THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 47, pp. 40494–40508, November 25, 2011

Chemical Structure of *Trichomonas vaginalis* Surface Lipoglycan

A ROLE FOR SHORT GALACTOSE (β1–4/3) N-ACETYLGLUCOSAMINE REPEATS IN HOST CELL INTERACTION[*] [S]

Received for publication, July 8, 2011, and in revised form, September 6, 2011 Published, JBC Papers in Press, September 7, 2011, DOI 10.1074/jbc.M111.280578

Christopher M. Ryan[‡], **Angela Mehlert**[§1], **Julia M. Richardson**[¶], **Michael A. J. Ferguson**[§2], **and Patricia J. Johnson**[‡3]

From the [‡]*Department of Microbiology, Immunology, and Molecular Genetics, UCLA, Los Angeles, California 90095-1489, the* [§]*Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 4HN, Scotland, United Kingdom, and the* [¶]*School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JR, Scotland, United Kingdom*

Background: *Trichomonas vaginalis* lipoglycan (TvLG) mediates interactions between the parasite and human host.
Results: TvLG is composed of a polyrhamnose backbone with branches of poly-*N*-acetyllactosamine that are involved in attachment to host epithelium.
Conclusion: TvLG has a unique structure among solved parasite glycans.
Significance: This work provides a template to analyze TvLG from *T. vaginalis* with different binding properties.

The extracellular parasite *Trichomonas vaginalis* contains a surface glycoconjugate that appears to mediate parasite-host cell interaction via binding to human galectin-1. This glycoconjugate also elicits cytokine production from human vaginal epithelial cells, implicating its role in modulation of host immune responses. We have analyzed the structure of this glycoconjugate, previously described to contain the sugars rhamnose (Rha), *N*-acetylglucosamine (GlcNAc), galactose (Gal), xylose (Xyl), *N*-acetylgalactosamine (GalNAc), and glucose (Glc), using gas chromatograph mass spectrometry (GC-MS), matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF), electrospray MS/MS, and nuclear magnetic resonance (NMR), combined with chemical and enzymatic digestions. Our data reveal a complex structure, named *T. vaginalis* lipoglycan (TvLG), that differs markedly from *Leishmania* lipophosphoglycan and *Entamoeba* lipopeptidophosphoglycan and is devoid of phosphosaccharide repeats. TvLG is composed of an α1–3 linked polyrhamnose core, where Rha residues are substituted at the 2-position with either β-Xyl or chains of, on average, five *N*-acetyllactosamine (-3Galβ1–4GlcNAcβ1-) (LacNAc) units and occasionally lacto-*N*-biose (-3Galβ1-3GlcNAcβ1-) (LNB). These chains are themselves periodically substituted at the Gal residues with Xyl-Rha. These structural analyses led us to test the role of the poly-LacNAc/LNB chains in parasite binding to host cells. We found that reduction of poly-LacNAc/LNB chains decreased the ability of TvLG to compete parasite binding to host cells. In summary, our data provide a new model for the structure of TvLG, composed of a polyrhamnose backbone with branches of Xyl and poly-LacNAc/LNB. Furthermore, the poly-LacNAc side chains are shown to be involved in parasite-host cell interaction.

Trichomonas vaginalis is an extracellular parasitic protozoan that inhabits the urogenital tract of men and women, infecting over 170 million people worldwide (1, 2). Infections by *T. vaginalis* are often asymptomatic or cause mild irritation and/or swelling of the urogenital tract and surrounding tissues (3, 4). However, infections can also lead to severe health complications, such as cervical erosion, premature birth, and infertility in men and women (2, 5–7). More recently, infection with this parasite has been associated with increased susceptibility to human immunodeficiency virus, cervical cancer, and aggressive prostate cancer (8–12).

As an extracellular parasite, molecules on the surface of *T. vaginalis* are critical for its interaction with host cells, competition with and consumption of other organisms inhabiting the human urogenital tract, and evasion of the immune response. The focus of most studies on *T. vaginalis* surface molecules has revolved around proteins and their possible roles in host-parasite interaction (13–15). The most extensively studied putative adhesion proteins are a controversial set of metabolic enzymes suspected of having dual localization inside the cells and on the surface (reviewed in Refs. 14 and 15). More recently, sequencing of the *T. vaginalis* genome has allowed for the identification of over 1000 predicted surface proteins, many containing domains that have homology to proteins implicated in pathogenesis in other organisms (14, 16). Following this analysis, a surface proteome for *T. vaginalis* was generated, which verified localization of several membrane proteins predicted by

* This work was supported, in whole or in part, by National Institutes of Health Grant R01 AI069058 (to P. J. J.). This work was also supported by a postdoctoral fellowship from the American Cancer Society and a travel grant from Boehringer Ingelheim Fonds (to C. M. R.). The Mass Spectrometry facility at the College of Life Sciences, University of Dundee, is supported by Wellcome Trust Strategic Award 083481.
⌘ *Author's Choice*—Final version full access.
[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1 and S2.
[1] Supported by Wellcome Trust Grant 085622 and a grant from the Medical Research Council, United Kingdom.
[2] To whom correspondence may be addressed: Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 4HN, Scotland, United Kingdom. Tel.: 1382-384219; Fax: 1382-348896; E-mail: m.a.j.ferguson@dundee.ac.uk.
[3] To whom correspondence may be addressed: 1602 Molecular Sciences Bldg., 609 Charles E. Young Drive East, Los Angeles, CA 90095-1489. Tel.: 310-825-4870; Fax: 310-206-5231; E-mail: johnsonp@ucla.edu.

the genome, identified other *T. vaginalis* surface protein families, and identified several unique hypothetical proteins that are found on the surface of *T. vaginalis* (13).

Along with surface proteins, several groups have examined the role of surface carbohydrates in the adhesion of *T. vaginalis* to human cells. Early experiments demonstrated that more virulent strains of *T. vaginalis* bound more strongly to soybean agglutinin than less adherent strains (17–19). Furthermore, treatment of *T. vaginalis* with glycosidases or periodate reduced binding of the parasite to epithelial cells (20, 21). These data implicated surface carbohydrates in adhesion of *T. vaginalis*, prompting isolation of the most abundant surface glycoconjugate from the parasite, which was described as a lipophosphoglycan (LPG)[4] because of its similarities to *Leishmania* LPGs (22, 23). Research on this LPG-like molecule has revealed a role in adhesion and immune interaction. In regard to adhesion, the isolated LPG-like material has been shown to inhibit binding of *T. vaginalis* to multiple human cell lines (24–26). In addition, parasites with truncated LPG-like molecules have reduced binding to vaginal epithelial cells, and their LPG-like glycans have lost the ability to compete parasite binding to host cells (25). Finally, galectin-1 has been shown to be the host cell receptor that binds the *T. vaginalis* LPG-like molecule (27). Although there is no direct evidence that LPG-galectin-1 interaction alters protein expression in the host cell, exposure of vaginal epithelial cells to purified *T. vaginalis* LPG-like molecules results in an increase in IL-8 and macrophage inflammatory protein 3α expression (26). *T. vaginalis* LPG-like molecules also activate nuclear factor-κB and extracellular-signal-regulated kinase 1/2 (28). These data suggest that *T. vaginalis* LPG-like glycan plays an important role in attachment of the parasite to the host and modulation of host cell gene expression.

Despite its variety of functions, the structure of the *T. vaginalis* LPG-like material is incompletely described. The LPG structure from *Leishmania*, upon which the *T. vaginalis* LPG-like material tends to have been modeled, is that of a glycoinositol phospholipid core of Galα1–6Galα1–3Gal$f\beta$1–3(±Glcα1-6)Manα1–3Manα1–4GlcNα1–6*myo*-inositol-1-P-*sn*-1-alkyl-2-*lyso*-glycerol attached through the 6-position of the non-reducing terminal Gal residue to several (from 10 to 30) phosphosaccharide repeats of -6Galβ1–4Manα1-P- (that can be substituted at the 3-position of the Gal residue with various carbohydrate side chains) and finally capped with di-, tri-, and tetrasaccharides (29–31). The GlcNα1–6*myo*-inositol-1-P-lipid motifs of the *Leishmania* LPGs further define them as members of the glycosylphosphatidylinositol superfamily of eukaryotic glycophospholipids (30). The pioneering work of Singh and colleagues (32, 33) clearly shows that the *T. vaginalis* LPG-like material has similarities to *Leishmania* LPGs (*i.e.* it is based on an inositol phospholipid (in this case an inositol phosphoceramide), it is devoid of protein components, and it is rich in carbohydrate). However, there are also significant differences in composition, with the *T. vaginalis* LPG-like material containing little or no mannose but containing other sugars that are absent from *Leishmania* LPGs, like *N*-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), xylose (Xyl) and rhamnose (Rha) (23, 25).

Recently, Singh and colleagues (21) defined a significant structural feature for the *T. vaginalis* LPG-like material by showing that it contains *N*-acetyllactosamine (-3Galβ1-4GlcNAcβ1-), LacNAc, repeats. These authors further proposed the current model of *T. vaginalis* LPG-like molecule as a glycosylphosphatidylinositol-based structure containing GlcN-inositol phosphoceramide and at least two domains separated by an acid-labile region. In this paper, we test parts of this model and provide additional structural information on this extremely complex molecule. Furthermore, we utilize our structural information to identify portions of the molecule that mediate adherence to host cells.

EXPERIMENTAL PROCEDURES

Cell Growth and Culture—*T. vaginalis* strain B7RC2 was grown in TYM medium as described previously (34). The human ectocervical cell line Ect1 E6/E7 (CRL 2614) was grown in Keratinocyte-SFM (Invitrogen) as described previously (35).

TvLG Isolation and Staining—TvLG was extracted from *T. vaginalis* utilizing an extraction procedure developed for purification of *Leishmania major* LPG with minor alterations (36). Parasites were washed and resuspended at 2×10^9 cells/ml in dH_2O and then mixed with methanol and chloroform to a final water/methanol/chloroform ratio of 0.8:2:1. The mixture was sonicated, incubated for 2 h at room temperature, and then centrifuged. The pellet was extracted with water, methanol, and chloroform at the same ratios as above, recollected by centrifugation, and dried under N_2. This delipidated pellet was extracted by resuspension in 9% 1-propanol in 100 mM ammonium acetate, sonication, and a 2-h incubation at room temperature. The supernatants were collected by centrifugation, pooled, and incubated with 2.5 μg/ml proteinase K at room temperature for 24 h. Samples were then subjected to octyl-Sepharose hydrophobic chromatography using a gradient from 5% 1-propanol in 100 mM ammonium acetate to 70% 1-propanol. The eluate fraction containing TvLG was determined by spotting on a HPTLC plate and staining with orcinol reagent. TvLG was also visualized by separation on SDS-PAGE and staining with Schiff reagent or Stains-All.

For adherence assays, TvLG was isolated using the method described in Bastida-Corcuera *et al.* (25) with the following modifications. After the final extraction and dialysis of TvLG, the sample was lyophilized, suspended in 5% 1-propanol in 100 mM ammonium acetate, and loaded onto an octyl-Sepharose column. After washing, TvLG was eluted with 45% 1-propanol and lyophilized. Dried TvLG was suspended in dH_2O, and size and lack of protein were confirmed by SDS-PAGE followed with periodic acid-Schiff staining and silver staining, respectively. Monosaccharide composition analysis was then performed as described below.

Monosaccharide Composition Analysis—The monosaccharide composition of TvLG was determined using methanolysis, TMS derivatization, and GC-MS as described by Ferguson (37).

[4] The abbreviations used are: LPG, lipophosphoglycan; TvLG, *T. vaginalis* lipoglycan; dH_2O, distilled H_2O; HF, hydrofluoric acid; Gal-ol, galactitol; GlcNAc-ol, *N*-acetylglucosaminitol; LacNAc, *N*-acetyllactosamine; LNB, lacto-*N*-biose.

