THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 266, No. 35, Issue of December 15, pp. 23670-23675, 1991
Printed in U.S.A.

Complete Structure of the Glycan of Lipopeptidophosphoglycan from *Trypanosoma cruzi* Epimastigotes*

(Received for publication, July 1, 1991)

Rosa M. de Lederkremer, Carlos Lima, and Maria I. Ramirez

From the Departamento de Quimica Organica, Facultad de Ciencas Exactas y Naturales, Universidad de Buenos Aires, Argentina

Michael A. J. Ferguson, Steve W. Homans, and Jane Thomas-Oates‡

From the Department of Biochemistry, University of Dundee, Dundee DD1 4HN, United Kingdom

The lipopeptidophosphoglycan is the major cell surface glycoconjugate of the epimastigote forms of the parasitic protozoan *Trypanosoma cruzi*. A detailed partial structure for this molecule has been reported (Previato, J. O., Gorin, P. A. J., Mazurek, M., Xavier, M. T., Fournet, B., Wieruszesk, J. M., and Mendonca-Previato, L. (1990) *J. Biol. Chem.* 265, 2518–2526). In this study, we complete the primary structure assignments and describe the microheterogeneity found in the lipopeptidophosphoglycan glycan, using a combination of ^{1}H and ^{31}P NMR, fast atom bombardment mass spectrometry, methylation linkage analysis, and exoglycosidase sequencing. The lipopeptidophosphoglycan is a glycosylated inositol-phosphoceramide with striking homology to glycosylphosphatidylinositol membrane anchors found attached to a wide variety of plasma membrane proteins throughout the eukaryotes.

The parasitic protozoan *Trypanosoma cruzi* is the causative agent of Chagas' disease (South American trypanosomiasis). The organism undergoes a complex life cycle between a wide variety of mammalian hosts and biting insect vectors (reduviid bugs). In the infected mammal, the parasite exists as intracellular dividing amastigote forms in tissues such as smooth and cardiac muscle. These forms give rise to extracellular nondividing bloodstream trypomastigote forms that spread the infection. The ingestion of trypomastigotes by the insect vector results in the differentiation of the parasite to the dividing epimastigote form, which efficiently colonizes the insect midgut. Migration of parasites to the insect hindgut results in their differentiation to metacyclic trypomastigote forms that are adapted for transmission to a mammalian host via fecal contamination of fresh wounds or mucus membranes.

The lipopeptidophosphoglycan (LPPG)[1] is the most abundant glycoconjugate of the epimastigote form of the parasite; a typical yield of LPPG is about 100 mg extracted from 2×10^{12} cells (Lederkremer *et al.*, 1990). Assuming a molecular weight of around 1890 (Previato *et al.* (1990) and this study) this suggests a minimum copy number of around 1.5×10^{7} LPPG molecules/epimastigote. The function of LPPG is unknown, but it appears to be restricted to the epimastigote forms of the parasite (Zingales *et al.*, 1982).

The LPPG fraction contains mannose, galactofuranose, 2-aminoethylphosphonate (AEP), *myo*-inositol, phosphate, long chain bases, and fatty acids, together with traces of glucose and amino acids (Lederkremer *et al.*, 1978; Ferguson *et al.*, 1981, 1985; Lederkremer *et al.*, 1985). The lipid component is an inositol-phosphoceramide containing mainly palmitoylsphinganine, palmitoylsphingosine, and lignoceroylsphinganine. The ceramide can be released by phosphatidylinositol (PI)-specific phospholipase C and the glycan chain is attached to the inositol ring via a glycosidic linkage to a non-*N*-acetylated glucosamine (GlcN) residue (Lederkremer *et al.*, 1990). These two latter features indicate that it is closely related to the glycosylphosphatidylinositol (GPI) membrane anchors common to many cell surface glycoproteins throughout the eukaryotes (reviewed recently by Thomas *et al.* (1990); Cross, 1990).

A detailed structure of the LPPG glycan was recently reported by Previato *et al.* (1990). In this paper, we confirm these structural features and complete the glycan structure by supplying (i) the nature of the glycosidic linkage between glucosamine and *myo*-inositol, (ii) the precise location of both galactofuranose residues, and (iii) the degree of heterogeneity in the glycan structure. The LPPG structure is discussed in the context of general and parasite-specific GPI metabolism.

DISCUSSION[2]

The analysis of the LPPG glycan moiety, generated by PI-specific phospholipase C cleavage of the ceramide lipid, presented some difficult problems. Attempts to fractionate the different glycan species by Dionex carbohydrate high pressure liquid chromatography produced a confusing array of peaks (data not shown) due to the presence of three major glycan species and compounded by the heterogeneity of the inositol phosphorylation state. Deamination and reduction was used to introduce a labeled 2,5-anhydromannitol terminus and to remove the inositol phosphate. However, Dionex high pres-

* This investigation received financial support from the United Nations Development Program/World Bank World Health Organization special program for research and training in tropical diseases, Secretaria de Ciencia y Tecnica, Consejo Nacional de Investigaciones, and Universidad de Buenos Aires (to R. M. L.) and from The Wellcome Trust (to M. A. J. F. and S. W. H). 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.

‡ Beit Memorial Research Fellow. Present address: Bijvoet Centre for Biomolecular Research, Faculty of Chemistry, University of Utrecht, P. O. Box 80083, 3508 TB Utrecht, The Netherlands.

[1] The abbreviations used are: LPPG, lipopeptidophosphoglycan; AEP, 2-aminoethylphosphonate; PI, phosphatidylinositol; GPI, glycosylphosphatidylinositol; FAB, fast atom bombardment; MS, mass spectrometry.

[2] Portions of this paper (including "Materials and Methods," "Results," Figs. 1–5, and Tables 1 and 2) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

sure liquid chromatography of this fraction also produced multiple products (data not shown) due to the simultaneous deamination of the AEP group to a mixture of ethanolphosphonate, phosphate, and other unidentified products. One useful piece of information that did arise from analysis of the deaminated material was the detection of the 2,5-anhydromannitol-6-ethanolphosphonate derivative by gas chromatograhy-MS, confirming the NMR assignment of $GlcNH_2$-6-AEP in the original material. Due to the complications described above, the PI-specific phospholipase C-generated head groups were examined as a mixture for most of the subsequent analyses.

