THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 18, Issue of May 2, pp. 15523–15531, 2003
Printed in U.S.A.

Functional Identification of Galactosyltransferases (SCGs) Required for Species-specific Modifications of the Lipophosphoglycan Adhesin Controlling *Leishmania major*-Sand Fly Interactions*

Received for publication, February 13, 2003
Published, JBC Papers in Press, February 24, 2003, DOI 10.1074/jbc.M301568200

Deborah E. Dobson‡§, Luella D. Scholtes‡, Kelli E. Valdez‡, Deborah R. Sullivan¶, Brenda J. Mengeling¶, Salvatore Cilmi‖, Salvatore J. Turco¶, and Stephen M. Beverley‡‖

From the ‡Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110, the ¶Department of Biochemistry, University of Kentucky Medical Center, Lexington, Kentucky 40536, and the ‖Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Lipophosphoglycan (LPG) is an abundant surface molecule that plays key roles in the infectious cycle of *Leishmania major*. The dominant feature of LPG is a polymer of phosphoglycan (PG) (6Galβ1,4Manα1-PO$_4$) repeating units. In *L. major* these are extensively substituted with Gal(β1,3) side chains, which are required for binding to midgut lectins and survival. We utilized evolutionary polymorphisms in LPG structure and cross-species transfections to recover genes encoding the LPG side chain β1,3-galactosyltransferases (βGalTs). A dispersed family of six *SCG* genes was recovered, whose predicted proteins exhibited characteristics of eukaryotic GalTs. At least four of these proteins showed significant LPG side chain βGalT activity; SCG3 exhibited initiating GalT activity whereas SCG2 showed both initiating and elongating GalT activity. However, the activity of *SCG2* was context-dependent, being largely silent in its normal genomic milieu, and different strains show considerable variation in the extent of LPG galactosylation. Thus the *L. major* genome encodes a family of *SCGs* with varying specificity and activity, and we propose that strain-specific LPG galactosylation patterns reflect differences in their expression.

The trypanosomatid protozoan parasite *Leishmania* infects over 12 million people worldwide, causing a variety of diseases that range from mild cutaneous lesions to fatal visceral infections (1). Within vertebrates *Leishmania* resides within acidified phagosomes of macrophages as the amastigote stage. A key step of the infectious cycle is the ability of the parasite to be transmitted to fresh hosts by an insect vector, phlebotomine sand flies. Several studies have emphasized the importance of lipophosphoglycan (LPG),[1] an abundant surface glycolipid of *Leishmania* promastigotes, in sand fly survival (reviewed in Refs. 2–4). Following a sand fly bite, *Leishmania* and the blood meal are enclosed by a midgut peritrophic matrix for several days, whereas parasites differentiate to the replicating procyclic promastigote stage. During this period LPG and other phosphoglycans (PGs) contribute to survival in the hydrolytic milieu of the midgut (3). After a few days the matrix is degraded and the remnants of the blood meal are excreted; at this time, promastigotes bind to midgut epithelium through an LPG-dependent interaction to avoid being excreted as well (5). As digestion is completed and the fly prepares to feed again, parasites differentiate to the infectious metacyclic stage, which synthesize a structurally modified metacyclic form LPG that is unable to bind the midgut (5–7). The detached metacyclic parasites are adapted for transmission and establishment of the infection in a new vertebrate host (8).

The basic "backbone" structure in all *Leishmania* consists of a 1-*O*-alkyl-2-*lyso*-phosphatidylinositol lipid anchor and heptasaccharide core, to which is joined a long PG polymer composed of 15–30 (Galβ1,4Manα1-PO$_4$) repeating units, terminated by a capping oligosaccharide (Fig. 1). In different species and/or developmental stages a variety of modifications of the LPG backbone have been observed, including changes in the terminal capping oligosaccharide of LPG, the addition of side chain (sc) sugar modifications to the prototypic PG (Gal-Man-P) repeating unit, and increases in the number of PG repeats (reviewed in Refs. 2–4). These modifications contribute in a species-specific manner to the binding and release of *Leishmania* promastigotes during development in the sand fly. In procyclic *Leishmania donovani* the PG repeats are unmodified, and midgut binding occurs through the Gal-containing capping oligosaccharide; in metacyclics, the number of PG repeats approximately doubles, resulting in a conformational change that precludes midgut binding (9). In *L. major* the PG repeats are modified by side chain β1,3 galactosyl residues (scβGal); in metacyclics, the number of PG repeats increases and the scβGal residues are further modified by addition of arabinose caps to block midgut binding (6, 7). These species-specific modifications also play important roles in the ability of the natural sand fly vector to transmit *Leishmania* species (2–4). For example, neither *L. donovani* nor *Leishmania major* mutants lacking LPG PG scβGal residues can be maintained in the natural *L. major* host *Phleobotomus papatasi* (10, 11).

A common theme in many protozoan parasites is their ability to alter their surface coats to ensure survival in both the insect vectors and mammalian hosts (12). In many respects, LPG functions as a stage- and species-specific adhesin in the sand fly. Adhesins have been extensively studied in bacteria (13) and

* This work was supported by National Institutes of Health Grants AI31078 (to S. J. T. and S. M. B.) and AI20941 (to S. J. T.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: 660 S. Euclid Ave., Box 8230, St. Louis, MO 63110.

[1] The abbreviations used are: LPG, lipophosphoglycan; PG, phosphoglycan; sc, side chain; Tn, transposon; SCG, side chain galactose; LmFV1, *Leishmania major* strain Friedlin V1; HPLC, high performance liquid chromatography; scβGalT, side chain β1,3-galactosyltransferases; Ld, *Leishmania donovani* Sudanese strain 1S2D; ORF, open reading frame; aa, amino acid(s); nt, nucleotide(s); βGalT, β1,3-galactosyltransferases; EST, expressed tag sequence.

fungi (14). Typically microbial adhesins are proteins that mediate attachment through interactions with specific carbohydrate, lipid, and/or protein moieties in host receptors. In *Leishmania*, this relationship is reversed because LPG is a glycolipid adhesin responsible for binding to a putative protein receptor in sand fly midguts (4, 15).

Because of its importance to parasite development and vector transmission in the well characterized *L. major–P. papatasi* model, we focused on genes affecting the attachment of the side chain Gal residues to the LPG PG repeats (Fig. 1). Previously, we utilized LPG-deficient mutants in transfection-based functional rescue approaches to identify genes affecting synthesis of the LPG backbone (16–19). Here, we used "cross-species" transfections to identify loci mediating LPG scβGal additions. The *SCG* (side chain galactose) genes identified represent the first "expression cloning" of a glycosylation gene family crucial for mediating midgut attachment of parasites, and suggest an approach for identifying genes involved in midgut attachment of other parasite species. Targeting genes that disrupt normal *Leishmania*-sand fly interactions may represent a novel approach for interrupting disease transmission and compromising virulence.

EXPERIMENTAL PROCEDURES

Leishmania Culture and Transfection—L. major strain Friedlin V1 (LmFV1) is a virulent clonal derivative of the Friedlin line (MHOM/IL/80/Friedlin) obtained from D. L. Sacks (National Institutes of Health). *L. donovani* Sudanese strain 1S2D (Ld) is a virulent clonal derivative (MHOM/S.D./00/1S-2D) obtained from D. Dwyer (National Institutes of Health). Cells were grown in M199 medium containing 10% heat-inactivated fetal bovine serum (20). Procyclic promastigotes were harvested from logarithmically growing (log phase) cultures and metacyclic promastigotes were isolated by the peanut agglutinin method (21) from cultures that had been in stationary growth phase for 2–3 days (stationary phase parasites). Infection of BALB/c mice and recovery and purification of lesion amastigotes were performed as described (22).

Parasites were transfected by electroporation and clonal lines were obtained by plating on semisolid M199 media (20), containing drugs appropriate for each selective marker (50 μg/ml HYG, 15 μg/ml G418). Agglutination assays were performed as described (23) using WIC79.3 monoclonal antibody (24).

Cosmid Library Transfection and WIC79.3 Monoclonal Antibody Panning—These studies were approved by the relevant institutional biosafety committees. An LmFV1 genomic DNA library constructed in the cosmid shuttle vector cLHYG (23) was introduced by 30 separate electroporations into Ld; 13,100 independent transfectants were obtained. These were combined into three independent pools, and transfectants bearing Gal-modified LPG PG repeats were isolated by panning with WIC79.3 antibody as described (18, 19). Three successive rounds of WIC79.3 antibody panning were performed, yielding a population that was strongly reactive. After plating to obtain single colonies, cosmid DNAs c*SCG1* (laboratory strain B3547), c*SCG2* (B3558), c*SCG2a* (B4876), and c*SCG4* (B3559) were recovered by transformation of *Escherichia coli* (18).

