2, A and B). On the basis of a $K_d$ for nonspecific DNA binding in the $10^{-4}$ to $10^{-6}$ M range (*24*), the protein occupancy was estimated to be 36% for ΔN331 and 27% for ΔN280 (Fig. 2, C and D).
18. DNA-only spectra of mixed DNA-protein samples containing specific DNA and either ΔN331 or ΔN280 show a difference in molar ellipticity between 250 and 295 nm as compared to the spectra of free DNA, which suggests a protein-induced change in the DNA. This difference was not observed in ΔN280 or ΔN331 samples mixed with nonspecific DNA (Fig. 2, C and D) (J. M. Petersen and B. J. Graves, unpublished results).
19. Specific DNA-only spectra in the presence of ΔN331 and ΔN280 were compared between 250 and 295 nm because ΔN331 and ΔN280 make no important contributions to their CD spectra in this region (Fig. 2, A through D). In contrast, specific DNA shows a strong CD signal in this region (Fig. 2E).
20. The $NH_2$-terminal deletion mutant ΔN336 and full-length Ets-1 show identical DNA contacts in protection and interference assays (*8*). DNA-binding studies of ΔN331 and ΔN280, including deoxyribonuclease I footprinting and binding site size selection, also indicate identical contact regions (Q. Xu and J. M. Petersen, unpublished results).
21. A time course of protease digestion was performed with fixed concentrations of trypsin (115 ng) or chymotrypsin (215 ng) in a buffer containing 50 mM $KPO_4$ (pH 7.9), 65 mM KCl, and 0.1 mM EDTA. Proteases were inactivated with 50 mM PMSF, 2% SDS, and 100 mM DTT. Polypeptides were fractionated on Tricine-SDS polyacrylamide gels [H. Schagger and G. von Jagow, *Anal. Biochem.* **166**, 368 (1987)] and visualized by Coomassie blue staining. Stained bands were quantified by video capture and computer-based densitometric analysis with NIH Image software (version 1.49). The concentrations of specific DNA (*16*) and protein that were sufficient to obtain 99% protein occupancy were 10 μM for reactions containing ΔN280 and 15 μM for reactions containing ΔN331. Nonspecific DNA (*23*) was used at equimolar concentrations relative to specific DNA. On the basis of a $K_d$ for nonspecific DNA binding in the $10^{-4}$ to $10^{-6}$ M range (*24*), the protein occupancy was calculated to be 46% for ΔN331 and 35% for ΔN280.
22. Fragments were electroblotted from Tricine-SDS polyacrylamide gels to polyvinylidenefluoride membranes and sequenced by Edman degradation with an Applied Biosystems automated sequencer.
23. A 21-bp synthetic DNA duplex (top strand, 5′-AGGCCTGATTGCCCAATTGTC-3′) was used as the nonspecific control DNA. This duplex lacks the core 5′-GGA-3′ motif required for Ets-1 sequence-specific DNA binding.
24. P. H. von Hippel, A. Revzin, C. A. Gross, A. C. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4808 (1974).
25. J. M. Petersen and B. J. Graves, data not shown.
26. J. J. Skalicky, J. M. Petersen, L. W. Donaldson, B. J. Graves, L. P. McIntosh, in preparation.
27. The interaction between the $NH_2$-terminal and COOH-terminal inhibitory regions is supported by the finding that the pattern of protease digestion within the $NH_2$-terminal inhibitory region is altered in the absence of the COOH-terminal inhibitory region (*13*).
28. J. A. Lefstin *et al.*, *Genes Dev.* **8**, 2842 (1994); R. V. Talanian *et al.*, *Science* **249**, 769 (1990); S. Tan and T. J. Richmond, *Cell* **62**, 367 (1990).
29. M. A. Weiss *et al.*, *Nature* **347**, 575 (1990); V. Saudek *et al.*, *Biochemistry* **30**, 1310 (1991); T. E. Ellenberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **71**, 1223 (1992); J. D. Shuman, C. R. Vinson, S. L. McKnight, *Science* **249**, 771 (1990); D. Zhao, C. H. Arrowsmith, X. Jia, O. Jardetzky, *J. Mol. Biol.* **229**, 735 (1993); Z. Otwinowski *et al.*, *Nature* **335**, 321 (1988); N. D. Clarke, L. J. Beamer, H. R. Goldberg, C. Berkower, C. O. Pabo, *Science* **254**, 267 (1991).
30. R. S. Spolar and M. T. Record Jr., *Science* **263**, 777 (1994).
31. Because the wHTH motif of Ets-1 does not change secondary structure upon DNA binding (*5*), we predict that the change in heat capacity ($\Delta C°_{assoc}$) that accompanies Ets-1 DNA binding will be affected by the unfolding of the α helix (residues 303 to 310). We propose that $\Delta C°_{assoc}$ will be smaller in magnitude than the values observed for protein-DNA interactions that display a substantial induced fit or rigid body association (*30*).
32. B. E. Kemp and R. B. Pearson, *Biochim. Biophys. Acta* **1094**, 67 (1991); T. R. Soderling, *J. Biol. Chem.* **265**, 1823 (1990); S. Mukherji, D. A. Brickey, T. R. Soderling, *ibid.* **269**, 20733 (1994).
33. M. Newman, T. S. Strzelecka, L. F. Dorner, I. Schildkraut, A. K. Aggarwal, *Science* **269**, 656 (1995).
34. We thank D. King for mass spectrometry analyses, use of equipment, and helpful discussions; B. Schackmann for peptide sequencing; and W. Sundquist and H. Nelson for critical comments on the manuscript. Supported by grants from NIH (GM 38663 to B.J.G.; CA 42014 to the Utah Cancer Center; GM 48958 to T.A.), and the National Cancer Institute of Canada, with funds from the Canadian Cancer Society (to L.P.M.).

