

Original article:

## Stage-specific expressions of four different ribonuclease H genes in *Leishmania*

Smita Misra, Anitra L. Farrow, Gautam Chaudhuri\*

Division of Microbial Pathogenesis and Immune Response, Department of Biomedical Sciences, Meharry Medical College, 1005 D. B. Todd, Jr. Blvd., Nashville, TN 37208. Phone # (615) 327-6499; Fax (615) 327-5559 email: [gchaudhuri@mmc.edu](mailto:gchaudhuri@mmc.edu) (\*corresponding author)

### ABSTRACT

The human pathogen of the genus *Leishmania* cause worldwide morbidity and infection of visceral organs by some species may be lethal. Lack of rational chemotherapy against these pathogens dictates the study of differential biochemistry and molecular biology of the parasite as compared to its human host. The ubiquitous enzyme ribonuclease H (RNase H, EC 3.1.26.4) cleaves the RNA from a RNA:DNA duplex and is critical for the replication of DNA in the nucleus and the mitochondria. We have characterized four RNase H genes from *Leishmania*: one is of type I (LRNase HI) and three others are of type II (LRNase HIIA, -HIIB and -HIIC). In contrast human cells have only one type I and one type II RNase H. All the four RNase H genes in *Leishmania* are single copy and located in discrete chromosomes. When expressed inside RNase H-deficient *E. coli*, all of the four *Leishmania* RNase H were capable to complement the genetic defect of the *E. coli*, indicating their identity as RNase H. The enzymes are differentially expressed in the promastigotes and the amastigotes, the forms that thrives in entirely different physico-chemical conditions in nature. Nucleotide sequences of the 5'-UTRs of three of these mRNAs have upstream open reading frames. Understanding the regulation of these four distinct ribonuclease H genes in *Leishmania* will help us better understand leishmanial parasitism and may help us to design rational chemotherapy against the pathogen.

**Keywords:** *Leishmania major*, *Trypanosoma brucei*, *Trypanosoma cruzi*, Amastigote, Promastigotes, Ribonuclease H, Upstream open reading frame

**Abbreviations:** LRNase H, *Leishmania* ribonuclease H; UTR, untranslated region; RACE, rapid amplification of cDNA ends; uORF, upstream open reading frame

### INTRODUCTION

*Leishmania* are the vector borne protozoan pathogens and the etiological agents for the various tropical diseases collectively known as leishmaniasis (Alexander et al. 1999; Desjeux, 1994). *Leishmania* is spread through the bite of the infected sandfly when it takes a blood

meal (Desjeux, 1994; Alexander et al. 1999). Leishmaniasis is a public health threat widely distributed over the Central to South America, some parts of Asia and Indian subcontinent. It infects 12 million individuals annually in over 88 countries of the world with 50 million at risk each year (<http://www.who.int/emc/disease/leish/leis.ht>

ml). *Leishmania* exists as extracellular flagellated promastigotes in the gut of the fly vector and as the intracellular non-motile amastigotes in the macrophages of the infected patients (Desjeux, 1994; Alexander et al. 1999). Amastigotes multiply in the macrophages and eventually lyse them. *Leishmania* and other related protozoa of the genus *Trypanosoma* (*T. brucei* and *T. cruzi*) are considered as primitive eukaryotes. Fortunately, although these early eukaryotes have many similarities in metabolic pathways and vital molecular structures with humans, they harbor many unique biochemical features that can be exploited for the development of rational chemotherapy against these human pathogens. The diseases caused by these pests are incurable in the chronic stage and often the disease is not detectable at the initial stages of infection using the currently available methods. Current drug treatments are effective only for the initial infection and have substantial toxicity. The available drugs are very costly and unaffordable by the poor and marginalized population, which are the major victims. Therefore, there is an acute need to exploit molecular targets unique to the parasite for the design of anti-trypanosomal drugs that are efficacious and safe.

Ribonuclease H (RNase H) has unique function to degrade specifically the RNA moiety of DNA/RNA duplex (Crouch and Toulme, 1998; Wu et al. 1999). RNase H activity has been suggested to be involved in DNA replication, recombination repair and transcription (Pileur et al. 2000; Engel and Ray, 1999). Genes encoding RNase H are ubiquitously present in all the kingdoms of life starting from viruses to the humans (Ohtani et al. 2004a; Ohtani et al. 2004b). As this enzyme is involved in the vital metabolism of the cell, any differential biochemistry and molecular biology of these enzymes in the parasites as compared to that of the host could be exploited

for rational drug design against leishmaniasis and trypanosomiasis (Misra et al. 2005).

RNase H are classified into three families: types I, II and III, depending upon their amino acid sequence similarity with the two *E. coli* RNase H enzymes and requirement of either  $Mg^{2+}$  or  $Mn^{2+}$  (Ohtani et al. 1999). Enzymes that require  $Mg^{2+}$  for its activity and share conserved amino acid residues/domains with that of *E. coli* RNase HI are classified as type I RNase H (Ohtani et al. 1999). Type II RNases H are those enzymes that share conserved domains with *E. coli* RNase HII and require  $Mn^{2+}$  for their activity (Ohtani et al. 1999). The third type of RNase H share conserved domains with *E. coli* RNase HII but requires  $Mg^{2+}$  for their activities (Ohtani et al. 1999). RNases H from different families have differences in their specific activities, divalent metal ion preferences and specificities for cleavage sites (Ohtani et al. 1999; Ohtani et al. 2000; Ohtani et al. 2004a; Ohtani et al. 2004b).

Most of the organisms have two RNase H genes. On the other hand *Caenorhabditis elegans* has four genes encoding RNase HI-related proteins and one gene for RNase HII (Arudchandran et al. 2002) and *Bacillus subtilis* has three RNases H genes (Ohtani et al. 1999). We report here four different RNase H in the parasitic protozoan *Leishmania*. One of them has a mitochondrial localization signal and is predominantly enriched in the mitochondria. All of them seem to be regulated during the differentiation of the parasite from promastigotes to amastigotes. Three of them have upstream open reading frames at their 5'-untranslated regions of their mRNAs and could be the potential cause for the regulations.

## MATERIALS AND METHODS

### *Leishmania* strains and growth conditions

*L. major* (Friedlin) promastigotes were cultured in M199 medium supplemented with 10% heat inactivated fetal bovine serum (HIFBS) at 26 °C with moderate shaking (Mishra et al. 2001a; Mishra et al. 2001b). The *in vitro* culture of the axenic amastigotes was done in M199 media with 25% HIFBS and 10 mM sodium succinate titrated to pH 5.5 at 37 °C (5% CO<sub>2</sub>) (Mishra et al. 2001a; Mishra et al. 2001b; Misra, Tripathi and Chaudhuri 2005).