Genomic Cosmid Library Screen for SCG Genes—The 774-bp radiolabeled *SCG* universal probe (described below; Fig. 2*B*) was used to screen the LmFV1 genomic cosmid library (23). Eighteen positive cosmids were identified, including c*SCG3* (B3979), c*SCG5* (B3985), and c*SCG6* (B3971). Mapping and limited sequence analysis showed that in their respective cosmids, *SCG1*, *SCG4*, *SCG5*, and *SCG6* are in the same transcriptional orientation as the *HYG* marker, whereas *SCG2* and *SCG3* are in the opposite orientation.

Molecular Constructs—Partial *Xho*I deletions of cosmids B3547/*SCG1* and B3558/*SCG2* (Fig. 2*A*) were generated by limiting digestion followed by circularization with T4 DNA ligase. A 7.6-kb *Hin*dIII fragment from B3558 was inserted into *Hin*dIII-digested pSNBR (25), yielding pSNBR-*SCG2* (B3743). pSNBR-*SCG2* was digested with *Bse*36I, blunted with T4 DNA polymerase, partially digested with *Sma*I to release a 2.8-kb 5′-flanking region fragment, then religated using T4 DNA ligase to create pSNBR-*SCG2*del1 (B3899). Digestion of pSNBR-*SCG2*del1 with *Xcm*I followed by circularization with T4 DNA ligase generated pXK-*SCG2* (B3900), containing the 2.4-kb *SCG2* coding region plus 48 bp of 5′ and 147 bp of 3′-flanking DNA. Relevant sequence of all constructs was verified using standard methods with an Applied Biosystems ABI-373 automated DNA sequencer.

In Vitro Transposon Mutagenesis—The *Xho*I deletion cAX21/*SCG1* (B3556) (Fig. 2*A*) was used as a target for *mariner* mosK transposon (Tn) mutagenesis as described (26). A total of 80 transposon insertions were mapped and eight that fell within the 5.6-kb *Leishmania* insert were analyzed by transfection into Ld parasites (Fig. 2*B* and data not shown). We also used pSNBR-*SCG2* (B3743) as a target for mosK Tn mutagenesis, mapping a total of 72 Tn insertions. Eight Tn insertions that fell within the 7.6-kb *Leishmania* insert were analyzed by transfection into Ld parasites (Fig. 2*B* and data not shown).

DNA Sequencing—Complete double-stranded DNA sequences were obtained for the 2.4-kb SCG coding and flanking regions from B3547/*SCG1* (GenBank™ AY230144), B3558/*SCG2* (GenBank™ AY230145), B3979/*SCG3* (GenBank™ AY230146), B3559/*SCG4* (GenBank™ AY230147), B3985/*SCG5* (GenBank™ AY230148), and B3971/*SCG6* (GenBank™ AY230149). In some cases, primers were labeled with [^{33}P]ATP (2000–4000 Ci/mmol) or [^{35}S]dCTP (1250 Ci/mmol) and templates sequenced manually. Single strand sequence from both ends of the *SCG1–6* cosmids was obtained using cLHYG-specific primers.

Northern and Southern Blot Analyses—Total *Leishmania* RNAs were prepared using the Trizol method (Invitrogen). RNA (5 μg) was analyzed by Northern blotting as described (27). RNA loading was normalized to ethidium bromide-stained rRNA. Genomic DNA was isolated and analyzed by Southern blotting as described (27). For molecular karyotype analyses, *Leishmania* chromosomes were prepared in agarose plugs and stored at 4 °C as described (28). Pulse field gel electrophoresis was performed in a Bio-Rad model CHEF-DR II apparatus using an electrophoresis field program separating 0.4–1.8 megabase DNAs.

SCG Probes—Probes were generated by PCR amplification in a PTC-200 thermocycler (MJ Research) using 20 pmol of the indicated primers, 50 ng of template, 0.25 mM dNTPs, and 1 unit of TAQ polymerase (Roche Diagnostics) in 50 μl of total volume following the manufacturer's directions. PCR-amplified DNAs were purified on a QIAquick PCR purification column (Qiagen) following the manufacturer's directions and probes were labeled as described above.

The 774-bp *SCG* "universal" coding region probe (corresponding to amino acids 1–258) was amplified as described above, using primers B890 (5′-CGCCATCGCAAACAGCATC) and B1240 (5′-gcggatccaccATGCGAGAGGAGAACAATGTGCCA; *Leishmania* sequences in uppercase) and pXK-*SCG2* template (B3900), 55 °C annealing temperature, and 2-min elongation time. The PCR product (1 μl) was used as template for an additional round of PCR amplification and purified before probe labeling.

Reverse Transcriptase-PCR Analyses—cDNA was prepared using 1 μg of LmFV1 RNA, random primers, and Superscript reverse transcriptase (Invitrogen) in a 20-μl reaction volume, following the manufacturer's directions. PCR reactions (50 μl), containing 1 μl of cDNA, 20 pmol of *L. major* miniexon (B936, 5′-AACGCTATATAAGTATCAGTTCTGTACTTTA) and common *SCG* (B871, 5′-GACGATGAGAGCAAGTAAGAC) primers, 0.25 mM dNTPs, and 1 unit of TAQ polymerase (Roche Diagnostics), were performed as described above using 62 °C annealing temperature and 2-min elongation time. Products were analyzed on a 5% acrylamide, 1× TBE (90 mM Tris, 90 mM borate, 2 mM EDTA, pH 8) gel run overnight at 2.3 V/cm. DNA was visualized by staining with SYBR Gold (Molecular Probes).

Purification and Analysis of LPG—LPG was extracted from logarithmically growing parasites ($4–8 \times 10^6$ cells/ml; 10^9 cells) in solvent E and purified by phenyl-coupled Sepharose chromatography (29). PG repeat units were generated by hydrolysis of purified LPG under mild acid conditions, recovered, and separated by HPLC Dionex chromatography (30). Aliquots of the PG repeats were dephosphorylated with *E. coli* alkaline phosphatase (0.1 unit, 16 h, 37 °C), desalted by passage through a two-layered column of AG 50W-X12 over AG 1-X8, labeled at the reducing ends with 8-aminonaphthalene-1,3,6-trisulfate and analyzed by GLYKO-FACE fluorophore-assisted carbohydrate electrophoresis according to manufacturer's specifications (Glyko Inc., Novato, CA). When [^{3}H]Gal-labeled LPG PG repeats were included, gels were visualized by UV illumination and the radioactivity eluted from excised bands was measured by scintillation counting. The presence of Gal in LPG side chains was shown by digestion with *E. coli* β-galactosidase prior to fluorescent labeling (10). Migration distances were compared with oligosaccharide standards. Strong acid hydrolysis (2 N trifluoroacetic acid, 2.5 h, 100 °C) of the repeat units followed by monosaccharide analysis indicated that >95% of the radiolabel remained as [^{3}H]Gal (data not shown).

LPG Side Chain β1,3-Galactosyltransferase (scβGalT) Assays—Mi-

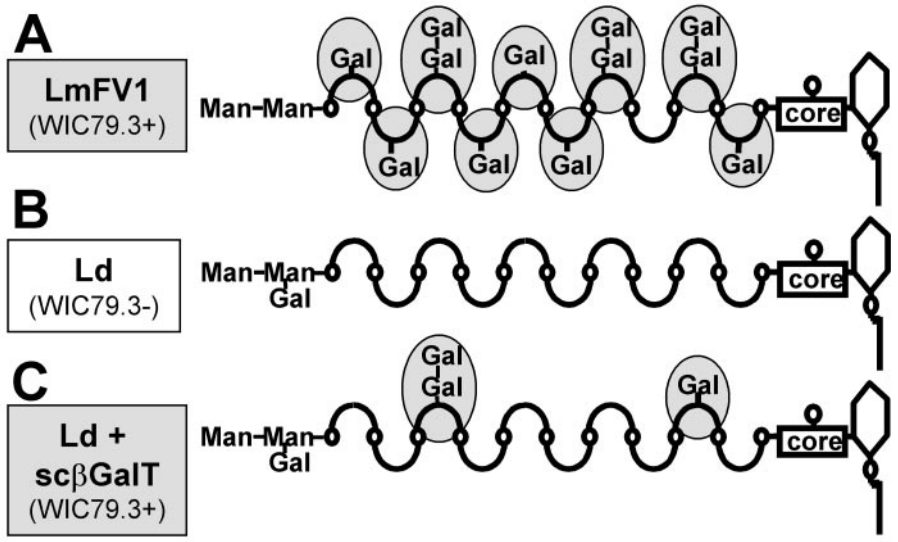

FIG. 1. **Predicted consequence of LPG side chain galactosyltransferase overexpression on *L. donovani* LPG structure.** The structures shown are modified from Refs. 6 and 48. *A*, *L. major* FV1 (*LmFV1*) LPG. The PG repeat unit backbone 6Gal(β1,4)Man(α1)-PO_4 is represented by *circles with curved lines*. The glycan "core" is Gal(α1,6)Gal-(α1,3)Gal_f(β1,3)[Glc(α1)-PO_4 $\rightarrow$ 6]Man-(α1,3)Man(α1,4)GlcN(α1,6) and is linked to a 1-*O*-alkyl-2-lyso-phosphatidylinositol anchor. Residues comprising WIC79.3 epitopes (Gal_n-(6Gal(β1,4)Man(α1)-PO_4) (24) are *shaded*. Note that the precise site of attachment of the Gal side chains in the PG repeats is heterogeneous. *B*, *L. donovani* (Ld) LPG. The structures shown are modified from Refs. 37 and 48, and abbreviations are defined in *panel A*. *C*, predicted LPG structure following LPG scβGalT expression in *L. donovani* LPG. The structures shown were observed in Ld *SCG* transfectants and are defined in *panel A*.

crosomes from logarithmically growing parasites ($4-8 \times 10^6$ cells/ml; 2×10^9 cells) were prepared by nitrogen cavitation and differential centrifugation, and LPG scβGalT assays were performed (10). Transfer of [^{3}H]Gal from UDP-[^{3}H]Gal to Ld LPG was verified by analysis of PG repeats by thin layer or Dionex HPLC chromatography (30).