TvLG was dried in a capillary tube with a known amount of *scyllo*-inositol internal standard then suspended in methanol and redried. The samples were then methanolyzed, under vacuum, with 0.5 M HCl in dry methanol at 80 °C for 4 h. In addition, a known mixture of monosaccharides and *scyllo*-inositol was used as an external sugar standard to compare the response factors of the different monosaccharides in the assay. After cooling and opening the tubes, the samples were neutralized with pyridine, and any *N*-acetylated sugars were re-*N*-acetylated by adding acetic anhydride, and then the samples were dried and redried from methanol in a SpeedVac. The resulting methyl glycosides were trimethylsilylated with a pyridine/hexamethyldisilazane/chlorotrimethlysilane mixture (10:3:1). The TMS-derivatized samples were analyzed by GC-MS on an HP-5 column (Agilent). For GC-MS, the splitless injector was held at 280 °C, and the initial oven temperature was 80 °C. The main part of the GC oven gradient was 140–250 °C at 5 °C/min, that temperature being held for 10 min. At the end of the run, the temperature was ramped to the final temperature of 280 °C. The GC-MS interface was held throughout the program at 280 °C. Material was detected by electron impact ionization using total ion scanning from *m*/*z* 40 to 500.

Aqueous Hydrofluoric Acid Treatment and Alditol Formation— h at 4 °C. Acid was removed by freeze drying the samples and redrying four times from dH_2O. Hydrolyzed samples were reduced with 0.5 M $NaBD_4$ for 12 h at 4 °C. The reductant was inactivated by the addition of 1 M acetic acid until reactivity ceased. The reduced samples were desalted by application to an AG50 column and then dried. These samples were washed and dried by successive washes with methanol in 1% acetic acid, methanol, and then toluene. This material was analyzed along with alditol standards using GC-MS as described for monosaccharide composition analysis.

Total Phosphate Analysis—Phosphate content of TvLG was determined using BIOMOL green (Enzo Life Sciences). TvLG was dried in acid-washed vials and hydrolyzed at 110 °C for 24 h in sealed tubes. The hydrolysis products were dried by SpeedVac, suspended in dH_2O, and transferred to a 96-well plate. Samples and standard curve containing supplied phosphate standard were then exposed to a 2× volume of BIOMOL green at room temperature for 20 min. Absorbance was measured at 620 nm, and the concentration of phosphate in TvLG was determined utilizing the standard curve.

NMR Spectroscopy—One-dimensional ^1H NMR spectra were acquired on dephosphorylated TvLG after exchange into ^2H_2O using a Bruker spectrometer operating at 500 MHz for ^1H and 300 K. The carbohydrate spin systems were assigned from two-dimensional COSY and TOCSY ^1H spectra acquired at 303 K on a Bruker Avance spectrometer operating at 800 MHz for ^1H. Two TOCSY spectra were acquired with mixing times of 60 and 120 ms. The TvLG sample used in these experiments was repurified on octyl-Sepharose after further digestion with β-hexosaminidase and β-galactosidase.

Methylation Linkage Analysis—The linkages between monosaccharides in TvLG were determined as previously described (37) with some modifications. In brief, TvLG was suspended in a slurry of NaOH in DMSO and then methylated by three successive treatments with methyl iodide. Chloroform and aqueous sodium thiosulfate were added to inactivate the methyl iodide, and the phases were separated. The organic phase was washed four times with dH_2O and dried. The sample was then hydrolyzed with H_2SO_4 in acetic acid for 4 h at 80 °C. Samples were cooled then neutralized with methanolic NaOH. To reduce these methylated monosaccharides, hydrolyzed products were dissolved in NH_4OH (with *scyllo*-inositol as an internal standard) at room temperature for at least 12 h with $NaBD_4$. Reductant was destroyed through the addition of acetic acid. Samples were dried, and these partially methylated alditols were acetylated by incubation with acetic anhydride followed by extraction with dichloromethane and water. The organic phase was extracted and concentrated under N_2 and then analyzed by GC-MS on an HP-5 column using the conditions described for monosaccharide analysis.

TFA Digestion—Mild TFA treatment to determine the extent of phosphorylation in TvLG was carried out by drying TvLG and then incubating in 40 mM TFA for 10 min at 100 °C. The sample was dried again and applied to an octyl-Sepharose column. The monosaccharide composition of unbound, wash, and bound fractions was determined as described above. For analysis by MALDI-TOF or ESI-MS/MS, TvLG was treated with 0.1 M TFA for 40 min or 1 h at 100 °C. Samples were then treated for analysis by mass spectrometry or reduced to identify reducing and non-reducing ends on the TFA fragments. For reduction, TFA-digested TvLG was incubated in 0.25 M $NaBD_4$ for 12 h at 4 °C. The reductant was destroyed by the addition of 1 M acetic acid, and the reaction was applied to an AG50 column to desalt. Eluted fractions were lyophilized and then co-evaporated three times with 2% acetic acid in methanol, three times with methanol, and twice with toluene. Samples were then treated for use in mass spectrometry.

Mass Spectrometry—The positive ion MALDI mass spectrum of native TvLG glycoconjugate was collected in linear mode on an Applied Biosystems Voyager DE-STR instrument using dihydroxybenzoic acid at 10 mg/ml in 0.1% TFA. Samples were mixed with matrix in a 1:1 ratio and then applied to the plate and dried. The accelerating voltage was 20,000 V, the extraction time was 600 ns, and the laser intensity was 2252 V. Data from 500 laser shots was averaged over a mass range of 1000–50,000 Da. TFA-digested TvLG, with or without reduction, was permethylated prior to analysis by mass spectrometry as described previously (36). Briefly, the treated TvLG samples were dried in an acid-washed vial, redried from methanol, and then suspended in 50 μl of DMSO. A slurry of NaOH in DMSO (120 mg/ml) was prepared, and 50 μl was added to each vial, and after 20 min, 10 μl of methyl iodide was also added. A further addition of 10 μl of methyl iodide was made after 10 min, and then a final addition of 20 μl was added after another 10 min. 20 min after the final addition of methyl iodide, the permethylated glycans were recovered, and the methyl iodide was inactivated by the addition of 500 μl of chloroform and 1 ml of 100 mg/ml sodium thiosulfate. After thorough mixing, the phases were separated, and the lower ($CHCl_3$) phase was extracted and washed five times with dH_2O and then dried. For MALDI-TOF analysis, the hydrolyzed permethylated TvLG was mixed 1:1 with dihydroxybenzoic acid matrix, as above. These samples were analyzed in the positive reflectron mode on

TABLE 1
Monosaccharide composition of *T. vaginalis* lipoglycan from this and previous reports
Percentages are expressed as mol % of total monosaccharide.

	Glucose[a]	Galactose	GlcNAc	GalNAc	Rhamnose	Xylose
Bastida-Corcuera *et al.* (25)	6.0	18.9	33.0	2.3	23.4	16.5[b]
Fichorova *et al.* (26)	2.5	10.3	38.7	2.5	30.8	15.2
This work	2.1 ± 1.3	28.8 ± 5.3	30.9 ± 8.1	Trace[c]	25.7 ± 4.4	12.4 ± 2.0

[a] Glucose is considered a common environmental contaminant and may not be a genuine component of TvLG.
[b] The assay used was unable to differentiate between xylose and mannose.
[c] GalNAc was only detected once in nine assays at 0.6%.

the MALDI-TOF mass spectrometer as above but using a mass range of 500–8000 Da. For electrospray mass spectrometry, samples were permethylated and dried as described above and then resuspended in 80% acetonitrile and mixed with an equal volume of 1 mM sodium acetate in 80% acetonitrile. The mixture was loaded into a nanotip (Micromass type F) and analyzed by positive ion ESI-MS on a Q-Star XL instrument (Applied Biosystems) using Analyst software.

Exoglycosidase Digestion—For enzymatic digestion, TvLG was incubated with 0.5 unit of β-galactosidase (bovine testes; Sigma), 15 units of β-*N*-acetylhexosaminidase (New England Biolabs), or a combination of the two at 37 °C for 24 h in 50 mM sodium citrate, pH 4.5. For mock treatment, equivalent amounts of TvLG, as used in digestion reactions, were incubated at 37 °C for 24 h in buffer only. Following digestion or mock treatment, the reaction was applied to an octyl-Sepharose column and washed with 5% 1-propanol, 100 mM ammonium acetate. TvLG was then eluted from the column by application of 45% 1-propanol. The eluted fraction was dried before additional treatment or analysis.

Digest with α-galactosidase (10 units) was carried out at 37 °C for 24 h in 50 mM sodium acetate, 5 mM $CaCl_2$, pH 5.5. Following each digestion, the reactions were applied to an octyl-Sepharose column and eluted as described above. The eluted fraction was then utilized for the next digest.

Smith Degradation—Smith degradation was carried out using a revised version of the protocol outlined by Abdel Aker *et al.* (38). TvLG in dH_2O was added to an equal volume of 400 mM sodium acetate, 0.1% 1-propanol, 60 mM sodium periodate. This mixture was incubated at 4 °C in the dark for 72 h. Ethylene glycol was added to the solution, and this reaction was incubated at 4 °C for 30 min followed by room temperature for 30 min. The products were then reduced by the addition of 1 M $NaBD_4$ in 1 M NaOH for 12 h at 4 °C. The reductant was eliminated by the addition of 2 M acetic acid. The reaction was then applied to an octyl-Sepharose column and eluted by the addition of 45% 1-propanol after several washes of 5% 1-propanol. The eluted fraction was dried, dissolved in 0.5 M TFA, and incubated for 24 h at 30 °C. The samples were dried again prior to further analysis or treatment. Control samples of TvLG were also subjected to mild acid treatment with 0.5 M TFA for 24 h at 30 °C.

Competition of Parasite Adherence to Vaginal Epithelial Cells by TvLG—Human ectocervical cell line *Ect1/E6E7* was grown as described (25, 35). For adherence assays, cells were grown on 12-mm coverslips in 24-well plates at 37 °C and 5% CO_2 until confluent. Adherence assays were carried out as described by Bastida-Corcuera *et al.* (25) with several modifications. Ectocervical cells grown on coverslips were washed with complemented Keratinocyte-SFM (Invitrogen) and then incubated at 37 °C and 5% CO_2 for 1 h with 0.64 μM mock-digested TvLG or TvLG digested with β-galactosidase and β-*N*-acetylhexosaminidase. Because the total amount of rhamnose should not change upon digest, concentration of both digested and mock-treated TvLG was determined by extrapolating the molarity of the total molecule from the experimentally determined molar amount of rhamnose present after treatment. Following incubation with TvLG, 5×10^4 parasites (*T. vaginalis* strain B7RC2) labeled with 10 μM cell tracker blue (Invitrogen) were added to the ectocervical cells. Parasite and cells were incubated together for 30 min at 37 °C under 5% CO_2. Coverslips were then washed with complemented Keratinocyte-SFM, fixed for 20 min in 4% paraformaldehyde, and mounted onto slides. Fifteen fields at ×10 magnification were analyzed per coverslip. Fluorescent parasites adhered to the host cells were counted using ImageJ version 1.41o for Windows. Each experiment was performed with three coverslips and repeated at least three times with different passages of epithelial cells and parasites as well as different TvLG digestions.

RESULTS

Purification and Monosaccharide Composition of T. vaginalis LPG-like Material—The first step toward defining the structure of *T. vaginalis* LPG-like material was its isolation. Here we used a procedure based on that of McConville *et al.* (36) for the isolation of *Leishmania* LPGs that has also been used for the isolation of *Trypanosoma brucei* procyclins (39). This method differs from previously used procedures (25, 26); however the monosaccharide composition of the isolated material was found to be similar with two differences (Table 1). The sugar GalNAc was not detected in our analyses, and there was a relative increase in the amount of Gal. These results indicate that the LPG-like molecules may differ between *T. vaginalis* strains. Similarly, our current and previous composition analyses for the B7RC2 strain have less rhamnose and more galactose than was reported by others using the UR1 strain (Table 1).