The structures shown in Fig. 6 are consistent with the NMR, FAB-MS, linkage composition, and exoglycosidase sequencing data and suggest that the LPPG fraction contains three major glycan structures. The largest of these, *Structure 1*, contains 4 Man and 2 Galf residues and constitutes about 65% of the molecules. This structure is comparable with the one described by Previato *et al.* (1990). However, in this case, we were able to assign (i) the $GlcNH_2$-*myo*-inositol linkage as $GlcNH_2\alpha$1–6*myo*-inositol from the NMR nuclear Overhauser effect spectroscopy spectrum of fraction A and methylation analysis of fraction D, (ii) the location of the terminal Galf residues exclusively to the nonreducing terminal and subterminal mannose residues by FAB-MS analysis of fraction C and the NMR COSY spectrum of fraction A, and (iii) the existence of the two smaller species. The existence of the two smaller species was defined by the FAB-MS spectrum of fraction B, and their structures were inferred from the methylation analysis of fraction A and the detection of a Man_3- as well as Man_4-containing mannan cores in fraction E. Assuming that all three LPPG species represent members of a common biosynthetic series, three independent or concomitant pathways to the major product (Structure 1) can be envisaged (Fig. 7). Pulse-chase experiments will be necessary to investigate these possibilities.

The nature of the association of the amino acids (3–5% by weight) that consistently co-purify with LPPG remains obscure. Their absence in fraction B, as judged by FAB-MS, suggests that if they are covalently associated, then the linkage(s) is base-labile. The possibility of ester linkage between the amino acid COOH groups and sugar hydroxyls has been previously discussed (Lederkremer *et al.*, 1985).

The LPPG molecules can be included as members of the GPI family since they contain the structural motif Manα1–4$GlcNH_2\alpha$1–6*myo*-inositol-1-PO_4-lipid. Indeed, they are very closely related to GPI protein-anchor structures (see Fig. 6), although LPPG represents the first example of a defined GPI core structure linked to a ceramide. The presence of AEP has not been found in any other fully defined GPI family members, but it has been described in the glycosylated inositolphosphoceramide of *Acanthamoeba castellani* (Dearborn *et al.*, 1976) and the glycolipids of various water animals such as sea hare (Hayashi, 1990). In the latter example, the AEP is found linked to the 6-position of Gal residues. Interestingly, the GPI anchor of the *T. cruzi* G-strain metacyclic trypomastigote 1G7-antigen glycoprotein is essentially identical with LPPG, except that it does not contain Galf and utilizes a glycerolipid rather than a ceramide.[3] The function of LPPG is unknown. However, it is noteworthy that the related kinetoplastid parasites of the *Leishmania* species also express GPI-related glycophospholipids (lipophosphoglycans and glycoinositolphospholipids) in equally high copy number, around 1–2 × 10^7 molecules/cell, in the insect-dwelling promastigote forms (McConville *et al.* (1990a, 1990b) and references therein). Indeed, it is likely that both *T. cruzi* epimastigotes and *Leishmania* promastigotes present a particularly carbohydrate-rich surface to their environment. This might serve a protective function in the harsh conditions of the digestive tract of their insect vectors. Whether or not LPPG is involved in the adhesion of parasites to the insect gut epithelia, as suggested for *Leishmania* lipophosphoglycan (Davies *et al.*, 1990), remains to be determined. In any case, the expression of LPPG by *T. cruzi* epimastigotes seems to represent another example of the adaptation of the ubiquitous GPI protein-anchor biosynthetic pathway in kinetoplastid parasites.

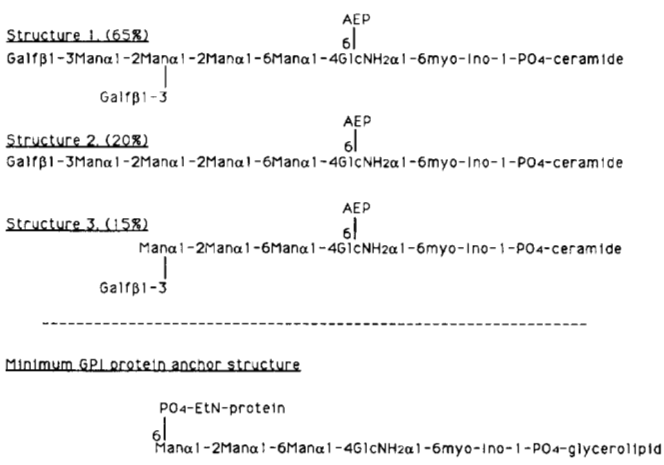


FIG. 6. **Primary structures of the three major LPPG species.** The structure of the simplest known GPI protein anchor (from *Leishmania major* promastigote surface protease (Schneider *et al.*, 1990)) is shown for comparison. *EtN*, ethanolamine.

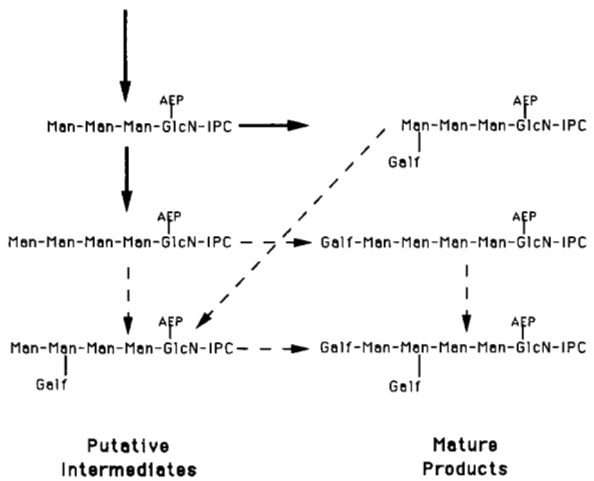


FIG. 7. **Possible biosynthetic interrelationships of the mature LPPG products and their putative intermediates.** *IPC*, inositol-phosphoceramide.

Acknowledgments—We are grateful to Dr. Walter Colli and colleagues (São Paulo, Brazil) for providing the *T. cruzi* cells and to Dr. Martin G. Low (Columbia University, New York) for a gift of phosphatidylinositol-specific phospholipase C from *Bacillus thuringiensis*. We thank Dr. Malcolm McConville for helpful advice.

