Novel O-Methylated Terminal Glucuronic Acid Characterizes the Polar Glycopeptidolipid of Mycobacterium habana Strain 5135*

(Received for publication, October 30, 1995, and in revised form, February 9, 1996)

Kay-Hooi Khoo, Delphi Chatterjeet, Anne Delling, Howard R. Morris, Patrick J. Brennann, and Philip Draper†

From the †Department of Microbiology, Colorado State University, Fort Collins, Colorado 80523 and the ‡Department of Biochemistry, Imperial College, London SW7 2AY and ¶National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom

Mycobacterium “habana” strain TMC 5135, which has been proposed as a vaccine against both leprosy and tuberculosis, is considered to be a strain of serotype I of the recognized species Mycobacterium simiae. We have now shown that each of these strains possesses characteristic polar glycopeptidolipids (GPL) which are sufficiently different to allow unequivocal strain identification. Thin layer chromatographic analysis demonstrated that M. habana synthesizes a family of apolar GPLs and three distinct polar GPLs (pGPL-I to -III) which exhibited migration patterns different from those of M. simiae serotype I (pGPL-Sim). Using a combination of chemical, mass spectrometric, and proton-NMR analyses, the GPLs from M. habana were determined to be based on the same generic structure as those from the M. avium complex, namely N-fatty acyl-o-Phe-(O-saccharide)-o-allo-Thr-o-ala-L-alaninyl-O-monosaccharide. The de-O-acetylated apolar GPLs contain a 3-O-Me-6-deoxy-Tal attached to the allo-Thr and either a 3-O-Me-Rha or a 3,4-di-O-Me-Rha attached to the alalinol. In the pGPLs, oligosaccharides were found to be attached to the allo-Thr. The oligosaccharidic aldilid reductively released from the least polar pGPL-I was fully characterized as 1-O-Fucp1-3-(6-O-Me)-o-Glcpi-1-3-(4-O-Me)-1-Rhap1-1-3-2-Rhap1-1-2-(3-O-Me)-6-deoxy-Tal. In pGPL-II and -III, the terminal Fuc residue is further 3-O-methylated and 4-O-substituted with an additional 2,4-di-O-Me-o-GlcA and 4-O-Me-o-GlcA, respectively. The corresponding oligosaccharide from pGPL-Sim was shown to be of identical molecular weight to pGPL-II but terminating with a 3,4-di-O-Me-GlcA. Enzyme-linked immunosorbent assay-based serological studies using anti-M. habana and anti-M. simiae sera against whole cells and purified pGPLs firmly established the polar GPLs as important antigens and indicated that the terminal epitopes 1-O-Fuc-, 2,4-di-O-Me-o-GlcA, and 4-O-Me-o-GlcA uniquely present in pGPL-I, -II, and -III, respectively, confer sufficient specificity for the identification of M. habana as a distinct serotype of M. simiae.

Leprosy and tuberculosis are grave diseases with large worldwide prevalences (Raviglione et al., 1995; Noordeen, 1994). If leprosy seems, for the present and at considerable cost, likely to be controlled by effective and available chemotherapy, tuberculosis incidence is rising and infections with novel multiple drug-resistant strains are effectively incurable (Bloom and Murray, 1992). Although BCG is a highly effective vaccine against both diseases in some populations, in others it is without detectable effect (Fine, 1995), and alternative vaccines are being sought. The need for novel tuberculosis vaccines is particularly urgent.

Mycobacterium habana strain TMC 5135 has been used successfully to vaccinate mice against tuberculosis (Gupta et al., 1979) and leprosy (Singh et al., 1985; Singh et al., 1989) and has been proposed as a human vaccine against both diseases (Singh et al., 1991). This strain was one of a group of 35 similar strains isolated in Havana from patients with lung disease (Valdivia-Alvarez et al., 1971). Although the symptoms resembled pulmonary tuberculosis, the M. habana strains differed markedly from Mycobacterium tuberculosis itself. It was originally proposed as a novel mycobacterial species, but subsequent taxonomic investigation using classical “biochemical” tests and seroagglutination indicated that M. habana strains did not differ sufficiently from serotype I of Mycobacterium simiae, a recognized mycobacterial species, to deserve the status of a separate species (Meissner and Schröder, 1975). However, these authors found that sera raised against M. habana, absorbed with M. simiae, retained some ability to agglutinate M. habana strains, indicating that these possess antigenic determinants absent from M. simiae. A similar conclusion was reached on the basis of reciprocal skin tests in guinea pigs and DNA hybridization measurements (Baess and Magnusson, 1982).