T. vaginalis LPG-like Material Is Not a Lipophosphoglycan—Preliminary studies were designed to exploit the predicted extreme acid lability of *T. vaginalis* LPG-like molecules at sugar-1-phosphate bonds. These experiments tested whether, similar to *Leishmania* LPGs, the *T. vaginalis* LPG-like molecules could be quantitatively depolymerized with mild acid (40 mM TFA, 100 °C for 8–10 min) (29). No significant fragmentation of the *T. vaginalis* LPG-like material was apparent using negative ion electrospray mass spectrometric analysis of this mild acid digest (data not shown). Furthermore, whereas GC-MS

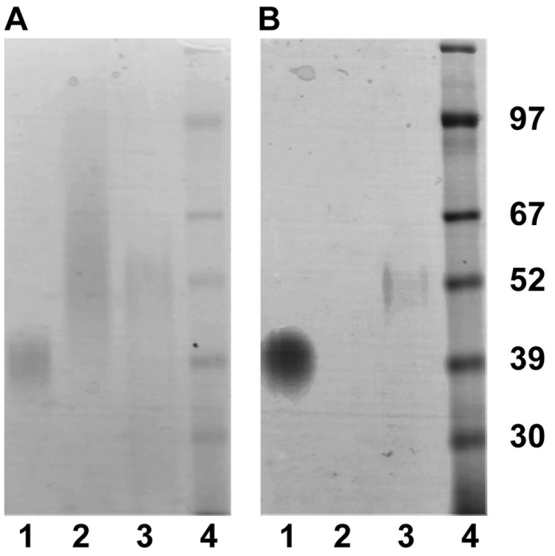


FIGURE 1. **Differential staining of parasite glycoconjugates.** Purified procyclin from *T. brucei* (*lane 1*), TvLG from *T. vaginalis* (*lane 2*), and LPG from *L. major* (*lane 3*) were resolved by SDS-PAGE along with a protein molecular mass marker (*lane 4*). Molecular mass in kDa is indicated on the *side* of the gel. The gels were stained with periodate-Schiff reagent (*A*), a general polysaccharide stain, and Stains-All (*B*), a polyanion-sensitive carbocyanine dye.

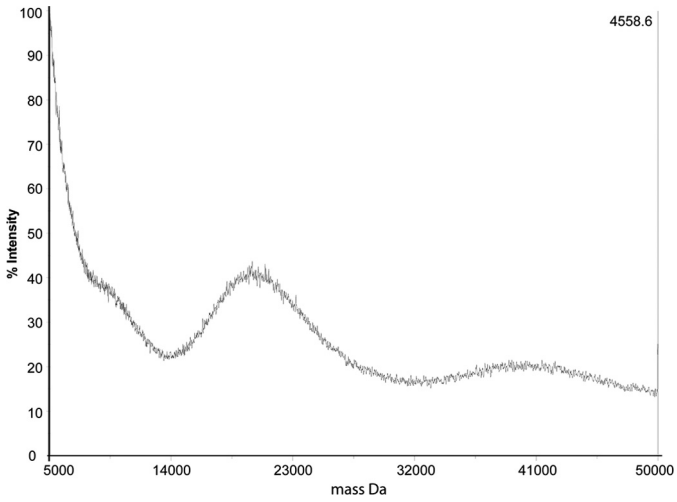


FIGURE 2. **Positive ion MALDI-TOF mass spectrum of the native purified *T. vaginalis* TvLG glycoconjugate.**

compositional analyses of *Leishmania* LPGs using the methanolysis/trimethylsilylation method generally results in the qualitative detection of Gal-6-P from the phosphosaccharide repeats (36), no sugar phosphate derivatives were detected for the *T. vaginalis* LPG-like material. These data caused us to question the LPG model of the *T. vaginalis* LPG-like material.

To test the LPG model further, we treated the material with 50% aqueous hydrofluoric acid (HF), under conditions that quantitatively cleave phosphate ester bonds (37), and then reduced the products with sodium borodeuteride. This method should convert every sugar linked to the rest of the structure through a sugar-1-phosphate bond into a neutral heavy isotope-labeled alditol. Analysis for monosaccharide composition by GC-MS revealed that less than 1% of the monosaccharides had been converted to alditols and that, in addition, there was no selectivity as to which monosaccharides were converted to alditols (data not shown). We concluded from these data that the tiny amounts of alditols produced were due to the expected minor amount of random acid-catalyzed degradation of the material during the aqueous HF step and that the original LPG model was probably incorrect.

We further tested the LPG model by comparing the staining properties of the *T. vaginalis* LPG-like material with those of *L. major* LPG and *T. brucei* procyclin on SDS-polyacrylamide gels. Under conditions where the periodate-Schiff reagent staining for carbohydrate of all three molecules was similar (Fig. 1*A*), only *L. major* LPG and *T. brucei* procyclin gave positive staining with the polyanion stain Stains-All (Fig. 1*B*). These data suggest that *T. vaginalis* LPG-like material is carbohydrate-rich but that, unlike *L. major* LPG and *T. brucei* procyclin, it has a very low density of anionic groups. Again, these data are inconsistent with an LPG model for the *T. vaginalis* LPG-like material.

Finally, *T. vaginalis* LPG-like material was completely hydrolyzed using sulfuric acid, and the resulting inorganic phosphate concentration was quantified utilizing BIOMOL green reagent. This analysis revealed that phosphate accounts for less than 0.5% of the total weight of the *T. vaginalis* LPG-like material (data not shown). In contrast, *L. major* LPG is 17% phosphate by weight (36).

Taken together, our results demonstrate that *T. vaginalis* LPG-like material does not contain significant amounts of phosphate in the glycan core, compared with other LPGs, and we propose that it should be renamed from here on as *T. vaginalis* lipoglycan (TvLG).

Although the aqueous HF procedure described above did not reveal a phosphosaccharide repeat structure for TvLG, it did convert 95% of the carbohydrate from being lipid-bound to lipid-free, as judged by monosaccharide compositional analysis of the bound and 45% propan-1-ol-eluted and flow-through fractions, respectively, on an octyl-Sepharose column (data not shown). In addition, solvent extraction after aqueous HF treatment and analysis of the lipid extract by GC-MS after trimethylsilylation revealed ceramide, principally in the form of *N*-palmitoyl-sphingosine.[5] This is consistent with the presence of a phosphate-linked ceramide lipid in TvLG, which, in turn, is consistent with the presence of an inositol phosphoceramide in TvLG, as postulated previously (32, 33). This aqueous HF-delipidated material was used for NMR analysis.

Mass Spectrometric Analysis of Native TvLG—Untreated TvLG was analyzed by positive ion MALDI-TOF mass spectrometry, revealing two wide peaks centered at *m/z* 20,000 (ranging from 15,000 to 25,000) and 40,000 (ranging from 30,000 to 50,000) (Fig. 2). The peak around *m/z* 20,000 probably represent doubly charged ions because the peak width is approximately half that of the peak centered around 40,000. These results show that TvLG is composed of a very large and polydisperse set of molecules with, on average, around 250 monosaccharides/molecule.

[5] Y. C. Liu, C. M. Ryan, A. Mehlert, T. Smith, I. Almeida, T. Hermoso, M. Field, P. Johnson, J. Mottram, and A. Acosta-Serrano, manuscript in preparation.

TABLE 2
MALDI-TOF analysis of *T. vaginalis* lipoglycan following TFA digestion

m/z (observed)	*m/z* (calculated)	Δ (obs − cal)	Xylose	Rhamnose	Galactose	GlcNAc
420.3077	420.2257	0.0820	1	1	0	0
450.3094	450.2364	0.0730	1	0	1	0
464.3314	464.2524	0.0790	0	1	1	0
535.3872	535.2892	0.0980	0	0	1	1
638.6499	638.3409	0.3090	0	2	1	0
668.3986	668.3516	0.0470	0	1	2	0
695.4237	695.3627	0.0610	1	0	1	1
709.4453	709.3783	0.0670	0	1	1	1
739.4769	739.3889	0.0880	0	0	2	1
780.5022	780.4152	0.0870	0	0	1	2
839.5322	839.4412	0.0910	1	2	0	1
855.5240	855.4360	0.0880	2	0	1	1
869.5344	869.4524	0.0820	1	1	1	1
899.5248	899.4628	0.0620	1	0	2	1
913.5453	913.4783	0.0670	0	1	2	1
954.6044	954.5054	0.0990	0	1	1	2
984.5995	984.5145	0.0850	0	0	2	2
1073.6316	1073.5476	0.0840	1	1	2	1
1100.6417	1100.5627	0.0790	2	0	1	2
1114.6722	1114.5782	0.0940	1	1	1	2
1144.6734	1144.5894	0.0840	1	0	2	2
1158.6723	1158.6043	0.0680	0	1	2	2
1188.6932	1188.6152	0.0780	0	0	3	2
1229.7318	1229.6418	0.0900	0	0	2	3
1304.7138	1304.6618	0.0520	2	0	2	2
1318.7297	1318.6777	0.0520	1	1	2	2
1348.7207	1348.6887	0.0320	1	0	3	2
1362.7040	1362.7039	0.0001	0	1	3	2
1433.7576	1433.7406	0.0170	0	0	3	3
1435.7551	1435.7461	0.0090	0	4	2	1
1478.7804	1478.7524	0.0280	2	1	2	2
1593.8605	1593.8145	0.0460	1	0	3	3
1607.8605	1607.8315	0.0290	0	1	3	3
1637.8650	1637.8410	0.0240	0	0	4	3
1638.8515	1638.8245	0.0270	3	1	2	2
1678.8771	1678.8681	0.0090	0	0	3	4
1753.9207	1753.8887	0.0320	2	0	3	3
1767.9372	1767.9042	0.0330	1	1	3	3
1811.9206	1811.9296	−0.0090	0	1	4	3
1869.9717	1870.9560	−0.9843	1	3	3	2
1882.9999	1882.9669	0.0330	0	0	4	4
1883.9691	1883.9521	0.0170	3	1	2	3
1977.0370	1977.0080	0.0290	0	5	5	0
2057.0556	2057.0566	−0.0010	0	1	4	4
2087.0487	2087.0657	−0.0170	0	0	5	4
2261.1778	2261.1878	−0.0100	3	2	3	3
2333.1541	2333.1781	−0.0240	3	1	3	4
2638.2074	2638.3504	−0.1430	1	6	3	3
2782.3789	2782.4039	−0.0250	3	1	4	5
2899.3323	2899.4713	−0.1390	2	4	4	4
3116.3467	3115.5817	0.7650	1	1	6	6

Assignment of Sugar Residue Anomericity by ^{1}H NMR Analysis of Dephosphorylated TvLG—The proton NMR spectra of the dephosphorylated TvLG fraction were extremely complex even after digestion with β-*N*-acetylhexosaminidase and β-galactosidase (supplemental Fig. S1, *A* and *B*). However, we were able to assign the anomericity of all or most of the Gal*p*, GlcNAc*p*, and Xyl*p* residues as β and all or most of the Rha*p* as α. Anomeric peaks assigned to α-Rha residues were seen at 5.57, 5.51, and 5.44 ppm. α-Gal residue anomeric signals were found at 5.24, 5.16, 5.15, and 5.14 ppm, and several anomeric signals between 4.45 and 4.57 were presumed to be from β-galactose. β-GlcNAc anomeric signals were found at 4.62, 4.59, and 4.57 ppm, and β-Xyl anomeric signals were found at 4.43, 4.41, and 4.43 ppm.

Analysis of TvLG Partial Acid Hydrolysis Fragments by Mass Spectrometry Reveals Polyrhamnose and Gal-GlcNAc Repeat Components—Before further analysis, the TvLG was subjected to partial acid hydrolysis (100 mM TFA, 100 °C, 40 min or 1 h) to fragment the structure into more tractable pieces. These fragments were analyzed by ESI-MS and ESI-MS/MS or by MALDI-TOF after either permethylation (which greatly increases glycan ionization and the information that can be obtained from MS/MS spectra) or after deutero-reduction followed by permethylation. The compositions of the various fragments detected are given in Tables 2 and 3, along with deduced compositions. In some cases, structures are proposed based on MS/MS data (see below).

