plasmid constructions were automatically sequenced in an ABI-PRISM 310 sequencer (Applied Biosystems). The *Ehago2-2* gene was subcloned into pRSET-A from the pKT-3M plasmid (generously provided by Dr. Upinder Singh, Stanford University, CA). Recombinant *EhXRN2*, *EhDCP2* and *EhAGO2-2* proteins were expressed as His6x-fusion polypeptides in *E. coli* strain BL21 (DE3) pLysS using 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) for 3 hours at 37°C. Cells were disrupted by sonication and soluble *EhXRN2*, *EhDCP2* and *EhAGO2-2* proteins were purified to near-homogeneity under native conditions with Ni-NTA sepharose chromatography (Qiagen). The identity and integrity of recombinant proteins were confirmed by Western blot assays using anti-His6x antibodies (1:10,000) and the ECL-Plus detection system (GE Healthcare).

Generation of antibodies and Western blot assays

After separation by SDS-PAGE, recombinant *EhXRN2*, *EhDCP2* and *EhAGO2-2* proteins were electro-eluted from the Coomassie Blue-stained gel and used to immunize pathogen-free BALB/c male mice (*EhXRN2*, *EhDCP2*) or a rabbit (*EhAGO2-2*) as follows. An initial dose of 20 µg (mice) or 150 µg (rabbit) of recombinant protein in complete Freund's adjuvant (Gibco) was injected subcutaneously inoculated into the animals. Next, three doses of 20 µg (mice) or 100 µg (rabbit) in incomplete Freund's adjuvant (Gibco) were injected via the same route at 21-day intervals. One week after the last immunization, mice were bled and antiserum was obtained. The specificity of each serum was confirmed by Western blot assays using recombinant polypeptides separated on 10% SDS-PAGE (30 µg/lane) and electrotransferred to nitrocellulose membranes. The membranes were then incubated with each immune serum (1:1000) in 2% non-fat dried milk, 0.05% Tween-20 in PBS pH 7.4 overnight at 4°C. Proteins were developed using peroxidase-conjugated secondary antibodies (1:5000, Zymed) and detected with the ECL-Plus system. Western blot assays using nuclear and cytoplasmic protein extracts from HM1-IMSS trophozoites were performed as described elsewhere [43].

Immunofluorescence confocal laser assays

Trophozoites were cultured on cover slides, fixed with 4% paraformaldehyde at 37°C for 40 min and permeabilized with 0.5% Triton-X100 in PBS 1× at 37°C for 40 min. Next, cells were blocked in 2% BSA for 60 min at 37°C, washed three times with PBS 1× and incubated with either anti-*Eh*CAF1 and anti-*Eh*AGO2-2 antibodies generated in the rabbit or anti-*Eh*XRN2, and anti-*Eh*DCP2 antibodies generated in the mouse (1:200) at 37°C for 1 h. Cells were washed three times and incubated with fluorescein-labeled secondary antibodies (1:200) at 37°C for 1 h. Lastly, DNA was counterstained with 4'-6-diamidino-2-phenylindole (DAPI, 5 mg/ml) for 1 min. Samples were analyzed under an inverted microscope connected to a laser confocal scanning system (TCS SP2, Leica Microsystems).

Immunoprecipitation assays

The immunoprecipitation protocol was essentially that described in [44] with minimal modifications. Briefly, cytoplasmic proteins (50 µg) were diluted in 300 µl HNTG buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.1% Triton, 10% glycerol) and incubated with 10 µl of protein G beads (Sigma) for 30 min at 4°C. Non-specifically interacting proteins were excluded by centrifugation at 3000 rpm at 4°C for 5 min. The supernatant was incubated with either anti-*Eh*CAF1 antibodies (1:100) or anti-*Eh*XRN2 antibodies (1:100) for 2 h at 4°C. Next, protein G-Sepharose beads were added to the samples and incubated for 16 h at 4°C. Immunoprecipitated proteins were collected by centrifugation, washed three times with HNTG buffer and then separated using 12% SDS-PAGE. Proteins were transferred to nitrocellulose membranes and detected by Western blot assays with rabbit anti-*Eh*CAF1 or mouse anti-*Eh*XRN2 antibodies.

Bacterial expression of double-stranded RNA and trophozoites transfection

Double-stranded RNA was obtained as described in [28]. Briefly, the *Eh*hpc4 full coding sequence (EHL_192520) was cloned into the *Sma*I and *Xho*I restriction sites of the pL4440 plasmid vector, which is flanked by T7 promoters [45]. Competent *E. coli* HT115 cells were transformed with pL4440-*Eh*hpc4 and spread onto LB-agar plates containing ampicillin and tetracycline. Isolated bacterial colonies were obtained after overnight culture at 37°C. Double-stranded RNA synthesis was induced by incubation with 2 mM IPTG for 4 h at 37°C. Cy3-labeling of purified dsRNA was performed with the Silencer siRNA labeling kit (Ambion), according to the manufacturer's instructions. The reaction mixture (containing 5 µg dsRNA, nuclease free-water, labeling buffer and Cy3 labeling reagent) was incubated for 1 h at 37°C in the dark. Labeled dsRNA was precipitated with ethanol, resuspended in DEPC-treated water and immediately used in colocalization assays.

RNA fluorescence *in situ* hybridization

To detect poly(A)⁺ RNA in fluorescence *in situ* hybridization (FISH) assays, parasites were harvested and allowed to adhere to microscope slides. Cells were fixed with 4% paraformaldehyde in PBS, denatured in 70% formamide at 70°C and incubated for 10 min with 25 mM NH₄Cl. The cells were permeabilized and

blocked for 30 min in 0.5% Triton X-100/2% BSA in PBS, followed by 2 h of prehybridization at room temperature in hybridization buffer (2% BSA, 4× SSC, 5% dextran sulfate, 35% deionized formamide and 10 U RNase inhibitor). Hybridization was performed overnight at room temperature in a humidified culture chamber in the presence of 300 ng/ml FITC-conjugated oligo-(dT)₃₀ in hybridization buffer. Slides were washed once in 4× SSC plus 35% deionized formamide and once in 4× SSC and were then counterstained with DAPI (1 mg/ml). Slides were mounted with Vectashield antifade (Vector). The subcellular localization was analyzed with a Leica TCS SP2 microscope. Merged images were obtained by superimposing image files using LCS Lite software (Leica Microsystems).

Supporting Information

Figure S1 XRN2 related proteins. (A) Multiple alignments of related XRN2 proteins. (B) Phylogenetic relationships between *Eh*XRN2 and XRN2 family members. The rooted tree was created with a neighbor-joining algorithm in the Mega 5.1 program, on the basis of ClustalW alignments of full amino acid sequences. Each protein's accession number is given in brackets. (PDF)

Figure S2 DCP2 related proteins. (A) Multiple alignments of related DCP2 proteins. (B) Phylogenetic relationships between *Eh*DCP2 and DCP2 family members. The rooted tree was created with a neighbor-joining algorithm in the Mega 5.1 program, on the basis of ClustalW alignments of full amino acid sequences. Each protein's accession number is given in brackets. (PDF)

Table S1 Predicted mRNA degradation proteins in *Entamoeba histolytica*. (PDF)

Table S2 Primers used in quantitative real-time PCR assays. (PDF)

Table S3 Comparison of *Eh*XRN2 with related proteins from several organisms. (PDF)

Table S4 Comparison of *Eh*DCP2 with related proteins from several organisms. (PDF)

Acknowledgments

We thank the Autonomous University of Mexico City for support and Dr. Upinder Singh (Stanford University, CA) for kindly providing the *Eh*ago2-2 gene clone.

Author Contributions

Conceived and designed the experiments: ILR EO LAM NG CPP CLC. Performed the experiments: ILR GGR CW ECT OHC. Analyzed the data: ILR LAM NG CPP CLC. Contributed reagents/materials/analysis tools: ILR EO GGR NG CW CPP CLC. Wrote the paper: ILR LAM CLC.

References

- Bevilacqua A, Ceriani MC, Capaccioli S, Nicolini A (2003) Post-transcriptional regulation of gene expression by degradation of messenger RNAs. *J Cell Physiol* 195:356–372.
- Chen CY, Shyu AB (2011) Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip Rev RNA* 2:167–183.
- Coller J, Parker R (2004) Eukaryotic mRNA decapping. *Annual Review of Biochemistry* Vol. 73: 861–890.
- Lykke-Andersen S, Tomecki R, Jensen TH, Dziembowski A (2011) The eukaryotic RNA exosome: same scaffold but variable catalytic subunits. *RNA Biol* 8:61–66.

5. Kulkarni M, Ozgur S, Stoecklin G (2010) On track with P-bodies. *Biochem Soc Trans* 38:242–251.
6. Aizer A, Shav-Tal Y (2008) Intracellular trafficking and dynamics of P bodies. *Prion* 2:131–134.
7. Eulalio A, Behm-Ansmant I, Izaurralde E (2007) P-bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* 8:9–22.
8. Sen GL, Blau HM (2005) Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nat Cell Biol* 7:633–636.
9. Liu J, Rivas FV, Wohlschlegel J, Yates JR, Parker R, et al. (2005) A role for the P-body component GW182 in microRNA function. *Nat Cell Biol* 7:1261–1266.
10. Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E (2007) P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. *Mol Cell Biol* 27: 3970–3981.
11. Sheth U, Parker R (2006) Targeting of aberrant mRNAs to cytoplasmic processing bodies. *Cell* 125: 1095–1109.
12. Cougot N, Babajko S, Séraphin B (2004) Cytoplasmic foci are sites of mRNA decay in human cells. *J Cell Biol* 165: 31–40.
13. Stalder L, Mühlemann O (2009) Processing bodies are not required for mammalian nonsense-mediated mRNA decay. *RNA* 15:1265–1273.
14. Chu CY, Rana TM (2006) Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol* 4(7):e210.
15. Stanley SL (2003) Amoebiasis. *Lancet* 361:1025–1034.
16. Abed M, Ankri S (2005) Molecular characterization of *Entamoeba histolytica* RNase III and AGO2, two RNA interference hallmark proteins. *Exp Parasitol* 110:265–269.
17. Zhang H, Ehrenkaufer GM, Pompey JM, Hackney JA, Singh U (2008) Small RNAs with 5'-polyphosphate termini associate with a piwi-related protein and regulate gene expression in the single-celled eukaryote *Entamoeba histolytica*. *PLoS Pathog* 4:e1000219.
18. Zhang H, Pompey JM, Singh U (2011) RNA interference in *Entamoeba histolytica*: implications for parasite biology and gene silencing. *Future Microbiol* 6:1103–117.
19. Weber C, Guigon G, Bouchier C, Frangeul L, Moreira S, et al. (2006) Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in *Entamoeba histolytica*. *Eukaryot Cell* 5:871–875.
20. López-Casamichana M, Orozco E, Marchat LA, López-Camarillo C (2008) Transcriptional profile of the homologous recombination machinery and characterization of the EhRAD51 recombinase in response to DNA damage in *Entamoeba histolytica*. *BMC Mol Biol* 9:35.
21. Ramos E, Olivos-García A, Nequiz M, Saavedra E, Tello E, et al. (2007) *Entamoeba histolytica*: apoptosis induced in vitro by nitric oxide species. *Exp Parasitol* 116:257–265.
22. Parker R, Sheth U (2007) P bodies and the control of mRNA translation and degradation. *Mol Cell* 25:635–646.
23. Xiang S, Cooper-Morgan A, Jiao X, Kiledjian M, Manley JL, et al. (2009) Structure and function of the 5'→3' exoribonuclease Rat1 and its activating partner Rai1. *Nature* 458:784–788.
24. She M, Decker CJ, Svergun DI, Round A, Chen N, et al. (2008) Structural basis of dcp2 recognition and activation by dcp1. *Mol Cell* 29:337–49.
25. López-Camarillo C, de la Luz García-Hernández M, Marchat LA, Luna-Arias JP, Hernández de la Cruz O, et al. (2008) *Entamoeba histolytica* EhDEAD1 is a conserved DEAD-box RNA helicase with ATPase and ATP-dependent RNA unwinding activities. *Gene* 414:19–31.
26. Aizer A, Brody Y, Ler LW, Sonenberg N, Singer RH, et al. (2008) The dynamics of mammalian P body transport, assembly, and disassembly in vivo. *Mol Biol Cell* 19:4154–4166.
27. López-Camarillo C, Luna-Arias JP, Marchat LA, Orozco E (2003) EhPgp5 mRNA stability is a regulatory event in the *Entamoeba histolytica* multidrug resistance phenotype. *J Biol Chem* 278:11273–11280.
28. Solis CF, Santi-Rocca J, Perdomo D, Weber C, Guillén N (2009) Use of bacterially expressed dsRNA to downregulate *Entamoeba histolytica* gene expression. *PLoS One* 4:e8424.
29. Jagannath A, Wood MJ (2009) Localization of double-stranded small interfering RNA to cytoplasmic processing bodies is Ago2 dependent and results in up-regulation of GW182 and Argonaute-2. *Mol Biol Cell* 20:521–9.
30. Cassola A, De Gaudenzi JG, Frasch AC (2007) Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. *Mol Microbiol* 65:655–670.
31. Schwede A, Ellis L, Luther J, Carrington M, Stoecklin G, et al. (2008) A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Res* 36:3374–3388.
32. Hotez FB, Correa A, Avila AR, Nakamura CV, Krieger MA, et al. (2007) Evidence of P-body-like structures in *Trypanosoma cruzi*. *Biochem Biophys Res Commun* 356:1062–1067.
33. Kedersha N, Anderson P (2002) Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem Soc Trans* 30:963–969.
34. Wilczynska A, Aigueperse C, Kress M, Dautry F, Weil D (2005) The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. *J Cell Sci* 118:981–992.
35. Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, et al. (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol* 169:871–884.
36. Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, et al. (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15:5383–5398.
37. Brengues M, Teixeira D, Parker R (2005) Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* 310:486–489.
38. Teixeira D, Sheth U, Valencia-Sanchez MA, Brengues M, Parker R (2005) Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* 11: 371–382.
39. Zhang H, Alramini H, Tran V, Singh U (2011) Nucleus-localized antisense small RNAs with 5'-polyphosphate termini regulate long term transcriptional gene silencing in *Entamoeba histolytica* G3 strain. *J Biol Chem* 286:44467–44479.
40. Takiff HE, Chen SM, Court DL (1989) Genetic analysis of the rnc operon of *Escherichia coli*. *J Bacteriol* 171:2581–2590.
41. Diamond LS, Harlow DR, Cunnick CC (1978) A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. *Trans R Soc Trop Med Hyg* 72:431–432.
42. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} Method. *Methods* 25:402–408.
43. Marchat LA, Gómez C, Pérez DG, Paz F, Mendoza L, et al. (2002) Two CCAAT/enhancer binding protein sites are cis-activator elements of the *Entamoeba histolytica* EhPgp1 (mdr-like) gene expression. *Cell Microbiol* 4:725–737.
44. Tovy A, Siman Tov R, Gaentzsch R, Helm M, Ankri S (2010) A New Nuclear Function of the *Entamoeba histolytica* Glycolytic Enzyme Enolase: The Metabolic Regulation of Cytosine-5 Methyltransferase 2 (Dnmt2) Activity. *PLoS Pathog* 6:e1000775.
45. Timmons L, Court DL, Fire A (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* 263:103–112.

EhCAF1, a functional deadenylase, colocalizes with LPS-endoribonuclease and Rrp41 mRNA degradation proteins in *Entamoeba histolytica*

Itzel López-Rosas¹, Beatriz Gallo Olvera², Laurence A. Marchat², Nancy Guillen^{3,4}, Christian Weber, Esther Orozco⁵, Guillermo Mendoza-Hernández⁶, Mavil López-Casamichana¹, Olga Hernández de la Cruz¹, César López-Camarillo^{1,*}

¹Autonomous University of Mexico City, Genomics Sciences Program. México DF.

²Institutional Program of Molecular Biomedicine and Biotechnology Program, National School of Medicine and Homeopathy, National Polytechnic Institute, México DF. México. ³United´ Biologie Cellulaire du Parasitisme, Institut Pasteur, Paris, France, ⁴INSERM U786, Paris, France. ⁵Departament of Infectomics y Molecular Pathogenesis, Investigation Center and Advanced Studies, National Polytechnic Institute, México DF. ⁶Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, México.

***Corresponding author:** Tel: (+52) 5488-6661 ext 15307.

Email: cesar.lopez@uacm.edu.mx (Dr. César López-Camarillo).

Keywords: *Entamoeba histolytica*, deadenylation, mRNA decay, EhCAF1 deadenylase, EhXRN2 exoribonuclease, P-bodies.

Running title: EhCAF1 deadenylase in *Entamoeba histolytica*

ABSTRACT (200)

Messenger RNA degradation is a crucial post-transcriptional event for gene expression regulation in eukaryotic cells. The degradation of transcripts initiates by poly(A) tail shortening which is performed by CAF1, CCR4, PAN2, PAN3 and PARN deadenylases. Remarkably, mRNA degradation occurs in cytoplasmic structures denoted as P-bodies, however, these mechanisms remain poorly understood in protozoan parasites. Here, we reported the characterization of the EhCAF1 protein in *Entamoeba histolytica*, the protozoan parasite responsible for human amoebiasis. Using biochemical analysis, we evidencing that EhCAF1 have RNA binding and deadenylase activities *in vitro*. After cellular stress induction using heath shock, DNA damage and nitrosative treatments, we found that EhCAF1 mRNA expression did not changed. However, cellular localization studies using immunofluoresce assays and confocal microscopy demonstrate that EhCAF1 was rapidly distributed and relocalized with EhXRN2 ribonuclease in P-bodies structures. In addition, using pull-down experiments and ESI-MS/MS tandem mass spectrometry, we identified five *EhCAF1*-interacting proteins. Interestingly, one protein corresponds to the endoribonuclease L-PSP. Moreover, we found that endoribonuclease L-PSP interacts with the *EhRrp41* 3'-to- 5' exoribonuclease, linking deadenylation to the core exosome machinery. Our data suggests the existence of a complex containing the EhCAF1 deadenylase, a putative L-PSP endoribonuclease and the *EhRrp41* exosome exoribonuclease. These results provide novel data in mRNA degradation mechanisms in the early branching *E. histolytica*.

Introduction

Polyadenylated mRNA turnover is a critical determinant in eukaryotic gene expression. It results from both synthesis and degradation rates, which are continuously adjusted to cell physiological requirements. In cytoplasm, mRNA degradation is initiated by poly(A) tail shortening by deadenylases, followed by 5'-end cap removal by DCP1 and DCP2 enzymes [1]. Then, mRNA is submitted to 5'-3' exonucleolytic digestion by XRN1. Alternatively, deadenylated mRNA can be degraded by exosome [2]. Alterations of any regulatory elements participating in these reactions can modify both mRNA stability and deadenylation rate, indicating that poly(A) tail shortening is a rate-limiting step in mRNA decay [3,4].

Deadenylation is performed by multiple poly(A) ribonucleases. CCR4 and POP2 proteins are the main deadenylases in *Saccharomyces cerevisiae* [5,6]. Higher eukaryotic cells possess one homologue of yeast CCR4, whereas two POP2 homologues (CAF1 and CALIF) have been identified in human [7-9]. PAN2/PAN3 and PARN are additional poly(A) ribonucleases involved in specific deadenylation events [10,11]. Interestingly, CAF1 and CCR4 have been involved in other cellular processes. In yeast, their association with five NOT proteins, CAF40 and CAF130 polypeptides, forms the CCR4-NOT complex, which interacts with TBP and TFIID factors, establishing a functional link between mRNA synthesis and degradation [12]. CAF1 also participates in cell cycle regulation interacting with BTG1/2 and ANA/BTG3 antiproliferative proteins [13,14]. mRNA deadenylase activity of CCR4-NOT complex contributes to DNA damage response in yeast, interacting with DUN1, MRC1, RAD9, and RAD17 checkpoint factors [15]. CAF1 also regulates mRNA poly(A) tail length of Crt1 transcription factor, which controls several DNA damage-induced genes [16]. In addition, PUF proteins that bind 3' untranslated region (UTR) triggering mRNA decay or translational repression, physically binds POP2 and stimulates recruitment of CCR4 and two other enzymes involved in mRNA regulation, Dcp1p and Dhh1p [17,18].

In *Entamoeba histolytica*, the protozoan parasite responsible for human amoebiasis that affects 50 million people worldwide [19], posttranscriptional gene expression regulation and mRNA degradation mechanisms are poorly understood. Recently, we have characterized the P-bodies like structures in *E. histolytica*, and we have detected that several proteins involved in mRNA degradation are enriched in these cytoplasmic foci (In press). Because of the relevance of poly(A) tail removal in mRNA decay, it is of prime interest to characterize the enzymes involved in deadenylation as they are likely to be target of regulators affecting mRNA stability. Here, we report the identification of the poly(A) ribonuclease genes in *E. histolytica* and the functional characterization of EhCAF1 deadenylase.

Results

E. histolytica has small deadenylation machinery

In higher eukaryotic cells CCR4 and CAF1 (POP2 in yeast) deadenylases associate with five NOT transcriptional factors in CCR4/CAF1/NOT complex, which perform the poly(A) tail removal from mRNA messenger (mRNA). In addition deadenylation it is performed in less extend by PARN, PAN2, and PAN3 enzymes. *In silico* screening of *E. histolytica* genome databases revealed the presence of two intron-less genes

encoding for CAF1 and CAF1-like proteins, but intriguingly no homologues for CCR4, PARN, PAN2 or PAN3 were identified (Table 1). *Ehcaf1* and *Ehcaf1-like* genes encode for CAF1 and CAF1-like proteins, which share 83% identity (Table S1). Both proteins present lowest e-values ($2e-39$ to $4e-61$), and 32-42% identity and 54-66% similitude with homologous eukaryotic CAF1 proteins ranging from mammals, plants and protozoan parasites (Table S1). We also found five NOT proteins in *E. histolytica* genome suggesting the existence of a CAF1/NOT complex lacking CCR4 subunit (Fig. 1). These data evidencing that amoeba has a reduced number of genes codified for minimal deadenylation machinery.