REFERENCES

Castellani, O., Ribeiro, L. V., and Fernandes, J. F. (1967) *J. Protozool.* **14,** 447–451

Cross, G. A. M. (1990) *Annu. Rev. Cell Biol.* **6,** 1–39

Davies, C. R., Cooper, A. M., Peacock, C., Lane, R. P., and Blackwell,

[3] M. L. S. Güther, M. L. Cardoso de Almeida, N. Yoshida, and M. A. J. Ferguson, manuscript in preparation.

J. M. (1990) *Parasitology* **101,** 337-343
Dearborn, D. G., Smith, S., and Korn, E. D. (1976) *J. Biol. Chem.* **251,** 2976-2982
Dell, A. (1987) *Adv. Carbohydr. Chem. Biochem.* **45,** 19-72
Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith F. (1956) *Anal. Chem.* **28,** 350-356
Ferguson, M. A. J. (1992) in *Lipid Modifications of Proteins: A Practical Approach* (Turner, A. J., and Hooper, N., eds) IRL Press, Oxford, in press
Ferguson, M. A. J., Allen, A. K., and Snary, D. (1982) *Biochem. J.* **207,** 171-174
Ferguson, M. A. J., Snary, D., and Allen, A. K. (1985) *Biochim. Biophys. Acta* **842,** 39-44
Ferguson, M. A. J., Homans, S. W., Dwek, R. A., and Rademacher, T. W. (1988) *Science* **239,** 753-759
Gerwig, G. J., de Waard, P., Kamerling, J. P., Vliegenthart, J. F. G., Morgenstern, E., Lamed, R., and Bayer, E. A. (1989) *J. Biol. Chem.* **264,** 1027-1035
Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F. G., Morag, E., Lamed, R., and Bayer E. A. (1991) *Eur. J. Biochem.* **196,** 115-122
Hayashi, A. (1990) *Trends Glycosci. Glyotechnol.* **2,** 381-386
Lederkremer, R. M., Casal, O. L., Tanaka, C. T., and Colli, W. (1978) *Biochem. Biophys. Res. Commun.* **85,** 1268-1274
Lederkremer, R. M., Casal, O. L., Alves, M. J. M., and Colli, W. (1985) *Biochem. Int.* **10,** 89-96
Lederkremer, R. M., Lima, C., Ramirez, M. I., and Casal, O. L. (1990) *Eur. J. Biochem.* **192,** 337-345
Mayor, S., Menon, A. K., Cross, G. A. M., Ferguson, M. A. J., Dwek, R. A., and Rademacher, T. W. (1990) *J. Biol. Chem.* **265,** 6164-6173
McConville, M. J., Homans, S. W., Thomas-Oates, J. E., Dell, A., and Bacic, A. (1990a) *J. Biol. Chem.* **265,** 7385-7394
McConville, M. J., Thomas-Oates, J. E., Ferguson, M. A. J., and Homans, S. W. (1990b) *J. Biol. Chem.* **265,** 19611-19623
Previato, J. O., Gorin, P. A. J., Mazurek, M., Xavier, M. T., Fournet, B., Wieruszesk, J. M., and Mendonca-Previato, L. (1990) *J. Biol. Chem.* **265,** 2518-2526
Schneider, P., Ferguson, M. A. J., McConville, M. J., Mehlert, A., Homans, S. W., and Bordier, C. (1990) *J. Biol. Chem.* **265,** 16955-16964
Thomas, J. R., Dwek, R. A., and Rademacher, T. W. (1990) *Biochemistry* **29,** 5413-5422
Turco, S. J., Orlandi, P. A., Jr., Homans, S. W., Ferguson, M. A. J., Dwek, R. A., and Rademacher, T. W. (1989) *J. Biol. Chem.* **264,** 6711-6715
Volwerk, J. J., Shashidhar, M. S., Kuppe, A., and Griffith, O. H. (1990) *Biochemistry* **29,** 8056-8062
Zingales, B., Marin, N. F., Lederkremer, R. M., and Colli, W. (1982) *FEBS Lett.* **142,** 238-242

Supplemental Material to:

Complete structure of the glycan of Lipopeptidophosphoglycan from *Trypanosoma cruzi* epimastigotes

Rosa M. de Lederkremer, Carlos Lima and Maria I. Ramirez, Departamento de Quimica Organica, Facultad de Ciencas Exactas y Naturales, Universidad de Buenos Aires, Argentina, and
Michael A.J. Ferguson, Steve W. Homans and Jane Thomas-Oates, Department of Biochemistry, University of Dundee, Scotland.

Materials and Methods

Parasite cultivation and purification of LPPG.

Epimastigote forms of *T. cruzi* (Y strain) were cultivated in LIT medium as described (Castellani *et al* 1967). Lyophilised cells were delipidated by extraction with chloroform/methanol. The LPPG was extracted with water saturated with 1-butanol and purified on octyl-Sepharose as described previously (Lederkremer *et al* 1990)

Preparation of fraction A: PI-PLC digestion and isolation of the LPPG glycan.

A sample of LPPG (13mg) was dissolved in 1.5ml of 50mM buffer Tris-HCl, pH 7.2, containing 0.1% deoxycholate and incubated with 6 units of phospholipase C from *Bacillus thuringiensis* for 3 h at 37°C. The lipid was extracted with ether (3 x 1ml) and the aqueous phase was rechromatographed on octyl-Sepharose. The glycan component (fraction A) was recovered from the exclusion volume. After desalting on BioGel P-2, analysis by the phenol-sulphuric acid method (Dubois *et al* 1956) of the LPPG glycan gave 6mg of neutral sugars.

Preparation of fraction B: N-acetylation and methylation.