Most species of mycobacteria form clumps spontaneously during growth so that seroagglutination techniques cannot be used to identify them. However, a few species, notably members of the so-called Mycobacterium avium-Mycobacterium intracellulare complex but also some others, can be successfully typed by seroagglutination (Schaefer, 1967). In these species the dominant surface antigens are the distinctive polar glycopeptidolipids (GPLs) (Brennan, 1988). These are members of a larger family of GPLs, all related to “mycoside C” (Jollès et al., 1961), which also includes apolar types devoid of significant

*This work was supported in part by National Institute of Health NIAID Grant AI 30669,‡‡D.A., a Medical Research Council Programme Grant, Biotechnology and Biological Sciences Research Council Project Grants, and a Wellcome Trust Grant 030826 (to H. R. M. and A. D.), and a Heiser Research Fellowship in Leprosy and Tuberculosis (to K.-H. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked ‘advertisement’ in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡‡To whom correspondence should be addressed. Tel.: 44-181-959-3666 (ext. 2346); Fax: 44-181-913-8528; E-mail: p-draper@nimr.mrc.ac.uk.

1 The abbreviations used are: GPL, glycopeptidolipid; ABTS, 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid); COSY, correlated spectroscopy; ELLISA, enzyme-linked immunosorbent assay; FAB, fast atom bombardment; GC, gas chromatography; MS, mass spectrometry; dI He, deoxyhexose; Hex, hexose; HexA, hexuronic acid; Rha, rhamnose; Me, methyl; Fuc, fucose.
serological activity. The polar types are distinguished by the possession of complex oligosaccharides which are the major antigenic determinants (Yanaghara et al., 1985). For strains of mycobacteria producing such polar GPLs, definitive identification is simplified and may be effected by TLC, ELISA, or physicochemical analysis (Denner et al., 1992), as well as the original seroagglutination methods.

In this paper, we have investigated the presence and structures of GPLs in M. habana TMC 5135. In addition to a family of simple apolar GPLs, we found three polar GPLs carrying complex oligosaccharides with unique sequences. Similar structures were identified in the polar GPLs of M. simiae serotype I but were clearly distinguishable from each of those of M. habana by the degree and positions of O-methylation on the nonreducing terminal glucuronic acid residue. Serological studies provided further insights into the known cross-reactivities, which can now be structurally rationalized to some extent. Identification of these unique molecules in M. habana TMC 5135 provides a simple serological or chromatographic method of defining the strain, which will be essential for quality control if it is to be used to protect humans against tuberculosis or leprosy. It also underlines the biochemical versatility of mycobacteria, and the strong effect of apparently minor variations in chemical structure on serological activity/specificity.

EXPERIMENTAL PROCEDURES

Mycobacteria—M. habana strain TMC 5135 was obtained from Dr. N. B. Singh and was maintained on Löwenstein-Jensen slopes. M. simiae serotype I was obtained from Dr. P. A. Jenkins, Mycobacterium Reference Unit, Cardiff, as strain 26110 of his reference collection. The following related strains were also obtained from Dr. J. Jenkins: M. simiae serotype I 26111, M. simiae serotype I (Habana) 26112, and M. simiae serotype I (Habana) 26113.

Preparation and purification of GPLs—For lipid production, mycobacteria were grown on slopes of Middlebrook 7H9 medium with OADC supplement (Difco) in 50 ml screw-cap polystyrene jars (R.S. Jars) for 4–6 weeks. The cultures were sterilized by 500 krad of γ radiation from 60Co before harvesting. Cells were washed off the surface of the medium by vortexing with water, washed once with water, and freeze-dried. Dry mycobacterial mass was then extracted with ethanol/ether (1:1 by volume) at room temperature for 24 h, followed by chloroform/methanol (2:1 by volume) under similar conditions. Solvent extracts were dried in a rotary evaporator, then deacylated by the method of Brennan and Goren (1979). Apolar and polar GPL were isolated by preparative TLC on 20 × 20-cm silica Gel 60 plates (Merck), using chloroform/methanol (9:1 by volume) and chloroform/methanol/water (65:25:4 by volume, respectively). Residues were identified in side bands cut from the plates, using α-naphthol-sulfuric acid, and corresponding iodine vapor-stained bands on the main part of each chromatographic plate were extracted from scraped off silica gel with chloroform/methanol (2:1 by volume) or chloroform/methanol/water (65:25:4 by volume). They were repurified by TLC until homogeneous. Identical methods of growth and processing were used to obtain polar GPLs from M. simiae serotype I and, on a smaller scale, from the related strains. Polar GPL from M. avium serovar 19 was prepared as described before (Chatterjee et al., 1988).

