Parasitology Research (2019) 118:517–538
https://doi.org/10.1007/s00436-018-6170-6

HELMINTHOLOGY – ORIGINAL PAPER

Identification of the gene encoding the TATA box-binding protein-associated factor 1 (TAF1) and its putative role in the heat shock response in the protozoan parasite *Entamoeba histolytica*

Bartolo Avendaño-Borromeo[1] · Ravi Kumar Narayanasamy[1] · Guillermina García-Rivera[2] · María Luisa Labra-Barrios[1] · Anel E. Lagunes-Guillén[2] · Bibiana Munguía-Chávez[2] · Carlos Alberto Castañón-Sánchez[3] · Esther Orozco[2] · Juan Pedro Luna-Arias[1]

Received: 13 July 2018 / Accepted: 29 November 2018 / Published online: 15 December 2018

Abstract

Transcription factor IID (TFIID) is a cornerstone in the transcription initiation in eukaryotes. It is composed of TBP and approximately 14 different subunits named TBP-associated factors (TAFs). TFIID has a key role in transcription of many genes involved in cell proliferation, cell growth, cell cycle, cell cycle checkpoint, and various other processes as well. *Entamoeba histolytica*, the protozoan parasite responsible for human amoebiasis, represents a major global health concern. Our research group has previously reported the genes coding the TATA box-binding protein (EhTBP) and TBP-related factor 1 (EhTRF1), which displayed different mRNA levels in trophozoites under different stress conditions. In this work, we identified the TBP-associated factor 1 (*Ehtaf1*) gene in the *E. histolytica* genome, which possess a well-conserved DUF domain and a Bromo domain located in the middle and C-terminus of the protein, respectively. The EhTAF1-DUF domain tertiary structure is similar to the corresponding HsTAF1 DUF domain. RT-qPCR experiments with RNA isolated from trophozoites harvested at different time points of the growth curve and under different stress conditions revealed that the *Ehtaf1* gene was found slightly upregulated in the death phase of growth curve, but under heat shock stress, it was found upregulated 10 times, suggesting that *Ehtaf1* might have an important role in the heat shock stress response. We also found that EhTAF1 is expressed in the nucleus and cytoplasm at 37 °C, but under heat shock stress, it is overexpressed in both the nucleus and cytoplasm, and partially colocalized with EhHSP70 in cytoplasm.

Keywords *Entamoeba histolytica* · TFIID · TAF1 · Heat shock stress · Cloning

Introduction

Entamoeba histolytica is the protozoan parasite that causes human amoebiasis, affecting 50 million people worldwide and causing more than 100 thousand deaths annually (Ali et al. 2008). This parasite shows a great variability in its virulence due to the enormous plasticity of its genome, haploid changes, amplification of non-programmed genes, and duplication events, which may be related to gene expression regulation (Gangopadhyay et al. 1997; Das and Lohia 2002). The *E. histolytica* genome has 20 Mb and contains 8201 protein-coding genes, including those genes that are involved in the virulence of this parasite, as well as canonical genes coding for transcription factors involved in the basal transcription machinery (Loftus et al. 2005; Lorenzi et al. 2010). Studies on the structure and function of some gene promoters have allowed the discovery of several core promoter elements,

Handling Editor: Julia Walochnik

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00436-018-6170-6) contains supplementary material, which is available to authorized users.

✉ Juan Pedro Luna-Arias
jpluna@cell.cinvestav.mx; jpluna@cinvestav.mx; jpluna2003@gmail.com

[1] Departamento de Biología Celular, Cinvestav-IPN, Av. IPN 2508, Col. San Pedro Zacatenco, 07360 Ciudad de México, Mexico

[2] Departamento de Infectómica y Patogénesis Molecular, Cinvestav-IPN, Av. IPN 2508, Col. San Pedro Zacatenco, 07360 Ciudad de México, Mexico

[3] Laboratorio de Investigación Biomédica, Subdirección de Enseñanza e Investigación, Hospital Regional de Alta Especialidad de Oaxaca, Aldama S/N, San Bartolo Coyotepec, 71256 Oaxaca, Mexico

including the TATA box (5′-GTATTTAAA-3′), the INR (5′-AAAAATTCA-3′), the GAAC box (5′-AATGAACT-3′), and the GAAC-like box (5′-GAACTACAAA-3′) (Bruchhaus et al. 1993; Purdy et al. 1996; Singh et al. 1997 2002; Manna et al. 2014).

This parasite possesses two *tbp* genes, named *Ehtbp1* (Gene ID in the *E. histolytica* HM-1:IMSS genome, EHI_020610) and *Ehtbp2* (Gene ID EHI_112050), which encode the TATA box-binding proteins EhTBP1 and EhTBP2 of 234 and 212 residues in length, respectively. The only difference between EhTBP1 and EhTBP2 is a peptide of 22 residues only present in the amino-terminal domain of EhTBP1. However, *Ehtbp2* is endogenously silenced (Ali et al. 2007). Remarkably, the *E. histolytica* genome contains another gene related to *Ehtbp*, named *TBP-related factor 1* or *TBP-like protein 1* (*Ehtrf1* or *Ehtbpl1*, respectively; Gene ID EHI_112050), which has 42.6% identity and 73.7% similarity to the full-length EhTBP, being their carboxy-terminal ends the most conserved regions. These basal transcription factors have the ability to bind the TATA box, several TATA box variants, the GAAC box, and the GAAC-like box (Luna-Arias et al. 1999; de Dios-Bravo et al. 2005; Castañón-Sánchez et al. 2010; Narayanasamy et al. 2018). Interestingly, they showed differential gene expression under different stress stimuli such as serum depletion, UV irradiation, and heat shock stress, and during the interaction with mammalian cells (Narayanasamy et al. 2018). However, other molecules involved in transcription initiation have not been described in this protozoan.

In eukaryotes, transcription of protein-coding genes begins with the binding of TFIID on promoters to recruit TFIIA, TFIIB, RNA polymerase II-TFIIF, TFIIE, TFIIH, and the Mediator complex to assemble the transcription pre-initiation complex (PIC) (Hantsche and Cramer 2017). The transcription factor TFIID, a well-conserved cornerstone of the eukaryotic transcription initiation machinery that has been characterized in *Drosophila*, human, and *Saccharomyces cerevisiae* (Dynlacht et al. 1991; Kokubo et al. 1994; Brou et al. 1993; Poon and Weil 1993; Reese et al. 1994; Moqtaderi et al. 1996), is made up of TBP and 13–14 TBP-associated factors (TAFs) depending on the species (Cler et al. 2009). TBP specifically binds to the TATA box, while some TAFs are directly or indirectly linked to other elements of the core promoter, such as the initiator (INR) or other downstream promoter elements (DPEs). On the other hand, TFIIB binds to either the upstream or downstream TFIIB recognition element (BRE), which flanks the TATA box (Burley and Roeder 1996; Lagrange et al. 1998; Juven-Gershon and Kadonaga 2010). Noteworthy that some TAFs are not exclusive of TFIID, id est, they are also part of other complexes including SAGA, pCAF, and TFTC (Ogryzko et al. 1998; Wieczorek et al. 1998; Martinez et al. 1998; Nagy and Tora 2007).

Initial studies in *Drosophila* revealed the presence of a core-TFIID complex constituted by dimers of TAF4, TAF5, TAF6, TAF9, and TAF12 in vivo (Wright et al. 2006). Further studies in yeast showed a quasi-symmetrical structure of the core-TFIID complex (Papai et al. 2009) as well as in human (Bieniossek et al. 2013). The tri-lobed symmetry of the human core-TFIID is broken by the binding of the TAF8/TAF10 heterodimer forming the so-called 7TAF complex. The TAF8/TAF10 dimer is imported to the nucleus and then it assembles in the 7TAF complex. Then, single copies of TAF1, TAF2, TAF3, TAF7, TAF11, TAF13, and TBP are recruited to assemble the holo-TFIID complex (Bieniossek et al. 2013). Moreover, the TAF8/TAF10 dimer is bound by TAF2 in cytoplasm and imported to the nucleus to form the 8TAF complex before the final assembly of the holo-TFIID complex (Trowitzsch et al. 2015), revealing a high degree of complexity and putative control points in the assembly of the functional holo-TFIID complex in the nucleus.

To continue with the characterization of basal transcription factors in *E. histolytica*, we report here the cloning of the *Ehtaf1* gene, the in silico characterization of the EhTAF1 protein, and its expression pattern under different stress conditions, including heat shock stress, serum depletion, UV irradiation, and erythrophagocytosis, revealing a putative key role in the heat shock response. EhTAF1 has a well-conserved DUF domain and a Bromo domain located in the middle and C-terminus of the protein, respectively, which are shared with other species. The EhTAF1-DUF domain tertiary structure is similar to the corresponding DUF domain of HsTAF1. The aforementioned data provide evidence for the nature of the protein coded by the *Ehtaf1* gene.

Materials and methods

Entamoeba histolytica cell culture

Trophozoites of *E. histolytica* HM1:IMSS clone A strain (Orozco et al. 1983) were axenically grown and maintained in 15 mL Kimax glass tubes with 13 mL of TYI-S-33 Diamond's medium supplemented with 20% adult bovine serum (ABS, Microlab, Mexico City, Mexico) at 37 °C in a 5200 NAPCO incubator (Darwin Chambers, Sain Louis MO, USA) (Diamond et al. 1978).

Determination of the *E. histolytica* growth curve

Two hundred trophozoites were collected at the exponential growth phase and inoculated in glass tubes containing 13 mL of TYI-S-33 medium supplemented with 20% ABS and incubated at 37 °C for 12, 18, 36, 72, 96, and 120 h. Cells were harvested at the indicated times by centrifugation at 1800× rpm (652×*g*) in a 5820R centrifuge (Eppendorf, Hamburg,

Germany) at 4 °C. The viable cell number was determined in 50 μL aliquots using 0.4% trypan blue dye in 1× PBS and a Neubauer chamber. Remaining cells were used for the isolation of total RNA as described below.

Culture of *E. histolytica* trophozoites in serum depletion

Trophozoites (1×10^6) collected at the exponential growth phase were inoculated in 13 mL of TYI-S-33 culture medium with or without 20% ABS at 37 °C for 12 h (Shrimal et al. 2010). Viability of cells was determined as described above and processed for morphology determination and RNA isolation. Cells were fixed in 2.5% glutaraldehyde in 1× PBS for 15 min at room temperature (RT), washed thrice by centrifugation with 1× PBS, and observed through a light microscope (Olympus BH-2, NY, USA). Microphotographs were obtained using a Pixera Penguin 600CL digital camera (PIXERA, Santa Clara, CA, USA) attached to the microscope.

Culture of *E. histolytica* trophozoites under heat shock stress

One million trophozoites collected at the logarithmic growth phase were inoculated in 15 mL glass tubes containing 13 mL of TYI-S-33 medium supplemented with 20% ABS and cultured at 37 °C or 42 °C for 4 h (Weber et al. 2006). Then, cells were processed to measure cell viability and isolate total RNA, and observed through a light microscope as shown above.

Exposure of trophozoites to UV irradiation

Trophozoites (two million) harvested in the logarithmic growth phase were deposited on the surface of a 100 mm Petri dish and let to adhere to the dish. Then, culture medium was completely removed to expose cells to 254 nm UV light (150 J/m^2) for 8 s (Weber et al. 2009) using an 1800 Stratalinker system (Stratagene, San Diego, CA, USA). At the end of exposure, 20 mL of fresh pre-warmed TYI-S-33 medium supplemented with 20% ABS was added to the cells and incubated at 37 °C. At 5, 15, and 30 min post-UV irradiation, cells were detached by incubation on ice and processed for total RNA isolation and cell viable count, and stored it at − 80 °C until its use. Trophozoites that were not exposed to UV irradiation were used as a control.

Erythrophagocytosis

Red blood cells (RBCs) of a human donor (O Rh negative) were obtained and immediately washed thrice with sterile Alsever 1× isotonic solution (0.42% NaCl, 2% D-glucose, 0.8% sodium citrate, 0.055% citric acid), and incubated with trophozoites in the ratio 50 RBCs/trophozoite, in serum-free TYI-S-33 culture medium for 2, 5, 10, 15, and 30 min at 37 °C (Orozco et al. 1983). After incubation, non-phagocytosed RBCs were burst with cold distilled water and trophozoites harvested as described and processed to isolate the total RNA. To determine the number of ingested erythrocytes per trophozoite, cells were stained with 8.4 mM diaminobenzidine and observed through a light microscope. One hundred trophozoites were randomly selected for counting ingested RBCs.

Determination of the mRNA levels by quantitative PCR (RT-qPCR)

Total RNA isolation from trophozoites was performed with Trizol (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions, and quantified using a Nanodrop 2000c system (Thermo Scientific). Total RNA (5 μg) was incubated with DNase I (Thermo Scientific) and samples of 100 ng were used for RT-qPCR reactions with the SYBR FAST One-Step RT-qPCR kit (Kapa Biosystems, Wilmington, MA, USA) and an Eco™ Real-Time PCR System (Illumina, San Diego, CA, USA) under the following conditions: 50 °C for 30 min, initial denaturation at 95 °C for 5 min, 40 cycles of denaturation at 95 °C for 10 s, annealing at 60 °C for 30 s, and extension at 72 °C for 20 s. The expression level of the target genes was normalized with the *Eh40Ss2* gene (GenBank ID: EHI_020280) encoding the 40S ribosomal protein S2, and the relative expression of each gene was measured using the equation $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen 2001) with the v4.02 EcoStudy program (Illumina, San Diego, CA, USA). Proper controls (no reverse transcriptase and un-tempered reactions) were included in each 48-well reaction plate. The oligonucleotides used in the RT-qPCR assays are shown in Table 1. The GraphPad Prism v6.0 program was used for all statistical analyses. The analysis was performed using an unpaired *t* test, and the values of significant difference were shown as mean ± standard deviation for three biological repeats, each by triplicate.