A sample of fraction A (100nmole) was dissolved in 200ul saturated $NaHCO_3$ and N-acetylated at 0°C by the addition of 10, 10 and 20ul of acetic anhydride at 10min intervals. After warming to room temperature over 30min the sample was desalted by passage through 0.2ml AG50X12(H^+) and dried twice from 50ul of toluene. The N-acetylated material was dissolved in 50ul dimethyl sulphoxide (DMSO) and treated with 50ul 120mg/ml NaOH suspended in DMSO for 20min with stirring. Methylation was performed by adding 10, 10 and 20ul of methyl iodide at 10min intervals with continual stirring. The reaction was terminated by the addition of 0.25ml chloroform and 1ml 100mg/ml sodium thiosulphate. The N-acetylated, methylated glycans (fraction B) were recovered in the chloroform phase which was washed three times with 1ml water. After drying under vacuum, fraction B was redissolved in methanol and 5% was taken for FAB-MS analysis, which was carried out as previously described (McConville *et al* 1990b).

Preparation of fraction C.

The remainder of fraction B was re-methylated as describao above. The N-acetylated, permethylated glycans were dried and redissolved in 100ul 20% acetonitrile and chromatographed on a C18 Hypersil-ODS column (25 x 0.46cm, Hichrom) eluted with a linear gradient of acetonitrile from 20 to 80% over 90min at 1ml/min. Fractions of 1ml were collected and 2ul aliquots were spotted onto an aluminium backed HPTLC plate. Carbohydrate containing fractions were revealed by α-naphthol, sulphuric acid staining of the HPTLC plate. The major carbohydrate fraction eluting at 53% acetonitrile (fraction C), was dried and redissolved in methanol and 10% used for FAB-MS analysis, which was carried out as described previously (McConville *et al* 1990b).

Preparation of fraction D: aq. HF dephosphorylation and N-acetylation.

A sample of fraction A (10nmole) was dephosphorylated with 50ul 48% aq. HF (BDH aristar) at 0°C for 60h. The digest was neutralised with 270ul frozen saturated LiOH. The LiF precipitate was removed by centrifugation and washed twice with 50ul water. The combined supernatants were added to 40mg $NaHCO_3$ and N-acetylated at 0°C by the addition of three aliquots of acetic anhydride (10ul) at 10min intervals. The products (fraction D) were desalted by passage through a tandem column of 1ml AG50X12(H^+) over 0.4ml AG3X4(OH^-) over 0.2ml QAE-Sephadex A25. After drying and evaporation with toluene, fraction D was taken for methylation linkage analysis.

Preparation of fraction E: nitrous acid deamination, NaB^3H_4 reduction and aq. HF dephosphorylation.

A sample of fraction A (20nmole) was dissolved in 15ul 0.1M sodium acetate buffer pH4.0 and deaminated by the addition of 15ul freshly prepared 0.5M sodium nitrite, 2h at room temperature. The deaminated products were reduced by the addition of 15ul 400mM boric acid, 12ul 1M NaOH and 15ul 30mM NaB^3H_4 (12Ci/mmole, NEN) in 100mM NaOH; final pH 11.2. After incubation at room temperature for 2h 20ul of 1M $NaBD_4$ was added and reduction was continued for 16h. Excess reductant was destroyed with acetic acid and the deaminated, reduced products were desalted by passage through 0.2ml AG50X12(H^+) and evaporation four times with 0.25ml methanol. The products were subjected to descending paper chromatography on Whatman 3MM paper in 1-butanol, ethanol, water (4:1:1) for 60h and the radiolabelled glycans were localised using a Raytest RITA linear analyser and recovered from within 2cm of the origin by elution with water.

The deaminated, reduced glycans were dephosphorylated with aq. HF as described above except that the combined supernatants after LiF precipitation were desalted by passage through 0.2ml AG50X12(H^+) over 0.3ml AG3X4(OH^-). The products were dried and subjected to high voltage paper electrophoresis (to remove residual radiochemical contaminants) at 80V/cm in pyridine, 60min, using acetic acid, water (3:1:387). The labelled neutral glycans (fraction E) were recovered from the origin by elution with water, passed through a column of 0.1ml Chelex 100(Na^+) over 0.2ml AG50X12(H^+) over 0.2ml AG3X4(OH^-) over 0.1 ml QAE-Sephadex A25 and filtered through a 0.2μ membrane. Aliquots of fraction E were taken for Dionex carbohydrate HPLC and for α-mannosidase digestions.

NMR spectroscopy

One- and two-dimensional ^1H-NMR spectra were obtained as previously described (McConville *et al* 1990b). ^{31}P NMR spectra were acquired at a spectrometer frequency of 202 MHz and a probe temperature of 300K, with a spectral width of 10,000 Hz. Phosphorus chemical shifts are referenced to external phosphoric acid (0.0 ppm at 300K), and proton shifts are referenced to internal acetone (2.225 ppm at 300K).

Gas chromatography-mass spectrometry and linkage analysis

Samples for methylation linkage analysis were methylated, as described above, dried and hydrolysed in 100ul 0.25M H_2SO_4, 93% acetic acid (80°C, 2.5h). Following neutralisation with 70ul 1M NaOH, the mixture was dried and residual acetic acid was removed by evaporation with toluene. The residue was dissolved in 50ul 1M NH_4OH and reduced with 200ul 0.25M $NaBD_4$ (3h at room temperature). Excess reductant was destroyed with acetic acid and the mixture was dried and evaporated twice with 0.25ml 5% acetic acid in methanol and twice with 0.25ml methanol, to remove boric acid. The products were acetylated with 0.25ml acetic anhydride (100°C, 2.5h). Acetic anhydride was removed under reduced pressure and the partially methylated alditol acetate (PMAA) derivatives were partitioned into dichloromethane and washed with water.

Samples of PMAA derivatives were analysed by GC-MS using a Hewlett-Packard 5890-MSD system equipped with an on-column injector using He as carrier gas at 1ml/min. The derivatives were separated using an SE-54 bonded phase (Econocap SE-54, 30m x 0.25mm, Alltech) and a Supelco SP2380 bonded phase column (30m x 0.25mm, Supelco) with the following temperature programme: 80°C (1min), 30°C/min to 140°C, 5°C/min to 250°C, hold for 20min. Mass spectra were recorded by linear scanning from m/z 40 to 350 using electron impact. Derivatives were identified by their characteristic retention times and mass spectra. Quantitation was performed by integration of the total ion current (see legend to Table 2).

Liquid chromatography and exoglycosidase digestions.

