RESEARCH NOTE

# *Leishmania panamensis*: a 44bp Deletion in gp63 Gene is Found in cDNA and Genomic Libraries

**RD Hoya, CE Trujillo, C Cardenas, F Puentes, ME Patarroyo, LA Murillo+**

Instituto de Immunologia, Hospital San Juan de Dios, Universidad Nacional de Colombia, Santafe de Bogotá, Colombia

Key words: *Leishmania panamensis* - major surface glycoprotein - GP63 - deletion - DNA sequence

Leishmanolysin (EC 3.4.24.36) (gp63) is a zinc metalloprotease (ND Rawlings et al. 1995 *Methods Enzymol 248*: 183-228) abundantly expressed in the promastigote form of *Leishmania* parasites, where it is attached to the plasma membrane by a glycosylphosphatidylinositol (GPI) anchor (P Schneider et al. 1990 *J Biol Chem 265*: 16955-16964, TJ Salvatore et al. 1992 *Annu Rev Microbiol 46*: 65-94). It has been suggested that gp63 becomes active at multiple steps during the process of macrophage invasion and its function seems to be required for the parasite's intraphagolysosomal survival (GD Rusell et al. 1989 *Immunol Today 10*: 328-333, AR Miller et al. 1990 *Mol Biochem Parasitol 39*: 267-274, A Brittingham et al. 1996 *J Immunol 155*: 3102-3111). The protein is encoded by genes present in multiple copies that are clustered in two to three tandem repeats usually located on a single chromosome (JR Webb et al. 1991 *Mol Biochem Parasitol 48*: 173-184, T Hanekemp et al. 1991 *Mol Biochem Parasitol 48*: 27-38) and which are highly conserved throughout the genus (LL Button et al. 1988 *J Exp Med 167*: 724-729, HB Steinkraus 1993 *Mol Biochem Parasitol 62*: 173-186, SC Roberts et al. 1993 *Mol Biochem Parasitol 62*: 157-172, E Medina-Acosta et al. 1993 *Mol Biochem Parasitol 57*: 31-46).

The gp63 familiy is constituted by a group of proteins with different molecular weights depending upon the species. The importance of the major surface glycoprotein of *Leishmania* promastigotes, gp63, in the binding of promastigotes to macrophages has been inferred largely due to its abundance, surface localization, and proteolytic activity (J Bouvier et al. 1985 *J Biol Chem 260*: 15504-15509, R Etges et al. 1986 *J Biol Chem 261*: 9098-9101, CS Chang 1986 *Proc Natl Acad Sci USA 83*: 100-104, Brittingham *loc. cit.*).