Immunofluorescence and confocal microscopy

E. histolytica trophozoites collected from the exponential growth phase were washed twice with 1× PBS, fixed and permeabilized in cold methanol for 12 min, washed thrice with 1× PBS, and incubated in blocking solution (10% horse serum, 3% bovine serum albumin, 10 mM glycine in 1× PBS) for 1 h at RT. Cells were washed with 1× PBS, incubated with the anti-EhTAF1 N-terminus rabbit polyclonal antibody (1:150 dilution, see below) and the mouse monoclonal anti-Hsp70 (Genetex, dilution 1:200) overnight at 4 °C, washed thrice with 1× PBS, incubated with the Alexa Fluor® 488-coupled goat anti-rabbit polyclonal antibody (1:300 dilution, Invitrogen, Carlsbad, CA, USA) and Alexa Fluor 594-coupled

Table 1 Oligonucleotides used in the evaluation of gene expression and cloning of *Ehtaf1*

Oligonucelotide	Sequence	Tm (°C)
EhTAF1 (EHI_093770)	Fw 5′-AATACCACCAGTCACTGTTTC-3′	60
	Rv 5′-GCAGAAATGAAAGACGCAAATA-3′	60
Ehtbp (EHI_020610)	Fw 5′-TAGTAGGGTCATGTGATGTG-3′	60
	Rv 5′-CTTCATCTTTAGCACCTGTAAG-3′	60
Ehtrf1 (EHI_077240)	Fw 5′-TATTTCCAGGTGTTGTATATCGT-3′	60
	Rv 5′-TTGTTCTATTTCTTTCCCTCCT-3′	60
Eh40Ss2 (EHI_020280)	Fw 5′-ATTCGGAAATAGAAGAGGAGG-3′	58
	Rv 5′-ACTAATCTTCCAAGCTTGGT-3′	58
Ehrad54 (EHI_103840)	Fw 5′-GACAAGGACCTAATGGAGAAC-3′	60
	Rv 5′-GTCTTAGCTCCAGAACATGAA-3′	60
Ehhsp70 (EHI_197860)	Fw 5′-GAATGAAAGTGATGATGAGAGGAG-3′	60
	Rv 5′-GTGAAATAACCAGGAACAGAAATAAC-3′	60
Ehtaf1		
CCG1 amoeba 1S	Fw 5′-CG**GGATCC**ATGTATCAAGCTTGTCATCG-3′	58
5512AS without *Xho* I	Rv 5′-AAATCATCTTCTGTTGGAAGGTCAAAT-3′	63
5512RACE	Fw 5′-ACTGTAAGAGTTACAGATTTTGTTGAT-3′	61
CCG1 amoeba 2AS	Rv 5′-AACTGCAGTTACATACAACTCTTACTATTTTTTGA-3′	58
EhTaf1 Δ901-3492	Fw 5′-CG**GGATCC**ATGTATCAAGCTTGTCATCGT-3′	59
	Rv 5′-GGAATTCTTACTCTTCTTTATTATTTTCAGGTTCTATT-3′	60

Fw, forward; Rv, reverse; Tm, melting temperature. Endonuclease sites: *Bam* HI (bold letters); *Pst* I (underlined); *Eco* RI (double underlined)

goat anti-mouse polyclonal antibody (1:300 dilution, Invitrogen, Carlsbad, CA, USA) for 40 min at RT, washed thrice with 1× PBS, incubated with 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI, Thermo Fisher, Mexico City, Mexico) for 5 min at RT, washed once with 1× PBS, mounted on a glass slide with Vectashield antifade mounting medium (Vector Laboratories, Burlingame, CA, USA), and observed through a Carl Zeiss LSM 700 confocal microscope (Carl Zeiss, Mexico City, Mexico) from the Departamento de Infectómica y Patogénesis Molecular, Cinvestav, Mexico, using the ZEN lite 2009 software (Carl Zeiss, Mexico City, Mexico).

Immunolocalization of EhTAF1 by immunoelectron microscopy (IEM)

E. histolytica trophozoites were fixed with 4% paraformaldehyde and 0.5% glutaraldehyde in 1× PBS for 1 h at RT, washed thrice with 1× PBS and dehydrated through increasing concentrations of ethanol (50%, 70%, 90%, and 100%, embedded in the LR white acrylic resin (London Resin, Polysciences, Warrington, PA, USA), and polymerized under UV irradiation at 4 °C overnight. Sections (60 nm width) were obtained with an ultramicrotome (Ultracut E, Reichert-Jung, Wien, Austria), mounted on Formvar-covered nickel grids, blocked with 10% fetal bovine serum in 1× PBS for 1 h at RT, incubated with rabbit anti-EhTAF1 N-terminus polyclonal antibodies (undiluted), washed three times with 1× PBS and incubated with goat anti-rabbit IgG secondary antibodies (dilution 1:60) conjugated to 30 nm colloidal gold particles (TED PELLA, Redding, CA, USA) at RT for 1 h, and washed thrice with 1× PBS. Finally, sections were contrasted with a solution of 5% uranyl acetate for 18 min at RT, washed with distilled water, incubated with 133 mM lead citrate for 90 s at RT, washed with plenty of distilled water, dried, and examined under a JEM1011 Transmission Electron Microscope (JEOL, Mexico City, Mexico) from the Departamento de Infectómica y Patogénesis Molecular, Cinvestav, Mexico.

Cloning of the *Entamoeba histolytica taf1* gene (Ehtaf1)

To clone the *Ehtaf1* gene, we first amplified two DNA fragments, fragment A of 1922 bp and fragment B of 1731 bp, using the oligonucleotides CCG1 amoeba 1S (forward) and 5512AS without *Xho* I (reverse), and 5512RACE (forward) and CCG1 amoeba 2AS (reverse), respectively (Table 1), all designed using the *Ehtaf1* sequence reported for the *E. histolytica* HM-1:IMSS genome database, in 100 μL PCR reaction mixtures containing 100 ng *E. histolytica* clone A genomic DNA, 0.4 μM each oligonucleotide, 0.2 mM each dNTP, 3 mM $MgSO_4$, 1 μg/μL bovine serum albumin (BSA), and 2 Units Deep Vent High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA). PCR reactions were performed in a Mastercycler personal thermocycler (Eppendorf, Hamburg, Germany) and the conditions used

were initial denaturation step at 94 °C for 5 min, followed by 35 cycles of 94 °C for 30 s, 58 °C for 30s, 72 °C for 4 min, and a final extension step at 72 °C for 10 min. Then, DNA fragments were separated in 1% agarose gel, excised from the gel, and purified with the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). Finally, the full-length *Ehtaf1* gene was assembled by PCR using the purified A and B DNA fragments as templates, which share a common region of 133 nucleotides located at the 3′ and 5′ ends, respectively, and the oligonucleotides CCG1 amoeba 1S (forward) and CCG1 amoeba 2AS (reverse) using the PCR conditions mentioned above. The full-length *Ehtaf1* gene was then added with one A at their 3′ ends with *Taq* DNA polymerase (Invitrogen, Carlsbad, CA, USA) and dATP, cloned into the pGEM-T-Easy vector (Promega, Madison, WI, USA), and sequenced with the BigDye Terminator kit version 3.1 (Applied Biosystems, Foster City, CA, USA) in an Automated DNA Sequencer (310 Genetic Analyzer, Applied Biosystems, Foster City, CA, USA). The complete DNA sequence was deposited in GenBank with the accession number KJ420971.1 (nucleotide sequence) and AIY53222.1 (protein sequence).

To clone the first 900 bp (*EhTaf1_Δ901-3492*) corresponding to the amino terminus of EhTAF1, this fragment was amplified by PCR using the oligonucleotides shown in Table 1, using as template the plasmid containing the full-length *Ehtaf1* and the PCR conditions mentioned above, but using 2 mM $MgCl_2$ and 2 Units of Phusion High-Fidelity DNA Polymerase (New England Biolabs, MA, USA). The amplicon was then purified as described and cloned into the TOPO PCR4 plasmid (ThermoFisher Scientific, Mexico City, Mexico) according to the manufacturer's instructions. Finally, it was subcloned into the *Bam* HI and *Eco* RI restriction sites of pRSET A vector (Invitrogen, Carlsbad, CA, USA).

Expression and purification of the recombinant EhTAF1 amino terminus in *Escherichia coli*

The pRSET A/*EhTaf1_Δ901-3492* plasmid was transformed into *Escherichia coli* BL21 (DE3) pLysS (Invitrogen, Carlsbad, CA, USA) competent cells and cultured overnight at 37 °C. One colony was inoculated in 10 mL Luria-Bertani (LB) medium containing 100 μg/mL ampicillin and cultured overnight at 37 °C at 250 rpm. Cells were then inoculated in 100 mL LB medium containing 100 μg/mL ampicillin to an initial optical density (OD) at 600 nm of 0.01, incubated at 37 °C and 250 rpm until the $OD_{600\ nm}$ value was between 0.6 to 1.0, and the recombinant polypeptide production was induced with 1 mM IPTG at 37 °C for 2 h and 250 rpm. Cells were harvested by centrifugation at 4500 rpm in a 5810R Eppendorf centrifuge at 4 °C and broken by sonication using a CPX130PB Ultrasonic Processor (Cole-Palmer, Vernon Hills, IL, USA) at 60% amplitude, subsequently centrifuged to obtain the clarified extract at 13000 rpm (12,000×*g*) in a Heraeus Biofuge Fresco microcentrifuge (DJB Labcare, Buckinghamshire, England). Finally, the rEhTAF1 amino terminus (rEhTAF1-Nt) was purified under denaturing conditions by immobilized metal affinity chromatography (IMAC) using a Ni^{2+}-NTA-agarose column (Qiagen, Mexico City, Mexico) as reported (Luna-Arias et al. 1999), and its identity verified by tandem mass spectrometry analysis using an AB SCIEX 4800 MALDI-TOF/TOF Mass Analyzer (Applied Biosystems, Foster City, CA, USA) from the Laboratorio Nacional de Servicios Experimentales (LaNSE) at Cinvestav-IPN, Mexico City, Mexico.

Generation of rabbit polyclonal antibodies against the EhTAF1 amino terminus

One hundred micrograms of the purified recombinant EhTAF1 amino terminus was mixed with 10% TiterMax Gold adjuvant (Sigma, St. Louis, USA) and subcutaneously inoculated in a male New Zealand rabbit to obtain specific antiserum against the protein, which was previously bled to obtain the pre-immune serum. Three additional doses of 50 μg of the recombinant protein mixed with 10% TiterMax Gold adjuvant were inoculated every 15 days. One week after the last immunization, the rabbit was bled. Finally, antibodies were further purified by precipitation with caprylic acid (Bio-Rad, Richmond, CA, USA) (Steinbuch and Audran 1969) and kept at − 30 °C.

Western blot

Proteins separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) in a 10% SDS-PAGE gel and a Mini-PROTEAN Tetra Cell system (Bio-Rad, Mexico City, Mexico) were transferred to a nitrocellulose membrane (GE Healthcare Life Sciences, Mexico City, Mexico) in a Mini Trans-Blot Cell (Bio-Rad, Mexico City, Mexico) using Tris-glycine buffer (48 mM Tris-HCl, 39 mM glycine, and 20% methanol). Membrane was blocked with 5% skim milk in PBS-Tween (1× PBS and 0.1% Tween 20) overnight at 4 °C in gentle agitation, incubated with 0.2 μg/mL rabbit anti-6His tag polyclonal antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA) or anti-EhTAF1 polyclonal antibodies (1:2500) in 3% BSA in PBS-Tween overnight at 4 °C. Membrane was washed 10 times with PBS-Tween, incubated with anti-rabbit IgG polyclonal antibodies conjugated with horseradish peroxidase (Zymed; 1:20,000) for 2 h at RT, washed 10 times each 10 min, revealed with the ECL plus detection kit (Amersham), and exposed to a High Performance Film (GE Healthcare, Little Chalfont, UK).

Bioinformatics analysis

The *E. histolytica* genome database at the National Center for Biotechnology Information (NCBI) was searched to identify TAF1-related sequences using as query the amino acid sequence of the human TAF1 (UniProtKB/SwissProt ID P21675). TAF1 orthologous were identified with the BLAST suite in the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using the amino acid sequence of the identified EhTAF1 (Access Number C4LYG8, gene EHI_093770). Domains were identified using the Protein Sequence Analysis and Classification program (InterPro, https://www.ebi.ac.uk/interpro/), the Protein Family database (Pfam, https://pfam.xfam.org/) from the European Molecular Biology Laboratory and the European Bioinformatics Institute (EMBL-EBI), and the Database of Protein Domains, Families and Functional Sites (PROSITE, https://prosite.expasy.org/scanprosite/) from the Swiss Institute of Bioinformatics (SIB) Bioinformatics Resource Portal. Multiple sequence alignments were performed using the Clustal Omega program (https://www.ebi.ac.uk/Tools/msa/clustalo/) from the EMBL-EBI. Alignment was edited and transformed to Nexus format with the Jalview program (http://www.jalview.org/). The phylogenetic tree was built with PAUP program version 4.0b10 (Sinauer Associates) using the neighbor-joining (NJ) method. NJ tree was bootstrapped by doing 2000 replicates, which was visualized with Dendroscope program (version 3.5.9).

The prediction of the EhTAF1 protein structure was performed by homology modeling with the Modeler program at the ModWeb Protein Structure Modeling Server (ModWeb version r194, https://salilab.org/) from the Laboratory of Andrej Sali at the University of California at San Francisco, selecting the best, longest scoring model, and slow modeling processes. The query protein (only residues 367–750 from the submitted 347–1075 residues) was modeled using as a template the corresponding region of TAF1 in the human TAF1-TAF7 complex obtained at a resolution of 2.30 Å (RCSB Protein Data Bank ID 4RGW), which showed an identity of 34.211 with EhTAF1. Protein structures were visualized with UCSF Chimera (version 1.10.2, build 40686, https://www.cgl.ucsf.edu/chimera/). A comparison of the structures was done with the Match Maker module of Chimera program, using the Needleman–Wunsch algorithm and BLOSUM62 matrix for best-aligning pair of chains. A quality model evaluation was performed with the Verify 3D, ERRAT, Prove, PROCHECK, and WHATCHECK programs at the Structural and Analysis Verification Server (SAVES, v5.0) from the Web Services of the Molecular Biology Institute at the University of California, Los Angeles (https://servicesn.mbi.ucla.edu/SAVES/). The EhTAF1 homology model was deposited in The Protein Model Database from CASPUR and the Biocomputing group of the Department of Biochemical Sciences of the University of Rome "La Sapienza" (https://bioinformatics.cineca.it/PMDB/main.php) with the identifier PM0081520. To determine the electrostatic potential (EP) of the EhTAF1 and HsTAF1 molecules, first the PDB files were converted to a PQR format using the PDB2PQR server version 2.1.1 at the National Biomedical Computation Resource web site (http://nbcr-222.ucsd.edu/pdb2pqr_2.1.1/) employing the default parameters in the program (force field, PARSE; optimization of hydrogen bonding network, pH 7.0). The EP of molecular models was determined with the Adaptive Poisson-Boltzmann Solver (APBS) software following the web link mentioned above. Finally, the EP on protein surface was visualized with Chimera (setting values between − 10 (red) and + 10 (blue) kcal/mol).

Immunoprecipitation

Forty million *E. histolytica* trophozoites were harvested and washed twice with cold 1 × PBS, resuspended in 1 × PBS containing a cocktail of 1× protease inhibitors and vortex mixed to lyse trophozoites, and the protein concentration determined using the Bradford method. Rabbit anti-rEhTAF1-Nt polyclonal antibodies (21 μg) were coupled to 3 μg Dynabeads® M-270 Epoxy for 48 h at 4 °C in a shaking roller, according to the protocol suggested by the manufacturer (Life technologies, Mexico City, Mexico), magnetically concentrated and washed according to the manufacturer. Then, antibody-coupled Dynabeads were allowed to interact with 2 mg of total trophozoite extract for 15 min at RT, washed with 1 × PBS, and eluted with 20 μL 20 mM glycine at pH 2 for 2 min, and neutralized with 2 μL of 1 M Tris-HCl. Protein concentration was determined with a Nanodrop 2000c system (Thermo Scientific), using the molar extinction coefficient of BSA at 280 nm. Subsequently, we send the samples to the mass spectrometry core facilities at Cinvestav.

