Stage-specific Proteophosphoglycan from *Leishmania mexicana* Amastigotes

STRUCTURAL CHARACTERIZATION OF NOVEL MONO-, DI-, AND TRIPHOSPHORYLATED PHOSPHODIESTER-LINKED OLIGOSACCHARIDES*

(Received for publication, November 6, 1997, and in revised form, March 3, 1998)

Intracellular amastigotes of the protozoan parasite *Leishmania mexicana* secrete a macromolecular proteophosphoglycan (aPPG) into the phagolysosome of their host cell, the mammalian macrophage. The structures of aPPG glycans were analyzed by a combination of high pH anion exchange high pressure liquid chromatography, gas chromatography-mass spectrometry, enzymatic digestions, electrospray-mass spectrometry as well as 1H and 31P NMR spectroscopy. Some glycans are identical to oligosaccharides known from *Leishmania mexicana* promastigote lipophosphoglycan and secreted acid phosphatase. However, the majority of the aPPG glycans represent amastigote stage-specific and novel structures. These include neutral glycans ([Glcβ1–3]1–Galβ1–4Man, Galβ1–3Galβ1–4Man, Galβ1–3Glcβ1–4Man, PO4-6-Galβ1–4Man), several monophosphorylated glycans containing the conserved phosphodisaccharide backbone (R-3-[PO4-6-Galβ1–4Man]–), and monophosphorylated aPPG tri- and tetrasaccharides that are uniquely phosphorylated on the terminal hexose (PO4-6-Galβ1–3Glcβ1–4Man, PO4-6-Glcβ1–3Glcβ1–4Man, PO4-6-Galβ1–3Glcβ1–4Man). In addition aPPG contains highly unusual di- and triphosphorylated glycans whose major species are PO4-6-Glcβ1–3Glcβ1–3PO4-6-Galβ1–4Man, PO4-6-Galβ1–3Glcβ1–3PO4-6-Galβ1–4Man, PO4-6-Galβ1–3Glcβ1–3PO4-6-Galβ1–4Man, PO4-6-Galβ1–3Glcβ1–3PO4-6-Galβ1–4Man, PO4-6-Glca1–3PO4-6-Galβ1–4Man, and PO4-6-Glca1–3PO4-6-Galβ1–4Man. These glycans are linked together by the conserved phosphodiester-R-Manα1-PO4-6-Gal-R or the novel phosphodiester-R-Manα1-PO4-6-Glc-R and are connected to Ser(P) of the protein backbone most likely via the linkage R-Manα1-PO4-6-Gal-Ser. The variety of stage-specific glycan structures in *Leishmania mexicana* aPPG suggests the presence of developmentally regulated amastigote glycosyltransferases which may be potential anti-parasite drug targets.

Protozoan parasites of the genus *Leishmania* are the causative agent of a spectrum of human diseases. *Leishmania* have a digenetic life cycle that encompasses the extracellular promastigote in the digestive tract of the parasite-transmitting insect vector, the sandfly, and the disease-causing intracellular amastigotes living in parasitophorous vacuoles of mammalian macrophages.

The parasites produce unusual glycoconjugates, which are thought to play crucial roles for survival, development, and virulence in both developmental stages of the parasite. The best characterized *Leishmania* glycoconjugate is the promastigote cell-surface glycolipid, lipophosphoglycan (LPG). The structure of LPG from five different *Leishmania* species has been determined. LPG contains a conserved *lyso*-alkylphosphatidylinositol membrane anchor linked to a phosphohexasaccharide core structure, a conserved backbone of up to 40 phosphodiester-linked disaccharide repeats ([PO4-6-Galβ1–4Manα1-]), and species-, strain-, and stage-specific components linked to the core and repeats as well as terminating neutral (cap) glycans at the non-reducing end of the molecule (reviewed in Refs. 1 and 2). In the sandfly, LPG serves as a ligand for the attachment of non-infectious procyclic promastigotes to the midgut wall lining and may protect the parasites against the hydrolytic environment of the insect’s digestive tract. A stage-specific form of LPG confers complement resistance to the highly infectious metacyclic promastigotes, which are injected by the sandfly into the skin of the mammalian host. LPG also acts as a receptor for the invasion of macrophages by metacyclic promastigotes and may protect this transient mammalian parasite stage against the initial microbicidal response of the host cell by acting as a radical scavenger and by modulating signal transduction and gene expression of the macrophage (reviewed in Refs. 3–5). Structure-function analysis of LPG and its fragments demonstrated the importance of the glycolipid anchor, the phosphoglycan chains, and the cap oligosaccharides for these functions (6–13). The biosynthetic pathway of LPG has been partially elucidated and has been implicated as a poten-

---

* This work was supported by a fellowship of the Deutsche Forschungsgemeinschaft (to T. I.). 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: Max-Planck-Institut für Biologie, Abteilung Membranbiochemie, Corrensstrasse 38, D-72076 Tübingen, Germany. Fax: 49-7071-62971; E-mail: thomas. ilg@tuebingen.mpg.de.

The abbreviations used are: LPG, lipophosphoglycan; PG, phosphoglycan; pPPG, promastigote proteophosphoglycan; aPPG, amastigote proteophosphoglycan; SAP, secreted acid phosphatase; GC-MS, combined gas chromatography-mass spectrometry; SABG, sweet almond aglycone; BTBG, bovine testes β-glucosidase; AP, calf intestine alkaline phosphatase; HPAGE-HPLC, high pH anion exchange HPLC; TFA, trifluoroacetic acid; ELISA, enzyme-linked immunosorbent assay; HPF, hydrofluoric acid; ES-MS, electrospray ionization mass spectrometry; TOCSY, total correlation spectroscopy; HMBC, heteronuclear multiple bond correlation.
tial target for the development of chemotherapeutic agents (14–20).

Although the crucial role of LPG for the promastigote stages of *Leishmania* is well established, its importance for the disease-causing amastigote stage in the mammalian host is less clear. *Leishmania donovani* and *Leishmania mexicana* amastigotes do not express LPG (21, 22). *Leishmania major* amastigotes synthesize low levels of a stage-specific LPG (23–25), which may be involved in host cell binding and uptake (26), but amastigote LPG does not form a protective surface glyocalyx like in promastigotes (27).

It has been demonstrated that some of the biologically active structural elements of LPG like the repetitive phosphoglycans and the neutral cap oligosaccharides are also present on *Leishmania* promastigote proteins like acid phosphatase (sAP) (28–33) and the filamentous proteophosphoglycan of promastigotes (pPPG) (32, 34). In these molecules, the glycans are linked to the protein backbone via phosphoserine residues (33, 34), a form of protein glycosylation not yet observed in mammalian cells. Recently, it has been shown that protein-bound phosphoglycans are also present in the amastigote stage of *L. mexicana* (22, 35). This amastigote proteophosphoglycan (aPPG) is secreted by the parasites in large amounts into the phagolysosomes of host macrophages, where it may accumulate to mg/ml concentrations (35). The massive secretion of aPPG by amastigotes may contribute to the expansion of the phagolysosomes to huge parasitophorous vacuoles, which are the hallmark of *L. mexicana* infections (36). It has also been demonstrated that *L. mexicana* aPPG is an activator of the complement cascade via the lectin pathway. This property may contribute to lesion development and pathology caused by *L. mexicana* (37). A preliminary analysis of aPPG showed that it is immunologically and chemically related to the promastigote proteophosphocyan antigens LPG and sAP but also exhibits distinct properties (35).