RESULTS

Isolation of Cosmids Conferring Addition of Gal Side Chains to LPG—The LPG PG repeating units from LmFV1 bear β1,3-linked Gal modifications, which confer reactivity to the antibody WIC79.3 (Fig. 1*A*) (24). In contrast, the unmodified *L. donovani* (Ld) LPG is unreactive with WIC79.3 (Fig. 1*B*). *In vitro*, Ld LPG can serve as substrate for *L. major* PG scβGalT activity (10, 31). We reasoned expression of LmFV1 scβGalTs in *L. donovani* would confer WIC79.3+ LPG reactivity, and designed a functional rescue strategy to exploit this (Fig. 1*C*).

An LmFV1 genomic library prepared in the *Leishmania* shuttle cosmid vector cLHYG (23) was transfected into *L. donovani*, yielding a library of 13,100 Ld transfectants that provided >10-fold coverage of the ~35 megabase *L. major* genome. WIC79.3+ Ld transfectants were recovered following 3 rounds of WIC79.3 antibody panning, and clonal lines were obtained by plating. From 24 lines we recovered four different cosmids, which upon retransfection into *L. donovani* conferred WIC79.3 reactivity. Restriction mapping showed that these contained 3 different loci that we termed *SCG*: *SCG1* (cosmid B3547), *SCG2* (cosmids B3558, B4876), and *SCG4* (cosmid B3559; Figs. 2 and 3, and data not shown).

Identification of SCG Genes—The active regions within *SCG1* and *SCG2* cosmids were identified by WIC79.3 reactivity tests of Ld transfectants bearing deletion derivatives, or following *mariner* Tn insertion mutagenesis (26) (Fig. 2 and data not shown). Analysis of 10 cosmid *SCG1*/B3547 *Xho*I deletions identified a 5.6-kb *SCG1* active region (*gray box*, Fig. 2*A*, and data not shown). Analysis of eight transposon insertions within this region showed that four had lost WIC79.3 reactivity (Tns 45, 31, 63, and 13; Fig. 2*B* and data not shown). The WIC79.3-unreactive Tn insertion sites were clustered in a 1.4-kb region; sequence analysis out to the nearest flanking WIC79.3+ Tn insertion sites revealed a 2.4-kb open reading frame (ORF) encoding a 814-amino acid protein (Figs. 2*B* and 4*A*). Notably, all WIC79.3-unreactive Tn insertions were mapped within the predicted *SCG1* ORF.

A similar approach was used to localize the active *SCG2* gene to a 7.2-kb region in cosmid B3558 deletion B1X17 (*gray box*, Fig. 2*A*). A 7.6-kb *Hin*dIII fragment encompassing this region was inserted into the *Leishmania* shuttle vector pSNBR (25) (pSNBR-*SCG2*), sequenced, and subjected to Tn mutagenesis as described above. The sequence revealed a 2.4-kb ORF encoding a 814-amino acid protein, and 2.1 kb upstream of this a histone H4 pseudogene with a highly divergent 5′ end bearing numerous deletions and frameshifts. As seen for *SCG1*, the three WIC 79.3-unreactive Tn insertions obtained clustered in an 1.5-kb region within the *SCG2* ORF (Fig. 2*B*). This *SCG2* ORF was closely related to the *SCG1* ORF and other *SCG* genes as discussed below. Expression of the *SCG2* ORF alone in an expression vector similar to the pX vectors (pXK-*SCG2*) yielded exclusively WIC79.3+ colonies when transfected into Ld parasites (Table I).

The SCG Family Comprises Six Independent Loci—Southern blot analysis of LmFV1 chromosomes or digested DNAs with *SCG1*/*SCG2* ORF probes indicated that there were additional *SCG* loci not recovered in our screen, suggesting it had not reached saturation (Fig. 3*A* and data not shown). These were recovered from the LmFV1 cLHYG genomic cosmid library by screening with a universal *SCG* coding region probe (Fig. 2*B*): from a total of 18 cosmids we obtained new representatives of *SCG1*, *SCG2,* and *SCG4*, as well as three new cosmid loci termed *SCG3* (B3979), *SCG5* (B3985), and *SCG6* (B3971; Fig. 3*A* and data not shown).

Southern blot analysis of LmFV1 chromosomes separated by pulsed-field electrophoresis revealed six chromosome bands hybridizing to the universal *SCG* probe (Fig. 3*B*). To assign each *SCG* locus to a given chromosome, locus-specific fragments were isolated from each cosmid and hybridized to chromosome blots. This assigned the *SCG3*, *SCG1*, *SCG2*, *SCG6*, *SCG5*, and *SCG4* loci to chromosomes of ~300, 550, 700, 850, 1600, and 2800 kb, respectively (Fig. 3*B* and data not shown). These data and comparisons of the X*ho*I digestion patterns of total LmFV1 genomic DNA against the *SCG* cosmid panel (Fig. 3*A* and data not shown) suggested that the entire family of LmFV1 *SCG* loci had been identified and recovered intact.

WIC79.3 tests of multiple clonal lines for each Ld cosmid *SCG* transfectant showed that all *SCG3*/B3979 parasites were reactive, whereas all *SCG5*/B3985 and *SCG6*/B3971 transfectants were unreactive. The WIC79.3−unreactive phenotype of *SCG5* and *SCG6* Ld transfectants was not because of alterations in *SCG* cosmid DNAs following transfection into *Leishmania* (data not shown). As a control, the three cosmids identified by functional complementation (*SCG1*/B3547, *SCG2*/

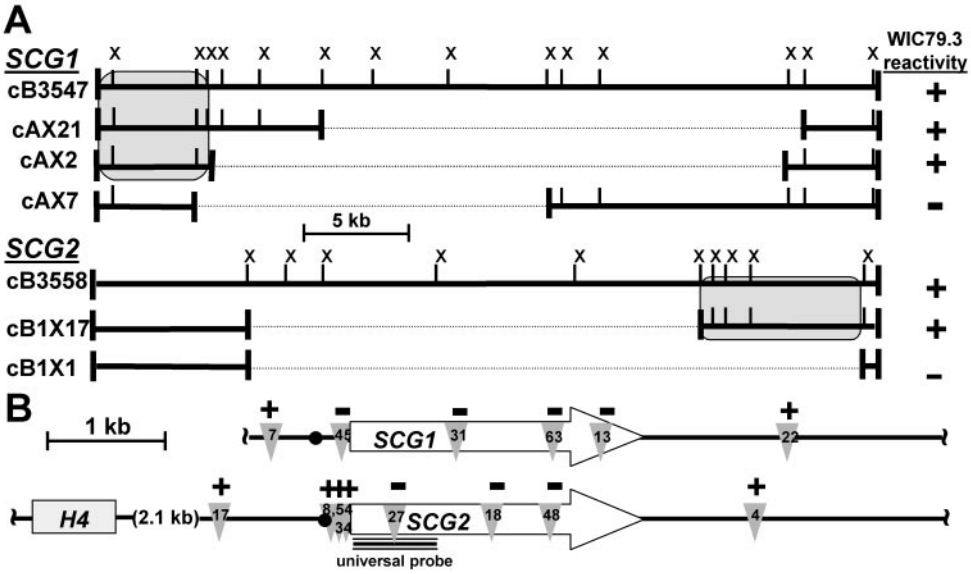

FIG. 2. **Functional mapping of *SCG1* and *SCG2* loci.** *A*, functional analyses of *SCG1* and *SCG2* cosmid deletions. A map of the *Leishmania* DNA present in cosmids c*SCG1*/B3547 (*top*) and c*SCG2*/B3558 (*bottom*) is shown. *Insets* from representative *Xho*I deletion cosmids are shown as *solid lines*; *dotted lines* represent the deleted region. *X*, *Xho*I restriction site. WIC79.3 reactivity in agglutination tests of Ld *SCG* transfectants expressing each cosmid is shown to the *right* of each construct ("+" = positive; "−" = negative). Predicted gene locations (*gray boxes*) are based on WIC79.3 reactivity of deletions. *B*, localization by transposon mutagenesis mapping. The location of the relevant *mariner* mosK Tn insertion sites (*gray triangles* with Tn number) within *SCG1* deletion cosmid AX21 (*top*) and *SCG2* plasmid pXK-*SCG2* (*bottom*) are shown. WIC79.3 reactivity of Ld *SCG* transfectants expressing each Tn insertion construct is indicated *above* the insertion site, as defined in *panel A*. *Open arrow*, predicted SCG ORFs; *H4*, histone H4 homolog. The position of the *SCG* "universal probe" is marked. The predicted SCG(X) splice acceptor is shown (●).