### *Amplification of RNase H genes from L. major for characterization*

*L. major* BLAST searches were performed at either the Sanger ([http://www.sanger.ac.uk/cgi-bin/blast/submitblast/l\\_major/omni](http://www.sanger.ac.uk/cgi-bin/blast/submitblast/l_major/omni)) or National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST/>) web sites. Query RNase H protein sequences were RNase HI, -HII of *E. coli*, *Saccharomyces pombe*, *S. cerevisiae*, *Crithidia*, *Trypanosoma brucei* and human. Pfam protein families' database (<http://pfam.wustl.edu>) was used for the prediction of the protein domains. The subcellular organelle targeting was predicted using the PSORTII program (<http://psort.nibb.ac.jp>). MacVector Program was used for the sequence analysis and primer design. Initial primers were designed to amplify overlapping cDNA fragments from each gene following 5'- and 3'-RACE protocols (see below). Full-length cDNA for each gene was generated by recombination PCR (Sambrook and Russell 2001) using the 5'- and 3'-RACE products as templates and the miniexon-specific primer and AUAP as primers (see below). Following terminal primers were used to amplify the coding sequence of each gene: For RNase HI (5'-CACCATGTCGGCCTCTCTCTCTGC-3'/5'-CTACGATCTGTGCAGCCG-3'); RNase

HIIA (5'-CACCATGGCGGGTCCGGTGGG C-3'/5'-TCACATTACCGGCCCTC-3'); RNase HIIB (5'-GGATCCATATGACGGAT ACGCCATCG-3'/5'-AAGCTTCTACACCAT GTGCC GCTTCGG-3'); RNase HIIC (5'-GGATCCATGGTCCACGCATGGCGC-3'/5'-GGATCCATG GTCACTTCCTCTGCAGCC-3'). PCR amplifications were done using Pfu DNA polymerase to avoid inadvertent mutations; the amplified products were treated with AmpliTaq and ATP to add 'A' at the 3'-ends and the products were cloned into pCRII-Topo using Topo-TA cloning techniques (Invitrogen). The nucleotide sequences of the inserts were determined by reading both strands following 'primer-walking' protocol (Sambrook and Russell 2001).

### *Southern blot analysis to determine gene copy number*

Genomic DNA was isolated from *L. major* promastigotes using Qtip-100 columns (Qiagen). Genomic DNA (5 µg) was digested with seven different restriction endonucleases. The restriction endonucleases chosen do not have a site in the probe DNA. The digests were resolved on a 0.8% agarose gel, transferred to Nytran nylon membrane (Schleicher and Schuell, Keene, NH) in denaturing buffer following recommended protocols and probed with <sup>32</sup>P-labeled respective LRNase H ORF DNA as probe (Sambrook and Russell 2001).

### *Genetic complementation in RNase H-deficient temperature sensitive E. coli*

*E. coli* mutant MIC1066 [*rnhA-339::cat recB270(Ts)*] (Cazenave, Mizrahi, and Crouch 1998) was a generous gift from Dr. R. J. Crouch, NIH, USA.. These cells have a temperature sensitive mutation, because of which they can form colonies at 30°C but cannot support the growth at 42°C in the absence of functionally active RNase H. The ORFs of the different RNase H were amplified with high fidelity Pfu polymerase (Stratagene)

and the terminal primers from the *Leishmania major* genomic DNA. Primers were synthesized to amplify the coding regions such that the amplified products can be cloned directly in a directional manner in the pENTR/SD/D-TOPO vector from Invitrogen. The pENTR/SD/D-TOPO vector provides the optimal expression of the PCR product by the T7 RNA polymerase after recombination with the Gateway destination vector pEXP2-DEST (Invitrogen). The primers used for LRNHI (5'-CACCATGTCGGCCTCTCTCTCTGC-3'/5'-CTACGATCTGTGCAGCCG-3'), LRNHIA (5'-CACCATGGCGGGTCCGGTGGGC-3'/5'-TCACATTACCGGCCCTC-3'), LRNHIIB (5'-CACCATGACGGATACGCCATCG-3'/5'-AAGCTTCTACACCATGTGCCGCTTCGG-3'), LRNHIIC (5'-CACCATGCTCCACGCATGGCGC-3'/5'-GGATCCATGGTCACTTCTCTG CAGCC-3') and LRNHIICΔMLS (5'-CACCCCGCCGACATGGAACGCGCTG-3'/5'-GGATCCATGGTCACTTCTCTG CAGCC-3'). The pEXP2-DEST vectors harboring different *rnh*-like coding regions were transformed into the MIC1066 *E. coli* strain (Cazenave, Mizrahi, and Crouch 1998). Transformants were plated on LB-amp plates at 30 and 42°C. Growth at 42°C indicates that a functional RNase H is present in MIC1066. The pEXP2-DEST vector with/without LBRH2 coding region was used as a control in this study.

#### *RT-PCR analysis to evaluate mRNA levels*

RNA was isolated from *Leishmania* promastigotes and amastigotes using Trizol reagent (Invitrogen) following the manufacturer's protocol. DNase-treated RNAs were made to cDNAs using Superscript II reverse transcriptase and random hexamers as primers. Specific RNase H cDNA was amplified using gene specific primer pairs. The primers used for the RT-PCR were: for RNase HI, 5'-CACCATGTCGGCCTCTCTCTCTGC-3'/5'-TCTCCACGGTTGTTCTGTCTG-3'; for RNase HIIA, 5'-CACCATGGCGGGTCCGGTGGGC-

CGGTGGGC-3'/5'- ACCACCACAAAGGCTA CTGC-3'; for RNase HIIB, 5'-TTCCCGCACGCCTATGACTTTG-3'/5'-AAGCTTCTACACCATGTGCCGCTTCGG-3'; and for RNase HIIC, 5'-GCAATGCCGATGGACCAG-3'/5'-GCACCCGAGCAGTTCAG-3'. Beta-actin mRNA was amplified (5'-CAACGTGCCGTCGCTG-3'/5'-GCACTGTTCTGTCATCTC-3') to use as a experimental control. To ensure that the amplified DNA is not coming from any spurious genomic DNA contamination, we used mock cDNA preparations (no reverse transcriptase added) as controls.

#### *RACE analysis*

To amplify the 5'- and 3'-UTR sequences of *Leishmania* RNase H mRNAs we used the standard reagents and protocols provided by the supplier (Invitrogen). For 5'-RACE, we made first strand cDNAs from total RNA using random hexamers as primers. On the other hand we used 5'-anchored oligo (dT) as the primer for 3'-RACE. In both cases we used Superscript II (Invitrogen) as the reverse transcriptase. Since all *Leishmania* mRNAs have a 39 nt miniexon sequence at their 5'-ends (Chaudhuri 1997), we used the miniexon specific sense primer (Mex2: 5'-GGATCCAGTTTCTGACTTTATTG-3') for the synthesis of second strand cDNA synthesis. For the amplification of the 5'-UTR sequence we used Mex2 as sense primer and RNase H gene-specific antisense primer from the coding sequences. The gene specific antisense primers for 5'-RACE were the following: RNase HI: 5'-TCTCCACGGTTGTTCTGTCTG-3'; RNase HIIA: 5'-CGTTCGTTGATGACAGAAAC-3'; RNase HIIB: 5'-AAGTAATCCTCGCTCACGC-3'; RNase HIIC: 5'-GCAATGCCGATGGACCA G-3'. Often gene-specific nested antisense primer was used for a second PCR run to ensure specific DNA amplification. The nested primers used for 5'-RACE were: RNase HI: 5'-ATGATGAGATGCCGCGAGTG-3'; RNase