E. histolytica deadenylase genes are differentially expressed in response to cellular stress.

Deadenylation is a key mechanism to modulate mRNA turnover and gene expression after cellular stress in eukaryotic cells. By semiquantitative RT-PCR assays, we analyzed *E. histolytica* poly(A) ribonucleases mRNA expression in response to stress represented by heat shock and UV-induced DNA damage. Results showed that *Ehcaf1* transcripts level remained unchanged after UV-C irradiation, whereas *Ehcaf1-like* mRNA expression was five-fold repressed at 0.5 and 3 h after UV-C irradiation (Fig. 1A and C). Heat shock did not induce any change in *Ehcaf1* transcription, whereas *Ehcaf1-like* mRNA expression was reduced about five-fold (Fig. 1B and D). *Actin* mRNA expression appeared unchanged in both experimental conditions.

EhCAF1 conserves the characteristic features of ribonuclease D family proteins.

Multiple alignment of POP2/CAF1 proteins showed that EhCAF1 shares well-conserved motifs and critical functional amino acids residues. EhCAF1 contain the ribonuclease D family signature represented by CAF1 nuclease domain (PF004857) and the conserved DEDD residues (D₈₄E₈₆D₂₀₆D₂₇₆) involved in 3'-to- 5' exonuclease activity in yeast POP2 and human CAF1 proteins [6]. A less conserved exonuclease domain (PF00929), which is present in some H-like ribonucleases [27], overlaps with ribonuclease D domain (Fig. 3A) in EhCAF1 amino acids sequence. Predicted EhCAF1 tertiary structure has the characteristic folding of yeast POP2 crystal [27], with nine- α helix bundle (α 1 to α 9) and five central β -sheet (β 1 to β 5). The conserved DEDD residues that are dispersed in CAF1 domain appear clustered to the central region of EhCAF1 (Fig. 3B). PSI-BLAST study evidenced a clear relation between homologous proteins from human, yeast and protozoan parasites, including *Trypanosoma gondii*, *Trichomonas vaginalis*, *Plasmodium falciparum*, *Giardia lamblia*, and *E. histolytica*. EhCAF1 is highly related to human CAF1 and other mammalian CAF1 proteins. EhCAF1 is related in less extent to the CAF1 from free-living protozoa *Tetrahymena thermophila* and is more divergent from plant and fungi homologues (Fig. 3C). Phylogenetic inference of CAF1 homologous showed that protozoan CAF1 proteins are distributed in a clade at the bottom of the tree. (Fig. 3D).

EhCAF1 has RNA binding and deadenylase activities

Binding to RNA substrates is the first step required for deadenylase activity. However, EhCAF1 has not an apparent RNA binding domain, thus in order to define if it has RNA binding activity we performed RNA electrophoretic mobility shift assays (REMSA) using the [α -³²P] UTP-labeled *EhPgp5* 3'UTR as probe (Lopez Camarillo *et al.*, 2008). Results evidencing that purified recombinant EhCAF1 binds radioactive

RNA probe in a specific manner (Fig. 5A). RNA-protein complexes disappeared when 350 molar excess of specific competitor was added to reaction, whereas complexes remain unchanged in the presence of unspecific competitor (Fig. 5A). Control assays showed that purified GST has not interaction with RNA probe. To investigate the participation of EhCAF1 protein in degradation of poly(A) tailed RNA, we examined its deadenylase activity using as substrate a polyadenylated RNA (A₃₀) generated by in vitro transcription. The results in Figure 5B show a progressive deadenylation of the A₃₀ polyadenylated RNA substrate, resulting in a fully deadenylated product after 60 minutes (lane 4). Non-adenylated control transcript was not affected during reaction (lane 5). These data showed that EhCAF1 is a functional deadenylase.

EhCAF1 is a cytoplasmic protein

Ehcaf1 gene was cloned in pRSETA vector to produce the recombinant EhCAF1 protein (rEhCAF1). In Coomassie blue stained gels rEhCAF1 appeared as a 41 kDa band that was recognized by anti-6x-His-tag antibodies in Western blot assays (Fig. 3E and F). Then purified rEhCAF1 was used as immunogen to obtain rabbit anti-EhCAF1 serum. In Western blot assays, EhCAF1 antibodies recognized the purified 41 kDa rEhCAF1 (Fig. 2G) and in cytoplasmic (CE) and nuclear extracts (NE), we detected a 37 kDa band in CE which corresponds to predicted molecular weight for endogenous EhCAF1, but no signal was detected in NE. EhPC4 antibodies (Lopez Rosas *et al.*, 2012) used as control identified EhPC4 in NE as expected indicating the absence of cross contamination of cellular fractions (Fig. 2H).

Western blot assays showed that antibodies recognized in cytoplasmic fraction two single 103 kDa and 27.5 kDa bands corresponding to the predicted molecular weight for *EhXRN2* and *EhDCP2*, respectively (Figure 4). In contrast, for both proteins we did not observed signal in nuclear fraction. As control, we used preimmune serum which did not give signal (data not shown), and antibodies that recognized *EhPC4* transcription factor predominantly in nuclei as reported [30], evidencing the absence of cross contamination in cellular fractions.

Immunofluorescence and confocal microscopy assays confirmed the localization of EhCAF1 in cytoplasm of trophozoites. In most cells, EhCAF1 signal appeared homogeneously in cytoplasm as a particular speckled pattern in discrete regions resembling P-bodies (Fig. 3D and H).

EhCAF1 is relocated with EhXRN1 after cellular stress in cytoplasmic P-bodies

We next analyzed the cellular localization of EhCAF1. We used EhXRN2 antibodies as a marker for P-bodies (Lopez Rosas *et al.*, 2012). As expected EhCAF1 and EhXRN2 colocalized in cytoplasmic foci

EhCAF1 interacts with the putative endoribonuclease L-LPS

In human CAF1 associates with CCR4 deadenylase and with NOT transcription factors in CCR4/CAF1/NOT degradation complex. We performed His tag protein interaction pull-down experiments to capture and purify proteins that interact with His-EhCAF1 used as bait and total protein extracts as prey. EhCAF1-interacting proteins were identified by ESI-MS/MS tandem mass spectrometry. Results showed that EhCAF1 interacts with some proteins representing by peroxiredoxin, profiling, actin and glyceraldehyde-3-phosphate dehydrogenase (Fig. 6A). Interestingly, we also identified the endoribonuclease L-PSP (EH_087570) as an EhCAF1 associated

protein (Fig. 6B), which seems to be product of a lateral gene transfer event. Phylogenetic analysis evidenced that nearby homologous proteins for endoribonuclease L-PSP have a bacterial origin (Fig. S2).

In silico and in vitro interaction of endoribonuclease L-PSP with the EhRrp41 3'-to-5' exoribonuclease

In silico predictions of endoribonuclease L-PSP interacting proteins using STRING algorithm showed that it potentially associated to exosome ribonucleases RRP4, RRP10, RRP40, RRP41, RRP42, RRP44, and RRP54 suggesting a link between CAF1 deadenylation machinery and exosome mRNA degradation complex. However, we need further experimentation to corroborate these assumptions.

4. Discussion

Poly(A) tail removal is the first step for eukaryotic transcripts degradation. Here, we demonstrated that the protozoan parasite *E. histolytica* has only two poly(A) ribonucleases, EhCAF1 and EhCAF1-like. Particularly, EhCAF1 is a well conserved orthologue of CAF1/POP2 deadenylases previously described, suggesting that it could be participating in poly(A) tail degradation in *E. histolytica*. The presence of two CAF1 and the absence of CCR4 in *E. histolytica*, confirmed that POP2/CAF1 proteins are the most relevant enzymes for deadenylation in animal cells [28-31], whereas CCR4 is the main deadenylase in yeast [9]. CCR4 is absent in many animal species including other protozoan parasites [29]. The PAN2/3 complex appears to have been lost many times in eukaryotic evolution; it is absent in *E. histolytica* and *Plasmodium*, while it is present in *Trypanosoma* [32].

Ehcaf1 and *Ehcaf1*-like genes are transcribed in basal culture conditions, suggesting a relevant role in gene expression regulation in *E. histolytica*. Additionally, EhCAF1 was immunodetected in cytoplasm, which is the cellular compartment where deadenylation reaction takes place. However, we cannot discard that anti-EhCAF1 antibodies also detect EhCAF1-like.

Ehcaf1 gene transcription appears unchanged after UV-C irradiation and heat shock, whereas *Ehcaf1*-like gene is inhibited, indicating that both deadenylases could target different genes. Particularly, the inhibition of *Ehcaf1*-like gene transcription could induce a reduced EhCAF1-like deadenylase activity, which could inhibit deadenylation of stress response genes mRNAs, increasing their stability and translational efficiency, so that trophozoites could rapidly produce required proteins for stress response. A rapid and strong inhibition of mRNA deadenylation and degradation has been observed after UV-B irradiation in HeLA cells [33]. In *Drosophila*, Mammals and yeast, inhibition of *hsp70* mRNA deadenylation has been associated to increased mRNA stability and translation during heat shock [34-36]. In *E. histolytica*, we have recently described that genes encoding DNA repair factors and Fe-S clusters containing proteins were up-regulated by UV-C irradiation [25,37,38]. In addition, heat shock induced transcriptional activation of genes coding for heat shock proteins and other factors [26]. Notably, overexpression of EhMLBP plays a protective role against heat shock [39]. It is possible that EhCAF1-like could be targeting mRNAs from genes that are directly or indirectly involved in UV-C induced DNA damage and heat shock response in *E. histolytica*, including those described above, to allow the rapid synthesis of proteins that are required for cell survival. Other molecular events, including inactivation of ARE-binding proteins

promoting deadenylation and inhibition of translation, might also be inhibited by stress, contributing to a reduced deadenylation [40].

In trophozoites, EhCAF1 was observed in small cytoplasmic dots, colocalizing with poly(A) RNA, which is consistent with its possible role in mRNA deadenylation [1]. In eukaryotic cells, mRNA degradation takes place in small cytoplasmic structures called P-bodies that are also involved in mRNA storage, mRNA surveillance and RNA-based gene silencing [41]. Colocalization of EhCAF1 and poly(A) RNA in cytoplasmic structures raised the possibility that these structures may correspond to P-bodies in *E. histolytica*. However, additional experiments currently in progress should help to confirm this hypothesis.

Experimental procedures

In silico identification of poly(A) ribonucleases

Screening of *E. histolytica* genome sequence at Amoeba DB database (<http://amoebadb.org/amoeba/>) was performed by BLAST using yeast and human CCR4, CAF1, POP2, PAN2, PAN3 and PARN protein sequences as queries. Putative *E. histolytica* orthologous proteins were confronted by BLAST against all databases using the Blosum 62 matrix. POP2/CAF1 amino acids (aa) sequences from diverse organisms (<http://www.pir.uniprot.org/>) were aligned by ClustalW (<http://www.ch.embnet.org/software/ClustalW.html/>). Pfam (<http://pfam.sanger.ac.uk/>) identified structural domains and sequence patterns. Structural and evolutionary relationships were analyzed through PSI-BLAST [22]. Phylogenetic analysis was done by the Neighbor-joining distance method; robustness of phylogenetic inference was tested by bootstrapping method, involving 1000 replications of data based on criteria of 50% majority-rule consensus using the MEGA software [23] (<http://www.megasoftware.net/>). Three-dimensional structure was predicted by Swiss model (<http://swissmodel.expasy.org//SWISS-MODEL.html>) using yeast POP2 crystal structure (PDB entry 1UOC; <http://www.rcsb.org/pdb/home/home.do>) as template.

E. histolytica cultures

Trophozoites (strain HM1-IMSS) were axenically cultured in TYI-S-33 medium [24]. For DNA damage studies, trophozoites (1.5×10^6) were irradiated with 254 nm UV-C light at 150 J m^{-2} for 8 s using a UV-Stratalinker 1800 device (Stratagene). Then, cells were incubated in fresh medium at 37°C for 0.5 and 3 h [25]. Heat shock was done by incubating trophozoites at 42°C for 3 h [26]. Cells in logarithmic phase of growth were used in all experiments.

Semi-quantitative RT-PCR assays

Total RNA from trophozoites harvested before or after UV-C irradiation or heat shock was obtained using Trizol reagent (Gibco). cDNA was synthesized using $1 \mu\text{g}$ total RNA, 100 ng oligo-dT₍₁₈₎, 10 mM dNTPs, 40 U SUPERase-in (Ambion) and 200 U Superscript II reverse transcriptase (Invitrogen) at 42°C for 90 min. Deadenylases genes amplification was done using *Ehcaf1b*-S (5'-CAA TAA TGC CAC CCA AT -3') and *Ehcaf1b*-AS (5' ACT GTT GAT GGA TTT ATT AAA-3'), *Ehcaf1-like*-S (5'-GATGTTAGTTTCCCAA-3') and *Ehcaf1-like*-AS (5'-TTTTGTAAAGAAATCATCAATAGAA-3') specific primers. *Actin* sense (5'-AGCTGTTCTTTCATTATATGC-3') and antisense (5'-

TTCTCTTTCAGCAGTAGTGGT-3') primers were used as control. Amplified products were separated by 6% PAGE, stained with ethidium bromide and submitted to densitometric analysis in a Fujifilm FLA-5100 apparatus using the MultiGauge software. Three independent experiments were done by duplicate. Pixels corresponding to *actin* were taken as 100% in each lane and used to normalize densitometric data.

Cloning of EhCaf1 gene

The full-length *EhCaf1* gene was PCR amplified using genomic DNA and *Ehcaf1*-S (5'-CCGGATCCTTTATGAATATTAACGCCTCA-3') and *Ehcaf1*-AS (5'-CCAAGCTTAAATTTCAATACAAGGCTAAT-3') primers (*Bam*HI and *Hind*III restriction sites are underlined). PCR was performed at 94°C for 5 min, 94°C for 30 s, 48°C for 30 s and 72°C for 60 s, for 30 cycles, with a final extension step at 72°C for 7 min using Taq DNA polymerase high fidelity (2 U). Amplified product was purified and cloned into pRSET-A expression vector (Invitrogen). The obtained pRSET-*Ehcaf1* plasmid was automatically sequenced in an ABI-PRISM 310 sequencer (Applied Biosystems).

Expression and purification of recombinant EhCAF1

The recombinant EhCAF1 protein (rEhCAF1) was expressed in *Escherichia coli* BL21 (DE3) pLysS strain as a 6x-His-tagged fusion protein. After induction with 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) for 3 h at 37°C, cells were disrupted by sonication at 4°C. Soluble rEhCAF1 was purified near to homogeneity under native conditions through Ni-NTA affinity chromatography according to manufacturer recommendations (Qiagen). Identity of purified rEhCAF1 was confirmed by Western blot assays using anti-6x-His-tag antibodies (Roche, 1:5000) and detected by ECL-Plus system (Amersham).

Generation of anti-EhCAF1 antibodies and Western blot assays

Purified rEhCAF1 was submitted to preparative SDS-PAGE, electroeluted from Coomassie stained gel and used to immunize a pathogen free New Zealand male rabbit. An initial dose of 200 µg in complete Freund's adjuvant (Sigma) was subcutaneously inoculated, followed by three 100 µg doses in incomplete Freund's adjuvant each 15 days by the same via. One week later, rabbit was bled, IgGs were purified through protein G sepharose chromatography and tested for reactivity against rEhCAF1 in Western blot assays. For immunolocalization of EhCAF1, trophozoites nuclear (NE) and cytoplasmic (CE) extracts [20] were separated in 10% SDS-PAGE (20 µg/lane), electrotransferred to nitrocellulose filters and incubated with immune sera (1:1000) in 5% nonfat dry milk and 0.05% Tween-20 in PBS pH 7.4 overnight at 4°C. Proteins were developed by peroxidase-conjugated secondary antibodies (1:1500) (Zymed) and detected by ECL-Plus system (Amersham).

Immunofluorescence and laser confocal microscopy assays

Trophozoites cultured on cover slides were fixed with 4% paraformaldehyde at 37°C for 40 min and permeabilized with 0.5% Triton-X100 in PBS at 37°C for 40 min. Then, cells were blocked with 2% BSA in PBS for 60 min at 37°C, washed with 0.2% Triton-X100 in PBS and incubated with anti-EhCAF1 antibodies (1:200) at 4°C overnight. Cells were washed and incubated with fluorescein-labeled secondary antibodies (1:200) (Jackson) at 37°C for 1 h. Finally, DNA was counterstained with 4',

6-diamidino-2-phenylindole (DAPI) (5 µg/ml). Samples were analyzed through laser confocal microscopy.

RNA mobility shift assays (REMSA)

RNA electrophoretic mobility shift assays (REMSA) REMSA were performed using a 256 nt RNA probe corresponding to the 3'UTR of the *EhPgp5* gene (Lopez-Camarillo et al., 2003). The RNA probe was [α -³²P] UTP- labeled by *in vitro* transcription using T7 RNA polymerase (Promega). Purified recombinant EhCAF1 protein (5 nM) and RNA probe (5X10⁵ cpm) were incubated in binding buffer for 15 min at 4 °C and RNA–protein complexes were resolved through 6% non-denaturing PAGE. Gels were vacuum-dried, and radioactive complexes were detected in a Phosphor Imager apparatus.

Preparation of deadenylation substrates

To study the deadenylation, *in vitro* synthesized RNA substrates are used. The preparation of deadenylation substrate was performed using a 256 nt RNA probe corresponding to the 3'UTR of the *EhPgp5* gene (Lopez-Camarillo et al., 2003) with a poly (A₅₀) tract. The RNA probe was obtained by *in vitro* transcription using T7 RNA polymerase (Promega). Reactions conditions were ??? ug, temp, purification ??

In vitro deadenylation assay

In vitro deadenylation assays were performed in 20 mM Tris/Cl (pH 7.0), 150 mM NaCl, 2 mM MgCl₂, 5 U RNasin (Promega), 5 ug RNA substrate and indicated amounts of purified EhCAF1. A volume of 20 µl of the reaction mixture was incubated at 30 °C for 1 h. Reactions that were stopped by the addition of formamide/EDTA buffer were loaded onto 7 M urea/18% acrylamide (19:1) gels that were stained with toluidine blue.

GST pull-down assay

For GST-EhCAF1 pull-down assays, 4 mg of purified fusion protein was added to 100 ug of crude *E. histolytica* extract. Protein extracts from HMI-IMSS trophozoites were performed as described elsewhere (Marchat et al, 2002).. Following addition of the glutathione sepharose-bound GST fusion protein, extracts were incubated at 4°C with gentle agitation for 3 h. Glutathione beads were collected by brief centrifugation, washed three times in was buffer (20 mM HEPES, pH 7, 0.01% Nonidet P40, 100 mM KCl, 4 mM MgAc, 1mM EDTA, 1 mM DTT) , resuspended in SDS-PAGE sample buffer and bound proteins were separated by SDS-polyacrylamide gel electrophoresis and identified by mass spectrometry.

Tandem Mass Spectrometry (LC/ESI-MS/MS)

Protein spots were excised from Coomassie blue stained polyacrylamide gels, washed with 50% (v/v) methanol, 5% (v/v) acetic acid for 2 h and then, with deionized water 3 times, 10 min each. The destained gels were soaked for 10 min in 100 mM ammonium bicarbonate, cut into small pieces, completely dehydrated with 100% acetonitrile and vacuum-dried. In gel digestion was performed by adding 30 µL of modified porcine trypsin solution (Promega, Madison, WI, USA) containing 20 ng/µL in 50 mM ammonium bicarbonate followed by overnight incubation at room temperature. Peptides were extracted with 50% (v/v) acetonitrile, 5% (v/v) formic acid twice for 30 min, and each time with sonication. The volume of the extracts was

reduced by evaporation in a vacuum centrifuge and then adjusted to 20 μ L with 1% (v/v) formic acid.

Mass spectrometric analysis was carried out on a 3200 Q TRAP hybrid tandem mass spectrometer (Applied Biosystems/MDS Sciex, Concord, ON, Canada), equipped with a nano electrospray ion source (NanoSpray II) and a MicroIonSpray II head. The instrument was coupled on-line to a nanoAcquity Ultra Performance LC system (Waters Corporations, Milford, MA). Mass calibration of the hybrid triple quadrupole linear ion trap spectrometer was done with polypropylene glycol standard solutions. The instrument was then tuned and tested using [Glu1]-fibrinopeptide B (Sigma). Samples were desalted by injection onto a Symmetry C₁₈ UPLC trapping column (5 μ m, 180 μ m x 20 mm, Waters Corporations) and washed with 0.1% formic acid in 100% Milli Q water at a flow rate of 15 μ L/min. After 3 min, the trap column was switched in-line with the analytical column. Peptides were separated on a BEH, C₁₈ UPLC column (1.7 μ m, 75 μ m x 100 mm, Waters Corporations) equilibrated with 2% acetonitrile, 0.1% formic acid using a linear gradient of 2-70% acetonitrile, 0.1% formic acid over a 60 min period, at a flow rate of 0.25 μ L/min. Spectra were acquired in automated mode using Information Dependent Acquisition (IDA), which involves switching from MS to MS/MS mode on detection of +2 to +4 charged species. The precursor ions were fragmented by collisionally-activated dissociation (CAD) in the Q2 collision cell. Collision voltages were automatically adjusted based upon the ion charge state and mass using rolling collision energy. Other instrument operation conditions were as described previously [16].

Data interpretation and protein identification were performed with the MS/MS spectra data sets using the MASCOT search algorithm (Version 1.6b9, Matrix Science, London, UK available at <http://www.matrixscience.com>). Searches were conducted using the *Homo sapiens* subset of the National Center for Biotechnology Information non-redundant database (NCBI nr, <http://www.ncbi.nih.gov>). Trypsin was used as the specific protease and one missed cleavage was allowed with tolerances of 0.5 Da for the precursor and 0.3 Da for the fragment ion masses. Carbamidomethyl-cysteine and methionine oxidation were used as the fixed and variable modifications, respectively. A protein "hit" was accepted as a valid identification when at least two MS/MS spectrum matched at the 95% level of confidence ($p < 0.05$). Ion score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event. The threshold ion score in the above conditions was 41 for $p < 0.05$.

Acknowledgments

This work was supported by UACM, CONACyT, COFAA-IPN and SIP-IPN grants (Mexico). This work was also supported by Mexico-France program (SEP-CONACYT ANUIES; ECOS NORD grant M08-S02).

