Lipophosphoglycan (LPG) and the identification of virulence genes in the protozoan parasite Leishmania

Stephen M. Beverley and Salvatore J. Turco

Protozoan parasites of the genus Leishmania are responsible for a spectrum of human diseases, termed leishmaniasis. Depending on the species involved, leishmaniasis appears clinically in three forms: cutaneous, mucocutaneous and visceral, the last being fatal if untreated. The parasites have a remarkable capacity to avoid destruction in the hostile environments encountered in their life cycle, alternating between intracellular macrophage parasitism and extracellular life in the gut of the sand fly vector (Fig. 1). As with other microbial pathogens, the development of genetic tools for the study of these parasites promises to help unravel the molecular details of how they persevere under such harsh circumstances.

The state of Leishmania genetics
Since the first report of transient transfection in the related protozoan Leptomonas1, great progress has occurred in the range of molecular genetic approaches available for the study of Leishmania and its trypanosome relatives. Table 1 summarizes current approaches and compares Leishmania with the reigning leader in eukaryotic genetics, the yeast Saccharomyces cerevisiae. The available methods include a variety of positive and negative markers, diverse episomal and integrating expression vectors, highly efficient homologous gene replacement, artificial chromosomes (LACs) and inducible expression systems. These techniques permit extensive testing of known parasite genes by reverse genetic methods.

An important limitation of reverse genetic approaches is that one must already have identified a gene, by any one of several routes (homology, expression, genome projects or immunoreactivity). As these properties do not guarantee that the target gene will turn out to be functionally ‘interesting’, the researcher is often placed in a position of expectancy about the phenotype of engineered mutations. In contrast, genetic approaches, which start with a mutant showing alterations in some important property, can provide some promise (again not absolute) that the genes involved will turn out to be ‘interesting’. Moreover, genetic screens are not limited by the researcher’s preconceptions as to which molecules are important; instead, the organism is given the opportunity to identify relevant processes without prejudice.

Leishmania and other trypanosomatids pose interesting challenges for prospective geneticists. First, sexual crosses are either impossible or difficult. Although the occasional occurrence of hybrid parasites has raised the prospect of experimental sexual exchange in Leishmania, population genetic approaches have not disclosed a high level of genetic exchange in nature. Several attempts to generate crosses between marked parasites have proved unsuccessful (summarized in
Ref. 3), although one cross between lines bearing episomal selectable markers has recently been reported (A. Laban, pers. commun.). If this could be performed routinely with chromosomal markers, it would be tremendously useful. Second, *Leishmania* and trypanosomes are diploid organisms. This poses no problem for dominant gain-of-function mutations, which may require only a single activating mutation, but poses a considerable problem in generating loss-of-function mutations, as two alleles must be disrupted by independent mutations or by one mutation followed by some mechanism involving loss of heterozygosity. Regardless of the mechanism, even after heavy mutagenesis *Leishmania* mutants are recovered at a low frequency, in the order of $10^{-7}$ (Refs 4–6). Experimentally, this means that selection experiments rather than screening assays must be used to identify mutants, placing a tremendous burden on those seeking to make avirulent mutants, whose phenotype is difficult to select for directly in vivo or in vitro in *Leishmania* using current methods (the number of mice required to screen for avirulent mutants would fill a room!). It should be noted that other eukaryotic microorganisms are also ‘asexual’ diploids, including species of *Candida*, creating similar obstacles for researchers. Fortunately, there is a candidate virulence determinant in *Leishmania* whose properties made it an ideal starting point for genetic manipulations.