31 May 1995; accepted 14 August 1995

# A Specialized Pathway Affecting Virulence Glycoconjugates of *Leishmania*

Albert Descoteaux,* Ya Luo, Salvatore J. Turco, Stephen M. Beverley†

For virulence and transmission, the protozoan parasite *Leishmania* must assemble a complex glycolipid on the cell surface, the lipophosphoglycan (LPG). Functional complementation identified the gene *LPG2*, which encodes an integral Golgi membrane protein implicated in intracellular compartmentalization of LPG biosynthesis. *lpg2*$^-$ mutants lack only characteristic disaccharide-phosphate repeats, normally present on both LPG and other surface or secreted molecules considered critical for infectivity. In contrast, a related yeast gene, *VAN2/VRG4*, is essential and required for general Golgi function. These results suggest that *LPG2* participates in a specialized virulence pathway, which may offer an attractive target for chemotherapy.

In *Leishmania*, developmentally regulated glycoconjugates are essential during the infectious cycle. LPG, the major cell surface glycoconjugate of promastigotes, is a multifunctional virulence determinant required for development within the fly, resistance to complement in the bloodstream, and establishment of the infection within the macrophage (*1*, *2*). In *L. donovani*, LPG consists of a phosphoglycan polymer of repeating disaccharide-$PO_4$ units [Gal(β1-4)Man(α1-$PO_4$-6)] (where Gal is galactose and Man is mannose) attached by means of a glycan core to an unusual phosphatidylinositol (PI)-lipid anchor (*2*). These components occur on other parasite molecules (sometimes with further modifications), and their expression can be constitutive or specific to a particular stage (*2*–*5*).

LPG is experimentally attractive because its biosynthesis can be studied in vitro and *lpg*$^-$ mutants can be readily generated (*6*–*8*). We recently demonstrated the feasibility of functional genetic complementation in *Leishmania* by identification of *LPG1*, which mediates addition of the $Gal_f$ moiety (where "f" indicates furanose) within the LPG glycan core (*9*, *10*). *LPG1* thus represents a class of genes (class I) encoding LPG biosynthetic enzymes. Here, we report a second class of LPG mutants that affect compartmentalization and LPG assembly. These genes are functionally distinct from those affecting protein secretion and glycosylation.

The mutant C3P0 synthesizes a truncated LPG containing only the glycan core and lipid anchor and lacking repeating units (*8*, *11*). A wild-type *L. donovani* cosmid library in the shuttle vector cLHYG was introduced into C3P0, and LPG$^+$ transfectants were recovered by antibody panning (*9*, *10*, *12*). From these, four cosmids

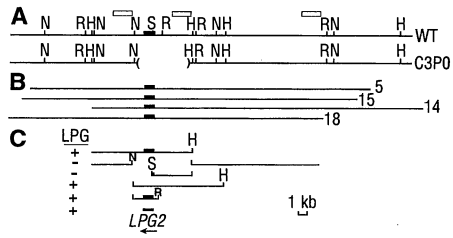


**Fig. 1.** Functional and physical mapping of *LPG2*. (**A**) Restriction map of the *LPG2* locus. C3P0 contained a deletion (indicated by parentheses) of 6.0 kb containing *LPG2* (black box). Repeating sequences are indicated by shaded boxes. WT, wild type. (**B**) Position of the four overlapping cosmids recovered from LPG$^+$ C3P0 transfectants. (**C**) Localization of *LPG2* by deletional analyses. The indicated restriction fragments were tested for LPG expression after cloning into appropriate *Leishmania* vectors (*13*). N, Not I; R, Eco RV; H, Hind III; S, Spe I.

A. Descoteaux and S. M. Beverley, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA.
Y. Luo and S. J. Turco, Department of Biochemistry, University of Kentucky Medical Center, 800 Rose Street, Lexington, KY 45036, USA.
*Present address: McGill Centre for the Study of Host Resistance, Montreal General Hospital, 1650 Cedar Avenue, Montreal, Quebec H3G 1A4, Canada.
†To whom correspondence should be addressed.