The MALDI-TOF and ESI-MS data for the permethylated products of a 1-h partial hydrolysis revealed several glycan fragments composed of up to six deoxyhexose (Rha) sugars, together with varying amounts of *N*-acetylhexosamine (GlcNAc), hexose (Gal), and pentose (Xyl) residues. The MS/MS spectra of some of these fragments show Rha residues in linear chains, which are then substituted with GlcNAc residues or GlcNAc-Gal disaccharides. Annotated representative spectra for two of these fragments are shown in (Fig. 3). We interpret the intense product ions from *m/z* 1184, corresponding to $[GlcNAc_1Rha_5 + Na]^+$, at *m/z* 619 and 591 as being due to $^{1,4}X_2$ and $^{1,5}X_2$ cross-ring cleavages occurring adjacent to the

TABLE 3
Observed masses (*m/z*) and structural breakdown of peaks from electrospray ionization mass spectrometry of TFA-digested TvLG and TFA-digested and reduced TvLG

	TFA-digested TvLG					TFA-digested and reduced TvLG			
m/z	Xylose	Rhamnose	Galactose	GlcNAc	*m/z*	Xylose	Rhamnose	Galactose	GlcNAc
488	0	1	0	1	450	1	0	1	0
518	0	0	1	1	464	0	1	1	0
662	0	2	0	1	535	0	0	1	1
692	0	1	1	1	709	0	1	1	1
722	0	0	2	1	739	0	0	2	1
763	0	0	1	2	780	0	0	1	2
822	1	2	0	1	855	2	0	1	1
836	0	3	0	1	869	1	1	1	1
852	1	1	1	1	913	0	1	2	1
866	0	2	1	1	940	1	0	1	2
896	0	0	3	1	954	0	1	1	2
937	0	1	1	2	984	0	0	2	2
967	0	0	2	2	1114	1	1	1	2
996	1	3	0	1	1144	1	0	2	2
1010	0	4	0	1	1158	0	1	2	2
1040	0	3	1	1	1188	0	0	3	2
1170	1	4	0	1	1229	0	0	2	3
1184	0	5	0	1	1433	0	0	3	3
1214	0	4	1	1	1607	0	1	3	3
1255	0	4	0	2	1637	0	0	4	3
1301	1	1	2	2					
1344	1	5	0	1					
1358	0	6	0	1					
1416	0	0	3	3					
1429	0	5	0	2					

site of GlcNAc elimination from the major molecular species (illustrated in Fig. 3*A*). The presence of other minor molecular species with the GlcNAc attached to one of the other four Rha residues can be inferred from the presence of the minor product ions at *m/z* 577, 751, and 630, corresponding to Rha_3, Rha_4, and $GlcNAc_1Rha_2$ fragments, respectively. The MS/MS spectrum of the ion at *m/z* 1214, corresponding to $[Gal_1GlcNAc_1Rha_4 + Na]^+$, shown in Fig. 3*B*, suggests that Gal and GlcNAc can be attached to the Rha backbone as a GlcNAc-Gal disaccharide.

Analysis of similar partial acid hydrolysis samples (40 min), with $NaBD_4$ reduction prior to permethylation, generated an ESI-MS spectrum of complementary ions because the ionization properties of non-reduced and reduced permethylated glycans are significantly different. Thus, the reduced/permethylated sample spectrum showed a series of linear glycans with alternating Gal and GlcNAc residues ending in either galactitol (Gal-ol) or *N*-acetylglucosaminitol (GlcNAc-ol). The largest of these at *m/z* 1433.7 corresponds to a mixture of GlcNAc-Gal-GlcNAc-Gal-GlcNAc-Gal-ol and Gal-GlcNAc-Gal-GlcNAc-Gal-GlcNAc-ol structures (Fig. 4*A*). The presence of the *m/z* 329 $^{3,5}A_2$ ion in the MS/MS spectrum of the Gal-GlcNAc-Gal-ol fragment (Fig. 4*B*) is consistent with a Gal1–4GlcNAc linkage, which, together with the NMR anomeric information and the methylation linkage data (below), provides strong support for the presence of Galβ1–4GlcNAc (LacNAc) units in TvLG.

Also of interest in the ESI-MS spectrum of the partial acid, deutero-reduced, and permethylation sample were fragments containing Rha and Xyl. The MS/MS spectrum of the fragment at *m/z* 709.4 (Fig. 5*A*) suggests that there are at least two isobaric structures present. However, the main structure appears to be based on a disubstituted Gal-ol, as evidenced by the abundant product ions at *m/z* 521 (loss of non-reducing terminal Rha) and *m/z* 450 (loss of non-reducing terminal GlcNAc) and the *m/z* 262 double cleavage Gal-ol ion. One of the minor structures present is most likely a linear structure with a GlcNAc-ol deutero-reduced terminus, giving rise to the *m/z* 299, 317, and 433 product ions, respectively. The principle *m/z* 709.4 component of Rha-(GlcNAc-)Gal-ol suggests that a few of the LacNAc units are substituted by Rha. This is further supported by the MS/MS spectrum of the fragment from the same partial acid, deutero-reduced, and permethylation sample at *m/z* 1114 (Fig. 5*B*). In this case, the main component appears to be GlcNAc-(Xyl-Rha-)Gal-GlcNAc-ol, suggesting that the Rha branch can be further substituted by Xyl. These occasional substitutions of the poly-LacNAc chains probably explain the inability of a mixture of β-galactosidase and β-hexosaminidase to completely remove Gal and GlcNAc from TvLG (see below).

Exoglycosidase Digestion of Native TvLG Suggests a Polyrhamnose Core with Partially Substituted Poly-LacNAc Side Chains—Thus far, our data indicated that TvLG is composed predominantly of a polyrhamnose and poly-LacNAc components. To test this model, we digested the TvLG with β-galactosidase and β-hexosaminidase alone and in combination and repurified the product by octyl-Sepharose chromatography. Neither enzyme, when used alone, had a significant effect on the structure, as judged by SDS-PAGE and periodate-Schiff reagent stain (data not shown), whereas the combined digestion caused a significant decrease in apparent molecular weight and polydispersity of TvLG (Fig. 6*A*). Monosaccharide analysis of the double-digested material showed the removal of greater than 60% of Gal and 50% of GlcNAc (Fig. 6*B*), approximately one-third of the total monosaccharides (Table 1). These data suggest that the polyrhamnose is the core of the molecule with exoglycosidase-accessible poly-LacNAc repeat side chains attached to it.

Methylation Linkage Analysis of TvLG before and after Exoglycosidase Digestion Supports a Polyrhamnose Core with Partially Substituted Poly-LacNAc/Lacto-N-biose Side Chains