In the present study we describe the structural analysis of the glycans from *L. mexicana* aPPG purified from infected mouse lesion tissue. We describe that, in addition to oligosaccharides also present in *L. mexicana* promastigote phagolysomes, and neutral and phosphorylated glycans. These glycans are linked to the protein backbone via phosphoserine residues. We also show that a large proportion of aPPG glycans is modified by two, three, or even four phosphate groups in diester linkages to other glycans.

**EXPERIMENTAL PROCEDURES**

**Parasites—*L. mexicana* promastigotes (strain MINYC/BZ/62/M379)** were grown in semi-defined medium 79 as described (38). *L. mexicana* amastigote-infected tissue was obtained from dorsal lesions of CBA mice infected 3–6 months previously with 5 × 10⁶ stationary phase promastigotes in the shaven rump at the base of the tail.

**Purification of *Phosphoglycan Antigens and Analytical Procedures—*L. mexicana* aPPG** was purified from infected mouse tissue as described earlier (35) with the modifications reported recently (36). Purification of *L. mexicana* LPG, sAP, PG, PPG, and L. major pPPG from promastigote culture supernatant was performed as described previously (33, 34, 39).

**Analytical Techniques—**Colorimetric carbohydrate, protein, and phosphate analysis, mild acid hydrolysis, 40% HF dephosphorylation, phosphoamino acid analysis, and SDS-polyacrylamide gel electrophoresis were performed as outlined previously (31, 33, 34, 39). Carbohydrate in *L. mexicana* aPPG was detected semiquantitatively by spotting 1-μl aliquots onto Silica Gel 60 plates (Merck, Darmstadt, FRG) followed by reaction with orcinol/H₂SO₄ (39). Two-site ELISA using the monoclonal antibody AP3 as trapping antibody was also performed as described before (35), except for using biotinylated AP3 followed by Extravidin coupled to calf intestine alkaline phosphatase (AP) (Sigma, Deisenhofen, FRG) or LT22 followed by goat anti-mouse IgG (γ-chain-specific) antibodies coupled to AP (Sigma) as detection systems. Combined gas chromatography-mass spectrometry (GC-MS) was performed using a Hewlett-Packard HP6895 GC-MS, fitted with either a 25-m × 0.3-mm CPSil low polarity column (Chrompack, Middleburg, The Netherlands) for trimethylsilyl derivatives and permethylated alditol acetates or a 25-m × 0.2-mm BPX-70 high polarity column (SGE, Ringwood, Australia) for alditol acetates and permethylated alditol acetates as described previously (34). Amino acid analysis and N-terminal protein sequencing were performed on automated systems (Applied Biosystems, models 420A and 477A, respectively) according to the manufacturer’s protocols.

**Monosaccharide and Methylolation Analysis—**Native and mild acid-hydrolyzed aPPG (10 μg) containing myo-inositol as an internal standard and monosaccharide standards were subjected to methanolysis, re-N-acetylated, and then fractionated by reversed-phase high-performance liquid chromatography (HPLC) using a 25-μm high performance LC (HPLC) column. The resulting neutral and phosphorylated monosaccharides were subjected to methanolysis (35). Neutral and phosphorylated monosaccharides in *L. mexicana* aPPG and *L. mexicana* pPPG, samples (10 μg) were hydrolyzed in 2 μl trifluoroacetic acid for 2 h at 105 °C. This treatment cleaves most glycosidic bonds involving pentoses and hexoses quantitatively (41), whereas the phosphomonoester bonds of hexose 6-phosphates are stable. The resulting neutral and phosphorylated monosaccharides were separated by HPLC using program 8 (see below). Pooled fractions were desalted by passage over AG50 × 12 (H⁺) and lyophilized; myo-inositol was added as an internal standard, and the phosphorylated glycans were dephosphorylated by AP treatment (see above). The resulting monosaccharides were reduced, acetylated, and analyzed as alditol acetates by GC-MS and quantitated relative to the internal standard. Methylation linkage analysis of dephosphorylated glycans (0.5–10 μg) was performed as described previously (25, 39, 43).

**High pH Anion Exchange HPLC—**Neutral and phosphorylated glycans released from *L. mexicana* aPPG by mild acid hydrolysis or 40% HF were separated by high-performance liquid chromatography on a Dionex BioLC carbohydrate analyzer (Dionex Corp. Sunnyvale, CA) using a Carbo-Pac PA1 column and pulsed amperometric detection using linear gradients of sodium acetate in 50 mM NaOH. Several gradient programs were used as follows: program 1, 0 mM for 6 min, raised to 50 mM over 18 min, to 125 mM over 7 min, and held at 125 mM for 14 min; program 2, program 1 followed a raise of the sodium acetate concentration to 175 mM over 30 min, held at 175 mM for 10 min, raised to 625 mM over 10 min, and held at 625 mM for 5 min; program 3, 0 mM for 6 min, raised to 50 mM over 18 min, to 125 mM over 15 min, and held at 125 mM for 15 min; program 4, 250 mM for 6 min, raised to 1000 mM over 1 min, and held at 1000 mM for 13 min; program 5, 250 mM for 6 min, raised to 385 mM over 1 min, held at 385 mM for 18 min, raised to 625 mM over 1 min, and held at 625 mM for 10 min; program 6, isocratic, 385 mM for 20 min; program 7, 385 mM for 15 min, raised to 625 mM over 30 min, and held at 625 mM for 10 min. For the separation of hexoses from hexose phosphates, the Carbo-Pac PA1 column was held at 100 mM NaOH for 10 min; NaOH and NaN₃ were then raised to 150 and 187.5 mM, respectively, over 2 min and held at that concentration for 18 min (program 8).

**Electrospray Ionization-Mass Spectrometry (ES-MS)**—Mass spectra of oligosaccharides were acquired on a Finnigan LCQ ES-MS. Samples were introduced into the electrospray source through a hydrodynamic injector with a 5-μl loop at a flow rate of 5 μl/min in either 25% methanol in H₂O for native oligosaccharides or 50% aqueous acetonitrile for permethylated oligosaccharides. Mass spectra were acquired both in the positive and the negative ion modes using the following conditions. The electrospray ionization source was set to 170 °C; the maximum trapping time was 500 ms; the capillary, tube lens, and needle voltages were 25, 50, and 4.5 V, respectively, and the number of microscans was set to 1. MS-MS scans were performed by triggering the ion of interest and ejecting all other ions outside a 3-atomic mass unit window centered around the parent ion. Collision energy was set such that the parent ion was attenuated between 95 and 99%. Data were collected as averages of four spectra.
**RESULTS**

**L. mexicana aPPG Glycans Are Linked to Phosphoserine of the Protein Backbone via Mild Acid-labile Phosphodiester Bonds—** L. mexicana aPPG purified from amastigote-infected mouse lesion tissue eluted on a Superose 6 gel filtration column as a broad peak between the 2000 and the 440 kDa markers as detected by phosphate and carbohydrate determination as well as two-site ELISA (Fig. 1A). On SDS-polyacrylamide gels aPPG from the D₂O solvent. Spectra were processed using an exponential line broadening function of 0.3 Hz. Chemical shifts are referenced to 2.2-dimethyl-2-silapentane-5-sulfonate at 0.00 ppm. One-dimensional ³¹P spectra were recorded with a spectral width of 5000 Hz, a pulse width of 5 μs (45°), and a relaxation delay between scans of 2 s. Typically 500 scans were acquired prior to Fourier transformation. Spectra were processed using an exponential line broadening function of 5–10 Hz. Chemical shifts were referenced to an external capillary of neat H₃PO₄ at 0.00 ppm. All two-dimensional homonuclear ¹H spectra were recorded in D₂O on Bruker ARX 500 or DMX 750 spectrometers for the intact aPPG, LPG, and aAP3 and LT22 in a two-site ELISA. Glycan was determined semi-quantitatively by reaction with orcinol/H₂SO₄ on TLC plates. The fraction size was 1 ml.