Analytical TLC—Lipids were analyzed by TLC using silica Gel 60 TLC and high performance TLC plates (Merck), and chloroform/methanol (9:1 by volume) and chloroform/methanol/water (65:25:4 by volume) as developing solvents for separating apolar and polar GPL, respectively. Lipids were detected with dichromate/sulfuric acid (generating Libbye), α-naphthol sulfurous acid (reagent for lipids), α-naphthol sulfuric acid (reagent for glycolipids) and exposure to chlorine gas followed by spraying with starch–potassium iodide (reagent for amine and imine groups).

Glycolipids—Most of the glycolipid activities among the mycobacteria/mycobacterium were ascribed (and subsequently demonstrated) to be attached to allo-Thr was released from GPLs by reductive elimination (500 μl of 10 mgl mL sodium borohydride, 0.5 n sodium hydroxide, 500 μl of ethanol, stirred at 60 °C for 16 h) as di-O-glycosyl alkylid and designated as OSE. The sample was then deacetylated directly with Dowex (50W X-8H) beads, centrifuged, and the supernatant repeatedly co-evaporated to dryness with 10% acetic acid in methanol under nitrogen.

Compositional Analysis—Amino acids in the GPLs were identified using a Beckman 6300 amino acid analyzer after hydrolysis of the glycolipids in 4 N HCl at 110 °C for 18 h in sealed tubes, and also by GC and GC-MS analysis of their heptakisdeoxybutyril derivatives (MacKenzie and Tenashuk, 1975) to determine the absolute configurations as described previously (Besra et al., 1993). Briefly, the N(O)-heptakis- deoxybutyril esters were prepared by hydrolysis in 6 N HCl at 110 °C for 12 h, followed by treatment with 100 μl of 3 N HCl in (R)- and (S)-2-butanone at 120 °C for 120 min, and then 100 μl of anhydrous ethyl acetate and 40 μl of heptakisdeoxybutyril anhydride at 150 °C for 5 min. The α-configuration of the O-methylated glycosyl residues were determined by GC-MS analysis of the trimethylsilyl derivatives of (R)-(−) and (S)+(−)2,3-buty1 glycosides (Gerwig et al., 1978) after initial de-O-methylation of the permethylated OSE with 1 N boron tribromide in dichloromethane at 0 °C for 16 h 30 min followed by 30°C. GC-MS was carried out on a Hewlett-Packard 5890 Gas Chromatograph connected to a Hewlett-Packard 5790 Mass Selective Detector. Partially methylated and trideuteromethylated alditol acetates were analyzed on a SP 2380 capillary column (30 m, 0.25-mm inner diameter, 0.20-μm film thickness; Supelco) using a temperature gradient of 150 °C for 4 min and then 2 °C/min to 150°C. GC-MS was carried out on a Hewlett-Packard 5890 instrument fitted with a DB-23 capillary column (15 m, 0.25-mm inner diameter, 0.25-μm film thickness; J&W Scientific) using a temperature gradient of 100 °C for 20 min, and then 10 °C/min to 215°C. The heptakisdeoxybutyryl derivatives of amino acids were resolved with a temperature gradient of 30 °C/min to 190°C and then 4 °C/min to 250°C. The trimethylsilyl derivatives of butyl glycosides were injected at 80°C and chromatographed on a DB-1 capillary column (J&W Scientific) with a temperature gradient of 30 °C/min to 100°C and then 10°C/min to 265°C. The heptakisdeoxybutyryl derivatives of amino acids were resolved with a temperature gradient of 30 °C/min to 100°C and then 8°C/min to 250°C.