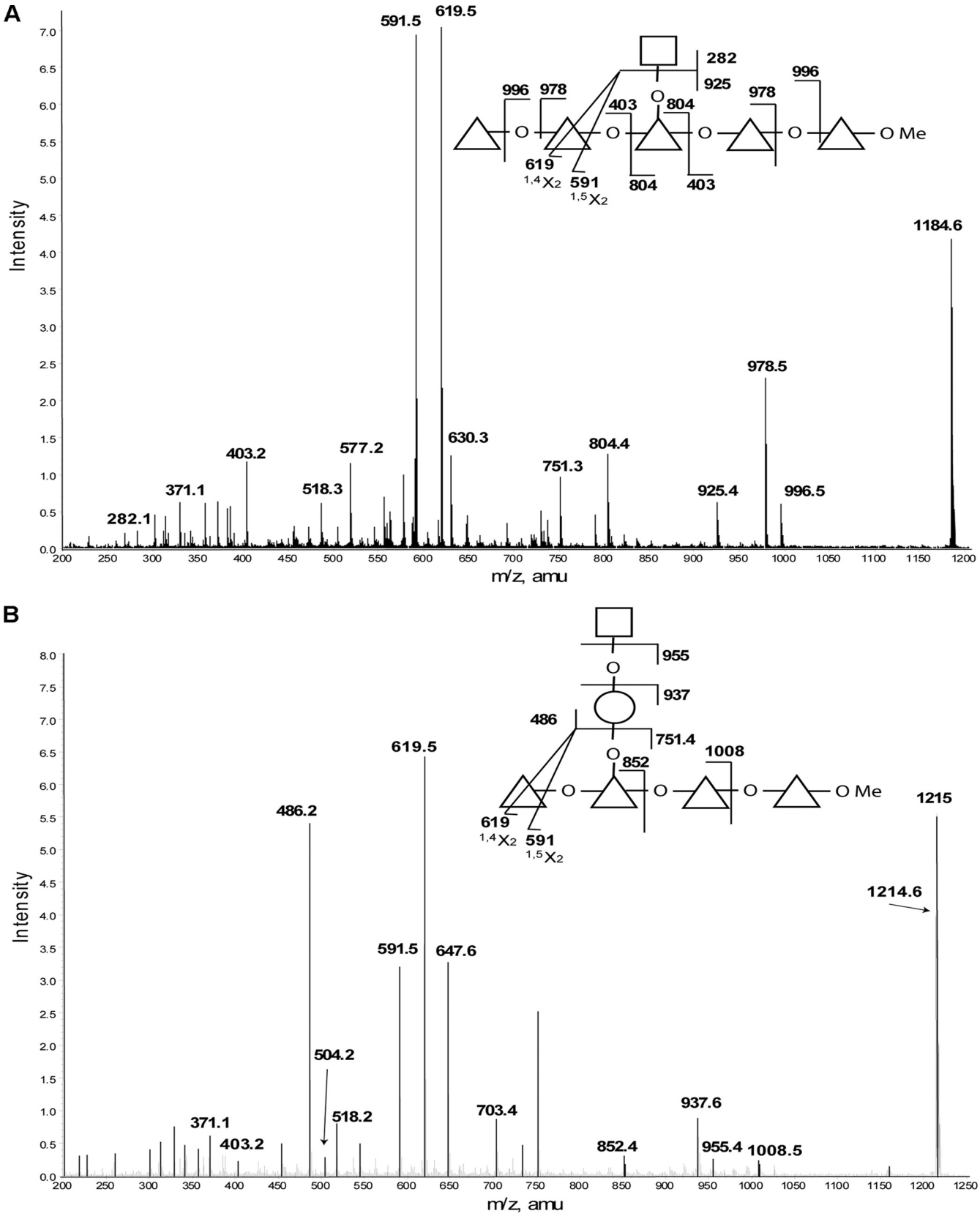
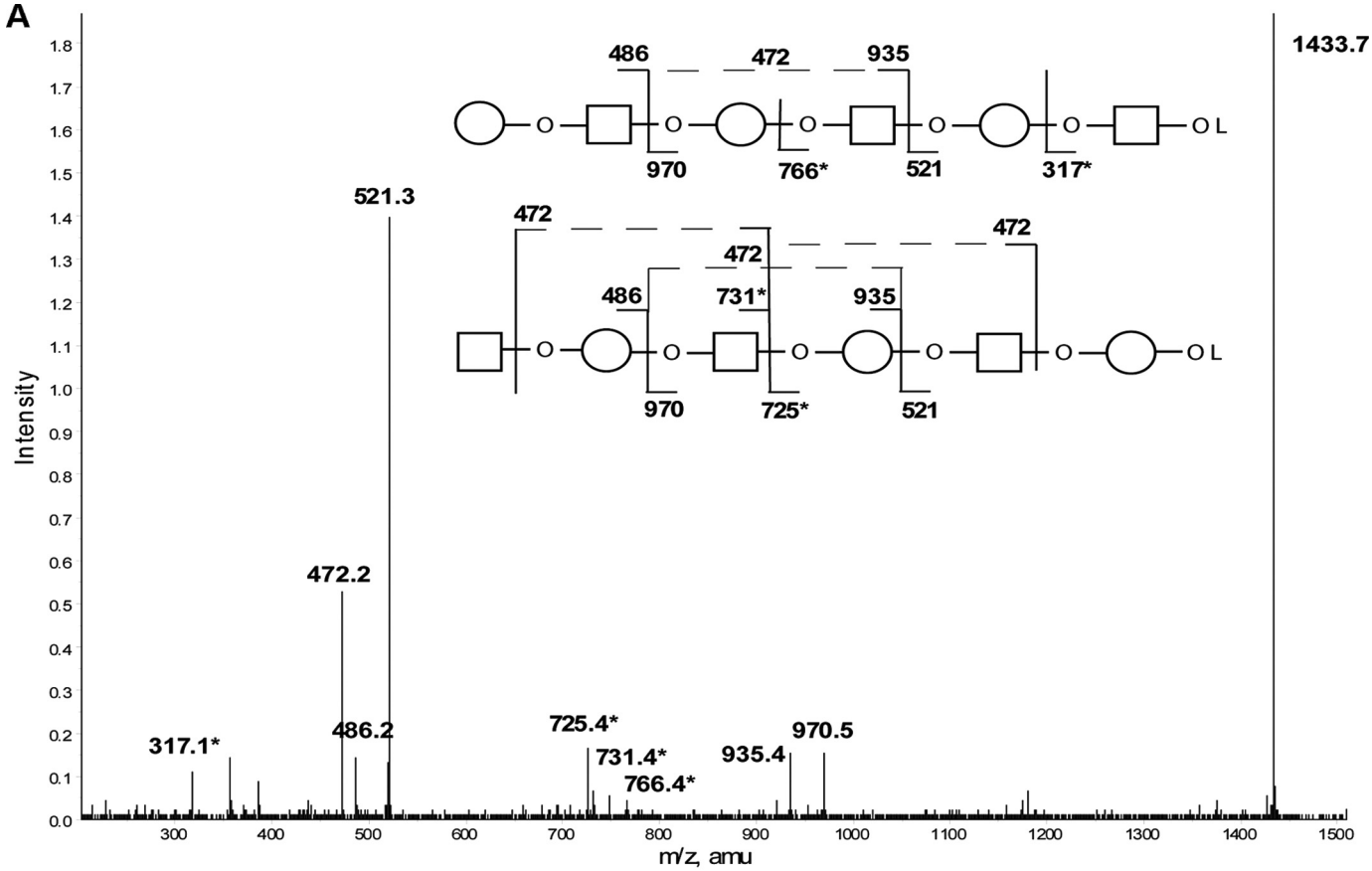
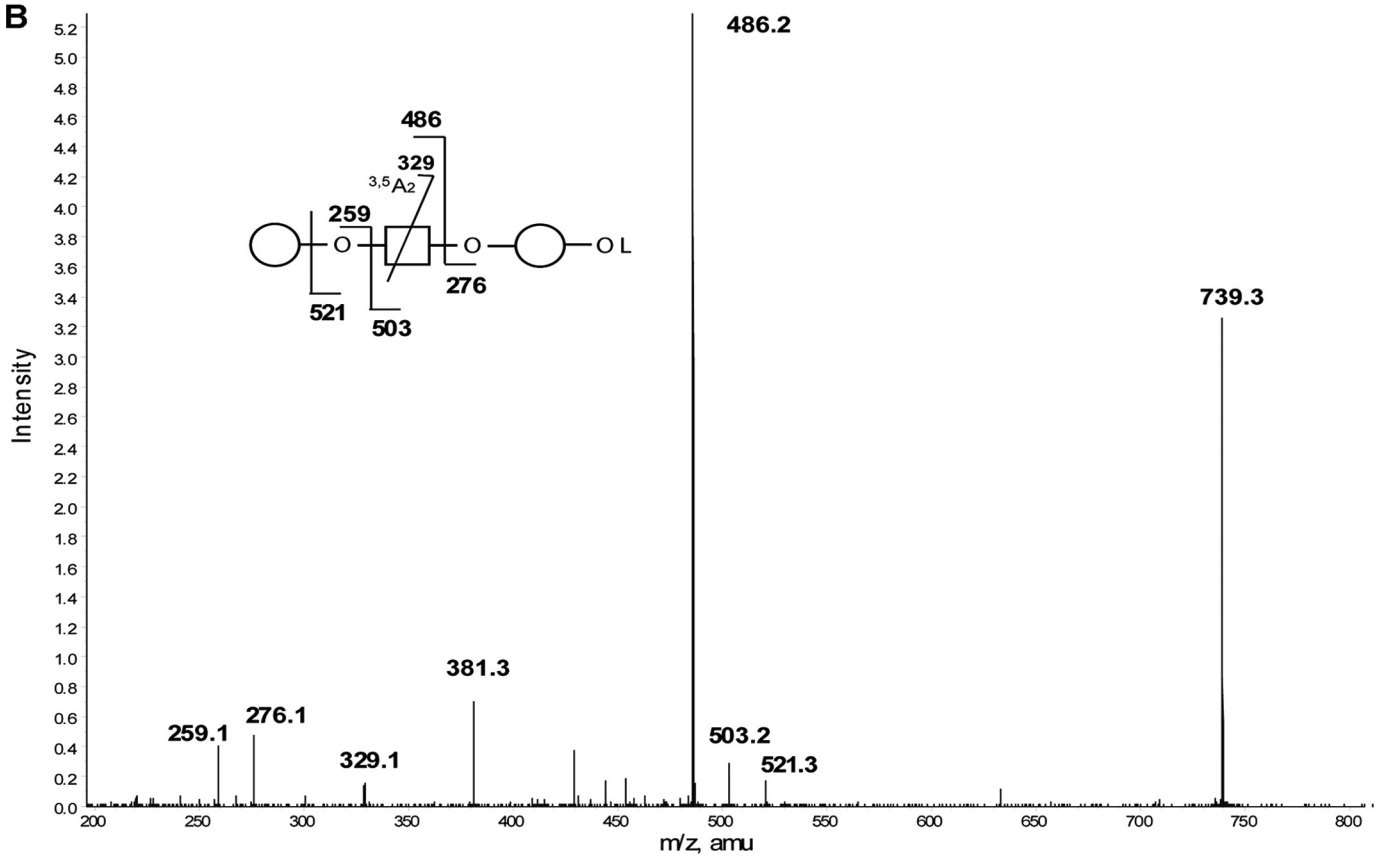


FIGURE 3. **Selected positive ion ESI-MS spectra of partially hydrolyzed and permethylated TvLG.** *A*, product ion spectrum of the $[M + Na]^+$ precursor ion at *m/z* 1184 from the partially hydrolyzed and permethylated TvLG sample. *B*, product ion spectrum of the $[M + Na]^+$ precursor ion at *m/z* 1215 from the partially hydrolyzed and permethylated TvLG sample. The annotated structures indicate the ion assignments. Cross-ring cleavages are indicated by a *line* bisecting the respective monosaccharide, and the type of cleavage is identified *below* the mass value. The residues Gal, GlcNAc, Xyl, and Rha are represented by *open circles*, *open squares*, *open pentagons*, and *open triangles*, respectively. The *O between* monosaccharides in the diagram indicates a glycosidic linkage, and *OMe* represents a methylated reducing terminus.


A
Intensity
m/z, amu
1433.7
521.3
472.2
486.2
317.1*
725.4*
731.4*
766.4*
935.4
970.5
486
472
935
970
766*
521
317*
472
472
472
486
731*
935
970
725*
521
OL


B
Intensity
m/z, amu
486.2
739.3
381.3
259.1
276.1
329.1
503.2
521.3
486
329
3,5A2
259
276
521
503
OL

Model—We performed GC-MS methylation linkage analysis (36) on the native and β-galactosidase plus β-hexosaminidase-digested TvLG. This method, which defines how the constituent monosaccharides are substituted by each other, is semi-quantitative because pure authentic standards of all of the relevant sugar derivatives (partially methylated alditol acetates) are not available. However, to a first approximation, the integrated GC-MS peaks for each derivative provide an idea of the proportion of each type of sugar in the overall structure (Table 4).

In native TvLG, about 70% of the GlcNAc is substituted at the 4-position, with the remainder present as either non-reducing terminal or 3-substituted GlcNAc. The presence of significant amounts of 3-substituted Gal along with the 4-substituted GlcNAc is consistent with the presence of LacNAc repeats of -3Galβ1–4GlcNAcβ1- (see above), although the presence of some 3-substituted GlcNAc might also indicate the presence of about one lacto-*N*-biose (-3Galβ1–3GlcNAcβ1-) (LNB) unit for every four LacNAc units. The presence of non-reducing terminal Gal residues is consistent with the ability of the TvLG to bind the lectin ricin (32), whereas the presence of some terminal GlcNAc would predict that TvLG may also bind wheat germ agglutinin. Assuming that all or most of the Gal is involved in LacNAc and LNB units, the relative amounts of terminal Gal, 3-substituted Gal, and 3,4-disubstituted Gal residues suggest that, on average, there are about four linear LacNAc/LNB units to one LacNAc/LNB unit bearing a substitution branch point, where another sugar or sugar chain is attached to the 4-position of that Gal residue. This branch point most likely corresponds to GlcNAc-(±Xyl-Rha-)Gal-GlcNAc-ol fragments described in Fig. 5. From these data (together with the identification of Gal-GlcNAc repeat fragments in partial acid digests of TvLG; see above), we can infer that the average number of LacNAc/LNB units per chain is about five and that, on average, each of these will be substituted once with a side chain starting Xyl-Rha-.

Methylation linkage analysis of the mixed exoglycosidase-digested molecule revealed significant changes (Table 4), including an almost complete loss of terminal Gal and a 75% reduction of 3-substituted Gal. Digestion also caused an almost complete loss of 4-substituted GlcNAc and 3-substituted GlcNAc and a small increase in (non-reducing) terminal-GlcNAc. These changes are consistent with the sequential actions of β-galactosidase and β-hexosaminidase to remove (poly-)LacNAc and LNB units down to the 3,4-disubsituted Gal branch points (occupied by Rha and Xyl-Rha; see above). The presence of terminal GlcNAc residues in the digested material suggests that β-hexosaminidase cannot remove the last β-GlcNAc residues from these unusual and sterically constrained branch points.

Consistent with a model of TvLG based on a polyrhamnose core, the majority of the Rha residues in TvLG are 2,3-disubstituted (Table 4). The simplest model is one where the Rha residues are linked together exclusively through either α1–2 or α1–3 linkages, with glycan side chains attached through the Rha 3- or 2-hydroxyls, respectively. Unfortunately, the mixed exoglycosidases are unable to digest the structure down to the polyrhamnose core, so there are no significant changes in substitution patterns of the Rha residues after digestion.

The least abundant of the four major monosaccharides, Xyl, appears to be exclusively terminal and 4-substituted (Table 4). Again, this substitution pattern was similar before and after mixed exoglycosidase digestion.

Smith Degradation Reveals Internal Structural Features of TvLG—TvLG is predicted to be largely resistant to Malprade periodate oxidation because only the non-reducing terminal Gal, GlcNAc, and Xyl and 4-substituted Xyl and 2-substituted Rha residues observed in the methylation analysis possess the necessary vicinal diols to undergo the reaction. We therefore performed a Smith degradation, which consisted of periodate oxidation and borohydride reduction and repurification of the product by octyl-Sepharose chromatography, followed by very mild acid hydrolysis, to destroy and remove all of the aforementioned residues. We checked the efficiency of this treatment by performing a composition analysis, which showed that the monosaccharide ratios were barely changed except for the almost complete destruction of Xyl (data not shown). A subsequent methylation linkage analysis of the Smith degraded and control sample (Table 5) was performed. The results were normalized to 3-substituted Gal that comes from the inter-LacNAc/LNB linkages (*i.e.* 3/4GlcNAcβ1–3Gal) and is therefore expected to be completely resistant to Smith degradation.