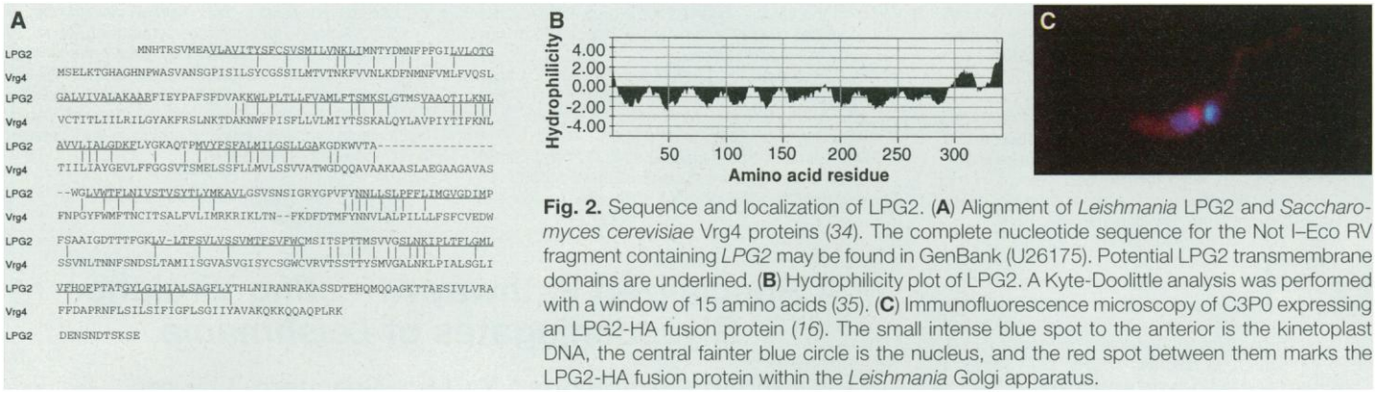


**Fig. 2.** Sequence and localization of LPG2. (**A**) Alignment of *Leishmania* LPG2 and *Saccharomyces cerevisiae* Vrg4 proteins (*34*). The complete nucleotide sequence for the Not I–Eco RV fragment containing *LPG2* may be found in GenBank (U26175). Potential LPG2 transmembrane domains are underlined. (**B**) Hydrophilicity plot of LPG2. A Kyte-Doolittle analysis was performed with a window of 15 amino acids (*35*). (**C**) Immunofluorescence microscopy of C3P0 expressing an LPG2-HA fusion protein (*16*). The small intense blue spot to the anterior is the kinetoplast DNA, the central fainter blue circle is the nucleus, and the red spot between them marks the LPG2-HA fusion protein within the *Leishmania* Golgi apparatus.

with overlapping inserts were recovered, and the functionally active gene was mapped to a single 3-kb Not I–Eco RV fragment (Fig. 1) (*13*). The sequence revealed a single open reading frame (ORF; termed *LPG2*) encoding a hypothetical protein of 341 amino acids (Figs. 1 and 2A). LPG2 was predicted to be hydrophobic and to possess up to 10 transmembrane domains (Fig. 2B). Sequence database searches showed that LPG2 was related to three yeast proteins (*14*). Of these, Vrg4/Van2 showed 28% identity (55% similarity) (Fig. 2) and has recently been found to be essential and required for correct protein N-glycosylation and maintenance of Golgi function and structure (*15*). Consistent with this, an active LPG2 protein tagged with an influenza hemagglutinin epitope at its COOH-terminus was localized to the *Leishmania* Golgi apparatus (Fig. 2C) (*16*).

Several lines of evidence implicated *LPG2* in the *lpg*$^-$ defect of the C3P0 mutant. First, transfection of an *LPG2* expression construct fully restored LPG levels to those present in the wild type (*17*). Second, molecular karyotype analysis showed that the *LPG2* locus was deleted in C3P0 (Fig. 3A). Southern (DNA) blot analysis revealed that C3P0 contained a homozygous deletion of about 6 kb encompassing *LPG2* (Fig. 1) (*18*). These studies showed that *LPG2* is flanked by repeated DNAs (Fig. 1) that may have served as the site of rearrangement (*18*). Although C3P0 was obtained after nitrosoguanidine mutagenesis, classic point mutagens can also induce DNA rearrangements. *Leishmania* are thought to be diploid, which was confirmed for *LPG2* specifically (*19*). The homozygous *LPG2* deletion in C3P0 may be the result of a two-step process, whereby an initial heterozygous deletion was rendered homozygous. Finally, *lpg2*$^-$ null mutants obtained by homologous gene replacement (*19*) lacked LPG and exhibited properties identical to those of the C3P0 mutant. *LPG2* is thus not required for viability.

