1. Introduction

Protozoan parasites of the genus *Leishmania* are the causative agent of the disease leishmaniasis, which infects more than 10 million people worldwide [1]. *Leishmania* parasites are transmitted through sand fly bites, where the flagellated promastigote cells are introduced into the mammalian host. Parasites are then taken up by macrophages and differentiate into the non-flagellated amastigotes. Depending on parasite species, *Leishmania* infections in humans cause manifestations from self-containing cutaneous lesions to lethal visceral infections.

The surface of the promastigote stages of *Leishmania* is coated with a variety of interrelated glycoconjugates including lipophosphoglycan (LPG), glycosylinositolphospholipids (GIPLs), proteophosphoglycan (PPG), and GPI-anchored proteins [2–4]. In the sand fly, stage-specific modifications of LPG are responsible for the attachment and release of promastigotes from the midgut [5–7]. In the mammalian host, LPG confers resistance to complement-mediated lysis, oxidative stress, and inhibits phagolysosomal fusion [8,9]. However, LPG is down-regulated in amastigotes, whose surface is dominated instead by the abundant GIPLs, which suggested that these play critical roles in amastigote survival and virulence [10,11]. This view was recently called into question by studies of a *Leishmania major* mutant lacking the enzyme alkyldihydroxyacetone phosphate synthase (ADS1), which is required for the synthesis of ether phospholipids. The *ads1* mutant lacked LPG and GIPLs but maintain the presence of GPI-anchored proteins, probably bearing a modified GPI anchor [12]. Remarkably, *ads1* mutants showed little phenotype beyond that attributable to loss of ABCs.
LPG (e.g., reduced ability to establish macrophage infections), and \(a d i^1\) amastigotes can replicate in macrophages and cause disease in susceptible mice [12]. In *L. mexicana*, there are conflicting data about the role of GIPLs in parasite viability and/or virulence [13–15]. A potential drawback is that the various *Leishmania* mutants studied were pleiotropic and affected a number of glycoconjugates, making inferences about GIPL function indirect. Thus, the identification of genes that specifically affect GIPL synthesis would simplify and strengthen genetic studies of their function. *Leishmania* GIPLs have different structures in different species, consisting of type I GIPLs whose structure resembles that of protein GPI anchors, type II GIPLs whose structure resembles that of the LPG core, and/or “hybrid” GIPLs [3,10,11]. *L. major* synthesizes primarily type II GIPLs, which contain galactosylfuranose (Gal\(^f\)) in the Gal\(_3\)-containing GIPL sequence, which is also found in the LPG core ([16,17], Fig. 1). We showed previously that the Gal\(_f\)-containing GIPLs [18]. Thus, there must be additional Gal\(_f\)Ts that are responsible for the synthesis of GIPLs. Given the similarity of the Gal\(_f\)/T acceptors for LPG and GIPLs, it seemed reasonable to postulate that these GPL-specific Gal\(_f\)/Ts could be related to *LPG1*. In this report we describe the characterization of a family of *LPG1* homologs in the *L. major* genome, and studied the effect of genetic inactivation of three of these singly or in combinations (*LPG1*, *LPG1L*, and *LPG1R*).

2. Materials and methods

2.1. Leishmania cultures

*L. major* LTV3 clone 5 (Rho/SU599P) cells were grown in M199 medium supplemented with 10% heat inactivated fetal calf serum at 26°C [19]. Selective drugs used in this study included G418 (10 \(\mu\)g/ml), hygromycin (50 \(\mu\)g/ml), puromycin (10 \(\mu\)g/ml), nourseothricin (80 \(\mu\)g/ml), blasticidin (10 \(\mu\)g/ml), and phleomycin (10 \(\mu\)g/ml).

2.2. Isolation of *LPG1L* and *LPG1R* open reading frames

At the time these studies were initiated the *L. major* genome was largely incomplete. Sequence data was obtained from the Sanger Institute website at http://www.sanger.ac.uk/Projects/L_major/. Searches with *LPG1* first identified a 537-bp region from the *Leishmania* Genome Database; this region was PCR amplified and used to identify a cosmids (B4308) containing the full length gene from a *LHYG* genomic cosmid library made from *L. major Friedlin* strain V1 (MHOM/IL/80/Friedlin). This gene was termed *LPG1L* (for *LPG1*-like) and its open reading frame and flanking regions were sequenced (GenBank accession number AY235572).

A second gene termed *LPG1R* (LPGL1-related) was similarly identified in the *Leishmania* Genome Database; this ORF and flanking sequences were recovered by PCR and their sequences were determined (AY235573). Preliminary sequence data from *T. cruzi* was obtained from The Institute of Genomic Research through their website at http://www.tigr.org.

