Stage-Specific Regulation of Gene Expression in *Leishmania*

Posttranscriptional gene regulation mediates morphological, biochemical changes when *Leishmania* move from sandflies into humans

**Barbara Papadopoulou, Xiao-Fang Huang, Nathalie Boucher, and François McNicoll**

*Leishmania* is an early-branching unicellular eukaryote belonging to the order *Kinetoplastidae* and family *Trypanosomatidae*. At least 20 *Leishmania* spp. cause leishmaniasis, vector-borne parasitic diseases with a large spectrum of clinical symptoms that depend on the species. *Leishmania* infects mononuclear phagocytes, key cellular components of the immune system. Intracellular survival of these parasites within human macrophages helps them to subvert the effector and regulatory functions of these cells.

Clinical manifestations of leishmaniasis are multifaceted, ranging from spontaneously resolving skin lesions to life-threatening visceral diseases, and depend on complex interactions between the host immune response and the infecting *Leishmania* species. Visceral leishmaniasis by *L. donovani*, *L. infantum*, or *L. chagasi* causes elevated morbidity and mortality. The cutaneous form is transmitted mainly by *L. major*, *L. tropica*, and *L. mexicana* and causes chronic skin ulcers. Mucosal leishmaniasis can be found in many Latin American countries and is caused by the *L. braziliensis* complex. The disease is endemic in 88 countries and is found in both tropical and subtropical regions, according to the World Health Organization (WHO). More than 15 million people are infected, and some 350 million people are at risk.

During its digenetic life cycle, *Leishmania* moves between the alimentary tract of the sandfly vector and the acidic phagolysosomes of mammalian macrophages. The ability of *Leishmania* to survive within phagolysosomes is heavily dependent on the developmental regulation of a variety of genes. Identifying the genes that are expressed preferentially during the parasite’s intracellular growth would help to elucidate the mechanisms controlling stage-specific gene regulation and intracellular survival of the parasite. Moreover, genes whose expression is modulated during the parasite’s cytodifferentiation may represent therapeutic targets. Therefore, considerable efforts are being directed toward identifying and characterizing these developmentally expressed genes and the molecular mechanisms underlying their regulation. The ongoing *Leishmania major* genome sequence project (www.ebi.ac.uk/parasites/leish.html) is contributing significantly to these efforts.

**Gene Organization and RNA Processing of *Leishmania*, an Ancient Eukaryote**

The trypanosomatid *Leishmania*, which diverged very early, is considered one of the most ancient eukaryotes. Compared to other eukaryotes, *Leishmania* are unusual in terms of gene organization, transcription, and mRNA processing. For example, several genes in *Leishmania* and *Trypanosoma* are arranged in tan-
dem arrays separated by intercistronic regions that help to regulate those genes and modify their transcripts.

As with other eukaryotes, *Leishmania* genes are transcribed by at least three different types of RNA polymerases (pol). RNA pol I and III transcribe the rRNA and tRNA genes, respectively, whereas RNA pol II, which is sensitive to low concentrations of the mushroom toxin α-amanitin, transcribes protein-coding genes. However, so far it has not been possible to map *Leishmania* transcription initiation sites for protein-coding genes, suggesting that its RNA pol II does not regulate transcription. How then does the *Leishmania* RNA pol II (and that of other trypanosomatids) transcribe genes without promoter elements?

Part of the answer is that *Leishmania* protein-coding genes are organized as polycistronic transcription units, making each primary transcript a polycistronic precursor RNA. Individual mRNAs are cleaved from the precursor by a trans-splicing reaction, which adds a capped 39-nucleotide-long spliced leader (SL) sequence to the 5′ end of every mRNA (Fig. 1). Meanwhile, the 3′ end of each mRNA is polyadenylated, despite an absence of dedicated polyadenylation signals. Instead, polyadenylation occurs at a fixed distance (100–400 nt) upstream of the splicing signals. Several pieces of experimental evidence indicate that trans-splicing and polyadenylation are mechanistically coupled and share common regulatory signals, mainly polypyrimidine-rich stretches within the intergenic regions. In addition to playing an essential role for mRNA maturation, these posttranscriptional RNA processing reactions may also help to control gene expression.