Northern (RNA) blot analysis revealed a single, 2.8-kb *LPG2* mRNA in wild-type but not C3P0 promastigotes (Fig. 3B). This mRNA was abundant in log phase promastigotes and down-regulated in stationary phase promastigotes and lesion amastigotes. Amastigotes exhibited reduced amounts of the 2.8-kb mRNA, as well as larger forms ranging up to 9 kb, which probably represent unprocessed precursors. *LPG2* mRNA amounts were correlated with the developmental expression of molecules such as LPG and the secretory acid phosphatase (sAP), which bear the repeating units (*2–5*).

The accumulation of a truncated LPG in C3P0 suggested a defect in the addition of the first mannose-phosphate residue (*11*). However, C3P0 microsomal membranes catalyzed the addition of repeating units onto endogenous LPG glycan core acceptors. C3P0 LPG synthesis was 65% of that in the wild type in the presence of detergent and was 11% that of the wild type in its absence (*20*). In contrast, microsomes from the R2D2 mutant, which is defective in the synthesis of the LPG core, synthesized only background levels of LPG (*6*, *9*, *21*). Most importantly, the labeled product made by C3P0 microsomes was the same size as LPG synthesized in vitro from wild-type extracts, and mild acid hydrolysis released repeating units labeled with $^{14}$C-labeled Man (*21*). Thus, the enzymatic machinery for LPG synthesis is present in C3P0, which indicates that LPG2 must affect the localization or compartmentalization of a key LPG biosynthetic precursor or enzyme.

Previous studies have shown that C3P0 completely lacks repeating units (*8*). For sAP, this is shown by altered electrophoretic mobility in nondenaturing gels (Fig. 4A) and its failure to react with an antibody to the repeating unit, CA7AE (Fig. 4B). Both electrophoretic mobility and repeating unit modification returned to levels found in the wild type when *LPG2* expression was restored in C3P0 (Fig. 4). Addition of repeating units to sAP is thought to occur in the Golgi apparatus, which is consistent with the localization of LPG2 (Fig. 2C) (*3*). The normal activity of the C3P0 sAP implies that it does not require repeating unit modifications for either secretion or catalysis (*3*, *8*, *22*).

Otherwise, the rate of growth, and the

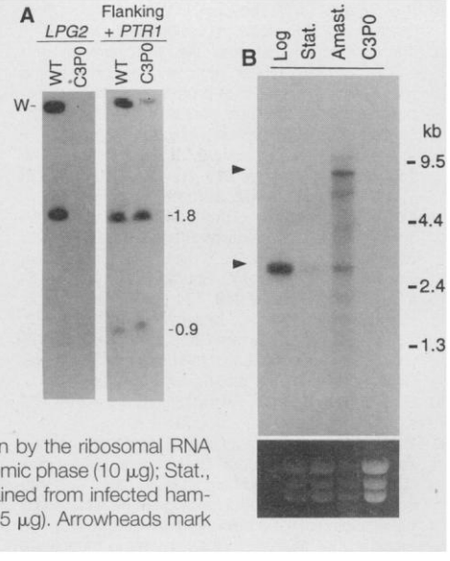


**Fig. 3.** Chromosomal mapping and expression of *LPG2*. (**A**) Molecular karyotype analysis of *LPG2*. Chromosomes were separated by pulse-field electrophoresis (Bio-Rad CHEF Mapper). The 1.8-Mb *LPG2* chromosome was identified by Southern hybridization with an *LPG2*-specific, $^{32}$P-labeled Spe I–Sph I fragment of 0.6 kb (left). The presence of the 1.8-Mb chromosome in C3P0 was determined by Southern hybridization with a $^{32}$P-labeled Bam HI–Eco RI fragment of 2.2 kb (flanking) located 3.0 kb 3′ of the *LPG2* coding region. To ensure equal loading of DNA, we simultaneously probed the membrane with a $^{32}$P-labeled *PTR1*-coding (*36*) region probe (right). W, wells. (**B**) Developmental regulation of *LPG2* mRNA. Total *L. donovani* RNAs were subjected to Northern blot analysis with the use of a $^{32}$P-labeled, 0.6-kb Spe I–Sph I fragment located within the *LPG2* protein-coding region (top). The integrity and loading of the RNA is shown by the ribosomal RNA stained with ethidium bromide (bottom). Log, logarithmic phase (10 μg); Stat., stationary phase (10 μg); Amast., amastigotes obtained from infected hamsters (10 μg); and C3P0, logarithmic phase C3P0 (15 μg). Arrowheads mark the 2.8- and 9-kb RNAs.