References

- [1] C.Y. Chen, A.B. Shyu, Mechanisms of deadenylation-dependent decay, Wiley Interdiscip Rev RNA. 2 (2011) 167-183.
- [2] C. Allmang, E. Petfalski, A. Podtelejnikov, M. Mann, D. Tollervey, P. Mitchell, The yeast exosome and human PM-Scl are related complexes of 3' $\rightarrow$ 5' exonucleases, Genes Dev. 13 (1999) 2148-2158.
- [3] G. Caponigro, R. Parker, Mechanisms and control of mRNA turnover in *Saccharomyces cerevisiae*, Microbiol. Rev. 60 (1996) 233-249.

- [4] G. Caponigro, R. Parker, mRNA turnover in yeast promoted by the MAT α 1 instability element, *Nucleic Acids Res.* 24 (1996) 4304-4312.
- [5] J. Chen, J. Rappsilber, Y.C. Chiang, P. Russell, M. Mann, C.L. Denis, Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex, *J. Mol. Biol.* 314 (2001) 683-694.
- [6] M.C. Daugeron, F. Mauxion, B. Séraphin, The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation, *Nucleic Acids Res.* 29 (2001) 2448-2455.
- [7] T.K. Albert, M. Lemaire, N.L. van Berkum, R. Gentz, M.A. Collart, H.T. Timmers, Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits, *Nucleic Acids Res.* 28 (2000) 809-817.
- [8] M. Tucker, M.A. Valencia-Sanchez, R.R. Staples, J. Chen, C.L. Denis, R. Parker, The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*, *Cell.* 104 (2001) 377-386.
- [9] M. Tucker, R.R. Staples, M.A. Valencia-Sanchez, D. Muhlrads, R. Parker, Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*, *EMBO J.* 21 (2002) 1427-1436.
- [10] C.E. Brown, A.B. Sachs, Poly(A) tail length control in *Saccharomyces cerevisiae* occurs by message-specific deadenylation, *Mol. Cell. Biol.* 18 (1998) 6548-6559.
- [11] J. Martínez, Y.G. Ren, P. Nilsson, M. Ehrenberg, A. Virtanen, The mRNA cap structure stimulates rate of poly(A) removal and amplifies processivity of degradation, *J. Biol. Chem.* 276 (2001) 27923-27939.
- [12] V. Badarinarayana, Y.C. Chiang, C.L. Denis, Functional interaction of CCR4-NOT proteins with TATAA-binding protein (TBP) and its associated factors in yeast, *Genetics.* 155 (2000) 1045-1054.
- [13] J.P. Rouault, D. Prévôt, C. Berthet, AM. Birot, M. Billaud, J.P. Magaud, L. Corbo, Interaction of BTG1 and p53-regulated BTG2 gene products with mCaf1, the murine homolog of a component of the yeast CCR4 transcriptional regulatory complex, *J. Biol. Chem.* 273 (1998) 22563-22569.
- [14] D. Prévôt, A.P. Morel, T. Voeltzel, M.C. Rostan, R. Rimokh, J.P. Magaud, L. Corbo, Relationships of the antiproliferative proteins BTG1 and BTG2 with CAF1, the human homolog of a component of the yeast CCR4 transcriptional complex: involvement in estrogen receptor α signaling pathway, *J. Biol. Chem.* 276 (2001) 9640-9648.
- [15] A. Traven, A. Hammet, N. Tennis, C.L. Denis, J. Heierhorst, Ccr4-Not Complex mRNA deadenylase activity contributes to DNA damage responses in *Saccharomyces cerevisiae*, *Genetics.* 169 (2005) 65-75.
- [16] R.N. Woolstencroft, T.H. Beilharz, M.A. Cook, T. Preiss, D. Durocher, M. Tyers, Ccr4 contributes to tolerance of replication stress through control of CRT1 mRNA poly(A) tail length, *J. Cell Sci.* 119 (2006) 5178-5192.
- [17] A.C. Goldstrohm, B.A. Hook, D.J. Seay, M. Wickens, PUF proteins bind Pop2p to regulate messenger RNAs, *Nat. Struct. Mol. Biol.* 13 (2006) 533-539.
- [18] A.C. Goldstrohm, D.J. Seay, B.A. Hook, M. Wickens, PUF protein-mediated deadenylation is catalyzed by Ccr4p, *J. Biol. Chem.* 282 (2007) 109-114.
- [19] S.L. Jr. Stanley, Amoebiasis, *Lancet.* 361 (2003) 1025-1034.
- [20] L.A. Marchat, C. Gómez, D.G. Pérez, F. Paz, L. Mendoza, E. Orozco, Two CCAAT/enhancer binding protein sites are cis-activator elements of the

- Entamoeba histolytica* EhPgp1 (mdr-like) gene expression, *Cell Microbiol.* 4 (2002) 725-737.
- [21] J.A. Hackney, G.M. Ehrenkaufer, U. Singh, Identification of putative transcriptional regulatory networks in *Entamoeba histolytica* using Bayesian inference, *Nucleic Acids Res.* 35 (2007) 2141-2152.
- [22] S.F. Alschult, T.L. Madden, A.A. Schaffer, J. Zang, Z. Zang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acid Res.* 25 (1997) 3389-3402.
- [23] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods, *Mol. Biol. Evol.* 28 (2011) 2731-2739.
- [24] L.S. Diamond, D.R. Harlow, C.C. Cunnick, A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*, *Trans. R. Soc. Trop. Med. Hyg.* 72 (1978) 431-432.
- [25] M. López-Casamichana, E. Orozco, L.A. Marchat, C. López-Camarillo, Transcriptional profile of the homologous recombination machinery and characterization of the EhRAD51 recombinase in response to DNA damage in *Entamoeba histolytica*, *BMC Mol. Biol.* 9 (2008) 35-50.
- [26] C. Weber, G. Guigon, C. Bouchier, L. Frangeul, S. Moreira, O. Sismeiro, C. Gouyette, D. Mirelman, J.Y. Coppee, N. Guillén, Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in *Entamoeba histolytica*, *Eukaryot. Cell.* 5 (2006) 871-875.
- [27] S. Thore, F. Mauxion, B. Séraphin, D. Suck, X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex, *EMBO Rep.* 4 (2003) 1150-1155.
- [28] C. Temme, S. Zaessinger, S. Meyer, M. Simonelig, E. Wahle, A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in *Drosophila*, *EMBO J.* 23 (2004) 2862-2871.
- [29] A. Schwede, L. Ellis, J. Luther, M. Carrington, G. Stoecklin, C. Clayton, An essential role for Caf1 in mRNA deadenylation in trypanosomes and human cells, *Nucleic Acids Res.* 36 (2008) 3374-3388.
- [30] D. Zheng, N. Ezzeddine, C.Y. Chen, W. Zhu, W. X. He, A.B. Shyu, Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells, *J. Cell Biol.* 182 (2008) 89-101.
- [31] L. Molin, A. Puisieux, C. elegans homologue of the Caf1 gene, which encodes a subunit of the CCR4-NOT complex, is essential for embryonic and larval development and for meiotic progression, *Gene.* 358 (2005) 73-81.
- [32] A. Schwede, T. Manful, B.A. Jha, C. Helbig, N. Bercovich, M. Stewart, C. Clayton, The role of deadenylation in the degradation of unstable mRNAs in trypanosomes, *Nucleic Acids Res.* 37 (2009) 5511-5528.
- [33] G. Gowrishankar, R. Winzen, F. Bollig, B. Ghebremedhin, N. Redich, B. Ritter, K. Resch, M. Kracht, H. Holtmann, Inhibition of mRNA deadenylation and degradation by ultraviolet light, *Biol. Chem.* 386 (2005), 1287-1293.
- [34] G. Gowrishankar, R. Winzen, O. Dittrich-Breiholz, N. Redich, M. Kracht, H. Holtmann, Inhibition of mRNA deadenylation and degradation by different types of cell stress, *Biol. Chem.* 387 (2006) 323-327.
- [35] N.G. Theodorakis, R.I. Morimoto, Posttranscriptional regulation of hsp70 expression in human cells: effects of heat shock, inhibition of protein synthesis,

- and adenovirus infection on translation and mRNA stability, *Mol. Cell Biol.* 7 (1987) 4357-4368.
- [36] R.P. Dellavalle, R. Petersen, S. Lindquist, Preferential deadenylation of Hsp70 mRNA plays a key role in regulating Hsp70 expression in *Drosophila melanogaster*, *Mol. Cell Biol.* 14 (1994) 3646-3659.
- [37] C. López-Camarillo, M. Lopez-Casamichana, C. Weber, N. Guillen, E. Orozco, L.A. Marchat, DNA repair mechanisms in eukaryotes: special focus in *Entamoeba histolytica* and related protozoan parasites, *Infect. Genet. Evol. J.* 9 (2009) 1051-1056.
- [38] C. Weber, L.A. Marchat, N. Guillen, C. López-Camarillo, Effects of DNA damage induced by UV irradiation on gene expression in the protozoan parasite *Entamoeba histolytica*, *Mol. Biochem. Parasitol.* 164 (2009) 165-169.
- [39] S. Katz, O. Kushnir, A. Tovy, R. Siman Tov, S. Ankri, The *Entamoeba histolytica* methylated LINE-binding protein EhMLBP provides protection against heat shock, *Cell Microbiol.* (2011), doi: 10.1111/j.1462-5822.2011.01697.
- [40] C. Bönisch, C. Temme, B. Moritz, E. Wahle, Degradation of hsp70 and other mRNAs in *Drosophila* via the 5' 3' pathway and its regulation by heat shock, *J. Biol. Chem.* 282 (2007) 21818-2128.
- [41] N. Cougot, S. Babajko, B. Séraphin, Cytoplasmic foci are sites of mRNA decay in human cells, *J. Cell Biol.* 165 (2004) 31-40.

Figures legends

Fig. 1. Expression and subcellular localization of EhCAF1. (A) Comparison of POP2/CAF1 proteins from diverse organisms. Upper panel, Schematic representation of POP2/CAF1 proteins. Lower panel, multiple sequence alignment of DEDD residues from POP2/CAF1 proteins. Black boxes, identical residues; grey boxes, conserved substitutions. Numbers are relative to the initial Met in each protein. Asterisks, conserved DEDD residues (B) Prediction of EhCAF1 tertiary structure using the Swiss Model software and yeast POP2 protein crystal structure (PDB entry 1UOC) as template. Model was displayed and refined using the Pymol PBD viewer. (C) Relationships between POP2/CAF1 proteins from various organisms through PSI-BLAST analysis. The width of connected lines represents similarity level according to e-value threshold. (D) Phylogenetic relationships between POP2/CAF1 proteins. The rooted tree of CAF1 proteins was created with the Mega 5.1 program using the Neighbor Joining algorithm based on ClustalW alignments of complete aa sequences. Origin and accession number of each protein are indicated. The scale is at the bottom. (E) Expression of rEhCAF1-6x-His-tagged protein in *E. coli* BL21(DE)pLysS. Bacterial proteins were analyzed through 10% SDS-PAGE and Coomassie blue staining. Lane 1, non-induced bacteria lysate; lane 2 and 3, IPTG induced bacteria lysate. (F) Purification of rEhCAF1 through affinity chromatography. Fractions were analyzed through 10% SDS-PAGE and Coomassie blue staining. Lane 1, pellet from IPTG induced bacteria; lane 2, clarified lysate from IPTG induced bacteria; lane 3, clarified lysate flow through; lane 4, rEhCAF1 bound to NI-NTA resin; lanes 5 and 6, washes; lanes 7 and 8, elution fractions containing purified rEhCAF1. (G) Western blot of rEhCAF1 using anti-6x-His-tag antibodies. Lane 1, clarified lysate from IPTG induced bacteria; lane 2, rEhCAF1 bound to NI-NTA resin; lane 3 and 4, purified rEhCAF1. (H) Immunodetection of EhCAF1 in

trophozoites using specific anti-EhCAF1 antibodies (lanes 1 and 2). Antibodies raised against EhPC4 were used as control (lanes 3 and 4). Lanes 1 and 3, CE; lanes 2 and 4, NE. Arrowhead, EhCAF1 (37 kDa) and EhPC4 (18 kDa) proteins. Molecular weight markers are at the left.

Fig. 2. Cellular localization of EhCAF1 in trophozoites. Fixed trophozoites were incubated with anti-EhCAF1 antibodies, treated with FITC-labeled secondary antibodies, counterstained with DAPI and analyzed through confocal immunofluorescence microscopy. (A-H) EhCAF1 immunodetection. (I-L) Control of immunodetection with pre immune serum. (A, E and I) Nomarsky images; (C, G and K) green channel (FITC); (B, F and J) blue channel (DAPI); (D, H and L) merge images. E, F, G and H, single cell (100 X magnification). Arrowheads, EhCAF1.

Fig. 3. Subcellular localization of EhCAF1 after induced stress.

Fig. 4. RNA binding and dedenylation assays for recombinant EhCAF1.

Fig. 5. Pulldown experiments and identification of CAF1 interacting proteins by ESI-MS/MS.

Fig. 6. Predicted interactome for EhCAF1 protein.

Fig. 7. Cellular localization of EhCAF1 with LPS endoribonuclease

Tables

Table 1. Comparison of *Entamoeba histolytica*, *Homo sapiens* and *Saccharomyces cerevisiae* deadenylases.

Table 2. Comparison of EhCAF1 with related proteins from other organisms.

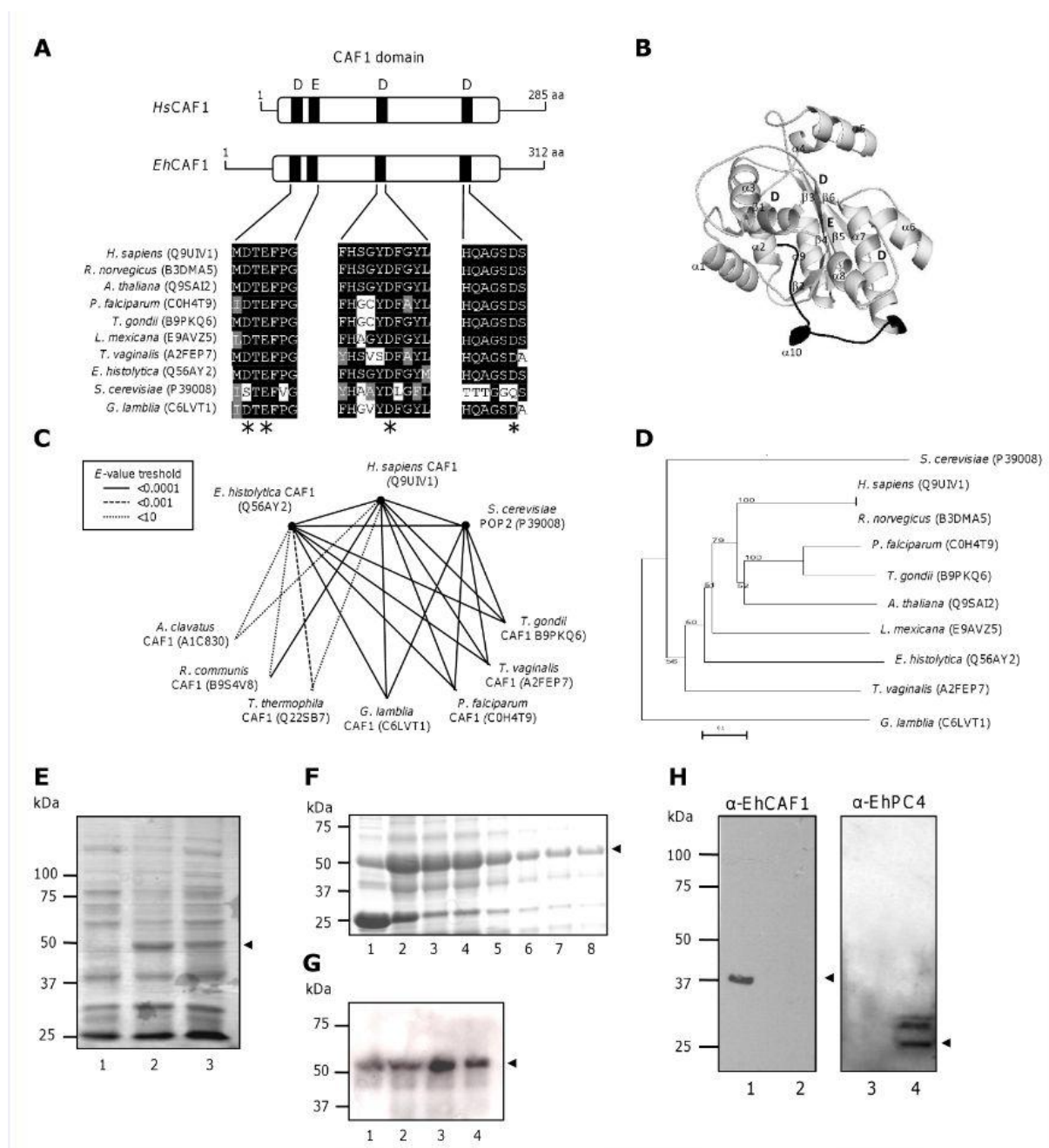


Figure 1. Expression and subcellular localization of EhCAF1.

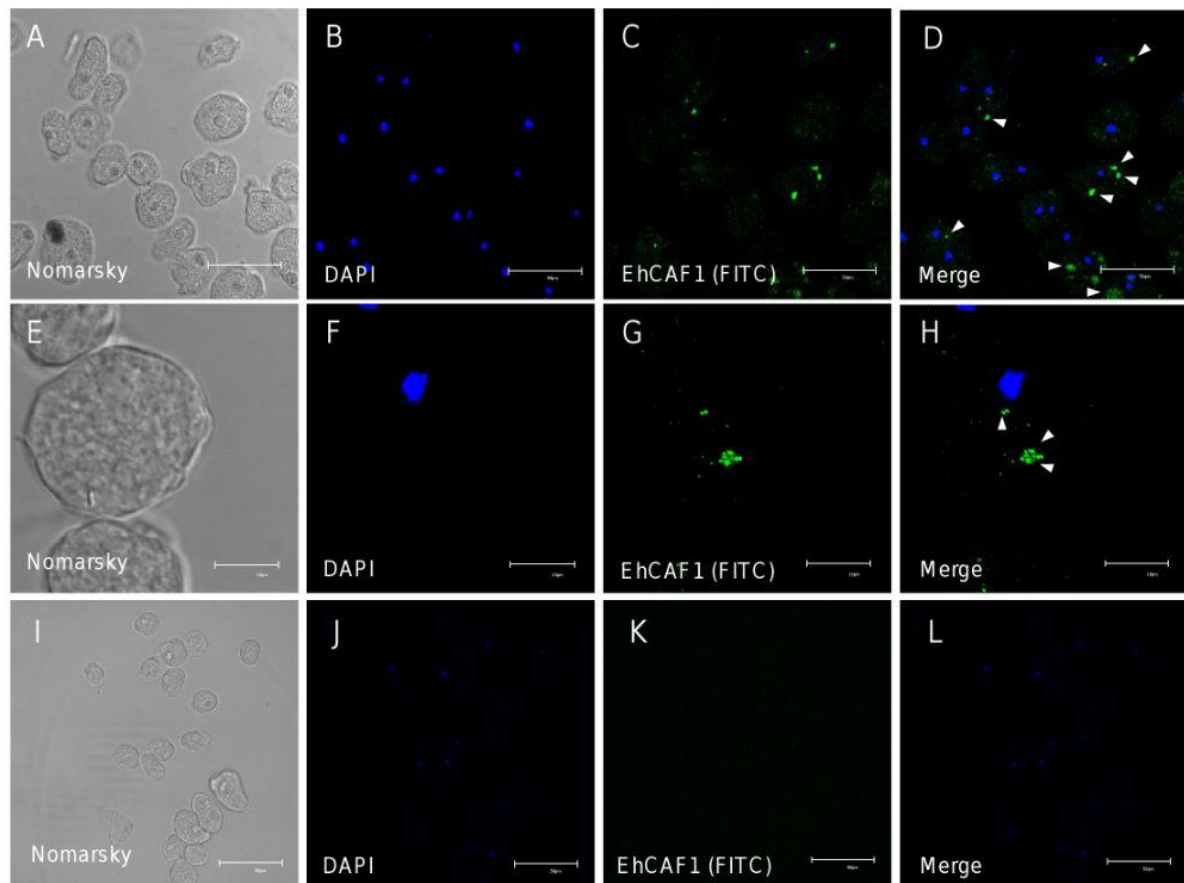


Figure 2. Cellular localization of EhCAF1 in trophozoites.

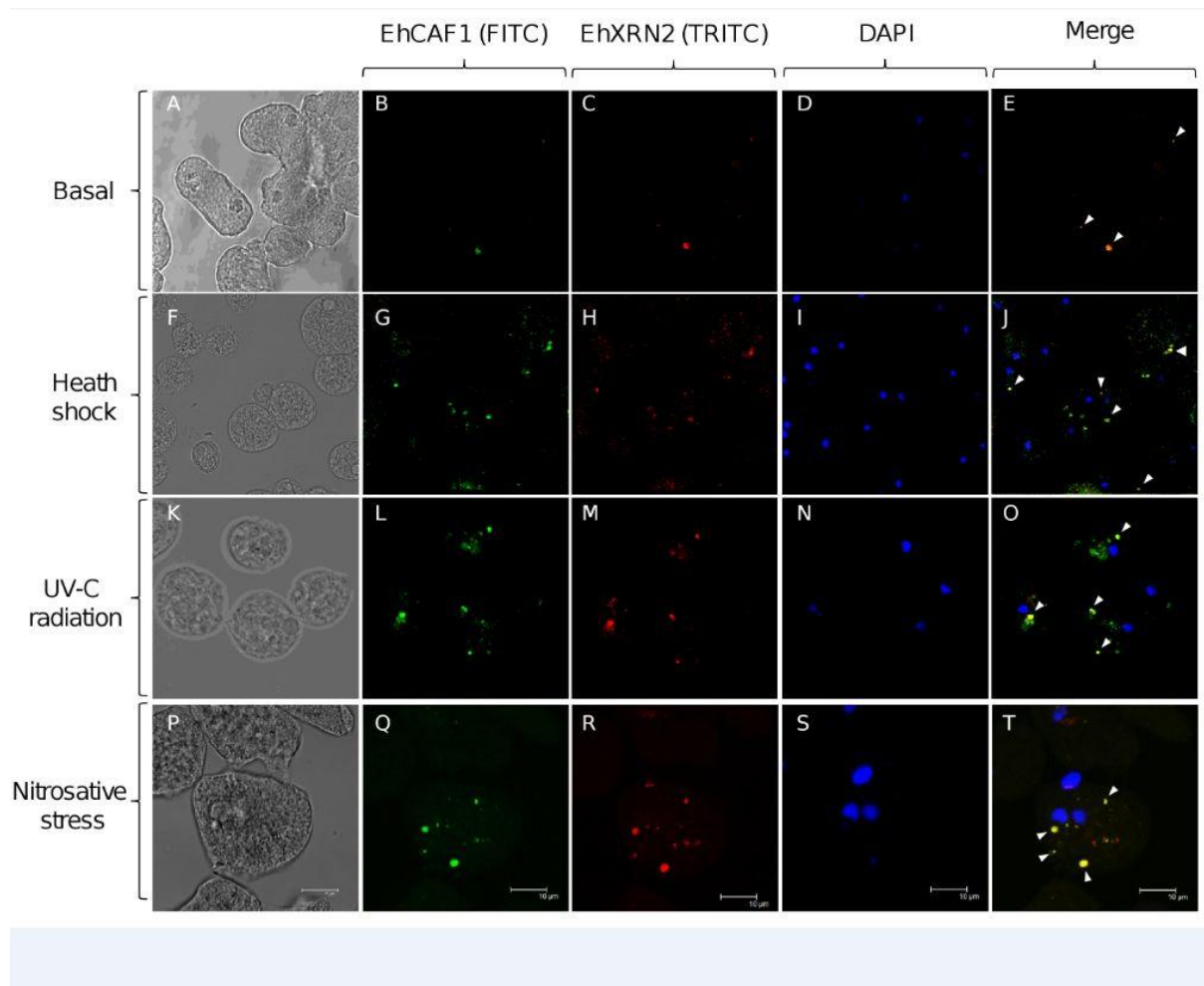


Figure 3. Subcellular localization of EhCAF1 after induced stress.

