Journal of Bacteriology and Parasitology
ISSN: 2155-9597



Research Article Open Access

# Transcriptional Activation by an URE4-like Sequence in the *EhPgp1* Gene Core Promoter

**ME Ramírez, DG Pérez, E Náder and C Gómez***

*Institutional Program of Molecular Biomedicine, School of Medicine and Homeopathy, National Polytechnic Institute (IPN), Guillermo Massieu Helguera No. 239, Fracc. La Escalera, Ticoman, D. F., Mexico*

## Abstract

*EhPgp1* is one of the multidrug resistance genes expressed in drug resistant trophozoites from *Entamoeba histolytica*. Previous studies in our laboratory have demonstrated that two C/EBP sites participate in the transcriptional activation of this gene. However there is other relevant region that also governs the regulation of *EhPgp1* expression in clone C2. In this report we provide evidence that transcription of the *EhPgp1* gene is at least partly regulated by the cis-acting R9 repeated sequences and EhEBP1 protein. Structural analysis of the region from -234 to -197 bp shows the presence of two repeated sequences of 9 bp [R9(1) and R9(2)] located at -226 to -203 bp. Deletions and mutations analysis of the R9 motifs significantly reduced promoter activity in trophozoites from clone C2. EMSA experiments revealed specific binding of nuclear proteins from *E. histolytica* to the R9 sequence. While competition assays showed that the presence of more than one R9 sequence is necessary for a strong DNA-protein interaction. Moreover, Western blot experiments with partially-purified proteins interacting with the R9 motif and antibodies against EhEBP1, recognized a 28 kDa protein. Interestingly, this antibody in supershift assays prevented the DNA-protein interactions formation, of the R9 sequences and nuclear proteins from amoeba, indicating that one of the proteins that interact with the R9 element is an EhEBP1-like one. In conclusion, we demonstrate that R9 motifs are recognized by an EhEBP1 protein and activate the *EhPgp1* gene expression.

**Keywords:** *EhPgp1* gene; MDR; *E. Histolytica; Cis*-acting element

## Introduction

*Entamoeba histolytica* is the protozoan responsible for human amoebiasis, it kills 70,000 humans each year around the world and is considered fourth in mortality after malaria, Chagas disease and leishmaniasis [1]. The parasite presents the multidrug resistance phenotype (MDR), due to the expression of a surface P-glycoprotein that transports the drug outside the cell, avoiding its therapeutic effects. In amoeba, there are four genes that code for Pgp proteins, *EhPgp1*, *EhPgp2*, *EhPgp5*, and *EhPgp6*. The *EhPgp1* and *EhPgp6* genes are constitutively expressed in drug-resistant mutants (clone C2). The *EhPgp5* gene is induced by the presence of the drug while the *EhPgp2* gene transcript is not detected [2,3]. Differential *EhPgp* genes expression suggests a specific control mechanism of the MDR phenotype in this parasite.

Cloning and transcriptional characterization of the *EhPgp1* and *EhPgp5* gene promoters from drug-sensitive and drug-resistant trophozoites showed that these were 99.7% identical, however differential complexes were formed when nuclear extracts from sensitive and resistant clones were used. These results suggest that specific transcriptional regulators may be involved in the expression of the *EhPgp* genes in drug-resistant cells [4,5]. Until now only some *cis*-regulatory elements [6-11], and very few transcription factors have been identified and characterized in gene expression of this parasite [9,12-16].

Analysis of the core promoter of 37 protein encoding genes of *E. histolytica* revealed three conserved regions: i) the putative TATA element located approximately at -30 (GTATTTAAA(G/C)), ii) the GAAC sequence, located between the TATA box and Inr sequence, and iii) the putative Inr region overlapping the transcription initiation site (AAAAATTCA) [7]. Five major upstream regulatory elements (UREs) are present in the *hgl5* gene promoter, four of them act as positive regulatory elements: URE1, URE2, URE4, URE5, whereas the URE3 motif performs a negative regulatory activity [7]. However URE3 function as a positive regulatory element in the ferredoxin (fdx1) promoter region [17]. Additionally URE1-like sequence was reported as a *cis*-acting element in the *EhRabB* gene promoter [18], and recently was identified as the protein that specifically binds to the URE1 sequence (EhURE1BP), which contains five SNase domains and one Tudor motif [19].

Gilchrist et al. [14] identified the protein that binds to the URE3 element (URE3-BP) and recently demonstrated that several genes of *E. histolytica* are regulated by URE3-BP. The URE3 motif was found in 54% and 39% of promoter regions of the genes modulated by URE3-BP *in vitro* and *in vivo*, respectively [20,21]. On the other hand, Schaenman et al. [13] reported that the URE4 sequence, composed of two 9 bp repeats, functions as an enhancer in the *hgl5* gene and it interacts with two URE4 enhancer-binding proteins of 18 and 28 kDa called EhEBP1 and EhEBP2 [22].

In the *EhPgp1* gene promoter, we identified by homology to consensus sequences reported so far, C/EBP, HOX, GATA-1 and OCT regulatory elements. Some specific oligonucleotides for these elements were able to compete against the DNA promoter in electrophoretic mobility shift assays [4]. Specific deletions of C/EBP elements, demonstrated that two CCAAAT/enhancer binding protein sites (-54 to -43 bp and -198 to -186 bp), were *cis*-acting elements of *EhPgp1* gene expression in both drug-sensitive and resistant trophozoites [9]. In addition, two nuclear proteins of 25 and 65 kDa that were specifically binding to C/EBP probe, share epitopes with the human C/EBP tran-

***Corresponding author:** C. Gómez, Institutional Program of Molecular Biomedicine, School of Medicine and Homeopathy, National Polytechnic Institute (IPN), Guillermo Massieu Helguera No. 239, Fracc. La Escalera, Ticoman, D. F., Mexico, Tel: (+46)-57 29 60 00; Fax: (+46) 57 29 60 00; E-mail: cgomezg@ipn.mx

**Received** August 06, 2012; **Accepted** August 29, 2012; **Published** August 30, 2012

**Citation:** Ramírez ME, Pérez DG, Náder E, Gómez C (2012) Transcriptional Activation by an URE4-like Sequence in the *EhPgp1* Gene Core Promoter. J Bacteriol Parasitol 3:148. doi:10.4172/2155-9597.1000148

**Copyright:** © 2012 Ramírez ME, 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.

scription factor. However functional activities of the *EhPgp1* promoter demonstrated that other sequences within -259 to -206 bp besides C/EBP are crucial for promoter activity [9]. Previously, we demonstrated that in the -234 to -197 bp region, putative cis-activator sequences for GATA-1, GAL 4, NIT-2 and C/EBP transcription factors are present [23]. Here, we report the presence of a *cis*-acting element located at -226 to -218 bp of the *EhPgp1* gene promoter and the putative transcription factor with which it interacts.