Analytical TLC—Lipid analysis of purified polar GPL was performed using a Hewlett-Packard 5890 instrument fitted with a DB-23 capillary column (15 m, 0.25-mm inner diameter, 0.20-μm film thickness; J&W Scientific) using a temperature gradient of 100°C for 20 min, and then 10°C/min to 250°C. The trimethylsilyl derivatives of butyl glycosides were injected at 80°C and chromatographed on a DB-1 capillary column (J&W Scientific) with a temperature gradient of 30°C/min to 100°C and then 10°C/min to 265°C. The heptakisdeoxybutyryl derivatives of amino acids were resolved with a temperature gradient of 30°C/min to 100°C and then 8°C/min to 250°C.

Analytical TLC—Lipid analysis of purified polar GPL was performed using a Hewlett-Packard 5890 instrument fitted with a DB-23 capillary column (15 m, 0.25-mm inner diameter, 0.20-μm film thickness; J&W Scientific) using a temperature gradient of 100°C for 20 min, and then 10°C/min to 250°C. The trimethylsilyl derivatives of butyl glycosides were injected at 80°C and chromatographed on a DB-1 capillary column (J&W Scientific) with a temperature gradient of 30°C/min to 100°C and then 10°C/min to 265°C. The heptakisdeoxybutyryl derivatives of amino acids were resolved with a temperature gradient of 30°C/min to 100°C and then 8°C/min to 250°C.

Analytical TLC—Lipid analysis of purified polar GPL was performed using a Hewlett-Packard 5890 instrument fitted with a DB-23 capillary column (15 m, 0.25-mm inner diameter, 0.20-μm film thickness; J&W Scientific) using a temperature gradient of 100°C for 20 min, and then 10°C/min to 250°C. The trimethylsilyl derivatives of butyl glycosides were injected at 80°C and chromatographed on a DB-1 capillary column (J&W Scientific) with a temperature gradient of 30°C/min to 100°C and then 10°C/min to 265°C. The heptakisdeoxybutyryl derivatives of amino acids were resolved with a temperature gradient of 30°C/min to 100°C and then 8°C/min to 250°C.

Analytical TLC—Lipid analysis of purified polar GPL was performed using a Hewlett-Packard 5890 instrument fitted with a DB-23 capillary column (15 m, 0.25-mm inner diameter, 0.20-μm film thickness; J&W Scientific) using a temperature gradient of 100°C for 20 min, and then 10°C/min to 250°C. The trimethylsilyl derivatives of butyl glycosides were injected at 80°C and chromatographed on a DB-1 capillary column (J&W Scientific) with a temperature gradient of 30°C/min to 100°C and then 10°C/min to 265°C. The heptakisdeoxybutyryl derivatives of amino acids were resolved with a temperature gradient of 30°C/min to 100°C and then 8°C/min to 250°C.
Polar Glycopeptidolipids of M. habana

12335

FIG. 1. High performance TLC of GPLs isolated from M. habana TMC 5335 and M. simiae serotype I. A, ethanol/ether (E, left panel) and chloroform/methanol (CH, right panel) extracts of M. habana, showing the apolar GPL; B, de-O-acetylated apolar GPLs of M. habana (aGPL-I to -V on lanes I-V, respectively) purified by preparative TLC; C, polar GPLs of M. habana (pGPL-I to -III, left panel) and M. simiae (pGPL-Sim, right panel). The high performance TLCs were developed using chloroform/methanol, 9:1 (A and B), or chloroform/methanol/water, 65:25:4 (C). GPLs were detected with α-naphthol sulfuric acid as purple bands.

Fig. 1. High performance TLC of GPLs isolated from M. habana TMC 5335 and M. simiae serotype I. A, ethanol/ether (E, left panel) and chloroform/methanol (CH, right panel) extracts of M. habana, showing the apolar GPL; B, de-O-acetylated apolar GPLs of M. habana (aGPL-I to -V on lanes I-V, respectively) purified by preparative TLC; C, polar GPLs of M. habana (pGPL-I to -III, left panel) and M. simiae (pGPL-Sim, right panel). The high performance TLCs were developed using chloroform/methanol, 9:1 (A and B), or chloroform/methanol/water, 65:25:4 (C). GPLs were detected with α-naphthol sulfuric acid as purple bands.