Mass spectrometry analysis

Samples were sequentially treated with 100 mM dithiothreitol (Sigma-Aldrich) and 300 mM iodoacetamide (Sigma-Aldrich) each in 50 mM ammonium bicarbonate, and digested in solution with trypsin (250 ng, mass spectrometry grade, Sigma-Aldrich) in 50 mM ammonium bicarbonate; resulting peptides were concentrated to 10 μL. Samples (4.5 μL) were then loaded into Symmetry C18 Trap V/M precolumn (Waters, Milford, MA; 180 μm × 20 mm, 100 Å pore size, 5 μm particle size) and desalted using as a mobile phase A, 0.1% formic acid (FA) in deionized water, and mobile phase B, 0.1% FA in acetonitrile (ACN) under the following isocratic gradient: 99.9% mobile phase A and 0.1% of mobile phase B at a flow of 5 μL/min during 3 min. Then, peptides were loaded and separated on HSS T3 C18 Column (Waters, Milford, MA; 75 μm × 150 mm, 100 Å pore size, 1.8 μm

particle size) using an UPLC ACQUITY M-Class (Waters, Milford, MA) with the same mobile phases under the following gradient: 0 min 7% B, 30.37 min 40% B, 32.03–35.34 min 85% B, 37–47 min 7% B at a flow of 400 nL/min and 45 °C. Spectra data were acquired in a mass spectrometer with electrospray ionization (ESI) and ion mobility separation (IMS) Synapt G2-Si (Waters, Milford, MA) using data-independent acquisition (DIA) approach by HDMSE mode (Waters, Milford, MA). The tune page for the ionization source was set with the following parameters: 2.75 kV in the sampler capilar, 30 V in the sampling cone, 30 V in the source offset, 70 °C for the source temperature, 0.5 Bar for the nano flow gas and 150 L/h for the purge gas flow. Two chromatograms were acquired (low and high-energy chromatograms) in positive mode in a range of m/z 50–2000 with a velocity of 0.5 scans/s. No collision energy was applied to obtain the low-energy chromatogram, while for the high-energy chromatograms, the precursor ions were fragmented in the transfer using a collision energy ramp of 19–55 V. Generated .raw files were deconvoluted and compared using ProteinLynx Global SERVER (PLGS) v 3.0.3 software (Waters, Milford, MA) against a reversed *Entamoeba histolytica* ATCC 30459/HM1:IMSS database (downloaded from UniProtKB web site, 7959 protein sequences). Workflow parameters were trypsin as a cut enzyme, one missed cleavage allowed; carbamidomethyl (C) as a fixed modification and acetyl (K), acetyl (N-term), amidation (N-term), deamidation (N, Q), oxidation (M), phosphoryl (S, T, Y) as variable modifications. Automatic peptide and fragment tolerance, minimum fragment ion matches per peptide: 2, minimum fragment ion matches per protein: 5, minimum peptide matches per protein: 1, and false discovery rate (FDR) of 4%. All identifications had ≥ 95% of reliability and Synapt G2-Si was calibrated with [Glu1]-Fibrinopeptide, [M + 2H]2 + = 785.84261 at ≤ 1 ppm.

Results

Identification of the gene coding the *Entamoeba histolytica* TBP-associated factor 1 (*Ehtaf1*)

To identify the *Ehtaf1* gene, we searched the *E. histolytica* genome using as query the *Homo sapiens* TAF1 sequence (UniProt Knowledgebase (UniProtKB) Swiss-Prot ID P21675.2). We found an open reading frame (ORF) of 3492 bp coding a polypeptide of 1163 amino acids, which showed 32% identity and 48% similarity to HsTAF1 in a region of 525 aa (Table 2). This nucleotide sequence has 100% identity to the corresponding sequence in the *E. histolytica* HM1:IMSS genome database, which was expected because clone A was isolated from the HM1:IMSS strain (Orozco et al. 1983). The EhTAF1 protein also showed 74%, 97%, and 99% identity to *E. invadens*, *E. dispar*, and *E. nuttalli* TAF1 proteins, respectively, and 28 to 38% identity to other species (Table 2), indicating conservation through evolution. EhTAF1 contains the DUF3591 domain of unknown function (Protein family PF12157.4 domain) located between residues 347 and 792, which was also found in all analyzed TAF1 sequences (Fig. 1). This domain is located in the central region of all TAF1 proteins and has an identity of 23.5 to 36% to all non-amoebic sequences (Fig. 1), and 78.2 to 99.6% identity to the *E. invadens*, *E. dispar*, and *E. nuttalli* corresponding domains. In addition, a single Bromo domain (PF00439.21) was identified in the carboxy-terminal end of TAF1 sequences from *A. thaliana*, *E. invadens*, *E. dispar*, *E. histolytica*, *E. nuttalli*, *T. vaginalis*, *D. discoideum*, *A. castellanii*, and *Schistosoma mansoni*. TAF1 from *C. elegans*, *Danio rerio*, *M. musculus*, *H. sapiens*, *D. melanogaster*, and *Anopheles darlingi* showed two Bromo domains located at their carboxy termini (Figs. 1 and 2). However, the TAF1 of *C. albicans* has no Bromo domain or it is highly divergent. The EhTAF1 Bromo domain is similar to the Bromo domain 2 (BD2) found in *C. elegans*, *D. rerio*, *H. sapiens*, *M. musculus*, *D. melanogaster*, and *A. darlingi* (Fig. 2a). It is known that the Bromo domain is involved in binding acetylated lysine residues such as those located in the N-terminal domain of histones. Residues that have been described in binding acetylated lysine are conserved, and correspond to V1532, V1537, Y1540, S1579, N1583, and Y1589 in the human TAF1 BD2 protein (UniProtKB/Swiss-Prot ID P21675.2; Conserved Domain Database cd05511, https://www.ncbi.nlm.nih.gov/cdd). These residues can also be found in the human TAF1 BD1 and correspond to V1409, V1414, Y1417, S1456, N1460, and L1466, although the last residue does not correspond to the Y found in position 1589 in the human TAF1 BD2 (Fig. 2a, asteriks). In EhTAF1, these residues correspond to V1079, I1084, Y1087, C1126, N1130, and K1136 (Fig. 2a), which are conserved residues in BD2 as observed in WebLogos (Fig. 2b). A zinc knuckle domain (PF15288.2) was also located between the DUF3591 and Bromo domains in *A. castellanii*, *S. mansoni*, *C. elegans*, *D. rerio*, *M. musculus*, *H. sapiens*, *D. melanogaster*, and *A. darlingi* sequences. Noteworthy, this domain is found in members of the FAM90A mammalian proteins and its function remains unknown. The CaTAF1 sequence also contained this domain located at its carboxy-terminal end (Fig. 1). The TBP-binding domain PF09247.7 was only identified in the amino terminus of *D. rerio*, *M. musculus*, *H. sapiens* (isoform 1), *D. melanogaster*, and *A. darlingi*. Moreover, we did not localize the protein kinase domains in all the sequences displayed in Fig. 1, that had been previously found in regions encompassing residues 1–434 and 1425–1893 in the human TAF1 isoform 1 (Dikstein et al. 1996). An additional domain was only identified in residues 576–647 of *A. thaliana* TAF1 that correspond to the ubiquitin domain IPR000626 (Fig. 1).

Table 2 Comparison of the *Entamoeba histolytica* TBP-associated factor 1 (EhTAF1; GenBank ID XP_657295.1) with TAF1s from other species

Protein	GB ID[a]	L (aa)[b]	I (%)[c]	S (%)[d]	aa compared	E-value[e]
EnTAF1	XP_008856912.1	1163	99	99	1163	0
EdTAF1	XP_001741907.1	1133	97	98	1133	0
EiTAF1	XP_004183817.1	1252	74	86	574	0
AcTAF1	XP_004338710.1	1803	34	51	540	3×10^{-68}
DdTAF1	XP_629749.1	2310	38	57	386	9×10^{-61}
MmTAF1	NP_001277658.1	1893	32	48	525	2×10^{-51}
HsTAF1	P21675.2	1872	32	48	525	3×10^{-56}
DrTAF1	NP_001038250.1	1947	32	49	526	3×10^{-51}
DmTAF1	AAB26991.2	2068	31	50	521	2×10^{-48}
AdTAF1	ETN64130.1	1916	31	49	525	1×10^{-47}
CeTAF1	NP_493426.2	1744	29	48	538	5×10^{-45}
SmTAF1	XP_018650728.1	2173	30	48	498	5×10^{-42}
AtTAF1b	NP_188534.2	1786	28	44	432	7×10^{-32}
TvTAF1	XP_001326171.1	1016	28	45	442	5×10^{-25}
CaTAF1	KHC83472.1	1261	29	51	342	2×10^{-27}

aa, amino acids; Ac, *Acantamoeba castellanii*; Ad, *Anopheles darlingi*; At, *Arabidopsis thaliana*; Ca, *Candida albicans*; Ce, *Caenorhabditis elegans*; Dd, *Dictyostelium discoideum*; Dm, *Drosophila melanogaster*; Dr, *Danio rerio*; Ed, *Entamoeba dispar*; Ei, *Entamoeba invadens*; En, *Entamoeba nuttalii*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Sc, *Schistosoma mansoni*; Tv, *Trichomonas vaginalis*

[a] GB ID, GenBank identifier

[b] L, full length

[c] I, identity

[d] S, similarity

[e] E-value, expected value

The EhTAF1 DUF domain is structurally related to the human TAF1

The tertiary structure of the EhTAF1 DUF domain obtained by homology modeling using as template the structure of TAF1 from the human TAF1-TAF7 complex revealed the presence of two main regions, an α-helix domain and a β-sheet domain. The first domain is composed of nine α-helices located in the carboxy terminus and corresponds to one half of the molecule (Fig. 3a), while the second domain has eight β-sheets in the amino terminus corresponding to the other half of the EhTAF1 DUF domain (Fig. 3a). The β-sheet domain has two structural-related regions, each composed of two parallel β-sheets and one anti-parallel β-sheet (Fig. 3a, circles), and a domain of two short anti-parallel β-sheets. This structural organization was observed in the human TAF1 DUF domain as well (Fig. 3b, circles). The evaluation of the EhTAF1 DUF domain model with VERIFY program indicated that 71.88% of the residues have an average 3D-1D score ≥ 0.2; ERRAT gave an overall quality factor of 43.0894; PROVE indicated that 5.4% of the protein atoms are buried outlier atoms; WHATCHECK indicated that the overall structure passed the analysis. Even though some programs indicated problems with the modeled structure, the Ramachadran plot for EhTAF1 DUF domain revealed that 98.7% of the residues are located in favorable and allowed regions, and only 5 residues (1.3%) were located in the outlier region, indicating in summary, that the model for EhTAF1 DUF domain has very high quality. In the case of the human TAF1 DUF domain structure (PDB 4RGW), VERIFY program indicated that 87.86% of the residues have an average 3D-1D score ≥ 0.2; ERRAT gave an overall quality factor of 90.6417 for TAF1 subunit and 89.2157 for TAF7 subunit; PROVE indicated that 1.7% of the protein atoms are buried outlier atoms; PROCHECK indicated that the structure passed the evaluation; WHATCHECK indicated that the overall structure passed the analysis. The Ramachandran plot for human TAF1-TAF7 complex indicated that 100% of residues are in favorable and allowed regions, indicating that it is an excellent model. When the EhTAF1 and HsTAF1 DUF domain structure models were superimposed, we observed very good fitting of both structures (Fig. 3c). However, differences between both models were better observed when the electrostatic potential on their surfaces was determined at pH 7.0 as described. The HsTAF1 DUF domain electrostatic potential has a higher positive value (blue region) than that of the EhTAF1

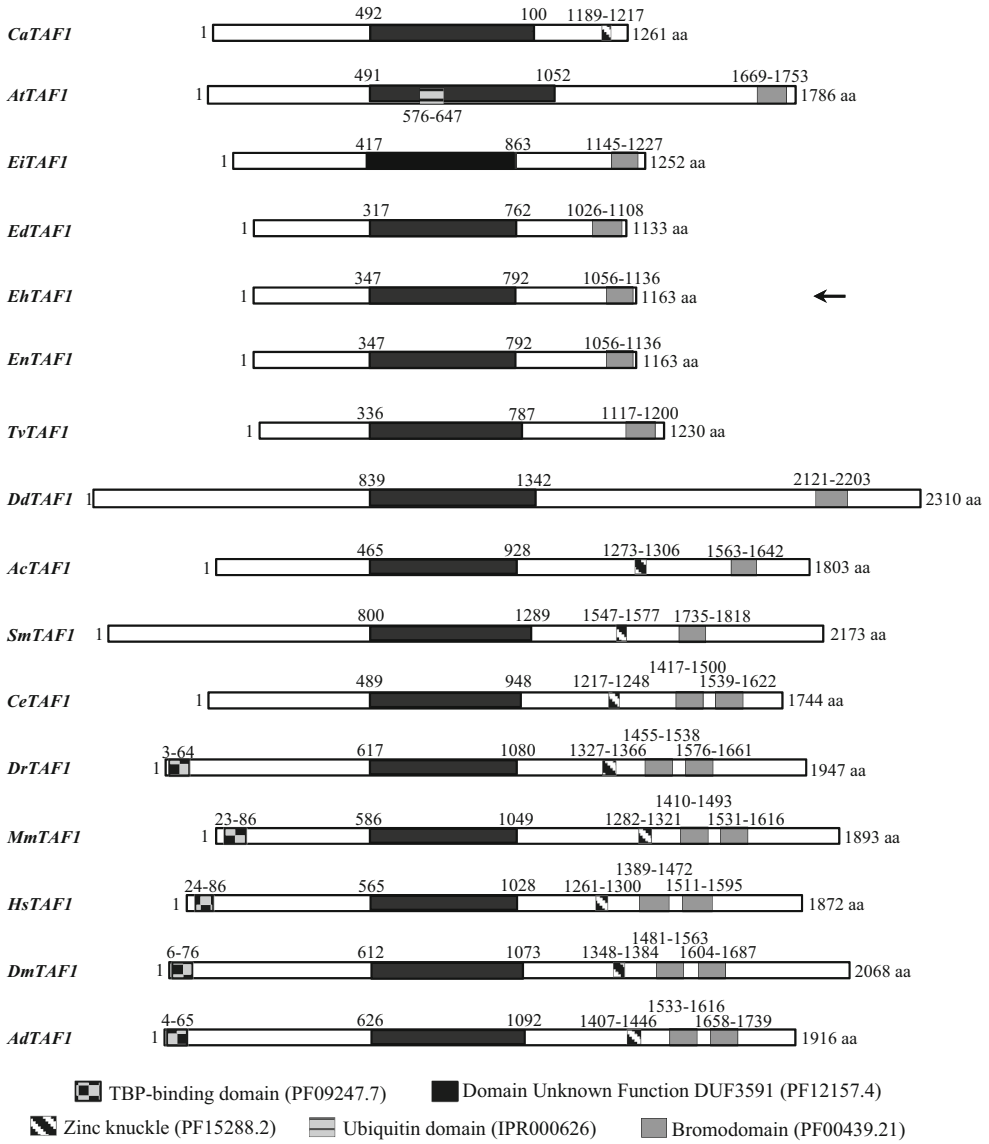


Fig. 1 Schematic representation of the domains found in TAF1 proteins aligned with the Clustal Omega program. Ac, *Acantamoeba castellanii*; Ad, *Anopheles darlingi*; At, *Arabidopsis thaliana*; Ca, *Candida albicans*; Ce, *Caenorhabditis elegans*; Dd, *Dictyostelium discoideum*; Dm, *Drosophila melanogaster*; Dr, *Danio rerio*; Ed, *Entamoeba dispar*; Eh, *Entamoeba histolytica*; Ei, *Entamoeba invadens*; En, *Entamoeba nuttalii*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Sc, *Schistosoma mansoni*; Tv, *Trichomonas vaginalis*. The complete list of GenBank Accession Numbers of TAF1 sequences used in this study is included in Table 2. Numbers above or below each box correspond to the boundaries of each domain found in the aligned TAF1 sequences. Arrow points the *Entamoeba histolytica* TAF1 protein

DUF domain (Fig. 3e, d). These differences are due to the degree of divergence of both proteins in evolution and suggest a high degree of structural conservation for these TAF1 proteins. Moreover, X-ray diffraction studies revealed that the DUF domain of human TAF1 has a conserved glycine-rich region that interacts with TAF7 in its conserved arginine-rich region (Wang et al. 2014). However, the glycine-rich region is not conserved in TAF1 proteins in protozoa. Interestingly, we have also found a gene encoding the *E. histolytica* TAF7 protein (gene EHI_137090 or protein XP_652589.1) that can have the potential of forming a TAF1-DUF domain-TAF7 dimer as reported for the human dimer (Mizzen et al. 1996; Wang et al. 2014). However, more studies are needed to test this hypothesis.