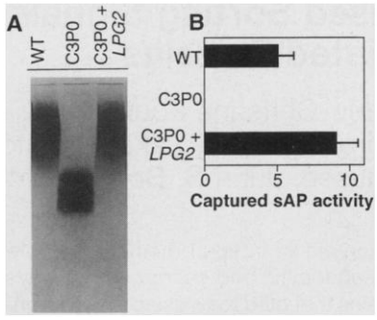


**Fig. 4.** *LPG2* restores repeating unit modification on C3P0 sAP. (**A**) Electrophoretic pattern of sAP. Nondenaturing electrophoresis of concentrated (8×) culture supernatants from the wild type (lane 1), C3P0 (lane 2), and C3P0 transfected with pX63HYG-LPG2 (lane 3) was performed in 7% acrylamide gels in the absence of SDS. The gel was stained for acid phosphatase activity (*37*). (**B**) Repeating unit modification of sAP. An antigen capture assay was performed with the CA7AE monoclonal antibody to repeating units (*38*). Culture supernatants (100 μl) from the wild type, C3P0, and C3P0 transfected with pX63HYG-LPG2 were tested.

expression or modifications of other membrane proteins and glycolipids, were unaffected in the *lpg2*$^-$ mutants. The levels and pattern of total membrane proteins labeled with $^3$H-Man were similar in C3P0 and wild-type cells (*8*). Protein immunoblot analysis of the glycosylated abundant protease gp63 from the surface membrane revealed only fully modified gp63 in both C3P0 and *lpg2*$^-$ knockout mutants (*23*). Similarly, sAP activity in the *lpg2*$^-$ mutants implied that sAP N-glycosylation was unaffected, as this is known to be required for activity (*23*). C3P0 contained normal amounts of a heterogeneous group of smaller glycoinositolphospholipids (GIPLs), which include the biosynthetic precursors of the GPI (glycosylphosphatidylinositol) lipid anchors of LPG and proteins (*2*, *11*). Staining with the membrane dye DiOC6 (*15*) showed only normal morphology.

Although *Leishmania LPG2* and yeast *VAN2/VRG4* are clearly related, their role or roles must be different. *VAN2/VRG4* is essential and required for general Golgi function and structure, whereas *lpg2*$^-$ cells lack only repeating units and are otherwise normal. This suggests that an ancestral gene has functionally diverged and specialized for LPG-related modifications in *Leishmania*. Consistent with this proposal, attempts to cross-complement *Leishmania lpg2*$^-$ mutants with *VRG4* and yeast *vrg4* mutants with *LPG2* have been unsuccessful (*24*). Alternatively, the different mutant phenotypes may reflect more stringent demands placed on a common, shared pathway in yeast than those in *Leishmania*, which differ in numerous aspects of protein secretion and glycolipid metabolism.

The specific molecular step or steps carried out by LPG2 are under investigation. Most probably, LPG2 mediates transport of an essential LPG precursor or biosynthetic enzyme into the secretory network. There are several proteins closely associated with LPG, such as the B protein and the strongly immunogenic protein KMP-11, both of which lack typical $NH_2$-terminal signal peptides (*25*). Because putatively LPG-deficient lines show alterations in the surface localization of the B protein (*26*), there could be an LPG-coupled pathway for translocation of proteins into the secretory network involving LPG2.

The pleiotropic roles of LPG2 have implications for genetic studies of the role of LPG in the *Leishmania* infectious cycle. In both *L. donovani* and *L. major*, LPG mutants show alterations in virulence, as assessed by survival in host macrophages and modulation of the immune response (*27*). Because genes such as *LPG2* affect multiple cellular components simultaneously, some caution must be exercised when attributing the phenotype of LPG mutants to specific molecules.

The properties of *LPG2* point to the existence of a previously unknown class (II) of *Leishmania* biosynthetic genes, a class that is specifically devoted to the correct targeting and assembly of LPG and related molecules implicated in parasite virulence. At least some portion of this pathway is distinct from that mediating the assembly, modification, and targeting of secreted proteins and smaller glycolipids. Preliminary studies of other *L. donovani* mutants suggest the existence of other class II genes that primarily affect LPG assembly and targeting (*28*). *Leishmania* thus resembles prokaryotes in possessing pathways focused on the targeting of complex surface polysaccharides involved in virulence (*29*), whose dissection promises to provide a rich ground for basic cellular research. As in other pathogens, this *Leishmania*-specific pathway may well offer an attractive target for chemotherapeutic intervention.