2.3. Molecular constructs

To generate the C-terminal GFP tagged proteins, full length open reading frames lacking stop codons were PCR amplified from *L. major* FV1 genomic DNA using primers SMB1307 (5'-TTATCGagctcACCATGAAAGGCAGAC-TAC'-3', lower case letters indicate added restriction sites) and SMB1310 (5'-ATATAggcgccgGAGCTGACAGC-CTGACAG for *LPG1L* and SMB1354 (5'-TATACGagctcAC-CAATGGAACATCAGGCTTGCG-3') for *LPG1R*). The resulting DNA fragments were digested with appropriate restriction enzymes (BamH I and NotI for *LPG1L*, BamH I and EcoRV for *LPG1R*) and cloned into the expression site of pXG<sup>-</sup>-(GFP:<sup>+</sup>) (a vector for making C-terminal GFP fusion proteins; strain B2863 [20]) to make pXG-*LPG1L-GFP* (B4400) and pXG-*LPG1R-GFP* (B4455).
To make deletion constructs for LPG1L, a 700-bp region immediately upstream of the start codon and a 900-bp region immediately downstream of the stop codon of LPG1L were PCR amplified from L. major genomic DNA and cloned into vector pX63PAC (B1129) [21] so that the end of the upstream region is linked to the beginning of the downstream region. Next a 900-bp NEO marker gene (conferring resistance to G418) was excised from vector pXG (B4090) [20] and inserted in between the upstream and downstream regions of LPG1L to make pX63PAC-KO-LPG1L-NEO (B4369). Similarly, a 1052-bp HVG marker gene (conferring resistance to hygromycin) was excised from vector pX63HYG (B617) [22] using SpeI and BamHI and inserted in between the upstream and downstream regions of LPG1L to make pX63PAC-KO-LPG1L-HYG (B4370). Linear deletion cassettes containing NEO and HVG markers flanked by the upstream and downstream regions of LPG1L were generated by SnaI digestion.

To make deletion constructs for LPG1R, first a 2.4-kb DNA fragment containing the LPG1R ORF (plus sequences 550 bp from upstream and 440 bp from downstream) was PCR amplified from L. major genomic DNA and cloned into vector pUC18 to give pUC-LPG1R (B4413). Then a SacI/SphI fragment containing the first 500 bp of LPG1R ORF was replaced by drug markers (PAC or SAT, conferring resistance to puromycin or nourseothricin, respectively) amplified by PCR. The resulting constructs, pUC-KO-LPG1R-PAC (B4451) and pUC-KO-LPG1R-SAT (B4513) were digested with EcoRI and HindIII to generate linear deletion cassettes for LPG1R.

Previously, a 4.2-kb fragment containing the LPG1 ORF and flanking sequences was cloned into the BamHI site of vector pUC18 (pUC-LPG1-Bam, B2880). To make the deletion construct for LPG1, the first 700 bp from LPG1 ORF was replaced by the BSD (confers resistance to blasticidin) marker gene. The resulting construct, pUC-KO-LPG1-BSD (B4573), was digested with BamHII to generate a linear knockout cassette for LPG1.

An overexpression construct for LPG1L, LPG1L, and LPG1R was made in three steps. First, the full-length LPG1L gene was cloned in the BglII site of vector pRI1PHLEO (B4054) derived from pRI1SAT (B3541) [23]) to make pRI1PHLEO-LPG1L (B4587). Second, the full-length LPG1R gene was cloned in the BamHI site of pRI1PHLEO-LPG1L (B4589). The resulting construct, pRI1PHLEO-LPG1L-LPG1R, was linearized with NoI and blunt ended, followed by ligation with the 4.2-kb LPG1 fragment (BamHI digestion of pUC-LPG1-Bam then blunt ended) to generate pRI1PHLEO-LPG1-LPG1L-LPG1R (B4789).

2.4. Genetic manipulations

Leishmania were transfected by electroporation as described [24]. Typically, 5–10 μg of DNA was used for each transfection. Transfected cells were plated on medium containing 1% noble agar and appropriate concentrations of selective drugs. To study the localization of LPG1L and LPG1R, L. major cells containing pX63HYG-LPG2-HA (a marker of the parasite Golgi apparatus; strain B1878 [22]) were transfected with pXG-LPG1L-GFP (B4400) or pXG-LPG1R-GFP (B4455) and selected with hygromycin and G418. The p[pg1l′-pg1r′]-triple knockout parasites were created using a loss-of-heterozygosity (LOH) approach [25]. Briefly, the first allele of LPG1 in the pg1l′-pg1r′ double knockout was replaced by the BSD marker. After two passages in the presence of 50 μg/ml of blasticidin, cells were incubated in media containing 10 μg/ml of galactose-binding lectin ricin RCA120 (Sigma) at 10^7 cells/ml. Cells that failed to agglutinate were plated, and ~80% of the resulting colonies had lost the second allele of LPG1. All knockout strains were confirmed by Southern-blot analysis.