Carbohydrate HPLC was performed using a Dionex BioLC system equipped with a CarboPac PA-1 analytical column (25 x 0.4cm), a pulsed amperometric detector (PAD) and an anion micro-membrane suppressor linked to a Raytest Ramona radioactivity flow monitor. All data were collected and processed using the Raytest Ramona interface and software. Neutral glycans were resolved using a linear gradient from 12.5mM to 50mM sodium acetate in 150mM NaOH over 50min followed by a wash cycle of 250mM sodium acetate in 150 mM NaOH for 10min. The flow rate was maintained at 0.6ml/min and 1min fractions were collected. All samples were co-injected with 300ug of a dextran partial acid hydrolysate, as a set of glucose oligomer internal standards, which were detected by the PAD detector. The radiolabelled neutral glycan species were detected by the Ramona detector and their elution positions were expressed in "Dionex Units" (Du) by linear interpolation of their elution positions between adjacent glucose oligomer internal standards. These Du values have no significance except as absolute chromatographic properties (Ferguson, 1991).

Aspergillus phoenicis α-mannosidase (1mU/ml, Oxford Glycosystems) and jack bean α-mannosidase (25U/ml, Boehringer) digestions were performed on 2nmole aliquots of fraction E in 20ul sodium acetate buffer pH 5.0 for 16h at 37°C. Digestions were terminated by heating, 100°C for 5min, and the products were desalted by passage through a tandem column of 0.1ml AG50X12(H^+) over 0.1ml AG3X4(OH^-) eluted with water.

Results

NMR analysis of fraction A.

The analytical strategy is shown in Fig. 1.

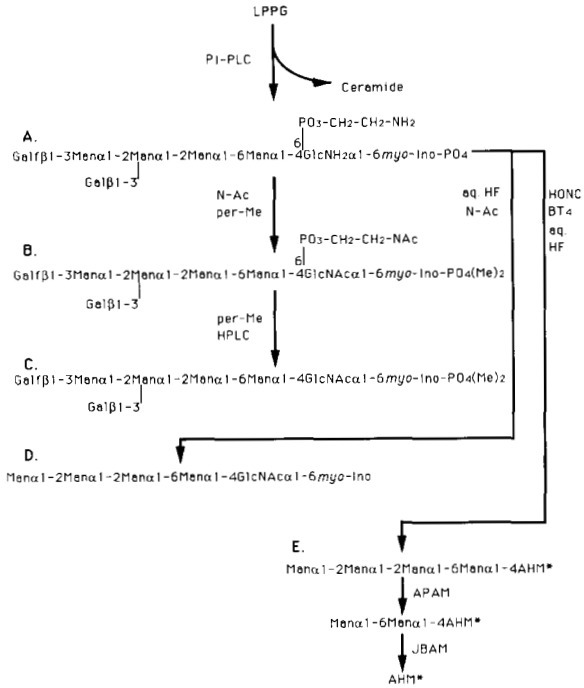


Figure 1. Summary of the chemical and enzymic modifications of LPPG.
The routes to fractions A to E are indicated using the major component (structure 1, see Fig. 6). Details of the modifications and analyses are given in the text.

Table 1: ^{1}H resonance assignments for reporter resonances of fraction A

Residue descriptor	Residue	Chemical shifts H1	H2	Coupling constant $J_{1,2}$(Hz)
N	αGlcNH$_2$	5.72	3.37	3.5
1	αMan	5.31	4.07	~2
2	αMan	5.14	4.05	~2
3	αMan	5.29	4.35	~2
4	αMan	5.13 5.20	4.25 4.16	~2 ~2
5	βGalf	5.17	4.16	~2

The LPPG molecule was cleaved with *B.thuringiensis* PI-PLC as previously described (Lederkremer *et al* 1990) to yield the glycophospholipid head groups, fraction A. This material was analysed by one dimensional ^{31}P-NMR (Figure 2a), and by two-dimensional ^{1}H-^{1}H COSY (not shown) and ^{1}H-^{1}H NOESY (Figure 2b). The ^{31}P spectrum shows a distinctive phosphonodiester resonance at 23.24ppm due to the AEP residue together with resonances at 17.58ppm (J=20Hz), and 4.56ppm (J=9.6Hz) which correspond to the phosphodiester of inositol-1,2-cyclic phosphate and a phosphomonoester, most likely of inositol-1-phosphate, respectively[4]. A connectivity network of the major species (arising from structure 1, Figure 6) could be traced in the ^{1}H-NOESY spectrum (Figure 2b) which, in combination with the methylation linkage composition data, (Table 2) suggests the sequence:

Galfβ1-3Manα1-2Manα1-2Manα1-6Manα1-4GlcNH2α1-6myo-Inositol

5 4 3 2 1 N I

where the residue descriptors are shown below the sequence. Residue types and anomericities were assigned according to characteristic chemical shifts and $J_{1,2}$ coupling constants, together with data derived from the1H-1H COSY spectrum of fraction A (not shown), Table 1. The tentative assignment of the Gal resonance as βGalf is based on comparison with published values for a variety of α and β galactofuranosides (Gerwig *et al* 1989,1991 and [5]) and is consistent with the same assignment made by Previato *et al* (1990) from ^{13}C-NMR. The downfield shifted GlcNH$_2$ H5 (4.27ppm) together with heteronuclear splitting of the GlcNH$_2$ H6 proton resonances at 4.13ppm (not shown) suggest that this is the site of AEP substitution. In addition, the large down-field shift of the M3 residue H2 resonance (4.35ppm, versus 4.11ppm for free α-Man), in comparison with M1 and M2, suggests that M3 may be substituted at the 3-position. A notable feature of the NOESY data (Figure 2b) is the presence of an additional inter-residue connectivity to Man-3 H2, derived from a resonance at 5.20 ppm (boxed in Figure 2b) and further characterised by other low intensity NOESY connectivities in Figure 2b. These are indicative of microheterogeneity in the chemical environment of the Man-4 residue, although the origins of this heterogeneity are not clear from the NMR data.