## Materials and Methods

### Culture of *E. histolytica* trophozoites

Trophozoites of *E. histolytica* clone C2 (emetine-resistant) were axenically cultured in TYI-S-33 medium [24].

### Search of putative consensus elements by *In silico* analysis

To find consensus sequences for transcription factors in the -234 to -197 bp *EhPgp1* promoter region, we used the transcription factors data base TF Search version 1.3 (http://www.cbrc.jp /research/db/TF-SEARCH.htlm.).

### Plasmid construction of *EhPgp1* promoter

Constructions containing different deletions and mutations of -234 to -197 bp *EhPgp1* promoter region were done. We amplified four different regions: I) from -234 to +24, II) from -222 to +24, III) from -211 to +24 and, IV) from -207 to +24bp, using the sense Pgp1F234 (5'-GCTCTAGAGGGAGTGTAAAAATGTTAT-3'), Pgp1F222 (5'-GCTCTAGATGTTATCTGAAAA-3'), Pgp1F211 (5'-GCTCTAGAAAAAATGTTATCTG-3') and Pgp1F207 (5'-GCTCTAGATGTTATCTGAATTG-3') oligonucleotides respectively, and the antisense Pgp1AS-33 (5'-CCAACGTTAAACTCACTTTCAGTTT-3') oligonucleotide. We generated fragments containing the mutations described below and marked in bold letters, and used the sense R7(3)m4 (5'-GCTCTAGAGGGATGATAAAT-3'),

R7(3)tm (5'-GCTCTAGAGGGAGTGTAAAAATGTTAT-3'),

R9(2)m5(5'-GCTCTAGATATCTGATATTCTAGTTATCT-GAAA-3'),

R9(2)m5(5'-GCTCTAGATATCTGATATTCTAGTTATCT-GAAA-3'), and

R9(2)tm (5'-GCTCTAGATATCTGATTTTCTAATGATCT-3') oligonucleotides, and the antisense Pgp1AS-33 oligonucleotide. We used 50 ng of the plasmid containing the core promoter (p268Pgp1) as template and 2U of Deep Vent Polimerase (New England Biolabs). The PCR products were cloned into XbaI and HindIII restriction sites of the pBSCAT-ACT plasmid (Nickel and Tannich, 1994). The sequence of all constructions was confirmed by DNA sequencing.

### Transfection and promoter activity

Transient transfection experiments were performed by electroporation [25], using $10^6$ trophozoites and 100 μg of plasmid DNA. Electroporated trophozoites were incubated at 37°C for 48 h. Total proteins were obtained and CAT activity was measured by two phase diffusion assays [6] using 100 μg of total proteins, 200 μl of chloramphenicol (1.25 mM), 10 μl of $C^{14}$-butyryl-CoA (4.15 mCi/mmol, NEN Life Science Products) and 4 ml of scintillation solution. The activities were measured at 2 h intervals in the linear range of assay. The assay of CAT activity was expressed as a percentage of the butyrylated derivatives. The background activities obtained from trophozoites transfected with the plasmid without promoter (pBSCAT-ACT) were subtracted from the activities of trophozoites transfected with different promoter constructions.

### Nuclear extracts preparation

Nuclear extracts (NE) from clone C2 trophozoites were obtained as described previously [4]. Protein concentration was determined by the Bradford method [26].

### Electrophoretic mobility shift assay (EMSA)

EMSA was performed as described [4]. Briefly, we used a double stranded Pgp1-226/218R3 oligonucleotide containing three copies of the R9(2) region, from -226 to -218 bp, marked in bold face (5'-**AAAAATGTT**ATCTGA**AAAAATGTT**ATCTG **AAAAAAT-GTT**-3'). Oligonucleotides were [γ-$^{32}$P]-ATP (3000 mCi/mmol) labeled using T4 polynucleotide kinase (Invitrogene). Binding reactions contained 1 ng of radiolabel probe, 20 μg of NE, 1 μg of poly d(I-C) (Amersham Pharmacia Biotech) and DNA-protein binding buffer. Competition assays were carried out using a 150 fold molar excess of the same unlabeled double stranded oligonucleotide or poly d(I-C). Mixtures were incubated 10 min on ice. The complexes were separated by electrophoresis on 6% non-denaturing poly-acrylamide gels in 0.5 X TBE, for 2.5 h at 120 V. Gels were dried and subjected to Phosphor Imager analysis (Bio-Rad). To determine the interaction of nuclear proteins with the sequence R9, we used also as a specific competitors three additional oligonucleotides, the Pgp1-226/218R2 containing two copies of these sequence (5'-**AAAAATGTT**ATCTG**AAAAAAT-GTT**ATCTGA-3'), the Pgp1-226/218R1 containing only one copy (5'-**AAAAATGTT**ATCTGA-3') and the Pgp1-234/197 oligonucleotides containing the wild type *EhPgp1* promoter sequence from -234 to -197 bp (5'-TATCTGAT**AAAAATGT T**ATCTG**AAAAA ATGT-T**ATCTGA-3').

### DNA binding protein purification

*E. histolytica* nuclear proteins that bind to the R9 region were partially purified using a DNA-binding protein purification kit (Roche). The Pgp1-226/218R3 double stranded oligonucleotide was used as DNA probe to obtain concatameric DNA using self primer PCR. These oligonucleotides undergo a self-priming reaction during the PCR, resulting in long concatamers with hundreds of specific binding sites. Concatameric DNA was bound to magnetic particles, then 50 μg of *E. histolytica* NE was added. The specific proteins were captured by the concatameric oligonucleotide with a high affinity constant, while non-specific proteins (with a lower affinity constant) did not bind. The specific DNA-binding proteins were eluted from the immobilized particle with a high ionic strength buffer. After removal of the elution buffer by filtration with centricon YM-10 (Millipore), the proteins were transferred to nitrocellulose membranes for Western blot assays. The eluted proteins were also evaluated by supershift assays.