**NMR Spectroscopy—** NMR spectra were recorded in D₂O on Bruker ARX 500 or DMX 750 spectrometers for the intact aPPG, LPG, and aPPG purified from amastigote-infected L. mexicana infected mouse lesion tissue eluted on a Superose 6 gel filtration column. Fractions were monitored for phosphate and for reactivity to the mAbs AP3 and LT22 in a two-site ELISA. Glycan was determined semi-quantitatively by reaction with orcinol/H₂SO₄ on TLC plates. The fraction size was 1 ml.

**Fig. 1.** Superose 6 gel filtration chromatography of L. mexicana aPPG. Purified aPPG before (A) and after mild acid hydrolysis (B) was chromatographed on Superose 6 in 250 mM ammonium acetate. Fractions were monitored for phosphate and for reactivity to the mAbs AP3 and LT22 in a two-site ELISA. Glycan was determined semi-quantitatively by reaction with orcinol/H₂SO₄ on TLC plates. The dot size indicates the relative intensity of staining. The elution positions of dextran blue (2000 kDa), thyroglobulin (667 kDa), ferritin (440 kDa), transferrin (80 kDa), and the inclusion volume (Vᵢ; CTP) of the column are indicated. The fraction size was 1 ml.
was detected in positive fractions as a smear migrating above the 200-kDa marker protein (Fig. 2A, lane 1 and Fig. 2B, lanes 9–14). After mild acid hydrolysis, known to be selective for hexose 1-phosphate bonds (43, 44), the high molecular weight aPPG (Fig. 2A, lane 1) disappeared on SDS-polyacrylamide gels, and a series of polypeptide bands between 40 kDa and the gel front were detected (Fig. 2A, lane 2). In Superose 6 chromatography, the majority of the carbohydrate (95%) and phosphate (90%) of mild acid-treated aPPG was found near the inclusion volume (Vt) of the column with small amounts of phosphate and traces of glycans in a second earlier eluting peak (Fig. 1B, fractions 15–18). This peak contained the polypeptides released by mild acid (Fig. 2C, lanes 15–19), which were pooled and subjected to amino acid analysis and protein sequencing. The amino acid composition of the pooled polypeptides (2.8% Asp/Asn, 5.7% Glu/Gln, 28.9% Ser, 3.3% His, 1.2% Arg, 9.9% Thr, 9.7% Ala, 7.7% Pro, 1.6% Tyr, 4.9% Val, 2.1% Ile, 3.7% Leu, 6.9% Phe, and 0.6% Lys) was similar to the published composition of intact aPPG (35), which suggests that they correspond to the deglycosylated aPPG protein backbone. N-terminal sequencing of the pooled polypeptides and one of the Superose 6 fractions (Fig. 2C, fraction 16) showed in both analyses the peptide sequence NPIFXXD (where X indicates ambiguities). This indicates that despite the complex pattern on SDS-polyacrylamide gels (Fig. 2A, lanes 2 and 4 and Fig. 2C, lanes 15–19), the aPPG protein backbone may be formed by either one or several closely related polypeptide species. Phosphoamino acid analysis showed that only serine residues are phosphorylated (>25%). Phosphoserine in aPPG was resistant to mild acid deglycosylation and to AP treatment, whereas the consecutive application of both treatments led to the loss (>90%) of the phosphorylated amino acid. On SDS-polyacrylamide gels, the polypeptides obtained after mild acid deglycosylation of aPPG were readily visualized by the cationic dye Stains-all (Fig. 2A, lane 2) with an intense blue color, and after dephosphorylation no staining occurred (Fig. 2A, lane 3). Coomassie Blue staining of the same gel revealed the dephosphorylated polypeptides between 65 and 50 kDa apparent molecular mass (Fig. 2A, lane 5). Taken together the results indicated that the majority of aPPG glycans are linked to a serine-rich protein backbone via mild acid labile phosphodiester bonds to serine. This interpretation was corroborated by 31P NMR spectroscopy, which shows that phosphate is exclusively present in diester linkages in intact aPPG (see below).

**The Majority of the Mild Acid Labile L. mexicana aPPG Glycans Are Stage-specific**—The mild acid-released aPPG glycans (Fig. 1B, fractions 20–24) were separated by HPAE-HPLC under conditions which resolve neutral, monophosphorylated, and multiply phosphorylated glycans (Fig. 3C). Mild acid-released glycans of _L. mexicana_ promastigote LPG (Fig. 3A) and...
sAP (Fig. 3B) served as standards and were resolved under identical conditions. Whereas *L. mexicana* aPPG contains the entire set of cap glycans and monophosphorylated glycans previously identified in LPG and sAP, the majority of its oligosaccharides are amastigote stage-specific and not detected in the promastigote phosphoglycan antigens.

**Structural Analysis of the Neutral Cap Oligosaccharides and the HF-dephosphorylated Oligosaccharide Backbones of *L. mexicana* aPPG.—**The mild acid-released neutral aPPG cap oligosaccharides and the neutral glycans obtained by 40% HF dephosphorylation of intact aPPG (corresponding to the entire set of oligosaccharide backbones) were isolated for structural analysis by HPAE-HPLC (compare Fig. 3C (0–48 min) and Fig. 4A, respectively). Their structures were determined by monosaccharide analysis, methylation linkage analysis, ES-MS (Table I), exoglycosidase digests (Table I, compare also Fig. 4, B and C), and coelution with authentic standards (Fig. 3, A and B). The results are summarized in Table II; Man, the mannooligosaccharide series N2a-N6a ((Manα1–2)1–5Man), N2b (Galβ1–4Man), and N3c (Glcβ1–3Galβ1–4Man) are known components of *L. mexicana* promastigote LPG and sAP (33, 39). N3b (Galβ1–3Galβ1–4Man) and N4c (Glcβ1–3Glcβ1–3Galβ1–4Man) were previously identified in *L. major* and *Leishmania tropica* LPG (2, 43) but not in *L. mexicana* glycoconjugates. N4b, N5b-d, N6b, and N6c, however, represent completely novel glycan backbones. BTBG digestion (product N3c), SABG digestion (product N2b), ES-MS of permethylated samples, and methylation analysis of N4b resulted in the proposed structure Galβ1–3Glcβ1–3Galβ1–4Man (Tables I and II, Fig. 4). The same analysis of N5b and N5d suggested the structures Galβ1–3Glcβ1–3Glcβ1–3Galβ1–4Man and Glcβ1–3Glcβ1–3Glcβ1–3Galβ1–4Man, respectively. N5c may be either Glcβ1–3Galβ1–3Glcβ1–3Galβ1–4Man or Glcβ1–3Glcβ1–3Galβ1–3Galβ1–4Man, which cannot be distinguished by the methods used. A similar situation arises with N6b and N6c, which have most likely the structure Glcβ1–3Hexβ1–3Hexβ1–3Hexβ1–3Galβ1–4Man, where two of the hexoses are Glc and one is Gal (Tables I and II, Fig. 4). The dominant glycan backbone in aPPG is N4c followed by Man, N2b, N4b, and N3c. In contrast in the two promastigote phosphoglycan antigens *L. mexicana* LPG and sAP, N2b and N3c are the major glycans (70–90 mol %; Table II and Ref. 33). Stage-specific backbone structures (N3b, N4b-c, N5b-d, N6b and -c) not previously observed in either of the promastigote glycoconjugates form the majority of the aPPG glycans (>55 mol %).