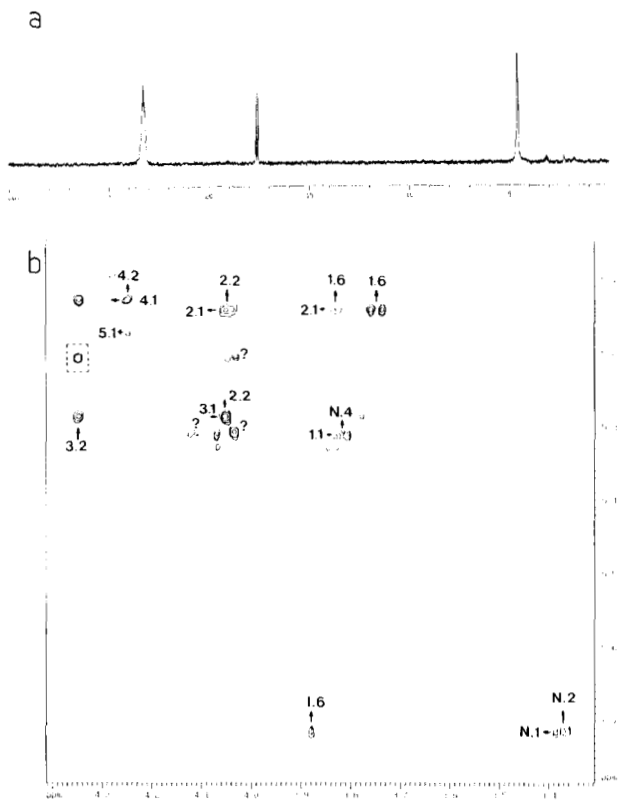


Figure 2. NMR analyses of LPPG fraction A.

a. ^{31}P NMR spectrum of fraction A.

b. Reporter region of ^{1}H-^{1}H NOESY NMR spectrum of fraction A. The abbreviations labelling each cross peak correspond to the residue descriptors shown in the text, followed by the proton assignment which is labelled according to its parent carbon. The origin of the boxed resonance is discussed in the text. The resonances marked ? derive from heterogeneity in the sample, and could not be assigned unambiguously.

FAB-MS analysis of fraction B.

Analysis of the N-acetylated, permethylated fraction B by FAB-MS produced the spectrum shown in Fig 3a. These data clearly show the existence of three glycan species. The molecular ion region of this spectrum contains clusters of molecular ions for three species. The major species is Hex$_6$HexNH$_2$,AEP,Inositol,PO$_4$: (M+H$^+$) at m/z 1991 and ions for the mono-undermethylated species at m/z 1977 (M+H$^+$) and 1999 (M+Na$^+$), and the di-undermethylated species at m/z 1963 (M+H$^+$). Two minor components were also observed with compositions Hex$_5$HexNH$_2$,AEP,Inositol,PO$_4$: (M+H$^+$) at m/z 1787 and undermethylated species at m/z 1773 and 1759, and Hex$_4$HexNH$_2$,AEP,Inositol,PO$_4$: (M+H$^+$) at m/z 1583 and an undermethylated component at m/z 1569. The spectrum contains two major series of fragment ions which give information on the sequence and branching patterns of the components. The first set of ions arises by single cleavage at the HexNAc residue to form A$^+$-type ions (Dell, 1987) at m/z 1239 for Hex$_4$,AEP,HexNAc$^+$,1443 for Hex$_5$,AEP,HexNAc$^+$ and 1647 for Hex$_6$,AEP,HexNAc$^+$. The second set of ions arises by single β-cleavage events (Dell, 1987) to form the ions at:

m/z 753 for [HO-(AEP)HexNAc,Inositol,PO$_4$]H$^+$,

957 for [HO-Hex,AEP,HexNAc,Inositol,PO$_4$]H$^+$,

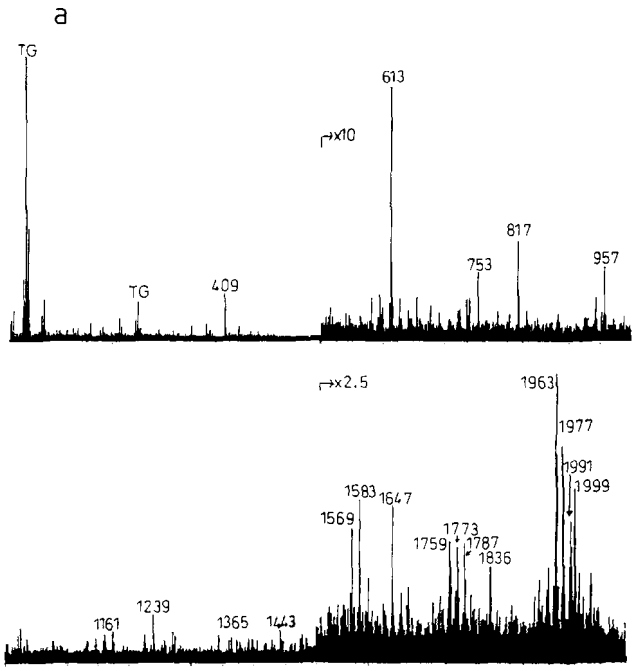
1161 for [HO-Hex$_2$,AEP,HexNAc,Inositol,PO$_4$]H$^+$,

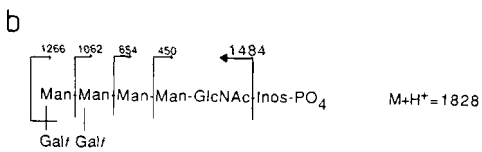
1365 for [HO-Hex$_3$,AEP,HexNAc,Inositol,PO$_4$]H$^+$,

1569 for [HO-Hex$_4$,AEP,HexNAc,Inositol,PO$_4$]H$^+$ and/or an undermethylated Hex$_4$-containing M+H$^+$ molecular ion, and m/z1773 for [HO-Hex$_5$,AEP,HexNAc,Inositol,PO$_4$]H$^+$ and/or an undermethylated Hex$_5$-containing M+H$^+$ molecular ion. Consideration of the A$^+$-type ions at m/z 1239 (Hex$_4$,AEP,HexNAc$^+$), 1443 (Hex$_5$,AEP,HexNAc$^+$), and 1647 (Hex$_6$,AEP,HexNAc$^+$), together with the β-cleavage ion at m/z 753 (HO-(AEP)HexNAc,Inositol,PO$_4$) allows asssignment of the AEP moiety to the HexNAc residue. However, the ambiguities introduced by the presence in the sample of three molecular species, each of which is undermethylated, made it impossible to assign further sequence or branching patterns from this spectrum.