### Western blot assays

Partially purified proteins and NE were separated on a 12% polyacrylamide gel, transferred to nitrocellulose membranes, and immunoblotted under standard conditions. Membranes were blocked with 4% non fat milk in PBS pH 7.4 /Tween 0.05%, for 2-3 h at room temperature, and then incubated with mouse anti-EhEBP1 antibody (1:600) (kindly supplied by Dr. Carol A. Gilchrist) for 1 h at 37°C. As control, we used the rabbit polyclonal antibody against the human C/EBPβ (Santa Cruz Biotechnology) (1:500). Immunoreactivity was detected

**Citation:** Ramírez ME, Pérez DG, Náder E, Gómez C (2012) Transcriptional Activation by an URE4-like Sequence in the *EhPgp1* Gene Core Promoter. J Bacteriol Parasitol 3:148. doi:10.4172/2155-9597.1000148



by a chromogenic method using anti-mouse and anti-rabbit peroxidase labeled secondary antibodies, respectively (Zymed laboratories) (1:3000) and revealed with $H_2O_2$ and 4-chloro-1-Napthol.

### Supershift assays

Supershift assays were performed as the EMSA described before, using the antibodies in the reaction mix. Briefly, we used as radio labeled probe 1 ng of the Pgp1-226/218R3 double stranded oligonucleotide, 20 µg of NE, mouse anti-EhEBP1, 1 µg of poly d(I-C) and DNA-protein binding buffer. As negative controls, anti-C/EBPβ human antibody was used (Santa Cruz Biotechnology).

## Results

### Identification of consensus sequences in the region from -234 to -197 bp

The structural analysis of the region from -234 to -197 bp in the *EhPgp1* gene core promoter showed three repeated sequences of 7 bp each, located at the positions -203 to -197; -218 to -212; and -234 to -228 bp (Figure 1), we have named these sequences R7(1), R7(2), and R7(3), respectively. Also, we identified two repeats of 9 bp each located at the positions -211 to -203 and -226 to -218 bp, these sequences were named R9(1) and R9(2), respectively. In silico analysis of the -234 to -197 bp region was performed to identify potential nuclear factor binding sites. Interestingly, in the R7 sequences we detected consensus sequences for GAL4, GATA-1, C/EBPβ and NIT-2 transcription factors, as was previously reported [23]. We also identified two new consensus sequences for the GATA-2 and NF-GMb transcription factors. Additionally we localized three consensus sequences for C/EBPβ and two sequences for EhEBP1 and EhEBP-2 transcription factors that overlap with the R9 motifs (Figure 1).

### The distal region of the *EhPgp1* promoter, between -234 to -197 bp contains an activator sequence

Using structural and *in silico* analysis of the promoter, we performed a series of deletions on the -234 to -207 bp promoter to locate *cis*-elements that could drive the *EhPgp1* gene expression. Four different plasmids (p258Pgp1, p246Pgp1, p235Pgp1, and p231Pgp1) carrying -234 to +24, -222 to +24, -211 to +24 and -207 to +24 bp sequence

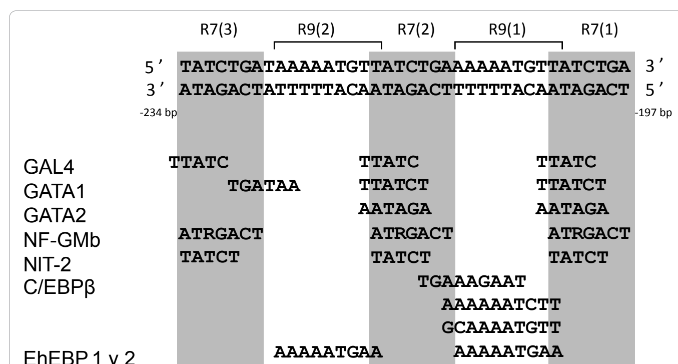


**Figure 1:** Putative consensus sequences for transcription factors in the region from -234 to -197 pb of EhPgp1 *E. histolytica* gene core promoter. Upper scheme shows the double stranded DNA from the -234 to -197 bp of the EhPgp1 promoter, relative to the transcription initiation site. Dash rectangles are the R7 repeated sequences. Brackets show the R9 repeated sequences. Down panel, the -234 to -197 bp region was searched against the TF Search database to detect potential consensus sequences. Previously and newly described protein-binding sites are located within the R7 or R9 repeated sequences. R, indicates A or G nucleotides.

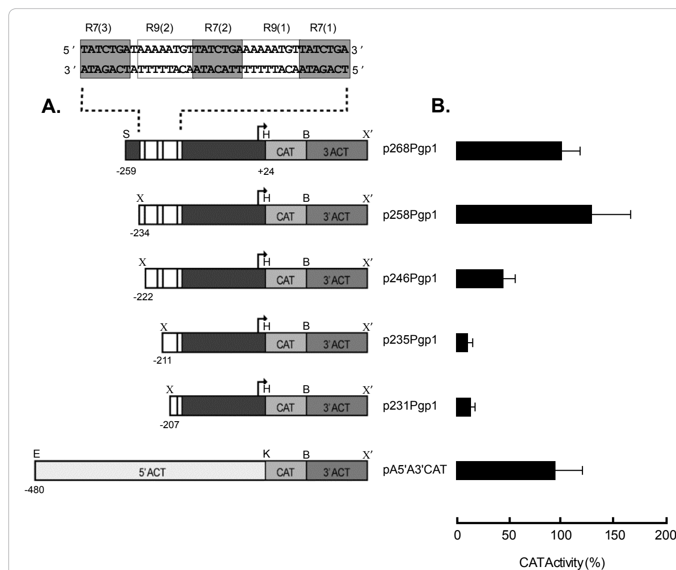


**Figure 2:** Activity of the EhPgp1 gene promoter (-234 to -197 bp) in drug resistant trophozoites. At the top are located the R7 and R9 repeated sequences. (A) Schematic representation of the plasmids containing different deletions of the EhPgp1 gene core promoter. The numbering refers to the transcription initiation site, designated as +1. All plasmids contain +24 bp downstream from the ATG of the EhPgp1 gene, the CAT reporter gene and the 3'-flanking actin region (3'ACT). The p268Pgp1 plasmid contains the EhPgp1 gene core promoter. The pA5'A3'CAT plasmid, used as positive control, contains 480 bp fragment of the actin gene promoter. Arrows, transcription initiation site, B, BamHI; E, EcoRI; H, HindIII; K, KpnI; S, SmaI; X, XbaI; X', XhoI. (B) CAT activity (%) of the trophozoites transfected with the corresponding plasmids. Bars, average of CAT activities ± S.D. representative of three independent experiments performed in duplicate.