**L. mexicana** aPPG Contains Gal-6-PO4 and Glc-6-PO4—Phosphohexose analysis of *L. mexicana* aPPG was performed using *L. mexicana* sAP as a reference compound. Qualitative analysis (GC-MS) showed that aPPG contains both Gal-6-PO4 and Glc-6-PO4, whereas sAP contains only Gal-6-PO4. None of the samples contained Man-6-PO4. To quantitate the degree of hexose phosphorylation in the two compounds, aPPG and sAP were hydrolyzed in 2% trifluoroacetic acid; neutral and phosphorylated monosaccharides were separated by HPAE-HPLC and their compositions were analyzed after AP digestion, reduction, and acetylation as alditol acetates by GC-MS (Fig. 5). *L. mexicana* sAP contained only Gal-PO4, and exhibited a Hex:Hex-PO4 ratio of 4.2:1. In contrast, *L. mexicana* aPPG contained both Gal-PO4 and Glc-PO4 (Fig. 5A) at a ratio of 1.4:1, and its Hex:Hex-PO4 ratio was 2.6:1.

**L. mexicana** aPPG Contains Conserved and Novel Monophosphorylated Glycans and Novel Di- and Triphosphorylated Oligosaccharides—HPAE-HPLC of mild acid-released aPPG glycans resulted in a variety of peaks in the region of the salt gradient where monophosphorylated glycans are expected to elute (Mono-P region, Fig. 3C). A further increase in salt concentration eluted two unresolved glycan peak areas (Di-P and Tri-/Tetra-P region, Fig. 3C) that are not observed in *L. mexicana* LPG and sAP (compare Fig. 3, A and B). All the aPPG phosphorylated glycans (Fig. 3C) were sensitive to AP, and the

---

2 The hexose fraction from 40% HF-treated intact aPPG contained exclusively mannose, whereas the respective fraction of mild acid-treated aPPG showed 83% Man, 6% Gal, and 11% Glc. The latter treatment can lead to some limited hydrolysis of glycosidic bonds in LPG with concomitant release of pentoses and hexoses (45), whereas 40% HF seems to be more selective for phosphate esters. Therefore, it appears likely that in intact aPPG, only Man of the hexose fraction is engaged in labile phosphodiester bonds via its anomeric hydroxyl group.

3 β-Glucosidase from sweet almonds cleaves terminal β-Glc residues. It can also cleave terminal β-1–3-galactosides (e.g. Galβ1–3Galβ1–4Man or Galβ1–3Galβ1–3Galβ1–4Man) but not Galβ1–4Man (Ref. 2 and this study).
corresponding dephosphorylation products coeluted with the neutral glycans N2b, N3b-c, N4b-c, N5b-d, and N6b-c (Fig. 6A). The structures of these glycan backbones were confirmed by negative ion ES-MS and methylation linkage analysis (Table II). The much higher complexity and the presence of novel structures in L. mexicana aPPG versus LPG were also apparent in the comparison of the ES-MS (M – H) pseudomolecular ions of mild acid hydrolysates. Whereas LPG gave rise only to the expected ions for Hex$_2$P and Hex$_3$P (not shown), aPPG showed, in addition, ions for phosphorylated tetra- and pentasaccharides (Hex$_4$P and Hex$_5$P, Fig. 7A). Surprisingly the most abundant molecular species in ES-MS of aPPG corresponded to novel diphosphorylated tetrasaccharides (Hex$_4$P$_2$, Fig. 7A). In addition dephosphorylated pentasaccharides (Hex$_5$P$_2$) and triphosphorylated pentasaccharides (Hex$_5$P$_3$) were detected (Fig. 7A).

**Table I**

<table>
<thead>
<tr>
<th>Glycan</th>
<th>Proposed structure</th>
<th>LPG relative abundance$^a$</th>
<th>aPPG relative abundance$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hex</td>
<td>Man1–2Man</td>
<td>6 (Glc + Man)</td>
<td>18.5 (Man)</td>
</tr>
<tr>
<td>N2a</td>
<td>Galβ1–4Man</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>N2b</td>
<td>Galβ1–4Man</td>
<td>65.5</td>
<td>12.5</td>
</tr>
<tr>
<td>N3a</td>
<td>(Man1–2),Man</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>N3b</td>
<td>Galβ1–3Galβ1–4Man</td>
<td>Trace</td>
<td>1.0</td>
</tr>
<tr>
<td>N3c</td>
<td>Glcβ1–3Galβ1–4Man</td>
<td>Trace</td>
<td>1.0</td>
</tr>
<tr>
<td>N4a</td>
<td>(Man1–2),Man</td>
<td>Trace</td>
<td>1.0</td>
</tr>
<tr>
<td>N4b</td>
<td>Galβ1–3Glcβ1–3Galβ1–4Man</td>
<td>10.5</td>
<td>36.0</td>
</tr>
<tr>
<td>N4c</td>
<td>Glcβ1–3Glcβ1–3Galβ1–4Man</td>
<td>Trace</td>
<td>0.9</td>
</tr>
<tr>
<td>N5a</td>
<td>(Man1–2),Man</td>
<td>Trace</td>
<td>0.9</td>
</tr>
<tr>
<td>N5b</td>
<td>Galβ1–3Glcβ1–3Glcβ1–3Galβ1–4Man</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>N5c</td>
<td>Glcβ1–3Hexβ1–3Galβ1–4Man</td>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>N6a</td>
<td>(Man1–2),Man</td>
<td>Trace</td>
<td>0.5</td>
</tr>
<tr>
<td>N6b</td>
<td>Glcβ1–3Hexβ1–3Galβ1–4Man</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>N6c</td>
<td>Glcβ1–3Hexβ1–3Galβ1–4Man</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

$^a$ Percent detector response (FAD) in HPAE-HPLC of mild acid-treated LPG and aPPG, respectively.

$^b$ One Hex is Glc, and one Hex is Gal.

$^c$ Two of the Hex are Glc, and one Hex is Gal.

Positive ion ES-MS, methylation analysis, and exoglycosidase digests of neutral aPPG cap oligosaccharides isolated after mild acid hydrolysis (Fig. 3C) and aPPG oligosaccharide backbones dephosphorylated by 40% HF treatment (Fig. 4A).