## REFERENCES AND NOTES


1. D. Sacks *et al.*, *Parasitology* **108**, S55 (1994).
2. S. J. Turco and A. Descoteaux, *Annu. Rev. Microbiol.* **46**, 65 (1992); M. McConville and M. A. J. Ferguson, *Biochem. J.* **294**, 305 (1993).
3. P. A. Bates, I. Hermes, D. M. Dwyer, *Mol. Biochem. Parasitol.* **39**, 247 (1990).
4. T. Ilg *et al.*, *Parasitology* **108**, S63 (1994).
5. T. Ilg *et al.*, *J. Cell Sci.* **99**, 175 (1991); C. L. Jaffe *et al.*, *J. Immunol.* **144**, 699 (1990); P. A. Bates *et al.*, *Exp. Parasitol.* **68**, 335 (1989).
6. M. A. Carver and S. J. Turco, *J. Biol. Chem.* **266**, 10974 (1991); C. Huang, *ibid.* **268**, 24060 (1993); M. A. Carver and S. J. Turco, *Arch. Biochem. Biophys.* **295**, 309 (1992).
7. D. L. King and S. J. Turco, *Mol. Biochem. Parasitol.* **28**, 285 (1988); K. Ng, E. Handman, A. Bacic, *Glycobiology* **4**, 845 (1994).
8. T. B. McNeely *et al.*, *Glycobiology* **1**, 63 (1990).
9. K. A. Ryan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8609 (1993).
10. A. Descoteaux *et al.*, in *Methods in Molecular Genetics*, K. W. Adolph, Ed. (Academic Press, New York, 1993), vol.1, pp. 22–48.
11. Preliminary studies show that C3P0 synthesizes a truncated LPG that contains only the glycan core Gal-Gal-Gal$_f$-[Glc-PO$_4$]-Man-Man-GlcNAc-PI (where Glc is glucose and Ac is acetate) (A. Descoteaux, Y. Luo, S. J. Turco, S. M. Beverley, unpublished results) (*2*). This truncated LPG disappears when *LPG2* expression is restored by transfection.
12. Independent transfectants (12,000) were obtained, providing about 10-fold coverage of the *Leishmania* genome. LPG$^+$ transfectants were recovered by two rounds of LPG panning with monoclonal antibody CA7AE (*10*).
13. Deletions shown in Fig. 1C were prepared either by partial digestion of cosmids cC3P0-14 and cC3P0-18 with Hind III or Not I or by cloning of specific fragments into the *Leishmania* shuttle vector pSNBR (*30*). Each construct was transfected into C3P0 and tested for LPG expression by agglutination (*10*).
14. BLAST database [S. F. Altschul, W. Gish, W. Miller, E. W. Meyers, D. J. Lipman, *J. Mol Biol.* **215**, 403 (1990)] searches yielded three *Saccharomyces cerevisiae* proteins: Van2/Vrg4 ($P = 6.9 \times 10^{-42}$) and two hypothetical proteins, YEM9 and YEA4 ($P = 3.1 \times 10^{-32}$ and $1.5 \times 10^{-3}$, respectively).
15. C. Kanik-Ennulat, E. Montalvo, N. Neff, *Genetics* **140**, 933 (1995); J. B. Poster and N. Dean, unpublished results.
16. LPG bearing the influenza hemagglutin (HA) epitope at its COOH-terminus was prepared by polymerase chain reaction (PCR) with the use of the oligonucleotides SMB-25 (5′-gcgggatcccatATGAACCATACTCGCTC-3′) and SMB-97 [gcgggatcc(ctacgcgtagtccgggacgtcgtacgggtagcc)CTCAGATTTGGAAGT]. Lowercase letters mark bases not coded by *Leishmania*, and parentheses in SMB-97 mark the HA epitope. The expected 1074–base pair product was digested with Bam HI, inserted into the Bgl II site of pX63HYG (*31*), and transfected into C3P0. LPG synthesis was shown by agglutination with CA7AE (*10*). Transfected cells were air-dried on a glass slide and fixed in 4% paraformaldehyde for 30 min at 4°C. After permeabilization with acetone (−20°C), the HA epitope was immunolocalized with the mouse monoclonal antibody 12CA5 (Boehringer Mannheim) and Texas red–conjugated antibody to mouse immunglobulin. The slide was also stained with DAPI to visualize the kinetoplast and nuclear DNA. Golgi localization of LPG2 is consistent with the report that monensin perturbs both Golgi function and the addition of the disaccharide phosphate repeating units to sAP (*3*).
17. The *LPG2* ORF was amplified by PCR with the oligonucleotides SMB-25 (*16*) and SMB-26 (5′-gcgggatccCTACTCAGATTTGGAAGTG-3′), digested with Bam HI, and inserted into the Bgl II site of the *Leishmania* expression vector pX63HYG. C3P0 transfectants were assayed for LPG expression as described (*10*). Surface localization was determined by immunofluorescence with the CA7AE monoclonal antibody (*32*).
18. Repeating elements (Fig. 1A) were mapped by Southern blot analysis of wild-type and C3P0 genomic DNAs with probes surrounding *LPG2* (not shown in the figure). The gap indicated for C3P0 marks only the absence of *LPG2* sequences; the rearrangement is more complex than a simple deletion and possibly involves duplication of adjacent sequences or insertion of non-adjacent DNA.
19. The 5.9-kb Eco RI–Xba I fragment containing the *LPG2* ORF was cloned into a pUC vector. The 0.6-kb Sph I–Spe I fragment from the *LPG2* ORF was replaced by a 2.0-kb Sal I–Bam HI fragment from pX63HYG containing the hygromycin resistance gene *HYG* and an upstream splice acceptor site. This construct was linearized with Not I and Xba I and transfected into the wild-type *L. donovani* 1S Ld4 line. Southern blot analysis of seven colonies showed that all possessed heterozygous *LPG2* replacements. For homozygous replacements, we used a loss-of-heterozygosity protocol (*33*) (F. Gueiros-Filho and S. M. Beverley, unpub-