of the *E. histolytica EhPgp1* gene core promoter were constructed and transfected into C2 trophozoites. CAT activities were measured and compared to the p268Pgp1 plasmid (positive control). Results showed a marked reduction in CAT activity (58%) after truncation of -234 to -222 bp (p246Pgp1 plasmid), suggesting that the R7(3) and R9(2) sequences are required for the *EhPgp1* gene expression (Figure 2).

As deletions progress toward the 3'-end of the promoter (p235Pgp1 and p231Pgp1 plasmids), in which almost all the repeated regions were eliminated, a decrease of 87 CAT activities were observed, with both constructions (Figure 2). These results provide evidence that the R7(3) and the R9(2) repeated sequences are involved in the *EhPgp1* transcriptional activation, but also showed the presence of another positive regulatory sequence, between the position -218 to -211 bp that correspond to the R7(2) sequence (Figure 2).

Promoter activity comparison of p258Pgp1 (-234 to +24 bp) and p268Pgp1 (-266 to +24 bp) demonstrated that 5' deletion up to -234 bp increase the promoter activity in 27%, suggesting that the region from -259 to -234 bp could contain negative *cis*-regulatory elements.

### A 9 bp repeated sequence is critical for driving *EhPgp1* gene expression

To determine if R7(3) or R9(2) or both repeated sequences produced the major effects on promoter activity, we performed point mutations into the R7(3) and R9(2) core sequences. Mutations of one or two bases on different positions of the repeated sequences did not significantly modify the CAT reporter gene activity (data no shown). Thus, we carried out constructions containing more than three point

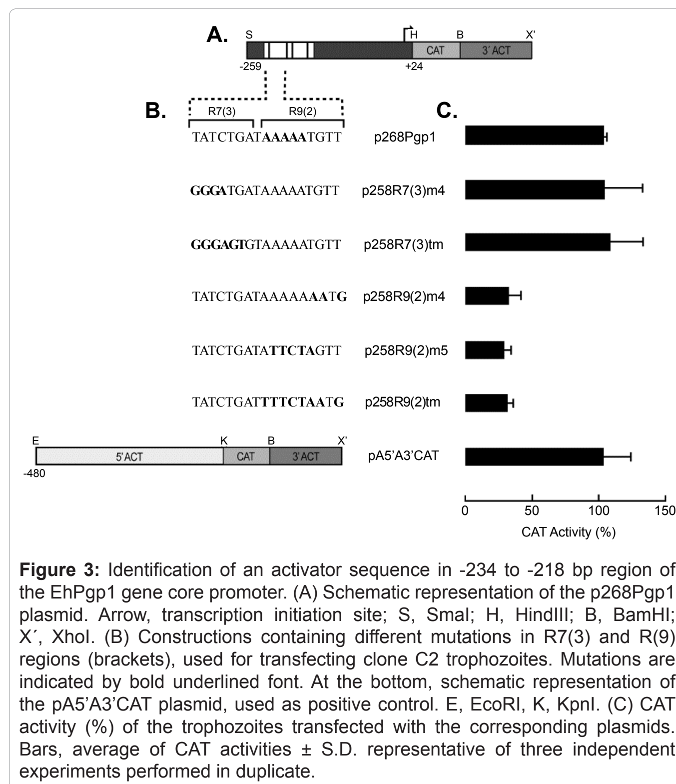


**Figure 3:** Identification of an activator sequence in -234 to -218 bp region of the EhPgp1 gene core promoter. (A) Schematic representation of the p268Pgp1 plasmid. Arrow, transcription initiation site; S, SmaI; H, HindIII; B, BamHI; X´, XhoI. (B) Constructions containing different mutations in R7(3) and R(9) regions (brackets), used for transfecting clone C2 trophozoites. Mutations are indicated by bold underlined font. At the bottom, schematic representation of the pA5'A3'CAT plasmid, used as positive control. E, EcoRI, K, KpnI. (C) CAT activity (%) of the trophozoites transfected with the corresponding plasmids. Bars, average of CAT activities ± S.D. representative of three independent experiments performed in duplicate.

mutations. Neither mutations of the R7(3) sequence (p258R7(3)m4 and p258R7(3)tm plasmids) significantly affect *EhPgp1* promoter activity (Figure 3). However, three, five or eight mutations in the R9(2) site (p258R9(2)m4, p258R9(2)m5, and p258R9(2)tm plasmids) reduced CAT activity by 70%, compared to a wild-type construction (p268Pgp1) (Figure 3). These data strongly suggest that the R9(2) sequence is crucial for *EhPgp1* gene transcription.

## R9 sequences DNA-protein interactions

Based on our observations that R9(2) could potentially up-regulate the transcriptional activation of the *EhPgp1* promoter, we investigated the ability of this repeated sequence to bind nuclear proteins from *E. histolytica*. Moreover, R9 is present two times in this region, we also analyzed the relevance of the presence of one, two or three repeated sequences in the DNA-protein complexes formation. To perform these, we generated a set of double stranded oligonucleotides, the Pgp1-226/218R3, the Pgp1-226/218R2 and the Pgp1-226/218R1, containing three, two, and one copy of the R9 repeated respectively and the same 6 bp present between each one, as the wild type promoter sequence. The Pgp1-226/218R3 oligonucleotide was used as probe to perform electrophoretic mobility shift assays, and the Pgp1-226/218R2, Pgp1-226/218R1 and Pgp1-234/197 (wild type promoter sequence) oligonucleotides were used as specific competitors.