Comparison of the Smith degraded and control samples showed the almost complete loss of terminal- and 4-substituted Xyl, a very significant reduction in 2,3-disubstituted Rha, and concomitant increase in 3-substituted Rha. These data suggest that Xylβ1–2Rha linkages are common and that, therefore, the polyrhamnose backbone is probably predominantly made of $(\text{Rha}\alpha1\text{–}3\text{Rha})_n$. Further, the substantial loss of terminal Gal residues and appearance of terminal GlcNAc residues after Smith degradation (Table 5) is consistent with the presence of multiple Gal-GlcNAc-terminating side chains. Finally, the appearance of terminal Rha residues after Smith degradation is consistent with the removal of the Xyl residues from the proposed Xyl-Rha- sequences branching off the poly-LacNAc/LNB chains, and the negligible change in the level of the 3,4-substituted Gal suggests that essentially all of the Rha residues branching off the poly-LacNAc/LNB chains are substituted by Xyl.

Structural Model for TvLG—The model presented in Fig. 7 is consistent with the data presented here and with previous observations of Singh and colleagues (21). The intact molecule and the products of the mixed β-galactosidase and β-hexosaminidase digest and of the Smith degradation experiments are

FIGURE 4. **Selected positive ion ESI-MS spectra of partially hydrolyzed, reduced, and permethylated TvLG.** *A*, product ion spectrum of the $[M + Na]^+$ precursor ion at *m/z* 1433 from the partially hydrolyzed, reduced, and permethylated TvLG sample. *Asterisks* in the *diagram* and *spectrum* indicate masses that are unique to the structures shown. *B*, product ion spectrum of the $[M + Na]^+$ precursor ion at *m/z* 793 from the partially hydrolyzed, reduced, and permethylated TvLG sample. The annotated structures indicate the ion assignments. Cross-ring cleavages are indicated by a *line* bisecting the respective monosaccharide, and the type of cleavage is identified *below* the mass value. The residues Gal and GlcNAc are represented by *open circles* and *open squares*, respectively. The *O between* monosaccharides in the diagram indicates a glycosidic linkage, and *OL* represents a methylated reduced terminus.

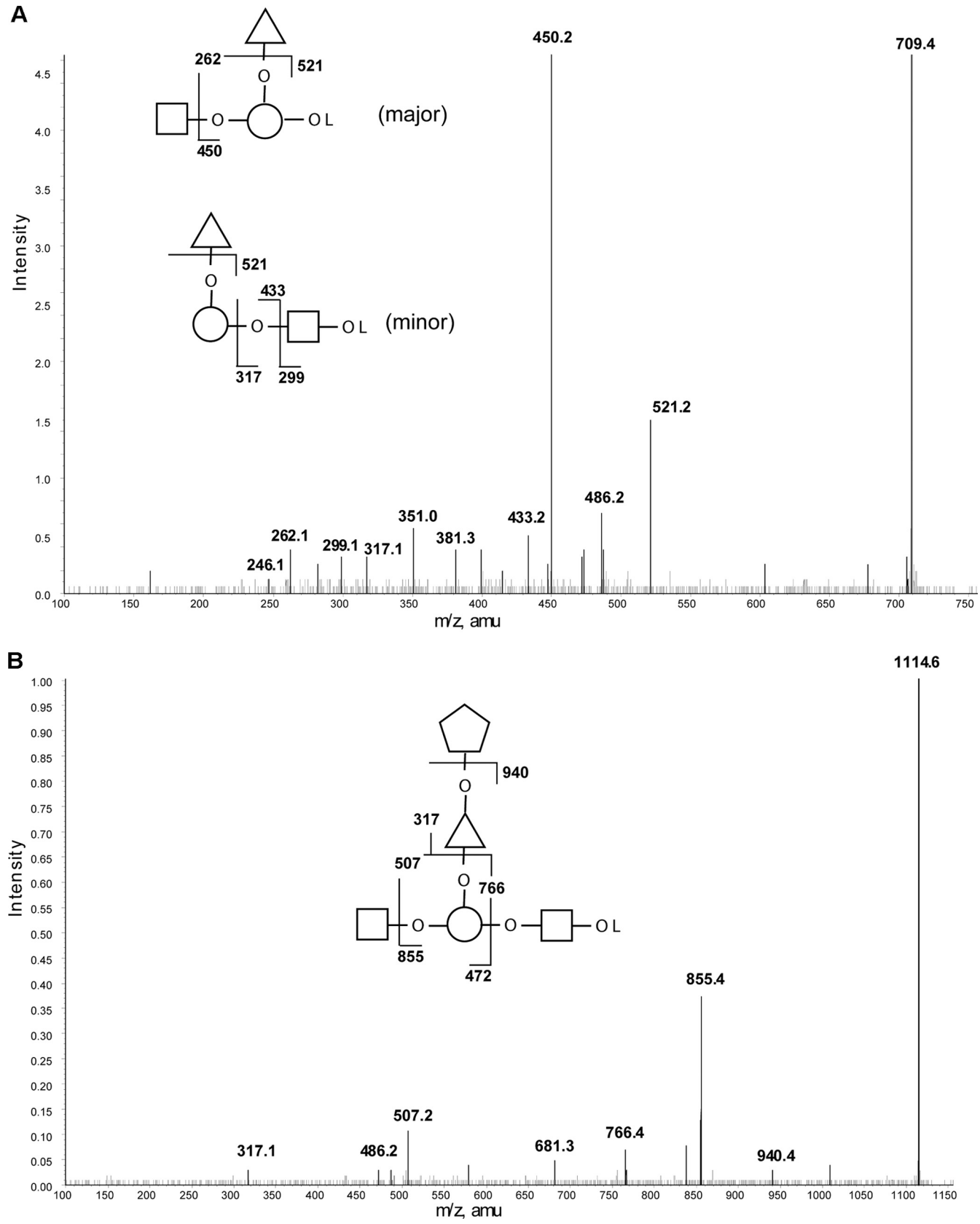


FIGURE 5. **Selected positive ion ESI-MS spectra of partially hydrolyzed, reduced, and permethylated TvLG.** *A*, product ion spectrum of the $[M + Na]^+$ precursor ion at *m/z* 709 from the partially hydrolyzed, reduced, and permethylated TvLG sample. *B*, product ion spectrum of the $[M + Na]^+$ precursor ion at *m/z* 1114 from the partially hydrolyzed, reduced, and permethylated TvLG sample. The annotated structures indicate the ion assignments. The residues Gal, GlcNAc, Xyl, and Rha are represented by *open circles*, *open squares*, *open pentagons*, and *open triangles*, respectively. The *O between* monosaccharides in the diagram indicates a glycosidic linkage, and *OL* represents a methylated reduced terminus.

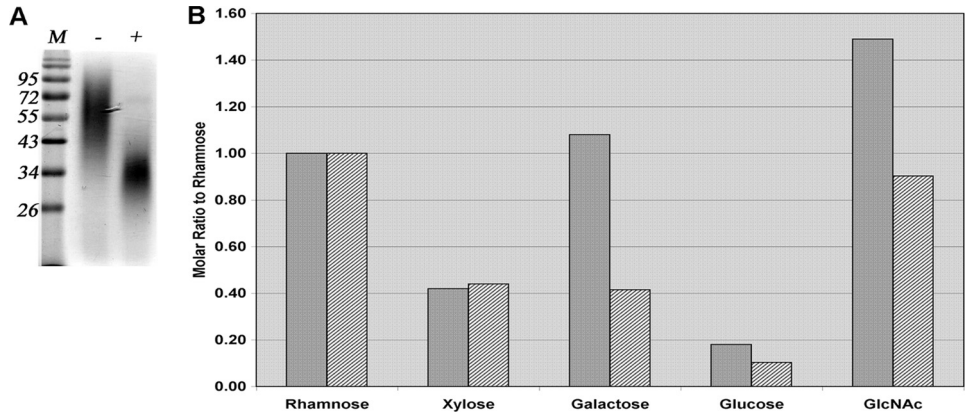


FIGURE 6. **The effects of exoglycosidase digestion on TvLG.** *A*, TvLG digested by β-galactosidase and β-*N*-acetylhexosaminidase (+) and mock-treated TvLG (−) were resolved by SDS-PAGE on a 12.5% acrylamide gel along with prestained molecular weight marker (*M*). The gel was stained with periodate-Schiff reagent to visualize the carbohydrates. *B*, the monosaccharide composition of digested (*hatched bars*) and mock-treated (*filled bars*) TvLG was determined, and the molar amount of each monosaccharide was calculated and presented as a ratio to rhamnose.

TABLE 4
Methylation linkage analysis of native and exoglycosidase-digested TvLG

Partially methylated alditol acetate derivative	Origin	Peak area ratio[a]	
		Native TvLG	Exoglycosidase[b]-digested TvLG
1,5-Di-*O*-acetyl-2,3,4-tri-*O*-methyl-1-[^{2}H]xylitol	Terminal Xyl	0.19	0.25
1,4,5-Tri-*O*-acetyl-2,3-di-*O*-methyl-1-[^{2}H]xylitol	4-Substituted Xyl	0.15	0.23
1,5-Di-*O*-acetyl-2,3,4-tri-*O*-methyl-1-[^{2}H]rhamnitol	Terminal Rha	Trace	0.1
1,2,5-Tri-*O*-acetyl-3,4-di-*O*-methyl-1-[^{2}H]rhamnitol	2-Substituted Rha	0.12	0.17
1,3,5-Tri-*O*-acetyl-2,4-di-*O*-methyl-1-[^{2}H]rhamnitol	3-Substituted Rha	0.09	0.12
1,2,3,5-Tetra-*O*-acetyl-4-*O*-methyl-1-[^{2}H]rhamnitol	2,3-Disubstituted Rha	1	1
1,2,4,5-Tetra-*O*-acetyl-3-*O*-methyl-1-[^{2}H]rhamnitol	2,4-Disubstituted Rha	0.18	0.24
1,5-Di-*O*-acetyl-2,3,4-tri-*O*-methyl-1-[^{2}H]galactitol	Terminal Gal	**0.2**	**Trace**
1,3,5-Tri-*O*-acetyl-2,4-di-*O*-methyl-1-[^{2}H]galactitol	3-Substituted Gal	**0.53**	**0.13**
1,3,4,5-Tetra-*O*-acetyl-2-*O*-methyl-1-[^{2}H]galactitol	3,4-Disubstituted Gal	0.17	0.19
1,5-Di-*O*-acetyl-3,4-di-*O*-methyl-2-*N*-methylacetamido-1-[^{2}H]glucosamine	Terminal GlcNAc	0.1	0.12
1,3,5-Tri-*O*-acetyl-,4-*O*-methyl-2-*N*-methylacetamido-1-[^{2}H]glucosamine	3-Substituted GlcNAc	**0.07**	**Trace**
1,4,5-Tri-*O*-acetyl-,3-*O*-methyl-2-*N*-methylacetamido-1-[^{2}H]glucosamine	4-Substituted GlcNAc	**0.36**	**Trace**

[a] Total ion current area relative to that for 1,2,3,5-tetra-*O*-acetyl-4-*O*-methyl-1-[^{2}H]rhamnitol.
[b] Digested with a mixture of β-galactosidase and β-hexosaminidase. The data in boldface type represent the largest changes.

shown in Fig. 7, *A*, *B*, and *C*, respectively. The salient features are that the molecule is probably based on an α1–3-linked polyrhamnose backbone substituted at the 2-position with single βXyl residues and/or βXyl residues further substituted at the 4-position with additional residues. The polyrhamnose backbone is also substituted with poly-LacNAc/LNB chains of varying length (on average around 5 Gal-GlcNAc units) that are occasionally substituted with Xyl-Rha side chains on the Gal residues, some of which could be further substituted at the 4-position of the Xyl residue. The details of how this oligosaccharide region is linked to the inositol phosphoceramide lipid component have not been resolved. Further, although the model presented probably accounts for the majority of the TvLG general structure, there will probably be less frequent structural variations and complexities in this complex molecule. For example, we have not accounted for the small amounts of 2,4-disubstituted Rha and 2-substituted Rha seen in the methylation linkage analysis of native TvLG or for the α-galactose residues observed by NMR. With respect to the latter, we used coffee bean α-galactosidase to investigate whether it could reduce the apparent molecular weight of intact TvLG or of the β-galactosidase/β-hexosaminidase-treated material along with various other enzyme combinations, but no effect was observed (supplemental Fig. S2).