HIIA: 5'-GATGTTGCGTTCGTTGATG-3'; RNase HIIB: 5'-AGACGCACACAGTGAA GGAG-3'; RNase HIIC: 5'-GATGTTGCTCT GAGGCAGCA-3'. For 3'-RACE analysis AUAP (5'-GGCCACGCGTCTGACTAGTAC-3') was used as the universal primer. The gene-specific primary sense primers and secondary nested primers used for 3'-RACE were the following: RNase HI: 5'-ACGCAAAGAAAT GGGTGC-3'/5'-ACAAGAGTGAGGGCGTT GAG-3'; RNase HIIA: 5'-TGCTTTGGATGG GTTCAGC-3'/5'-ACCACCACAAAGGCTAC TGC-3'; RNase HIIB: 5'-GCATCGCTTTTTT AGTTTCC-3'/5'-TTCCCGCACGCCTATGA CTTTG-3'; RNase HIIC: 5'-GAGGCGACAT CGTTGACG-3'/5'-ACTACTTCACCAGCGT TCG-3'. Amplified PCR products were cloned into the pCRII-Topo vector (Invitrogen) and the inserts of several clones were sequenced.

#### *Pulsed Field Gel Electrophoresis*

Late-log phase *Leishmania* promastigotes were harvested and suspended in PSG buffer (75 mM sodium phosphate (pH 8.0), 65 mM sodium chloride, 1% glucose) at a concentration of  $1.7 \times 10^8$  cell/ml. This suspension was incubated at 37 °C for 10 min and then diluted with one volume of 2.0% low melting agarose cooled to 37 °C. Aliquots (60 µl) of this suspension was then quickly transferred to a block mold to cool at 4 °C for 10 min. Solidified blocks of agarose were transferred to lysis mixture containing 0.5 M sodium N-laurylsarcosinate and proteinase K (2 mg /ml) and were then incubated at 45 °C for 48 h. The proteinase K-digested blocks were washed with several changes of 1x TBE and stored at 4 °C in 1x TBE. The agarose blocks were placed in the slots of the 1% agarose gel and the chromosomal DNA was electrophoresed through the gel by CHEF-DRII Pulsed Field Electrophoresis system from BioRad at the switch time of 60-90 sec for 17 h followed by 90-120 sec switch time for 7 h at 6v/cm and at 6°C. The DNA in the gel was partially depurinated prior to transfer

on the membrane by soaking the gel twice in 0.1 N HCl for 20 min each time with gentle shaking at room temperature. The transfer of the DNA was done in 0.4 N NaOH and 1.5 M NaCl solution for 24 h with upward capillary rise (Schleicher and Schuell). The membrane was neutralized with 0.5 M Tris.Cl, pH 7.0 for 5 min followed by with 2x SSC wash. Hybridization and autoradiography was done as described above.

## RESULTS AND DISCUSSION

### *Leishmania has four distinct RNase H genes*

RNase H (EC 3.1.26.4) is a ubiquitous enzyme. Most of the organisms so far studied have two, one type I and the other type II, RNase H genes. Recently, five (four of type I and one of type II) RNase H genes have been identified and characterized from *C. elegans* (Arudchandran et al. 2002). Since RNase H is a component of the DNA replication/recombination machinery (Desjeux 2004; Crouch 1998), we became interested to explore whether such enzyme in the parasitic protozoan *Leishmania* could be used as a chemotherapeutic target against the parasite infection. Database search followed by PCR-based amplification methods led us to identify four different RNase H genes in *Leishmania* (Table 1, Figure 1). One of the enzymes, LRNase HI, is a type I ribonuclease with strong similarity with other type I RNases H. It has two canonical domains characteristic of such enzymes. One is the RNA/RNA or RNA/DNA duplex binding domain (residues 5-48, Figure 1A). The other domain (residues 163-377, Figure 1A) is responsible for the catalytic activity of the protein. Three other *Leishmania* RNases H share conserved domains with the known type II enzymes (Figure 1). We are not sure about the metal ion requirements of each of these enzymes. Thus, the classification is simply based on sequence similarities. We named *Leishmania* type II RNases H as LRNase HIIA, -HIIB and -HIIC

based on the size of the protein (Table 1). LRNase HIIA and –HIIB are of very similar sizes but other than sharing conserved signature residues, they do not share any significant homology between themselves or with LRNase HIIC (Figure 1). Among the four RNases H, LRNase HIIC is the largest (~53 kDa), possesses a mitochondrial localization signal (Fig. 1) and is concentrated in the mitochondria of the parasite (unpublished data, S. Misra and G. Chaudhuri). LRNase HI, -HIIB and HIIC have their equivalents in related parasitic protozoan *T. brucei* and *T. cruzi* (Table 1). LRNase HI also has significant similarity with *Crithidia* RNase HI (Ray and Hines 1995; Engel and Hines 2001). The structural features of *T. brucei* RNase HI (Hesslein and Campbell 1997) are also conserved in LRNase HI (Figure 1). There are apparent distinctions between the RNase H

profile of non-parasitic and parasitic trypanosomes. In *Crithidia*, a trypanosome non-pathogenic to human or mammals, only one type I RNase H is shown to meet the DNA replication needs both in the nucleus and in the mitochondria (Ray and Hines 1995; Engel and Hines 2001). On the other hand, in the three pathogenic trypanosomes, *Leishmania*, *T. brucei* and *T. cruzi* multiple RNases H may be doing these jobs. The presence of LRNase HIIA exclusively in *Leishmania* and not in other trypanosomes is noteworthy. This protein is strongly homologous to bacterial RNase HII. The biological roles played by this different LRNases H in *Leishmania* and related protozoan parasites are yet to be determined. All four cloned *Leishmania* RNase H proteins could complement RNase H-deficient *E. coli*.

**Table 1:** Characteristics of *L. major* RNase H genes.

Type	Accession Number	ORF size (bp)	Size of the <i>Leishmania</i> protein (kDa)	Size of the equivalent <i>T. brucei</i> protein (kDa)*	Size of the equivalent <i>T. cruzi</i> protein (kDa)*	Length of 5'-UTR (bp) <sup>a</sup>	Length of 3'-UTR (bp) <sup>b</sup>
LRNase HI	AY835988	1146	42 (381 aa)	33 (301 aa)	36 (323 aa)	878	160
LRNase HIIA	AY835989	891	33 (296 aa)	Absent	Absent	120	82
LRNase HIIB	AY835990	906	33 (301 aa)	36 (326 aa)	36 (324 aa)	417	825
LRNase HIIC	AF441859	1425	53 (474 aa)	54 (477 aa)	54 (479 aa)	519	515

\*Obtained from GeneDB of the respective parasite.

<sup>a</sup>Includes 39 nt miniexon sequence. Determined by 5'-RACE analysis. See Fig. 5 for nucleotide sequences.