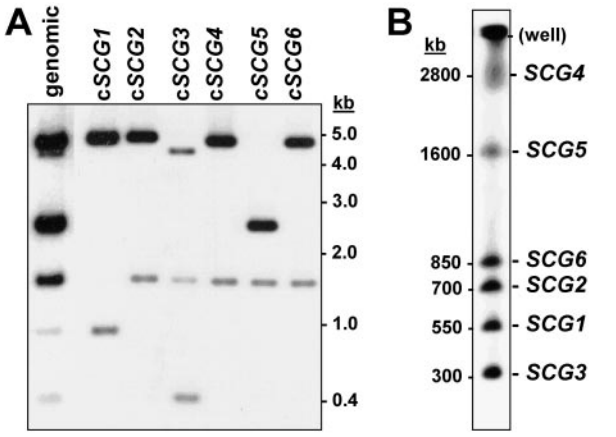

FIG. 3. **Arrangement of *SCG* loci in LmFV1.** *A*, *SCG* gene structures. Genomic DNA from LmFV1 (genomic) and DNAs from LmFV1 *SCG1–6* cosmids (c*SCG1–6*) were digested with *Xho*I, separated by gel electrophoresis, transferred to GeneScreen Plus membrane, and hybridized to a radiolabeled universal SCG probe (Fig. 2*B*). Positions of DNA size standards are marked. *B*, molecular karyotype analysis of *SCG* genes. Chromosomes of LmFV1 were prepared, separated by pulsed field electrophoresis, transferred to GeneScreen Plus, and hybridized with a radiolabeled universal *SCG* probe. The approximate size of SCG-hybridizing chromosomes was determined by comparison to mobility of *S. cerevisiae* chromosomes. The location of each *SCG* gene was determined by stripping the blot and hybridizing the unique flanking region radiolabeled probes from each *SCG* cosmid (data not shown).

B3558, and *SCG4*/B3559) again generated WIC79.3+ Ld transfectants.

Properties of the Predicted SCG Proteins—The regions of the *SCG3–6* cosmids that hybridized to the SCG1/2 ORF probes were mapped and sequenced (Fig. 4 and data not shown), revealing the presence of ORFs showing strong homology to those of *SCG1/2*. The predicted SCG1, -2, -3, and -6 ORFs were 814 amino acids (aa), whereas SCG4 had an N-terminal 32-amino acid extension arising from a single nt change; comparisons among these showed from 92 to 96% aa identity (Fig. 4, *A* and *B*). The SCG5 ORF encoded a protein of 816 residues, whose first 596 amino acids were highly homologous (94–97% aa identity) to the remaining SCG ORFs. Thereafter, the presence of numerous nucleotide substitutions/insertions beginning at SCG5 ORF nt 1789 caused higher divergence, resulting in 50% aa and 55% nt identity over the terminal 220 amino acids (Fig. 4, *A* and *B*, and data not shown). The 5′-flanking regions of the *SCG* genes were highly homologous, except for *SCG1* that diverged from other *SCG* genes 122 nt upstream of the predicted conserved start codon. Similarly, the 3′-flanking regions were also highly homologous, except for *SCG5* as noted above (data not shown).

All six predicted SCG proteins contained a single "DXD" sequence motif (DDD at aa 538–540 in SCG1, -2, -3, -5, -6, or 570–572 in SCG4; Fig. 4, *A* and *C*), a motif common to many glycosyltransferases that is implicated in catalytic activity (32). This motif was located within a region of weak homology (45% aa identity, 68% aa similarity) to eukaryotic β-galactosyltransferases (GalT, Fig. 4, *A* and *C*; GenBank™ conserved domain data base pfam01762). As shown below, transfection of *SCG* cosmids or an SCG2 ORF expression construct into *L. donovani* confers elevated LPG scβGalT activity. All predicted SCG proteins have the topology of type II membrane proteins (33), with a single predicted transmembrane domain (TM, Fig. 4*A*) preceded by an N-terminal signal anchor sequence of 108–141 aa (34). There were five potential *N*-linked glycosylation sites conserved in all SCG family members (*, Fig. 4*A*). These data suggested that the SCG proteins encode βGalTs with a lumenal catalytic domain, a conclusion supported by enzymatic studies of *SCG* transfectants (below).

Expression of SCG mRNAs—In log phase LmFV1 procyclic promastigotes, an SCG universal ORF probe identified a 3.8-kb mRNA, which will be referred to as *SCG(X)* (Fig. 5*A*). Lower amounts of a 5.5-kb transcript were also observed, which may reflect a processing intermediate arising from the polycistronic transcriptional mechanism employed by trypanosomatid protozoans (35). The 5′ end of the *SCG(X)* transcripts was mapped by reverse transcriptase-PCR to a position 264-nt upstream of

TABLE I
Activity of SCG genes in Ld transfection assay

Line[a]	WIC79.3 agglutination[b]	LPG side chain profile[c]	Microsomal LPG-scβGalT activity[d]
Ld c*SCG1*	+	3% (Gal)	66 ± 2
Ld c*SCG2*	++	2% (Gal) + <1% (Gal_2)	28 ± 17
Ld pXK-*SCG2*	+++	35% (Gal) + 35% (Gal_2) + 11% (Gal_3)	353 ± 66
Ld c*SCG3*	+++	68% (Gal) + 1% (Gal_2)	123 ± 11
Ld c*SCG4*	++	<1% (Gal)	30 ± 4
Ld c*SCG5*	−	0	6 ± 2
Ld c*SCG6*	−	0	44 ± 15
Ld	−	0	10 ± 1
LmFV1	+++	47% (Gal) + 14% (Gal_2) + 3% (Gal_3)	133 ± 60

[a] DNAs transfected into Ld are: "c," cosmids containing *SCG1* (B3547), *SCG2* (B3558), *SCG3* (B3979), *SCG4* (B3559), *SCG5* (B3985) or *SCG6* (B3971); pXK-*SCG2* (B3900) contains the SCG2 ORF in a *Leishmania* plasmid expression vector.

[b] Relative reactivity with WIC79.3 antibody in agglutination assays as an indication of LPG PG Gal side chains: "+" = 10–25%, "++" = 25–50%, "+++" = 100% agglutination.

[c] Percentage of LPG PG repeats that contain the indicated number of Gal side chains as determined by Dionex HPLC chromatography (Fig. 6). Representative of two independent experiments.

[d] Activity reported as counts/min of [^{3}H] Gal substituted Ld LPG PG repeats/h/mg protein ± SD. Representative of two independent experiments.