1/1000, the color developer was ABTS (Sigma) and absorbance of wells were measured at 405 nm in a Biotek EL312 plate reader. Tween 20 was not included in any of the washing solutions or diluents when glycolipids were being used as antigens. For cross-absorption experiments with pGPL-I to -III and pGPL-Sim, glycolipid suspensions at 40 μg/ml were mixed with the diluted sera (1/100 dilution) for 2 h at 37 °C before adding 0.1 ml of the serum to wells containing 3.5 μg of glycolipid. For experiments involving GPL from M. avium serotype 19, sera at 1200 dilution were absorbed with 20 and 40 μg/ml of glycolipid suspension for 66 h before being added to wells containing 1 μg of glycolipid.

RESULTS

The GPLs of M. habana—TLC analysis showed that M. habana possessed a family of apolar GPLs, a GPL of intermediate polarity (pGPL-I) and two very polar GPLs (pGPL-II and III) (Fig. 1). The distribution of polarities of these glycolipids was comparable to that reported for the two archetypal GPL families (polar and apolar) of the M. avium complex (Brennan, 1988). The apolar components and two of the polar lipid components, pGPL-I and pGPL-II, were largely extracted from the dry bacteria by ethanol/ether (1:1, v/v), whereas pGPL-III was found also in the chloroform/methanol (9:1, v/v) extract. GC and GC-MS analysis of the amino acid composition, in comparison to authentic standards, showed that both the apolar and polar GPLs contained d-Ala, d-Phe, d-Ala-dThr, and L-alaninol in approximately equimolar amounts, indicating that these GPLs might share the tetrapeptide core sequence found in all mycobacterial GPLs characterized so far except those from Mycobacterium xenopi (Aspinall et al., 1992, 1993; Khoo et al., 1995; Rivière et al., 1993).

Apolar GPLs—Mild deacylation reduced the mobilities of all bands of apolar GPL observable on chromatograms, indicating that these substances were O-acetylated in their native state. The deacylated apolar GPLs could be resolved into at least 5 major bands on TLC, which were isolated and designated aGPL-I to -V (Fig. 1B). A number of minor apolar components were present, but these were not further investigated.

The glycolipid compositions of the five aGPLs isolated fell into 2 groups: the more mobile species (aGPL-I to -III) contained 3-O-Me-6-deoxy-Tal and 3,4-di-O-Me-Rha whereas the less mobile ones (aGPL-IV and -V) contained 3-O-Me-6-deoxy-Tal and 3-O-Me-Rha. As described previously in the FAB-MS sequencing of the apolar GPLs from Mycobacterium smegmatis (Besra et al., 1994) and Mycobacterium butyricum (Kho et al., 1995), the perdeuteromethylated derivatives afforded diagnostic fragment ions which allowed definitive assignment of the tetrapeptide sequence and the identity of the two monosaccharide residues attached to allosThr and alaninol, respectively. From the m/z values of these (Table I) and the molecular ions (Table II), it could be concluded that: 1) each of the bands isolated by TLC comprised a mixture of molecular species of aGPL whose heterogeneity resided in their N-fatty acyl chains. The molecular ions and all fragment ions containing the fatty acyl chains comprised distinctive clusters of signals dominated by two or more major species. In contrast, fragment ions which did not contain the fatty acyl chains were characterized by single signals. 2) The monosaccharide attached to allosThr was invariably mono-O-methylated (3-O-Me-6-deoxy-Tal), whereas that attached to alaninol was mono-O-methylated (3-O-Me-Rha) in aGPL-IV and V but di-O-methylated in aGPL-I, -II, and -III (3,4-di-O-Me-Rha). 3) aGPL-I and aGPL-IV had similar fatty acyl heterogeneity but differed in the degree of O-methylation on the dHex attached to the alaninol, a similar structural relationship existed between aGPL-II and aGPL-V. The apolar GPL family is thus based on the same generic Structure I, a proportion of which carries an additional 4-O-methyl substituent on the 3-O-Me-Rha residue.