**Table II**

<table>
<thead>
<tr>
<th>Glycan</th>
<th>Proposed structure</th>
<th>LPG relative abundance$^a$</th>
<th>aPPG relative abundance$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hex</td>
<td>Man1–2Man</td>
<td>6 (Glc + Man)</td>
<td>18.5 (Man)</td>
</tr>
<tr>
<td>N2a</td>
<td>Galβ1–4Man</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>N2b</td>
<td>Galβ1–4Man</td>
<td>65.5</td>
<td>12.5</td>
</tr>
<tr>
<td>N3a</td>
<td>(Man1–2),Man</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>N3b</td>
<td>Galβ1–3Galβ1–4Man</td>
<td>Trace</td>
<td>1.0</td>
</tr>
<tr>
<td>N3c</td>
<td>Glcβ1–3Galβ1–4Man</td>
<td>Trace</td>
<td>1.0</td>
</tr>
<tr>
<td>N4a</td>
<td>(Man1–2),Man</td>
<td>Trace</td>
<td>1.0</td>
</tr>
<tr>
<td>N4b</td>
<td>Galβ1–3Glcβ1–3Galβ1–4Man</td>
<td>10.5</td>
<td>36.0</td>
</tr>
<tr>
<td>N4c</td>
<td>Glcβ1–3Glcβ1–3Galβ1–4Man</td>
<td>Trace</td>
<td>0.9</td>
</tr>
<tr>
<td>N5a</td>
<td>(Man1–2),Man</td>
<td>Trace</td>
<td>0.9</td>
</tr>
<tr>
<td>N5b</td>
<td>Galβ1–3Glcβ1–3Glcβ1–3Galβ1–4Man</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>N5c</td>
<td>Glcβ1–3Hexβ1–3Galβ1–4Man</td>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>N6a</td>
<td>(Man1–2),Man</td>
<td>Trace</td>
<td>0.5</td>
</tr>
<tr>
<td>N6b</td>
<td>Glcβ1–3Hexβ1–3Galβ1–4Man</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>N6c</td>
<td>Glcβ1–3Hexβ1–3Galβ1–4Man</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>
downfield shift (−0.06 ppm from H-4 in neutral glycans (43)) is also consistent with phosphorylation at C-6. The remaining anomeric signal at 4.68 ppm with a large coupling is assigned to the terminal β-Glc residue (33, 43). Remarkably some αPPG-specific monophosphorylated glycans exhibited novel phosphorylation on alternative positions of the same glycan as follows: P4c9 and P4c (Fig. 3C) both yielded N4c after dephosphorylation suggesting that they have the same neutral glycan backbone. ES-MS-MS sequencing of P4c (Table III, Fig. 8) resulted in a relative intensity of fragmentations very similar to that of L. major P4b (Table III, Fig. 8; Galβ1–3Galβ1–3(PO4-Gal)β1–4Man (43)), which indicates either the sequence Hex-Hex[PO4-Hex]Hex or Hex[PO4-Hex]Hex-Hex.5 Although P4c9 has the same fragmentation ions as P4c, their relative distribution was different. In particular a much stronger fragment ion 259 in ES-MS-MS experiments (25 versus 3%, Fig. 7B and Table III) is indicative of a terminal position of PO4-Hex in P4c9, either (PO4-Hex)Hex-Hex-Hex or Hex-Hex-Hex(PO4-Hex). The alternative sequences shown above are indistinguishable by MS-MS.5 However, the absence of Man-PO4 in αPPG (Fig. 5) taken together with methylation and exoglycosidase analysis and the finding that P4c9 is resistant to SABG, while P4c is degraded to P2,6 suggest the structures PO4-Glcβ1–3Glcβ1–3Galβ1–4Man for P4c9 and Glcβ1–3Glcβ1–3(PO4-Gal)β1–4Man for P4c. This interpretation was corroborated and extended by 1H NMR studies which established that the phosphorylation sites were at the C-6 of Gal in P4c9 and C-6 of Glc in P4c. This interpretation was corroborated and extended by 1H NMR studies which established that the phosphorylation sites were at the C-6 of Gal in P4c9 and C-6 of Glc in P4c9; the anomeric region of the NMR spectrum for each of these monophosphorylated tetrasaccharides is shown in Fig. 9. This region of the spectrum of P4c (Fig. 9A) is similar to that of P3c (whose shifts are marked on Fig. 9A) except for the presence of an additional anomeric signal ant to SABG, while P4c is degraded to P2.6 suggest the structures PO4-Glcβ1–3Glcβ1–3Galβ1–4Man for P4c9 and Glcβ1–3Glcβ1–3(PO4-Gal)β1–4Man for P4c. This interpretation was corroborated and extended by 1H NMR studies which established that the phosphorylation sites were at the C-6 of Gal in P4c9 and C-6 of Glc in P4c. This interpretation was corroborated and extended by 1H NMR studies which established that the phosphorylation sites were at the C-6 of Gal in P4c9 and C-6 of Glc in P4c9; the anomeric region of the NMR spectrum for each of these monophosphorylated tetrasaccharides is shown in Fig. 9. This region of the spectrum of P4c (Fig. 9A) is similar to that of P3c (whose shifts are marked on Fig. 9A) except for the presence of an additional anomeric signal

5 Attempts to resolve this question by reduction of P4c9 and P4c with NaBD4 (2 M NH3, 1 M NaBD4, 12 h, 4 °C) were unsuccessful, because the complex ES-MS-MS fragmentation patterns obtained after reduction suggested that phosphate migration had occurred under the alkaline conditions. Similar observations were made with the diphosphorylated glycans (T. Ilg and G. Currie, unpublished results). Therefore the assignment of the sequence relies on the fact that no Man-6-PO4 was detected in αPPG (see Fig. 5). Thus the hexose at the reducing end (i.e. Man) cannot be phosphorylated.

6 β-Glucosidase from sweet almonds contains low amounts of a contaminating acid phosphatase, which converts P4c9 and P4c into N4c. N4c is then rapidly degraded to N2b by the glycosidase. However, short term incubations of P4c9 result in the formation of P2, whereas no P2 formation is observed with P4c. Interestingly, P4c9 also seems to be more resistant to the contaminating phosphatase than P4c.
corresponding to the extra Glc residue (δ = 4.77 ppm, J = 7.5 Hz). Phosphorylation at C-6 of the Gal residue is confirmed from the characteristic shift of Gal H-4. By contrast, this peak is shifted upfield in P4c, indicating that Gal C-6 is not phosphorylated in this isomer. Phosphorylation in this case is at C-6 of a Glc residue as deduced by the appearance of two multiplets at 4.15 and 4.05 ppm, corresponding to the separate 6-CH2 protons of this Glc residue (Fig. 9B). In addition to the MS evidence noted above, phosphorylation at the terminal Glc is supported by a 0.02 ppm upfield shift for H-1 of the terminal Glc in P4c relative to P4c but no change in the H-1 shift of the nonphosphorylated penultimate Glc residue.