[4] The generation of inositol-1,2-cyclic phosphate is characteristic of the action of bacterial PI-specific phospholipase C enzymes, and subsequent opening of the cyclic phosphate can lead to an additional mixture of inositol 1-phosphate and inositol 2-phosphate. In this case, only one inositol-monophosphate signal is seen in addition to the inositol 1,2-cyclic phosphate in the ^{31}P NMR spectrum. Analysis of a mixture of unsubstituted inositol 1-phosphate and inositol 2-phosphate indicated that the two forms give rise to well resolved signals at 5.28 and 6.06 ppm, respectively. The results indicate that, under the conditions used in this paper, *B. thuringiensis* PI-specific phospholipase C can generate both the 1,2-cyclic phosphate and the 1-phosphate from the LPPG substrate. This result is in agreement with the observations made for *B. cereus* PI-specific phospholipase C (Volwerk *et al.*, 1990) that showed that this enzyme has an intrinsic, regio-specific, cyclic phosphodiesterase activity.

[5] J. R. Thomas, J. E. Thomas-Oates, M. J. McConville, S. W. Homans, M. A. J. Ferguson, P. A. J. Gorin, K. Greis, and S. J. Turco, manuscript in preparation.





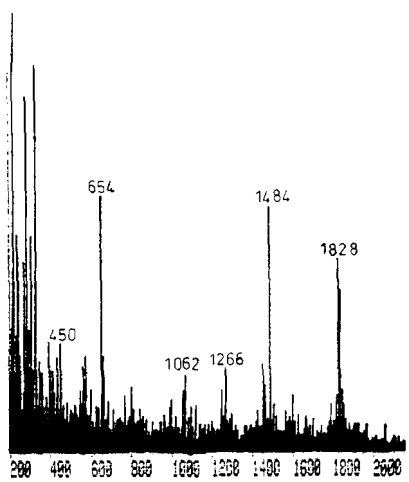


Figure 3. Fast atom bombardment mass spectrometric analyses of LPPG fractions B and C.

a. Positive ion FAB mass spectrum of fraction B.

b. Positive ion FAB mass spectrum of fraction C.

FAB-MS analysis of fraction C.

In order to resolve the ambiguities described above, fraction B was permethylated a second time to ensure full methylation prior to fractionation by reverse-phase HPLC. The main carbohydrate-containing peak, fraction C, was analysed by FAB-MS and gave the spectrum shown in Fig 3b. The re-methylation step caused loss of the N-acetylated AEP residue and an $M+H^+$ molecular ion corresponding to fully methylated $Hex_6HexNAc.Inositol.PO_4$ was observed at m/z 1828. An A^+-type ion arising by single cleavage was observed at m/z 1484 corresponding to $Hex_6HexNAc^+$, as well as a series of double cleavage ions arising from this ion by β-cleavage events. These double cleavage ions are observed at m/z 1266 ($HO\text{-}Hex_5HexNAc^+$), 1062 ($HO\text{-}Hex_4HexNAc^+$), 654 ($HO\text{-}Hex_2HexNAc^+$), and 450 ($HO\text{-}Hex_1HexNAc^+$). The absence of an ion at m/z 858, together with the presence of the other ions in this series, is crucial in assigning the branching pattern shown in Fig 3b.

Table 2: Methylation linkage compositions of fractions A, C and D

PMAA derivative	Origin	Corrected molar ratio* Fraction A	Fraction C	Fraction D
1,4-di-acetyl-2,3,5,6-tetra-methyl-galactitol	terminal Galf	1.52	2.00	†
1,2,5-tri-acetyl-3,4,6-tri-methyl-mannitol	2-subs.-Man	1.20	1.00	1.75
1,3,5-tri-acetyl-2,4,6-tri-methyl-mannitol	3-subs.-Man	1.05	1.00	trace
1,5,6-tri-acetyl-2,3,4-tri-methyl-mannitol	6-subs.-Man	1.00	1.00	1.00
1,2,3,5-tetra-acetyl-4,6-di-methyl-mannitol	2,3-disubs.Man	0.65	1.00	trace
1,4,5-tri-acetyl-3,6-di-methyl-2-methylacetamido-glucitol	4-subs.-GlcNAc	trace	1.00	1.10
6-acetyl-1,2,3,4,5-penta-methyl-myo-Inositol	6-subs-myo-Inositol	0.00	0.00*	0.85

*The total ion current peak areas were corrected using empirically derived molar response factors calculated from the anlyses of fraction C, a pure compound of known composition from FAB-MS analysis (Figure 3). The response factors were 0.75, 0.92, 0.92, 1.00, 0.98 and 0.62 for terminal Galf, 2-subs.-Man, 3-subs.-Man, 2,3-disubs.-Man and 4-subs.-GlcNAc respectively.

*Not detected in its phosphorylated state.
All values are the means of at least 2 analyses.

†Galf released by the aq. HF treatment was detected as a mutarotation mixture of terminal Galp and terminal Galf.

GC-MS methylation analyses and exoglycosidase sequencing.

The GC-MS methylation linkage analysis of the remainder of fraction C which was left after the FAB-MS analysis (Table 2) is entirely consistent with the NMR and FAB-MS assignments of the major structure. This material contains two terminal Galf residues and a 2,3-di-O-substituted-Man branch-point residue. In addition, the removal of the AEP moiety during the re-methylation step is witnessed by the appearance of the 4-O-substituted-GlcNAc residue in fraction C (c.f. fraction A, where the GlcNAc residue is cryptic in the methylation analysis due to its phosphonylated state). Evidence for the position of the $AEP\text{-}GlcNH_2$ linkage was obtained from the methylation analysis of deaminated and $NaBD_4$ reduced fraction A. The deaminated, reduced products were fractionated by Dionex HPLC and a complex series of peaks was observed (data not shown). GC-MS linkage analysis of several of the major peaks revealed the novel derivative shown in Figure 4. The mass spectrum indicates that this was derived from 4-O-substituted-AHM-6-AEP, which in turn was derived from 4-O-substituted-$GlcNH_2$-6-AEP prior to deamination and reduction.