lished results). Briefly, a heterozygote was grown in 50 μg/ml of hygromycin for 10 passages, and LPG+ cells were removed by agglutination with monoclonal antibody CA7AE. Several clonal lines containing homozygous *HYG/HYG* replacements were recovered (termed *lpg2*− knockouts).
20. In vitro LPG synthesis was assayed with uridine diphosphate–Gal, guanosine diphosphate–$^{14}$C-labeled Man, and *Leishmania* microsomal membranes as described (*6*). In the presence of 0.1% Triton X-100, 1580 ± 220 cpm was incorporated with the wild type and 950 ± 350 cpm with C3P0 (65%). Incorporation (1.3%) was observed previously with R2D2 (*6*). Without detergent, 4760 ± 660 cpm was incorporated with the wild type and 500 ± 80 cpm (11%) with C3P0.
21. S. J. Turco, M. A. Wilkerson, D. R. Clawson, *J. Biol. Chem.* **259**, 3883 (1984).
22. J. K. Lovelace and M. Gottlieb, *Mol. Biochem. Parasitol.* **22**, 19 (1987).
23. Protein immunoblot analyses with the mouse monoclonal antibody to gp63, 235, showed identical mobilities for gp63 from the wild type, C3P0, and the *lpg2*− knockout, all of which which were clearly distinct from nonglycosylated forms obtained after tunicamycin treatment.
24. A. Descoteaux, N. Dean, S. M. Beverley, unpublished data.
25. A. Jardim, D. L. Tolson, S. J. Turco, T. W. Pearson, R. W. Olafson, *J. Immunol.* **147**, 3538 (1991); A. Jardim *et al.*, *Biochem. J.* **305**, 305 (1995); H. M. Flinn, D. Rangarajan, D. F. Smith, *Mol. Biochem. Parasitol.* **65**, 259 (1994).
26. D. F. Smith and D. Rangarajan, *Glycobiology* **5**, 161 (1995).
27. E. Handman, L. F. Schnur, T. W. Spithill, G. F. Mitchell, *J. Immunol.* **137**, 3608 (1986); T. B. McNeely and S. J. Turco, *ibid.* **144**, 2475 (1990); M. Elhay *et al.*, *Mol. Biochem. Parasitol.* **40**, 255 (1990); S. H. Reiner, S. Zheng, Z. F. Wang, L. Stowring, R. M. Locksley, *J. Exp. Med.* **179**, 447 (1994); R. Cappai *et al.*, *Parasitology* **108**, 397 (1994).
28. A. Descoteaux, S. J. Turco, S. M. Beverley, unpublished data.
29. C. Whitfield and M. A. Valvano, *Adv. Microb. Physiol.* **35**, 135 (1993); P. Reeves, in *Bacterial Cell Wall*, J.-M. Ghuysen and R. Hakenbeck, Eds. (Elsevier, New York, 1994), pp. 281–317.
30. H. L. Callahan and S. M. Beverley, *J. Biol. Chem.* **266**, 18427 (1991).
31. A. Cruz, C. M. Coburn, S. M. Beverley, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7170 (1991).
32. D. L. Tolson, S. J. Turco, R. P. Beecroft, T. W. Pearson, *Mol. Biochem. Parasitol.* **35**, 109 (1989).
33. R. M. Mortensen, D. A. Conner, S. Chao, A. A. T. Geisterfer-Lowrance, J. G. Seidman, *Mol. Cell. Biol.* **12**, 2391 (1992).
34. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
35. J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
36. A. Bello, B. Nare, D. Freedman, L. Hardy, S. M. Beverley, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11442 (1994).
37. U. K. Laemmli, *Nature* **227**, 680 (1970); K. Katakura and A. Kobayashi, *Infect. Immun.* **56**, 2856 (1988).
38. T. Ilg, D. Harbecke, M. Wiese, P. Overath, *Eur. J. Biochem.* **217**, 603 (1993).
39. We thank H. Flinn for assistance, D. Russell for advice on cellular localization, N. Dean for extensive advice and communication of submitted manuscripts, D. Dwyer for discussions, and S. Cilmi, D. Dobson, F. Gueiros-Filho, and S. Singer for comments. The mouse monoclonal antibody to gp63, 235, was provided by D. Russell and W. R. McMaster. This work was supported by NIH grant AI31078. A.D. is the recipient of a Centennial Fellowship from the Medical Research Council of Canada, and S.M.B. and S.J.T. are Burroughs-Wellcome Scholars in Molecular Parasitology.