N-Fatty acyl-o-Phe-o-allosThr-o-Ala-o-alaninol-O-(3-O-Me-Rha)

3-O-Me-6-deoxy-Tal-0

STRUCTURE I

Polar GPLs—Chromatographic mobilities of the polar GPLs were unaffected by mild deacylation, implying that O-acetyl groups were not present. As with the apolar GPLs, their amino acid composition indicated a conventional N-acyl-o-Phe-o-allosThr-o-Ala-o-alaninol core. However, glycosyl composition analysis clearly showed that the polar GPLs possessed a much more complex saccharide content. GC and GC-MS analysis of the alditol acetates derived from pGPL-I showed the presence of 6 distinct sugar residues in approximately equimolar amounts, namely 3,4-di-O-Me-Rha, 3-O-Me-6-deoxy-Tal, 4-O-Me-Rha, Rha, Fuc, and a 6-O-Me-Hex. pGPL-II and -III had a common alditol acetate composition which differed from that of pGPL-I in having an additional 3-O-Me-dHex and a smaller amount of Fuc. Oligoglycosyl alditols, referred to as OSE-I, -II, and -III, were released from each of the pGPLs by reductive elimination and fully characterized as follows.

OSE-I from pGPL-I—The FAB-mass spectrum of the perdeuterocarboxyl derivative of OSE-I afforded an (M + Na)+ molecular ion at m/z 1326 together with prominent A-type oxonium ions at m/z 282, 548, 753, and 989, thus establishing the sequence as shown in Structure II.

This sequence was further corroborated by the perdeuteromethylated derivative which yielded an [M + Na]+ molecular ion at m/z 1018 and A-type ions at m/z 585 (dHex1O-Me)H3X(O-Me)H3X(+) ) and 765 (dHex2(O-Me)H3X(O-Me-Hex)(+)); upon complete hydrolysis, borodeuteride reduction, and acetylation, the resulting partially O-trideuteromethylated, O-methylated-alditol acetates gave four major peaks in
times as compared to authentic standards. The third and the fourth set of ions confirmed that the dHex residue attached to alaninol is either mono- or di-acylated ions afforded the peptide sequence as well as defining the mono-acyl-containing ions. *Thr and Alaninol* represent glycosylated allo-Thr and alaninol, respectively. ∆Thr represent the allo-Thr residue after eliminating its saccharide moiety.

**Table I**

Selected fragment ions observed in the FAB-mass spectra of perdeuteromethylated aGPLs

| Signal assignment | aGPLs | -I | -II | -III | -IV | -V |
|-------------------|-------|----|-----|------|-----|----|
| N-Acyl-Phe-Co⁺    | 705/719| 658/672| 610/624/691| 705/719| 658/672 |
| N-Acyl-Phe-Thr-CO⁺| 1017/1031| 970/984| 922/936/1003| 1017/1031| 970/984 |
| N-Acyl-Phe-Thr-Ala-CO⁺| 823/837| 776/780| 728/742/809| 823/837| 776/780 |
| N-Acyl-Phe-Thr-Ala-CO⁺| 893/907| 846/860| 799/812/879| 893/907| 846/860 |

* Elimination of the glycosyl residue was concomitant with cleavage of the peptide backbone at the N-C bond of Thr and acquisition of a sodium cation.

**Table II**

Molecular ions afforded by the aGPLs and deduced m/z values for the fatty acyl chains

| Ions | -I | -II | -III | -IV | -V |
|------|----|-----|------|-----|----|
| (M + Na)⁺ (native) | 1271/1285| 1227/1241| 1179/1193| 1257/1271| 1213/1227 |
| (M + Na)⁺ (d₃₅-permethylated) | 1410/1424| 1363/1377| 1315/1329| 1413/1427| 1366/1380 |
| Fatty acyl-Co (native)⁵ | 521/535| 477/491| 429/443| 521/535| 477/491 |
| Fatty acyl-Co (d₃₅-permethylated)⁶ | 541/555| 494/508| 446/460| 527 |

⁵ m/z values calculated from the molecular ions after subtracting the known mass of the peptide and saccharide moieties.

⁶ m/z values calculated from the ion clusters attributed to N-fatty acyl-Phe-Co⁺ (Table I). The presence of free OH group(s) can be determined from the total number of trideuteromethyl group(s) incorporated in the perdeuteromethyl derivatives, with each trideuteromethyl giving an increment of 17 units. The additional 3-unit increment observed for m/z 521/535 indicates the possible presence of a methyl group which can be replaced by a trideuteromethyl under the methylation conditions.