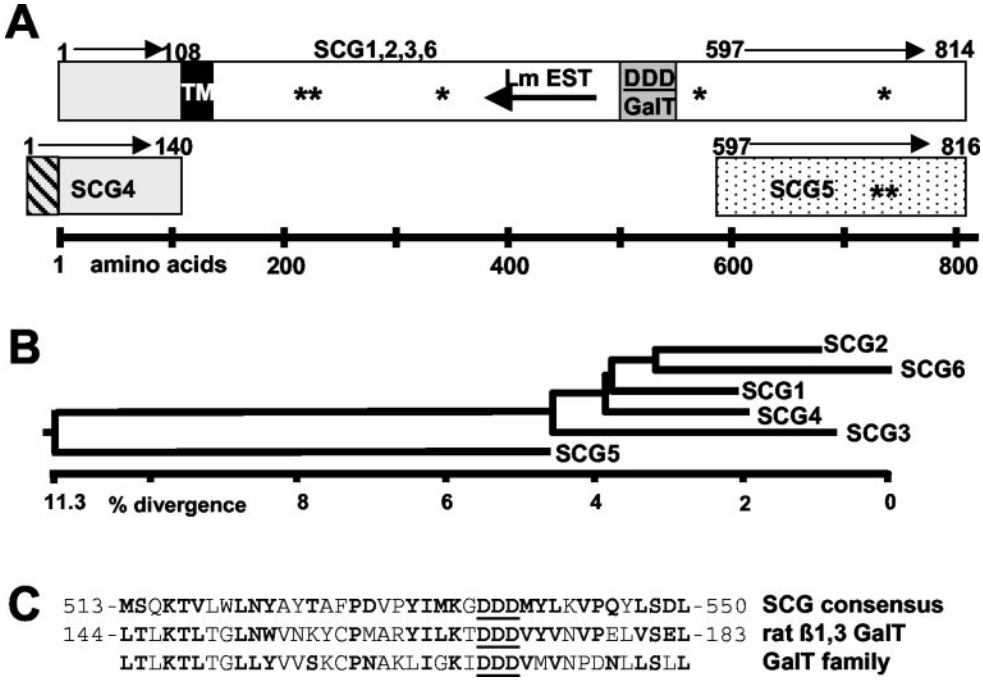

FIG. 4. **Properties of SCG proteins.** *A*, predicted SCG proteins. The conserved ORF identified in each *SCG* cosmid is represented as an *open box*. The locations of the predicted cytoplasmic domains (*light gray*), transmembrane domains (*TM*, *black*), galactosyltransferase catalytic domain (*DDD/GalT*, *dark gray*), *N*-glycosylation sites (*), and LmEST 0269 homology on antisense strand (*Lm EST*, arrow) are shown. Relevant amino acid positions are noted *above* the *box*. Positions of divergent regions are shown *below* the conserved SCG ORF: the C-terminal region of SCG5 (*stippled box*) and the N-terminal 32 amino acid extension of SCG4 (*striped box*). *B*, phylogenetic analysis of SCG ORFs. The amino acid sequence of the predicted SCG ORFs was analyzed using the Clustal method with the identity residue weight table (DNAstar); the percent aa sequence divergence is plotted. The unique N-terminal 32 amino acids in SCG4 were not included for this analysis. *C*, alignment of GalT domains. Alignment of the GalT homology domain for *SCG1–6* (*SCG* consensus), rat β1,3-GalT (GenBank™ AB003478), and GalT family consensus (GenBank™ pfam01762) is presented; *numbers* refer to the position in the respective proteins. *SCG5* differs from the *SCG* consensus at position 539 (M > T). Identical and conservative residues are in *bold* and the DDD catalytic motif (32) is *underlined*.

the conserved ATG in the SCG ORFs (Figs. 5*B* and 2*B*), and both *SCG(X)* transcripts were sufficiently large to encode the predicted SCG proteins (Fig. 4*A*). Relative to logarithmic growth phase procyclic parasites, *SCG(X)* transcript levels increased slightly in stationary growth phase parasites (about 1.5-fold for both transcripts) and more in metacyclic parasites (3- and 7.4-fold for 3.8- and 5.5-kb transcripts, respectively; Fig. 5*A*). This may be related to the doubling in the LPG PG repeat number known to occur as parasites differentiate from procyclic to metacyclic promastigotes upon entering stationary growth phase (6, 7). The presence of *SCG(X)* transcripts in amastigotes (Fig. 5*A*) may reflect the synthesis of phosphoglycans other than LPG, such as PPGs, that also bear Gal side chain modifications (36).

Searches of the *L. major* EST data base revealed an EST (lmEST0269; GenBank™ H64199) that was highly homologous (83–85% nucleotide identity) to the antisense strand of the *SCG1–6* genes (nt 1465–1162 in the SCG1, -2, -3, -5, -6 ORFs; nt 1561–1258 in SCG4; Fig. 4*A*). However, searches of the *L. major* genome did not yield a sequence identical to the EST. Potentially the sequence divergence could arise from technical sources (rapid EST sequencing) or polymorphisms between the Friedlin V1 strain studied here and the LV39 line studied in the EST project. Interestingly, these two strains differ in the degree of LPG PG galactosylation.[2] It was also surprising that the EST arose from the antisense strand of the SCG ORFs. Again, this could have a technical origin arising

[2] D. E. Dobson, B. Mengeling, S. Cilmi, S. Hickerson, S. Turco, and S. M. Beverley, manuscript in preparation.

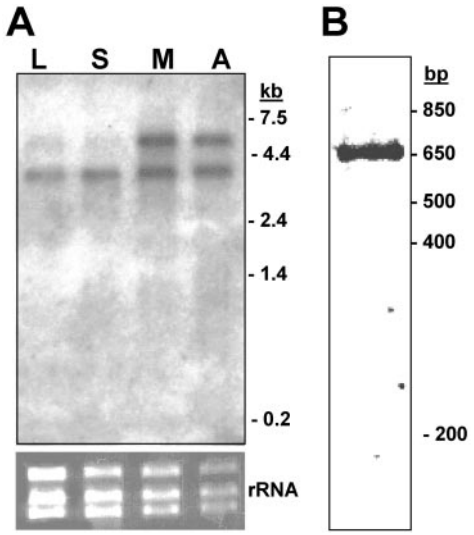

FIG. 5. **Structure and expression of *SCG(X)* transcripts.** *A*, Northern blot analysis. Total LmFV1 RNAs (5 μg) from logarithmicallly growing (*L*) and stationary phase (*S*) promastigotes, purified metacyclics (*M*), and lesion amastigotes (*A*) were separated by electrophoresis, transferred to GeneScreen Plus, and hybridized with a radiolabeled *SCG* universal probe (Fig. 2*B*). Positions of RNA size standards are marked. Ribosomal RNA was used as a loading control (*lower panel*). *B*, mapping transcript start site by reverse transcriptase-PCR. cDNA from log-phase LmFV1 promastigotes was subjected to reverse transcriptase-PCR using miniexon and conserved *SCG* primers, and the products were separated by electrophoresis. Positions of DNA size standards are marked. The data in both panels are representative of three independent experiments.

during cDNA library construction. However, antisense transcripts could play a role in *SCG* regulation through any one of a variety of mechanisms, although *Leishmania* appear to be deficient in the RNA-interference pathway (38).[3] At present we have no data addressing the reality or role of this EST.

LPG Side Chain Galactosylation in SCG Cosmid L. donovani Transfectants—The WIC79.3 reactivity of c*SCG1–4* Ld transfectants suggested that they synthesized LPGs with PG repeats containing βGal side chains (Fig. 1). Purified LPGs were subjected to mild acid hydrolysis and dephosphorylation, and the PG repeats were separated by Dionex HPLC chromatography (Fig. 6 and data not shown). As expected, Ld LPG yielded primarily the unsubstituted Gal-Man PG repeat (*peak G–M* in Fig. 6) (6, 29), whereas the majority of LmFV1 PG repeat units were substituted with 1–3 Gal residues (*peaks* G_{2-4}*-M* in Fig. 6). Ld-c*SCG3* cosmid transfectants showed significant levels of Gal-substituted PG repeat units (~77% of total repeat units), as expected given their strong reactivity with WIC79.3 (Table I). In contrast, only trace levels of Gal-substituted PG repeats were evident in the c*SCG1*, c*SCG2*, or c*SCG4* transfectants (1–4% of the total PG repeat units were so modified), and none were evident in the *SCG5* and *SCG6* cosmid transfectants (Fig. 6, Table I). Whereas the results with c*SCG5* and c*SCG6* were expected given their WIC79.3-negative phenotype, those for c*SCG1*, c*SCG2*, and c*SCG4* were surprising given their clear reactivity with WIC79.3 (Table I). Because LPG is highly abundant ($>10^6$ molecules/cell), presumably even a low level of βGal side chain addition to PG repeats is sufficient to confer strong reactivity with multivalent antibodies or agglutinins.

The Dionex HPLC results were confirmed by analysis of the LPG repeating unit structures by fluorophore labeling and electrophoresis (Fig. 7). LPG samples from Ld and Ld-vector control transfectants exhibited a single band corresponding to unsubstituted PG repeats (G-M). Ld-c*SCG1,* -c*SCG3,* and -c*SCG4* cosmid transfectant samples showed an additional band corresponding to PG repeats containing a single βGal side chain (Gal-G-M), as confirmed by their susceptibility to digestion with β-galactosidase (data not shown) and reactivity with WIC79.3 (Table I). In contrast, Ld4-c*SCG5* and -c*SCG6* samples showed only a single band corresponding to unsubstituted PG repeats (Fig. 7). Thus, *SCG5* and *SCG6* were unable to add either βGal or any other side chain sugar to LPG PG repeats.