<sup>b</sup>Does not include poly A tail. Determined by 3'-RACE analysis. Nucleotide sequence included in the Entrez data base (see the accession numbers)

**(A) LRNase HI**

5 Double-stranded RNA/RNA-DNA hybrid binding domain 48

MKPSFYAVAVGKARGIYSTWAQCSEQVTGYPGCIYKGFRTLDEARSFLAAHPLPSTGSAVSLGGSSSS  
VAAPYSSSEHAALGSNYFKGRLKRETRELLAGEESVAERQARRRRADHAQIAPPELSSMSASLSA  
ATPS

**163 RNase HI domain**

TTASKPLATTTTTSSPGAFEVYVDGACSHNGTSQARAGYGGYYGSMSPRNFSCPVPLTESQTNNRG  
EVRAVIHAIVQAFVDAGAPADALEATHLVDPRAWRLVDLTRPLPHLIYTDSTRYVIDGLTRYAKKWVQNGF  
VLSTNEPVQNQDLWRQLIRLRDRYNTLYARQQHDQQRACQWGDGHHSEADLAQLVRYTCQNARNDKS  
EGVELIH

**377**

VRGHSKVHGNEMADSLAVMGSRLHRS

**(B) LRNase HIIA**

17 RNase HII domain

MDTTERDAGLRPSIRYVIGVDEAGRGC MAGPVGVAATLYRISATPVRYAVMDSKKMKESDRAETLEA  
MCAELSDAGVEGFVGTTRVDGACCLVRYQRRKPLQAVAAALVDVSVINERNILNATLEGMSAVCEAVV  
RSYNAVSGLGTLTPLTCAILIDGNRVPWTFLSGEERQRVCVQAVRKVEARKRIGVEYPALDGFSCCTTVVK  
GDATLLS

**281**

IAAASIVAKVARDAYAVSVMHPAFPQYGFDDHHKGYCTAAHKAELARWGPSRFHRLDYAPVKGTGGQR  
GDA

ERTPASLEGPVM

**(C) LRNase HII B**

45 RNase II domain

MTDTPSALSARLSTRLSGTKRPRESLIEFRVKDIMESLGVSEDYLVLGIDEAGRGPVIGPMVYTGAVISLG  
EHDDL\_VRLCHVADSKMLNERRRLASLQQLRQLKTFRSFTVCVSPPEISKMTMTGRSGRNSNTLSHETAIQI  
ISEATLASAGKLCAAAYVDTVGPPEYQARLAGRFPHLRVTVAKKADSIFPIVSAASIVAKTVRDTAIEALG  
E

**252**

NIGSGYPSDPRTMEWLRSHVHRFFSFPAYDFARHSWGPLVQLANDPAVRVPVCEQDLEGARQQGR  
GGDSRQKLSFAKPSKRHMV

**(D) LRNase HII C**

1 Mitochondrial localization signal 40

MLHAWRFVPLARSTSALGSGLLRPSRVLKRLAYSPPFRPPTWNALQKVEVGHQRQDTGEPTLDGTL  
D

**134**

LVGYQLSRDGLLPQTNILGRTPVPDTSGISAAATRRKLERERDVAVAHMKRHGVRLRRSIKETAVTIGC

RNase II domain I

DEAGRGLAGPVVGAAVSRIPVSSFNNEFDQLYEAPEQFQIFDSKSVSERQRDLVFAMITGHVDFDIA

**208**

SCKKFVWHHCAGDEATSLTLSSKLD SHVKLSKLPFKKLLSMQTPYLITYHGYNSAGNYVYFWSIGIAN  
HTYIDEYNIYNASMNTMHRSAQSIWHMLSDARFSHEAAPRPRSSSIAQYLFSRFCTAANHDNEKRYQVP

**363 RNase HII domain II**

HHLELLKGATEYFDFEPIQPPLVLIDGHAVPGPSYDYFTSVRIGGDVQPIIEGDKRSLSIAAASCLAKV

**474**

TRDEL MN YIDALYPGYGFRDNKGYPVEQHMKYVAKNGLCPIHRKTYRPCR TVLEKGLQRK

**Figure 1: Conserved domains in *Leishmania* RNases H.** The amino acid sequences of LRNase HI, -HIIA, -HII B and HII C are shown. The sequences that are in the conserved domains (Marchler-Bauer and Bryant 2004) are in grey. The grey areas are also numbered. Amino acid residues identical or similar to the conserved domain amino acid sequences are in bold and underlined. (A) Two distinct domains of LRNase HI are shown: one is the double-stranded RNA- or RNA:DNA-duplex binding domain (residues 5-48) and the other is the RNase HI domain (residues 163-377). (B) The large RNase

HII homology domain (residues 17-281) in LRNase HIIA is shown. The amino acid residues said to be critical for the RNase H activity are conserved within this domain. (C) The RNase II homology domain in LRNase HIIB is shown (residues 45-252). (D) The putative mitochondrial localization signal in LRNase HIIC is shown (residues 1-40). The conserved RNase HII domain is split into two distinct domains: domain I is from residue 134-208 and the domain II is from residue 363-474. We compared the canonical sequences with the Cluster of Orthologous Groups (COG) of RNase H-like proteins COG0328 and COG3341 for LRNase HI and COG0164 for the other LRNases H (Tatusov et al. 2000).

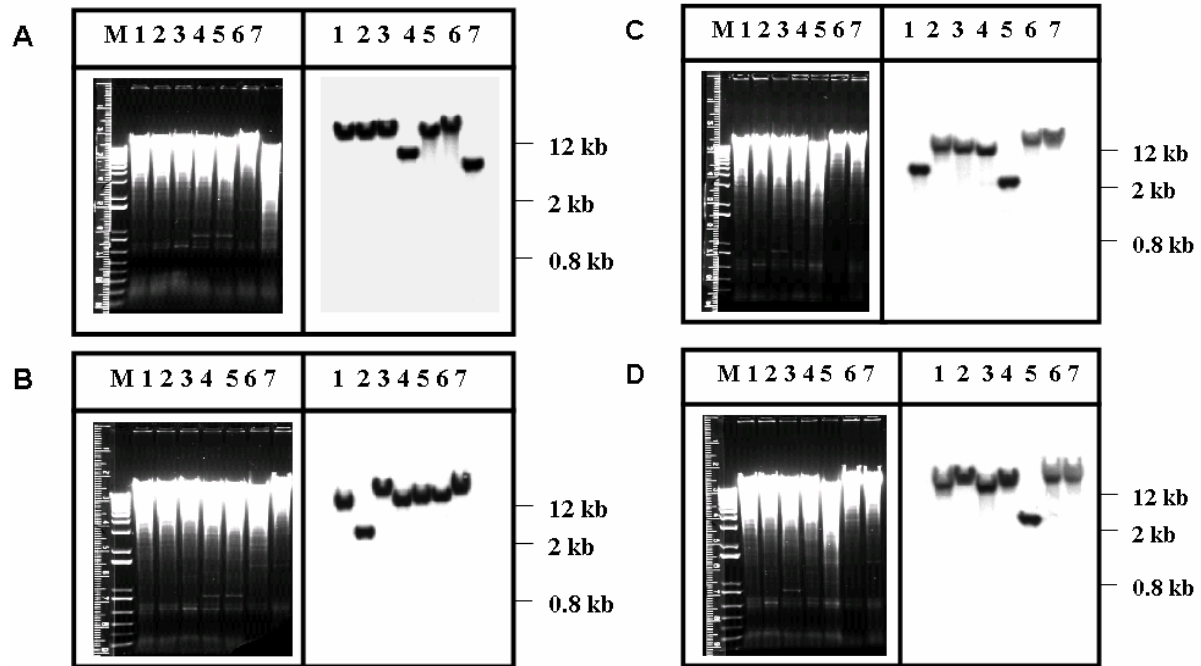
*Each of the genes has a single copy in the genome*