2.5. Molecular biology techniques

Southern-blot and Northern-blot analyses were performed as previously described [26]. The Southern blot shown in Fig. 9 was subsequently washed first in 2× SSPE/0.5% SDS for 15 min at room temperature, and then three times in 0.2× SSPE/0.5% SDS at 60 °C for 15 min each. RT-PCR analysis was used to map the mRNA splicing acceptor sites of LPG1L and LPG1R. Briefly, first strand synthesis was performed using gene-specific antisense primers SMB1228 (5′-CCGTTGACCACTATGTGGGCAG) for LPG1L and SMB1369 (5′-CAGCCAGCCGCACAGCTGCT) for LPG1R. Reverse transcription was performed at 42 °C with AMV reverse transcriptase (Sigma) and 2 μg of total RNA, followed by PCR amplification with the first strand primers plus the universal miniexon primer SMB936 (5′-AAGCGCTAATATAGTATACGATTCTGTACCTTTA) for L. major [27]. The resulting fragments were sequenced to map the mRNA splice acceptor sites.

2.6. Western blots and fluorescence microscopy

Log phase (1–10^6 cells/ml) cell extracts and culture supernatants were resolved by SDS/PAGE and electroblotted onto Hybond ECL nitrocellulose membranes (Amer sham Biosciences). Mouse monoclonal antibody WIC79.3 [28] was used to detect LPG and PPG (1:10000 dilution). Mouse anti-L. major GP63 monoclonal antibody [29] was used to detect GP63 (1:10000 dilution). Mouse monoclonal anti-a-tubulin antibody (TS168, Sigma) was used to detect α-tubulin (1:10000 dilution). An enhanced chemiluminescence (ECL) detection system (Amersham) was used to detect signals. For fluorescence microscopy, log phase cells were immobilized on polylysine coated cover slips, followed by fixation with 3.5% formaldehyde in PBS. After being permeated by 100% ethanol at 4 °C, cells were incubated with dilutions of primary and secondary antibodies as described [30].
2.7. Mouse footpad infections

Parasite virulence was evaluated by mouse footpad infections. Briefly, all parasite lines were first inoculated at high levels (2 × 10^7) into BALB/c mice in order to guard against effects arising from transfection/culture passage associated loss of virulence. They were recovered after 1 month prior to emergence of pathology, and converted into promastigotes in vitro. After three passages in culture, stationary phase parasites (day 3) were resuspended in DEMEM media at 2 × 10^7 cells/ml and 10^8 were injected into the footpads of 8-week-old female BABL/c mice (5–6 mice per group). Lesion sizes were measured weekly with a Vernier caliper and parasite numbers in the infected foot pads were determined by limiting dilution assay [31].

2.8. GPIPL analysis

Exponentially growing cells were metabolically labeled with [3H]galactose (50 μCi) for 8 × 10^8 cells) for 16h, and [3H]GPIPLs were purified as described [32]. The [3H]GPIPLs were treated with 0.04 M trifluoroacetic acid for 1 h at 100 °C to cleave the labile galactofuranosidic bonds, and the [3H-Gal]fragments were chromatographed on paper as described [33]. Other aliquots of unlabeled GPIPLs were subjected to nitrous acid deamination [33] to remove the lipid anchors, tagged at the reducing end with the fluorophore 9-aminonaphthalene 1,3,6-trisulfate, and eluted on paper as described [33]. Other aliquots of unlabeled GPIPLs were subjected to nitrous acid deamination [33] to remove the lipid anchors, tagged at the reducing end with the fluorophore 9-aminonaphthalene 1,3,6-trisulfate, and eluted on paper as described [33]. Other aliquots of unlabeled GPIPLs were subjected to nitrous acid deamination [33] to remove the lipid anchors, tagged at the reducing end with the fluorophore 9-aminonaphthalene 1,3,6-trisulfate, and eluted on paper as described [33]. Other aliquots of unlabeled GPIPLs were subjected to nitrous acid deamination [33] to remove the lipid anchors, tagged at the reducing end with the fluorophore 9-aminonaphthalene 1,3,6-trisulfate, and eluted on paper as described [33].

2.9. LPG and glycoprotein analysis for the LPG1-LPG1L-LPG1R overexpressors

Parasites (5 × 10^6 cells/ml) were metabolically labeled with [3H]galactose (50 μCi) in 5 ml of M199 culture medium supplemented with 10% fetal calf serum and 1 μg/ml of biotin. [3H]LPG was solubilized by differential organic solvent extraction and removed by chromatography on phenyl-Sepharose as described [38]. The insoluble residue containing [3H]glycopeptides was converted to [3H]glycopeptides by digestion with Pronase as described [34] and desalted by chromatography on Sephadex G25. Maltohexose (1 μmole) was added to the protein residue prior to Pronase digestion as an internal standard and was quantitated by capillary electrophoresis using conditions: 20°C, 20 psi, and 20 min.