**Lipophosphoglycan (LPG): a multi-functional virulence determinant**

The promastigote form of *Leishmania* is covered with a thick glyocalyx, composed primarily of a single molecule termed lipophosphoglycan (LPG; see Fig. 2). LPG contains a repeating polymer of disaccharide-phosphate repeating units, anchored to the surface of the parasite by a glycosphingolipidinositol (GPI) anchor. In many ways, LPG seems to be functionally and structurally analogous to the lipopolysaccharide of many prokaryotes. The abundance of LPG on the parasite surface, the site of the primary interface with the host, suggests a central role for the glycoconjugate in its infectious cycle. So far, LPG has been implicated in binding and release of the parasite in the midgut.
of binding to lectins such as ricin agglutinin. By subjecting heavily mutagenized *Leishmania donovani* to multiple rounds of agglutination with ricin and enriching for non-agglutinated cells, nonreactive parasites are readily obtained. These survivors lack detectable LPG (Ref. 6) and are thus termed lpg- mutants. In *Leishmania major*, which possesses a more complex LPG with many side-chain galactose residues attached to the disaccharide-phosphate repeats (Fig. 2), lectin- or anti-LPG antibody 

In *L. donovani*, lpg- parasites could arise from the loss of any step of the LPG biosynthetic pathway (from addition of the galactosyl cap through to the membrane phosphadiyninositol anchor) (Fig. 2). The first four lpg- mutants characterized (R2D2, C3PO, OB1 and JEDI) each show different defects in LPG structure, with R2D2 affecting the glycan core and the others affecting different aspects of the disaccharide-phosphate repeats (Table 2; Fig. 2). Many other lpg- mutants have been obtained and are now being characterized. From the structure of LPG, we estimate that at least 25 distinct biosynthetic enzymes are necessary, and current data suggest that a similar number of proteins will be required for the correct compartmentalization and targeting of LPG to the cell surface. Thus, LPG genetics is still in its infancy.

With the availability of mutants, the next task was to develop a functional rescue methodology similar to that used by geneticists studying other microbial systems. A genomic library has been constructed in an extrachromosomal *Leishmania-Escherichia coli* shuttle vector, cLHYG (Ref. 12), which has a capacity of 40 kb and thus requires only 1000 cosmids to cover the 35-50-Mb *Leishmania* genome minimally. Following transfection into lpg- mutants, parasites are identified by lectin or anti-LPG antibody selection by panning 

Analysis of the cosmids carried by these lpg- transfectants has confirmed that they all encode genes affecting LPG biosynthesis.

### Properties of LPG genes

Four *L. donovani* LPG genes have been characterized (Table 2; Fig. 2). These can be divided into two classes, based on their mode of action. LPG class I genes (LPG1 and LPG4A) encode glycosyltransferases and other enzymes required for the biosynthesis of LPG linkages and/or components. These enzymes and the novel linkages they synthesize offer an exciting opportunity for glyobiologists studying enzyme specificity and action, as well as to pharmaceutical companies seeking to design neoglycoconjugates or control leishmaniasis. For example, LPG1 appears to encode a glycosyltransferase required for the addition of galactosylfuranose (Gal) within the LPG core (Table 2; Fig. 2). Gal is common in many microbial pathogens (in various bacteria, including *Mycobacterium*, as well as in *Trypanosoma cruzi* and several pathogenic fungi) but is not found in mammals, thus making it a potential target for rational chemotherapy.

Class II LPG genes (LPG2 and LPG3) encode proteins involved in the compartmentalization and assembly of LPG, guiding it through the secretory pathway to its final surface destination. LPG2 encodes a GDP-mannose translocase activity that functions in the

### Table 1. Genetic tools in Leishmania and the yeast *Saccharomyces cerevisiae*

<table>
<thead>
<tr>
<th>Property</th>
<th><em>S. cerevisiae</em></th>
<th><em>Leishmania</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Colonies</td>
<td>Yes (1–2 days)</td>
<td>Yes (1–2 weeks)</td>
</tr>
<tr>
<td>Positive markers</td>
<td>Many</td>
<td>&gt;6</td>
</tr>
<tr>
<td>Negative markers</td>
<td>Many</td>
<td>&gt;4</td>
</tr>
<tr>
<td>Expression vectors</td>
<td>Episomal, chromosomal</td>
<td>Episomal, chromosomal</td>
</tr>
<tr>
<td>Artificial chromosomes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Efficient gene targeting/knockout</td>
<td>Yes</td>
<td>Yes&lt;sup&gt;38,42&lt;/sup&gt;</td>
</tr>
<tr>
<td>Inducible expression</td>
<td>Yes</td>
<td>Emerging</td>
</tr>
<tr>
<td>Transposable elements</td>
<td>Yes</td>
<td>Yes&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td>Functional rescue</td>
<td>Yes</td>
<td>Yes&lt;sup&gt;13,14,25,43&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sex</td>
<td>Yes</td>
<td>'Chaste'</td>
</tr>
<tr>
<td>Ploidy</td>
<td>Diploid or haploid</td>
<td>Diploid</td>
</tr>
</tbody>
</table>