At the chromosomal level, *Leishmania* genes are organized differently from those of many other eukaryotes. Along several *L. major* Friedlin chromosomes, for example, long polycistronic units are transcribed in either a “convergent” or “divergent” manner. Analyzing
transcription on Leishmania chromosome I with irradiation and nuclear run-on assays, Peter Myler of the Seattle Biomedical Research Institute, Seattle, Wash., and his collaborators identified a site where transcription initiates bidirectionally within two oppositely oriented units.

In some cases, copies of tRNA genes are found between convergent units of protein-coding genes, potentially juxtaposing RNA pol II and pol III during transcription. Whether their convergence at such sites significantly affects gene expression is not known. In Leishmania, pol II may also be transcribing the noncoding strand of the chromosome randomly and at a low level in a nonspecific manner. In general, the molecular nature of the RNA pol II transcription complex remains one of the major mysteries of trypanosomatid molecular biology.

The Digeneric Life Cycle of Leishmania and Stage-Specific Gene Expression

Leishmania parasites exist in two major developmental stages—free-living flagellated promastigotes in the alimentary tracts of sandflies, which serve as vectors for the parasite, and intracellular amastigotes residing in the phagolysosomes of mammalian macrophages that serve as hosts for the parasites.

Inside the sandfly, the parasite exists in several forms and reaches maturation to become the infective form, called metacyclic promastigotes, that move to the proboscis of the insect, from which they may be transferred to the mammalian host during a blood meal. Metacyclic promastigotes attach to macrophages through receptor-mediated binding. Once the parasites are ingested, the macrophages form phagolysosomes in which the promastigotes differentiate into a nonflagellated intracellular form, called amastigotes, that continue to multiply and spread throughout macrophages of the reticuloendothelial system while clinical symptoms develop. The life cycle is completed when a sandfly takes a blood meal containing amastigotes, which transfer into the insect gut.

During its complex life cycle, Leishmania experience drastic environmental changes. Inside the sandfly gut, for instance, promastigotes are exposed to neutral pH and widely fluctuating temperatures. Amastigotes, on the other hand, reside in acidic phagolysosomes, experience near-constant temperatures, and encounter host factors that subject them to oxidative, proteolytic, and metabolic stresses.

In withstanding these conditions, amastigotes express a variety of developmentally regulated genes. Few of these stage-specific genes have been identified (Table 1). Using axenic culture systems, researchers can obtain and study amastigote-like forms of several Leishmania species. Such amastigotes express distinct surface antigens compared to promastigotes, including abundant amastin surface proteins. Their differential expression on amastigote plasma membranes could mean that these proteins help to protect the parasite against lysosomal attack, possibly by modulating macrophage functions.