23 May 1995; accepted 8 August 1995

# Interaction of Tyrosine-Based Sorting Signals with Clathrin-Associated Proteins

Hiroshi Ohno, Jay Stewart, Marie-Christine Fournier, Herbert Bosshart, Ina Rhee, Shoichiro Miyatake, Takashi Saito, Andreas Gallusser, Tomas Kirchhausen, Juan S. Bonifacino*

Tyrosine-based signals within the cytoplasmic domain of integral membrane proteins mediate clathrin-dependent protein sorting in the endocytic and secretory pathways. A yeast two-hybrid system was used to identify proteins that bind to tyrosine-based signals. The medium chains ($\mu_1$ and $\mu_2$) of two clathrin-associated protein complexes (AP-1 and AP-2, respectively) specifically interacted with tyrosine-based signals of several integral membrane proteins. The interaction was confirmed by in vitro binding assays. Thus, it is likely that the medium chains serve as signal-binding components of the clathrin-dependent sorting machinery.

Targeting of integral membrane proteins to endosomes, lysosomes, the basolateral plasma membrane, and the trans-Golgi network (TGN) is largely mediated by sorting signals contained within the cytoplasmic domain of the proteins [reviewed in (*1*)]. Many of these sorting signals consist of continuous sequences of four to six amino acids containing a critical tyrosine residue. A subset of tyrosine-based signals conforms to the canonical motif YXXØ, where Y is tyrosine, X is any amino acid, and Ø is an amino acid with a bulky hydrophobic side chain (*1*). Although much has been learned in recent years about the structure and function of tyrosine-based signals, the molecular mechanisms involved in their recognition are still poorly understood. Previous studies have provided evidence for an association of cytoplasmic domains bearing tyrosine-based signals with clathrin-associated protein complexes (*2*). However, the exact identity of the signal-binding proteins and the molecular details of the recognition event remain to be established.

We decided to search for proteins that interact with tyrosine-based sorting signals, using a yeast two-hybrid approach (*3*). As a "bait" in the two-hybrid system, we used a triple repeat of the tyrosine-containing sequence SDYQRL (*4*, *5*) from the cytoplasmic tail of the integral membrane protein TGN38 (*6*). This sequence has the characteristics of a YXXØ motif and mediates both internalization from the cell surface and localization to the TGN (*7*). Screening of a mouse spleen complementary DNA (cDNA) library (~$2.5 \times 10^6$ clones) resulted in the isolation of two clones that interacted specifically with the $(SDYQRL)_3$ bait sequence (*8*). The two clones (termed 3M2 and 3M9) corresponded to the medium chain ($\mu_2$) of the plasma membrane, clathrin-associated protein complex AP-2 (*9*). In addition to $\mu_2$, the AP-2 complex contains two large chains (α- and β-adaptin) and one small chain ($\sigma_2$) (*10*).

Using growth on histidine-deficient (–His) plates as an assay (*11*), we found that proteins encoded by both 3M2 and 3M9 interacted not only with the $(SDYQRL)_3$ repeat but also with a single SDYQRL sequence and with the full-length TGN38 cytoplasmic tail (Fig. 1A). Mutation of the tyrosine (Y) residues in all three contexts abolished interaction with the $\mu_2$ clones (Fig. 1A). The binding specificity of $\mu_2$ was further characterized by mutation of each residue of the SDYQRL sequence individually to alanine. Only the Y and L residues were absolutely required for interaction with 3M9, whereas mutation of the S, D, and Q residues had no detectable effect, and mutation of the R residue decreased but did not completely abolish the ability to grow on –His plates (Fig. 1B). Thus, $\mu_2$ was capable of interacting with the sequence SDYQRL in various contexts and under sequence requirements that were consistent with those defined in studies in vivo (*7*).

To corroborate the results obtained with the two-hybrid system, we tested whether in vitro–translated, $^{35}$S-methionine–labeled $\mu_2$ was capable of interacting with various sequences appended to glutathione-S-transferase (GST) (Fig. 2). We observed that both the 3M2 and 3M9 forms of $\mu_2$ bound to GST-$(SDYQRL)_3$ but not to GST-$(SDGQRL)_3$ or to GST (Fig. 2A). In vitro–translated luciferase, used as a negative control, did not interact with any of the GST fusion proteins tested (Fig. 2A). Binding of

H. Ohno, M.-C. Fournier, H. Bosshart, I. Rhee, J. S. Bonifacino, Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA.
J. Stewart, A. Gallusser, T. Kirchhausen, Department of Cell Biology and Center for Blood Research, Harvard Medical School, Boston, MA 02115, USA.
S. Miyatake and T. Saito, Division of Molecular Genetics, Center for Biomedical Science, Chiba University School of Medicine, Chiba 260, Japan.

*To whom correspondence should be addressed.