Incubation of the probe with NE from clone C2 trophozoites resulted in the formation of three DNA-protein complexes called a, b and c of varying intensity (Figure 4, lane 2). To demonstrate the specificity of binding by this R9 repeated, we added an unspecific competitor which failed to compete off any DNA-protein complexes formation (Figure 4, lane 3). However, all the complexes were specifically competed with the same cold-fragment (Figure 4, lane 4). To also define whether one or more R9 sequences are necessary for the DNA-protein complexes formation, we performed competition assays using the Pgp1-226/218R2 and the Pgp1-226/218R1 double stranded oligonucleotides, both completely competed the formation of the complex a and produced the formation of a new complex with a minor electrophoretic mobility (complex d) (Figure 4, lanes 5 and 6). While the intensity of the complex b was reduced by 52% when two R9 repeats are present and a reduction of only 4% was detected when only one repeat was used in the competitor (Figure 4, lanes 5 and 6). Similarly, the intensity of complex c was reduced by 71 and 10% with two and one R9 motifs respectively (Figure 4, lanes 5 and 6). To further characterize if the *EhPgp1* promoter wild type sequence (-234 to -197 bp) form the same DNA-protein complexes observed with Pgp1-226/218R3, we added as specific competitor the Pgp1-234/197 double stranded oligonucleotide. Interestingly, all the complexes competed. Together these results demonstrate that the R9 repeated could serve as a recognition sequence for a DNA binding protein in the parasite extract.

## An EhEBP1 recognized the R9 repeated sequences

As was observed before in Figure 1, the sequences R9 contain almost the complete *EhEBP1* recognition sequence (URE4). To identify whether this protein was a component of the gel shift complexes of R9 DNA with NE from the trophozoites, we partially-purified the proteins interacting with it by DNA affinity chromatography.

Thus, we performed a Western blot assay using NE from trophozoites, the partially-purified fraction and the antibody against the EhEBP1 protein (kindly supplied by Dr. Carol A. Gilchrist). The results revealed the presence of a specific band of 28 kDa in NE and in the protein fraction (Figure 5A, lanes 1 and 2) indicating that one of the proteins that bind to R9 is an EhEBP1-like protein. In contrast, unrelated antibodies against to human C/EBPβ did not produce any detection (Figure 5A, lane 3).

To confirm the observation that R9 repeated sequence is recognized by an EhEBP1, we performed supershift assay using anti-EhEBP1 antibody. The results did not show a supershifted band, but caused the

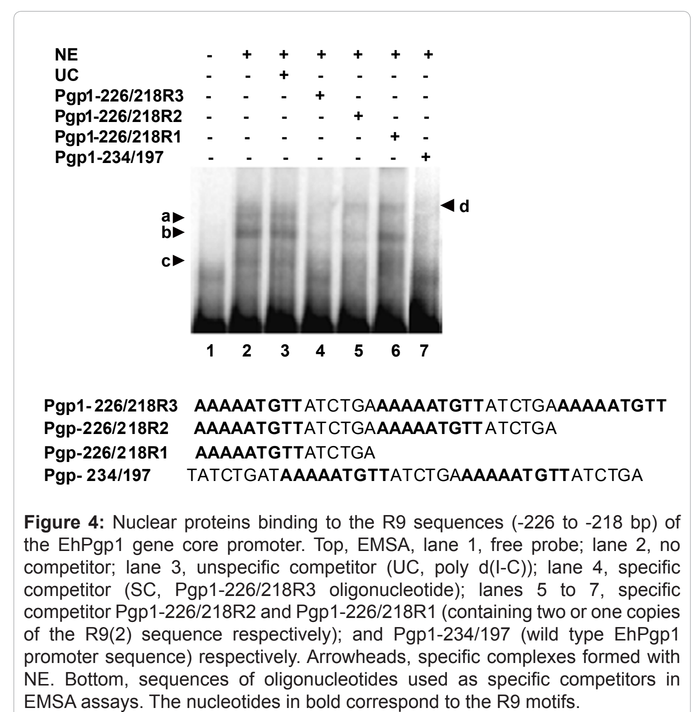


**Figure 4:** Nuclear proteins binding to the R9 sequences (-226 to -218 bp) of the EhPgp1 gene core promoter. Top, EMSA, lane 1, free probe; lane 2, no competitor; lane 3, unspecific competitor (UC, poly d(I-C)); lane 4, specific competitor (SC, Pgp1-226/218R3 oligonucleotide); lanes 5 to 7, specific competitor Pgp1-226/218R2 and Pgp1-226/218R1 (containing two or one copies of the R9(2) sequence respectively); and Pgp1-234/197 (wild type EhPgp1 promoter sequence) respectively. Arrowheads, specific complexes formed with NE. Bottom, sequences of oligonucleotides used as specific competitors in EMSA assays. The nucleotides in bold correspond to the R9 motifs.