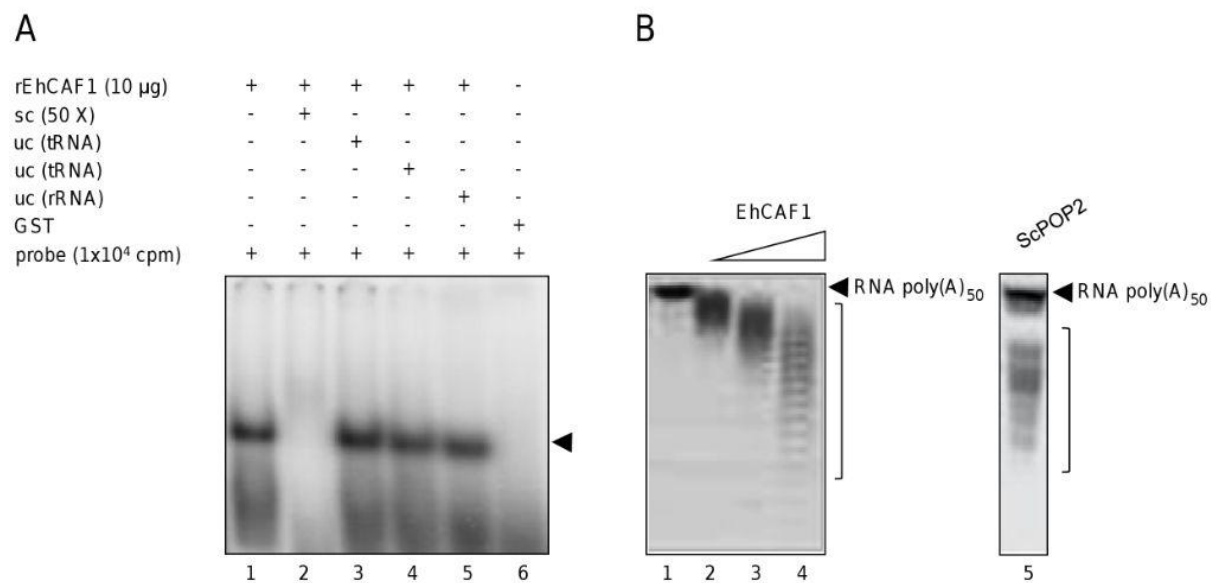


Figure 4. RNA binding and dedenylation assays for recombinant EhCAF1.

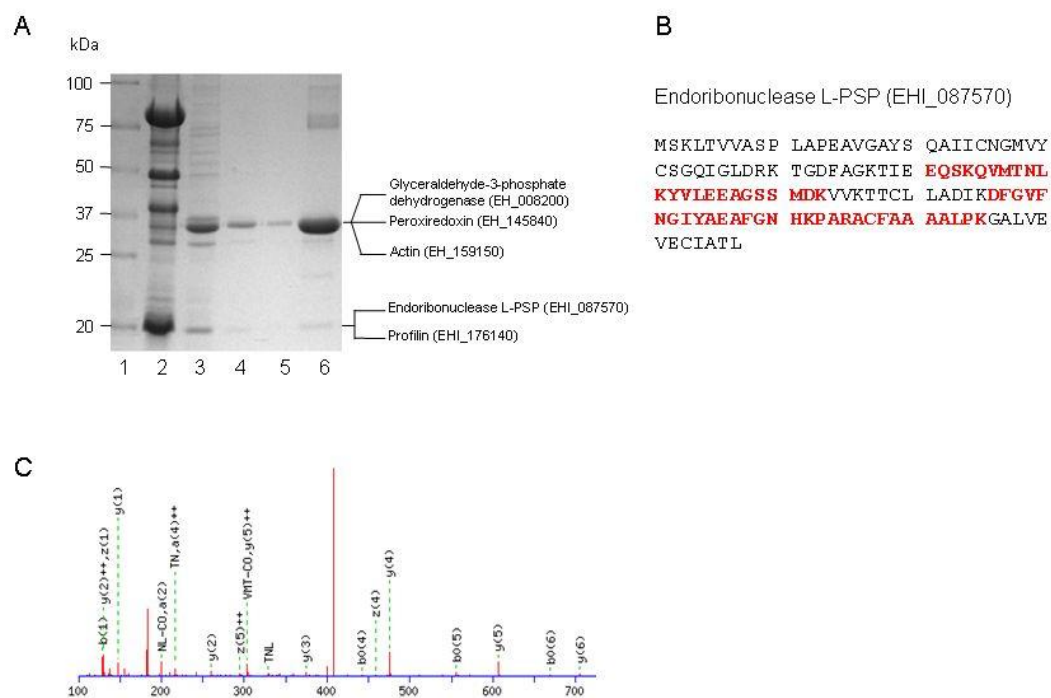


Fig. 5. Pulldown experiments and identification of CAF1 interacting proteins by ESI-MS/MS.

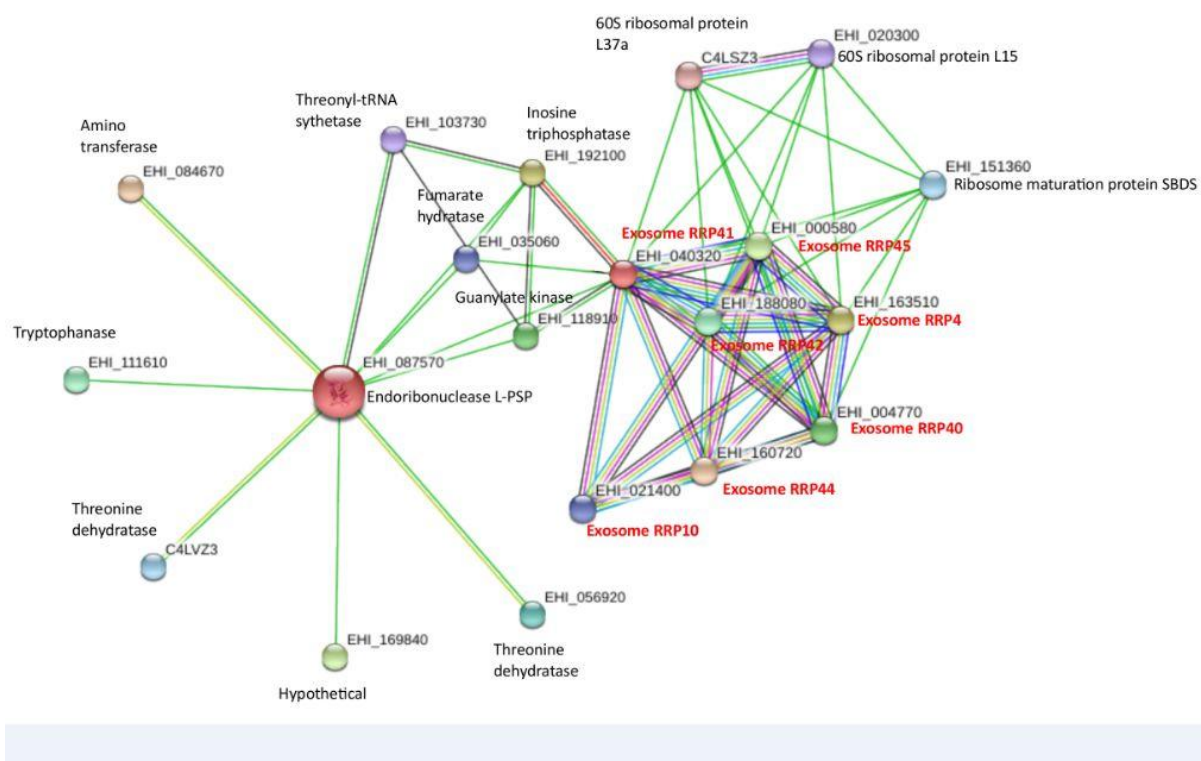


Figure 6. Predicted interactome for EhCAF1 protein

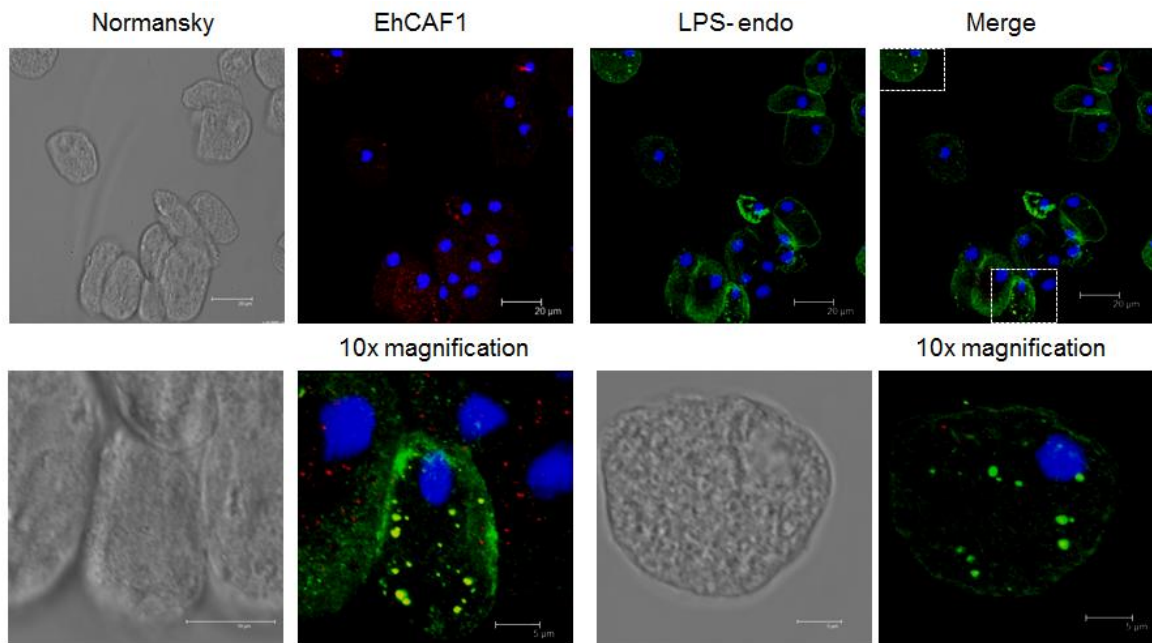


Figure 7. Cellular localization of EhCAF1 with LPS endoribonuclease

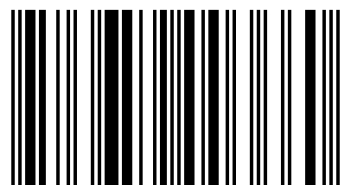
Degradación del RNAm en eucariontes

Las infecciones gastrointestinales asociadas a parásitos protozoarios tienen altos índices de morbilidad y mortalidad en el mundo. Desde hace algunos años una de las principales áreas de estudio en parásitos se ha enfocado en la investigación de los mecanismos que regulan la expresión de genes involucrados en virulencia, patogenicidad, diferenciación, resistencia a fármacos y la evasión de la respuesta inmune. La comprensión de dichos mecanismos moleculares puede ayudar en el desarrollo de pruebas de diagnóstico, tratamientos más eficientes y la generación de vacunas. El control de la degradación de RNAm es un evento importante en la regulación de la expresión génica. La cantidad de RNAm presente en la célula está determinada por su tasa de síntesis y degradación, las cuales regulan la expresión final de los genes al proveer de RNA mensajeros a la maquinaria de traducción. En este libro presentamos una revisión de los aspectos básicos de la degradación del RNAm en eucariontes. También se resumen los mecanismos de degradación del RNAm en *Entamoeba histolytica* el parásito protozoario causante de la amibiasis humana y se establece su importancia en el proceso infeccioso.



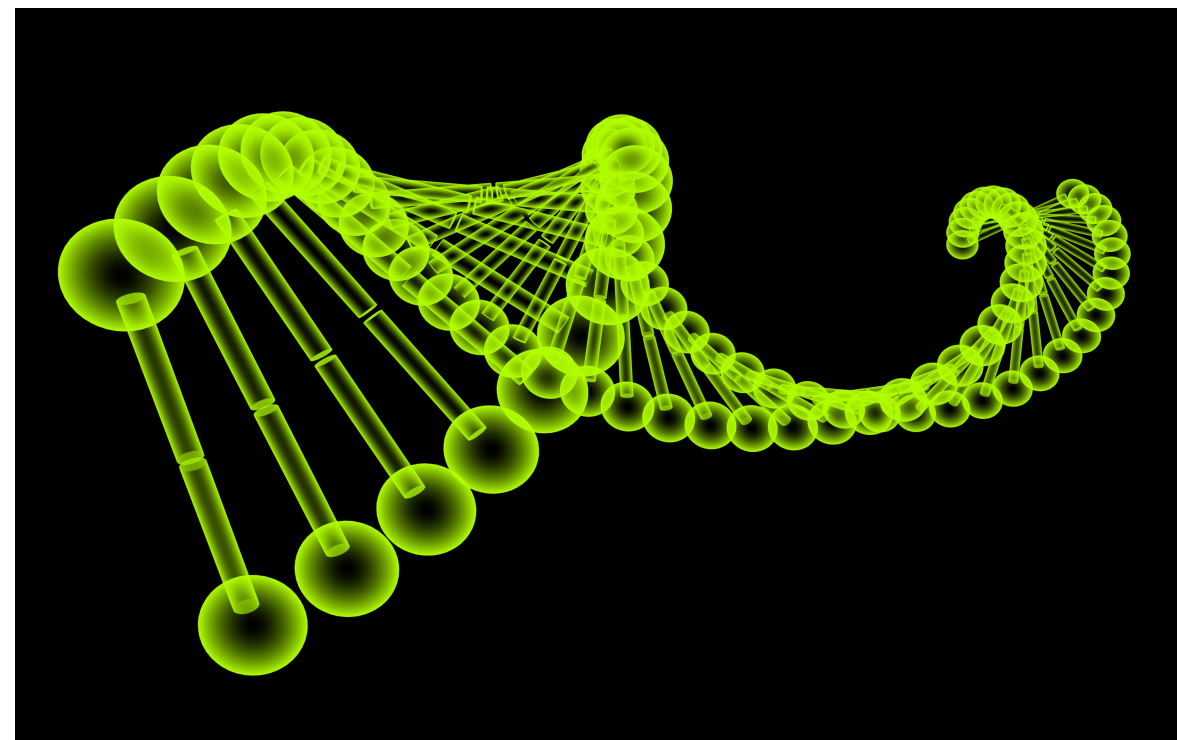
Mario César López Camarillo

Doctor en Biomedicina Molecular por la Escuela Nacional de Medicina y Homeopatía-IPN, México. Autor de 30 artículos de investigación en el área de la amibiasis y el cáncer de mama y de 12 capítulos de libros internacionales. Actualmente es Profesor Investigador en el Posgrado en Ciencias Genómicas de la Universidad Autónoma de la Ciudad de México.



978-3-659-00366-0

editorial académica española



Mario César López Camarillo · Itzel López-Rosas · Laurence A. Marchat

Degradación del RNAm en eucariontes

el caso del parásito *Entamoeba histolytica*

Degradación del RNAm en eucariontes: el caso del parásito *Entamoeba histolytica*.

Itzel López Rosas¹, Laurence A. Marchat², César López-Camarillo^{1*}

¹Universidad Autónoma de la Ciudad de México, Posgrado en Ciencias Genómicas, México D.F., México. ²Escuela Nacional de Medicina y Homeopatía del IPN, Programa Institucional en Biomedicina Molecular y Red en Biotecnología, México D.F., México.

Palabras clave: *Entamoeba histolytica*, amibiasis, genoma, degradación del RNAm.

***Autor responsable:** Tel: +(55) 54886661 ext. 15307.

Email: cesar.lopez@uacm.edu.mx

Genomics and proteomics approaches to understand virulence of *Entamoeba histolytica*.

C. López-Camarillo^{1*}, E. Azuara-Liceaga¹, A. Zamorano², O. Hernández de la Cruz¹, I. López Rosas¹ and L. Marchat²

¹Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México. San Lorenzo 290, Col del Valle. 03100. México DF. ²Programa Institucional de Biomedicina Molecular, Doctorado en Biotecnología en Red, ENMyH-IPN., Guillermo Massieu Helguera 239, Ticoman. 07320. México DF.

*Corresponding author. Email: cesar.lopez@uacm.edu.mx

Entamoeba histolytica is the protozoan parasite responsible for human amoebiasis. This infection has a world-wide distribution affecting more than 50 million people each year, mainly in developing countries. *E. histolytica* causes fulminating intestinal dysentery, and bloody diarrhea. In some cases, virulent trophozoites can cross the intestinal epithelia to reach different organs of the human body, usually the liver (but also the lungs, brain or spleen) to provoke liver abscesses, which can result in 70,000 -100,000 deaths a year. Virulence variability of trophozoites has been related to genome plasticity and *in vivo* changes in gene expression induced during invasion. Some virulence factors have been identified as key factors in pathogenesis, however most molecular mechanisms relevant for infection establishment are still poorly understood.

Genome sequence represents an invaluable tool to dissect the molecular mechanisms modulating virulence and gene expression in parasites. The fully sequenced genome of *E. histolytica* provides new insights into parasite biology and genome evolution and it also opens the *omics* research era in this major human pathogen. Microarrays made from synthetic oligonucleotides based on annotated genes have been successfully applied in *E. histolytica* to obtain transcriptional profiling in some *in vivo* and *in vitro* conditions. In addition, proteomic profiles have been obtained from virulent and non-virulent trophozoites providing novel exciting novel data about factors potentially involved in virulence. Here, we reviewed the actual efforts based in genomics and proteomics approaches to understand the pathogenesis of this early branch eukaryotic parasite.

Keywords *Entamoeba histolytica*; amoebiasis; virulence; genomics; proteomics; gene regulation.

1. Introduction

The protozoan parasite *Entamoeba histolytica* has a world-wide prevalence but it is often endemic in developing tropical countries, such as India, Mexico, Central and South America because of poor hygienic and sanitation conditions. About 10% of the world's population is infected with *E. histolytica*, which represents more than 50 million people each year. Amoebiasis is the third most common cause of death from parasitic infections, after Schistosomiasis and Malaria [1]. *E. histolytica* is usually transmitted by the fecal-oral route, through contaminated food and water. Infection begins with the ingestion of the cyst form of the parasite that is able to survive in the environment due to its protective cell wall. Cyst undergoes excystation in the intestine and produces the proliferative trophozoite forms that cause amoebic dysentery or amoebic colitis. If the parasite reaches the bloodstream, it can spread through the body and invade other organs, mainly the liver (but also the lungs, brain or spleen), to provoke abscesses that can result in 70,000 -100,000 deaths a year. On the other hand, about 90 percent of infections remain asymptomatic [2]. The virulence variability of trophozoites has been related to genome plasticity and *in vivo* changes in gene expression induced during host invasion. Some virulence factors have been identified as key factors in pathogenesis, however most molecular mechanisms relevant for infection establishment are still poorly understood.

The elucidation of molecular mechanisms modulating virulence and gene expression is of major interest for researchers working with *E. histolytica* in order to understand parasite biology and be able to identify biochemical targets or vaccine candidates that could help to control amoebiasis. Here, we reviewed the actual efforts based in genomics and proteomics approaches to understand the pathogenesis of this early branch eukaryotic parasite. Particularly, we focused on the recent sequencing of the parasite genome that has revealed a variety of metabolic adaptations and the existence of gene families associated with virulence. We also described data obtained from transcription and proteomic profiling that altogether contributed to the knowledge about factors potentially involved in virulence.

2. *Entamoeba histolytica* genome

The draft sequence of the complete genome of *E. histolytica* was published in 2005 making it one of the first protist genomes to be sequenced by The Institute for Genomic Research (<http://www.tigr.org/>) and the Sanger Institute (<http://www.sanger.ac.uk/>). A whole-genome shotgun method was used to sequence parasite genome as a better choice for sequencing [3]. This protist has an A/T-rich genome which have made difficult its sequencing and assembly. *E. histolytica* genome is 20.8 Mb in size and is carried on 14–17 chromosomes. The length variation of homologous chromosomes is thought to be caused by expansion and contraction of the sub-telomeric region, which might be formed by tRNA-containing arrays that contain between one and five tRNA types per repeat [3, 4]. The genome analysis was carried out on a 12.5-fold coverage genome assembly consisting of 23,751,783 base pairs (bp) distributed among 888 scaffolds. Initially 9,938 genes were predicted and automatically annotated. But recently *E. histolytica* genome was re-assembled and re-annotated using a combination of manual and automated methods. The new genome assembly consists of 20 Mb of sequence organized into 1,496 scaffolds. This new assembly showed a higher fragmentation and a reduction in genome size with respect of data published earlier. The new assembly contains less predicted protein coding genes. The main reason for gene number reduction is the elimination of genes within repetitive regions, artifactual tandem duplications, and the removal of genes smaller than 300 bp without any supporting evidence [5]. In summary the *E. histolytica* genome size is 20.8 Mb with identification of approximately 8300 predicted genes, each averaging 1260.9 pb in size, and corresponding to 49.7% of the genome. The shortest gene is 147 pb and the longest gene is 15,210 pb. Introns are contained in 24.4 % of the predicted genes. The intergenic regions mean length is about 708.7 bp with a 20.5% GC content [5].