---

**Fig. 2.** Structure of lipophosphoglycan (LPG). The four domains in the *Leishmania donovani* LPG are (1) a 1-O-alkyl-2-lyso-phosphatidyl(myo)inositol anchor (PI), (2) a heptasaccharide glycan core, (3) multiple repeating disaccharide–phosphate units and (4) a small oligosaccharide cap. LPGs from all species of *Leishmania* have an identical lipid anchor and glycan core. The Gal–Man–PO<sub>3</sub>, backbone of the repeating units is also conserved but additionally contains branching 'side-chain' sugars in other species of *Leishmania*. There can also be minor variations in cap structure. During metacyclogenesis (see Fig. 1), the number of repeating units on LPG approximately doubles and, in some species, additional modifications to the side-chain sugars occur. These play important roles in the ability of LPG to modulate binding and release from the midgut wall and in resistance to complement. The location of the defects in four lpg- mutants (JEDI, OB1, R2D2 and C3PO) are shown. Abbreviations: Gal, galactose; Gal<sub>α</sub>, galactosylfuranose; Glc, glucose; Gn, glucosamine; Man, mannose; P, phosphate.
transport of this key substrate into the Golgi lumen, where GDP–mannose is used in assembling the disaccharide–phosphate repeats of LPG and other parasite molecules 13,15,16 (Fig. 2). LPG3 encodes a member of the GRP94/HSP90 family of eukaryotic chaperones (A. Descoteaux, S.J. Turco and S.M. Beverley, unpublished; Table 2; Fig. 2). Unlike most chaperones, which usually show pleiotropic effects when mutated, lpg3 knockout mutants exhibit no defects other than loss of the characteristic disaccharide–phosphate repeats that are added to LPG and related molecules (Table 2; Figs 2, 3). This suggests that the LPG3 chaperone has a restricted substrate specificity, which potentially recognizes nascent LPG motifs.

As LPG is developmentally regulated, being expressed at high levels in the promastigote but not the macrophage amastigote stage (Fig. 1), a third group (class III) of LPG mutations affecting regulatory pathways may exist. Theoretically, mutations that alter the promastigote developmental program towards that of the amastigote could confer an lpg phenotype. Mutants of this class would be particularly informative, as little is known about the mechanism of stage-specific gene expression in Leishmania. Although no class III genes have yet been recovered, they may emerge when more lpg mutations are characterized; a pessimistic alternative is that they may be lethal.

Examination of three of the LPG mutants has shown that C3PO and JEDI possess deletions of their respective LPG gene, whereas OB1 contains point mutations (R2D2 has not been studied intensively, although Southern blot analysis shows it to possess an unarranged LPG1 locus; Table 2). Interestingly, the two deletions in C3PO and JEDI appear to be homozygous, suggesting that mutation followed by loss of heterozygosity may be a common mechanism of diploid mutation. In every case studied, L. donovani lpg mutants appear to be loss-of-function mutants, and their rescue by cosmid transfection occurs by restoration of the defective gene by genetic complementation (rather than repression).

Is LPG a virulence determinant?