Determination of the copy number and chromosomal localization (physical context of the gene) of a gene in a cell is important to understand its regulation and also to design experiments to knock out the gene by homologous recombination. We digested genomic DNA isolated from *L. major* promastigotes with seven different restriction endonucleases and analyzed the digestion products by Southern blotting using specific RNase H ORF DNA as probe. There were only one restriction fragment in each digest hybridized with the probe (Figure 2), suggesting that all the four RNases H in *Leishmania* exist as single copy genes. We also have verified the chromosomal localizations of these genes by pulsed field agarose gel electrophoresis. LRNase HI, -HIIA, -HIIB and HIIC are located in chromosomes 6, 13, 36 and 36, respectively (Figure 3). These observations conform to the *L. major* GeneDB data. It is interesting to note that although LRNase HIIB and -HIIC are located in the same chromosome they are transcribed from opposite strands.

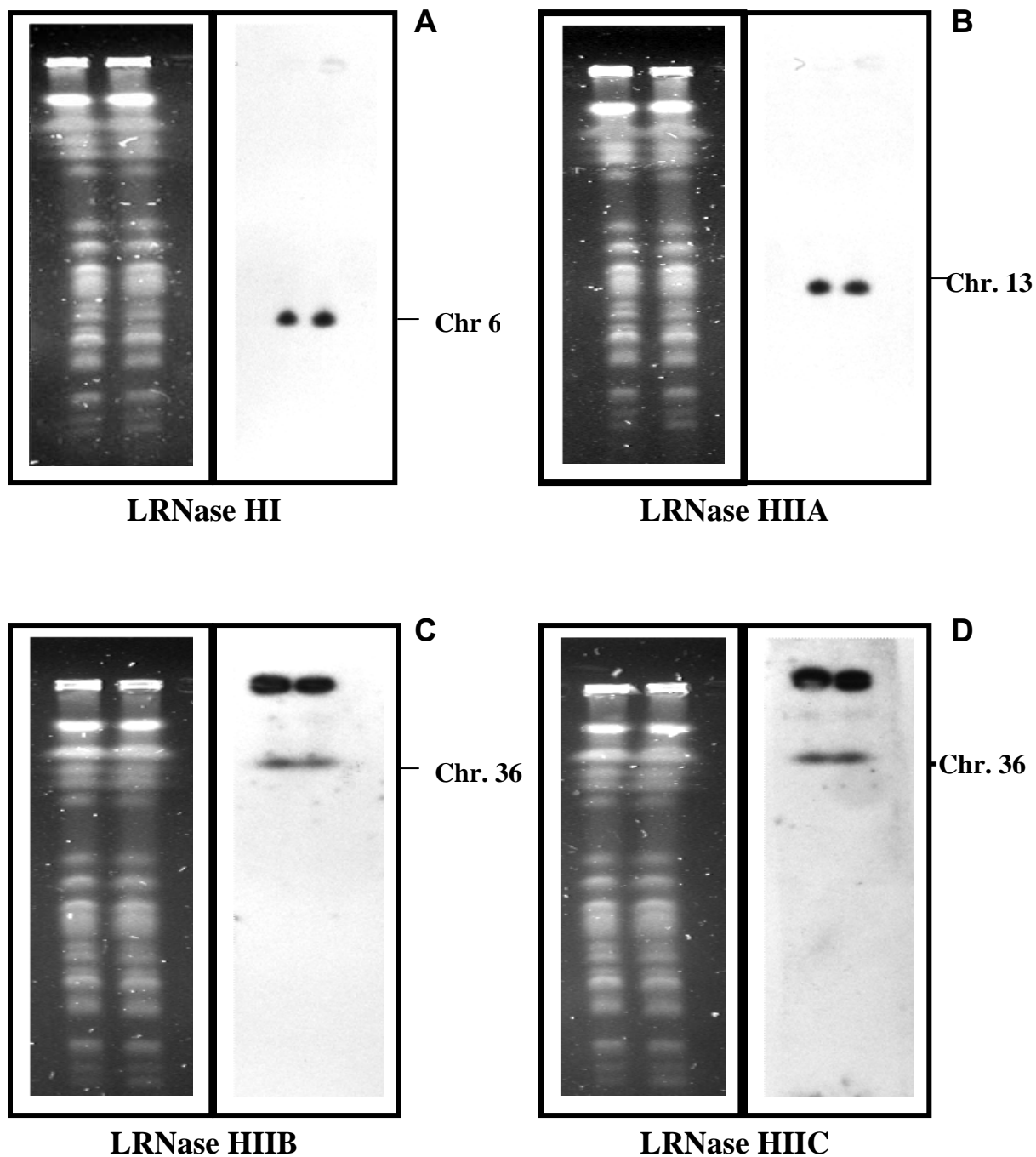
*Leishmania RNase H genes are expressed differentially in promastigotes and amastigotes*  
*Leishmania* go through radical changes in environmental cues during its natural life cycle (Alexander et al. 1999; Desjeux, 1994). One of

these variations is the change in growth temperature of the parasite from ambient (25 °C) to 35-37 °C. It thus may need different proteins suitable in different temperature environment to support important metabolic functions. There are many example of stage-dependent expression of genes in this parasite (Boucher et al. 2002; Mishra et al. 2003; Larreta et al. 2004). The surface protease gp63 of *Leishmania* is an example where different isotypes are expressed at different stages of growth of the parasite (Brittingham et al. 2001). To understand whether different RNase H genes in *Leishmania* are expressed in a stage-specific manner, we have studied the expressions of all four LRNases H in the promastigotes and the amastigotes of the parasite by RT-PCR analysis (Figure 4). LRNase HIIA transcripts are 5-6 folds more abundant in the amastigote stage of the parasite than that in the promastigotes (Figure 4B). All the other three RNase H are expressed in both of the stages but there are differences in the degrees of their expressions (Figure 4). The expressions of LRNase HI and LRNase HIIC are 2-3 folds higher in the amastigotes than in the promastigotes (Figure 4A and 4D). On the other hand LRNase HIIB is expressed 2-3 folds more in the promastigotes than in the amastigotes (Figure 4C). The mechanisms and the consequences of these differential expressions are under study.