The linkage analyses of fraction A (Table 2) confirmed the heterogeneity of the sample observed both by NMR (fraction A) and FAB-MS (fraction B). The absence of terminal Man residues shows that all the structures possess non-reducing Galf termini while the 2,3-di-O-subs-Man figure implies that the major species described above accounts for about 65 mole % of the structures. This figure is in reasonable agreement with the amount of terminal Galf observed in this fraction. The FAB-MS analysis of fraction B indicated that the remainder was composed of Hex_5 and Hex_4 containing structures. The linkage analysis of fraction A (Table 2), particularly the absence of terminal Man and the elevated level of 2-O-substituted-Man, further suggested that the Hex_5 structure is most likely:
Galβ1-3Manα1-2Manα1-2Manα1-6Manα1-4(AEP-6)$GlcNH_2$α1-6*myo*-Inositol-PO_4. However, this assignment relies on the assumption that the inter-mannose glycosidic linkages are the same as for the major species, the linkages of which were determined by NMR.

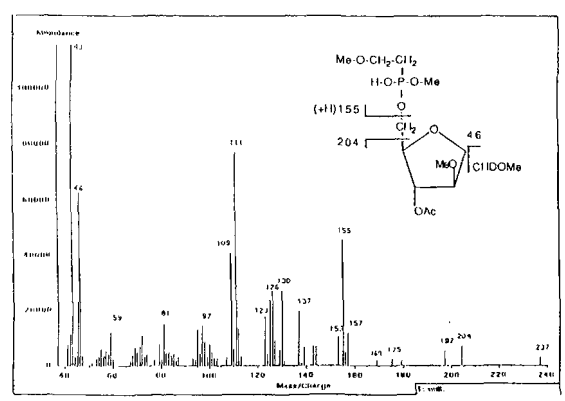


Figure 4. Electron impact mass spectrum of a novel derivative, 4-O-acetyl-1,3-di-O-methyl-1-deutero-2,5-anhydromannitol-6-(methyl) phosphonyl-ethylmethyl ether. This compound was derived from the methylation linkage analysis of deaminated and reduced fraction A. The ion at m/z 155 probably corresponds to $[CH_3OCH_2CH_2\text{-}P(OH)_2OCH_3]^+$.

To investigate the structure(s) of the underlying mannose backbone(s) directly a sample of fraction A was dephosphorylated with cold aq. HF, under conditions which cleave β-Galf glycosidic bonds but not α-Manp bonds (Ferguson *et al* 1988, Turco *et al* 1989, Mayor *et al* 1990), to yield fraction D after N-acetylation. This material was subjected to methylation linkage analysis (Table 2) which revealed the loss of terminal Galf, 3-O-substituted-Man and 2,3-di-O-substituted-Man and the appearance of terminal-Man and extra 2-O-substituted-Man, as predicted. In addition the 4-O-substituted-GlcNAc residue and the 6-O-substituted-*myo*-Inositol were observed, supporting the NMR NOESY assignment of the sequence Manα1-4GlcNH$_2$α1-6*myo*-Inositol. Fraction E, produced by deamination, NaB^{3}H$_4$ reduction and aq. HF dephosphorylation (with concomitant loss of the Galf residues), was resolved into two species by Dionex HPLC with elution positions of 2.5Du (15%) and 3.0Du (85%) respectively, Figure 5. These elution positions are identical to those of authentic Manα1-2Manα1-6Manα1-4AHM and Manα1-2Manα1-2Manα1-6Manα1-4AHM standards respectively (Ferguson 1991). Fraction E was digested with *Aspergillus phoenicus* Manα1-2Man specific α-mannosidase to yield a single compound which eluted at 2.2Du. This latter elution position is identical to that of authentic Manα1-6Manα1-4AHM, confirming that the original components terminated in one and two Manα1-2Man linked α-Man residues respectively. Finally the Fraction E material was digested with jack bean α-mannosidase to produce a compound eluting at 1.0Du, which corresponds to free AHM (Ferguson 1991). Taken together with the linkage analysis of fraction A (which shows only terminal Galf residues) and the FAB-MS data of fraction B (which shows the presence of Hex$_4$ containing glycans) the presence of a small amount (15%) of Manα1-2Manα1-6Manα1-4AHM suggests that the minor Hex$_4$ containing structure in fraction B is derived from:

Galfβ1-3Manα1-2Manα1-6Manα1-4(AEP-6)GlcNH2α1-6*myo*-Inositol-PO$_4$.

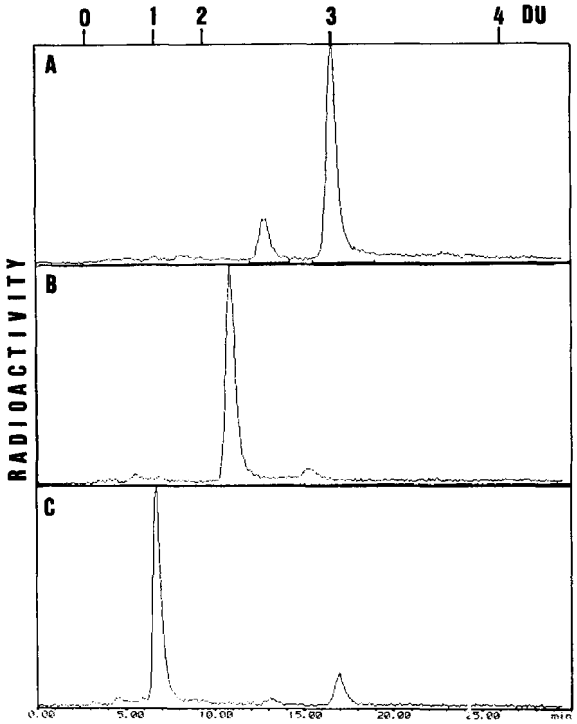


Figure 5. Dionex HPLC analysis and exoglycosidase sequencing of fraction E.
Panel A: Fraction E, untreated.
Panel B: Fraction E digested with Manα1-2Man specific *A.phoenicus* α-mannosidase.
Panel C: Fraction E digested with broad specificity jack bean α-mannosidase.
The numbers at the top indicate the elution positions of the glucose oligomer internal standards (Du values).