The LacNAc Side Chains Mediate the Interaction of T. vaginalis with Human Vaginal Epithelial Cells—Several studies have demonstrated that purified TvLG inhibits the binding of *T. vaginalis* to host cells *in vitro* (25, 26). Because our previous studies demonstrated that host cell galectin-1, a protein that preferentially binds poly-LacNAc/LNB (40), plays a role in host-parasite interactions (27), we postulated that the poly-LacNAc/LNB chains, shown by our data and model in this work, are involved in the *T. vaginalis* binding to host cells. To test this hypothesis, we utilized *in vitro* adherence assays (25, 26) in which isolated TvLG is preincubated with vaginal ectocervical cells prior to the addition of labeled parasites. Previous studies have shown that TvLG inhibits up to 40% of *T. vaginalis* binding to host cells (25). To investigate the role of poly-LacNAc/LNB chains in binding inhibition, we performed assays using TvLG isolated by methods previously used for competition assays (25, 26). This TvLG was digested with β-galactosidase and β-hexosaminidase to remove the outermost LacNAc/LNB chains and was shown to have the same decrease in apparent molecular weight and comparable loss of galactose and GlcNAc as observed in Figs. 2 and 7 (data not shown). The ability of digested and undigested TvLG to compete for parasite binding was compared. We found that the addition of undigested TvLG to adhesion assays caused an ~40% decrease in

TABLE 5
Methylation linkage analysis of control and Smith degraded TvLG

Partially methylated alditol acetate derivative	Origin	Peak area ratio[a]	
		Control TvLG	Smith degraded TvLG
1,5-Di-*O*-acetyl-2,3,4-tri-*O*-methyl-1-[^{2}H]xylitol	Terminal Xyl	**0.34**	**0**
1,4,5-Tri-*O*-acetyl-2,3-di-*O*-methyl-1-[^{2}H]xylitol	4-Substituted Xyl	**0.27**	**0**
1,5-Di-*O*-acetyl-2,3,4-tri-*O*-methyl-1-[^{2}H]rhamnitol	Terminal Rha	**Trace**	**0.39**
1,2,5-Tri-*O*-acetyl-3,4-di-*O*-methyl-1-[^{2}H]rhamnitol	2-Substituted Rha	0.23	0.17
1,3,5-Tri-*O*-acetyl-2,4-di-*O*-methyl-1-[^{2}H]rhamnitol	3-Substituted Rha	**0.11**	**0.5**
1,2,3,5-Tetra-*O*-acetyl-4-*O*-methyl-1-[^{2}H]rhamnitol	2,3-Disubstituted Rha	**1.37**	**0.36**
1,2,4,5-Tetra-*O*-acetyl-3-*O*-methyl-1-[^{2}H]rhamnitol	2,4-Disubstituted Rha	0.18	0.24
1,5-Di-*O*-acetyl-2,3,4-tri-*O*-methyl-1-[^{2}H]galactitol	Terminal Gal	**1.25**	**0.23**
1,3,5-Tri-*O*-acetyl-2,4-di-*O*-methyl-1-[^{2}H]galactitol	3-Substituted Gal	1	1
1,3,4,5-Tetra-*O*-acetyl-2-*O*-methyl-1-[^{2}H]galactitol	3,4-Disubstituted Gal	0.42	0.43
1,5-Di-*O*-acetyl-3,4-di-*O*-methyl-2-*N*-methylacetamido-1[^{2}H]glucosamine	Terminal GlcNAc	**0.36**	**0.59**
1,3,5-Tri-*O*-acetyl-,4-*O*-methyl-2-*N*-methylacetamido-1-[^{2}H]glucosamine	3-Substituted GlcNAc	0.19	0.17
1,4,5-Tri-*O*-acetyl-,3-*O*-methyl-2-*N*-methylacetamido-1-[^{2}H]glucosamine	4-Substituted GlcNAc	**1.38**	**0.88**

[a] Total ion current area relative to that for 1,3,5-tri-*O*-acetyl-2,4-di-*O*-methyl-1-[^{2}H]galactitol. The data in boldface type represent the largest changes.

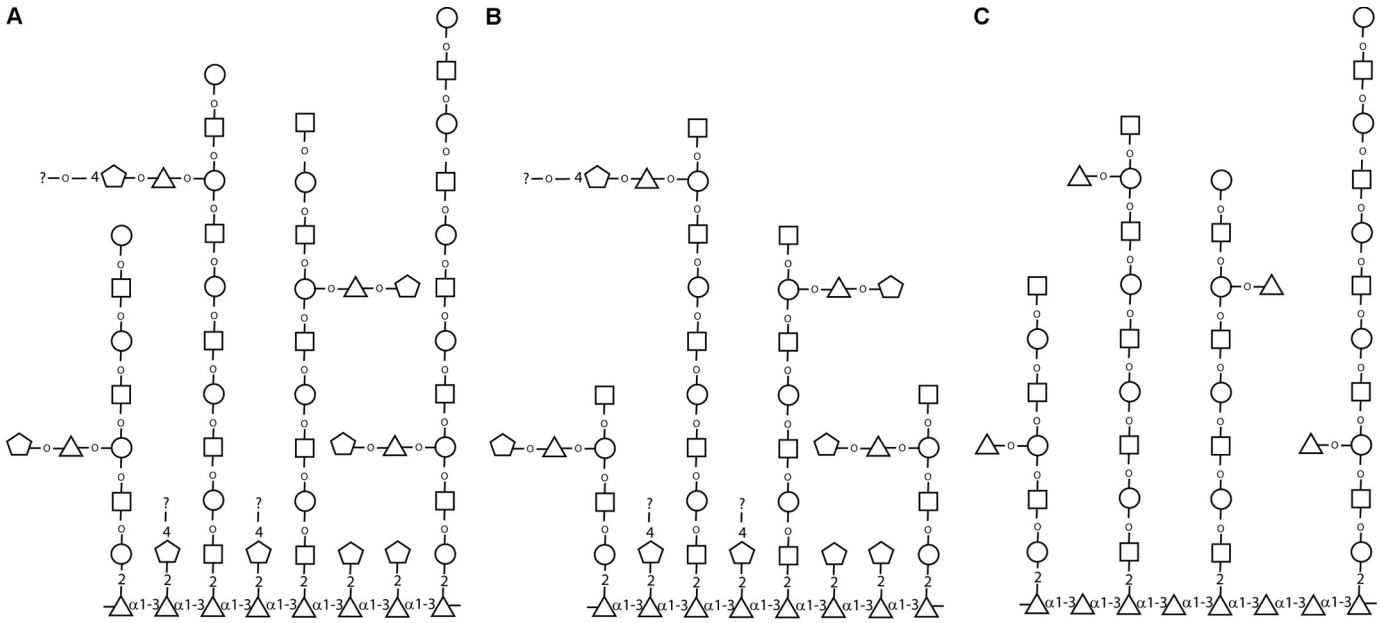


FIGURE 7. **Structural model for TvLG.** Presented are representative structures of TvLG: of native TvLG (*A*), of TvLG following digestion by β-galactosidase and β-hexosaminidase (*B*), and of TvLG following Smith degradation (*C*). Gal, GlcNAc, Xyl, and Rha are represented by *open circles*, *squares*, *pentagons*, and *triangles*, respectively. Anomericity and substitutions are indicated where known.

binding of *T. vaginalis* to vaginal ectocervical cells, whereas digested TvLG decreased binding by less than 10% (Fig. 8). This greater than 4-fold difference in binding inhibition indicates that the poly-LacNAc/LNB chains on TvLG play a role in parasite binding to host cells. To ensure a lack of contaminating proteins that might interfere with binding, the TvLG samples used in the assays were electrophoresed, and silver staining showed that no contaminating protein was present in any of them. Furthermore, heat treatment, additional proteinase treatment, or octyl-Sepharose purification of TvLG had no effect on inhibition of attachment (data not shown), indicating the lack of inhibition by contaminating protein.

DISCUSSION

The presence of glycans on the surface of various viruses, bacteria, fungi, and protozoa is critical for mediating interactions of pathogenic microbes with their host cells (40). Here we have determined the partial structure of the major surface glycoconjugate of the parasitic protozoan *T. vaginalis* and further investigated the interaction of this molecule with host cells. We demonstrate that this surface glycan, termed TvLG, is probably composed of an α1–3-linked polyrhamnose backbone substituted at the 2-position with either single βXyl residues or βXyl further substituted at the 4-position by additional residues and by poly-LacNAc/LNB chains of varying length that are occasionally substituted with Xyl-Rha side chains on the Gal residues. Furthermore, we present evidence that the poly-LacNAc/LNB chains are directly involved in attachment of *T. vaginalis* to host cells.

The *Trichomonas* glycan was originally identified as an LPG similar to *Leishmania* LPG (22). However, *Leishmania* LPG and the *T. vaginalis* glycan have since been found to have little in common beyond being lipid-anchored surface glycans. The compositions of the two differ significantly, with the *Leishmania* LPG being a phosphoglycan chain composed of repeating PO_4-6Galβ1–4Man units (41). The structure of TvLG, presented here, is considerably larger and more complex. Moreover, the TvLG polysaccharide chain does not contain phosphate as previously thought, hence the change in nomenclature

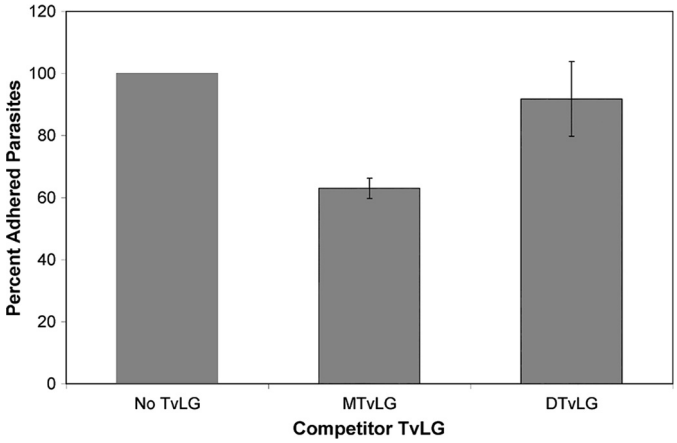


FIGURE 8. **Adherence of *T. vaginalis* strain B7RC2 to human vaginal ectocervical cells is inhibited by full-length but not partially digested TvLG.** Human vaginal ectocervical cells were preincubated in media alone (*No TvLG*), mock-treated TvLG (*MTvLG*), or TvLG digested with β-*N*-acetylhexosaminidase and β-galactosidase (*DTvLG*) for 1 h before the addition of labeled parasites as described under "Experimental Procedures." The average number of parasites per coverslip was calculated, and the number determined in the absence of TvLG (*No TvLG*) was set as 100% adherence. Results presented are the average of three different experiments performed with a total of nine coverslips. *Error bars*, S.D.

from LPG to TvLG. This *T. vaginalis* molecule is also unlike the LPPG found on the parasite *Entamoeba histolytica*. This surface glycan is composed of a Gal_1Man_2GlcN-myoinositol backbone modified with 1–20 αGal residues and linked by an ethanolamine to a highly acidic polypeptide backbone with serine residues modified with chains of $[Glc\alpha1-6]_nGlc\beta1-6Gal$ (n = 2–23) (42). The finding that *T. vaginalis*, *E. histolytica*, and *Leishmania* surface glycans vary greatly in sugar composition and structure is consistent with the evolutionary divergence of these parasites and their different modes of transmission and localization within hosts (43, 44).