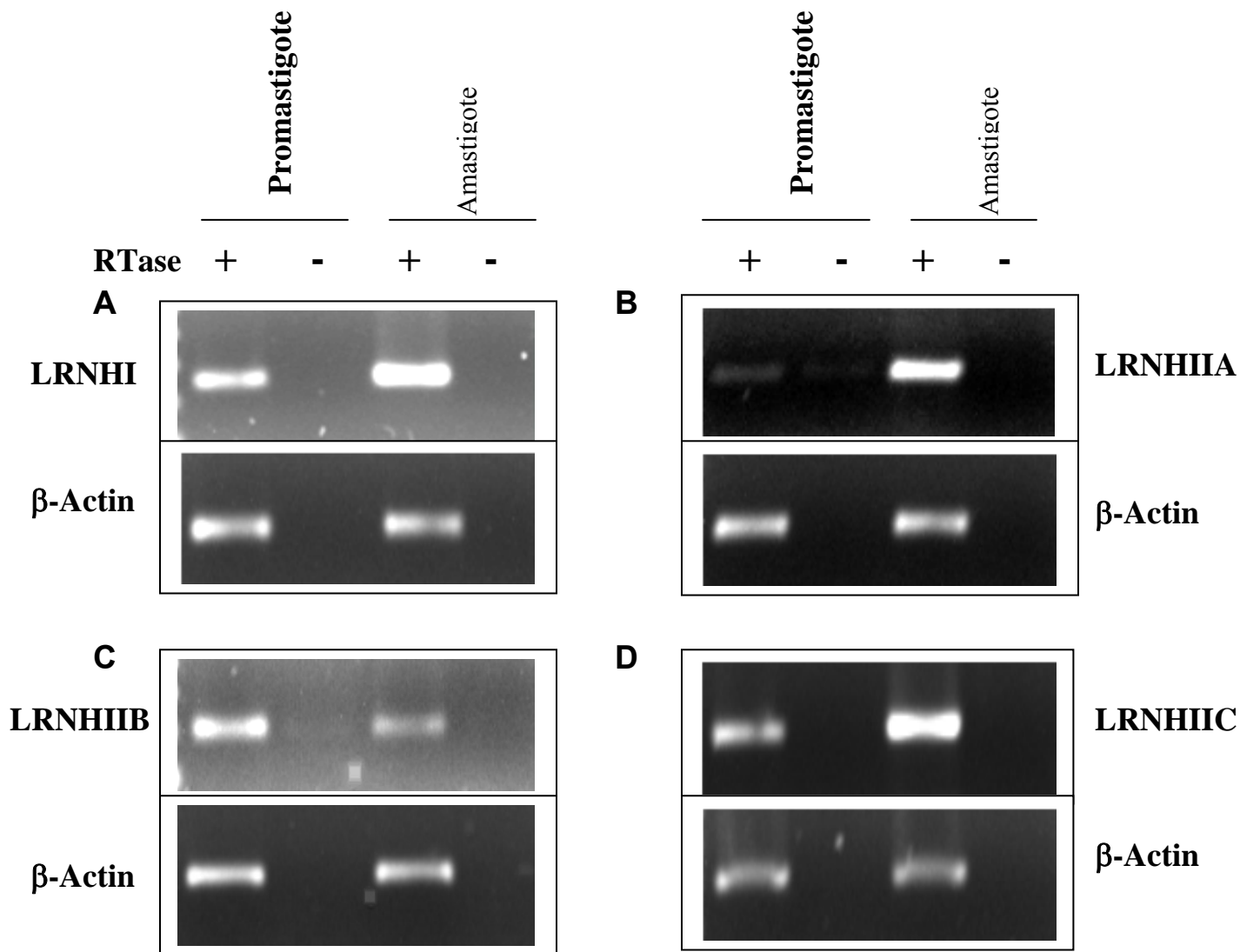




**Figure 2: Southern analysis of *L. major* genomic DNA showing single copy of each of the RNase H genes.** Left panel in each of (A) to (D): Ethidium bromide-stained agarose gel photograph. Lane M; 1 kb plus DNA ladder (Invitrogen); Right panel: corresponding autoradiograms. (A) LRNase HI: lanes 1-7, DNA was digested with Bam HI, Bgl II, Eco RI, Eco RV, Hind III, Xba I, and Xho I, respectively. (B) LRNase HIIA: lanes 1-7, DNA was digested with Bam HI, Bgl II, Eco RI, Eco RV, Hind III, Kpn I, and Xba I, respectively. (C) LRNase HIIB: lanes 1-7, DNA was digested with Bam HI, Eco RI, Hind III, Nde I, Nsi I, Spe I and Xba I, respectively. (D) LRNase HIIC: lanes 1-7, DNA was digested with Bam HI, Eco RI, Hind III, Nde I, Nsi I, Spe I and Xba I, respectively.



**Figure 3: Chromosomal localization of RNase H genes in *L. major* by pulsed field gel electrophoresis followed by Southern hybridization.** Left panel in each of (A) to (D): Ethidium bromide-stained pulsed field gel photograph; Right panel: corresponding autoradiograms. Probe used were: (A) LRNase HI; (B) LRNase HIIA; (C) LRNase HIIIB; and, (D) LRNase HIIC. Identifications of *L. major* chromosomes are from published pulsed field electrophoresis data (Bastien et al. 1992). The bright band below the loading well in the EtBr-stained gel may be a non-DNA artifact, as it does not hybridize with *Leishmania* total DNA probe (data not shown). Identical samples are loaded in two lanes.



**Figure 4: RT-PCR evaluation of the relative mRNA levels of different RNase H in *L. major* amastigotes and promastigotes.** (A) LRNase HI; (B) LRNase HIIA; (C) LRNase HIIB; and (D) LRNase HIIC. Beta actin was used as a loading control. RT, reverse transcriptase; (+): RNAs were treated with Superscript II RT to make the cDNA before PCR; (-): RNAs were not treated with Superscript II RT and thus cDNAs were not made before PCR. The sizes of the amplified products were 239 bp, 276 bp, 191 bp, 307 bp and 312 bp for RNase HI, RNase HIIA, RNase HIIB, RNase HIIC and  $\beta$ -actin, respectively.