Another amastigote-expressed gene of L. do-

---

**Table 1:** Known developmentally regulated genes in the intracellular amastigote form of Leishmania

<table>
<thead>
<tr>
<th>Genes</th>
<th>Posttranscriptional mechanisms of stage-regulated gene expression</th>
<th>Accession number</th>
<th>Presence of the conserved 450-nt 3’UTR element</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. donovani HSP83 gene</td>
<td>3’UTR</td>
<td>M92925</td>
<td>ND</td>
</tr>
<tr>
<td>L. donovani proton motive P-type ATPase 1b</td>
<td>ND</td>
<td>AF109296</td>
<td>ND</td>
</tr>
<tr>
<td>L. donovani A2 gene</td>
<td>3’UTR</td>
<td>s69693</td>
<td>Yes</td>
</tr>
<tr>
<td>L. donovani HSP100 gene</td>
<td>ND</td>
<td>294053</td>
<td>Yes</td>
</tr>
<tr>
<td>L. infantum HSP70 gene</td>
<td>3’UTR+IR</td>
<td>Y08020</td>
<td>ND</td>
</tr>
<tr>
<td>L. infantum triose phosphate isomerase gene</td>
<td>ND</td>
<td>AF459814</td>
<td>NF</td>
</tr>
<tr>
<td>L. infantum amastin gene family</td>
<td>3’UTR</td>
<td>AF195531</td>
<td>Yes</td>
</tr>
<tr>
<td>L. chagasi cysteine protease CYS1 gene</td>
<td>ND</td>
<td>AF004592</td>
<td>NF</td>
</tr>
<tr>
<td>L. pifanoi CYS2 cysteine proteinase gene</td>
<td>ND</td>
<td>M97695</td>
<td>NF</td>
</tr>
<tr>
<td>L. mexicana CPB3-CPB18 genes</td>
<td>3’UTR+IR</td>
<td>Z14061-Z49963</td>
<td>NF</td>
</tr>
<tr>
<td>L. mexicana AF032464</td>
<td>ND</td>
<td>AF032464</td>
<td>NF</td>
</tr>
<tr>
<td>L. mexicana A600 gene</td>
<td>ND</td>
<td>AF345946</td>
<td>NF</td>
</tr>
<tr>
<td>L. major HSPs (A1, A2, B) genes</td>
<td>NDA/237587</td>
<td>NF</td>
<td></td>
</tr>
<tr>
<td>L. major histone H1 gene</td>
<td>3’UTR</td>
<td>AJ223680</td>
<td>Yes</td>
</tr>
<tr>
<td>L. major-tubulin gene</td>
<td>3’UTR?</td>
<td>X93566</td>
<td>NF</td>
</tr>
<tr>
<td>L. chagasi peroxidoxin 1 (LoPxn1)</td>
<td>ND</td>
<td>AF134161</td>
<td>NF</td>
</tr>
<tr>
<td>L. donovani mitochondrial chaperonin cpn60.2</td>
<td>ND</td>
<td>AF100775</td>
<td>NF</td>
</tr>
</tbody>
</table>

3’UTR: 3’ untranslated region; IR: intercistronic region; ND: not determined; NF: not found.
novani, whose protein product is designated A2, is associated with virulence and infectivity. Amastigotes also harbor several additional stage-specific proteins, including HSP70, HSP83, and HSP100, which are induced upon heat stress inside human macrophages, cathepsin L-like cysteine proteinases, chaperonins, and histones (Table 1).

In addition to these developmentally regulated protein-coding genes, several small noncoding RNAs are also regulated in a stage-specific manner in *Leishmania*. These include the spliced leader SL-RNA that is essential for pre-mRNA trans-splicing in these parasites, RNAs encoded by tandemly repeated subtelomeric sequences of unknown function, and a number of small nucleolar RNAs (snoRNAs) that guide 2’-O-ribose methylation of rRNA sequences. Some of these noncoding RNAs are polyadenylated in amastigotes, even though they ordinarily are not polyadenylated. Polyadenylation of these RNAs might affect RNA processing reactions that could down-regulate general mRNA translation in response to environmental stresses.

**Untranslated Regions Help To Control Stage-Specific Gene Expression**

Sequences within 3’-untranslated regions (3’UTR)—acting posttranscriptionally—play a key role in regulating differentially expressed trypanosomatid genes. Indeed, the lack of promoter elements for RNA pol II and the unusually long 3’UTR sequences provide the molecular basis for this type of control. Posttranscriptional regulation mechanisms for gene expression include differential processing of polycistronic transcripts, alternating rates of mRNA decay, and protein translation, each of which is associated with such 3’UTR sequences.

We are studying molecular mechanisms underlying stage-specific gene expression in *Leishmania*, applying differential hybridization screening techniques to identify, among others, a developmentally regulated gene family coding for the amastin surface proteins. Our findings indicate that the last 770 nt of the amastin 3’UTR are responsible for stage-specific regulation.

Moreover, sequences within the 770-nt region of the amastin 3’UTR are involved in developmentally regulating the amastin transcript—increasing mRNA translation without affecting its stability. This finding represents the first example of a conserved 3’UTR element being directly implicated in mRNA translation in these parasites. Several other 3’UTRs implicated in stage-specific gene regulation in *Trypanosomatidae* appear to play a role in mRNA stability and translation, according to Michal Shapira and...
her collaborators at Ben Gurion University of the Negev, Israel, and also to Christine Clayton and collaborators at the University of Heidelberg in Germany.