3. Results

3.1. Identification of a family of six LPG1-related genes in L. major

Search for sequences with similarity to L. major LPG1 revealed the presence of fragments from related sequences in the L. major (Friedlin V1 strain) genome project data, which was then incomplete. The sequence for two of these genes, designated LPG1L, (LPG1-like) and LPG1R (LPG1-related), was determined. Southern-blot analysis and gene deletion studies showed that these genes occurred in a single copy (data not shown or below). As these studies were being finished, the emerging L. major genome sequence revealed three additional related homologous genes that were named LPG1G, whose predicted amino acid sequences were identical and located on chromosomes 5, 19, and 32 (LmjF05.1230, LmjF19.650, and LmjF32.3900; sequence information is available at http://www.gendb.org/genedb/leish/index.jsp). Southern-blot analysis confirmed the presence of multiple LPG1G genes in L. major strain LV39 clone 5 as well (data not shown). Interestingly, all three of the L. major Friedlin strain LPG1G genes appear to be located adjacent to telomeres. LPG1 has been described in several Leishmania species previously [35,36]; preliminary analysis of data arising from shotgun sequence of L. infantum suggest that this species also has homologs of LPG1L, LPG1R, and at least one copy of LPG1G, all of which show >99% identity to the homologous gene of L. major (data not shown).

3.2. Properties of predicted LPG1-family proteins

The predicted open reading frames for LPG1, LPG1L, LPG1R, and LPG1G contained 434, 592, 480, and 599 amino acids, respectively. All four genes (LPG1, LPG1L, LPG1R, and LPG1G) were predicted to encode type-II transmembrane proteins with short, basic cytoplasmic tails (~20 amino acids) at the N-termini, followed by single transmembrane domains (Fig. 2A). Predictions were made using the TMHMM Server at http://www.cbs.dtu.dk/services/TMHMM). Relative to LPG1, there were insertions in LPG1L, (encoding amino acid 43–149), and LPG1G (encoding amino acid 43–149) immediately after the transmembrane domain, that were not seen in LPG1 or LPG1G (Fig. 2A and B). The overall similarity among these four genes was modest (~20% amino acid identity, Fig. 2B). Nonetheless, the predicted luminal domains of all four genes contained several highly conserved regions as shown in the alignment in Fig. 2B. One of these regions (G-F/Y-F/L-D-E-N-F/Y-Y/F-P-A/I-Y/L-G/Y/F/M-E/D-I/T/V/L-D-Y/W/L) is likely to represent the catalytic site, since it contained the metal-binding DXD motif conserved amongst glycosyltransferases (Fig. 2A and B).

In all following studies we focused exclusively on the characterization of LPG1L and LPG1R.

3.3. Detection of LPG1L and LPG1R mRNA in L. major

To test whether LPG1L and LPG1R were expressed in L. major, total RNA was prepared from early log phase, late log phase, stationary phase, metacyclic phase, and amastigote
Fig. 2. LPG1 gene family. (A) Schematic representations of LPG1, LPG1L, LPG1R, and LPG1G. Numbers indicate amino acids. Positions of the conserved DXD catalytic motifs are indicated. CT, cytoplasmic tail; TM, transmembrane domain. Compared to LPG1 and LPG1R, LPG1L and LPG1G contain additional "inserts" (from amino acid 43–161 for LPG1L and 43–149 for LPG1G) after the transmembrane domain, as indicated. Our analysis indicates that the three copies of LPG1G on chromosomes 5, 19, and 32 are identical. (B) Alignment of LPG1, LPG1L, LPG1R, and LPG1G from *L. major* (∼20% identity among these genes at amino acid level) and a preliminary partial open reading frame (contig) from *T. cruzi* (*T. cruzi* ctg). Alignments were performed using the ClustalW algorithm as implemented in the program Megalign (DNA Star). Transmembrane domains were highlighted. The underlined region contains the DXD catalytic motif (marked by asterisks).
Fig. 3. LPG1L and LPG1R are expressed throughout L. major life cycle. Northern-blot analysis was performed as described in Section 2. Probes were made from coding regions of LPG1L and LPG1R. E, early log phase; L, late log phase; S, stationary phase; M, metacyclic phase (prepared by negative selection for binding with peanut agglutinin); A, amastigotes (from infected mouse foot pads). Expression of rRNA was used as controls for loading.

stage cells and subjected to Northern-blot analysis using the coding regions of LPG1L and LPG1R as probes. The 2.8-kb LPG1L and 2.0-kb LPG1R mRNAs were expressed at similar levels throughout parasite life cycle (Fig. 3). RT-PCR analysis mapped the site of trans-splicing sites for LPG1L and LPG1R to positions 129 and 160 nt upstream of the ATG start codon for LPG1L and LPG1R, respectively (data not shown).