**(A) LRNase HI**  
 -878 Mini-exon -840  
 AACUAACGCUAUAUAAGUAUCAGUUUCUGUACUUUAUUGUGCAAGCCGAUCACAGGAAAUACAAGUGCGCGAAG  
 AGGGAGGGCGGUGCAAAGCAAGAUCCAGUCGCUUCACUGAGAACAGGCCUAACGCAACACCCGUGCUCGCCUC  
 CUCUCCUUCGCGUUUACGUCGCAUCACCUCUCGCCGCGACCACUUCUGUGGCUGCUUCUGUAUAAACAGCG  
 CGAUACAGGCUGACAUACACACAUACACACCGACACACAUACGUGUACGAGUGCGGAGCGGAGCGCGCCGUUUU  
 CUUUGCUCUUGCUUAUCGCGCUACUCUGCCGUUUAACGUUAGCUUGCCACUGUCUGGUGCUAUCGCCACCG  
 UACCACACAACGCCUCAGAACCCGUCUCCUUCGCCUCUCUCCGGCAC  
 -420 uORF1  
 CGCCGUGUCUUCUUUGGAAAAGUGCCGUACGCUGUGCAUGGGUCUUUUCGUCUGUUGUUGGCGUGCCACCUC  
 CCCUGUUAUCGCGUCUCCUCAGUGGGGUCUCCGGUCACGAAAGACUCCGUCACGUGCCGCGGUCGCCGUCGUC  
 CGCGUCGUAUCGUGCGCUUG  
 -232 -213  
 CCUCGUCACCCGGUCUCUCCUCCUACCCAGCGCGGUCUCCUCGGCAGGGUAGAGUAGAGGAGCCGGGAGGG  
 GAUAUGGAGG  
 uORF2 -175  
 GCGUGCAGGCCGAUGUGAAGCCGGCAUCUGAGCGUGAGGGCGGGGGGGGAGACUGGCUUCGAUCCAAUCGUC  
 GCGUCAGGCAA  
 GCUGGAAGGCCUGCAGAGCUUCUUCUUGUCCGGUCUCUUGGCCUCUCUCAGCGGAACGCAUCGCCACCGUCACC  
 GCCGCCGCG  
 +1  
 CUCCUGCUAAGACCACGGGCGUAUCCUCUGCGUCUGCAUG

**(B) LRNase HIIA**  
 -120 Mini-exon -82  
 AACUAACGCUAUAUAAGUAUCAGUUUCUGUACUUUAUUGAAAAAGAAGACUACCACGACCAGCGGACACGUGCGU  
 GCACGCCU  
 +1  
 GUGCCUUAUCUCUGAGUGACAUG

**(C) LRNase HIIIB**  
 -417 Mini-exon -379  
 AACUAACGCUAUAUAAGUAUCAGUUUCUGUACUUUAUUGACACCGCCUGUACCCAAGUGUCUCUACCAUCGCCGCA  
 CCUACAAAAGGCAGGAAGGAUACGUCACGCCGUGAAGGCGACAGGGGAGGGAUCAGCCAGCAGACUUUCUCUCUC  
 GUACACGGCUGAUUGC  
 -187 uORF  
 UGACCCUCCGUAGCAGCAGCAGCCGCGCCACGUCCUCAACCUGCUCGAGGGUGCGGCAGUUAUGCAGCAGAAGG  
 GGAAGCUCC  
 -158  
 CUGUUUAGAGGCGGCGGACGAGUGUGCAAGCGGUGCCUGCACGCAGAGCGGAGCCCCUCUUCUUGCCUGCUUG  
 UGUGCGUGU  
 +1  
 CGAGGCCUUUCUUUCGGAGAAGCAAAGAACUGCGGUGGCGGAUACAGACGCGCAUCACCUCCACAUUAGCGAGGCA  
 GGGCAUG

**(D) LRNase HIIIC**  
 -519 Mini-exon -481  
 AACUAACGCUAUAUAAGUAUCAGUUUCUGUACUUUAUUGCGACACACCAACCACUCGAGGUCUAACAGGAGCACUA  
 CCCUCCCUCCUCUUGAGCGGUGGCGUUUUAACAGUCCCUUGUCGUCGCACUCACACUUCACAUACACAUACGCCU  
 GGUCACGCGCAGGCC  
 -290 uORF -279  
 AGAGAUAGAGAUAGACCUUUUAGCGCACGAGAGCUCUUUGACACACCCUGCCCAUACACUCAUGCGCUCAUAGACA  
 UCACUCGGCCCGAGUACCUCUCUUGUGUUUUUUGUCUUUUUCCAUCUGCACCGUUCGUCUGCUAAGGCGCGUCU  
 UCACACUGUUGUGGUUUUCUUCUUUUGGUUUUCGCGUGCCCAAAAAGGAGUUUCAUCGCCGUGUAGCGCUGGCUU  
 GGACGAGCUUCAGGGCGCCGACGCGCCUCCGUGAAGCUGUGACUUAACGAUUCACACACCACACAUUACUCUC  
 ACACGCACUAGCGUGCGGCCGACGUCACUCAA  
 +1  
 AAGUGACUGCCUAUCAUG

**Figure 5: Nucleotide sequences of the 5'-UTRs from LRNase HI, LRNase HIIIB and LRNase HIIIC of *L. major* showing the uORFs.** Nucleotide sequence of the 5'-UTR of RNase HIIA is also shown but there is no uORF in this sequence. The 39 nt minixon sequence at the 5'-ends, the uORFs and the AUG codon of the main ORFs are in grey shades. The start and the stop codons are underlined. The numbers indicate the distance of the element from the AUG codon of the RNase H protein coding ORF considering the 'A' in this AUG as +1.

### *The 5'-UTRs of three out of four LRNase H mRNAs have uORFs*

Transcriptional regulations of the protein-coding genes in *Leishmania* and other related protozoan are apparently absent. Thus, majority of the regulations of gene expression in *Leishmania* appears to be at the levels of trans-splicing, mRNA translation and stability (Quijada et al. 2000; Brittingham et al. 2001; Boucher et al. 2002; Mishra et al. 2003; Larreta et al. 2004). Translatability and stability of an mRNA are largely determined by elements at the 5'- and 3'-UTR of the mRNA (Quijada et al. 2000; Brittingham et al. 2001; Boucher et al. 2002; Mishra et al. 2003; Larreta et al. 2004). In order to understand possible regulation of the LRNase H mRNAs, we determined the nucleotide sequences of their 5'- and 3'-UTRs (Table 1). The nucleotide sequences of the 5'-UTRs are shown in Figure 5. Complete nucleotide sequences of the full-length cDNAs of these proteins are submitted to the GenBank (see Table 1 for accession numbers). While the lengths of the UTRs of LRNase HIIA mRNA are relatively short, the mRNAs of other three LRNases H have relatively large UTRs (Table 1). One of the interesting features in the 5'-UTRs of LRNase HI, -HIIB and HIIC is that they contain one or more upstream open reading frames (uORFs) in them (Figure 5). Particularly interesting is the 5'-UTR sequence of LRNase HI mRNA. It has two relatively large uORFs (Fig. 5). One of the uORFs (149 nt, -420 to -232) is unusually large and codes for a 62 amino acid peptide (lp62). The other uORF at the 5'-UTR of LRNase HI (39 nt, -213 to -175) codes for a 12 amino acid peptide (Fig. 5A). Whether lp62 behave similarly as AAP in *Neurospora* (Fang et al. 2000) to regulate the translation of the mRNA, needs to be evaluated. LRNase HIIB has one uORF at -187 to -158 that may potentially code for a 9 amino acid peptide (Figure 5C). Lastly, LRNase HIIC has one uORF from -290 to -279 and may potentially code for a tripeptide (Figure 5D). The role of uORF in the

regulation of eukaryotic mRNA translation is well studied in many systems including the fungus (Vilela and McCarthy 2003). Several variables define the function of a uORF in such regulations (Vilela and McCarthy 2003). They include, proximity of the AUG of the uORF to the 5'-cap and the context of the AUG, length of the uORF, environment of the stop codon of the uORF, length and structure of the downstream intercistronic sequence and the nature of the encoded peptide (Vilela and McCarthy 2003). We are studying the role of individual uORFs in the regulation of its cognate mRNA translation by mutation analysis.