E. histolytica genome is also characterized by its repetition and redundancy [6]. Redundancy can be seen in the predicted proteins, which can be organized into protein families. A total of 897 protein families were identified from the 8,201 predicted polypeptides. Among the families, 247 families have no homology to any known Pfam or TIGRfam domain [5]. Analysis of the genome reveals redundancy in genes encoding virulence factors. For example, *E. histolytica* genome contains a total of 86 genes coding for putative peptidases [4, 6]. These peptidases comprise 50 cysteine peptidases; some of them have a predicted N-terminal transmembrane anchor, which might allow them to be localized on the *Entamoeba* cell surface. Additionally, four aspartic, 10 serine and 22 metallo peptidase genes were also identified in *Entamoeba* genome [6]. Furthermore, 16 genes coding for putative saposin-like proteins were identified. The proteins encoded by these genes could be related to amoebapores [4]. Finally there are over 75 genes encoding leucine-rich tandem repeats of the type found in BspA-like proteins. These proteins generally have a surface location and may be involved in cell–cell interactions [5].

Vesicular trafficking is important in pathogenesis through phagocytosis and delivery of proteases and amoebapores to cell surface. Rab and Arf protein family expansions reflect the increased complexity and number of vesicle fusion and recycling steps that have been associated with this process. *E. Histolytica* has families with more than 50 members, such as small GTP binding proteins, which control a number of processes involving the actin cytoskeleton [3]. Signal-transduction-related proteins like protein kinases (270 gene encoding proteins), protein phosphatases (100 gene encoding proteins), which dephosphorylate proteins, numerous putative seven-transmembrane receptors, trimeric G proteins, Ras-family proteins were also identified in *Entamoeba* genome [3]. This represents the most varied set of signal-transduction-related proteins yet described in a single-celled eukaryote. Another important feature of *Entamoeba* genome is the lateral gene transfer (LTG) primarily from Bacteroidetes phyla (Cytophaga–Flavobacterium–Bacteroides). Initially 96 genes were identified using phylogenetic analyses [3], but only 41 remain as strongly supported [4]. Most of these genes encode proteins involved in carbohydrate and protein metabolism increasing the range of substrates available for energy generation. Most of the proteins encoded by these genes are specific to the parasite and may be used as potential drug targets for the design of new anti-amebic drugs.

Finally, *Entamoeba* genome and DNA microarrays expression data is organized in AmoebaDB database (<http://amoebadb.org/amoeba/>). This database belongs to the National Institute of Allergy and Infectious Diseases (NIAID) funded EuPathDB (<http://EuPathDB.org>) Bioinformatics Resource Center family of integrated databases. AmoebaDB contains the genomes of *E. dispar*, *E. invadens* besides *E. histolytica*. This database contains functional genomic databases for the scientific community studying eukaryotic pathogens [8].

3. Genomic expression profiling in *E. histolytica*

3.1 Genomics basics

Genomics is a recent biology discipline focused in the study of genomes. Genome science is focused in the study of the structure, content, and evolution of genomes [9]. The field includes the initial intensive efforts to determine the entire nucleotides sequence of uni- and multicellular organisms, as well as the development of sophisticated equipments and bioinformatic tools. Functional genomics attempts to understand the relationships between complete genome sequences of organisms and gene (and protein) functions. To gain insights into cell function, genomics exploits the genome-wide

approach involving high-throughput methods rather than a more traditional “gene-by-gene” method. Notably, the availability of genome sequences from parasites of medical importance for human, such as *E. histolytica*, provides opportunities to study genomic expression profiles in pathogens under diverse conditions.

3.2 DNA microarrays

The study of mRNA expression in the genomics era was accelerated by the development and use of high-throughput tools to measure at genome-wide scale the expression of thousand genes at the same time. These methodologies were initially represented by the Serial Analysis of Gene Expression (SAGE), which is based on cloning of small tags that correspond to fragments of transcripts followed by clones sequencing [10]. To date DNA microarray (also commonly known as DNA chip or biochip) are the most popular methodology to determine genome-wide expression profiles (transcriptome). DNA microarrays represent a very versatile tool that allows measuring mRNA expression of large numbers of genes simultaneously. DNA microarrays are a collection of microscopic DNA spots (represented by short or long sequences of nucleotides) attached to a solid surface commonly glass slides. Each DNA spot contains a specific DNA sequence (a gene or a non-coding DNA sequence), known as probe that is hybridized to a complementary DNA (cDNA) called target. To determine the relative abundance of nucleic acid sequences in the target (which represents mRNA levels) targets are first coupled to a fluorophore, which permits the detection of probe-target hybridization using a scanner. By using statistical approaches, global mRNA gene expression can be accurately determined. In addition, clustering methods may be used to identify genomic fingerprints which represent co-regulated genes [Figure 1].

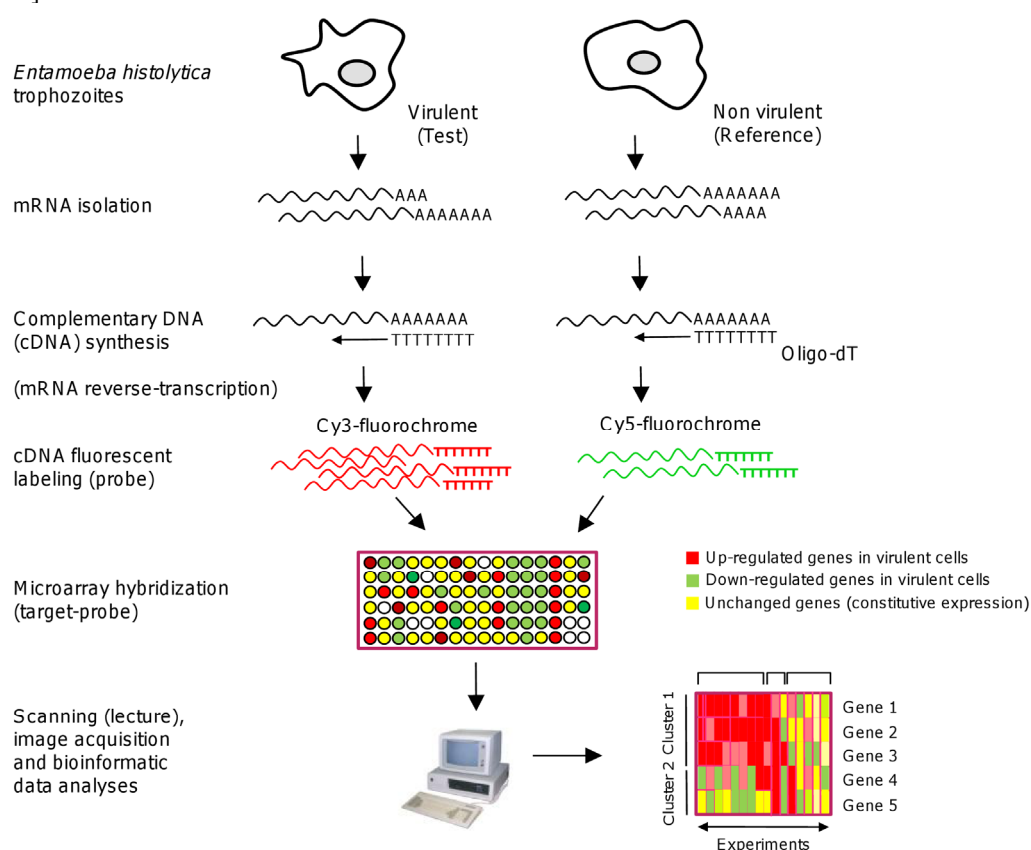


Figure 1. Principle of DNA microarrays assay of gene expression.

Microarrays technology has been successfully applied in *E. histolytica* to obtain transcriptional profiling in several conditions, including virulence, stress and DNA damage, during cyst-trophozoites conversion and to evaluate the role of epigenetic control on gene expression [11-23]. Next, we discuss some of the genomic studies performed to understand different topics of *E. histolytica* biology.

3.3 Genome-wide expression profiling in response to L-cysteine deprivation

L-cysteine is a sulfur-containing amino acid which plays an essential role in different cellular processes including stability, structure, regulation of catalytic activity, and posttranslational modification of proteins. Due to the ability of its thiol group to undergo redox reactions, it plays an important role in antioxidative defense and is used for biosynthesis of glutathione, the major endogenous antioxidant produced by the cells, participating directly in the neutralization of free radicals and reactive oxygen compounds. In *E. histolytica*, L-cysteine is required for growth, attachment, survival, and protection from oxidative stress [24] and L-cysteine deprivation leads to drastic changes in metabolic pathways, including energy, amino acid, and phospholipid metabolism [25]. To better understand the role of L-cysteine in *E. histolytica*, Husain et al. [26] performed DNA microarrays analysis of gene expression at different times after L-cysteine deprivation, and reported that 290 genes with functions in metabolism, signaling, DNA/RNA regulation, electron transport, stress response, membrane transport, vesicular trafficking/secretion, and cytoskeleton were ≥ 3 fold differentially expressed at one or more time points upon L-cysteine deprivation; 129 genes were up-regulated and 167 genes were down-regulated, while 6 genes showed both up- and down-regulation depending upon the time. This study also confirmed that most of the L-cysteine deprivation-mediated metabolomic changes in amino acid, central energy, and phospholipid metabolism are not associated with changes in the expression of the corresponding genes. This general lack of correlation between metabolome, proteome, and transcriptome appears to be a general characteristic in various organisms including *E. histolytica*, indicating that they have complex mechanisms of gene expression regulation.

3.4 Effects of DNA methylation inhibitors on global gene expression

Methylation is an enzymatic modification of DNA that epigenetically contributes to transcriptional regulation of gene expression. DNA methylation can silence genes by blocking the interaction of transcription factors to their regulatory sequences; it can also attract methyl-binding protein, which recruits histone deacetylases and histone methyltransferases resulting in an inactive chromatin structure. In *E. histolytica*, a DNA methyltransferase (*Eh*meth) has been characterized [27], and an amoebic protein which preferentially binds to methylated DNA has been identified (*Eh*MLBP) [28]. Interestingly, when *E. histolytica* strain HM1:IMSS is grown with 5-azacytidine (5-AzaC), a potent inhibitor of DNA methyltransferase, the *in vitro* and *in vivo* virulence is significantly reduced. To identify the genome-wide effects of DNA methylation in *E. histolytica*, Ali et al. [29] used a short oligonucleotides microarray representing 9,435 genes and compared the expression profile of HM-1:IMSS strain parasites with those treated with 23 μ M 5-AzaC for up to one week. Overall, 2.1% of genes tested were transcriptionally modulated under these conditions. 68 genes were up-regulated and 131 genes down-regulated. Sodium-bisulfite treatment and genes sequencing indicated that there were at least two subsets of genes with genomic DNA methylation in *E. histolytica*: genes that were endogenously silenced by genomic DNA methylation and for which 5-AzaC treatment induced transcriptional de-repression, and genes that have genomic DNA methylation, but were not endogenously silenced by methylation. Interestingly, authors showed that a cysteine proteinase and a lysozyme, both having known roles in amoebic pathogenesis, were down-regulated by 5-AzaC treatment. Decreased expression of these genes in the 5-AzaC treated *E. histolytica* may account in part for the reduced cytolytic activities in parasite.

3.5 Genomic studies of a Myb domain-containing protein involved in expression regulation of stage specific genes

Myb family transcription factors are important in regulating cell proliferation, differentiation, stage conversion and cell cycle progression and are widely found in eukaryotes. Conversion between cyst and trophozoite stage is essential to disease transmission and pathogenesis in many protozoa. *E. histolytica* differentiates into infectious cysts to survive outside of the host. Previous work has identified the transcriptome of *E. histolytica* cysts and determined two subsets of developmentally regulated genes, including 672 cyst-specific and 767 trophozoite-specific genes, respectively [30]. However, the molecular basis of encystation regulation is not known. The search for genes involved in the regulation of stage conversion allowed the identification of the first *E. histolytica* developmentally regulated transcription factor called *Eh*Myb-dr. Overexpression of the nuclear *Eh*Myb-dr resulted in a set of regulated genes that significantly overlapped with the expression profile of amoebic cysts, suggesting a role for *Eh*Myb-dr in the regulation of stage-specific genes expression and its participation in *E. histolytica* development. Data from expression profiling allowed the identification of 117 genes that were up-regulated by *Eh*Myb-dr overexpression, 40 overlapped with cyst-specific genes; of the 88 genes that were down-regulated by *Eh*Myb-dr overexpression, 30 overlapped with trophozoite-specific genes. Some genes of interest in these lists include the chitin synthase, the developmentally regulated protein and chitinase. Moreover, bioinformatic methods allowed the identification of transcriptional regulatory networks and CCCCC DNA binding motifs to which amoebic nuclear proteins bind in a sequence-specific manner in a subset of cyst-specific genes. This DNA microarrays-based study provided evidence for the involvement of *Eh*Myb-dr in the differentiation of *E. histolytica* trophozoites into cysts [31].

3.6 Genomic analysis of URE3-BP transcription factor

The transcription factor named upstream regulatory element 3-binding protein (URE3-BP) is a calcium-responding regulator of *E. histolytica* Gal/GalNAc lectin and ferredoxin genes, both implicated in virulence. To investigate the impact of URE3-BP regulation on virulence, the occurrence of URE3 consensus motif in promoter of genes specifically induced during *in vivo* infection, was measured by microarrays [32]. Amoeba expressing the dominant positive mutant form of URE3-BP was more virulent in two different animal models of amebiasis. An increase in liver abscess size was observed when trophozoites expressing the dominant positive mutant URE3-BP were injected into the gerbil liver. In the mouse model of amebic colitis, a competitive advantage at the host tissue interface was conferred upon amoebae expressing the dominant positive mutant URE3-BP protein. URE3-BP therefore promoted the expression of the *E. histolytica* virulence phenotype. Distinct patterns of *E. histolytica* gene expression have been observed under a variety of experimental conditions. A comparison of the transcripts expressed during *in vitro* culture trophozoite and one day after infection in the mouse model of amebiasis identified changes in parasite gene expression. Consistent with a role for URE3-BP in virulence, a statistically significant enrichment of the URE3 motif was observed in promoters of *in vivo*. However, further studies are required to understand the particular contribution of URE3-BP-regulated proteins to this complex process and their contribution to trophozoite virulence in different host environments [32]. However, it is clear that URE3-BP plays an important role in the virulence processes.

3.7 Genome-wide analysis of *E. histolytica* DNA damage response

The maintenance of DNA integrity is essential for cells viability and the life of organisms. The genetic material is constantly under attack from external environmental factors and endogenous metabolic products that can alter its chemical structure and nucleotides sequence. External and internal DNA damaging agents can induce changes in genome by producing single-strand breaks, double strand breaks (DSBs), inter- and intra-strand cross-links in the form of cyclobutane pyrimidine dimers and (6-4)-photoproducts, oxidation and alkylation of bases, or formation of bulky chemical adducts. Exogenous agents are represented by ionizing (IR) and ultraviolet (UV) irradiation, or chemicals, whereas cellular endogenous physiological events include restoration of collapsed replication forks in the course of DNA synthesis, telomere maintenance, or programmed events, such as V(D)J recombination and meiotic exchange [33, 34]. In order to determine changes in global mRNA expression profile in response to DNA damage in *E. histolytica*, Weber *et al.* (2008) performed a DNA microarrays study [16]. They used UV (150 J/m²) irradiated trophozoites. DNA damage and double strand-breaks were previously confirmed in irradiated cells by TUNEL and comet assays and evaluation of the EHH2AX histone phosphorylation status [35]. The overall expression analysis identified *E. histolytica* genes that may contribute to adaptation and survival of trophozoites in response to DNA damage response. Data showed that 11.6% (350 ORFs) and 17.2% (522 ORFs) of genes were modulated at 5 min and 3 h after UV irradiation, respectively. Surprisingly, most genes were less than 2-fold modulated evidencing a weak transcriptional activation. Genes encoding RAD52-epistasis group DNA repair genes (RAD54, MRE11, RAD50, and RAD52) were slightly regulated in trophozoites submitted to UV irradiation. Functional classification of regulated genes by Gene Ontology showed that they were involved in potential DNA damage response pathways, including cell cycle, signal transduction, and oxidative stress response. However, most genes were involved in pathways including RNA binding and processing, cell structure, protein synthesis and degradation, energy metabolism, adhesion and vesicle trafficking. Several hypothetical genes were also regulated [16]. Most of these genes have not been previously associated with DNA damage response in other organisms. These data provided insights about the potential roles of the *E. histolytica* RAD52 epistasis group genes and novel ORFs in DNA damage response in this ancient eukaryotic parasite.

3.8 Genomic-wide approach to understand *E. histolytica* virulence *in vivo*

A genomic transcriptional analysis of *E. histolytica* was performed in trophozoites isolated from the colon of six infected mice and from *in vitro* culture cells [36]. In this study, authors designed an Affymetrix platform gene expression array that included probe sets for 9435 ORFs. Interestingly, 523 transcripts (5.2% of all *E. histolytica* genes) were significantly modulated in trophozoites obtained from mice intestine on days 1 and 29 after infection, which represents the early and late responses to parasite colonization. Genes modulated correspond to cell signalling proteins, including transmembrane kinases, ras and rho family GTPases, and calcium binding proteins. A decrease in mRNA levels for genes involved in glycolysis and an increase in lipases were consistent with changes in energy metabolism. As expected, a decrease in oxygen detoxification pathways was also observed in the anaerobic colonic lumen. Other known virulence factors that showed significant changes include a 20–35-fold increase in a cysteine proteinase four-like gene, and a 2–3-fold decrease in two members of the Gal/GalNAc lectin light subunit family. Quantitative real-time reverse transcriptase PCR confirmation for 11/12 genes was performed. This study represents the first genome-wide analysis of the transcriptome of *E. histolytica* during infection in an animal model [36].

4. Proteomic approaches to study amoebiasis

4.1 Proteomics basics

Proteomics is the study of whole protein set obtained from a given tissue or cell type. These analyses are usually performed by two-dimensional gel electrophoresis (2-DE) followed by protein identification by mass spectroscopy (MS) and database queries. Although 2-DE dates back to 1970s [37] characterization by MS became a reality in the last 10 years with soft ionization methods, such as matrix-assisted laser-desorption ionization (MALDI) [38] and electrospray ionization (ESI) [39]. The success of this technique would not be possible without the development and availability of databases. Identification of proteins by MS uses two different approaches: in the peptide mass fingerprinting (PMF), protein spot of interest is in-gel digested with a specific enzyme [40-43] and in the second approach, peptides after in-gel digestion are fragmented in the mass spectrometer yielding partial amino acid sequences from peptides (sequence tags). PMF is usually performed by MALDI-TOF [44, 45]. In turn, 2-DE includes a proteins separation by isoelectric focusing (IEF) in an immobilized pH gradient (IPG) gel according to their isoelectric point [46], followed by a separation according to molecular weight through SDS-polyacrylamide electrophoresis. Finally, protein pattern is revealed through gel staining by Coomassie, silver or fluorescence and the pattern of spots can be analyzed quali- and quantitatively by imaging programs.

4.2 Proteomics profiles in the study of amoebiasis

Variations in the severity of *E. histolytica* infections are influenced by host immunity, host nutritional status, gut microbiota, and genetic divergence between isolates. These alternative effects are related to distinct protein expression profiles where proteomic studies have been a relevant tool to identify target molecules and propose their interactions. Despite the effort required, the significance of proteomic results is well acknowledged. Nowadays it is accepted that mRNA expression and protein levels do not necessarily correlate [47, 48]. Thus, post-translational modifications and their physiological role can be investigated by MS [49].

Studies of 2-DE for *E. histolytica* trophozoites have been developed and modified, so that pH ranges now permit to evidence more significant landmarks [50]. Because degradation of extracellular matrix by amoebic proteases, production of toxic factors, activation of host immune system cells, as well as killing and phagocytosis of human cells are features of amoebiasis, the characterization of proteins participating in these processes is of particular interest. In addition, host cells exhibit an innate immune response, which activates NF- κ B and secretes cytokines to recruit neutrophils and macrophages to the invasion site [51, 52]. During tissue inflammation, phagocytosis is an important mechanism for clearance of apoptotic cells, and a variety of molecules on the surface of the phagocytic cells participate in the adhesion of apoptotic bodies, the engulfment process, as well as their later destruction. In terms of apoptotic cells, CD14, CD68, avb3 integrin, MER (tyrosine kinase from monocytes and tissue of epithelial and reproductive origin) and PSR (phosphatidylserine receptor) promote the adhesion of dying cell to phagocyte [53, 54]. In the Jurkat human T lymphocytes cell line, activation of Caspase-3 and the exposition of phosphatidylserine from inner to the outer leaflet of the plasma membrane, as well as DNA fragmentation, take place after amoebic contact. Moreover, blocking of the formation of amoebic liver abscesses via treatment with pan-caspase inhibitor showed that apoptosis of host cells is critical for parasite virulence [55]. In particular, *E. histolytica* trophozoites can induce cell death by Caspase-8 independent pathways, but they can also destroy Caspase-8 deficient cells treated with Caspase-9.