3.4. Cellular localization of LPG1L and LPG1R

Previous data showed that an LPG1::GFP fusion protein was localized in the Golgi apparatus in Leishmania [20]. To examine the cellular localization of LPG1L and LPG1R gene products, we created similar fusion constructs joining the green fluorescent protein (GFP) to the C-termini of LPG1L and LPG1R. When expressed in Leishmania from episomal pXG vectors, both LPG1L::GFP and LPG1R::GFP fusion proteins showed strong green fluorescence in a region close to the kinetoplast, the location of the Golgi apparatus. To confirm Golgi localization, we introduced a construct expressing a hemagglutinin (HA)-tagged LPG2 gene, which encodes a Golgi GDP-mannose transporter [22,37]. Immunofluorescence analysis of the LPG2::HA showed co-localization with the LPG1L::GFP and LPG1R::GFP signals (Fig. 4). This result was further confirmed by immuno-EM study using anti-GFP antibody (data not shown). Therefore, similar to LPG1, the gene products of LPG1L and LPG1R were also localized at the Golgi apparatus, consistent with the cellular location of most glycosyltransferases.

3.5. Generation of double and triple LPG1, LPG1L and LPG1R knockouts

Because Leishmania has a diploid genome, two rounds of replacement are required to inactivate any given gene. This presents an obstacle in genetically assessing the function of all six members of the LPG1 gene family, as currently just six suitable markers exist. Here we studied the effects of inactivation of LPG1L and LPG1R, singly, together, or in combination with LPG1. An outline of the gene inactivation protocol is shown in Fig. 5A. Briefly, LPG1L alleles were replaced by HYG and NEO markers as (lpg1l−), while LPG1R alleles were replaced by PAC and SAT markers (as lpg1r−), and all four alleles were replaced in the double lpg1r−/lpg1l− mutant. The triple knockout parasites (lpg1r−/lpg1l−/lpg1−) were generated through a LOH approach [25], by first replacing one LPG1 allele of the lpg1−/lpg1l−/lpg1− mutant with the BSD marker, then selecting for the loss of LPG. In all
Fig. 5. Deletions of LPG1L, LPG1R, and LPG1 in L. major. (A) Procedure to generate the lpg1−/lpg1l−/lpg1r− triple knockout parasites. (B) Southern-blot analysis to confirm the lpg1−/lpg1l−/lpg1r− triple knockout parasites (nos. 4-1 and 5-1 are two independent clones). Genomic DNA was digested with HindIII and hybridized with probes made from the coding regions of LPG1, LPG1L, and LPG1R.

lines, Southern-blot analysis confirmed that the planned replacements had occurred; data for the lpg1−/lpg1l−/lpg1r− triple knockout parasite is shown in Fig. 5B.

The single, double and triple knockout parasites grew at similar rates as L. major wild-type cells in culture (data not shown). Western-blot analysis showed that all lines tested synthesized wild-type levels of PPG and GP63, and that the lpg1l−, lpg1r−, and lpg1l−/lpg1r− mutants synthesized WT levels of LPG (Fig. 6 and data not shown). These data indicated that LPG1L and LPG1R were not involved in the synthesis of large surface glycoconjugates such as LPG, PPG, or the GPI-anchored protein GP63. Additionally, expression of LPG1L and LPG1R from episomal vectors did not restore the synthesis of LPG in the lpg1− mutant, arguing that LPG1L and LPG1R could not substitute for the function of LPG1 even when overexpressed (data not shown).

3.6. Deletion of LPG1L and/or LPG1R did not affect the biosynthesis of GPIs

The effect of LPG1-family mutants on GPI synthesis was evaluated in two ways. First, total GPI fractions from wild type and mutant cells were extracted, dilipidated, fluorophore-labeled, and analyzed by FACE (fluorescent-assisted carbohydrate electrophoresis). GPIs from single (lpg1l−, lpg1r−) or double (lpg1l−/lpg1r−) knockout showed a pattern similar to the GPI pattern seen for wild type (Fig. 7A and data not shown). In contrast, the

Fig. 6. LPG1L and LPG1R are not required for the synthesis of LPG, PPG, or GPI-anchored proteins. Western blots of cell extracts (for LPG, GP63, and α-tubulin) and culture supernatants (for PPG) were performed as described in Section 2. For each cell type (WT, lpg1l−, and lpg1r−), materials from 2.5 × 10⁵, 5 × 10⁵, and 10⁶ cells (from left to right) were loaded.
GIPL pattern for the triple knockout (lpg1<sup>−</sup>/lpg1l<sup>−</sup>/lpg1r<sup>−</sup>) containing the deletion of LPG1 showed a pattern similar to the lpg1<sup>−</sup> knockout alone (Fig. 7A). Accumulation of the expected LPG biosynthetic intermediate (Glc-P-Man<sub>2</sub>-GlcN) was observed in all lines bearing the homozygous lpg1<sup>−</sup> deletions. Second, cells were metabolically labeled with [<sup>3</sup>H]galactose and total GIPLs were extracted and treated with trifluoroacetic acid to cleave the Galα-Man linkage in GIPLs with the structure Galα₂-Galα-Man-Man-GlcN. After separation by paper chromatography, similar profiles were seen in WT and all mutant lines (Fig. 7B and data not shown). Therefore, it is unlikely that any of genes (LPG1, LPG1L, and LPG1R) were involved in the biosynthesis of the Galα-containing type-II GIPLs in <i>L. major</i>.