Phylogenetic analysis of the EhTAF1 polypeptide

The phylogenetic analysis of EhTAF1 located this protein in a branch containing the TAF1 proteins of *E. nuttalli*, *E. dispar*, and *E. invadens* (Fig. 4c). EhTAF1 is more related to EnTAF1, a species that is a pathogen of non-human primates, than EdTAF1, being this finding similar

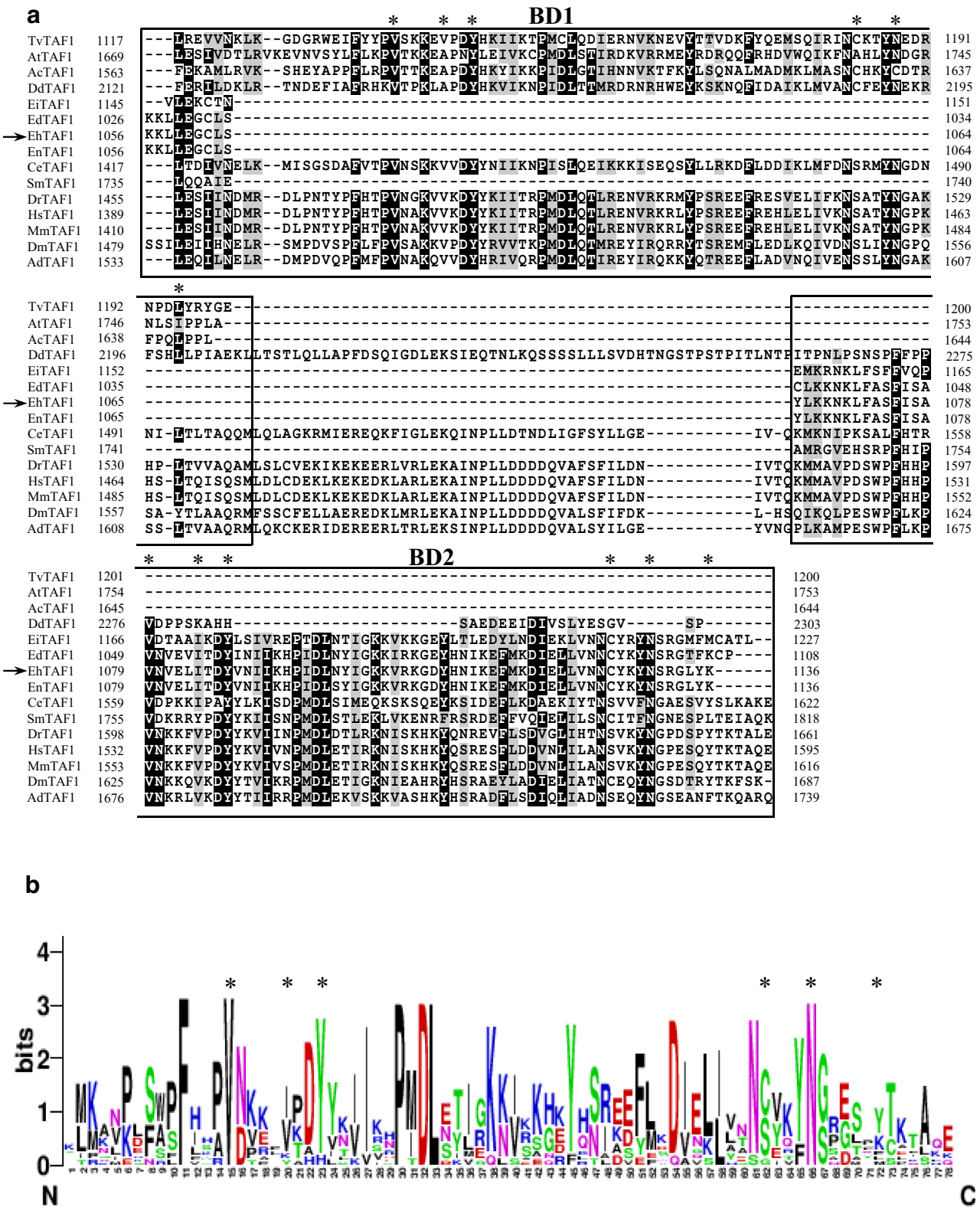


Fig. 2 Alignment of the EhTAF1 Bromo domain with TAF1 Bromo domains from other species. **a** Highly conserved regions in Bromo domains are highlighted with rectangles. Arrows points the EhTAF1 sequence. **b** WebLogos of BD2 (**a**) to show the highly conserved amino acids involved in binding acetylated residues (asterisks). BD1, Bromo domain 1; BD2, Bromo domain 2

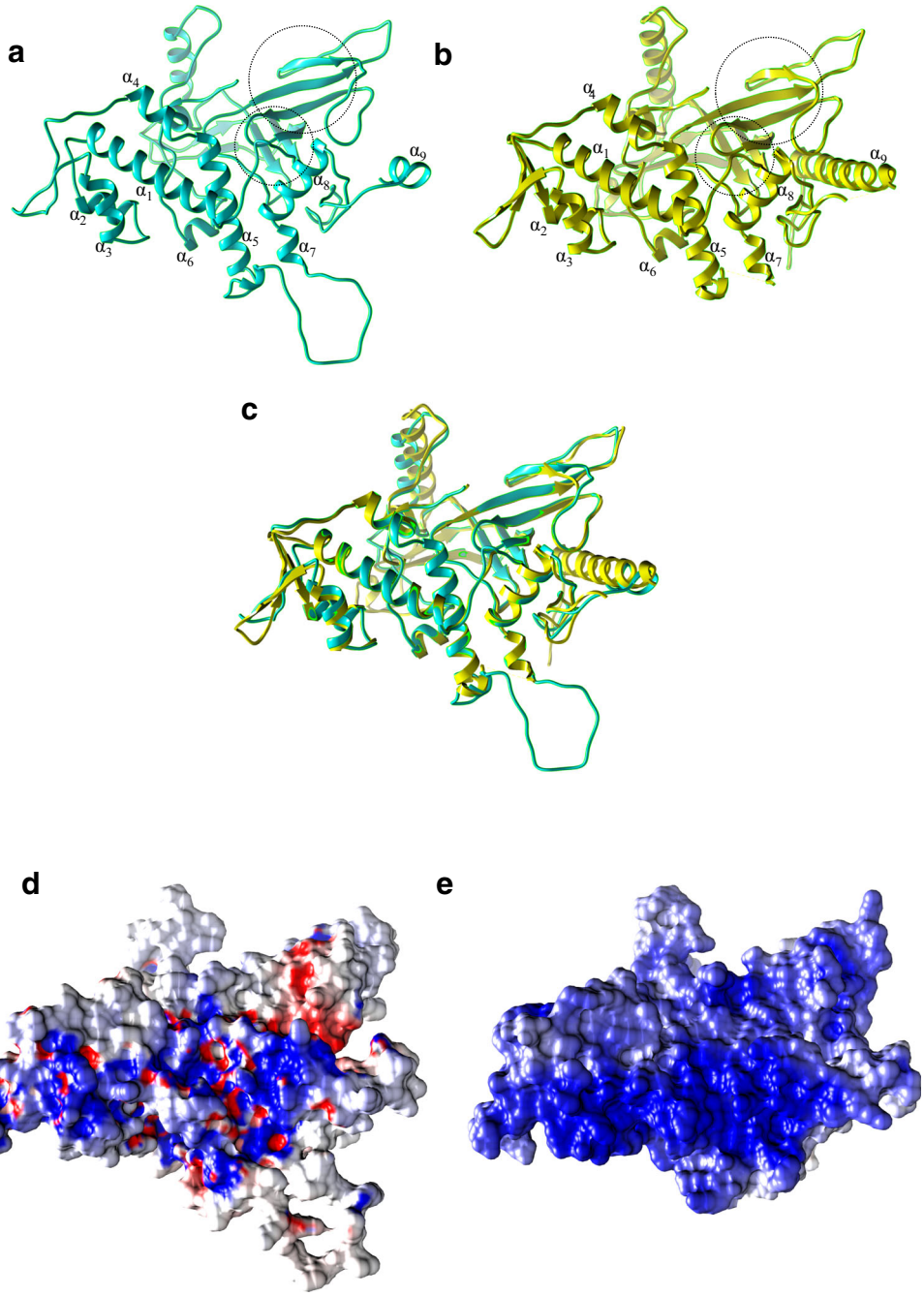


Fig. 3 Structural comparison of the molecular models of the DUF domains of EhTAF1 and HsTAF1 polypeptides. The structure of the EhTAF1 DUF domain was theoretically determined by homology modeling using the Modeler program in the ModWeb server, using as template the crystal structure of the a human TAF1-TAF7 complex obtained at a resolution of 2.3 Å (RCSB PDB 4RGW) and visualized with the Chimera program. EhTAF1 structure was only compared to the HsTAF1 subunit. **a**, **b** Cyan and yellow ribbon models for EhTAF1 and HsTAF1 proteins, respectively. The β-sheet domains are shown in circles. **c** Superimposed ribbon structure models for EhTAF1 and HsTAF1 polypeptides. **d**, **e** Electrostatic potential on the surface of EhTAF1 (**d**) and HsTAF1 (**e**) models calculated at pH 7.0

to that found by others (Tachibana et al. 2013). These species are also related to the slime mold *D. discoideum* and the ciliate *Paramecium tetraurelia*, as well as to the human parasite *T. vaginalis*, forming a branch containing members of the phyla *Ascomycota* (Fig. 4a) and *Streptophyta* (Fig. 4b). The clade described above is also evolutively related to *Acanthamoeba castellanii*. All the species mentioned above are also related to branches containing *Platyhelminthes* and *Nematoda* (Fig. 4d), *Cnidaria*, and *Arthropoda* (Fig. 4e), which all constitute a clade separated from the remaining species, including *Chordatas*, such as mammals, rodents, reptiles, and some fishes. Interestingly, *Tetraodon nigroviridis* and *Takifugu rubripes* evolutively diverged early from the species shown in tree.

The mRNA level of *Ehtaf1* increases in the death phase of growth curve of *E. histolytica* trophozoites

To determine if the *Ehtaf1* gene is being expressed in *E. histolytica* trophozoites, the expression profile of the

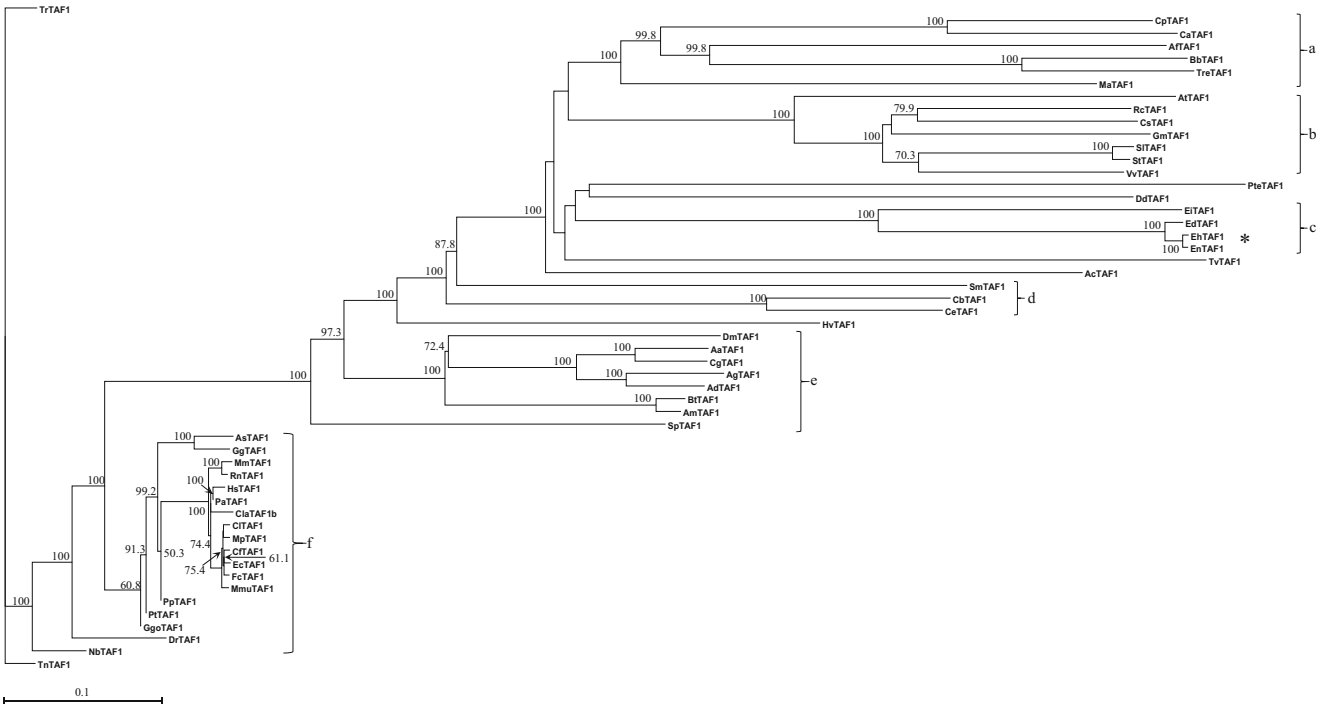


Fig. 4 Phylogenetic analysis of TAF1 proteins. Phylogenetic tree for 53 species was built using the neighbor-joining (NJ) method and bootstrapped by doing 2000 replicates as described in Materials and Methods. Values greater than 50% are only shown in the tree. Letters group species according to the phylum they belong. **a** *Ascomycota*. **b** *Streptophyta*. **c** Family of *Entamoebas*. **d** *Platyhelminthes* and *Nematoda*. **e** *Arthropoda*. **f** *Chordata*. Scale bar represents the number of substitutions per site. Please refer to electronic Supplementary Table 1 to see the complete list of abbreviations used and GenBank Accession Numbers of TAF1 protein sequences