Of the minor glycans, the monophosphorylated pentasaccharide P5d (Table IV) resulted in N5d after dephosphorylation, whereas its ES-MS-MS spectrum was very similar to the spectrum of P5b from L. major LPG (Table III, Fig. 8, Galβ1–3Galβ1–3Galβ1–3(PO4-6-Gal)β1–4Man, (43)). Therefore we tentatively assign the phosphate to the Gal residue in P5d (Glcβ1–3Glcβ1–3Glcβ1–3(PO4-Gal)β1–4Man, Table IV). Another minor component, the monophosphorylated trisaccharide P3b (Fig. 3C), gave rise to N3b (Table II) after enzymatic dephosphorylation. Coelution with an authentic standard from L. major suggests that its structure may be Galβ1–3(PO4-6-Gal)β1–4Man. Dephosphorylation of P3c yielded N3c, the same product as obtained from P3c (not shown). However, although the ES-MS-MS spectra of P3c derived from either aPPG or sAP were very similar, the spectrum of P3c was quite distinct and showed an abundant fragment ion of 259 atomic mass units, indicative of a terminal position of PO4-Hex (Table III, Fig. 8). These results suggest that P3c is PO4-Glcβ1–3Galβ1–4Man, a phosphoisomer of P3c. For the monophosphorylated tetrasaccharide P4b, N4b was identified as glycan backbone (Table II). Its ES-MS-MS spectrum was similar to the spectrum of P4c which suggests the structure PO4-Galβ1–3Glcβ1–3Galβ1–4Man (Table III, Fig. 8). Two other peaks (eluting on a Carbo-Pac PA1 column between P4b and P4c, Fig. 3C) were identified as monophosphorylated tetrasaccharides by ES-MS (Table IV), but insufficient amounts of material precluded more detailed studies.

Structure of Novel Diphosphorylated aPPG Glycans—ES-MS of the unresolved aPPG glycans eluting in the first large peak under high salt conditions in HPAE-HPLC (Di-P, Fig. 3C) revealed two major ion species corresponding to the novel diphosphorylated glycans Hex4P2 and Hex5P2 (not shown, see also Fig. 7A). After dephosphorylation the neutral glycan backbones were purified by HPAE-HPLC (Fig. 6B) and analyzed by positive ion ES-MS of the permethylated glycans as well as by methylation linkage analysis. The major oligosaccharide corresponded to N4c, whereas minor glycan backbones were N4b, N5b, and N5d (Table II). Individual diphosphorylated glycans were separated by HPAE-HPLC under isocratic conditions. Three glycans were obtained in a purified form (>90%), Di-P4b, Di-P4c, and Di-P5b (Fig. 3D), whose dephosphorylation products corresponded to N4b, N4c, and N5b, respectively (Table II). The main fragmentation series in ES-MS-MS (Fig. 7C, Fig. 8, and Table III) together with the absence of Man-PO4 in the...
hexose phosphate analysis (Fig. 5B) suggested the structures PO₄-Galβ₁–3Glcβ₁–3(PO₄-Gal)β₁–4Man for Di-P₄b and PO₄-Galβ₁–3Glcβ₁–3(PO₄-Gal)β₁–4Man for Di-P₄c (Table IV). These results were confirmed and extended by ¹H NMR, ³¹P NMR, and HMBC spectroscopy on Di-P₄b and Di-P₄c (Fig. 9, C–F). Phosphorylations at C-6 of both the Gal residue and the terminal Glc in Di-P₄c are indicated by the characteristic downfield shift of Gal H-4 (4.25 ppm) and the two multiplet signals for the 6-CH₂ protons of the terminal Glc (Fig. 9C) at 4.05 and 4.18 ppm. These two multiplets simplify on ³¹P decoupling, confirming a Glc-6-CH₂-O-P linkage. The Di-P₄b glycan is different from Di-P₄c in that the terminal Glc is replaced by Gal. As expected, the characteristic Glc-6-CH₂-O-P proton signals are missing from the Di-P₄b spectrum (Fig. 9D). The HMBC spectra of these two diphosphorylated derivatives (Fig. 9, E and F), which show cross-peaks only for those protons which have a long range spin-spin coupling interaction with ³¹P nuclei (in this case three bond couplings, compare Ref. 34), confirm these proposed assignments. For Di-P₄b two signals are detected, with ³¹P shifts characteristic of monoester phosphate groups and ¹H shifts characteristic of Gal-6-CH₂. These signals occur at 4.02 and 3.98 ppm for the two Gal residues, each corresponding to two near-degenerate CH₂ chemical shifts, and this leads to two signals (each split by H-H coupling) as indicated in Fig. 9E. The Gal-6-CH₂ protons remain close to degenerate and yield a single cross-peak. The fact that a single ³¹P nucleus yields HMBC cross-peaks to two protons unequivocally confirms the phosphorylation site for the terminal Glc to be at 6-CH₂ rather than at one of the ring positions, which have only single protons.

It should be noted that the ES-MS-MS spectrum of both Di-P₄b and Di-P₄c showed also fragment ions of Hex₂P₂ (483 and 501 atomic mass units) (Table III). This suggests that the phosphoisomer PO₄-Hexβ₁–3(PO₄-Hex)β₁–3Hexβ₁–4Hex is also present in Di-P₄b and Di-P₄c. However, NMR spectroscopy does not indicate phosphate position heterogeneity in the two glycans (Fig. 9). If present, components phosphorylated at the penultimate Glc would readily be detected in the spectrum of the Di-P₄b fraction via signals in the region 4.00–4.20 ppm but none are present at a detection limit of <5%. Similarly, minor components phosphorylated at the penultimate Glc would be detected in the Di-P₄c fraction as slightly shifted Glc H-1 signals, and none are present at this detection limit. Although the presence of small amounts of the respective PO₄-Hexβ₁–3(PO₄-Hex)β₁–3Hexβ₁–4Hex compound cannot be completely ruled out in Di-P₄b and Di-P₄c, a more likely explanation for the fragment ions 483 and 501 atomic mass units could be intramolecular migration of PO₄ in ES-MS prior to fragmentation. Intramolecular PO₄ migration from Hex to Hex in oligosaccharides has been observed by us after incubation in mildly alkaline solutions.⁵ ⁷ The main fragmentation pattern of negative ion ES-MS-MS on Di-P₅b (Table III and Fig. 8) and the results of hexose phosphate analysis are consistent with the structure PO₄-Galβ₁–3Glcβ₁–3Glcβ₁–3(PO₄-Gal)β₁–4Man (Table IV). Since the ES-MS-MS spectrum of the molecular ion 987.5 atomic mass units of the pooled diphosphorylated glycans was very

² Intramolecular phosphate migration in the gas phase may also explain the occurrence of low amounts of the "unexpected" fragment ions 241 and 259 atomic mass units in negative ion ES-MS-MS of P₄b and P₅b of L. major LPG, P₃c of L. mexicana LPG and sAP, P₃c and P₄c of L. mexicana aPPG (Table III and Fig. 7), as well as the low amount of unexpected ions identified in all other samples investigated in this study (Table III and Fig. 7).
similar to that of Di-P5b (Table III), it can be assumed that Di-P5d has the structure PO₄-Glc₁⁻³Glc₁⁻³Glc₁⁻³(PO₄-Gal)₁⁻⁴Man (Table IV). Similarly as for Di-P4b and Di-P4c, low levels of ions indicative for terminal Hex₂P₂ (483 and 501 atomic mass units) and Hex₃P₂ (645 and 663 atomic mass units) were found in both spectra (Table III) that may have been formed by PO₄ migration in ES-MS prior to fragmentation (see above).