L. major SCG Expression Confers LPG scβGalT Activity in L. donovani—An *in vitro* assay for LPG scβGalT activity (10) was used to test whether the *SCG* genes exhibited the predicted enzymatic activities. Parasite microsomes were incubated with the nucleotide sugar donor UDP-[^{3}H]Gal and purified unsubstituted *L. donovani* LPG acceptor, and the transfer of [^{3}H]Gal to LPG was measured. As expected, LmFV1 microsomes showed 13-fold more LPG scβGalT activity relative to Ld microsomes (Table I) (10). Significantly, LPG scβGalT activity was 12-fold higher in Ld-c*SCG3* cosmid transfectants relative to *L. donovani* controls. The microsomal LPG scβGalT activity in the other *SCG* cosmid transfectants was lower (1–6-fold above the Ld background; Table I), consistent with the low levels of galactosylation observed in purified LPGs (Figs. 6 and 7). In combination, the sequence, structural, and enzymatic data suggest the active *SCG* genes likely encode the LPG side chain β1,3-galactosyltransferases themselves.

SCG2 Expression Is Context-dependent and Shows Both Initiating and Elongating scβGalT Activity—Given the high sequence similarity among the predicted SCG proteins, the variation in PG repeat galactosylation in the *SCG* cosmid transfectants was unexpected (Figs. 6 and 7). One explanation was that the context of each *SCG* gene relative to the cosmid vector backbone led to differences in expression. However, restriction mapping and end sequencing of these cosmids showed that each contained all flanking sequences necessary to generate the 3.8-kb *SCG* mRNA (Fig. 5*A*), with the C terminus of each SCG ORF located 2.1–4.1 kb from the cloning site (data not shown). Although *SCG* genes in these cosmids were not all in the same orientation with respect to the selectable HYG marker (*SCG1, 4–6* = same, *SCG2,3* = opposite orientation), there was no correlation between *SCG* orientation and activity. Studies of other genes have shown that the effect of the cosmid vector orientation has little effect (39).

Another explanation involves location of cosmid-borne *SCG* genes within their normal genomic context. As described above, we expressed the SCG2 ORF using an expression vector similar to the widely used pX vectors (40). Unlike the c*SCG2* transfectants, the pXK-*SCG2* plasmid transfectants showed extensive β-galactosylation of LPG in both Dionex HPLC and electrophoretic analyses of fluorophore-labeled repeat units (Figs. 7 and 8), comparable or greater than the most active c*SCG3* cosmid Ld transfectant (Figs. 6 and 7; Table I). Similarly, the LPG scβGalT activity of the pXK-SCG2 Ld transfectants was 35-fold higher than untransfected *L. donovani*, or nearly 3-fold higher than the most active c*SCG3* cosmid transfectants (Table I).

A second finding from this study was that unlike the cosmid transfectants, the LPG synthesized in the pXK-*SCG2* Ld transfectant contained PG repeats bearing both single and oligo-Gal side chains (Figs. 7 and 8). Approximately 80% of the LPG PG repeats bore Gal modifications with about 46% consisting of oligo-Gal substitutions (Figs. 7 and 8; Table I). Because the pXK-*SCG2* plasmid and c*SCG3* cosmid Ld transfectants had similar overall levels of LPG β-galactosylation substitution (Fig. 7 and Table I, LPG side chain profile), the differences in

[3] K. Robinson and S. M. Beverley, submitted for publication.

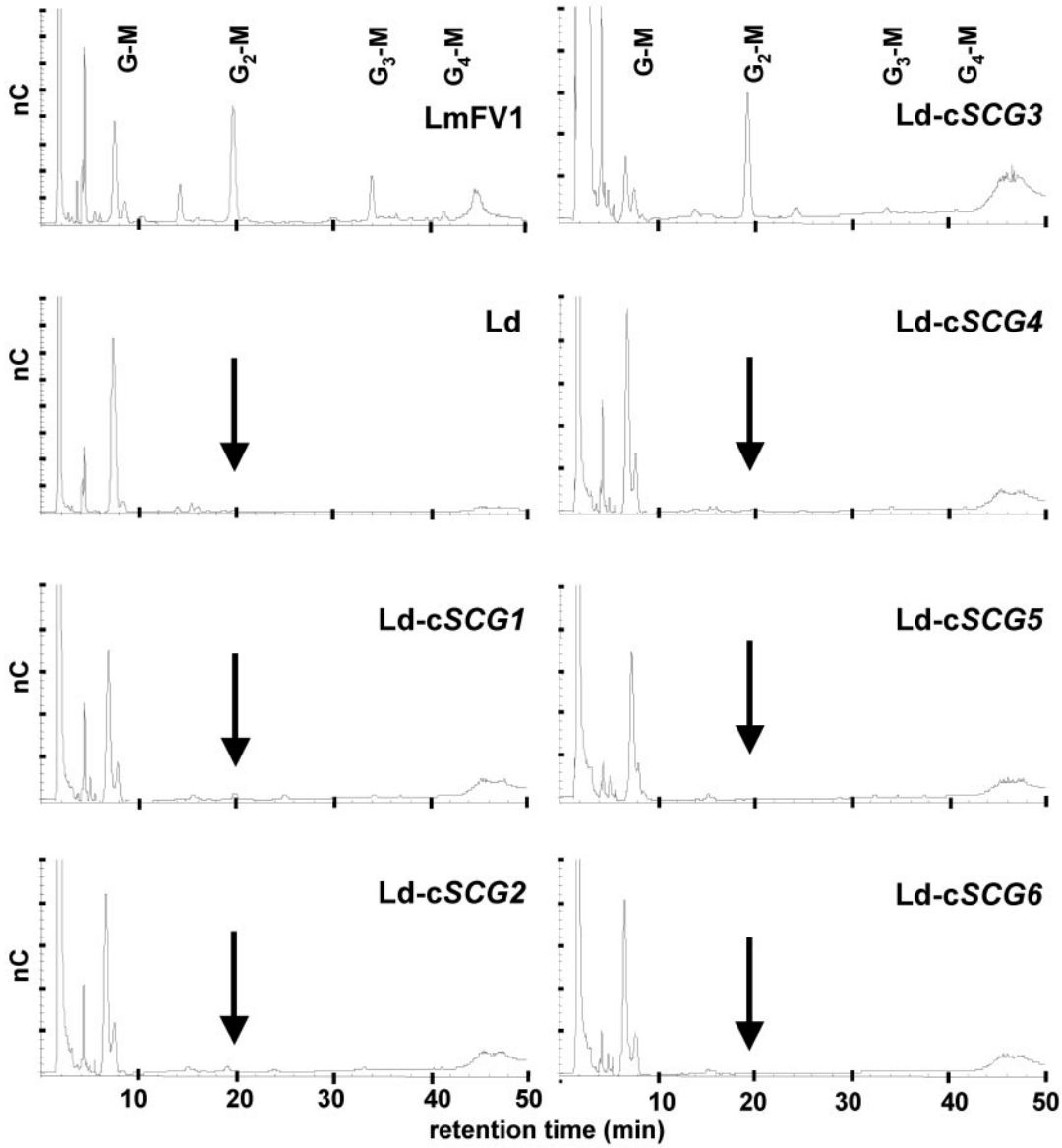

FIG. 6. **Effect of *SCG* expression in *L. donovani* on LPG structure.** PG repeating units from LPG isolated from untransfected *L. donovani* (*Ld*), Ld *SCG*-transfectants, or LmFV1 were fractionated by Dionex HPLC. The elution positions of unsubstituted PG repeats (*G-M*) and PG repeats containing one (G_2-*M*), two (G_3-*M*), or three (G_4-*M*) Gal side chains are marked. An *arrow* denotes the position of PG repeats containing a single Gal side chain (G_2-*M*). The abundance of each PG repeat species was quantitated by ED40 electrochemical detection (*nC*). Unlabeled peaks are derived from the LPG cap (Fig. 1). Two independent experiments gave similar results.

side chain length must reflect inherent differences in the specificity of these two LPG scβGalTs.

DISCUSSION

For *L. major* Friedlin V1 parasites, the β1,3-Gal side chains on the prominent LPG PG repeats (Fig. 1) are recognized by lectins in the *P. papatsi* midgut, and participate in parasite binding and survival in its sand fly host. We exploited a species-specific LPG structural polymorphism recognized by WIC79.3 antibodies (24) to isolate the LmFV1 *SCG* gene family that mediates addition of βGal side chains to LPG PG repeats. Three *SCG* genes were identified using functional genetic complementation of WIC79.3-negative *L. donovani* parasites (*SCG1*, *-2*, *-4*; Figs. 1–3), and the SCG family was completed by homology/library screening (*SCG3*, *-5*, *-6*; Fig. 3). This functional approach has great potential for identifying genes involved in other species-, strain-, or stage-specific LPG structural polymorphisms that are distinguishable by lectin or antibody reactivity. For example, we have adapted this protocol to recover a candidate LPG PG side chain capping β1,2-arabinosyltransferase.[2]

Although *SCG* genes were dispersed on distinct chromosomes (Fig. 3*B*), they exhibited 82–96% overall amino acid identity (Fig. 4, *A* and *B*). The predicted SCG proteins show features expected for eukaryotic βGalTs, including a short conserved GalT region including a potential "DDD" catalytic motif, and have a type II membrane protein topology with a large lumenal domain (Fig. 4, *A* and *C*). Functional evidence that many SCGs encode PG side chain βGalTs comes from enzymatic assays of Ld *SCG* transfectants, as some *SCGs* conferred synthesis of scβGal-modified LPG PG repeats (Figs. 6–8), and showed elevated LPG scβGalT activity when expressed in *L. donovani* (Table I) and when expressed using baculovirus vectors in insect cells.[4] The SCG proteins are most likely asso-

[4] D. R. Sullivan, S. M. Beverley, and S. J. Turco, manuscript in preparation.