To confirm that a gene or molecule is a ‘virulence’ gene, a set of ‘molecular’ Koch’s postulates should be satisfied17,18: (1) the molecule must be reasonably associated with pathogenicity or infectivity; (2) inactivation of the gene should lead to a significant loss of virulence and (3) restoration of gene function should fully restore pathogenicity. LPG readily satisfies the first criterion and, for the second, several Leishmania lpg mutants are significantly compromised in their ability to survive in

---

**Table 2. LPG mutants and genes**

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Affected LPG domain</th>
<th>Gene</th>
<th>Mutation</th>
<th>Mode of rescue</th>
<th>Role</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2D2</td>
<td>LPG core</td>
<td>LPG1</td>
<td>Possible point mutations</td>
<td>Unknown</td>
<td>Putative Gal, transferase</td>
<td>43,44</td>
</tr>
<tr>
<td>C3PO</td>
<td>Disaccharide–phosphate repeats</td>
<td>LPG2</td>
<td>Deletion</td>
<td>Complementation</td>
<td>Goi GDP–mannose transport</td>
<td>13,15</td>
</tr>
<tr>
<td>OB1</td>
<td>Disaccharide–phosphate repeats</td>
<td>LPG3</td>
<td>Point mutations</td>
<td>Complementation</td>
<td>GRP94/HSP90 homolog; chaperone?</td>
<td>b</td>
</tr>
<tr>
<td>JEDI</td>
<td>Disaccharide–phosphate repeats</td>
<td>LPG4A</td>
<td>Deletion</td>
<td>Complementation</td>
<td>Putative mannosyl–phosphate transferase</td>
<td>c</td>
</tr>
</tbody>
</table>

---

*Abbreviations: LPG, lipophosphoglycan; Gal, galactosylfuranose.*


*b. H. Xu et al., unpublished.*
macrophages, cause infections in animals or survive in the insect vector. However, in none of these studies has the third postulate been applied, owing to a lack of cloned LPG biosynthetic genes. This poses a serious problem, as Leishmania (and many other pathogens) are known to lose virulence during in vitro culture; moreover, LPG mutants are invariably obtained after heavy mutagenesis, the effects of which cannot be countered by sexual back-crossing. Culture or transfection-associated loss of virulence has also been a serious problem in the study of other presumptive virulence molecules in Leishmania, as the ‘add-back’ of the gene to the mutant frequently fails to restore virulence fully. The solution to this problem is an important area for future studies.

Preliminary studies of some lpg mutants suggest that they do not satisfy the third postulate above. These mutants were obtained following mutagenesis and selection, or by targeted gene knockouts, in either L. major or L. donovani. Thus, for these mutants genes a role (or lack thereof) in the infectious cycle cannot be proven. However, other LPG mutations do pass the third postulate, as their behavior in animal infectivity or fly survival tests is restored nearly or entirely to wild-type levels following restoration of LPG gene function. These findings emphasize the need for a rigorous examination of ‘cause-and-effect’ when studying asexual diploid parasites. The development of inducible expression systems will simplify this task considerably.

One complication in the study of Leishmania glycoconjugates is that LPG components appear on molecules other than LPG. For example, the disaccharide-phosphate repeats occur on secreted proteins, such as acid phosphatase (sAP) and proteophosphoglycan (PPG), and as a secretory polymer (PG; Fig. 3). Examination of lpg mutants (C3PO or homozygous knockouts) reveals that they have lost all disaccharide-phosphate repeat-bearing molecules. Thus, the phenotype of these mutants should be more precisely attributed to loss of repeats, rather than LPG alone. Similar concerns apply to other LPG components, such as the gycan core or phosphatidylinositol anchor (Fig. 2). As more LPG genes are identified, it seems likely that ones specifically affecting LPG domains will emerge, permitting more discrimination between the roles of these glycoconjugates.

**New horizons for functional gene identification**

So far, it has always been possible to perform successful functional genetic rescue if loss-of-function mutations can be obtained. Mutant hunts can be performed directly on wild-type parasites or on parasite populations bearing engineered reporter molecules designed to reveal prospective mutant phenotypes. In this way, a mutant (gim1-1) and then the gene (GIM1) affecting uptake of glycosomal proteins has been obtained. This study illustrates the potential range for functional approaches through mass screening; although no selective method existed for the wild-type phenotype, examination of several thousand individual cosmid-transfected colonies has permitted identification of the one bearing GIM1.