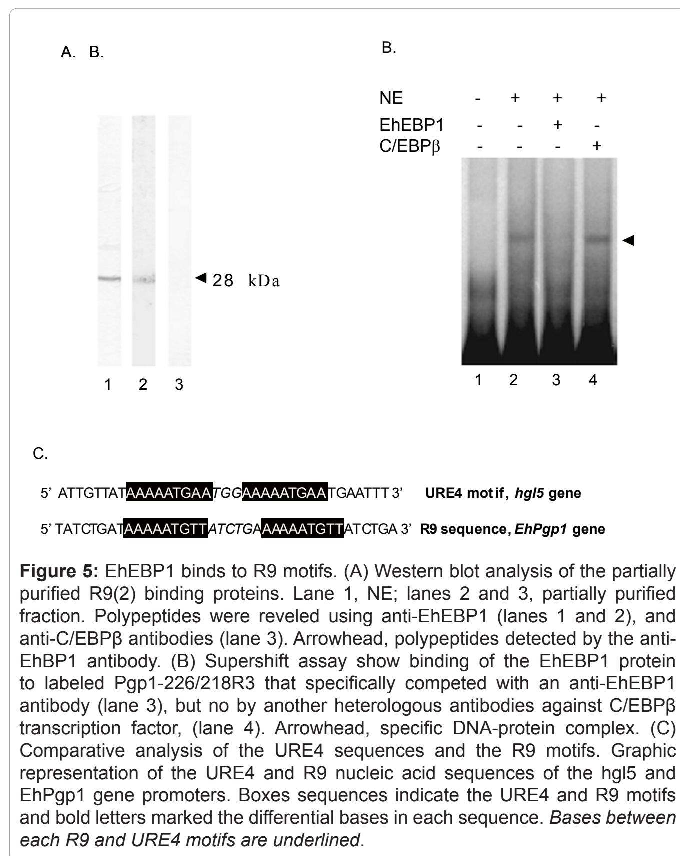


**Figure 5:** EhEBP1 binds to R9 motifs. (A) Western blot analysis of the partially purified R9(2) binding proteins. Lane 1, NE; lanes 2 and 3, partially purified fraction. Polypeptides were reveled using anti-EhEBP1 (lanes 1 and 2), and anti-C/EBPβ antibodies (lane 3). Arrowhead, polypeptides detected by the anti-EhBP1 antibody. (B) Supershift assay show binding of the EhEBP1 protein to labeled Pgp1-226/218R3 that specifically competed with an anti-EhEBP1 antibody (lane 3), but no by another heterologous antibodies against C/EBPβ transcription factor, (lane 4). Arrowhead, specific DNA-protein complex. (C) Comparative analysis of the URE4 sequences and the R9 motifs. Graphic representation of the URE4 and R9 nucleic acid sequences of the hgl5 and EhPgp1 gene promoters. Boxes sequences indicate the URE4 and R9 motifs and bold letters marked the differential bases in each sequence. *Bases between each R9 and URE4 motifs are underlined.*

fading of DNA-protein complexes (Figure 5B, lane 3), while the addition of heterologous antibody against C/EBPβ transcription factor (Figure 5B, lane 4) had no effect, confirming the binding of EhEBP1 to the R9 repeated sequence and the specificity of immunoreactions with anti-EhEBP1 antibody.

## Discussion

In this study, we have identified a *cis*-acting element that controls the *EhPgp1* gene expression in drug resistant trophozoites (clone C2) of *E. histolytica*. Since *EhPgp1* gene is implicated in the multidrug resistance of this parasite, the study of its transcriptional control in the clone C2 trophozoites provide an excellent *in vitro* system for studying molecular basis of *EhPgp1* gene regulation. Our previous work has defined that transcriptional regulation of the *EhPgp1* gene promoter depends on two CCAAT/enhancer binding sites (-54 to -43 and -198 to -186 bp) and other motifs present in the 53 bp upstream of the C/EBPIII site [9]. Moreover into the 53 bp of this promoter we delimited a functional region of 38 bp (-234 to -196 bp) that interact with nuclear proteins from *E. histoytica* and by *in silico* analysis showed the presence of GATA-2 and NF-GMb sequences as well as the GAL4, NIT2, GATA-1, and C/EBP binding sites [23]. However, although we identified several consensus sequences in this region, in the present work two interesting types of repeat sequences were also located. One of them are the R7 repeats which are present three times (-203 to -197, -218 to-212 and -228 to -234 bp) and the others are the R9 repeats located in two positions (-203 to -211 and -218 to -226 bp), at the *EhPgp1* promoter.

Deletions of the *EhPgp1* region (-234 to -197 bp) showed that the R7(3) and/or R9(2) repeated sequences could be necessary for the *EhPgp1* gene expression. One or two point mutations into these sequences did not produce any CAT activity modification, indicating that the DNA-protein interactions were not modified by these point changes. Four or more mutations into R7(3) did not affect promoter activity, suggesting that R7(3) is not important for the *EhPgp1* gene expression; however, we could not ignore the possible participation of the R7(2) and R7(1) sequences in the transcriptional regulation of the *EhPgp1* gene. On the other hand, we found that the R9(2) repeat was important for promoter activity, because its deletion or mutations in three bases at the 3' side (AAAAAAATG), five mutations at the middle of the sequence (ATTCTAGTT) or eight mutations of the nine bases (TTTCTAATG) produced a 70% reduction in CAT activity. These results clearly demonstrate that the R9 repeated sequences are necessary for the *EhPgp1* gene expression. Interestingly, similar results were observed when the URE4 sequence was identified and characterized in the *E. histolytica hgl5* promoter. Four mutations of the middle residues (AATCTAGAA) or in the 3' side (AAAAAAATG) within the upstream repeat or mutations into the two repeated sequences produced an 85% and 93% reduction in luciferase activity in the last two conditions, respectively [22]. Additionally, they found that the upstream repeat is more relevant for the promoter activity than the downstream one, because mutations into the upstream repeat diminish 85% luciferase reporter gene of wild type levels, while mutations into the downstream repeat only decreased 39% reporter gene. They suggested that the downstream repeat may play a role of supporting binding of factors to the upstream repeat, as evidenced by the fact that separating the repeats by seven base pairs decreased reporter gene activity to 22% [22]. These results are consistent with our findings, because the deletion of the R9(2) repeat drastically diminish CAT reporter gene activity (58%), while the elimination of both R9 repeats produced an additional activity reduction of 29%. The existence of a synergistic and accumulative R9 effect involved in the *EhPgp1* gene control may be possible.

Interestingly, the appearance of R9 repeat sequences in URE4 motif is not only important for their functional role, but also by the times