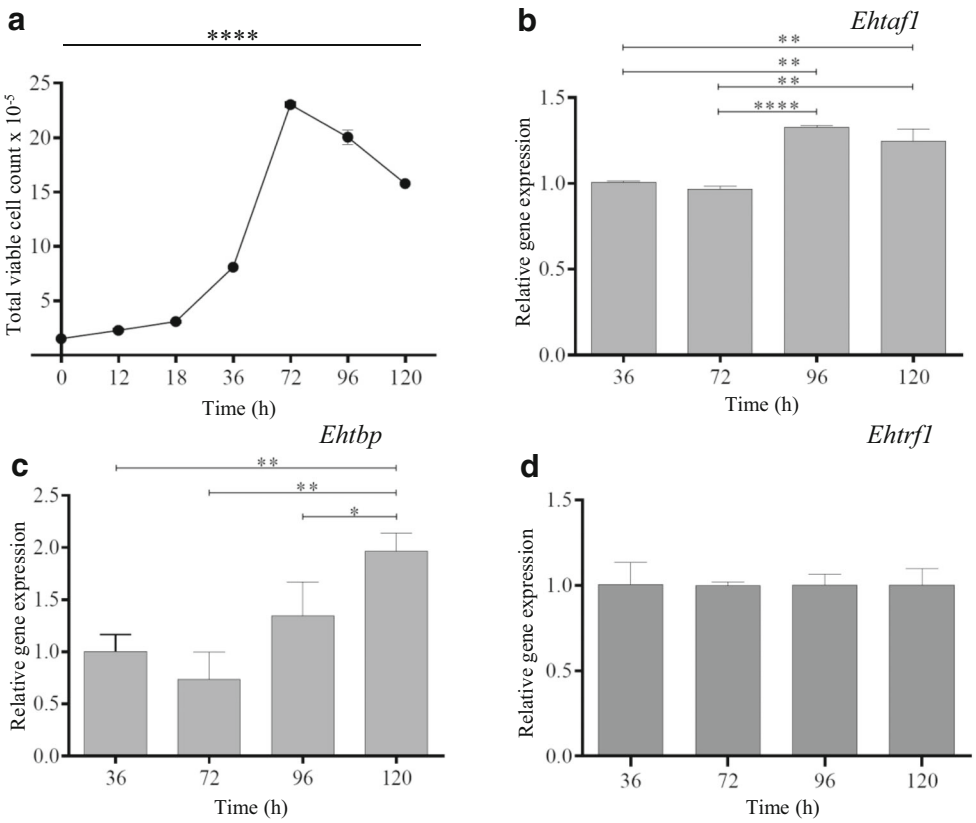


Fig. 5 The mRNA expression level of *Ehtaf1* slightly increases during nutrient deprivation. **a** Growth curve of trophozoites. Viable cell count was determined at 12, 18, 36, 72, 96, and 120 h. Relative mRNA levels determined by RT-qPCR for *Ehtaf1* (**b**), *Ehtbp* (**c**), and *Ehtrf1* (**d**) at different time points of the growth curve. The *Eh40Ss2* gene was used as a control in the analysis. Statistical unpaired *t* test: $*P < 0.05$, $**P < 0.01$. All values obtained in the growth curve were statistically significant ($****P < 0.0001$)

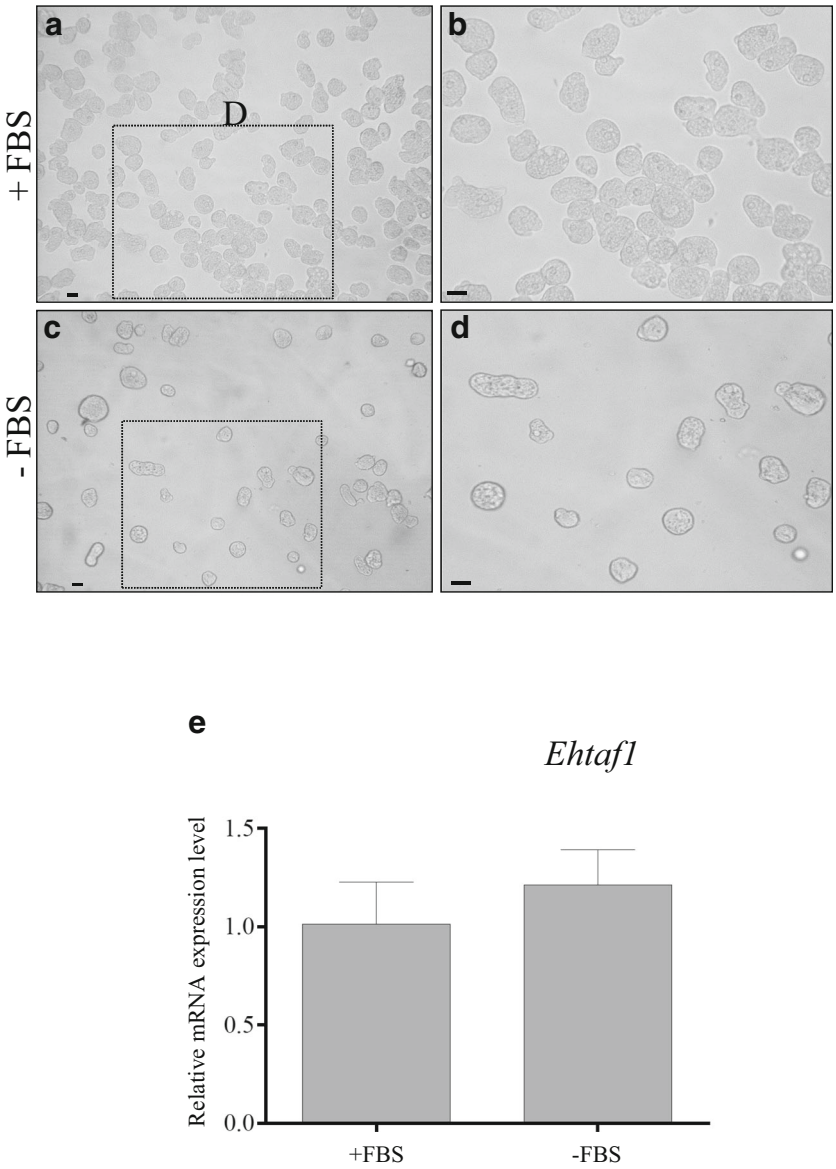


Fig. 6 The mRNA level of *Ehtaf1* does not change in serum depletion. *E. histolytica* trophozoites were grown in TYI-S-33 medium for 12 h at 37 °C with FBS (**a**, **b**) or without FBS (**c**, **d**). Image amplification of cells squared in **a** and **c** is shown in **b** and **d**, respectively. **e** RT-qPCR analysis of the *Ehtaf1* gene in trophozoites grown with (+FBS) or without FBS (−FBS). The *Eh40Ss2* gene was used as a control in the analysis. Scale bar 10 μm

Ehtaf1 gene was determined through the *E. histolytica* growth curve. Since there were no differences in the expression level of *Ehtaf1* at 12, 18, and 36 (data not shown), we selected 36 h as the start point in the analysis. The *Ehtaf1* mRNA level did not change through the exponential growth phase (18 to 72 h); however, an increase of 38.9% was observed at 96 h and 20.7% at 120 h, which correspond to the decline or death phase of the growth curve. At this phase, we observed a reduction of the cell viability in 13.1% and 31.5% at 96 and 120 h, respectively (Fig. 5a, b). As a control, the mRNA levels of *Ehtbp* and *Ehtrf1* genes were measured at the selected time points of the growth curve. *Ehtbp* mRNA level increased 74% at 120 h only, while *Ehtrf1* mRNA level showed no change through the growth curve (Fig. 5c, d). This behavior was similar to that shown for these genes in a previous report (Narayanasamy et al. 2018). Therefore, *Ehtaf1* could play a role under different stress conditions.

The mRNA level of *Ehtaf1* does not change in *E. histolytica* trophozoites cultured in serum depletion

Under this stress condition, cells did not divide, but were viable (97%) and their morphology changed from amoebic to rounded cells (Fig. 6a–d) as shown by others (Narayanasamy et al. 2018). However, the mRNA level of *Ehtaf1* did not change (Fig. 6f). This result was similar to that found for *Ehtrf1*, which showed no change in its mRNA level,

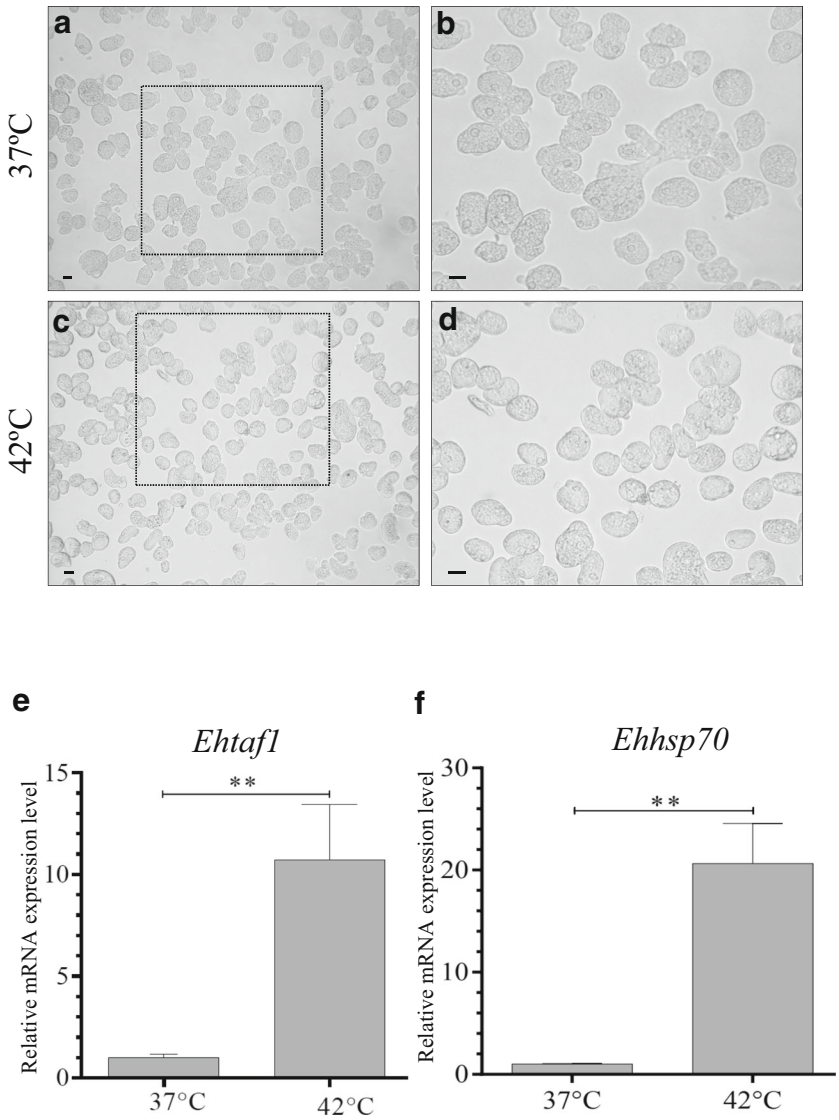


Fig. 7 The *Ehtaf1* mRNA level increases in trophozoites under heat shock stress. Microphotographs of trophozoites incubated at 37 °C (**a**, **b**) or 42 °C (**c**, **d**) for 4 h in TYI-S-33 medium at × 200 magnification through the light microscope. Image amplification of cells squared in (**a**) and (**c**) is shown in (**b**) and (**d**), respectively. Relative mRNA expression levels of *Ehtaf1* (**e**) and *Ehhsp70* (**f**) genes determined by RT-qPCR. The *Eh40Ss2* gene was used as a control in the analysis. Statistical significance (Student *t* test) defined as $^{**}P < 0.01$. Scale bar 10 μm

but contrary to the overexpression of 231% found for *Ehtbp* (Narayanasamy et al. 2018).

The mRNA expression level of the *Ehtaf1* gene is upregulated under heat shock stress

Another stress condition that was evaluated on trophozoites was the heat shock stress at 42 °C. Under this condition, cells were viable (97%) and their morphology changed from amoebic (Fig. 7a, b) to round shape (Fig. 7c, d), and the *Ehtaf1* mRNA expression level increased 970% in relation to the normal growth condition at 37 °C (Fig. 7e). As a control experiment, we measured the mRNA level of *Ehhsp70* gene and found an augment of 1960% (Fig. 7f), which was similar to that previously reported (Narayanasamy et al. 2018).

The *Ehtaf1* mRNA expression level remains unchanged during erythrophagocytosis and UV irradiation

E. histolytica trophozoites are specialized cells in phagocytosis and have great capacity of ingestion of RBCs. During this process, trophozoites ingested an increasing number of erythrocytes at 2, 5, 10, 15, and 30 min, reaching an average of 12 ingested erythrocytes per trophozoite at 30 min (Fig. 8a, b). This average number was obtained by the random selection of 100 trophozoites at each time point. Under this condition, the mRNA level of *Ehtaf1* did not change through the experiment (Fig. 8c).

The last stress condition tested was the irradiation with UV light, where no changes were observed up to 30 min of culture after the UV light doses applied to trophozoites as described in