4.3 Proteomic analysis of phagosomes in *E. histolytica*

Adhesion of *E. histolytica* to colonic mucins and host epithelial cell is essentially carried out through the Gal-GalNAc lectin. The Gal-GalNAc lectin is constituted by a heavy subunit (Hgl) and a covalent attached glycosyl phosphatidyl inositol (GPI) anchored light chain (Lgl) [56]. Proteomic analysis has already identified groups of novel proteins involved in phagosome biogenesis [57]. Particularly, 85 phagosome proteins participating in surface recognition, cytoskeleton, vesicular trafficking, and degradation were detected in primary phagosomes, but kinetics of maturation and biogenesis remained unclear [58]. Other proteomics studies with purified phagosomes or affinity-purified Gal/GalNAc binding proteins using reversed-phase capillary liquid chromatography and ion trap tandem MS allowed the identification of proteins involved in phagosome biogenesis [59-61]. Some examples include proteins acting on phagosome recruitment derived from endoplasmic reticulum and mitochondrion-like organelle, as well as proteins involved in vesicular trafficking and recruitment of hydrolytic enzymes. Likewise, via protein isolation using magnetic beads and sequencing by tandem MS, peptides from all of the lectin subunits (Hgl, Igl, and Lgl) were identified in GalNAc-BSA bead purified proteins. Other proteins were also isolated, including cytoskeletal-associated proteins such as actin and myosin, a talin homologue, homologues of Rab11-B and calreticulin, and *E. histolytica* cysteine proteinase 2. Interestingly, the most abundant protein isolated with GalNAc beads incubated with amoebic lysates was myosin. A categorization led by MS of phagosome-associated proteins proposed groups, such as lectins and surface proteins,

vesicular trafficking regulators and other small GTPases, hydrolytic enzymes and degradative proteins, and calcium and proton pumps. Predominantly, 90% of these peptides were assigned to 85 proteins.

The identification of molecules responsible for a specific phenotype can be addressed by differential expression, so that protein profile in a particular context can be contrasted with another. These studies have been conducted in *E. histolytica*, some of them have explored the protein composition of purified phagosomes. Several proteins involved in uptake-processes were recognized. Furthermore, in the same study, approximately 1,500 protein spots were detected by silver-staining and 10 landmark spots of the respective Coomassie-stained gels were identified by MALDI-TOF or protein sequencing. To construct a 2-DE reference map for *E. histolytica*, and after successful solubilisation, 100 proteins were subjected to analysis by MALDI-TOF mass spectroscopy. The resultant peptide mass fingerprints were analyzed by contrasting with data available in *E. histolytica* genome and NCBI databases. As a result, these proteins could be clustered into the following categories: a) cytoskeleton proteins; b) glycolysis; c) surface-associated proteins; d) RNA/DNA metabolism; e) ubiquitin-proteosome pathway; f) vesicular trafficking and signal transduction [62].

4.4 Proteomic analysis of trophozoites with diverse virulence degrees

In terms of virulence traits, the comparison of two genetically closely related *E. histolytica* cell lines with substantial difference in virulence were studied by 2-DE [63]. As a result of this investigation, only a very limited number of differentially synthesized proteins were detected. Preponderantly, the proteins found to be differentially synthesized are involved in cytoskeletal organization. In addition, three differentially up-regulated antioxidants were exclusively found in the pathogenic cell line, Fe-hydrogenase 2, peroxiredoxin and superoxide dismutase (SOD). Notably, only for two differentially regulated proteins, differential expression was also seen at the RNA level. Therefore, this implies that the regulation of these proteins often occurs at the posttranscriptional level.

A similar study took advantage of the ability to compare the proteome of *E. histolytica* to those of *E. dispar*, in order to investigate virulence factors of *E. histolytica* [64]. This report used *E. histolytica* HM-1:IMSS and *E. dispar* SAW760 to identify proteins that are differentially expressed between these two species. Results evidenced the relevance of the alcohol dehydrogenase 3 (EhADH3). Specially, *E. histolytica* possesses a higher amount of NADP-dependent alcohol dehydrogenase than *E. dispar* and EhADH3 can be detected in the surface of *E. histolytica*. Based on the same methodology, it has been reported that increased grainin levels may contribute to reduce virulence phenotype. This finding was the result of a comparative analysis of *E. histolytica* Rahman and *E. histolytica* HM-1:IMSS. Particularly, grainin 2 was expressed at significantly higher levels in *E. dispar* when compared to *E. histolytica*. Another protein expressed at higher levels in *E. histolytica* contained a LIM domain, which is a cysteine and histidine rich domain composed of two zinc fingers which has been related to differentiation, cytoskeleton reorganization, and regulation of transcription [65, 66]. Likewise a comparative proteomics of *E. histolytica* HM-1:IMSS and *E. histolytica* Rahman to identify virulence factors showed differential levels of two molecules linked to resistance to host oxidative defenses, peroxiredoxin, and SOD, and three proteins of unknown function, including grainins 1 and 2 [67].

5. Conclusions

The power of genomics and proteomics approaches lies in their ability to evidence gene expression networks and therefore molecular mechanisms and biochemical pathways that are related to different contexts of virulence in *E. histolytica*. Particularly, they allowed the determination of changes in gene expression level (transcripts and proteins) the study of protein- protein interactions, and the identification of post-translational epigenetic regulation, which could help to understand the pathogenesis of this parasite.

Acknowledgments This work was supported by UACM (México), CONACyT (project number 79293, México), COFAA-IPN (Mexico), SIP-IPN (Mexico) and European Community grants.

References

- [1] WHO. Entamoeba taxonomy. Bull. World Health Organ. 2007;75:291-294
- [2] Haque R, Huston CD, Hughes M, Hout E, Petri WA Jr. Amebiasis. *N Engl J Med*. 2003;348:1565-1573.
- [3] Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, Nozaki T, Suh B, et al. The genome of the protist parasite *Entamoeba histolytica*. *Nature*. 2005; 433:865–868.
- [4] Clark CG, Alsmark UC, Tazreiter M, Saito-Nakano Y, Ali V, et al. Structure and content of the *Entamoeba histolytica* genome. *Adv Parasitol*. 2007;65: 51–190.
- [5] Lorenzi HA, Puiu D, Miller JR, Brinkac LM, Amedeo P, Hall N, Caler EV. New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. *PLoS Negl Trop Dis*. 2010;4(6):e716.
- [6] Stanley SL Jr. The *Entamoeba histolytica* genome: something old, something new, something borrowed and sex too? *Trends Parasitol*. 2005;21(10):451-453.

- [7] Tillack M, Biller L, Irmer H, Freitas M, Gomes MA, Tannich E, Bruchhaus I. The *Entamoeba histolytica* genome: primary structure and expression of proteolytic enzymes. *BMC Genomics*. 2007;8:170.
- [8] Aurecochea C, Barreto A, Brestelli J, Brunk BP, Caler EV, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, Heiges M, Iodice J, Kissinger JC, Kraemer ET, Li W, Nayak V, Pennington C, Pinney DF, Pitts B, Roos DS, Srinivasamoorthy G, Stoeckert CJ Jr, Treatman C, Wang H. AmoebaDB and MicrosporidiaDB: functional genomic resources for Amoebozoa and Microsporidia species. *Nucleic Acids Res*. 2011;39(Database issue):D612-9.
- [9] Gregg Gibson and Spencer V. Muse. (2004). A primer to genome science. Second edition. ISBN 0-87893-232-1.
- [10] Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. *Science*. 1995; 270 (5235):484-487.
- [11] Debnath A, Tashker JS, Sajid M, McKerrow JH. Transcriptional and secretory responses of *Entamoeba histolytica* to mucins, epithelial cells and bacteria. *Int J Parasitol*. 2007;37:897-906.
- [12] Debnath A, Das P, Sajid M, McKerrow JH. Identification of genomic responses to collagen binding by trophozoites of *Entamoeba histolytica*. *J Infect Dis*. 2004;190:448-457.
- [13] Gilchrist CA, Houpt E, Trapaidze N, Fei Z, Crasta O, Asgharpour A, Evans C, Martino-Catt S, Baba DJ, Stroup S, Hamano S, Ehrenkaufer G, Okada M, Singh U, Nozaki T, Mann BJ, Petri WA Jr. Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. *Mol Biochem Parasitol*. 2006;147:163-176.
- [14] MacFarlane RC, Singh U. Identification of differentially expressed genes in virulent and nonvirulent *Entamoeba* species: potential implications for amebic pathogenesis. *Infect Immun*. 2006;74:340-351.
- [15] Weber C, Guigon G, Bouchier C, Frangeul L, Moreira S, Sismeiro O, Gouyette C, Mirelman D, Coppee JY, Guillen N. Stress by heat shock induces massive down regulation of genes and allows differential allelic expression of the Gal/GalNAc lectin in *Entamoeba histolytica*. *Eukaryot Cell*. 2006;5:871-875.
- [16] Weber C, Marchat LA, Guillen N, López-Camarillo C. Effects of DNA damage induced by UV irradiation on gene expression in the protozoan parasite *Entamoeba histolytica*. *Mol Biochem Parasitol*. 2009;164(2):165-169.
- [17] Ehrenkaufer GM, Haque R, Hackney JA, Eichinger DJ, Singh U. Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite. *Cell Microbiol*. 2007a;9:1426-1434.
- [18] Ehrenkaufer GM, Eichinger DJ, Singh U. Trichostatin A effects on gene expression in the protozoan parasite *Entamoeba histolytica*. *BMC Genomics*. 2007b;8: 216.
- [19] Ali IK, Ehrenkaufer GM, Hackney JA, Singh U. Growth of the protozoan parasite *Entamoeba histolytica* in 5-azacytidine has limited effects on parasite gene expression. *BMC Genomics*. 2007;8:7-11.
- [20] Davis PH, Schulze J, Stanley SL Jr. Transcriptomic comparison of two *Entamoeba histolytica* strains with defined virulence phenotypes identifies new virulence factor candidates and key differences in the expression patterns of cysteine proteases, lectin light chains, and calmodulin. *Mol Biochem Parasitol*. 2007;151:118-128.
- [21] Vicente-Joaó B, Ehrenkaufer GM, Saraiva LM, Teixeira M, Singh U. *Entamoeba histolytica* modulates a complex repertoire of novel genes in response to oxidative and nitrosative stress: implications for amebic pathogenesis. *Cell Microbiol*. 2008;11:51-69.
- [22] Santi-Rocca J, Weber C, Guigon G, Sismeiro O, Coppee JY, Guillen N. The lysine- and glutamic acid-rich protein KERP1 plays a role in *Entamoeba histolytica* liver abscess pathogenesis. *Cell Microbiol*. 2008;10: 202-217.
- [23] Biller L, Davis PH, Tillack M, Matthiesen J, Lotter H, Stanley SL Jr, Tannich E, Bruchhaus I. Differences in the transcriptome signatures of two genetically related *Entamoeba histolytica* cell lines derived from the same isolate with different pathogenic properties. *BMC Genomics*. 2010;11:63.
- [24] Gillin FD, Diamond LS. *Entamoeba histolytica* and *Giardia lamblia*: Effects of cysteine and oxygen tension on trophozoite attachment to glass and survival in culture media. *Exp Parasitol*. 52;1:9-17.
- [25] Husain A, Sato D, Jeelani G, Mi-ichi F, Ali V, Suematsu M, Soga T, Nozaki T. Metabolome analysis revealed increase in S-methylcysteine and phosphatidylisopropanolamine synthesis upon L-cysteine deprivation in the anaerobic protozoan parasite *Entamoeba histolytica*. *J Biol Chem*. 2010;285(50):39160-39170.
- [26] Husain A, Jeelani G, Sato D, Nozaki T. Global Analysis of Gene Expression in Response to L-Cysteine Deprivation in the Anaerobic Protozoan Parasite *Entamoeba histolytica*. *BMC Genomics*. 2011;12:275.
- [27] Banerjee S, Fisher S, Lohia A, Ankri S. *Entamoeba histolytica* DNA methyltransferase (EhMeth) is a nuclear matrix protein that binds EhMRS2, a DNA that includes a scaffold/matrix attachment region (S/MAR). *Mol Biochem Parasitol*. 2005, 1:91-97.
- [28] Lavi T, Siman-Tov R, Ankri S. EhMLBP is an essential constituent of the *Entamoeba histolytica* epigenetic machinery and a potential drug target. *Mol Microbiol*. 2008;69:55-66.
- [29] Ali IKM, Ehrenkaufer GM, Hackney JA, Singh U. Growth of the protozoan parasite *Entamoeba histolytica* in 5-azacytidine has limited effects on parasite gene expression. *BMC Genomics*. 2007;8:7.
- [30] Ehrenkaufer GM, Haque R, Hackney JA, Eichinger DJ, Singh U. Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite. *Cell Microbiol*. 2007b;9:1426-1444.
- [31] Ehrenkaufer GM, Hackney JA, Singh U. A developmentally regulated Myb domain protein regulates expression of a subset of stage-specific genes in *Entamoeba histolytica*. *Cell Microbiol*. 2009;11(6):898-910.
- [32] Gilchrist CA, Moore ES, Zhang Y, Bousquet CB, Lannigan JA, Mann BJ, Petri WA. Regulation of Virulence of *Entamoeba histolytica* by the URE3-BP Transcription Factor. *MBio*. 2010;18:1(1). pii: e00057-10.
- [33] Gellert M, Hesse JE, Hiom K, Melek M, Modesti M, Paull TT, Ramsden DA, van Gent DC. V(D)J recombination: links to transposition and double-strand break repair. *Cold Spring Harb Symp Quant Biol*. 1999;64:161-167.
- [34] Neale MJ, Keeney S. Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature*. 2006;442:153-158.
- [35] López-Casamichana M, Orozco E, Marchat LA, López-Camarillo C. Transcriptional profile of the homologous recombination machinery and characterization of the EhRAD51 recombinase in response to DNA damage in *Entamoeba histolytica*. *BMC Molecular Biology*. 2008;9(1):35.

- [36] Gilchrist CA, Houpt E, Trapaidze N, Fei Z, Crasta O, Asgharpour A, Evans C, Martino-Catt S, Baba DJ, Stroup S, Hamano S, Ehrenkaufer G, Okada M, Singh U, Nozaki T, Mann BJ, Petri WA Jr. Impact of intestinal colonization and invasion on the *Entamoeba histolytica* transcriptome. *Mol Biochem Parasitol.* 2006;147(2):163-176.
- [37] O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem.* 1975;250(10):4007-4021.
- [38] Karas M, Bachmann D, Bahr U, Hillenkamp F. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int J Mass Spectrom and Ion Proc.* 1987;78:53-68.
- [39] Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science.* 1989;246(4926):64-71.
- [40] Henzel WJ, Billeci TM, Stults JT, et al. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci USA.* 1993;90(11):5011-5015.
- [41] Mann M, Højrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol Mass Spectrom.* 1993;22(6):338-345.
- [42] James P, Quadroni M, Carafoli E, Gonnet G. Protein identification by mass profile fingerprinting. *Biochem Biophys Res Commun.* 1993;195(1):58-64.
- [43] Mann M, Wilm M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem.* 1994;66(24):4390-4399.
- [44] Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom.* 1994;5(11):976-989.
- [45] Bjellqvist B, Ek K, Righetti PG, et al. Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods.* 1982;6(4):317-339.
- [46] Görg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. *Proteomics.* 2004;4(12):3665-3685.
- [47] Anderson L, Seilhamer J. A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis.* 1997;18(3-4):533-537.
- [48] Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol.* 1999;19(3):1720-1730.
- [49] Görg A, Boguth G, Obermaier C, Weiss W. Two-dimensional electrophoresis of proteins in an immobilized pH 4-12 gradient. *Electrophoresis.* 1998;19(8-9):1516-1519.
- [50] Leitsch D, Radauer C, Paschinger K, et al. *Entamoeba histolytica*: analysis of the trophozoite proteome by two-dimensional polyacrylamide gel electrophoresis. *Exp Parasitol.* 2005;110(3):191-195.
- [51] Seydel KB, Li E, Zhang Z, Stanley SL Jr. Epithelial cell-initiated inflammation plays a crucial role in early tissue damage in amebic infection of human intestine. *Gastroenterology.* 1998;115(6):1446-1453.
- [52] Seydel KB, Stanley SL Jr. *Entamoeba histolytica* induces host cell death in amebic liver abscess by a non-Fas-dependent, non-tumor necrosis factor alpha-dependent pathway of apoptosis. *Infect Immun.* 1998;66(6):2980-2983.
- [53] Marion S, Guillén N. Genomic and proteomic approaches highlight phagocytosis of living and apoptotic human cells by the parasite *Entamoeba histolytica*. *Int J Parasitol.* 2006;36(2):131-139.
- [54] Fadok VA, Chimenti G. The phagocytosis of apoptotic cells. *Semin Immunol.* 2001;13(6):365-372.
- [55] Yan L, Stanley SL Jr. Blockade of caspases inhibits amebic liver abscess formation in a mouse model of disease. *Infect Immun.* 2001;69(12):7911-7914.
- [56] Cheng XJ, Hughes MA, Huston CD, et al. Intermediate subunit of the Gal/GalNAc lectin of *Entamoeba histolytica* is a member of a gene family containing multiple CXXC sequence motifs. *Infect Immun.* 2001;69(9):5892-5898.
- [57] Okada M, Huston CD, Mann BJ, et al. Proteomic analysis of phagocytosis in the enteric protozoan parasite *Entamoeba histolytica*. *Eukaryotic Cell.* 2005;4(4):827-831.
- [58] Hamon Y, Chambenoit O, Chimenti G. ABCA1 and the engulfment of apoptotic cells. *Biochem Biophys Acta.* 2002;1585(2-3):64-71.
- [59] Okada M, Huston CD, Oue M, et al. Kinetics and strain variation of phagosome proteins of *Entamoeba histolytica* by proteomic analysis. *Mol Biochem Parasitol.* 2006;145(2):171-183.
- [60] Marion S, Laurent N, Guillén N. Signalization and cytoskeleton activity through myosin IB during the early steps of phagocytosis in *Entamoeba histolytica*: a proteomic approach. *Cell Microbiol.* 2005;7(10):1504-1518.
- [61] McCoy JJ, Mann BJ. Proteomic analysis of Gal/GalNAc lectin-associated proteins in *Entamoeba histolytica*. *Exp Parasitol.* 2005;110(3):220-225.
- [62] Tolstrup J, Krause E, Tannich E, Bruchhaus I. Proteomic analysis of *Entamoeba histolytica*. *Parasitology.* 2007;134(Pt 2):289-298.
- [63] Biller L, Schmidt H, Krause E, et al. Comparison of two genetically related *Entamoeba histolytica* cell lines derived from the same isolate with different pathogenic properties. *Proteomics.* 2009;9(17):4107-4120.
- [64] Davis PH, Chen M, Zhang X, et al. Proteomic comparison of *Entamoeba histolytica* and *Entamoeba dispar* and the role of *E. histolytica* alcohol dehydrogenase 3 in virulence. *PLoS Negl Trop Dis.* 2009;3(4):e415.
- [65] Bruchhaus I, Roeder T, Lotter H, Schwerdtfeger M, Tannich E. Differential gene expression in *Entamoeba histolytica* isolated from amoebic liver abscess. *Mol Microbiol.* 2002;44(4):1063-1072.
- [66] Kadmas JL, Beckerle MC. The LIM domain: from the cytoskeleton to the nucleus. *Nat Rev Mol Cell Biol.* 2004;5(11):920-931.
- [67] Davis PH, Zhang X, Guo J, Townsend RR, Stanley SL Jr. Comparative proteomic analysis of two *Entamoeba histolytica* strains with different virulence phenotypes identifies peroxiredoxin as an important component of amoebic virulence. *Mol Microbiol.* 2006;61(6):1523-1532.

Recent Insights in Pre-mRNA 3'-End Processing Signals and Proteins in the Protozoan Parasite *Entamoeba histolytica*

César López-Camarillo^{a*}, Olga N. Hernández de la Cruz^a, Jessica García Vivas^b, Jorge Fernández Retana^b, Marisol Pezet Valdez^b, Itzel Lopez Rosas^a, Elizabeth Alvarez-Sánchez^a and Laurence A. Marchat^b

^aUniversidad Autónoma de la Ciudad de México, Posgrado en Ciencias Genómicas, México D.F., México. ^bEscuela Nacional de Medicina y Homeopatía del IPN, Programa Institucional de Biomedicina Molecular, México D.F., México

Abstract: The messenger RNA precursors (pre-mRNA) 3'-end processing occurs in a two-step co-transcriptional coupled reaction, denoted as cleavage and polyadenylation. Both processes depend on *trans*-acting factors interacting in a coordinated manner with *cis*-sequence motifs located at the 3' untranslated region of transcripts. In this paper, we reviewed mechanisms involved in pre-mRNA processing in eukaryotic organisms, including our own findings about sequences and proteins potentially involved in mRNA 3'-end formation in the protozoan parasite *Entamoeba histolytica*. Interestingly, protein sequence comparisons among *E. histolytica*, yeast, and human pre-mRNA processing machineries showed that amoeba pre-mRNA 3'-end processing machinery appears to be in an intermediate evolutionary position between mammals and yeast. In addition, the presence of non canonical poly(A) polymerases family recently identified in *E. histolytica*, adds more complexity to the mRNA 3'-end formation process in this ancient eukaryote.

Keywords: *Entamoeba histolytica*, amoebiasis, posttranscriptional regulation, pre-mRNA processing, polyadenylation.

INTRODUCTION

Eukaryotic messenger RNA precursors (pre-mRNA) undergo extensive co-transcriptional processing in the nucleus before they can be translated into proteins in cytoplasm. These modifications include the addition of a 7-methylguanosine at the 5'-end (capping), introns removal (splicing) and cleavage/polyadenylation of 3'-untranslated region (3'-UTR). Although the molecular machineries involved in these events were separately identified, it is now well known that these processes are intimately coupled, as well as with mRNA transcription and translation [1,2]. The nuclear polyadenylation of transcripts is a complex reaction that involves pre-mRNA 3'-end cleavage followed by adenosine residues addition. Both processes depend on *trans*-acting factors interacting in a coordinated way with *cis*-sequence motifs in pre-mRNA 3'-UTR. First, the primary transcript is cleaved at the poly(A) site. Then, several adenosine residues are added to RNA 3'-end to form a poly(A) tail, producing mature mRNA that can be translated [1]. Studies performed in yeast and higher eukaryotic cells have demonstrated that 3'-end processing is also required for the efficient export of transcripts from the nucleus to the cytoplasm [3]. Notably, the poly(A) tail and the subsequent binding of poly(A) binding protein (PABP) contribute to mRNA stability, protecting the 3'-end from the attack of ribonucleases [4]. Moreover, PABP interaction with the 5'-cap binding proteins promotes an efficient translation [5,6]. It has been shown that 3'-end processing factors are physically associated to the nuclear cap-binding complex (CBC) formed by CBP20 and CBP80 proteins [7] as well as

the 65 kDa subunit of the U2 snRNP auxiliary factor (U2AF) complex at an early step of spliceosome assembly [8], providing an unsuspected functional link between capping, splicing and polyadenylation reactions. In this work, we focused on the review of mechanisms involved in the formation of the poly(A) tail at pre-mRNA 3'-end described in eukaryotic organisms, and in our own findings about RNA sequences and proteins potentially involved in pre-mRNA polyadenylation in the protozoan parasite *Entamoeba histolytica*.