**Structure of Triphosphorylated Glycans and Evidence for Tetraphosphorylated Glycans**—The aPPG glycans of the second HPAE-HPLC high salt peak (Tri- and Tetra-P, Fig. 3C) were rechromatographed to remove residual diphosphorylated glycans and to resolve the triphosphorylated glycans (Tri-P, Fig. 3E) from other components. ES-MS on the pooled triphosphorylated glycans revealed three major ions corresponding to Hex⁴P₃ (905.8 atomic mass units), Hex⁵P₃ (1068.5 atomic mass units), and Hex⁶P₃ (1229.7 atomic mass units) (not shown, compare also Fig. 7). After dephosphorylation, N₄c, N₅b, N₅c, N₅d, N₆b, and N₆c (Table II) were detected in HPAE-HPLC (Fig. 6C), and their structures were confirmed by methylation analysis. The ES-MS-MS fragmentation pattern of Tri-P₄c from the native glycan mixture (Table III and Fig. 8) was consistent with the structure PO₄-Glcβ₁⁻³(PO₄-Glc)β₁⁻³(PO₄-Gal)β₁⁻⁴Man (Table IV). Partial purification of the pooled glycans yielded three fractions (fractions 1–3, Fig. 3E). Fraction 3 contained mainly Tri-P₅d and only small amounts of Tri-P₆b as shown by AP digestion followed by HPAE-HPLC (not shown). ES-MS-MS on Tri-P₅d (Table III and Fig. 8), the results of the hexose phosphate analysis and methylation analysis indicated the structure PO₄-Glcβ₁⁻³(PO₄-Glc)β₁⁻³(PO₄-Gal)β₁⁻⁴Man. ES-MS-MS spectra of the Hex₃P₃ molecular ion of the mixture of triphosphorylated glycans or fractions enriched for either Tri-P₅b or Tri-P₅c (relative abundance, fraction 2: Tri-P₅b : Tri-P₅c; fraction 1: Tri-P₅b : Tri-P₅c) were very similar to that of Tri-P₅d (Table III and Fig. 8), suggesting that the structures of Tri-P₅b and Tri-P₅c could be PO₄-Galβ₁⁻³(PO₄-Glc)β₁⁻³(PO₄-Gal)β₁⁻⁴Man and PO₄-Glcβ₁⁻³(PO₄-Gal)β₁⁻³(PO₄-Gal)β₁⁻⁴Man, respectively. The fragment ions for Hex₃P₂ (725 and 743 atomic mass units) that were present in all Tri-P₅ spectra at low abundance (Table III) were most likely PO₄ migration products as discussed above for the Di-P glycans.\(^5\)\(^7\)

Tri-P₆b (highly enriched in fraction 1) showed an ES-MS-MS spectrum (Table III and Fig. 8) suggestive of either the sequence PO₄-Hex-Hex-(PO₄-Hex)Hex-(PO₄-Hex)Hex or PO₄-Hex-Hex-(PO₄-Hex)Hex-(PO₄-Hex)Hex-(PO₄-Hex)Hex (Table IV), which are indistinguishable by MS. ES-MS-MS on a mixture of Tri-P₆b and Tri-P₆c oligosaccharides yielded fragment ions of the same type and similar abundance as Tri-P₆b alone, which suggests a
similar distribution of the phosphates on the hexoses of the oligosaccharide chain in both compounds (Tables III and IV).

In HPAE-HPLC of the triphosphorylated glycans, an additional peak eluting later in the gradient was observed (Tetra-P, Fig. 3E). Negative ion ES-MS of this peak yielded ions indicative of tetraphosphorylated glycans. The ES-MS-MS fragmentation pattern of the two most abundant molecular species Hex₇P₄ (1471.5 amu) and Hex₈P₄ (1634.5 atomic mass units) is shown in Table III. Whereas the fragmentation of Hex₇P₄ is consistent with the structure P-Hex-Hex-(P-Hex)-(P-Hex)-Hex-(P-Hex)-Hex containing only phosphomonoesters (Table IV, compare Fig. 8), Hex₈P₄ is most likely composed of two diphasphorylated tetrasaccharides linked together by a phosphodiester bond (compare Fig. 8). The high abundance of the fragment ions 825 and 807 atomic mass units (Table III), which could arise from a preferred cleavage of the labile phosphodiester, is consistent with the proposed structure.

**FIG. 9.** NMR spectroscopy of phospho-oligosaccharides isolated from mild acid-hydrolyzed *L. mexicana* aPPG. Anomeric region of the ¹H NMR spectra of phosphorylated tetrasaccharides; A, P₄c; B, P₄c; C, Di-P₄c; D, Di-P₄b. The chemical shifts of corresponding signals in P₃c are shown as filled circles in A. E and F are HMBC spectra of Di-P₄c and Di-P₄b, respectively. The splitting of the two inequivalent CH₃ signals for Glc-6-PO₄ is indicated in E.

**L. mexicana Amastigote Proteophosphoglycan**
trum, which indicates that the multiple $^{31}$P nuclei present all have close to the same chemical environment. For aPPG the $^{31}$P NMR peak is much broader and contains a discernible shoulder, suggesting a superimposition of different $^{31}$P environments (Fig. 10B). This is confirmed in the HMBC spectrum (Fig. 10C); the advantage of this two-dimensional representation is that the various phosphorous environments are partially resolved by the chemical shift of their J-coupled protons. Three distinct $^{31}$P shifts (1.10, 1.18, and 1.35 ppm) are resolved based on non-overlapped proton shifts and correlate with the $^1$H shift of 6-CH$_2$-Gal (4.08 ppm), the $^1$H shift of H-1 of 2-Man$_1$PO$_4$ (5.68 ppm), and with the $^1$H shifts of the non-degenerate 6-CH$_2$-Glc protons (4.10 and 4.20 ppm), respectively. The strongest cross-peak in the HMBC spectrum has a single $^1$H chemical shift corresponding to the overlapped H-1 signals of 4-Man$_1$PO$_4$ and unsubstituted Man$_1$PO$_4$, but is split into two components of approximately equal intensity in the $^{31}$P dimension. As the two components have $^{31}$P shifts corresponding to -PO$_4$-6-Gal (1.10 ppm) and -PO$_4$-6-Glc (1.35 ppm), the most likely explanation for the splitting is due to the different $^{31}$P environments brought about by a linkage to the respective monosaccharide via the phosphodiester bond. The alternative explanation of the two $^{31}$P components reflecting differences on the other side of the phosphodiester bond (Man$_1$- versus 4-Man$_1$-) is less likely since the environmental differences between these two groups as far as the phosphodiester bond is concerned is very small, as demonstrated by their coincident $^1$H shifts. There was no evidence for the presence of PO$_4$-2-, whose position of the second phosphate residue is ambiguous. It could be located on either the third or the fourth Hex. NA: not applicable.