3.7. LPG1L and LPG1R were not required for metacyclogenesis nor mouse infectivity

The percent metacyclic promastigotes (infectious form) present in stationary phase cultures showed very little difference amongst the wild type and various knockout cells (data not shown). To test if LPG1L and/or LPG1R were implicated in virulence, stationary phase parasites were inoculated in susceptible BABL/c mice. There was no significant difference in lesion formation between wild type and various knockout cells (lpg1l<sup>−</sup>, lpg1r<sup>−</sup>, and lpg1l<sup>−</sup>/lpg1r<sup>−</sup>), arguing that LPG1L and LPG1R were not required for virulence in the mammalian host (Fig. 8).

3.8. Potential role of LPG1L and LPG1R in adding Galα to glycoproteins

Since neither LPG1L nor LPG1R appeared to function in GIPL biosynthesis, we asked whether these genes might participate in other glycoconjugate synthetic pathways. Galα residues are not synthesized by the mammalian hosts, but are found in glycoconjugates synthesized by a number of pathogens including Mycobacterium, <i>T. cruzi</i>, and <i>Aspergillus</i> [38–40]. In <i>Leishmania</i>, Galα residues have only been found in glycolipids [41]. In contrast, other Kinetoplastid species such as <i>T. cruzi</i>, <i>Crithidia fasciculata</i>, <i>Lepptomonas samuelii</i>, and <i>Endotrypanum schaudi</i> contain abundant amounts of Galα residues in their glycoproteins.
LV39 WT

Fig. 8. LPG1L and LPG1R are not required for \textit{L. major} virulence. Mouse footpad infections were performed as described in Section 2. Stationary (WT, lpg1\textsuperscript{−}, lpg1r\textsuperscript{−}, and lpg1l\textsuperscript{−}/lpg1r\textsuperscript{−}) cells (10\textsuperscript{6}) were injected into each mouse. Five to six mice were used in each experiment. Error bars represent standard deviations.

LV39 WT

Fig. 9. Homologues of LPG1L and LPG1R exist in other Kinetoplastida parasites. Genomic DNAs from \textit{L. major} LV39 (lane 1), \textit{Trypanosoma cruzi} (lane 2), \textit{Crithidia fasciculata} (lane 3), \textit{Leptomonas samuli} (lane 4), and \textit{Endotypanum schaudinni} (lane 5) were digested with restriction enzymes and hybridized with probes corresponding to the coding regions of LPG1, LPG1L, and LPG1R. The blots were washed under ‘moderately’ stringent conditions as described in the Section 2 (0.2\times SSPE/0.5% SDS at 60 °C).

often β1, 3-linked to mannose as in the LPG and GIPLs in \textit{L. major} [38,42–44]. Southern-blot analysis was performed on the genomic DNA from the species mentioned above using the coding regions of LPG1, LPG1L, and LPG1R from \textit{L. major} as probes. After several washes at moderate stringy (Section 2), strong signals were seen for several of these LPG1 family members in all species (Fig. 9). Additionally, LPG1-related genes have been identified in the emerging genome sequence of \textit{T. cruzi} (Fig. 2).

Since Gal\textsubscript{f} had never been observed previously in \textit{Leishmania} proteins, we undertook an overexpression approach to determine whether LPG1 and its homologues (LPG1L and LPG1R) had any effect on protein galactosylation. In these studies, we generated a construct which simultaneously expressed LPG1, LPG1L, and LPG1R and integrated it into the ribosomal RNA locus of \textit{L. major}, a position which gives high levels of expression [23]. WT and transfectants were metabolically labeled with \[^{3}H\text{galactose, and incorporation of the radiolabel into glycoproteins was measured. The results showed a small but reproducible increase in protein galactosylation (Fig. 10A). Efforts to assign this activity to either LPG1, LPG1R, or LPG1L in single gene overexpressors were inconclusive, possibly due to the low level of incorporation.