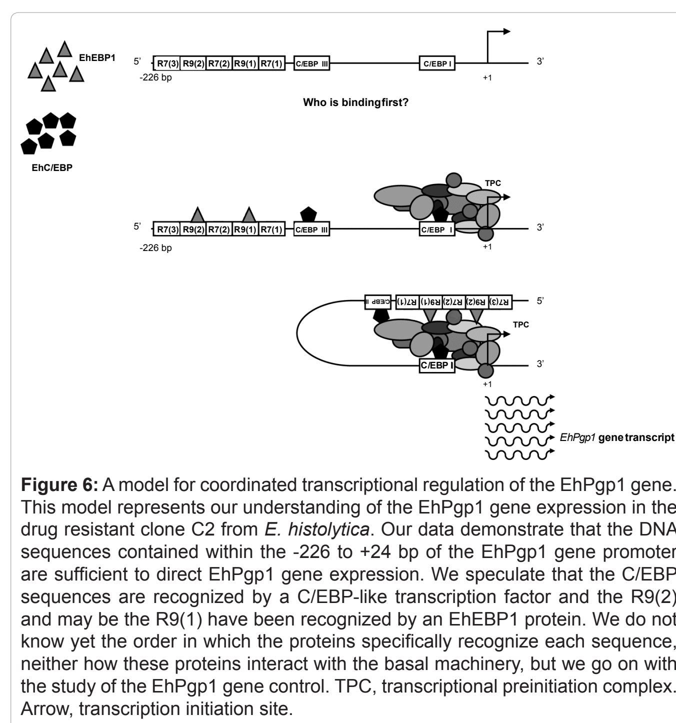


**Figure 6:** A model for coordinated transcriptional regulation of the EhPgp1 gene. This model represents our understanding of the EhPgp1 gene expression in the drug resistant clone C2 from *E. histolytica*. Our data demonstrate that the DNA sequences contained within the -226 to +24 bp of the EhPgp1 gene promoter are sufficient to direct EhPgp1 gene expression. We speculate that the C/EBP sequences are recognized by a C/EBP-like transcription factor and the R9(2) and may be the R9(1) have been recognized by an EhEBP1 protein. We do not know yet the order in which the proteins specifically recognize each sequence, neither how these proteins interact with the basal machinery, but we go on with the study of the EhPgp1 gene control. TPC, transcriptional preinitiation complex. Arrow, transcription initiation site.

that are present in the *EhPgp1* promoter and by its sequence. Each R9 repeat contains 9 bp (AAAAATGTT) and is separated from the other by 6 bp. These elements present a similar arrangement to the URE4 motif identified in the *E. histolytica* Gal/GalNac lectin heavy subunit hgl5 gene promoter [22]. The URE4 sequence is also composed of two 9 bp repeats (AAAAATGAA) but separated by only 3 bp. Two main differences between URE4 and R9 sequences are the last two bases in the sequences (AA/TT) and the distance between them, nevertheless they are very similar, suggesting that the R9 repeats form an URE4 element (Figure 5C).

Moreover, our DNA-protein interaction assays strongly suggest that R9 motifs have been specifically recognized by nuclear proteins from amoeba and that these proteins required at least the presence of two R9 sequences. Thus because in competition assays the DNA-protein complexes disappeared or their intensity was strongly diminished (more than 70%) when three or two R9 motifs were used as competitors. Whereas one copy of this motif did not modify the DNA-protein complexes formation except the complex a, indicating that one R9 is able to interact with amoeba proteins. However, this may not happen with the same affinity and may diminish the stability of the DNA-protein interaction. In addition we observed the formation of a new low mobility complex showing that DNA-protein interactions with one, two or three R9 motifs generate different mobility complexes. Our results provide evidence that probably, the transcription factors bind in an independent fashion on each R9 sequence or at the same time to allow protein-protein interactions for the formation of a stable dimeric complex. A similar point of view was suggested by Shaenman et al. [13] because they observed a decrease in the *hgl5* promoter activity when they modified the spacing between the URE4 repeats. In other systems this kind of interactions has been also reported like the protein binding to the retinoic acid response element or the participation of the NIT2 nitrogen regulatory protein of *N. crassa* in the turning on expression of different structural genes through its binding to GATA [27]. The monomeric form of this protein can bind to the GATA sequence with very low affinity, however when two GATA sequences are in close proximity, two monomeric NIT2 proteins can interact with each other to stabilize the interaction [27]. These results demonstrate that the presence of two R9 motifs favors its recognition by nuclear proteins from *E. histolytica* and may be activating the *EhPgp1* gene expression.

Finally, because the R9 repeats may be an URE4 element (Figure 5C), it is possible that the EhEBP1 protein can bind to the R9 motifs of the *EhPgp1* promoter to induce its expression. This assumption is supported by two findings: i) the Western blot assays performed with the partially purified R9 binding proteins using the anti-EhEBP1 antibody revealed a 28 KDa protein, and ii) in the super shift assays, DNA-protein complexes totally competed with the anti-EhEBP1 antibody. Consistent with these results, previous reports [13,22] showed that antibodies against EhEBP1 recognized two proteins of 28 and 18 KDa in *E. histolytica* extracts and were able to compete the URE4-protein interaction of the hgl5 promoter. Interestingly, overexpressions of the 28 KDa EhEBP1 repress the *hgl5* gene transcription through its ability to recognize the URE4 motif. In our case, EhEBP1 could be participating in the transcriptional activation of the *EhPgp1* gene through its interaction with the R9 repeats. Several transcription factors like Sp3, GATA and C/EBP between others have been reported to act both as a positive or negative regulator of transcription depending on promoter and cell type [28-30]. A similar role seems to be occurring with the EhEBP1 in the amoeba; however, more in depth investigation will be required to know more about this protein and its functional role in the transcriptional control of the *EhPgp1* and other genes.

Taking these results together and in concordance with our previous reports [9,23], we proposed a model to address molecular insights into regulation of the *EhPgp1* gene. In this model, we showed that transcriptional control of the *EhPgp1* gene is coordinated by C/EBPI, C/EBPIII and R9 motifs that have been recognized by C/EBP-like and EhEBP1 transcription factors respectively. Then, these complexes may interact with basal machinery of transcription and activate the *EhPgp1* gene (Figure 6). However, the chronological events that drive the binding of each transcription factor and how these interact to enhance the *EhPgp1* transcriptional activation, poses an open question to be solved.


## Acknowledgement

This research was supported by CONACYT project P39986M and by SIP projects. We would like to thank Dr. Carol A. Gilchrist from Department of Internal Medicine, University of Virginia, USA, for the anti-EhEBP1 antibody, Dr. Kevin McGuigan from Royal College of Surgeons, Ireland, for critical reading of the manuscript, and Veronica Castillo Chavez for graphical design.