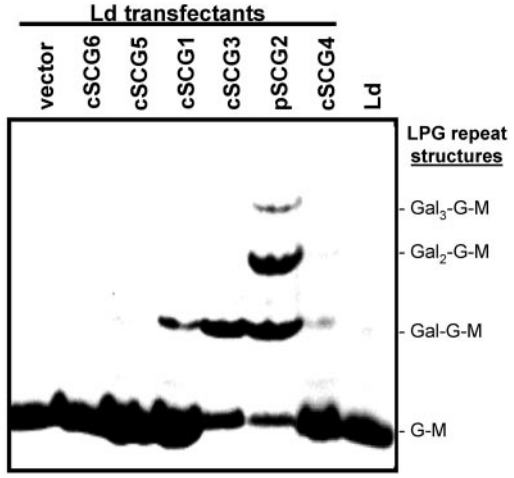

FIG. 7. **Electrophoretic analysis of LPG repeat units.** Dephosphorylated PG repeat units from Ld and transfectants (*vector*, cLHYG; *c*, cosmid; p*SCG2*, pXK-*SCG2*) were fluorophore-labeled, separated by electrophoresis, and visualized with UV light. LPG repeat side chain structures corresponding to each major band are noted, with *G-M* = unsubstituted PG repeat. Two independent experiments gave similar results.

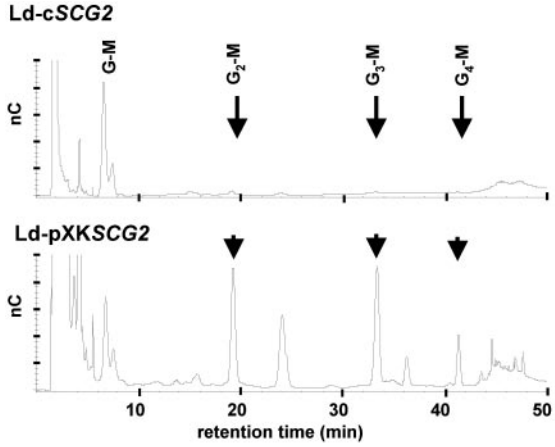

FIG. 8. **Context-dependent activity of *SCG2*.** LPG PG repeat units isolated from *SCG2* cosmid (Ld-c*SCG2*) and pXK-*SCG2* plasmid (Ld-pXK*SCG2*) Ld transfectants were analyzed by Dionex HPLC as described in the legend to Fig. 6. Two independent experiments gave similar results.

ciated with the parasite Golgi apparatus, because other LPG proteins involved in PG synthesis and side chain modification are found in this compartment (17).[2]

Transfection of the *SCG3* cosmid into *L. donovani* resulted in the synthesis of an LPG bearing predominantly single βGal modifications on PG repeats, whereas the remaining *SCG* cosmid Ld transfectants did not show significant levels of activity, as monitored by PG repeat structures characterized by HPLC or fluorophore labeling and electrophoresis (Figs. 6 and 7). However, expression of the *SCG2* coding region alone using a standard *Leishmania* expression vector approach yielded transfectants that synthesized an LPG with abundant PG side chain β-galactosylation. This suggests that expression of the individual *SCG* genes is context-dependent, which is discussed more extensively below. Notably, whereas the overall degree of PG repeat galactosylation was similar in the transfectants bearing cosmid *SCG3* or the *SCG2* expression vector, the pattern differed markedly (Figs. 6–8, Table I). The LPG synthesized by Ld c*SCG3* transfectants bore predominantly a single βGal side chain, similar to that of the parental *L. major* Friedlin V1 line, and thus displayed initiating scβGalT activity. In contrast, LPG synthesized by the pXK-*SCG2* Ld transfectants contained PG side chains of 1–3 βGal residues, and thus displayed both initiating and elongating scβGalT activities. Because our ability to detect PG polygalactosylation is compromised by low levels of activity (as shown in comparisons of *SCG2* cosmid *versus* expression vector patterns; Fig. 8), determination of the exact specificities of *SCG1* and *SCG4* will require similar ORF expression vector tests. These studies, however, establish for the first time the presence of at least two classes of LPG scβGalTs in the *Leishmania* genome.

Unexpectedly, two of the six *SCG* genes were inactive in the sensitive Ld transfection assay and synthesized no detectable βGal-modified LPG PG repeats (*SCG5* and *-6*; Table I, Fig. 6). We can exclude the possibility that these mediated transfer of sugars other than Gal by the fluorophore labeling experiment (Fig. 7). Although the molecular basis underlying this heterogeneity was not pursued, the SCG ORFs and flanking regions showed numerous polymorphisms, which may contribute to differences in mRNA and/or protein expression, activity, specificity, or regulation. For example, the C-terminal region of *SCG5* is highly divergent from the other *SCGs*, and the predicted transmembrane domain of *SCG6* contains a charged aspartyl residue at position 113. Another contributing factor may be the use of a heterologous *Leishmania* species for these assays. Whereas in general *Leishmania* signals are recognized across species, some species specificity has been found (41). However, preliminary studies where *L. major* strains or mutants lacking Gal-modified LPG were transfected with the *SCG* cosmids have yielded similar results to those obtained with Ld.[5] An attractive possibility is that *SCG5* and SCG6 (and possibly other SCGs) lack significant initiating LPG scβGalT activity, and possess only elongating scβGalT activity. Tests of this model will require expression of these gene products simultaneously with an initiating scβGalT such as *SCG3*.

Why does the *Leishmania major* genome encode so many LPG scβGalT genes, given that expression of *SCG3* alone in *L. donovani* is apparently sufficient to generate an LPG side chain Gal modification pattern similar to that of the parental *L. major* Friedlin V1 strain (Fig. 6)? One possibility invokes differences in developmental expression. Whereas gene specific probes for *SCG2*, *-3*, *-4*, and *-6* are unavailable, preliminary data suggest this may be the case for *SCG1* and *SCG5*.[5] Interestingly, *L. major* amastigote LPG (which is present at very low levels) contains long polymeric Gal side chains (42, 43), which would require both initiating and elongating LPG scβGalT activities as seen for *SCG2*. Alternatively, perhaps different *SCG* scβGalTs show differences in PG acceptor specificity, which include LPG- and PG-modified proteins such as PPG and secreted acid phosphatase (which is a substrate for PG galactosylation when expressed in *L. major*) (44, 45). However, preliminary Western blot analyses do not provide support for differential modification of LPG relative to other PGs in the Ld transfectants studied here (data not shown).

A second class of models (not exclusive from those above) considers the role of LPG PG repeat side chain modifications and polymorphisms in parasite biology. Significant variation exists in the degree of LPG side chain galactosylation in *L. major*; whereas the Freidlin V1 strain shows primarily single Gal LPG side chains, strain LV39 clone 5 synthesizes LPG bearing polygalactosyl modifications,[2] and the Seidman strain lacks Gal modifications (46). Notably, LPG side chain galactosylation is associated with the ability of *L. major* to survive in its sand fly vector *P. papatasi* (10, 47). Thus, intraspecific LPG

[5] D. E. Dobson, L. D. Scholtes, P. Myler, S. J. Turco, and S. M. Beverley, unpublished observations.

polymorphic modifications play important roles in parasite-sand fly interactions, as seen previously for inter-specific LPG differences (3, 4).

We propose that strain-specific patterns in LPG side chain β-galactosylation depend on the pattern of expression and specificity of *SCG* genes, which in turn play key roles in sand fly survival. Thus in the Friedlin V1 strain only initiating LPG scβGalTs are expressed (such as that encoded by *SCG3*), while in the LV39 strain both initiating and elongating LPG scβGalTs are expressed (such as that encoded by *SCG2*), and Seidman does not express any SCG activity. Because the number of *SCG* genes appears to be comparable in most *L. major* strains,[5] there must be mechanisms for regulating the expression and/or activity of the *SCG* gene repertoire during evolution (supported by the context-dependent activity of *SCG2*). Note that this model does not necessarily imply expression of only one *SCG* gene at a time, only that those expressed collectively yield the final LPG side chain β-galactosylation pattern.