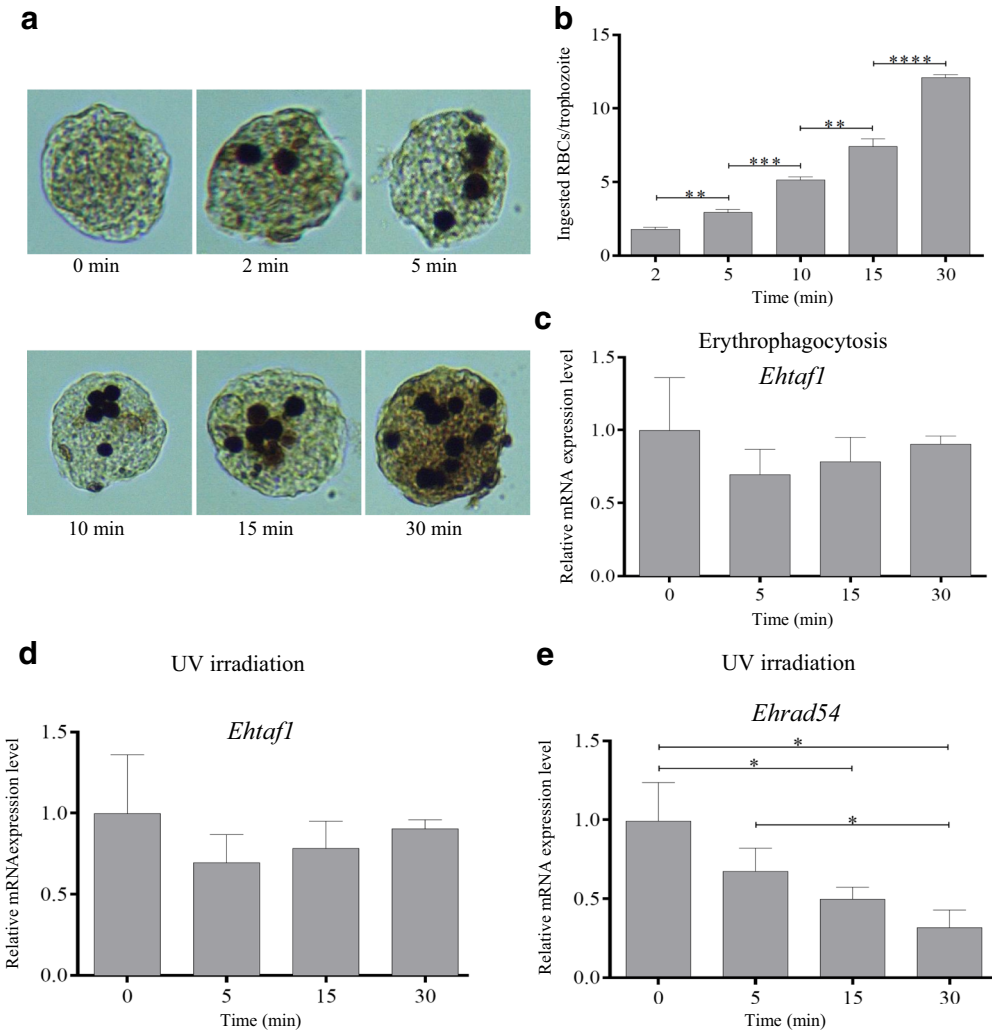


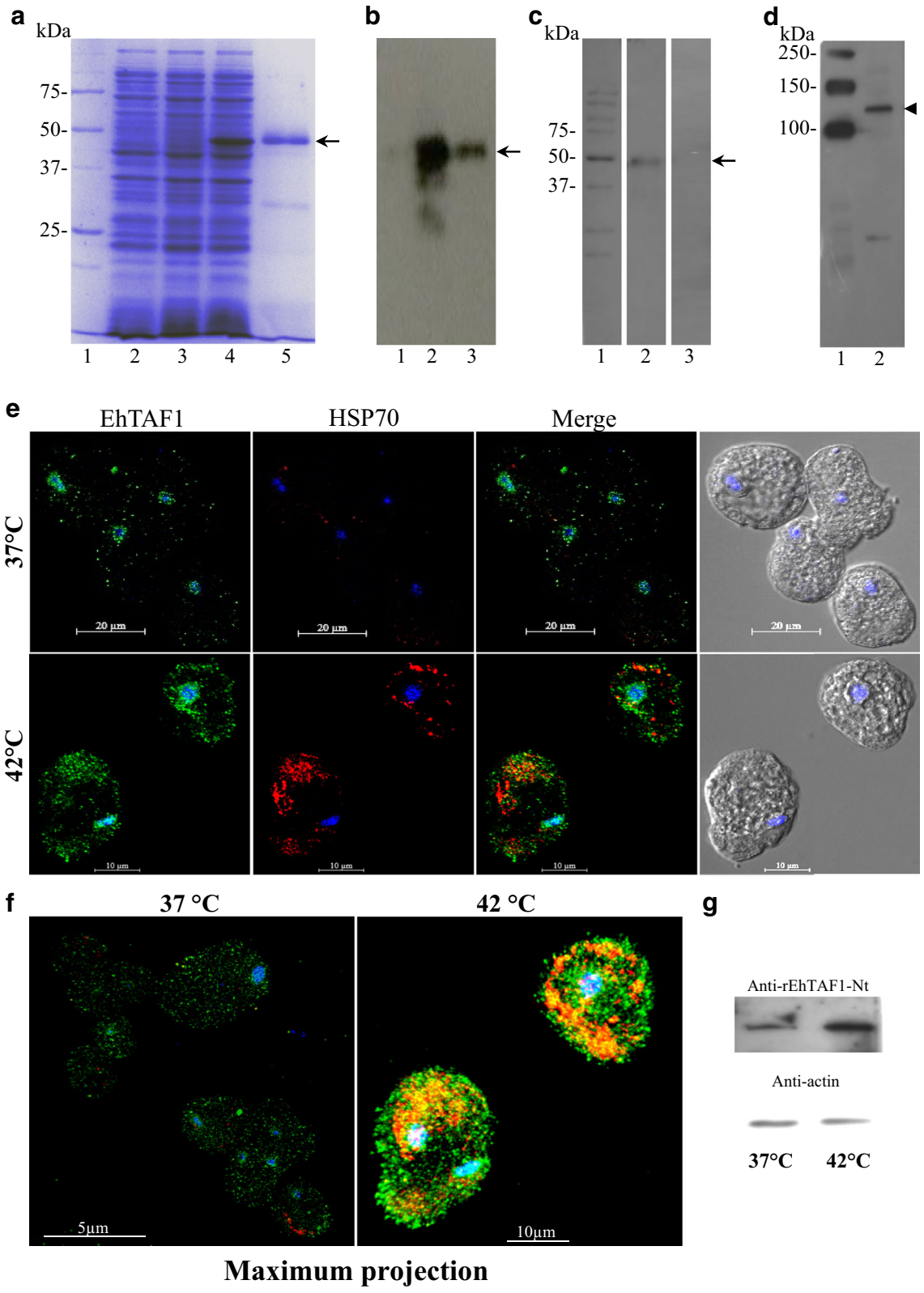
Fig. 8 The relative mRNA expression level of *Ehtaf1* does not vary during erythrophagocytosis and UV irradiation. **a** Light field microscope images showing erythrocytes ingested by trophozoites at different time points. Photomicrographs were obtained at × 200 magnification. **b** Graph of the average number of erythrocytes ingested by trophozoites at 2, 5, 10, 15, and 30 min of incubation of trophozoites with RBCs. **c** Relative *Ehtaf1* mRNA level during erythrophagocytosis. **d**, **e** Relative *Ehtaf1* and *Ehrad54* mRNA levels in trophozoites during UV irradiation. The *Eh40Ss2* gene was used as a control in the analyses. Statistical analysis (Student *t* test): $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$

Materials and Methods (Fig. 8d). As a control, we measured the *Ehrad54* mRNA level (Fig. 8e), a gene which is involved in DNA repair (López-Casamichana et al. 2008), and observed a significant reduction of 49% and 69% at 15 and 30 min post-irradiation, respectively, as previously reported (Narayanasamy et al. 2018).

The EhTAF1 is overexpressed in the cytosol and nuclei of trophozoites under heat shock stress

We cloned the *EhTaf1_Δ901-3492* DNA fragment into pRSET A plasmid, and induced the production of the recombinant EhTAF1 N-terminus (rEhTAF1-Nt, amino acids 1–300) polypeptide with 1 mM IPTG in *E. coli* BL21(DE3)pLysS (Fig. 9a). The rEhTAF1-Nt showed a molecular mass of 45 kDa and was purified by IMAC through a Ni^{2+}-NTA-agarose column under denaturing conditions (Fig. 9a). The identity of the rEhTAF1-Nt polypeptide was verified by Western blot, using the rabbit anti-6His tag polyclonal antibody (Fig. 9b) and by tandem mass spectrometry (data not shown). The purified recombinant protein was then used to produce rabbit polyclonal antibodies, which recognized the rEhTAF1-Nt polypeptide (Fig. 9c), and a 130-kDa band in whole protein extracts of trophozoites that corresponds to the full-length EhTAF1, and a band of 25 kDa that might be produced by digestion of the full-length EhTAF1 with proteinases (Fig. 9d).

Since the only stress condition that induced the overexpression of *Ehtaf1* gene was heat shock, we immunolocalized EhTAF1 in trophozoites grown at 37 °C and 42 °C. EhTAF1 was mainly localized in the nuclei at 37 °C (Fig. 9e). However, when cells were under the heat shock stress, an overexpression of EhTAF1 was seen mainly in cytoplasm (Fig. 9e). The maximum intensity projection of all collected optical planes along Z-axis allowed to see the overexpression of EhTAF1 under this stress condition (Fig. 9f). This was also confirmed by Western blot (Fig. 9g).


a
kDa
75-
50-
37-
25-
1 2 3 4 5
b
1 2 3
c
kDa
75-
50-
37-
1 2 3
d
kDa
250-
150-
100-
1 2
e
EhTAF1
HSP70
Merge
37°C
42°C
20 µm
10 µm
f
37 °C
42 °C
5µm
10µm
Maximum projection
g
Anti-rEhTAF1-Nt
Anti-actin
37°C
42°C

◀ **Fig. 9** Intracellular localization of EhTAF1 protein by immunofluorescence and confocal microscopy. **a** Induction and purification of the recombinant EhTAF1-Nt polypeptide. 12% SDS-PAGE gel stained with Coomassie Brilliant blue: lane 1, molecular weight markers; lane 2, extracts of bacteria transformed with pRSET A plasmid; lanes 3 and 4, whole protein extracts of transformed bacteria with the pRSET A/*EhTaf1_Δ901-3492* plasmid grown with no IPTG (lane 3) or with 1 mM IPTG (lane 4); lane 5, purified rEhTAF1-Nt polypeptide by IMAC through a NTA-Ni^{2+} column. **b** Western blot using the anti-6His tag monoclonal antibody; lane 1, uninduced bacterial extract; lane 2, induced bacterial extract; lane 3, purified rEhTAF1-Nt polypeptide. **c** Western blot using the anti-rEhTAF1-Nt polyclonal antibodies; lane 1, molecular weight marker; lane 2, purified rEhTAF1-Nt polypeptide; lane 3, Western blot of purified rEhTAF1-Nt polypeptide using rabbit preimmune serum as a control. **d** Western blot of whole protein extract of *E. histolytica* trophozoites using the rabbit anti-rEhTAF1-Nt polyclonal antibody; lane 1, molecular weight markers; lane 2, whole protein extracts. Arrow indicates the rEhTAF1-Nt polypeptide; arrowhead indicates the full-length EhTAF1 endogenous protein. **e** Trophozoites cultured at 37 °C or under heat shock stress at 42 °C for 4 h were permeabilized and incubated with rabbit anti-rEhTAF1 polyclonal antibodies labeled with Alexa 488 (Green Channel) and anti-mouse Hsp70 monoclonal antibody labeled with Alexa 594 (red channel). Nuclei were stained with DAPI (blue channel) and the samples were observed through a confocal microscope. Light field images are shown on the right. **f** Maximum projection images of trophozoites grown at 37 °C and 42 °C. **g** Western blot using the anti-rEhTAF1-Nt polyclonal and anti-actin monoclonal antibodies and whole protein extracts of cells grown at 37 °C and 42 °C

Subcellular localization of the *Entamoeba histolytica* TAF1 in trophozoites grown at 37 °C and under heat shock stress by immunoelectron microscopy (IEM)

We have determined that EhTAF1 is localized in both the nucleus and cytoplasm in *E. histolytica* trophozoites by immunofluorescence and confocal microscopy (Fig. 9). To get a more detailed intracellular localization, we carried out immunoelectron microscopy using the rabbit polyclonal antibodies against the rEhTAF1-Nt polypeptide (Fig. 10). EhTAF1 was localized in the nucleus, in regions of euchromatin and heterochromatin (Bowers and Korn 1968). In cytosol, it was found associated with small cytoplasmic vesicles and vacuoles (Fig. 10c, e). At 42 °C, the EhTAF1 protein amount increased in these compartments, mainly in cytoplasm (Fig. 11). There was no EhTAF1 protein recognition when cells were incubated with preimmune serum (Figs. 10 and 11a, b). Furthermore, some structures known as spherules or corpuscles (Chévez et al. 1972) were observed within the nucleus of trophozoites grown either at 37 °C or 42 °C. In addition, a highly ordered structure and crystalloid-like appearance was also observed (Fig. 11a, b). The nature and function of these structures remain unknown.

Discussion

In this work, the *Ehtaf1* gene was identified in the *E. histolytica* genome database, which codes a polypeptide with a theoretical molecular mass of 135 kDa. EhTAF1 showed an identity of 74 to 99% to TAF1 proteins from *E. nuttalli*, *E. dispar*, and *E. invadens*, revealing a high conservation between these species. *E. nuttalli* is the causative agent of invasive amoebiasis in primates, which can cause hemorrhagic dysentery, hepatic abscess, or extraintestinal amoebiasis. *E. dispar* is found in humans and is considered an avirulent commensal, while *E. invadens* is a commensal protozoan living in the gastrointestinal tract of reptiles (Loomis et al. 1983; Levecke et al. 2010; Weedall and Hall 2011).

The human *CCG1* was cloned by the complementation of two temperature sensitive mutants of the BHK cell line (tsBN462 and ts13) that arrest the cell cycle in late G1 phase at the restrictive temperature (Sekiguchi et al. 1988, 1991). Later, the cloning of the gene coding the human TAF250 subunit (renamed TAF1) showed that this gene is identical to the *CCG1* gene (Ruppert et al. 1993). *TAF1* gene performs essential functions in mammals, *S. cerevisiae* and *D. melanogaster* (Ruppert and Tjian 1995; Dikstein et al. 1996; Lee et al. 2000; Matangkasombut et al. 2000). A temperature conditional TAF1 yeast mutant showed that the expression of approximately 80% of protein-coding genes is deeply affected, revealing a central role of TAF1 in the control of gene expression (Lee et al. 2000). TAF1 binds to TBP (Dikstein et al. 1996), RAP74 (a subunit of TFIIF), and TAF7 (Gegonne et al. 2001; Wang et al. 2014), another subunit of TFIID. In addition, there has been reported that TAF1 possesses phosphorylation and acetyltransferase enzymatic activities (Dikstein et al. 1996; Mizzen et al. 1996). In relation to the latter activity, the human TAF1 acetylates histone H4 (Jacobson et al. 2000). Moreover, it also displays the binding capacity to acetylated histones through its tandem Bromo domains localized at its carboxy-terminal domain (Matangkasombut et al. 2000), and the DNA-binding capacity to promoters through its zinc knuckle domain, which has a critical role for the TAF1 DNA-binding capacity (Curran et al. 2018).

The EhTAF1 protein sequence contains the DUF3591 domain localized at the center of the protein, which has been found conserved in all TAF1 analyzed sequences (Fig. 1). The tridimensional structure of the human TAF1 DUF domain co-crystalized with the human TAF7 confirmed previous results for this molecular interaction (Wang et al. 2014), which affects the acetyltransferase activity of TAF1 (Gegonne et al. 2001). EhTAF1 has a conserved DUF domain structure, revealing a conserved function through evolution (Fig. 1 and Fig. 3). Interestingly, we also found a sequence related to the human TAF7 in the *E. histolytica* genome (our unpublished results), which might reveal a conserved TAF1-TAF7 dimer structure in this parasite. However, more studies are needed to address this issue.