Unlike several other protozoans, the presence of Gal\textsubscript{f} in N-linked carbohydrate chains in \textit{L. mexicana} has been investigated and was not detected [45,46], possibly due to the low abundance of the monosaccharide. To examine whether \textit{L. major} might possess Gal\textsubscript{f}-containing proteins, \[^{3}H\text{-Gal}glycopeptides were prepared by Pronase digestion of glycoproteins from wild-type cells and aliquots were incubated in the presence or absence of exo-β-galactofuranosidase [47], desalted by anion exchange that removes most of the glycopeptides, and analyzed.
creased incorporation of \([^{3}H]\)galactose in glycoproteins. L V39 WT cells residues in \(L.\ major\) Fig. 10. (A) Overexpression of \(LPG1\), \(LPG1L\), and \(LPG1R\) caused increased incorporation of \([^{3}H]\)galactose in glycoproteins. LV39 WT cells and cells overexpressing \(LPG1\), \(LPG1L\), and \(LPG1R\) (triple OE; triple overexpressor) were metabolically labeled with \([^{3}H]\)galactose and glycopeptides were extracted as described in Section 2. Radioactivity in these glycopeptides was determined using a scintillation counter; several experiments were done and the results of one with three replicas is shown; error bars represent standard deviation. (B) Demonstration of terminal Gal \(f\) residues growth, differentiation or the ability to induce disease loss of \(LPG1\) in \(L.\ major\) showed similarly that loss of \(LPG1\) in \(L.\ major\) caused in- netic markers, we focused functional studies on the \(LPG1\) gene, there was no effect on GIPL biosynthesis three mutants were constructed, the first such triple knockout mu- tants reported in \(L.\ major\). In this study we described the \(LPG1\) gene family of \(L.\ major\), and reported detailed characterization of two of the \(LPG1G\) and \(LPG1T\). In being type II transmembrane \(LPG1L\) and \(LPG1R\)-GFP fusion proteins localized to the parasite Golgi apparatus (Fig. 4). Thus, these genes were reasonable candidates to be the GIPL-specific Gal \(f\) transferase postulated previously [9], and their modest degree of overall sequence conservation (~20% identity) provides the requisite opportunity for divergence in Gal \(f\) acceptor specificity.

Due to the late emergence of the three \(LPG1\) gene sequences and the availability of only six selectable ge- netic markers, we focused functional studies on the \(LPG1\), \(LPG1L\), and \(LPG1R\) genes. Single and double mutants of \(LPG1L\) and \(LPG1R\) were made, and a mutant lacking all three was also constructed, the first such triple knockout mu- tant reported in \(Leishmania\) (Fig. 5). These studies showed collectively that \(LPG1L\) and \(LPG1R\) had no effect on parasite growth, differentiation or the ability to induce disease in animals (Fig. 8). Analysis of glycoconjugate synthesis showed similarly that loss of \(LPG1L\) and \(LPG1R\) had no ef- fect on the synthesis of the abundant glycoconjugates GP63, PPG, or LPG (Fig. 6). Lastly, while the triple mutant showed the expected adverse effect on LPG synthesis due to deletion of the \(LPG1\) gene, there was no effect on GIPL biosynthesis (Fig. 7). These data suggest that none of these three genes are essential for the GIPL-specific Gal \(f\)/T. Obviously, the proposal that the \(LPG1\) genes encode this activity is at- tractive, but the possibility that a completely unrelated gene encodes this activity cannot be excluded. It is interesting that

4. Discussion

In this study we described the \(LPG1\) gene family of \(L.\ major\), and reported detailed characterization of two of the genes, \(LPG1L\) and \(LPG1R\). The five new genes predict pro- teins showing strong similarity to the active LPG-specific Gal \(f\)/T encoded by \(LPG1\), in being type II transmembrane glycoproteins with a short N-terminal cytoplasmic domain and a luminal C-terminal domain bearing a canonical glycosyltransferases DXD motif (Fig. 2A). Both \(LPG1R\) and \(LPG1L\) are constitutively expressed in stable mRNA throughout the parasite life cycle (Fig. 3). Furthermore, epi- somally expressed \(LPG1L\) and \(LPG1R\)-GFP fusion proteins localized to the parasite Golgi apparatus (Fig. 4). Thus, these genes were reasonable candidates to be the GIPL-specific Gal \(f\) transferase postulated previously [9], and their modest degree of overall sequence conservation (~20% identity) provides the requisite opportunity for divergence in Gal \(f\) acceptor specificity.

Percentage of incorporation relative to WT

<table>
<thead>
<tr>
<th>WT</th>
<th>triple OE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>50</td>
<td>1500</td>
</tr>
<tr>
<td>10</td>
<td>2000</td>
</tr>
<tr>
<td>15</td>
<td>2500</td>
</tr>
<tr>
<td>20</td>
<td>3000</td>
</tr>
<tr>
<td>25</td>
<td>3500</td>
</tr>
</tbody>
</table>