### ACKNOWLEDGEMENTS

We thank Dr. Manish K. Tripathi for reading the manuscript. This research was supported by National Institute of Health, USA grants 5-R01AI42327-03 and 3 SO6 GM008037-33S1 to GC.

### REFERENCES

- Alexander J, Satoskar AR, Russell DG. *Leishmania* species: models of intracellular parasitism. *J Cell Sci* 1999;112:2993-3002.
- Arudchandran A, Cerritelli SM, Bowen NJ, Chen X, Krause MW, Crouch RJ. Multiple ribonuclease H-encoding genes in the *Caenorhabditis elegans* genome contrasts with the two typical ribonuclease H-encoding genes in the human genome. *Mol Biol Evol* 2002;19:1910-19.
- Bastien P, Blaineau C, Pages M. *Leishmania*: Sex, lies and karyotype. *Parasitol Today* 1992;8:174-7.
- Boucher N, Wu Y, Dumas C, Dube M, Sereno D, Breton M, Papadopoulou B. A common mechanism of stage-regulated gene expression in *Leishmania* mediated by a conserved 3'-untranslated region element. *J Biol Chem* 2002;277:19511-20.

- Brittingham A, Miller MA, Donelson JE, Wilson ME. () Regulation of GP63 mRNA stability in promastigotes of virulent and attenuated *Leishmania chagasi*. *Mol Biochem Parasitol* 2001;112:51-9.
- Chaudhuri G. Scavenger receptor-mediated delivery of anti-miniexon antisense phosphorothioate oligonucleotide to *Leishmania*-infected macrophages: Selective and efficient elimination of the parasite. *Biochemical Pharmacol* 1997;53:385-91.
- Crouch RJ, Toulme JJ. Ribonuclease H. INSERM, Paris, 1998.
- Desjeux, P. Leishmaniasis. *Nat. Rev. Microbiol.* 2004;2:692-3.
- Engel ML, Hines JC, Ray DS. The *Crithidia fasciculata* RNH1 gene encodes both nuclear and mitochondrial isoforms of RNase H. *Nucleic Acids Res* 2001; 29:725-31.
- Engel M.L, Ray DS. The kinetoplast structure-specific endonuclease I is related to the 5' exo/endonuclease domain of bacterial DNA polymerase I and co-localizes with the kinetoplast topoisomerase II and DNA polymerase beta during replication. *Proc Natl Acad Sci USA* 1999;96:8455-60.
- Fang P, Wang Z, Sachs MS. Evolutionarily conserved features of the arginine attenuator peptide provide the necessary requirements for its function in translational regulation. *J Biol Chem* 2000;275:26710-9.
- Hesslein DG, Campbell AG. Molecular cloning and expression of a ribonuclease H from the kinetoplastid, *Trypanosoma brucei*. *Mol Biochem Parasitol* 1997;86:121-6.
- Larreta R, Soto M, Quijada L, Folgueira C, Abanades DR, Alonso C, Requena JM. The expression of HSP83 genes in *Leishmania infantum* is affected by temperature and by stage-differentiation and is regulated at the levels of mRNA stability and translation. *BMC Mol Biol* 2004;5:3.
- Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* 2004;32:327-31.
- Mishra KK, Holzer TR, Moore LL, LeBowitz JH. A negative regulatory element controls mRNA abundance of the *Leishmania mexicana* Paraflagellar rod gene PFR2. *Eukaryot Cell* 2003;2:1009-17.
- Mishra M, Bennett JR, Chaudhuri G. Increased efficacy of antileishmanial antisense phosphorothioate oligonucleotides in *Leishmania amazonensis* overexpressing ribonuclease H. *Biochem Pharmacol* 2001a;61:465-74.
- Mishra M, Porter-Kelley J, Singh PK, Bennett JR, Chaudhuri G. Enhanced activity of antisense phosphorothioate oligos against *Leishmania* amastigotes: augmented uptake of oligo, ribonuclease H activation, and efficient target intervention under altered growth conditions. *Biochem Pharmacol* 2001b,62:569-80.
- Misra S, Bennett J, Friew YN, Abdulghani J, Irvin-Wilson CV, Tripathi MK, Williams S, Chaudhuri M, Chaudhuri G. A type II ribonuclease H from *Leishmania* mitochondria: An enzyme essential for the growth of the parasite. *Mol Biochem Parasitol* 2005 Jun 21; [Epub ahead of print].
- Misra S, Tripathi MK, Chaudhuri G. *Leishmania* infection of macrophages: Down regulation of 7SL RNA expression and impairment of vesicular protein transport pathways. *J Biol Chem.* 2005 Jun 14; [Epub ahead of print]
- Ohtani N, Yanagawa H, Tomita M, Itaya M. Cleavage of double-stranded RNA by RNase HI from a thermoacidophilic archaeon, *Sulfolobus tokodaii* 7. *Nucleic Acids Res* 2004a;32:5809-19.
- Ohtani N, Yanagawa H, Tomita M, Itaya M. Identification of the first archaeal Type 1 RNase H gene from *Halobacterium* sp. NRC-1: archaeal RNase HI can cleave an RNA-DNA junction. *Biochem J* 2004b;381:795-802.
- Ohtani N, Haruki M, Morikawa M, Crouch RJ, Itaya M, Kanaya S. Identification of the genes

encoding Mn<sup>2+</sup>-dependent RNase HII and Mg<sup>2+</sup>-dependent RNase HIII from *Bacillus subtilis*: classification of RNases H into three families. *Biochemistry* 1999;38:605-18.

Ohtani N, Haruki M, Muroya A, Morikawa M, Kanaya S. Characterization of ribonuclease HII from *Escherichia coli* overproduced in a soluble form. *J Biochem* 2000;127:895-9.

Pileur F, Tolome JJ, Cazenave C. Eukaryotic ribonucleases HI and HII generate characteristic hydrolytic patterns on DNA-RNA hybrids: further evidence that mitochondrial RNase H is an RNase HII. *Nucleic Acids Res* 2000;28:3674-83.

Quijada L, Soto M, Alonso C, Requena JM. Identification of a putative regulatory element in the 3'-untranslated region that controls expression of HSP70 in *Leishmania infantum*. *Mol Biochem Parasitol* 2000;110:79-91.

Ray DS, Hines JC. Disruption of the *Crithidia fasciculata* RNH1 gene results in the loss of two active forms of ribonuclease H. *Nucleic Acids Res* 1995;23:2526-30.

Sambrook J, Russell DW, Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001.

Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 2000;28:33-36.

Vilela C, McCarthy EG. Regulation of fungal gene expression via short open reading frames in the mRNA 5'-untranslated region. *Mol Microbiol* 2003;49:859-67.

Wu H, Lima WF, Crooke ST. Properties of cloned and expressed human RNase H1. *J Biol Chem* 1999;274:28270-8.