Another domain that has been found in some TAF1 proteins from other species is the TBP-binding domain, which

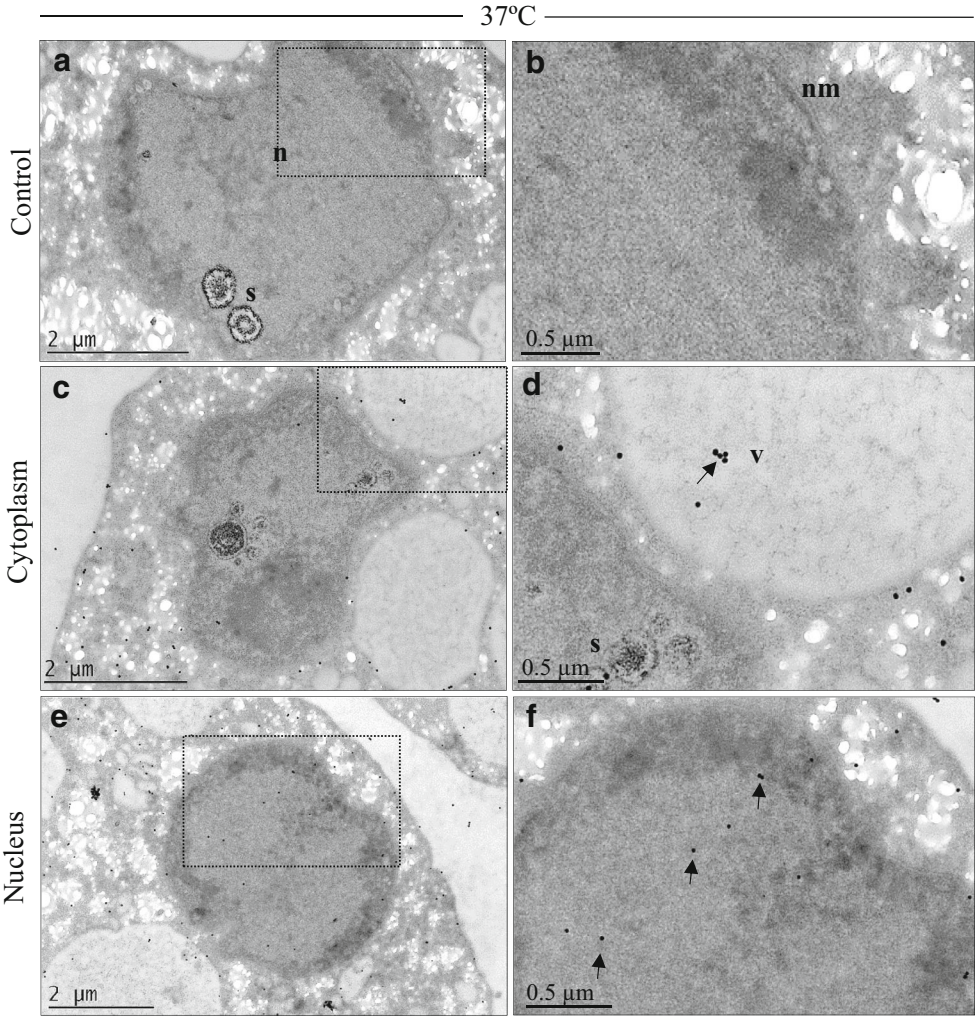


Fig. 10 Intracellular localization of EhTAF1 in *Entamoeba histolytica* trophozoites cultured at 37 °C by immunoelectron microscopy (IEM). Thin sections of trophozoites were incubated with the anti-rEhTAF1-Nt antibodies and then with IgG anti-rabbit polyclonal antibodies conjugated to 30 nm gold particles. **a**, **b** Preimmune serum used as a control. **c**–**f** Immunolocalization of EhTAF1 in vacuoles (**d**) and nucleus (**e**, **f**) is indicated with arrows. Nucleus (n), nuclear membrane (nm), spherules or corpuscles (s), vacuole (v). Figures (**b**, **d**, and **f**) are the digital zoom of figures

was not found in the EhTAF1 protein, which might indicate that this region is not conserved through evolution, in spite of the great conservation observed in TBP proteins. However, we do not discard the possible interaction between the N-terminal domain of EhTAF1 and EhTBP. The zinc knuckle domain has also been found in some TAF1 proteins of other species (Fig. 1), which is located between the DUF and Bromo domains. The zinc knuckle domain was not found in any of the TAF1 proteins from amoebic organisms (Fig. 1).

In mammals, the carboxy-terminal domain of TAF1 binds to the tail of acetylated histone H4 in vitro (Jacobson et al. 2000). EhTAF1 has a single Bromo domain at its carboxy-terminal end, unlike some TAF1 molecules of other species, which showed two Bromo domains arranged in tandem (Figs. 1 and 2). Interestingly, the *C. albicans* TAF1 sequence has no Bromo domains, but it is presumed that another protein, known as BDF1, which contains a Bromo domain, could be replacing the lack of a Bromo domain in CaTAF1, allowing the binding to acetylated nucleosomes by TFIID (Matangkasombut et al. 2000; Ladurner et al. 2003; Matangkasombut and Buratowski 2003; Lawit et al. 2007).

Furthermore, we did not identify the protein kinase domains in all the sequences shown in Fig. 1, which are located in the amino-terminal and carboxy-terminal ends in the human TAF1 isoform 1 (Dikstein et al. 1996). An additional domain was only identified in TAF1 of *A. thaliana*, which corresponds to the ubiquitin domain (Fig. 1). With all these findings, we provide evidences that *E. histolytica* contains the *Ehtaf1* gene in its genome.

This parasite has sets of genes that respond to different stimuli to display specialized functions, including adhesion, lysis of host cells, phagocytosis, trogocytosis, encystment, and stress responses (Ravdin and Guerrant 1981; Orozco et al. 1983; Ralston et al. 2014). These mechanisms are regulated by genes, including transcription factors that are activated in response to the stress stimuli (Kathryn et al. 2004). However, the participation of basal factors in these processes is unclear. We explored the expression levels of *Ehtaf1* mRNA under several stress conditions. In the death phase of the growth curve, cells changed their morphology to round shape and

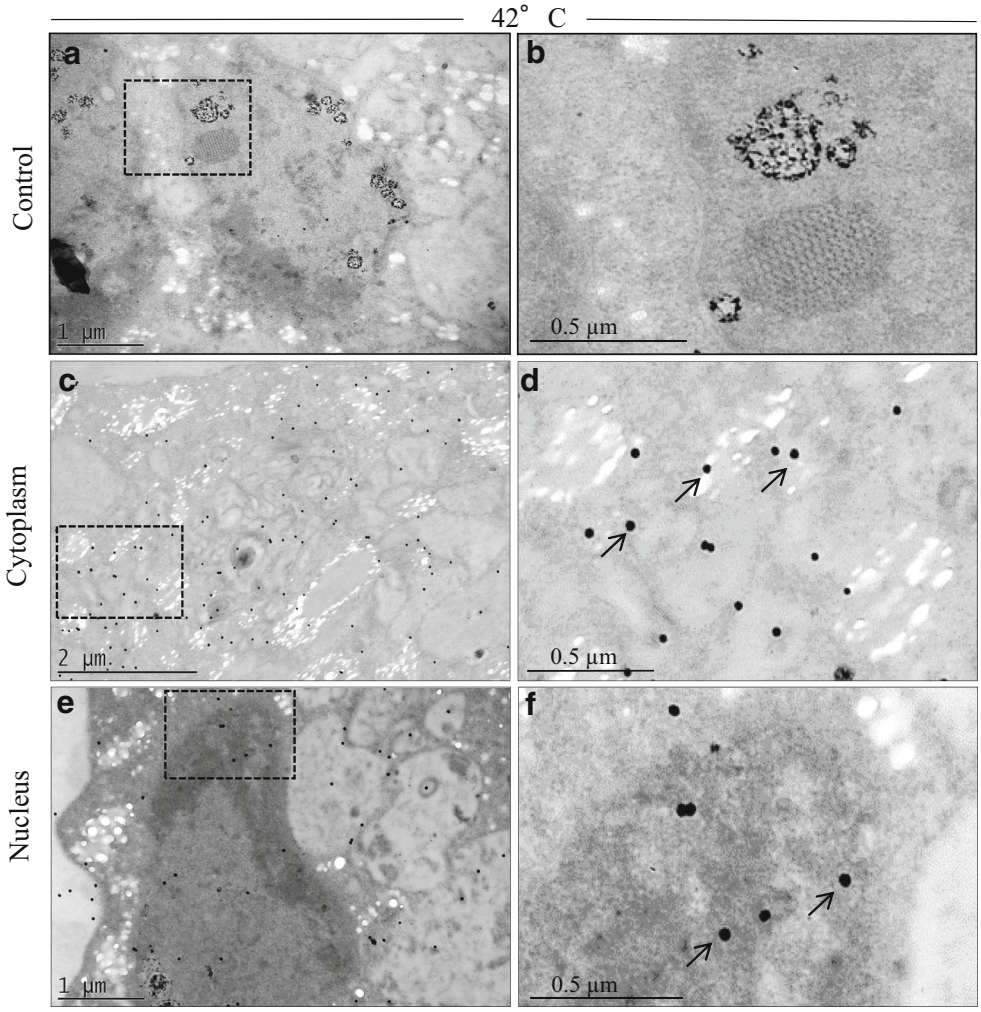


Fig. 11 Intracellular localization of EhTAF1 in *Entamoeba histolytica* trophozoites cultured at 42 °C by immunoelectron microscopy (IEM). Thin sections of trophozoites were incubated with the anti-rEhTAF1-Nt antibodies and then with IgG anti-rabbit polyclonal antibodies conjugated to 30 nm gold particles. **a**, **b** Preimmune serum used as a control. **c**, **d** Immunolocalization of EhTAF1 significantly increased in cytoplasm. Arrows show EhTAF1 located in small vesicles. **e**, **f** We also observe staining of EhTAF1 in the nucleus of this parasite. Figures (**b**, **d**, and **f**) are the digital zoom of figures

the *Ehtaf1* mRNA levels increased slightly (38.9%) at 96 h (Fig. 5a, b). Under this condition, the *Ehtbp* increased 74% at 120 h of growth culture with no change in the mRNA level of *Ehtrf1* gene. These results were similar to those described previously (Narayanasamy et al. 2018). This result suggests that *Ehtaf1* could have a role in stress responses. However, the mRNA level of *Ehtaf1* did not change during serum depletion, UV irradiation, and erythrophagocytosis, but only under heat shock stress at 42 °C. This behavior was different to that previously reported for the basal transcription factors *Ehtbp* and *Ehtrf1* in similar stress conditions (Narayanasamy et al. 2018). To our knowledge, there are no reports on the participation of TAF1 in the heat shock stress response.

We localized EhTAF1 protein in the nucleus and cytoplasm in trophozoites grown in normal conditions. Interestingly, under the heat shock stress at 42 °C, the amount of TAF1 protein increased both in the nucleus and cytoplasm (Fig. 9). Moreover, TAF1 partially colocalized with EhHSP70 and

Fig. 12 Identification of EhTBP and EhTRF1 proteins by immunoprecipitation with anti-rEhTAF1-Nt polyclonal antibodies. In bold letters are de amino acids identified in peptides (underlined) by tandem mass spectrometry. IDs in parenthesis correspond to the UniProtKB Accession Numbers for the proteins

>EhTRF1(A7UFC2)
MSSQSSPISLTNNDLEISSNNNKPIIKQLPVPTNEIEKPVIQNIVATVELDCTINLQ
DVVRRVRNAEYNPKRFGALIIRITNPK**TTALVFHSGK**LVVTGGKTVDDSRLAGRKYA
RILQRLGYNVKFNHFK**IQNVVASCDMK**FAISLKELIQLAPKITK**YEPEIFPGVVYR**L
ADPKMVLLIFASGKIVFTGGKEIEQINKAFSEIYK**ILLQVANNDN**

>EhTBP(P52653)
MSTPGDFSLSPFILGGAVDPRSMSQLGNICHADYMSTSTESQERSLNNPNDTHPEIV
NVVSTFQLGVKLELRKIVQKARNAEYNPKRFAGAIMRISSPK**STALIFQTGK**IVCTG
TRSIEESKIASKKYAKIIKK**IGYPIHYSNFNVQNIVGSCDVK**FQIALRTLVDSYLAF
CQYEPEVFPGLVYRMASPK**VTLLVFSTGKVVLTGAKDEESLNLAYK**NIYPILLANRK
EDISNQ

accumulated in structures known as stress granules in the cytoplasm (Fig. 9e, f) (Katz et al. 2014). This interaction could protect EhTAF1 from heat and help in its refolding by EhHSP70 or could be used as a scaffold for interacting with other proteins involved in the heat shock stress response. There has been shown that mRNAs, excluding those mRNAs encoding stress-induced heat shock proteins, and RNA-binding proteins accumulate in these granules (Anderson and Kedersha 2006). One interesting finding obtained by transmission electron microscopy was the observation of spherules and highly ordered structures in the nucleus of trophozoites (Fig. 11b) that have also been reported (Bowers and Korn 1968; Chévez et al. 1972), but their functions remain unveiled.

Recent immunoprecipitation assays using the rabbit anti-rEhTAF1-amino terminus polyclonal antibodies and tandem mass spectrometry analysis conducted by our group revealed that EhTAF1 interacts with EhTBP and EhTRF1, a finding that suggests the possibility of the existence of two TFIID complexes in *E. histolytica* trophozoites (Fig. 12; electronic Supplementary Material Table 2).

Acknowledgments The authors are grateful to Dr. Javier Cázares-Apátiga for his technical support in confocal microscopy experiments, Emmanuel Ríos-Castro MSc. for his technical support in MS analysis at the Unidad de Genómica, Proteómica y Metabolómica (UGPM), LaNSE, Cinvestav-IPN, and Cesar Isaac Bazán-Méndez for spelling and grammar review of manuscript.

Author contributions BA-B, performed, designed the experiments, discussed the results, and wrote the manuscript; RKN, designed the experiments; GG-R, helped in some experiments with *E. histolytica* trophozoites, MLL-B, participated in the generation of antibodies and Western blots; AEL-G, performed the electron microscopy experiments; BCHM, designed and performed the electron microscopy experiments; CAC-S, discussed experiments and results; EO, supplied the *E. histolytica* strain and laboratory facilities for the development of some experiments; JPL-A, conducted experiments, discussed experiments and results, wrote and revised the manuscript.

Funding We are grateful to the National Council of Science and Technology of Mexico (CONACyT) for the support to this work with the Grant number CB-2009-132020-B to Dr. Juan Pedro Luna-Arias, and for the PhD fellowship to Bartolo Avendaño Borromeo (CVU 229747).

Compliance with ethical standards

Ethics statement Cinvestav fulfills the standard of the Mexican Official Norm (NOM-062-ZOO-1999) "Technical Specifications for the Care and Use of Laboratory Animals" based on the Guide for the Care and Use of Laboratory Animals "The Guide," 2011, NRC, USA with the Federal Register Number BOO.02.03.02.01.908. The Institutional Animal Care and Use Committee (IACUC/ethics committee) from Cinvestav, as the regulatory office for the approval of research protocols involving the use of laboratory animals and in fulfillment of the Mexican Official Norm, has reviewed and approved all animal experiments.

Conflict of interest The authors declare that they have no conflict of interest.