### DISCUSSION

Previous studies have shown that *Leishmania* promastigotes synthesize lipid-bound (LPG), free (PG), and protein-bound (sAP, pPPG) phosphoglycan antigens, which may play crucial roles for virulence and transmission of this parasite life stage (1, 3, 5, 46, 47). In the amastigote form, which causes disease in
the mammalian host, LPG expression is strongly down-regulated (23–25, 48), in most species to undetectable levels (21, 22). However, L. mexicana amastigotes do synthesize large amounts of a stage-specific proteophosphoglycan (aPPG) (22, 35). In this study we have elucidated the main structural features of this novel parasite antigen; aPPG consists of a defined polypeptide backbone, which is modified by a variety of carbohydrate structures via Ser(P) residues. We demonstrate that aPPG contains all glycans previously identified in L. mexicana promastigote phosphoglycan antigens LPG and sAP (33, 39), which include mannose, the manno-oligosaccharide series N2a-6a and N2b, P2 and P3c (Table IV). However, the majority of the glycans (Tables II and IV) have not been detected previously in L. mexicana (N3b, P3b, P4c) or represent completely novel structures (N4b, N4c, P4b, P4c, P5d, Di-P4b, Di-P4c, Di-P5b, Di-P5d, Tri-P4c, Tri-P5b, Tri-P5c, Tri-P5d, Tri-P6b, Tri-P6c, and Tetra-P7, Tables II and IV). Another surprising feature of aPPG is the presence of phosphoisomers of some monophosphorylated glycans (P3c versus P3c and P4c versus P4c) and the presence of novel multiphosphorylated glycans. These glycans are phosphorylated at the 6-position of either Gal or Glc residues, or both. To our knowledge, Glc-6-P has not been previously observed in glycoconjugates from any source. Neither the promastigote phosphoglycan antigens from L. mexicana (LPG and sAP, this study; PG and pPPG) nor from promastigotes of other Leishmania species (L. major LPG, PG, pPPG and L. donovani LPG, PG, sAP) contain the novel structure elements described above. The neutral and phosphorylated glycans are linked by phosphodiester bonds of the conserved structure R-Manα1-PO4-6-Gal-R and the newly identified linkage R-Manα1-PO4-6-Glc-R (compare Fig. 11). Taken together, our results suggest that the aPPG glycans

FIG. 10. NMR spectroscopy of native L. mexicana aPPG and LPG. A, anomeric region of the 500 MHz 1H NMR spectrum of aPPG at 288 K. The 1H NMR spectrum was recorded at 288 K to avoid overlap of the residual solvent signal with the anomeric protons marked I. The assignments for well resolved peaks are indicated. The peak envelopes I–III contain many overlapping signals, including the following: I, t-Glcβ1-3-(33), t-Galβ1-3- (34, 43), and based on the shifts for the isolated glycans in Fig. 9, 3-Glcβ1-3-, 3-[PO4-6-Glcβ1-3-], PO4-6-Galβ1-3-, and PO4-6-Glcβ1-3-; II, 3-Galβ1-4 and 3-[PO4-6-Galβ1-4-]; III, PO4-6-Galβ1-4-, t-Galβ1-4-. B, 31P NMR spectrum of aPPG at 298 K. The lower trace is the corresponding 31P spectrum of L. mexicana LPG; C and D are HMBC spectra of aPPG and LPG, respectively.

FIG. 11. Schematic model for L. mexicana aPPG. The model depicts possible structural arrangements of some of the aPPG oligosaccharides listed in Table IV. On average, each glycan chain is composed of 6 phospho-oligosaccharides and 4 terminating cap structures. The glycan chains are linked to the protein core via Ser(P). However, the detailed structure of the individual glycan chains in aPPG and the degree of site-specific microheterogeneity is unknown.

A. T. Ilg, unpublished results.
could be highly branched chains as proposed in the structure model shown in Fig. 11. This structural arrangement is quite distinct from promastigote phosphoglycans like LPG or sAP (compare Refs. 33 and 39). Based on the results shown in Table IV, it can be calculated that, on average, each aPPG glycan chain contains approximately six phosphorylated oligosaccharides, which are capped by four neutral oligosaccharides. These glycan chains are linked to the protein backbone most likely via the basic structure R-Man1-PO4-Ser (compare Fig. 11). However, the sequence, length, and branching of the glycan chains on individual glycosylation sites remain to be determined.

In contrast to conventional N- and O-glycosylation via glycosidic linkages to Asn and Ser/Thr, respectively (49), glycosylation of proteins via phosphoamino acids is much less common. This type of protein-glycan linkage has only been reported in the slime mold Dictyostelium discoideum (via Ser(P) (50)) and in the parasitic protozoa Leishmania (via Ser(P) (33, 34)) and Trypanosoma cruzi (via Thr(P) (51)) but appears to be absent in vertebrates. Phosphodiester linkages of the type Man1-PO4-(via Thr(P) (51)) and Man1-PO4-(via Ser(P) (50)) and Dictyostelium discoideum appear to be absent in vertebrates. Phosphodiester linkages of the type Man1-PO4-Ser, 6-Man-R are common in cell wall mannoproteins as well as in vesicular traffic and exocytosis from the living infected host cell upon rupture of infected macrophages and possibly also by vesicular traffic and exocytosis from the living infected host cell (35). Released aPPG activates the complement cascade via the C3 convertase (33, 39). Based on the results shown in Table IV, it can be calculated that, on average, each aPPG glycan (compare Refs. 33 and 39). Based on the results shown in Table IV, it can be calculated that, on average, each aPPG glycan chain contains approximately six phosphorylated oligosaccharides, which are capped by four neutral oligosaccharides. These glycan chains are linked to the protein backbone most likely via the basic structure R-Man1-PO4-Ser (compare Fig. 11). However, the sequence, length, and branching of the glycan chains on individual glycosylation sites remain to be determined.

In contrast to conventional N- and O-glycosylation via glycosidic linkages to Asn and Ser/Thr, respectively (49), glycosylation of proteins via phosphoamino acids is much less common. This type of protein-glycan linkage has only been reported in the slime mold Dictyostelium discoideum (via Ser(P) (50)) and in the parasitic protozoa Leishmania (via Ser(P) (33, 34)) and Trypanosoma cruzi (via Thr(P) (51)) but appears to be absent in vertebrates. Phosphodiester linkages of the type Man1-PO4-Ser, 6-Man-R are common in cell wall mannoproteins as well as in vesicular traffic and exocytosis from the living infected host cell upon rupture of infected macrophages and possibly also by vesicular traffic and exocytosis from the living infected host cell (35). Released aPPG activates the complement cascade via the C3 convertase (33, 39). Based on the results shown in Table IV, it can be calculated that, on average, each aPPG glycan (compare Refs. 33 and 39). Based on the results shown in Table IV, it can be calculated that, on average, each aPPG glycan chain contains approximately six phosphorylated oligosaccharides, which are capped by four neutral oligosaccharides. These glycan chains are linked to the protein backbone most likely via the basic structure R-Man1-PO4-Ser (compare Fig. 11). However, the sequence, length, and branching of the glycan chains on individual glycosylation sites remain to be determined.
L. mexicana Amastigote Proteophosphoglycan

Stage-specific Proteophosphoglycan from *Leishmania mexicana* Amastigotes: STRUCTURAL CHARACTERIZATION OF NOVEL MONO-, DI-, AND TRIPHOSPHORYLATED PHOSPHODIESTER-LINKED OLIGOSACCHARIDES

Thomas Ilg, David Craik, Graeme Currie, Gerd Multhaup and Antony Bacic

doi: 10.1074/jbc.273.22.13509

Access the most updated version of this article at http://www.jbc.org/content/273/22/13509

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 63 references, 34 of which can be accessed free at http://www.jbc.org/content/273/22/13509.full.html#ref-list-1