## References


1. Espinosa-Cantellano M, Martínez-Palomo A (2000) Pathogenesis of intestinal amebiasis: from molecules to disease. Clin Microbiol Rev 13: 318-331.
2. Descoteaux S, Ayala P, Samuelson J, Orozco E (1995) Increase in mRNA of multiple Eh pgp genes encoding P-glycoprotein homologues in emetine-resistant *Entamoeba histolytica* parasites. Gene 164: 179-184.
3. Orozco E, Pérez DG, Gómez MC, Ayala P (1995) Multidrug resistance in *Entamoeba histolytica*. Parasitology Today 11: 473-475.
4. Gómez C, Pérez DG, López-Bayghen E, Orozco E (1998) Transcriptional analysis of the EhPgp1 promoter of *Entamoeba histolytica* multidrug resistant mutant. J Biol Chem 273: 7277-7284.
5. Pérez DG, Gómez C, López-Bayghen E, Tannich E, Orozco E (1998) Transcriptional analysis of the EhPgp5 promoter of *Entamoeba histolytica* multidrug resistant mutant. J Biol Chem 273: 7285-7292.
6. Buss H, Lioutas C, Dobinsky S, Nickel R, Tannich E (1995) Analysis of the 170-kDa lectin gene promoter of *Entamoeba histolytica*. Mol Biochem Parasitol 72: 1-10.
7. Purdy J, Pho L, Mann B, Petri W Jr (1996) Upstream regulatory elements controlling expression of the *Entamoeba histolytica* lectin. Mol Biochem Parasitol 78: 91-103.
8. Hidalgo E, Orozco E (1997) Structural characterization of the *Entamoeba histolytica* enolase gene promoter. Arch Med Res 28: 46-48.
9. Marchat LA, Gomez C, Perez D, 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.
10. Nieto A, Guillermo Pérez D, Orozco, E, Paz F, Gómez C (2005) *Entamoeba histolytica* EhPgp5 transcriptional activation depends on putative emetine response elements. Exp Parasitol 110: 233-237.
11. Ehrenkaufer GM, Hackney JA, Singh U (2009) A developmentally regulated Myb domain protein regulates expression of a subset of stage-specific genes in *Entamoeba histolytica*. Cell Microbiol 11: 898-910.
12. Luna-Arias JP, Hernández-Rivas R, de Dios-Bravo G, García J, Mendoza L, 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: 33-40.
13. Schaenman JM, Gilchrist CA, Mann BJ, Petri WA Jr (2001) Identification of two *Entamoeba histolytica* sequence-specific URE4 enhancer-binding proteins with homology to the RNA-binding motif RRM. J Biol Chem 276: 1602-1609.
14. Gilchrist CA, Holm CF, Hughes MA, Schaenman JM, Mann BJ, et al. (2001) Identification and characterization of an *Entamoeba histolytica* upstream regulatory element 3 sequence-specific DNA-binding protein containing EF-hand motifs. J Biol Chem 276: 11838-11843.

15. de Dios-Bravo G, Luna-Arias JP, Riverón AM, Olivares-Trejo JJ, López-Camarillo C, et al. (2005) *Entamoeba histolytica* should be in italics TATA-box binding protein binds to different TATA variants in vitro. FEBS J 272: 1354-1366.

16. Garcia CG, Arguelles MLM, Ishiwara DGP, Tecla MS, Martinez AN, et al. (2007) A novel heat shock transcription factor family in *Entamoeba histolytica*. Am J Infect Dis 3: 115-122.

17. Gilchrist CA, Mann BJ, Petri WA Jr (1998) Control of ferredoxin and Gal/Gal/Nac lectin gene expression in *Entamoeba histolytica* by a cis-acting DNA sequence. Infect Immun 66: 2383-2386.

18. Romero-Díaz M, Gomez C, Lopez-Reyes I, Martinez MB, Orozco E, et al. (2007) Structural and functional analysis of the *Entamoeba histolytica* EhrabB gene promoter. BMC Mol Biol 8: 82.

19. Calixto-Gálvez M, Romero-Díaz M, García-Muñoz A, Salas-Casas A, Pais-Morales J, et al. (2011) Identification of a polypeptide containing Tudor and staphyloccocal nuclease-like domains as the sequence-specific binding protein to the upstream regulatory element 1 of *Entamoeba histolytica*. Int J Parasitol 41: 775-782.

20. Gilchrist CA, Baba DJ, Zhang Y, Crasta O, Evans C, et al. (2008) Targets of the *Entamoeba histolytica* transcription factor URE3-BP. PLoS Negl Trop Dis 2: e282.

21. Gilchrist CA, Moore ES, Zhang Y, Bousquet CB, Lannigan JA, et al. (2010) Regulation of virulence of *Entamoeba histolytica* by the URE3-BP transcription factor. MBio 1: e00057-10.

22. Schaenman JM, Driscoll PC, Hockensmith JW, Mann BJ, Petri WA Jr (1998) An upstream regulatory element containing two nine basepair repeats regulates expression of the *Entamoeba histolytica* hgl5 lectin gene. Mol Biochem Parasitol 94: 309-313.

23. Ramírez ME, Pérez DG, Náder E, Gómez C (2005) *Entamoeba histolytica*: Functional characterization of the -234 to -196 bp promoter region of the multidrug resistance EhPgp1 gene. Exp Parasitol 110: 238-243.

24. Diamond LS, Harlow DR, Cunnick CC (1978) A new medium for axenic cultivation of *Entamoeba histolytica* and other Entamoeba. Trans R Soc Trop Med Hyg 72: 431-432.

25. Nickel R, Tannich E (1994) Transfection and transient expression of chloramphenicol acetyltransferase gene in the protozoan parasite *Entamoeba histolytica*. Proc Natl Acad Sci U S A 91: 7095-7098.

26. Bradford MM (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.

27. Feng B, Xiao X, Marzluf GA (1993) Recognition of specific nucleotide bases and cooperative DNA binding by the trans-acting nitrogen regulatory protein NIT2 of *Neurospora crassa*. Nucleic Acids Res 21: 3989-3996.

28. Sapetschnig A, Koch F, Rischitor G, Mennenga T, Suske G (2004) Complexity of translationally controlled transcription factor Sp3 isoform expression. J Biol Chem 279: 42095-42105.

29. Yu M, Riva L, Xie H, Schindler Y, Moran TB, et al. (2009) Insights into GATA-1 mediated gene activation versus repression via genome-wide chromatin occupancy analysis. Mol Cell 36: 682-695.

30. Bedi R, Du J, Sharma AK, Gomes I, Ackerman SJ (2009) Human C/EBP-epsilon activator and repressor isoforms differentially reprogram myeloid lineage commitment and differentiation. Blood 113: 317-327.