References

Ali IKM, Ehrenkaufer GM, Hackney JA, Singh U (2007) Growth of the protozoan parasite *Entamoeba histolytica* in 5-azacytidine has limited effects on parasite gene expression. BMC Genomics 8(1):7. https://doi.org/10.1186/1471-2164-8-7

Ali IK, Clark CG, Petri WA (2008) Molecular epidemiology of amebiasis. Infect Genet Evol 8:698–707. https://doi.org/10.1016/j.meegid.2008.05.004

Anderson P, Kedersha N (2006) RNA granules. J Cell Biol 172(6):803–808. https://doi.org/10.1083/jcb.200512082

Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, Papadopoulos P, Tora L, Schultz P, Berger I (2013) The architecture of human general transcription factor TFIID core complex. Nature 493(7434):699–702. https://doi.org/10.1038/nature11791

Bowers B, Korn ED (1968) The fine structure of *Acanthamoeba castellanii*. J Cell Biol 39(1):95–111. https://doi.org/10.1083/jcb.39.1.95

Brou C, Chaudhary S, Davidson I, Lutz Y, Wu J, Egly JM, Tora L, Chambon P (1993) Distinct TFIID complexes mediate the effect of different transcriptional activators. EMBO J 12(2):489–499

Bruchhaus I, Leippe M, Lioutas C, Tannich E (1993) Unusual gene organization in the protozoan parasite Entamoeba histolytica. DNA Cell Biol 12:925–933. https://doi.org/10.1089/dna.1993.12.925

Burley SK, Roeder RG (1996) Biochemistry and structural biology of transcription factor IID (TFIID). Annu Rev Biochem 65:769–799. https://doi.org/10.1146/annurev.bi.65.070196.004005

Castañón-Sánchez CA, Luna-Arias JP, de Dios-Bravo MG, Herrera-Aguirre ME, Olivares-Trejo JJ, Orozco E et al (2010) *Entamoeba histolytica*: a unicellular organism containing two active genes encoding for members of the TBP family. Protein Expr Purif 70(1):48–59. https://doi.org/10.1016/j.pep.2009.12.007

Chévez A, Segura M, Corona D, Iturbe-Alessio I (1972) Juxtanuclear spherules of *Entamoeba histolytica*. Morphological and dynamic study. Arch Invest Med (Mex) 2(Suppl 2):257–264

Cler E, Papai G, Schultz P, Davidson I (2009) Recent advances in understanding the structure and function of general transcription factor TFIID. Cell Mol Life Sci 66(13):2123–2134. https://doi.org/10.1007/s00018-009-0009-3

Curran EC, Wang H, Hinds TR, Zheng N, Wang EH (2018) Zinc knuckle of TAF1 is a DNA binding module critical for TFIID promoter occupancy. Sci Rep 8:4630. https://doi.org/10.1038/s41598-018-22879-5

Das S, Lohia A (2002) Delinking of S phase and cytokinesis in the protozoan parasite *Entamoeba histolytica*. Cell Microbiol 4:55–60. https://doi.org/10.1046/j.1462-5822.2002.00165.x

de Dios-Bravo G, Luna-Arias JP, Riverón AM, Olivares Trejo JJ, López Camarillo C, Orozco E (2005) *Entamoeba histolytica* TATA box binding protein binds to different TATA variants *in vitro*. FEBS J 272(6):1354–1366. https://doi.org/10.1111/j.1742-4658.2005.04566.x

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. https://doi.org/10.1016/0035-9203(78)90144-X

Dikstein R, Ruppert S, Tjian R (1996) TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. Cell 84(5):781–790. https://doi.org/10.1016/S0092-8674(00)81055-7

Dynlacht BD, Hoey T, Tjian R (1991) Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional

activation. Cell 66(3):563–576. https://doi.org/10.1016/0092-8674(81)90019-2

Gangopadhyay SS, Ray SS, Kennady K, Pande G, Lohia A (1997) Heterogeneity of DNA content and expression of cell cycle genes in axenically growing *Entamoeba histolytica* HM1: IMSS clone A. Mol Biochem Parasitol 90:9–20. https://doi.org/10.1016/S0166-6851(97)00156-4

Gegonne A, Weissman JD, Singer DS (2001) TAFII55 binding to TAFII250 inhibits its acetyltransferase activity. Proc Natl Acad Sci U S A 98(22):12432–12437. https://doi.org/10.1073/pnas.211444798

Hantsche M, Cramer P (2017) Conserved RNA polymerase II initiation complex structure. Curr Opin Struct Biol 47:17–22. https://doi.org/10.1016/j.sbi.2017.03.013

Jacobson RH, Ladurner AG, King DS, Tjian R (2000) Structure and function of a human TAFII250 double bromodomain module. Science 288:1422–1425. https://doi.org/10.1126/science.288.5470.1422

Juven-Gershon T, Kadonaga JT (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol 339(2):225–229. https://doi.org/10.1016/j.ydbio.2009.08.009

Kathryn L, Huisinga B, Pugh F (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. Mol Cell 13(4):573–585. https://doi.org/10.1016/S1097-2765(04)00087-5

Katz S, Trebicz-Geffen M, Ankri S (2014) Stress granule formation in *Entamoeba histolytica*: cross-talk between EhMLBP, EhRLE3 reverse transcriptase and polyubiquitinated proteins. Cell Microbiol 16(8):1211–1223. https://doi.org/10.1111/cmi.12273

Kokubo T, Gong DW, Wootton JC, Horikoshi M, Roeder RG, Nakatani Y (1994) Molecular cloning of Drosophila TFIID subunits. Nature 367(6462):484–487. https://doi.org/10.1038/367484a0

Ladurner AG1, Inouye C, Jain R, Tjian R (2003) Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. Mol Cell 11:365–376. https://doi.org/10.1016/S1097-2765(03)00035-2

Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH (1998) New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev 12:34–44. https://doi.org/10.1101/gad.12.1.34

Lawit S, O'Grady K, Gurley W, Czarnecka Verner E (2007) Yeast two hybrid map of Arabidopsis TFIID. Plant Mol Biol 64:73–87. https://doi.org/10.1007/s11103-007-9135-1

Lee TI, Causton HC, Holstege FC, Shen WC, Hannett N, Jennings EG, Et a (2000) Redundant roles for the TFIID and SAGA complexes in global transcription. Nature 405(6787):701–704. https://doi.org/10.1038/35015104

Levecke B, Dreesen L, Dorny P, Verweij JJ, Vercammen F, Casaert S, Vercruysse J, Geldhof P (2010) Molecular identification of *Entamoeba spp*. in captive nonhuman primates. J Clin Microbiol 48(8):2988–2990. https://doi.org/10.1128/JCM.00013-10

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C (T)). Methods 25(4):402–408. https://doi.org/10.1006/meth.2001.1262

Loftus B, Anderson I, Davies R, Alsmark UCM, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, Nozaki T, Suh B, Pop M, Duchene M, Ackers J, Tannich E, Leippe M, Hofer M, Bruchhaus I, Willhoeft U, Bhattacharya A, Chillingworth T, Churcher C, Hance Z, Harris B, Harris D, Jagels K, Moule S, Mungall K, Ormond D, Squares R, Whitehead S, Quail MA, Rabbinowitsch E, Norbertczak H, Price C, Wang Z, Guillén N, Gilchrist C, Stroup SE, Bhattacharya S, Lohia A, Foster PG, Sicheritz-Ponten T, Weber C, Singh U, Mukherjee C, el-Sayed NM, Petri WA, Clark CG, Embley TM, Barrell B, Fraser CM, Hall N (2005) The genome of the protist parasite *Entamoeba histolytica*. Nature 433:865–868. https://doi.org/10.1038/nature03291

Loomis MR, Britt JO Jr, Gendron AP, Holshuh HJ, Howard EB (1983) Hepatic and gastric amebiasis in black and white colobus monkeys. J Am Vet Med Assoc 183(11):1188–1191

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. https://doi.org/10.1186/1471-2199-9-35

Lorenzi HA, Puiu D, Miller JR, Brinkac LM, Amedeo P, Hall N, Caler EV (2010) New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. PLoS Negl Trop Dis 4:e716. https://doi.org/10.1371/journal.pntd.0000716

Luna-Arias JP, Hernandez-Rivas R, de Dios-Bravo G, Garcia J, Mendoza L, Orozco E (1999) The TATA-box binding protein of *Entamoeba histolytica*: cloning of the gene and location of the protein by immunofluorescence and confocal microscopy. Microbiol 145(1):33–40. https://doi.org/10.1099/13500872-145-1-33

Manna D, Ehrenkaufer GM, Singh U (2014) Regulation of gene expression in the protozoan parasite *Entamoeba invadens*: identification of core promoter elements and promoters with stage-specific expression patterns. Int J Parasitol 44(11):837–845. https://doi.org/10.1016/j.ijpara.2014.06.008

Martinez E, Kundu TK, Fu J, Roeder RG (1998) A human SPT3-TAFII31-GCN5-L acetylase complex distinct from transcription factor IID. J Biol Chem 273:23781–23785. https://doi.org/10.1074/jbc.273.37.23781

Matangkasombut O, Buratowski S (2003) Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. Mol Cell 11(2):353–363. https://doi.org/10.1016/S1097-2765(03)00033-9

Matangkasombut O, Buratowski RM, Swilling NW, Buratowski S (2000) Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. Genes Dev 14(8):951–962. https://doi.org/10.1101/gad.14.8.951

Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T et al (1996) The TAF (II) 250 subunit of TFIID has histone acetyltransferase activity. Cell 87(7):1261–1270. https://doi.org/10.1016/S0092-8674(00)81821-8

Moqtaderi Z, Yale JD, Struhl K, Buratowski S (1996) Yeast homologues of higher eukaryotic TFIID subunits. Proc Natl Acad Sci U S A 93(25):14654–14658. https://doi.org/10.1073/pnas.93.25.14654

Nagy Z, Tora L (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. Oncogene 26:5341–5357. https://doi.org/10.1038/sj.onc.1210604

Narayanasamy RK, Castañón-Sanchez CA, Luna-Arias JP, García-Rivera G, Avendaño-Borromeo B, Labra-Barrios ML, Valdés J, Herrera-Aguirre ME, Orozco E (2018) The *Entamoeba histolytica* TBP and TRF1 transcription factors are GAAC-box binding proteins, which display differential gene expression under different stress stimuli and during the interaction with mammalian cells. Parasit Vectors 11(1):153. https://doi.org/10.1186/s13071-018-2698-7

Ogryzko VV, Kotani T, Zhang X, Schiltz RL, Howard T, Yang XJ, Howard BH, Qin J, Nakatani Y (1998) Histone-like TAFs within the PCAF histone acetylase complex. Cell 94:35–44. https://doi.org/10.1016/S0092-8674(00)81219-2

Orozco E, Guarneros G, Martinez-Palomo A, Sánchez T (1983) Entamoeba histolytica. Phagocytosis as a virulence factor. J Exp Med 158(5):1511–1521. https://doi.org/10.1084/jem.158.5.1511

Papai G, Tripathi MK, Ruhlmann C, Werten S, Crucifix C, Weil PA, Schultz P (2009) Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. Structure 17(3):363–373. https://doi.org/10.1016/j.str.2009.01.006

Poon D, Weil PA (1993) Immunopurification of yeast TATA-binding protein and associated factors. J Biol Chem 268(21):15325–15328

Purdy JE, Lana TP, Mann BJ, Petri WA (1996) Upstream regulatory elements controlling expression of the *Entamoeba histolytica* lectin. Mol Biochem Parasitol 78:91–103. https://doi.org/10.1016/S0166-6851(96)02614-X

Ralston KS, Solga MD, Mackey-Lawrence NM, Somlata BA, Petri WA (2014) Trogocytosis by *Entamoeba histolytica* contributes to cell killing and tissue invasion. Nature 508:526–530. https://doi.org/10.1038/nature13242

Ravdin JI, Guerrant RL (1981) Role of adherence in cytopathogenic mechanisms of *Entamoeba histolytica*. Study with mammalian tissue culture cells and human erythrocytes. J Clin Invest 68:1305–1313. https://doi.org/10.1172/JCI110377

Reese JC, Apone L, Walker SS, Griffin LA, Green MR (1994) Yeast TAFIIS in a multisubunit complex required for activated transcription. Nature 371(6497):523–527. https://doi.org/10.1038/371523a0

Ruppert S, Tjian R (1995) Human TAFII250 interacts with RAP74: implications for RNA polymerase II initiation. Genes Dev 9:2747–2755. https://doi.org/10.1101/gad.9.22.2747

Ruppert S, Wang Edith H, Tjian R (1993) Cloning and expression of human TAFII250: a TBP-associated factor implicated in cell-cycle regulation. Nature 362:175–179. https://doi.org/10.1038/362175a0

Sekiguchi T, Miyata T, Nishimoto T (1988) Molecular cloning of the cDNA of human X chromosomal gene (CCG1) which complements the temperature-sensitive G1 mutants, tsBN462 and ts13, of the BHK cell line. EMBO J 1988(6):1683–1687

Sekiguchi T, Nohiro Y, Nakamura Y, Hisamoto N, Nishimoto T (1991) The human CCG1 gene, essential for progression of the G1 phase, encodes a 210-kilodalton nuclear DNA-binding protein. Mol Cell Biol 11(6):3317–3325. https://doi.org/10.1128/MCB.11.6.3317

Shrimal S, Bhattacharya S, Bhattacharya A (2010) Serum-dependent selective expression of EhTMKB1-9, a member of *Entamoeba histolytica* B1 family of transmembrane kinases. PLoS Pathog 6(6):e1000929. https://doi.org/10.1371/journal.ppat.1000929

Singh U, Rogers JB, Mann BJ, Petri WA (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 94:8812–8817. https://doi.org/10.1073/pnas.94.16.8812

Singh U, Gilchrist CA, Schaenman JM, Rogers JB, Hockensmith JW, Mann BJ et al (2002) Context-dependent roles of the *Entamoeba histolytica* core promoter element GAAC in transcriptional activation and protein complex assembly. Mol Biochem Parasitol 120(1):107–116. https://doi.org/10.1016/S0166-6851(01)00441-8

Steinbuch M, Audran R (1969) The isolation of IgG from mammalian sera with the aid of caprylic acid. Arch Biochem Biophys 134(2):279–284. https://doi.org/10.1016/0003-9861(69)90285-9

Tachibana H, Yanagi T, Lama C, Pandey K, Feng M, Kobayashi S, Sherchand JB (2013) Prevalence of *Entamoeba nuttalli* infection in wild rhesus macaques in Nepal and characterization of the parasite isolates. Parasitol Int 62(2):230–235. https://doi.org/10.1016/j.parint.2013.01.004

Trowitzsch S, Viola C, Scheer E, Conic S, Chavant V, Fournier M et al (2015) Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. Nat Commun 14(6):6011. https://doi.org/10.1038/ncomms7011

Wang H, Curran EC, Hinds TR, Wang EH, Zheng N (2014) Crystal structure of a TAF1-TAF7 complex in human transcription factor IID reveals a promoter binding module. Cell Res 24(12):1433–1444. https://doi.org/10.1038/cr.2014.148

Weber C, Guigon G, Bouchier C, Frangeul L, Moreira S, Sismeiro O, Gouyette C, Mirelman D, Coppee JY, Guillén N (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. https://doi.org/10.1128/EC.5.5.871-875.2006

Weber C, Marchat LA, Guillén N, Lopez-Camarillo C (2009) Effects of DNA damage induced by UV irradiation on gene expression in the protozoan parasite *Entamoeba histolytica*. Mol Biochem Parasitol 164:165–169. https://doi.org/10.1016/j.molbiopara.2008.12.005

Weedall GD, Hall N (2011) Evolutionary genomics of *Entamoeba*. Res Microbiol 162(6):637–645. https://doi.org/10.1016/j.resmic.2011.01.007

Wieczorek E, Brand M, Jacq X, Tora L (1998) Function of TAF (II)-containing complex without TBP in transcription by RNA polymerase II. Nature 393:187–191. https://doi.org/10.1038/30283

Wright KJ, Marr MT 2nd, Tjian R (2006) TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. Proc Natl Acad Sci U S A 103(33):12347–12352. https://doi.org/10.1073/pnas.0605499103