We wondered whether other differentially expressed genes in Leishmania amastigotes are also regulated at the level of translation with a similar mechanism. To address this question, we took the 770-nt sequence of the amastin 3'UTR and looked for homologous sequences in Leishmania major databases (about 70% of the genome is sequenced), and found more than 100 BLAST entries that share 65–79% identity to

the first 450 nt of this 3'UTR region. These sequences correspond to the 3' end of some 100 protein-coding genes (Fig. 2).

Almost half of these protein-coding genes are not annotated or belong to unclassified hypothetical proteins. However, the rest of the proteins with a predicted function can be classified either as surface proteins; amastin homologues and transporters; histones and chaperones; and proteins involved in energy metabolism, signaling pathways, gene regulation like spliceosome assembly and translation initiation, cell adhesion, or ubiquitin-dependent protein degradation (Fig. 2). Remarkably, among these protein-coding genes, some correspond to known amastigote-specific proteins, suggesting a correlation between the highly conserved 3'UTR element and amastigote-regulated gene expression.

Reporter gene assays combined with Northern blot and Western blot hybridization studies show a very good correlation between the conserved 3'UTR 450-nt element introduced downstream of a reporter gene and differential translational regulation of the reporter mRNA. Moreover, the 450-nt element enhances mRNA translation specifically in amastigotes by enabling them to associate more tightly with polysomes, which is a hallmark of highly translated mRNAs. Hence, the widespread distribution of the 450-nt conserved element among numerous leishmanial transcripts and its role in regulating mRNA translation in a stage-specific manner suggest a regulatory mechanism that is shared by a specific class of differentially expressed genes in Leishmania.

**Translational Control Provides Several Adaptive Advantages**

Regulating genes through a translational control mechanism confers valuable advantages, including speed and a readily reversible response to changing physiological conditions. That speediness and versatility enable cells to modulate membrane composition, activate signal

---

**FIGURE 3**

Sequence alignment of highly conserved regions within the 3'UTRs of a selected number of leishmanial amastigote-specific transcripts using the GCG Pileup program. These include the known amastigote-specific genes; L. infantum amastin gene and L. major histone H1 gene (AC008242), the L. major Lmflchr31_02 and Lmflchr34_00 amastin gene homologues, and the 3-ketoacyl-CoA thiolase (AL117263) and elf4E-E leishmanial orthologues.
transduction pathways, regulate general and specific mRNA translational efficiency, and adjust needs for energy metabolism and nutrient requirements when changing conditions dictate changing needs.

Several genes regulated by the 3′UTR 450-nt conserved element fall within these categories and could very well be regulated at the level of translation. Although the underlying molecular mechanisms are not understood, investigators speculate that 3′UTR elements regulate translation by looping between 5′ and 3′UTR sequences, deadenylating target mRNA molecules, or by controlling the distribution of those molecules within the cytoplasm.

These regulatory elements, with the exception of mRNA localization signals, contain short AU-rich sequences and are not similar in sequence to the conserved 450-nt element found in leishmanial mRNAs. Given the extensive length of this 3′UTR element and its conserved sequence (Fig. 3), it seems likely that it adopts a particular tertiary structure and regulates gene activity through protein-RNA interactions. Indeed, a computer-based structural analysis using the mfold program is consistent with this possibility. That analysis reveals that the conserved 450-nt elements in the 3′UTR segments of several amastigote-specific mRNAs fold into a similar bipartite Y-shaped stem-loop structure, which might enable them to bind a common trans-acting factor.

However, not all the amastigote-specific transcripts in Leishmania harbor this conserved 450-nt element in their 3′UTR (Table 1), and even in its presence it is not rare that additional sequences within the 3′UTRs could regulate mRNA stability and/or translation, suggesting that stage-regulated gene expression in Leishmania is a complex process involving a variety of mechanisms. Stage-specific gene expression in Leishmania can be controlled either at the level of mRNA stability or at the level of translation by a bipartite mechanism involving distinct 3′UTR elements that mediate regulation at the level of mRNA abundance and translation, respectively. We now are using two-dimensional gel electrophoresis and mass spectrometry to identify other differentially expressed proteins to map 3′UTR signals and characterize these and other mechanisms that control stage-specific gene expression in these parasites.