A key feature of our structural model of TvLG is an α1–3-linked polyrhamnose backbone with side chains of β-Xyl and LacNAc/LNB repeats, which, to our knowledge, is unique to *Trichomonas* and the first example of a polyrhamnan described outside bacteria and plants. Rhamnose is a critical component of the peptidoglycan cell wall of bacteria, and, in the case of *Mycobacterium tuberculosis*, rhamnose synthetic enzymes are being developed as drug targets (45). A similar rhamnose synthesis pathway is found in *T. vaginalis* (25) and is absent in the human host. The lack of rhamnose synthesis by humans makes its synthesis and transfer an appealing target for chemotherapeutic attack. Therefore, identifying and confirming the synthesis pathways as well as transferases for rhamnose and xylose may allow the development of new drugs to target *T. vaginalis*. Targeting rhamnose and xylose synthesis and/or transferases is likely to radically disrupt TvLG, either destabilizing or truncating the structure. As discussed below, loss or alteration of TvLG would prevent or greatly inhibit attachment of this extracellular parasite to host cells. Changes in TvLG structure could also expose immunogenic surface molecules that are normally hidden by the dominant glycocalyx. Unlike rhamnose, xylose is synthesized in humans; however, it is generally found modifying proteins (proteoglycans) linked β-1 to serine with carbohydrate substitutions at the 4-position (46). This linkage to serine rather than rhamnose may allow for targeting the xylosyltransferase(s) of *T. vaginalis* without damaging the host.

The presence of *N*-acetyllactosamine (-3Galβ1-4GlcNAcβ1-), or LacNAc, repeats in TvLG was first suggested by the finding that TvLG acted as a substrate of galectin-1 binding (27). Later, LacNAc was observed directly through ESI-MS/MS (28). Here we have extended these analyses by demonstrating the presence of LacNAc in chemically fragmented TvLG by ESI-MS and conducting methylation linkage analyses that suggest the presence of small amounts of lacto-*N*-biose (-3Galβ1–3GlcNAcβ1-), LNB. These data are consistent with the presence of poly-LacNAc or LNB of varying length with an average size of 5 units. Furthermore, enzymatic digestions demonstrate that these poly-LacNAc/LNB chains are on branches radiating from a central rhamnose core.

Our structural model predicts that the reduction of poly-LacNAc/LNB chains by enzymatic digestion will result in a corresponding loss in the ability of isolated TvLG to inhibit binding of parasites to host cells. Indeed, this was observed when comparing the ability of undigested and digested TvLG to inhibit parasite binding (Fig. 8). These data demonstrate that LacNAc repeats are important for the attachment of *T. vaginalis* to host epithelium. This finding is consistent with and corroborates our previous observations that an increase in galectin-1 in human epithelial cells increases parasite binding, whereas a decrease in galectin-1 decreases the binding of *T. vaginalis* to host cells (27). Together these data indicate that the interaction between galectin and the TvLG polylactosamine repeats is critical for the attachment of *T. vaginalis* to host cells.

In summary, we have provided a detailed model of the structure of the major surface glycan of *T. vaginalis*. Its unique structure is unusually complex and provides insight into mechanisms underlying the attachment of this parasite to its human host. This study sets the stage for future experiments focused on determining whether the TvLG of highly adherent and poorly adherent *T. vaginalis* strains exhibit significant differences.

Acknowledgments—We thank Giulia Bandini and Gina Mackay for technical support.

REFERENCES

1. World Health Organization (1996) *Sexually Transmitted Diseases (STDs)*, Fact Sheet 110, World Health Organization, Geneva
2. Petrin, D., Delgaty, K., Bhatt, R., and Garber, G. (1998) *Clin. Microbiol. Rev.* **11,** 300–317
3. Soper, D. (2004) *Am. J. Obstet. Gynecol.* **190,** 281–290
4. Gardner, W. A., Jr., Culberson, D. E., and Bennett, B. D. (1986) *Arch. Pathol. Lab. Med.* **110,** 430–432
5. Minkoff, H., Grunebaum, A. N., Schwarz, R. H., Feldman, J., Cummings, M., Crombleholme, W., Clark, L., Pringle, G., and McCormack, W. M. (1984) *Am. J. Obstet. Gynecol.* **150,** 965–972
6. Cotch, M. F., Pastorek, J. G., 2nd, Nugent, R. P., Hillier, S. L., Gibbs, R. S., Martin, D. H., Eschenbach, D. A., Edelman, R., Carey, J. C., Regan, J. A., Krohn, M. A., Klebanoff, M. A., Rao, A. V., and Rhoads, G. G. (1997) *Sex. Transm. Dis.* **24,** 353–360
7. Lloyd, G. L., Case, J. R., De Frias, D., and Brannigan, R. E. (2003) *J. Urol.* **170,** 924
8. Sorvillo, F., and Kerndt, P. (1998) *Lancet* **351,** 213–214
9. Van Der Pol, B., Kwok, C., Pierre-Louis, B., Rinaldi, A., Salata, R. A., Chen,

P. L., van de Wijgert, J., Mmiro, F., Mugerwa, R., Chipato, T., and Morrison, C. S. (2008) *J. Infect. Dis.* **197,** 548–554
10. Yap, E. H., Ho, T. H., Chan, Y. C., Thong, T. W., Ng, G. C., Ho, L. C., and Singh, M. (1995) *Genitourin. Med.* **71,** 402–404
11. Zhang, Z. F., and Begg, C. B. (1994) *Int. J. Epidemiol.* **23,** 682–690
12. Sutcliffe, S., Giovannucci, E., Alderete, J. F., Chang, T. H., Gaydos, C. A., Zenilman, J. M., De Marzo, A. M., Willett, W. C., and Platz, E. A. (2006) *Cancer Epidemiol. Biomarkers Prev.* **15,** 939–945
13. de Miguel, N., Lustig, G., Twu, O., Chattopadhyay, A., Wohlschlegel, J. A., and Johnson, P. J. (2010) *Mol. Cell Proteomics* **9,** 1554–1566
14. Hirt, R. P., Noel, C. J., Sicheritz-Ponten, T., Tachezy, J., and Fiori, P. L. (2007) *Trends Parasitol.* **23,** 540–547
15. Alderete, J. F., Millsap, K. W., Lehker, M. W., and Benchimol, M. (2001) *Cell Microbiol.* **3,** 359–370
16. Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., Wortman, J. R., Bidwell, S. L., Alsmark, U. C., Besteiro, S., Sicheritz-Ponten, T., Noel, C. J., Dacks, J. B., Foster, P. G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G. J., Westrop, G. D., Müller, S., Dessi, D., Fiori, P. L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F. D., Simoes-Barbosa, A., Brown, M. T., Hayes, R. D., Mukherjee, M., Okumura, C. Y., Schneider, R., Smith, A. J., Vanacova, S., Villalvazo, M., Haas, B. J., Pertea, M., Feldblyum, T. V., Utterback, T. R., Shu, C. L., Osoegawa, K., de Jong, P. J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S. B., Logsdon, J. M., Jr., Henze, K., Gupta, A., Wang, C. C., Dunne, R. L., Upcroft, J. A., Upcroft, P., White, O., Salzberg, S. L., Tang, P., Chiu, C. H., Lee, Y. S., Embley, T. M., Coombs, G. H., Mottram, J. C., Tachezy, J., Fraser-Liggett, C. M., and Johnson, P. J. (2007) *Science* **315,** 207–212
17. Warton, A., and Honigberg, B. M. (1980) *J. Protozool.* **27,** 410–419
18. Wartoń, A., and Honigberg, B. M. (1983) *Z. Parasitenkd.* **69,** 149–159
19. Choromańśki, L., Beat, D. A., Nordin, J. H., Pan, A. A., and Honigberg, B. M. (1985) *Z. Parasitenkd.* **71,** 443–458
20. Warton, A., Kon, V. B., and Papadimitriou, J. M. (1988) *J. Electron Microsc.* **37,** 134–140
21. Gilbert, R. O., Elia, G., Beach, D. H., Klaessig, S., and Singh, B. N. (2000) *Infect. Immun.* **68,** 4200–4206
22. Singh, B. N. (1993) *Mol. Biochem. Parasitol.* **57,** 281–294
23. Singh, B. N. (1994) *Parasitol. Today* **10,** 152–154
24. Singh, B. N., Lucas, J. J., Beach, D. H., Shin, S. T., and Gilbert, R. O. (1999) *Infect. Immun.* **67,** 3847–3854
25. Bastida-Corcuera, F. D., Okumura, C. Y., Colocoussi, A., and Johnson, P. J. (2005) *Eukaryot. Cell* **4,** 1951–1958
26. Fichorova, R. N., Trifonova, R. T., Gilbert, R. O., Costello, C. E., Hayes, G. R., Lucas, J. J., and Singh, B. N. (2006) *Infect. Immun.* **74,** 5773–5779
27. Okumura, C. Y., Baum, L. G., and Johnson, P. J. (2008) *Cell Microbiol.* **10,** 2078–2090
28. Singh, B. N., Hayes, G. R., Lucas, J. J., Sommer, U., Viseux, N., Mirgorodskaya, E., Trifonova, R. T., Sassi, R. R., Costello, C. E., and Fichorova, R. N. (2009) *Glycoconj. J.* **26,** 3–17
29. McConville, M. J., Thomas-Oates, J. E., Ferguson, M. A., and Homans, S. W. (1990) *J. Biol. Chem.* **265,** 19611–19623
30. McConville, M. J., and Ferguson, M. A. (1993) *Biochem. J.* **294**, 305–324
31. Turco, S. J., and Descoteaux, A. (1992) *Annu. Rev. Microbiol.* **46,** 65–94
32. Singh, B. N., Beach, D. H., Lindmark, D. G., and Costello, C. E. (1994) *Arch. Biochem. Biophys.* **309,** 273–280
33. Costello, C. E., Glushka, J., van Halbeek, H., and Singh, B. N. (1993) *Glycobiology* **3,** 261–269
34. Clark, C. G., and Diamond, L. S. (2002) *Clin. Microbiol. Rev.* **15,** 329–341
35. Fichorova, R. N., Rheinwald, J. G., and Anderson, D. J. (1997) *Biol. Reprod.* **57,** 847–855
36. McConville, M. J., Bacic, A., Mitchell, G. F., and Handman, E. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84,** 8941–8945
37. Ferguson, M. A. J. (1992) in *Lipid Modifications of Proteins: A Practical Approach* (Turner, A. J., and Hooper, N. M., eds) pp. 191–230, IRL Press, Oxford
38. Abdel Akher, M., Cadotte, J. E., Montgomery, R., Smith, F., Vancleve, J. W., and Lewis, B. A. (1953) *Nature* **171,** 474–475
39. Ferguson, M. A., Murray, P., Rutherford, H., and McConville, M. J. (1993) *Biochem. J.* **291**, 51–55
40. Vasta, G. R. (2009) *Nat. Rev. Microbiol.* **7,** 424–438
41. McConville, M. J., Schnur, L. F., Jaffe, C., and Schneider, P. (1995) *Biochem. J.* **310**, 807–818
42. Moody-Haupt, S., Patterson, J. H., Mirelman, D., and McConville, M. J. (2000) *J. Mol. Biol.* **297,** 409–420
43. Keeling, P. J., Burger, G., Durnford, D. G., Lang, B. F., Lee, R. W., Pearlman, R. E., Roger, A. J., and Gray, M. W. (2005) *Trends Ecol. Evol.* **20,** 670–676
44. Hampl, V., Hug, L., Leigh, J. W., Dacks, J. B., Lang, B. F., Simpson, A. G., and Roger, A. J. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 3859–3864
45. Ma, Y., Stern, R. J., Scherman, M. S., Vissa, V. D., Yan, W., Jones, V. C., Zhang, F., Franzblau, S. G., Lewis, W. H., and McNeil, M. R. (2001) *Antimicrob. Agents Chemother.* **45,** 1407–1416
46. Wilson, I. B. (2004) *Cell Mol. Life Sci.* **61,** 794–809