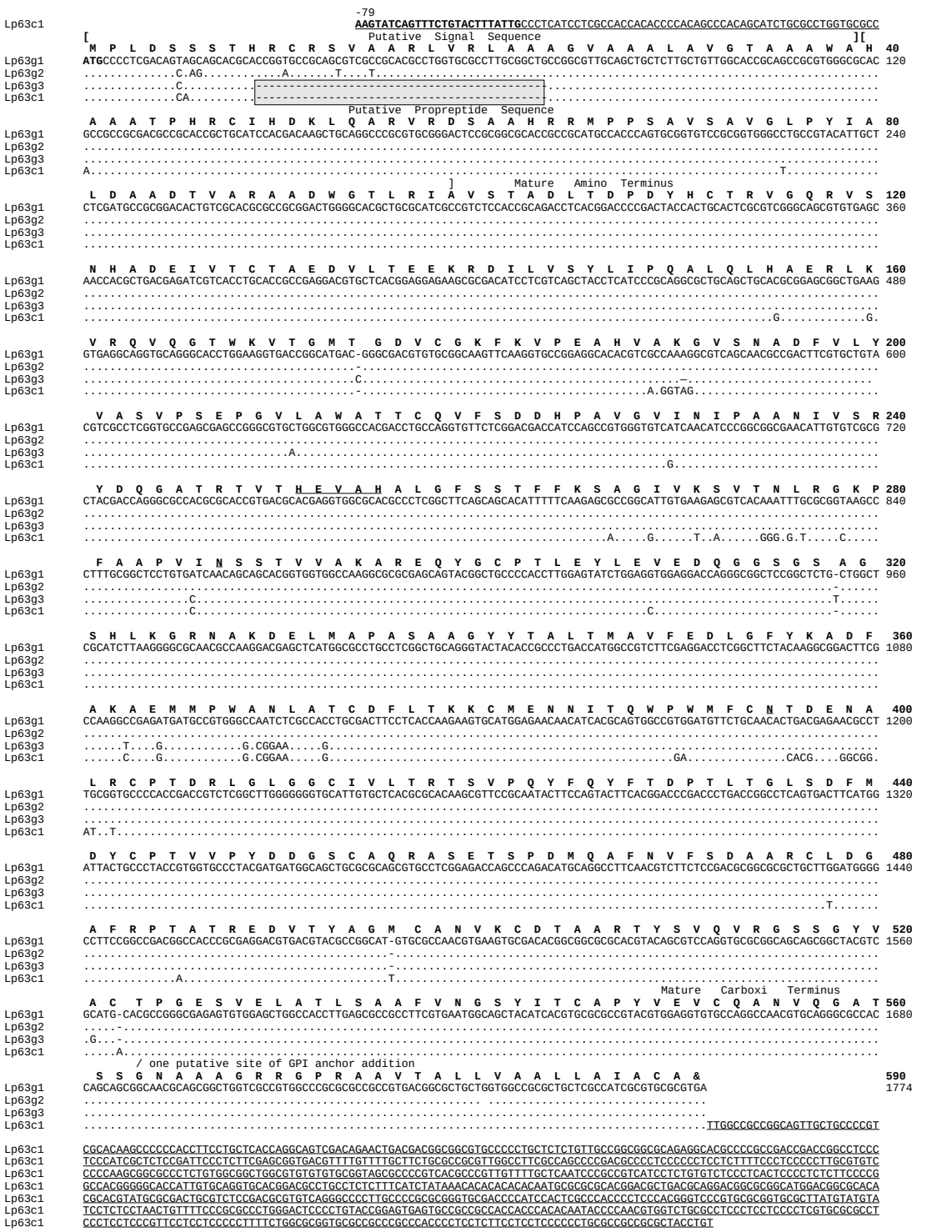
In the present work, we report for the first time the molecular characterization of gp63 gene in *L. panamensis*, which is the most prevalent *Leishmania* species in Colombia, and also the presence of copies containing deletions of this gene either on cDNA or DNA material.

In order to identify the gp63 gene from *L. panamensis*, an internal 940 bp long probe was generated by PCR, using two consensus primers designed from the alignment of all *Leishmania* gp63 nucleotide sequences reported to date. Lpan1: 5'TACGTCGCCTCGGTGCCGA3' and Lpan2: 5´GCACCTGGACGCTGTACG3' primers were synthesized by the solid phase phosphito-triester method. This probe (U62634) was radiolabelled and used for hybridization assays and library screening.

A cDNA library was constructed from *L. panamensis* infective clone M/HOM/PA/71/LS/74 stationary phase promastigotes. cDNA was synthesized using a cDNA synthesis kit (Amersham) and ligated into the Eco RI site of bacteriophage λgt11. Several clones were isolated and sequenced either by manual (Sequenase v.2.0; United States Biochemical) or automatic (Applied Biosystems 373; Perkin Elmer) sequencing using both DNA strands. Of those, Lp63c1, represents a 2648-nt long cDNA sequence that is homologous to other gp63 sequences and includes the 1728 nucleotides coding region, 79 nt of the 5' untranslated region, and 841 nt of the 3' untranslated region (Figure). However, this clone exhibits a 44 bp deletion 26 nt downstream to the start codon causing a frameshift compared to other previously described gp63 genes from *Leishmania* spp.

In order to detect the presence of this gp63 form in *L. panamensis* genome, we built a Sal I genomic library containing restriction fragments ranging from 2.8-3.2 kb in size into the pMOS vector

This research was supported by the Presidency and Ministry of Public Health of Colombia, Colciencias, and the German Leprosy Relief Association.
The nucleotide sequence data reported in this paper will appear in the GenBank Nucleotide Sequence Database under accession Nos. U62634, AF038028, AF037165, AF037166 and AF037167
+Corresponding author. Fax: +57-1-2803999. E-mail: mepatarr@bacata.usc.unal.edu.co
Received 9 November 1998
Accepted 23 March 1999