Closely related strains of the other *Leishmania* species also show differences in LPG side chain modification. In *Leishmania tropica* and *Leishmania aethiopica*, more than 10 different patterns have been found involving PG repeat side chain modifications other than galactose (48) and the Indian strains of *L. donovani* bear sc-glucosyl modifications similar to those described in *L. mexicana* (30). Whereas the gene(s) and mechanisms responsible for LPG modifications in these species have not been identified, it seems likely that variation in the expression of the relevant glycosyltransferases is responsible, as proposed here for *L. major* LPG scβGalTs.

In many respects the model for differential LPG side chain modifications during evolution is reminiscent of other systems of antigenic variation in microorganisms such as *Trypanosoma brucei*, *Plasmodium falciparum*, or *Borrelia* (49–51). A significant difference is that in these organisms variation concerns expression of a family of surface protein antigens, rather than glycosyltransferases responsible for glycocalyx synthesis, and that variation is induced in response to antigenic pressure. In *Leishmania*, it seems more likely that variation occurs because of changes in the sand fly population, perhaps in response to selective pressures exerted by *Leishmania* or other microbes, or because of transport of *Leishmania* into regions with differing sand fly fauna by mobile vertebrate hosts. However, a role for species-specific PG modifications in vertebrate infections has also been suggested (52). A unique feature of LPG PG modifications is that by simultaneously expressing glycosyltransferases with differing specificities, combinatorial diversity can be generated. In the microbial antigenic variations systems, members of the surface antigen family are encoded at many places within the genome (as for the *SCGs*), and expression is controlled by mechanisms leading to the exclusive transcription of a single active mRNA encoding each antigen. How this is accomplished in *Leishmania* remains to be determined, as context-dependent expression of *SCG2* activity could be controlled at the levels of transcript abundance and/or translation. The molecular, biochemical, and biological predictions of this model are testable, and future studies will focus on the mechanism(s) of differential regulation and activity of *SCG* expression.

Acknowledgments—We thank Ashley Bray Mahoney, David Sacks, and members of the Beverley laboratory for discussions.

REFERENCES

1. Ashford, R. W., Desjeux, P., and deRaadt, P. (1992) *Parasitol. Today* **8,** 104–105
2. Ilg, T. (2001) *Med. Microbiol. Immunol.* (*Berl.*) **190,** 13–17
3. Sacks, D., and Kamhawi, S. (2001) *Annu. Rev. Microbiol.* **55,** 453–483
4. Sacks, D. L. (2001) *Cell Microbiol.* **3,** 189–196
5. Pimenta, P. F., Turco, S. J., McConville, M. J., Lawyer, P. G., Perkins, P. V., and Sacks, D. L. (1992) *Science* **256,** 1812–1815
6. McConville, M. J., Turco, S. J., Ferguson, M. A., and Sacks, D. L. (1992) *EMBO J.* **11,** 3593–3600
7. Sacks, D. L., Brodin, T. N., and Turco, S. J. (1990) *Mol. Biochem. Parasitol.* **42,** 225–233
8. Sacks, D. L. (1989) *Exp. Parasitol.* **69,** 100–103
9. Sacks, D. L., Pimenta, P. F., McConville, M. J., Schneider, P., and Turco, S. J. (1995) *J. Exp. Med.* **181,** 685–697
10. Butcher, B. A., Turco, S. J., Hilty, B. A., Pimenta, P. F., Panunzio, M., and Sacks, D. L. (1996) *J. Biol. Chem.* **271,** 20573–20579
11. Pimenta, P. F., Saraiva, E. M., Rowton, E., Modi, G. B., Garraway, L. A., Beverley, S. M., Turco, S. J., and Sacks, D. L. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 9155–9156
12. Roditi, I., and Liniger, M. (2002) *Trends Microbiol.* **10,** 128–134
13. Mulvey, M. A., and Hultgren, S. J. (2000) *Adhesion, Bacterial, Encyclopedia of Microbiology*, Second Ed., Vol. 1, Academic Press, New York
14. Sundstrom, P. (1999) *Curr. Opin. Microbiol.* **2,** 353–357
15. Dillon, R. J., and Lane, R. P. (1999) *Parasitology* **118,** 27–32
16. Beverley, S. M., and Turco, S. J. (1998) *Trends Microbiol.* **6,** 35–40
17. Descoteaux, A., Luo, Y., Turco, S. J., and Beverley, S. M. (1995) *Science* **269,** 1869–1872
18. Ryan, K. A., Garraway, L. A., Descoteaux, A., Turco, S. J., and Beverley, S. M. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 8609–8613
19. Descoteaux, A., Avila, H. A., Zhang, K., Turco, S. J., and Beverley, S. M. (2002) *EMBO J.* **21,** 4458–4469
20. Kapler, G. M., Coburn, C. M., and Beverley, S. M. (1990) *Mol. Cell. Biol.* **10,** 1084–1094
21. Sacks, D. L., Hieny, S., and Sher, A. (1985) *J. Immunol.* **135,** 564–569
22. Titus, R. G., Gueiros-Filho, F. J., de Freitas, L. A., and Beverley, S. M. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 10267–10271
23. Ryan, K. A., Dasgupta, S., and Beverley, S. M. (1993) *Gene* (*Amst.*) **131,** 145–150
24. Kelleher, M., Curtis, J. M., Sacks, D. L., Handman, E., and Bacic, A. (1994) *Mol. Biochem. Parasitol.* **66,** 187–200
25. Callahan, H. L., and Beverley, S. M. (1991) *J. Biol. Chem.* **266,** 18427–18430
26. Tosi, L. R., and Beverley, S. M. (2000) *Nucleic Acids Res.* **28,** 784–790
27. Cunningham, M. L., and Beverley, S. M. (2001) *Mol. Biochem. Parasitol.* **113,** 199–213
28. Beverley, S. M. (1988) *Nucleic Acids Res.* **16,** 925–938
29. Orlandi, P. A., Jr., and Turco, S. J. (1987) *J. Biol. Chem.* **262,** 10384–10391
30. Mahoney, A. B., Sacks, D. L., Saraiva, E., Modi, G., and Turco, S. J. (1999) *Biochemistry* **38,** 9813–9823
31. Ng, K. H. E. B. A. (1996) *Biochem. J.* **317,** 247–255
32. Wiggins, C. A., and Munro, S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 7945–7950
33. Singer, S. J. (1990) *Annu. Rev. Cell Biol.* **6,** 247–296
34. Al-Qahtani, A., Teilhet, M., and Mensa-Wilmot, K. (1998) *Biochem. J.* **331,** 521–529
35. Clayton, C. E. (1992) *Prog. Nucleic Acids Res. Mol. Biol.* **43,** 37–66
36. Ilg, T., Stierhof, Y. D., Craik, D., Simpson, R., Handman, E., and Bacic, A. (1996) *J. Biol. Chem.* **271,** 21583–21596
37. Thomas, J. R., McConville, M. J., Thomas-Oates, J. E., Homans, S. W., Ferguson, M. A., Gorin, P. A., Greis, K. D., and Turco, S. J. (1992) *J. Biol. Chem.* **267,** 6829–6833
38. Beverley, S. M. (2003) *Nat. Rev. Genet.* **4,** 11–19
39. Pedrosa, A. L., and Cruz, A. K. (2002) *Mol. Biochem. Parasitol.* **122,** 141–148
40. Ha, D. S., Schwarz, J. K., Turco, S. J., and Beverley, S. M. (1996) *Mol. Biochem. Parasitol.* **77,** 57–64
41. LeBowitz, J. H., Coburn, C. M., McMahon-Pratt, D., and Beverley, S. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 9736–9740
42. Moody, S. F., Handman, E., McConville, M. J., and Bacic, A. (1993) *J. Biol. Chem.* **268,** 18457–18466
43. Turco, S. J., and Sacks, D. L. (1991) *Mol. Biochem. Parasitol.* **45,** 91–99
44. Wiese, M., Gorcke, I., and Overath, P. (1999) *Mol. Biochem. Parasitol.* **102,** 325–329
45. Späth, G. F., Epstein, L., Leader, B., Singer, S. M., Avila, H. A., Turco, S. J., and Beverley, S. M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9258–9263
46. Sacks, D. L., and da Silva, R. P. (1987) *J. Immunol.* **139,** 3099–3106
47. Joshi, P. B., Sacks, D. L., Modi, G., and McMaster, W. R. (1998) *Mol. Microbiol.* **27,** 519–530
48. McConville, M. J., Schnur, L. F., Jaffe, C., and Schneider, P. (1995) *Biochem. J.* **310,** 807–818
49. Barbour, A. G., and Restrepo, B. I. (2000) *Emerg. Infect. Dis.* **6,** 449–457
50. Kyes, S., Horrocks, P., and Newbold, C. (2001) *Annu. Rev. Microbiol.* **55,** 673–707
51. Barry, J. D., and McCulloch, R. (2001) *Adv. Parasitol.* **49,** 1–70
52. Pelletier, I., and Sato, S. (2002) *J. Biol. Chem.* **277,** 17663–17670