**Questions for future research**

- What approaches could be used to increase the frequency of mutants in asexual diploids such as Leishmania?
- What are the forces and affected genes involved in loss of virulence during prolonged culture?
- How is lipophosphoglycan (LPG) synthesis compartmentalized within the cell, relative to enzymes involved in protein glycosphingolipid synthesis?
- What regulates LPG synthesis during development?
- Could similar approaches be used to identify mutants genes affecting the synthesis of glycoconjugates found on the surfaces of other parasites, such as Trypanosoma cruzi, Entamoeba and Trichomonas?

The need for defined mutations and the difficulties in their recovery in the asexual diploid parasite Leishmania are strategic limitations at present. Thus, the development of methods to facilitate the identification of relevant mutants is a critical need. Basic studies of mutation, recombination, loss of heterozygosity, chromosome stability and ploidy, and treatments that could elucidate these in a useful manner, will provide important insights into overcoming these obstacles.

There are other routes to the functional identification of ‘interesting’ genes currently under development in studies on Leishmania. One of these is multi-copy suppression, where overexpression of inappropriate expression of a wild-type protein bypasses a mutational defect. As episomal vectors such as cLHYG are present in multiple copies, transfected cosmids libraries are ideally suited to this mode of rescue. An obvious application of this strategy is in the identification of drug resistance genes, whose detection following gene amplification has been a popular topic in Leishmania genetics for some time. Indeed, we have isolated numerous genes implicated in resistance to a wide spectrum of selective agents using this approach (P.C. Cotrim and S.M. Beverley, unpublished). Similarly, identification of gain-of-function mutations can be effected by transfection of cosmids libraries derived from mutant line DNA into wild-type lines, followed by appropriate selections.

Another approach is to use gene fusions. Recently, an active Drosophila transposable element, mariner, has been introduced into Leishmania in an active form. This ‘designer’ insertionional mutagen offers the possibility of generating ‘tagged’ gain-of-function or loss-of-function mutations, as well as gene fusions to drug resistance genes and reporter molecules, such as the green fluorescent protein (GFP, Ref. 29). GFP fusions in particular can be selected in various ways to yield molecules that show interesting patterns of stage-specific expression and/or cellular localization.

**Conclusions**

Genetic methods in Leishmania have progressed to the point that a variety of screens and selections designed to yield genes likely to play important roles in the infectious cycle can now be contemplated. Similar efforts are under way for other protozoan parasites, including the related kinetoplastid protozoans Trypanosoma brucei and T. cruzi, the apicomplexan parasites...
Plasmodium and Toxoplasma, Entamoeba, Giardia and Trichomonas. In some of these parasites, genetic methods are well advanced and the feasibility of functional rescue has been clearly illustrated. Although each of these parasites has its own peculiarities and hurdles, it seems likely that in the near future our understanding of virulence in protozoan parasites is likely to undergo a revolution.

Acknowledgements

We thank P. Cotzin, A. Descoteaux, L. Epstein, B. Mengeling, D. Sacks and H. Xu for providing unpublished results, D. Dobson and B. Mengeling for helpful comments on this manuscript, and H. Avila and A. Bray for assistance in preparing the figures. This work was supported by NIH grants AI29646 and 31078.

References

20 Opat, A. et al. (1996) Glycolobiology 6, 387–397
38 Laban, A. et al. (1990) Nature 343, 572–574

In the other Trends journals

A selection of recently published articles of interest to TIM readers.

• Genome sequencing and analysis, by J.D. Grausz – Drug Discovery Today 2, 510–511
• Pathogen avirulence and plant resistance: a key role for recognition, by P.J.G.M. De Wit – Trends in Plant Science 2, 452–458
• The genetic reconstruction of BCG as a new immunotherapeutic tool, by M.A. O’Donnell – Trends in Biotechnology 15, 512–517
• Aetiology of Kaposi’s sarcoma: current understanding and implications for therapy, by C. Bosshoff and R.A. Weiss – Molecular Medicine Today 3, 488–494
• Archaeal introns: splicing, intercellular mobility and evolution, by J.L. Anderson et al. – Trends in Biochemical Sciences 22, 326–331