Sequence of nucleotides and deduced amino acids of gp63 gene in *Leishmania panamensis*: aligned nucleotide sequences of *L. panamensis* gp63 cDNA Lp63c1 and genomic DNA Lp63g1, Lp63g2, Lp63g3 clones. Start and stop codons and spliced leader partial sequence are represented by bold upper case. The 44bp deleted in Lp63g3 and Lp63c1 are boxed. Sequence identity with *L. panamensis* Lp63g1 is denoted by a dot. Gaps in the sequences (indicated by a dash) have been introduced for maximum sequence identity. 5'UTR and 3'UTR sequences of gp63 cDNA clone are underlined. Putative signal sequence (residues 1-39), propeptide (residues 40-99), amino terminus and carboxy terminus of the mature protein (residues 100-563), and GPI anchor addition site are indicated. Potential glycosylation sites and the proposed active site/zinc-binding site domain (residues 251-255) are underlined.

(Amersham). Out of three clones isolated, one (Lp63g3) displays the same 44 bp deletion observed in clone Lp63c1. This copy has also some nucleotide changes including an insertion of a cytosine at position ± 520, a thymine at position ± 954, and a deletion of two nucleotides at position ± 570 which puts it out of frame several times (Figure). This suggests the presence of at least two non-functional copies of gp63 gene. The remaining two clones (Lp63g1 and Lp63g2) show no deletions, so they would theoretically be expected to code for complete, functional proteins. Nucleotide sequence of Lp63c1, Lp63g1, Lp63g2 and Lp63g3 display 95% identity among them (Figure). Furthermore, the coding sequence of the Lp63g1 and Lp63g2 clones give 80% and 70% identity with the *L. guyanensis, L. major, L. donovani, L. chagasi,* and *L. infantum* reported genes by nucleotide and deduced amino acid sequence, respectively (Button *loc. cit.*, Steinkraus *loc. cit.*, Roberts *loc. cit.*, Medina-Acosta *loc. cit.*). However, the protein sequence corresponding to the mature protein (residues 100-563, Figure) indicates that *L. panamensis* sequences are 75% identical to L. *guyanensis* and 70.5% identical to *L. major, L. donovani, L. chagasi,* and *L. infantum*.

The gp63 protein from *L. panamensis* shares some notable features with respect to several putative functional regions in other *Leishmania* species. There is a putative signal sequence cleavage site between amino acids Ala-39 and His-40 and a potential propeptide cleavage site between amino acids Ala-99 and Val-100. Also, the proposed active site/zinc binding site domain is homologous to the one proposed for other zinc metalloproteinases, including some members of the matrix metalloproteinases (RW McMaster et al. 1994 *Parasitology 108*: S29-S36). Two potential N-glycosylation sites (Asn-287 and Asn-395) out of the three defined for *L. guyanensis* and *L. major* gp63 are retained in *L. panamensis*. As observed in other species, *L. panamensis* gp63 contains a hydrophobic carboxyl terminal region where the membrane anchor motif is attached. The above mentioned sites have been verified experimentally by site-specific mutation (BS McGwire & KP Chang 1996 *J Biol Chem 271*: 7903-7909). The presence of deletions inside the coding region of some *L. panamensis* gp63 gene copies indicates the existence of pseudogenes within the gp63 gene cluster probably due to mechanisms such as gene duplication or recombination, processes that could certainly operate in trypanosomatids. Given that *Leishmania* bears so many gp63 genes in close proximity, pseudogenes could easily appear between functional genes. Gene clusters are supposed to be translated as policystronic RNA messages where pseudogenes would be included; post-transcriptional processing would then cause these pseudogene transcripts or translation products to degrade and become non-functional. Nevertheless, clone Lp63c1 has an spliced leader partial sequence (25 nt. out of 35 nt.), common to all trypanosomatids species, 54 nt. upstream from the initiation codon (Figure). *L. panamensis* gp63 gene sequences could represent differentially expressed copies, or they could even form a repertoire of concurrently expressed gp63 proteins as it has been already suggested (LL Button et al. 1989 *Mol Biochem Parasitol 32*: 271-284, Webb et al. *loc. cit.*, Steinkraus *loc. cit.*, Roberts et al. *loc. cit.*, Medina-Acosta et al. *loc. cit.*).

The high conservation of the gp63 gene observed throughout *Leishmania* species, suggests a strong selective pressures to maintain the role this protein might play for the parasite's survival (Button et al. *loc. cit.*), making it relevant for the design of leishmanial synthetic vaccine candidates.