Radioactivity in these glycopeptides was determined using a scintillation counter; several experiments were done and the results of one with three replicas is shown; error bars represent standard deviation. (B) Demonstration of terminal Gal \(f\) residues growth, differentiation or the ability to induce disease loss of \(LPG1\) in \(L.\ major\) showed similarly that loss of \(LPG1\) in \(L.\ major\) caused in- netic markers, we focused functional studies on the \(LPG1\) gene, there was no effect on GIPL biosynthesis three mutants were constructed, the first such triple knockout mu- tants reported in \(Leishmania\) (Fig. 5). These studies showed collectively that \(LPG1L\) and \(LPG1R\) had no effect on parasite growth, differentiation or the ability to induce disease in animals (Fig. 8). Analysis of glycoconjugate synthesis showed similarly that loss of \(LPG1L\) and \(LPG1R\) had no ef- fect on the synthesis of the abundant glycoconjugates GP63, PPG, or LPG (Fig. 6). Lastly, while the triple mutant showed the expected adverse effect on LPG synthesis due to deletion of the \(LPG1\) gene, there was no effect on GIPL biosynthesis (Fig. 7). These data suggest that none of these three genes are essential for the GIPL-specific Gal \(f\)/T. Obviously, the proposal that the \(LPG1\) genes encode this activity is at- tractive, but the possibility that a completely unrelated gene encodes this activity cannot be excluded. It is interesting that
L. infantum appears to contain an LPG1G homolog, while the GPIL composition of this species isolate has not been investigated, its close relative L. donovani lacks Gal$_f$-containing GPILs [11]. While we were unable to test the effect of LPG1G knockouts in these studies, our repertoire of six selectable markers will be sufficient to permit the eventual inactivation of all three LPG1G genes simultaneously in the future.

In L. major Gal$_f$-containing glycolipids such as LPG- and GPIs-, whereas in other trypanosomatid species other glycoconjugates contain Gal$_f$, similarly linked to Man [38,42–44]. This suggested that perhaps one or more of the LPG1 family members participate in Gal$_f$/T reactions involving other acceptors including the carbohydrate chains of proteins. Although Parodi and coworkers reported that L. mexicana lacks Gal$_f$-containing glycoproteins [45,46], a low abundance of this substituent was not excluded. Consistent with the possibility of Gal$_f$ in Leishmania glycoproteins, several species known to express Gal$_f$-containing proteins possess genes homologous to LPG1, LPG1L, and LPG1R (Fig. 9), and the T. cruzi genome project also has revealed a family of LPG1-homologs (http://www.tigr.org/tdb/e2k1/tca1/). Significantly, simultaneous overexpression of LPG1, LPG1R, and LPG1L resulted in increased incorporation of galactose into L. major proteins (Fig. 10), at least some of which was in the form of terminal Gal$_f$. Although we were unable to assign this function to a single gene, it is likely that one or more of the LPG1-family genes encode proteins with protein-glycoconjugate Gal$_f$/T activity. This would be especially relevant to studies of T. cruzi, where the abundant mucins are highly modified by Gal$_f$ and are thought to contribute to parasite virulence [48,49].

One scenario consistent with our data is that Leishmania LPG1L and LPG1R may be cryptic, relatively “silent counterparts” of genes highly active in other trypanosomatids. This further suggests that the potential genomic glyco-synthetic repertoire of these parasites may be larger than revealed by biochemical assays, which typically focus on the most abundant molecules or subsets thereof. A similar finding was reached in studies of the SCG family of LPG side chain galactosyltransferases, as only a subset of the seven genes of this family have shown significant Gal$_f$/T activity [6]. The presence of a number of silent but potentially active glycosyltransferases genes is intriguing and raises a number of speculative possibilities. Potentially, these genes represent a source for the emergence of novel or modified glycoconjugate synthases during evolution, which may contribute in any number of ways to parasite virulence and host range. In protozoa such as Plasmodium or Trypanosoma variant surface protein(s) represent an interesting analogy, as gene families in these organisms comprise a group of cryptic genes that are expressed in various ways amongst strains and species and contribute to pathology [50–52]. In these species many such genes are located at telomeres and it is interesting to note that the three LPG1G genes appear to be located at telomeres in L. major and L. infantum (unpublished data; Leishmania Genome Project). Presumably during the evolutionary divergence of trypanosomatids, different lineages have chosen to emphasize the synthesis of particular glycoconjugates relevant to their biological niche within the mammalian and insect hosts. This raises the possibility that this cryptic glyco-synthetic capability may be reactivated and expressed at higher levels in the future, with consequences best understood at present by the parasite.

Acknowledgements

We thank A. Capul, D. Dobson, and K. Robinson for comments on the manuscript, M. Cunningham for pIR1PHLEO, W. Beatty for immuno-EM studies, everyone in our laboratories for helpful discussions, the Leishmania and T. cruzi Genome Projects for generously making available preliminary partial sequence of LPG1L, LPG1R, and the sequences LPG1G, and Maria Julia Manso Alves for the gift of exo-β-galactofuranosidase. This work was supported by NIH grant AI 31078. Sequencing of the L. major genome was accomplished as part of the Leishmania Genome Network with support by The Wellcome Trust. Sequencing of T. cruzi was accomplished by TIGR with support from the NIAID.

References

[9] Spath GF, Garraway LA, Turco SJ, Beverley SM. The role(s) of lipophosphoglycan (LPG) in the establishment of Leishmania major infections in mammalian hosts. Proc Natl Acad Sci USA 2003;100:9536–41.


Ilg T. Lipophosphoglycan is not required for infection of macrophages or mice by Leishmania mexicana. EMBO J 2000;19:1953–62.