Environmental Signals, Stage-Specific Regulation when Leishmania Differentiates

When Leishmania promastigotes in the fly transform into obligatory intracellular amastigotes in human macrophages, the parasite is subjected to abrupt environmental changes, including a drop in pH and a sharp rise in temperature. These external signals can trigger Leishmania cell differentiation even when cells are raised in axenic culture systems.

Heat shock induces several Leishmania chaperones and heat shock proteins, including HSP70, HSP83, and HSP100. Heat shock can cause misfolding of a fraction of cell proteins, thus altering gene regulation at the level of transcription and translation, according to Anthony Fink of the University of California, Santa Cruz. Meanwhile, subjecting cells to acidic pH conveys a stress signal that elicits transduction pathways in various microorganisms, including acid-tolerant bacteria and fungi, according to Holly Hall and collaborators at the University of South Alabama, Mobile. These microorganisms respond to fluctuating pH via specific sensors and transmitters on their surface membranes. The sensor typically is a transmembrane protein possessing intrinsic kinase activity that responds to pH changes in part through conformational changes. The transmitter kinase continues to pass the signal by phosphorylating intracellular components that regulate gene expression.

The survival of Leishmania parasites within acidic phagolysosomes of mammalian macrophages proves that they adapt quickly to sharply changed conditions that are encountered there. Although we do not know whether Leishmania have such a pH-detecting and transmitting system, acidic pH helps to trigger Leishmania cell differentiation into the amastigote form. Moreover, in axenic culture, acidic pH is essential for the parasite’s differentiation. Growing such parasites at acidic pH rapidly induces 450-nt conserved element-controlled translational regulation of amastigote-specific mRNAs, including the amastin gene, several amastin gene homologues, and the A2 gene.

Thus, acidic pH apparently is an important signal for initiating 450-nt 3′UTR element-dependent translational control. The pH signaling pathway in Leishmania might also be controlled by phosphorylation-dephosphorylation
mechanisms similar to those in other organisms, and we plan several experiments to address this possibility. In addition to acidic stress, we are studying whether oxidative and metabolic stresses affect translational regulation in these parasites.

Perspectives

Unlike more complex eukaryotes, gene expression in Leishmania is regulated mainly at the posttranscriptional level. While studying the amastigote-specific amastin gene family, we identified a conserved 3′UTR element among a set of developmentally regulated genes encoding diverse functions. Our data support the possibility that a common mechanism mediated by a conserved 3′UTR element controls stage-specific translational regulation of a specific subset of transcripts in Leishmania. To better understand this mechanism, we plan to identify and characterize the regulatory protein(s) that interact with this conserved 3′UTR element. Preliminary data from our RNA binding assays indicate that at least one such amastigote-specific protein binds to this conserved element.

Both acidic pH and elevated temperature help to trigger Leishmania cytodifferentiation, leading to stage-specific expression of a variety of genes. However, acidic pH appears more important for activating a pathway leading to differential expression of selected mRNAs harboring the conserved 450-nt 3′UTR element through a mechanism that involves translational regulation. To understand stage-specific translational control, we need to identify the pH signaling events underlying this process, knowledge of which eventually could help researchers to design better means for controlling these parasites.

ACKNOWLEDGMENTS

We thank Carole Dumas and Marthe Dubé for permitting the citation of unpublished data and Marc Ouellette for helpful comments on the manuscript.

This work was supported by a Canadian Institutes of Health Research (CIHR) grant # gr-14500 to BP. NB and FM hold a CIHR and NSERC studentships, respectively. BP is a Fonds de Recherche En Santé de Quebec (FRSQ) Senior Scholar and a Burroughs Wellcome Fund New Investigator in Molecular Parasitology.

SUGGESTED READING