SEQUENCES ELEMENTS IN THE 3'-UTR THAT DIRECT PRE-mRNA 3'-END PROCESSING IN MAMMALS AND YEAST

In mammal cells, the cleavage and polyadenylation site (poly(A) site) is generally denoted by the CA dinucleotide [9], which is flanked by the hexameric A(A/U)UAAA polyadenylation signal (PAS) and the GU/U-rich downstream element (DSE) that are located 10-30 nucleotides (nt) upstream to 10-30 nucleotides (nt) downstream of the poly(A) site, respectively. Additional elements without a consensus sequence have also been associated with poly(A) site, including U-rich (UUUU, UGUA or UAUA) and G-rich upstream elements, as well as the G-rich downstream element [10, 11].

The *cis*-elements described in yeast are less conserved; furthermore they exhibit significant differences in sequence and location from higher eukaryotes. The A-rich positioning element (PE), whose more efficient sequence is AA(U/A)AAA, is the counterpart of the mammalian PAS [12]. The PE sequence is preceded by the AU-rich efficiency element (EE) that is mainly represented by the UAUAUA sequence as demonstrated by a large scale analysis of 3'-UTR [13]. The cleavage site that generally contains a pyrimidine fol-

*Address correspondence to this author at the Universidad Autónoma de la Ciudad de México, Posgrado en Ciencias Genómicas, México D.F., México; Tel: (52)+55 58-50-19-01; Ext: 15307; E-mail: genomicas@yahoo.com.mx

lowed by multiple adenosines is flanked by upstream (UUE) and downstream (DUE) U-rich elements [14].

THE PRE-mRNA 3'-END PROCESSING NUCLEAR MACHINERY IN EUKARYOTES

The mammalian pre-mRNA 3'-end processing complexes include the cleavage and polyadenylation specificity factor (CPSF), cleavage stimulating factor (CstF), cleavage factor I (CF Im) and II (CF IIm), poly(A) polymerase (PAP), poly(A) binding protein (PABP), symplekin and carboxy-terminal domain of the RNA polymerase II largest subunit (RNA pol II CTD). During pre-mRNA synthesis, CPSF and CstF polyadenylation factors are associated to elongating RNA pol II, together with the transcriptional positive coactivator PC4. When pre-mRNA 3'-UTR polyadenylation signals are transcribed, CPSF, CstF and PC4 factors dissociate from holoenzyme, which is subsequently destabilized, allowing transcription termination. All polyadenylation factors, except PABP, are required for *in vitro* pre-mRNA cleavage, whereas CPSF, PAP and PABP are sufficient for polyadenylation [1].

CPSF contains five subunits, CPSF-160, CPSF-100, CPSF-73, CPSF-30 and hFip1. The 160 kDa subunit recognizes the consensus PAS [15]. It also interacts with TFIID and RNA pol II CTD, playing an exciting role in transcription initiation, elongation and termination [16]. CPSF-73 contains the metallo- β -lactamase domain previously identified in β -CASP proteins that are involved in nucleic acid processing [17]. Biochemical and structural studies evidenced that CPSF-73 directly binds to the poly(A) site and possesses a Ca^{2+} -dependent ribonuclease activity, which suggests that it could be the endoribonuclease responsible for the pre-mRNA cleavage. hFip1 binds to CPSF-160, CPSF-73 and CstF-77; it also interacts with PAP to tether it to RNA [18]. Symplekin that is often considered as a component of CPSF is thought to facilitate the recruitment of CPSF and CstF to RNA molecule [19].

The trimeric CstF binds to the GU/U-rich DSE through the N-terminal RNA recognition motif (RRM) of its 64-kDa subunit [20] that also interacts with symplekin [19] and PC4 bound to RNA pol II CTD. CstF-77 has a HAT domain that seems to mediate homodimer formation, as well as interactions with CPSF-160 and RNA pol II CTD [21,22]. CstF-50 binds to RNA pol II CTD with a higher efficiency; moreover gene silencing experiments using RNA interference showed that it also interacts with the splicing co-activator SRm160, establishing another link between mRNA 3'-end processing and splicing [23].

CF Im functions as a heterodimer that is composed by the 25-kDa subunit and one of the 59-, 68- or 72-kDa subunits, which share high sequence similarity. The RNP-type binding domain at CF Im-68 N-terminus is essential for CF Im-25 binding, whereas the C-terminus interacts with spliceosomal SR proteins [24,25]. CF Im bound to hFip1 and PAP, interacts with the UGUAA sequence that is located upstream PAS, suggesting that it may be participating in the determination of the correct poly(A) site [26,27]. CF IIm contains hPcf11 and hClp1p subunits. The same region of hPcf11 interacts with the phosphorylated RNA pol II CTD and RNA, which probably contributes to the release of 3'-end

processing factors from transcriptional machinery [28]. In addition, hPcf11 also interacts with hClp1p that tethers CPSF with CF Im [29].

After pre-mRNA 3'-end cleavage, the poly(A) tail synthesis in the nucleus is catalyzed by the nuclear PAP that belongs to the DNA polymerases β -like nucleotidyl-transferase superfamily. PAP presents a tripartite domain organization that includes the catalytic domain with three invariant aspartate residues involved in nucleotide transfer and a highly conserved F/YGS motif responsible for ATP binding. The PAP C-terminal end contains a RNA-binding domain (RBD) and binds hFip1 and CPSF during early poly(A) tail addition reaction [30]. Then, PABP binds to the nascent poly(A) tract increasing the efficiency of polyadenylation [31].

Most mammalian pre-mRNA 3'-end processing factors have homologues in *Saccharomyces cerevisiae*, suggesting that mRNA 3'-end processing is conserved through evolutionary scale. CPSF corresponds to the cleavage and polyadenylation factor (CPF) in yeast, which can be separated into two subcomplexes: the cleavage factor II (CF II) contains Cft1p/Yhh1p, Cft2p/Ydh1p, Brr5p/Ysh1p and Pta1p that are homologous to CPSF-160, CPSF-100, CPSF-73 and symplekin, respectively, as well as the poly(A) polymerase Pap1p; whereas the polyadenylation factor I (PF I) is formed by Yth1p and Fip1p that are homologous to CPSF-30 and hFip1, respectively. It also contains the Pf2sp factor without homologue in mammals. CstF and CF IIm correspond to the Cleavage Factor IA (CF IA) containing Rna14p, Rna15p, Pcf11p and Clp1p that are homologous to CstF-77, CstF-64, hPcf11p and hClp1p, respectively. Particularly, Cft1p/Yhh1p and Rna15p bind to cleavage site and A-rich PE, respectively, whereas their mammals homologues, CPSF-160 and CstF-64 are associated with PAS and GU-rich DSE, respectively. The yeast machinery also includes a PABP homologue known as Pab1p. Notably, yeast lacks homologues for CstF-50 and CF IIm subunits. On the other hand, CPF has many additional proteins, such as Hrp1p that forms the cleavage factor IB (CF IB), as well as Mpel1p, Ref2p, Syc1p, Swd2p, Ssu72p, Glc7p and Pti1p, which are thought to play a regulatory function [1]. Thus, CF IA, CF IB and CF II are required for cleavage, whereas CPF, CF IA, CF IB and Pab1p are necessary for polyadenylation.

MOLECULAR EVENTS OF CYTOPLASMIC POLYADENYLATION IN EUKARYOTES

Recently, it has been demonstrated that new rounds of mRNA polyadenylation occur in cytoplasm [32]. This readenylation process is performed by a novel family of cytoplasmic non canonical PAPs known as TRF4, GLD-2 and Cid proteins in human and yeast [33-36]. Intriguingly, cytoplasmic PAPs do not share high sequence homology with canonical PAP; they have the conserved central domain with the amino acids required for ATP binding and catalytic activity, but lack the C-terminal RBD. Indeed, the interaction with GLD3/Airp1-2 proteins containing RBD reconstitutes the tripartite domain organization of canonical PAPs to restore catalytic activity [33,34]. Although they share some characteristics, nuclear and cytoplasmic polyadenylation reactions are readily distinct: i) cytoplasmic CPSF complex

lacks the 70 kDa subunit [37,38]; ii) the cytoplasmic reaction is not associated with any cleavage reaction; iii) it is restricted to a few species of transcripts containing the cytoplasmic polyadenylation (CPE) motif; and finally, iv) the cytoplasmic PAP is activated in mitosis, whereas the nuclear PAP is inhibited in this phase [39].

The cytoplasmic polyadenylation of mRNA was initially studied in *Xenopus*, *Mus musculus* and *Drosophila* development [32]. It requires the participation of two *cis*-regulatory sequences: PAS, which is also essential for nuclear polyadenylation, and the CPE motif. The phosphorylation of the cytoplasmic polyadenylation element binding protein (CPEB), a zinc finger protein with an RRM, enhances its affinity to CPE and the cytoplasmic form of CPSF-160. This interaction stabilizes CPSF binding to PAS and induces the recruitment of a cytoplasmic non canonical PAP to initiate the poly(A) tail elongation [40]. The co-immunolocalization of CstF-77, CPSF-100 and eIF4E in punctuate cytoplasmic foci with CPBE and PAP, suggested that they are also involved in cytoplasmic polyadenylation [41].

THE PROTOZOAN PARASITE *ENTAMOEBA HISTOLYTICA*

E. histolytica is the protozoan parasite causative of human amoebiasis, which has a world-wide distribution with a higher prevalence in developing countries, affecting more than 50 million people each year. The parasite causes intestinal dysentery and hepatic abscesses that result in 70,000-100,000 deaths a year, making it a leading cause of parasitic death in humans [42]. Some virulence factors have been identified as key factors in pathogenesis, however most molecular mechanisms that are relevant for infection establishment are still unknown and they are under extensive investigation [43].

The *E. histolytica* genome sequence has already been published [44]. The parasite genome assembly is 23.7 million base pairs in size, which predicts 9,938 genes that have an average size of 1.17 kb. The gene density is around 2.4 kb/gene and the average length of predicted proteins is ~389 amino acids [45]. Most genes only comprise a single exon; however as many as 25% of genes may be potentially spliced and 6% contains two or more introns. *E. histolytica* genes are usually short, mainly due to the loss of introns and shortened 5'- and 3'-UTRs. Repeat sequences consisting of long and short interspersed repetitive elements account for 6% of the genome. Genome sequence analysis revealed a variety of metabolic adaptations, including reduction or elimination of mitochondrial metabolic pathways and the use of oxidative stress enzymes associated with anaerobic prokaryotes. Moreover, evidence for lateral gene transfer of bacterial genes into the *E. histolytica* genome has been provided [44].

The knowledge of parasite genome sequence is an extraordinary tool to identify and functionally characterize novel genes that are important for pathogen survival and host infection. For instance, data mining of genome sequence allowed the detection of a large number of a variety of gene families associated with virulence, including cysteine, metallo-proteinases and amoebapores. Particularly, it also made possible the identification of conserved proteins and *cis*-

regulatory elements that are relevant for pre-mRNA 3'-end processing in this primitive eukaryote.

PUTATIVE SEQUENCES ELEMENTS INVOLVED IN PRE-mRNA 3'-END PROCESSING IN *E. HISTOLYTICA*

As mentioned above, eukaryotic poly(A) tail formation involves multiprotein complexes that interact with specific sequences at 3'-UTR of transcripts. Multiple alignment of a limited number of *E. histolytica* genomic and cDNA 3'-ends allowed the detection of two putative poly(A) sequences: the PAS (UA(A/U)UU) previously reported by Bruchhaus *et al.*, [46] and an U-rich motif downstream the poly(A) site [47]. Extensive computational analysis of a larger EST and genomic sequences collection confirmed that the region between the stop codon and the poly(A) site contains a AU-rich domain corresponding to the consensus UA(A/U)UU PAS or variants, which is followed by a novel U-rich region located before the poly(A) site. Data also evidenced the presence of a novel A-rich element that is located after the poly(A) site. This putative array was validated through the analysis of genomic sequences and predicted mRNA folding of genes with known poly(A) sites [48]. The fact that PAS was included in the stop codon in several genes may support the hypothesis that the stop codon could represent the PAS ancestor. The presence of multiple PAS sequences also suggested that the use of distinct signals could be an important regulatory event for *E. histolytica* gene expression. Interestingly, PAS is the conserved A(A/U)UAAA hexanucleotide in plants, animals and yeast, whereas it appears to be species-specific in the few protozoan parasites studied, corresponding to AGU(A/G)AAA in *Giardia lamblia* [49], UAAA in *T. vaginalis* [50], AUUAAA in the *Plasmodium* pgs28 gene [51] and UA(A/U)UU in *E. histolytica* (See Fig. (1)). These data showed that the molecular organization of pre-mRNA 3'-end processing signals is roughly conserved through evolutionary scale, whereas the polyadenylation signal seems to be species-specific in protozoan parasites and the novel A-rich element is unique for the primitive eukaryote *E. histolytica*.

THE PRE-mRNA 3'-END PROCESSING MACHINERY IN *E. HISTOLYTICA*

The *E. histolytica* genome contains 16 genes whose predicted products are homologous to members of the pre-mRNA 3'-end processing machinery described in other organisms [47]. Putative *E. histolytica* cleavage/polyadenylation factors include the four subunits of EhCPSF (160, 100, 73 and 30 kDa) with EhFIP1 and EhPAP, the three subunits of EhCstF (77, 64 and 50 kDa), the two subunits of EhCFIm (CIP1 and PCF11), and the 25 kDa subunit of CFIm, as well as additional factors, such as EhMpe1, EhPsf2, EhSsu72 and EhPC4. Notably, *E. histolytica* pre-mRNA 3'-end processing machinery does not seem to contain homologous genes coding for human Symplekin, CFIm59, and CFIm68 proteins, neither sequences related to yeast Ptalp and Hrp1p.

Interestingly, protein sequence comparison among *E. histolytica*, yeast, and human cleavage/polyadenylation factors shows little variation in their functional domains through evolutive scale and suggests that *E. histolytica* pre-mRNA

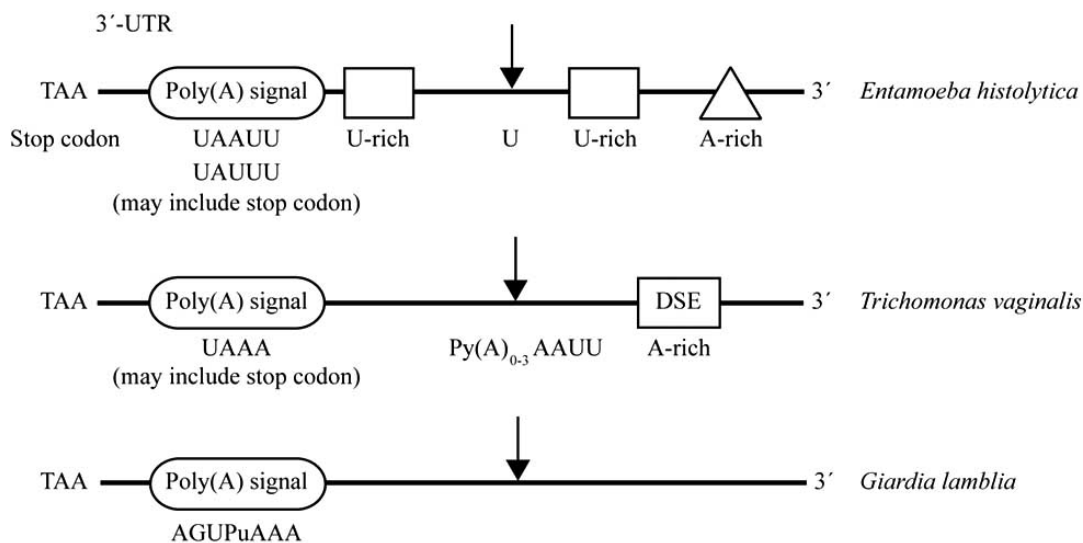


Fig. (1). Comparison of pre-mRNA 3'-UTR processing signals in *E. histolytica* and related protozoan parasites. DSE, downstream element; EE, efficiency element; FUE, far-upstream element; Arrow, cleavage site.

3'-end processing machinery appears to be in an intermediate evolutionary position between yeast and mammals (See Fig. (2)). Therefore, EhCPSF complex could bind to the polyadenylation signal UA(A/U)UU; EhCstF complex may recognize the U-rich motif and recruit EhCFIm-25, EhClp1 and EhPcf11 proteins to perform RNA cleavage at the poly(A) site. Finally, EhPAP could synthesize the poly(A) tail, in a reaction stimulated by the interaction with other processing factors. Additional interactions with EhMpe1, EhPC4/Sub1, EhSsu72, and EhPfs2 factors, could also contribute to the efficient cleavage and polyadenylation reaction [46].

CHARACTERIZATION OF THE POLY(A) POLYMERASE EhPAP OF *E. HISTOLYTICA*

The poly(A) polymerase EhPAP is thought to be responsible for poly(A) addition at *E. histolytica* pre-mRNA 3'-ends. It is a conserved 522 amino acids polypeptide, which share 25-32% identity and 43-52% similarity with yeast, protozoa, nematodes, plants and mammals nuclear PAPs [52]. Like other canonical nuclear PAPs, EhPAP presents the highly conserved PAP central catalytic domain, with the three invariant active aspartate residues (D76, D78, and D130) involved in nucleotide transfer and the F/YGS motif (63-65 aa) responsible for ATP binding in eukaryotic PAPs, which showed that EhPAP is a member of the polymerases β -like superfamily of nucleotidyl-transferases. EhPAP C-terminus contains a RNA-binding domain (RBD: 322-444 aa), which is highly divergent from the one described in homologous proteins. Moreover, no typical nuclear localization signal was detected in EhPAP, suggesting that *E. histolytica* could use other sequences for protein nuclear translocation.

Like most polymerases, 3D-structure of EhPAP was modeled as a U-shaped molecule with a tripartite domain organization [53]. Catalytic and RNA-binding domains form a groove in which the central domain is thought to orient the incoming ATP and RNA molecule for poly(A) tail extension.

EhPAP C-terminal domain shares topological similarities with RNA binding proteins that lack the conserved RNA recognition motif and bind RNA through β -sheet surfaces. Since EhPAP has not a conserved RBD, it is possible that it uses the β -sheets structures at the C-terminal domain to bind RNA. Indeed, the purified recombinant EhPAP was able to form four complexes with the 3'UTR of the *EhPgp5* gene, indicating that the lack of a conserved RBD does not greatly alter its RNA binding activity. Intriguingly, the number of complexes correlates with the number of putative polyadenylation and cleavage sites identified in the 3'UTR fragment used as probe [53], although no specific binding sequences have been reported for other PAPs.

EhPAP was expressed as a 60 kDa protein in nuclear and cytoplasmic extracts of trophozoites, as well as an additional polypeptide of 63 kDa, which could result from phosphorylation at the putative sites identified by computational analysis [52]. In immunofluorescence and microscopy laser confocal assays, EhPAP appeared as punctuate nuclear *foci* and in cytoplasmic dots in trophozoites, as it has been described for several factors involved in mRNA transcription and polyadenylation. Additionally, *EhPap* mRNA expression was modulated through cell cycle progression; RT-PCR assays showed that *EhPap* mRNA was mainly transcribed in G1 and S phases, whereas it was not detected in G2 and M phases [52]. On the other hand, other data suggested that increased EhPAP expression could contribute to the generation of the MDR phenotype in *E. histolytica* trophozoites. Apparently, emetine induces an increase in EhPAP expression, which allows the synthesis of a longer poly(A) tail to the multidrug resistance *EhPgp5* mRNA and therefore an increase in *EhPgp5* mRNA stability in trophozoites growing at higher emetine concentrations [54].

NON CANONICAL POLY(A) POLYMERASES IN *E. HISTOLYTICA*

The *E. histolytica* genome also encodes six predicted proteins that present 21-30% identity and 44-52% similarity with *H. sapiens* and *S. pombe* unconventional PAPs,

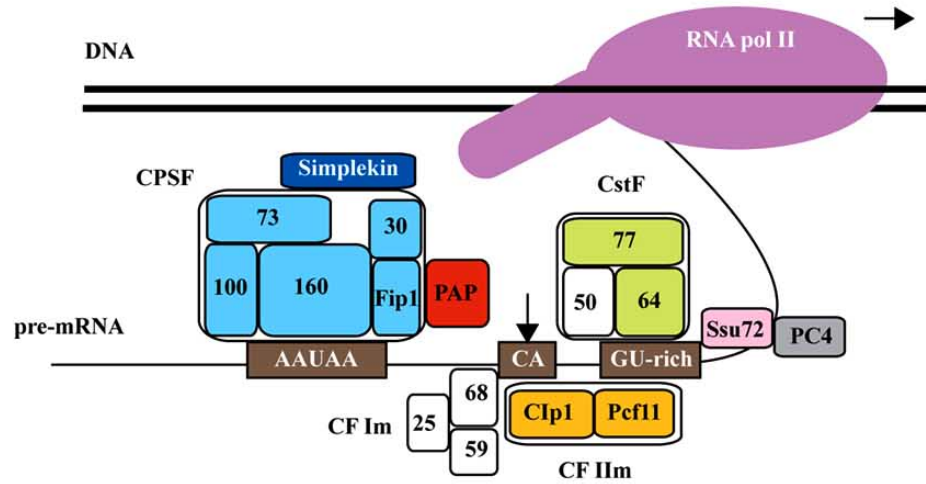
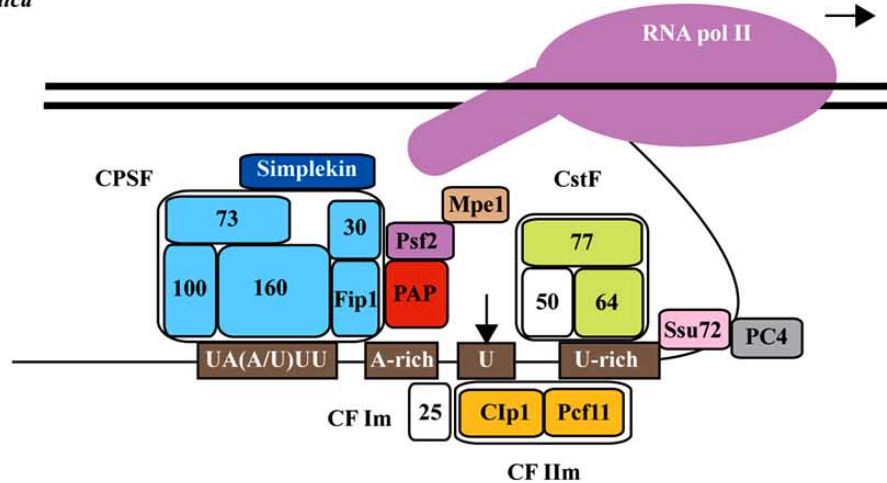
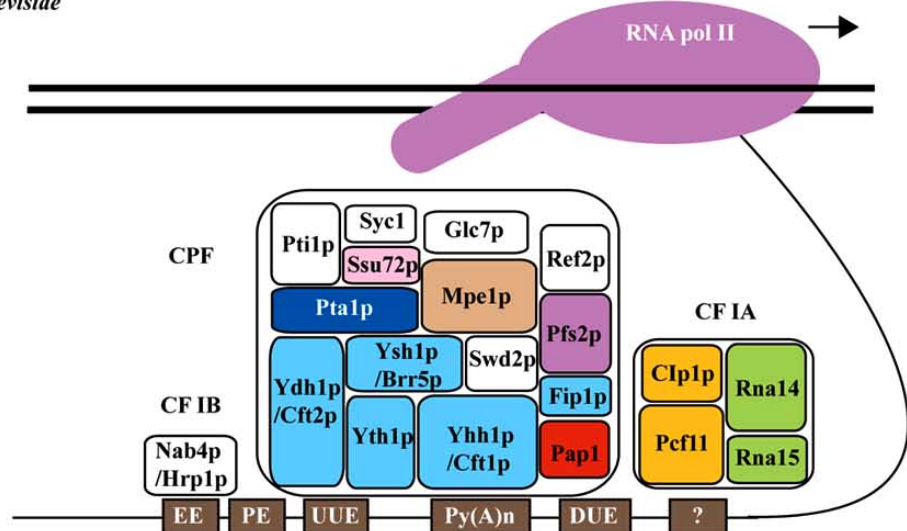
Homo sapiens*Entamoeba histolytica**Saccharomyces cerevisiae*

Fig. (2). Comparison between human, *Entamoeba histolytica* and yeast pre-mRNA processing machineries and *cis*-acting elements. Homologous proteins are represented by the same color in the three organisms. Double line corresponds to DNA molecule; thin line denotes nascent pre-mRNA. Brown boxes indicate the polyadenylation signals at 3'-UTR. Horizontal arrows indicate the sense of transcription. Vertical arrows denote the polyadenylation site.

including GLD-2, Cid11, Cid13, Cid14 and TRF-4 proteins [53]. Three proteins, named as EhGLD-2a, EhGLD-2b and EhGLD-2c, were considered as GLD-2 homologues, and three proteins, named as EhTRF4-a, EhTRF4-b and EhTRF4-c, were classified as TRF4 homologues. RT-PCR assays showed that EhTRF4-a gene is highly expressed; EhTRF4-c, EhGLD-2a and EhGLD-2c have a moderate expression; whereas EhTRF4-b and EhGLD-2b are poorly expressed. Structural and evolutionary relationship between canonical and non canonical PAPs evidenced that EhTRF4-a is highly related to EhTRF4-b, EhTRF4-c, EhGLD-2a and EhGLD-2c and presents less similarity with EhGLD-2b. As expected, *E. histolytica* unconventional PAPs do not have relationship with the nuclear EhPAP, which is highly related to other eukaryotic canonical PAP proteins, whereas the bacterial PAP remains isolated.

Notably, *E. histolytica* non canonical and canonical PAPs, as well as bacteria and other eukaryotic canonical PAPs, are all related to the minimal nucleotidyl-transferase (MNT) domain. Interestingly, this conserved functional domain was initially identified in eukaryotic members of the polymerase β -like superfamily of nucleotidyl-transferases, and then it was also detected in bacterial polymerases, suggesting that the evolution of this superfamily may have included bursts of rapid divergence linked with the emergence of new functions as well as a number of horizontal gene transfer events [55]. The presence of this ancestral bacterial MNT domain in *E. histolytica* canonical and non canonical PAPs confirm that these proteins belong to the polymerases β -like superfamily of nucleotidyl-transferases. The functional characterization of unconventional PAPs will help us to understand their biological role in *E. histolytica* gene expression regulation.

THE *E. HISTOLYTICA* CLEAVAGE FACTOR EhCFIM-25

As mentioned above, *E. histolytica* only has the 25 kDa subunit of the CF Im complex, whereas four polypeptides have been described in mammals and none in yeast [1]. The *EhCF Im-25* gene is an intron-less 768 bp sequence that encodes a 256 amino acids polypeptide, which shares 27-35% identity with homologous proteins from plants to human, including protozoan parasites. Phylogenetic inference indicated that EhCFIm-25 clusters with protozoan homologous proteins and is evolutionary close to higher eukaryote proteins. The predicted EhCF Im-25 has a nuclear localization signal at the amino terminus as well as the central NUDIX domain that has been previously described in human CF Im-25 and other Nudix proteins with distinct functions [56]. Notably, the Nudix box of EhCF Im-25 (138-161 aa) lacks three of the four glutamate residues that are important for catalytic function and metal binding, which could affect its function. Three-dimensional modeling of the EhCF Im-25 indicates that it conserves the typical fold described for human CF Im-25, including two overlapped regions involved in RNA and PAP/PABP binding. These similarities suggest that EhCF Im-25 could similarly act in nucleotide substrate binding but not hydrolysis [57]. In Western Blot assays, polyclonal antibodies raised against EhCF Im-25 recognized

a 55 kDa band in both nuclear and cytoplasmic extracts of *E. histolytica* trophozoites, suggesting that EhCF Im-25 could be interacting with another protein through a biochemical link that could not be disrupted under the reducing and denaturing electrophoresis conditions used. In addition, functional assays showed that the recombinant EhCF Im-25 was able to efficiently bind an RNA substrate confirming the *in silico* prediction (manuscript in preparation).

THE TRANSCRIPTION AND POLYADENYLATION FACTOR EhPC4 OF *E. HISTOLYTICA*

In silico analysis of the *E. histolytica* genome allowed the detection of a putative EhPC4, which exhibits high similarity and identity to human PC4 (65 and 61%, respectively) and SUB1 yeast (44 and 40%, respectively) homologues. The intron-less *Ehpc4* gene (477 bp) predicts a 151 amino acids polypeptide, which is evolutionary conserved. *Ehpc4* gene is overexpressed in virulent trophozoites of clone A, but it was not detected in non-virulent clone L6, suggesting a potential role in virulence. In Western blot assays, specific anti-EhPC4 antibodies mainly detected the endogenous protein in nuclei, as expected for a factor involved in transcriptional termination and polyadenylation. Functional assays evidenced that the recombinant EhPC4 protein has DNA-binding activity *in vitro* (manuscript in preparation). Generation and overexpression of negative-mutant EhPC4 proteins in trophozoites, in progress, will help us to define its precise role in *E. histolytica* gene expression.

CONCLUDING REMARKS

The protozoan parasite *E. histolytica* has a well conserved pre-mRNA processing machinery, which shows that general posttranscriptional gene expression regulation mechanisms are present in primitive eukaryotes like amoeba. In addition, the identification of a large poly(A) adding enzymes family that is constituted by one canonical poly(A) polymerase and six unconventional poly(A) polymerases, indicates that nuclear and cytoplasmic mRNA poly(A) addition is a regulatory event that is conserved through evolutive scale. This also suggests that mRNA poly(A) tail synthesis and mRNA turnover are important control points for translation efficiency in *E. histolytica*. The further characterization of pre-mRNA processing machinery members and non canonical PAPs would provide new insights into the molecular mechanisms involved in gene expression regulation in this protozoan parasite. The present review also highlights the useful of computer-based methods to determine sequences with a potential role in biological process. Our data will contribute to the understanding of gene expression regulation in *E. histolytica*, providing new insights into pre-mRNA 3'-end polyadenylation mechanisms in this pathogen. The functional relevance of some *E. histolytica* polyadenylation signals and proteins reviewed here is currently under experimental investigation.

ACKNOWLEDGEMENTS

This work was supported by Mexican grants from UACM, CONACyT (39888 and 59481), COFAA-IPN and SIP-IPN.

ABBREVIATIONS

3'-UTR	=	3'-untranslated region
RNA pol II CTD	=	Carboxy-terminal domain of the RNA polymerase II
CPSF	=	Cleavage and polyadenylation specificity factor
CF Im	=	Cleavage factor I
CF IIm	=	Cleavage factor II
CstF	=	Cleavage stimulating factor
pre-mRNA	=	Messenger RNA precursor
PABP	=	Poly(A) binding protein
PAS	=	Polyadenylation signal
PAP	=	Poly(A) polymerase
RBD	=	RNA-binding domain

REFERENCES

- [1] Zhao, J.; Hyman, L.; Moore, C. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.*, **1999**, *63*, 405-445.
- [2] Mandel, C.; Bai, Y.; Tong, L. 2008. Protein factors in pre-mRNA 3'-end processing. *Cell Mol. Life Sci.*, **2008**, *65*, 1099-1122.
- [3] Huang, Y.; Carmichael, G. Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell Biol.*, **1996**, *16*, 1534-1542.
- [4] Tucker, M.; Valencia-Sánchez, M.A.; Staples, R.R.; Chen, J.; Denis, C.L.; Parker, R. The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell*, **2001**, *104*, 377-386.
- [5] Tarun, S.Z Jr.; Sachs, A.B. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J.*, **1996**, *15*, 7168-7177.
- [6] Tarun, S.Z Jr.; Wells, S.E.; Deardorff, J.A.; Sachs, A.B. Translation initiation factor eIF4G mediates *in vitro* poly(A) tail-dependent translation. *Proc. Natl. Acad. Sci. USA*, **1997**, *94*, 9046-9051.
- [7] Flaherty, S.M.; Fortes, P.; Izaurralde, E.; Mattaj, I.W.; Gilmartin, G.M. Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc. Natl. Acad. Sci. USA*, **1997**, *28*, 11893-11898.
- [8] Millevoi, S.; Loulengue, C.; Dettwiler, S.; Zeineb, K.S.; Keller, W.; Antoniou, M.; Vagner, S. An interaction between U2AF 65 and CF Im links the splicing and 3' end processing machineries. *EMBO J.*, **2006**, *18*, 4854-4864.
- [9] Sheets, M.D.; Ogg, S.C.; Wickens, M.P. Point mutations in AAUAAA and the poly(A) addition site: Effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro*. *Nucleic Acid Res.*, **1990**, *18*, 5799-5805.
- [10] Beaudoin, E.; Freier, S.; Wyatt, J.R.; Claverie, J.M.; Gautheret, D. Patterns of variant polyadenylation signal usage in human genes. *Genome Res.*, **2000**, *10*, 1001-1010.
- [11] Hu, J.; Lutz, C.S.; Wilusz, J.; Tian, B. Bioinformatic identification of candidate *cis*-regulatory elements involved in human mRNA polyadenylation. *RNA*, **2005**, *11*, 1485-1493.
- [12] Guo, Z. and Sherman, F. 3'-end-forming signals of yeast mRNA. *Trends Biochem. Sci.*, **1996**, *21*, 477-481.
- [13] Guo, Z.; Russo, P.; Yun, D. F.; Butler, J. S.; Sherman, F. Redundant 3' end-forming signals for the yeast CYC1 mRNA. *Proc. Natl. Acad. Sci. USA*, **1995**, *92*, 4211-4214.
- [14] Graber, J.H.; Cantor, C.R.; Mohr, S.C.; Smith, T.F. Genomic detection of new yeast pre-mRNA 3'-end processing signals. *Nucleic Acid Res.*, **1999**, *27*, 888-894.
- [15] Keller, W.; Bienroth, S.; Lang, K.M.; Christofori, G. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J.*, **1991**, *10*, 4241-4149.
- [16] Proudfoot, N.J. New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr. Opin. Cell Biol.*, **2004**, *16*, 272-278.
- [17] Callebaut, I.; Moshous, D.; Mornon, J.P.; deVillartay, J.P. Metallo-beta-lactamase fold within nucleic acids processing enzymes: The beta-CASP family. *Nucleic Acid Res.*, **2002**, *30*, 3592-3601.
- [18] Ryan, K.; Calvo, O.; Manley, J.L. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA*, **2004**, *10*, 565-573.
- [19] Takagaki, Y.; Manley, J.L. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell Biol.*, **2000**, *20*, 1515-1525.
- [20] Takagaki, Y.; Manley, J.L. RNA recognition by the human polyadenylation factor CstF. *Mol. Cell Biol.*, **1997**, *17*, 3907-3914.
- [21] Bai, Y.; Auperin, T.C.; Chou, C.Y.; Chang, G.G.; Manley, J.L.; Tong, L. Crystal structure of murine CstF-77: Dimeric association and implications for polyadenylation of mRNA precursors. *Mol. Cell*, **2007**, *25*, 863-875.
- [22] McCracken, S.; Fong, N.; Yankulov, K.; Ballantyne, S.; Pan, G.; Greenblatt, J.; Patterson, S.D.; Wickens, M.; Bentley, D.L. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, **1997**, *385*, 357-361.
- [23] McCracken, S.; Longman, D.; Johnstone, I. L.; Caceres, J.F.; Blencowe, B.J. An evolutionarily conserved role for SRm160 in 3'-end processing that functions independently of exon junction complex formation. *J. Biol. Chem.*, **2003**, *278*, 44153-44160.
- [24] Dettwiler, S.; Aringhieri, C.; Cardinale, S.; Keller, W.; Barabino, S.M.L. Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *J. Biol. Chem.*, **2004**, *279*, 35788-35797.
- [25] Ruegsegger, U.; Blank, D.; Keller, W. Human premRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted *in vitro* from recombinant subunits. *Mol. Cell*, **1998**, *1*, 243-253.
- [26] Brown, K.M.; Gilmartin, G.M. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol. Cell*, **2003**, *12*, 1467-1476.
- [27] Venkataraman, K.; Brown, K.M.; Gilmartin, G.M. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev.*, **2005**, *19*, 1315-1327.
- [28] Hollingworth, D.; Noble, C.G.; Taylor, I. A.; Ramos, A. RNA polymerase II CTD phosphopeptides compete with RNA for the interaction with Pcf11. *RNA*, **2006**, *12*, 555-560.
- [29] de Vries, H.; Ruegsegger, U.; Hubner, W.; Friedlein, A.; Langen, H.; Keller, W. Human pre-mRNA cleavage factor IIm contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J.*, **2000**, *19*, 5895-5904.
- [30] Martin, G.; Keller, M.; Doublet, S. Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J.*, **2000**, *19*, 4193-4203.
- [31] Kerwitz, Y.; Kuhn, U.; Lilie, H.; Knoth, A.; Scheuermann, T.; Friedrich, H.; Schwarz, E.; Wahle, E. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J.*, **2003**, *22*, 3705-3714.
- [32] Richter, J.D. Cytoplasmic polyadenylation in development and beyond. *Microbiol. Mol. Biol. Rev.*, **1999**, *63*, 446-456.
- [33] Wang, L.; Eckmann, C.R.; Kadyk, L.C.; Wickens, M.; Kimble, J. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. *Nature*, **2002**, *419*, 312-316.
- [34] Kwak, J.E.; Wang, L.; Ballantyne, S.; Kimble, J.; Wickens, M. Mammalian GLD-2 homologs are poly(A) polymerases. *Proc. Natl. Acad. Sci. USA*, **2004**, *101*, 4407-4412.
- [35] Saitoh, S.; Chabes, A.; McDonald, W.H.; Thelander, L.; Yates, J.R.; Russell, P. Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. *Cell*, **2002**, *109*, 563-573.
- [36] Haracska, L.; Johnson, R.E.; Prakash, L.; Parcash, S. Trf4 and Trf5 proteins of *Saccharomyces cerevisiae* exhibit Poly(A) RNA polymerase activity but no DNA polymerase activity. *Mol. Cell Biol.*, **2005**, *25*, 10183-10189.
- [37] Stebbins-Boaz, B.; Richter, J.D. Translational control of the cleavage and polyadenylation specificity factor in *Xenopus* during early development. *Crit. Rev. Eukaryot Gene Expr.*, **1997**, *7*, 73-94.

- [38] Dickson, K.S.; Bilger, A.; Ballantyne, S.; Wickens, M.P. The cleavage and polyadenylation specificity factor in *xenopus laevis* oocytes is a cytoplasmic factor involved in regulated polyadenylation. *Mol. Cell Biol.*, **1999**, *19*, 5707-5717.
- [39] Colgan, D.F.; Murthy, K.G.; Prives, C.; Manley, J.L. Cell-cycle related regulation of poly(A) polymerase by phosphorylation. *Nature*, **1996**, *384*, 282-285.
- [40] Mendez, R.; Murthy, K.G.; Ryan, K.; Manley, J.L.; Richter, J.D. Phosphorylation of CPEB by Eg2 mediates the recruitment of CPSF into an active cytoplasmic polyadenylation complex. *Mol. Cell*, **2000**, *6*, 1253-1259.
- [41] Rouget, C.; Papin, C.; Mandart, E. Cytoplasmic CstF-77 protein belongs to a masking complex with CPEB in *Xenopus* oocytes. *J. Biol. Chem.*, **2006**, *281*, 28687-28698.
- [42] Jackson, T.F. In: *Epidemiology*. Ravdin, J.I. Ed. Amebiasis. Imperial College Press, London, **2000**, pp. 47-63.
- [43] Baxt, L.A.; Singh, U. New insights into *Entamoeba histolytica* pathogenesis. *Curr. Opin. Infect. Dis.*, **2008**, *21*, 489-494.
- [44] Loftus, B.; Anderson, I.; Davies, R.; Alsmark, U.C.; Samuelson, J. *et al.* The genome of the protist parasite *Entamoeba histolytica*. *Nature*, **2005**, *433*, 865-868.
- [45] Clark, G.C.; Alsmark, U.C.M.; Hofer, M.; Saito-Makano, Y.; Aloï, V. *et al.* Structure and content of the *entamoeba histolytica* genome. *Adv. Parasitol.*, **2007**, *65*, 51-190.
- [46] Bruchhaus, M.; Leippe, C.; Lioutas, C.; Tannich, E. Unusual gene organization in the protozoan parasite *Entamoeba histolytica*. *DNA Cell Biol.*, **1993**, *12*, 925-933.
- [47] Lopez-Camarillo, C.; Orozco, E.; Marchat, L.A. *Entamoeba histolytica*: comparative genomics of the pre-mRNA 3' end processing machinery. *Exp. Parasitol.*, **2005**, *110*, 184-190.
- [48] Zamorano, A.; López-Camarillo, C.; Orozco, E.; Weber, C.; Guillen, N.; Marchat, L.A. In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for *Entamoeba histolytica* mRNA polyadenylation. *Comput. Biol. Chem.*, **2008**, *32*, 256-263.
- [49] Peattie, D.A.; Alonso, R.A.; Hein, A.; Caulfield, J.P. Ultrastructural localization of giardins to the edges of disk microribbons of *Giardia lamblia* and the nucleotide and deduced protein sequence of alpha giardin. *J. Cell Biol.*, **1989**, *109*, 2323-2335.
- [50] Espinosa, N.; Hernandez, R.; Lopez-Griego, L.; Lopez-Villasenor, I. Separable putative polyadenylation and cleavage motifs in *Trichomonas vaginalis* mRNAs. *Gene*, **2002**, *289*, 81-86.
- [51] Cann, H.; Brown, S.V.; Oguariri, R.M.; Golightly, L.M. 3'UTR signals necessary for expression of the *Plasmodium gallinaceum* ookinete protein, Pgs28, share similarities with those of yeast and plants. *Mol. Biochem. Parasitol.*, **2004**, *137*, 239-245.
- [52] Garcia-Vivas, J.; Lopez-Camarillo, C.; Azuara-Liceaga, E.; Orozco, E.; Marchat, L.A. *Entamoeba histolytica*: cloning and expression of the poly(A) polymerase EhPAP. *Exp. Parasitol.*, **2005**, *10*, 226-232.
- [53] López-Camarillo, C.; Marchat, L.A.; Orozco, E.; Azuara-Liceaga, E.; García-Vivas, J. Canonical and non canonical poly(A) polymerases: from human to the protozoan parasite *Entamoeba histolytica*. In: *Messenger RNA Research Perspectives*. Nova Publishers Editors, USA, **2007**, pp. 125-144.
- [54] López-Camarillo, C.; Luna-Arias J.P.; Marchat L.A.; Orozco E. EhPgp5 mRNA stability is a regulatory event in the *entamoeba histolytica* MDR phenotype. *J. Biol. Chem.*, **2003**, *278*, 11273-11280.
- [55] Aravind, L.; Koonin, E.V. DNA polymerase β -like nucleotidyl-transferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res.*, **1999**, *27*, 1609-1618.
- [56] Ranatunga, W.; Hill, E.E.; Mooster, J.L.; Holbrook, E.L.; Schulze-Gahmen U.; Xu, W.; Bessman, M.J.; Brenner, S.E.; Holbrook, S.R. Structural studies of the nudix hydrolase DR1025 from *deinococcus radiodurans* and its ligand complexes. *J. Mol. Biol.*, **2004**, *339*, 103-116.
- [57] Coseno, M.; Martin, G.; Berger, C.; Gilmartin, G.; Keller, W.; Doublé S. Crystal structure of the 25 kDa subunit of human cleavage factor Im. *Nucleic Acids Res.*, **2008**, *36*, 3474-3483.

El jurado designado por la Universidad Autónoma de la Ciudad de México aprobó esta tesis el día 27 de junio del 2013, en la Ciudad de México, D.F. para optar al Grado de Doctora en Ciencias Genómicas, a la M. en C. Itzel López Rosas.

Dr. Mario César López Camarillo

Dra. María Elizbeth Álvarez Sánchez

Dra. Mavil López Casamichana

Dra. Laurence Annie Marchat

Dr. Israel López Reyes