**Table III**

NMR Analysis of OSE-I and pGPL-I—The 500 MHz ¹H one-dimensional NMR analysis of OSE-I in [¹H]₅O at 25°C revealed (i) four distinct anomic protons at δ.5.42 (J₁₂ = 3.5 Hz), 5.25, 5.15 (J₁₂ = 10 Hz, indicative of a anomers in the manno configuration) and 4.85 (a doublet, J₂₅ = 7.7 Hz, indicative of a β-hexosyl unit); (ii) three proton singlets between 3.52–3.75 assignable to three 6-deoxyhexoses, suggesting the presence of 4 different 6-deoxyhexoses. When further analyzed by two-dimensional ¹H-¹H correlated spectroscopy (COSY), the four anomic protons of unit intensity and their corresponding H₁-H₂ cross-peaks were observed in the low-field region of the spectrum. The two cross-peaks at ω₁ = 5.09 ppm, ω₂ = 4.31 ppm (J₁₂ = 1.2 Hz), and ω₁ = 4.99 ppm, ω₂ = 4.09 ppm (J₁₂ = 1 Hz) confirmed the presence of two deoxyhexoses in their manno configuration, i.e. the 4-O-Me-Rha and Rha, whereas the most downfield cross-peak at ω₁ = 5.35 ppm, and ω₂ = 3.85 ppm (J₁₂ = 3.5 Hz) indicated the presence of a terminal α-Fuc residue. Finally, the cross-peaks at ω₁ = 4.71 ppm, ω₂ = 3.59 ppm (J₁₂ = 7.5 Hz) established that the 6-O-Me-Glc was in the β-anomeric configuration (Table III). Additional information on the anomeric configuration of the 3-O-Me-6-deoxy-Tal and 3,4-di-O-Me-Rha was obtained from NMR analysis of the parent pGPL-I using the ¹H detected two-dimensional [¹H,¹³C] heteronuclear multiple quantum correlation experiment (Fig. 2A) and measured the J⁺ values by inverse-detection two-dimensional-NMR. Thus the proton resonating at δ.505 was assigned to the C₁α at δ.541 (J₁₂ = 3.5 Hz, from the ¹H NMR and J⁺ values). The 1H-1H connectivity of H-2, H-3, H-4, H-5, and H-6,6' identified the glycosyl residue to be an α-Fuc residue. The anomeric protons resonating at δ.541, 4.93, 4.68 have C1s resonating at δ.510, 102.2, and 96.69, respectively.

The broad singlet at δ.493, which appeared to be a single proton, could be further resolved.
The chemical shifts are reported relative to chloroform (δ 7.0 ppm). The J₁₂ values for the Rha are approximate since the signals are broad and approaching the line widths.

| Glycosyl residue | H-1 ppm | H-2 ppm | J₁₂ Hz |
|------------------|---------|---------|--------|
| OSE-I            |         |         |        |
| Fuc (α₁-2)       | 5.05    | 3.65    | 3.5    |
| Rha (α₁-2)       | 4.99    | 4.06    | ~1     |
| 2×Rha (α₁-2)     | 4.93    | 3.90, 3.89 | ~1 |
| 6→Tal (α₁-2)     | 4.68    | 3.93    | ~2     |
| 6-O-Me-Glc (β₁-2) | 4.51    | 3.41    | 7.5    |
| OSE-II           |         |         |        |
| Fuc (α₁-2)       | 5.35    | 3.85    | 3.5    |
| Rha (α₁-2)       | 5.09    | 4.31    | ~1.4   |
| Rha (α₁-2)       | 4.99    | 4.09    | ~1.0   |
| 6-O-Me-Glc (β₁-2) | 4.71    | 3.59    | 7.5    |
| OSE-III          |         |         |        |
| Fuc (α₁-2)       | 5.48    | 4.02    | 3.5    |
| Rha (α₁-2)       | 5.18    | 4.40    | ~1.2   |
| Rha (α₁-2)       | 5.08    | 4.15    | ~1.2   |
| 3,4-Di-O-Me-GlcA (β₁-2) | 4.82   | 3.65 | 7.3 |
| 6-O-Me-Glc (β₁-2) | 4.65    | 3.65    | 7.3    |

Further information was obtained from FAB-MS analysis of the perdeuteromethylated derivative which afforded a pair of molecular ions at m/z 1214 (M + H)⁺ and 1236 (M + Na)⁺ (Fig. 3). The presence of a less abundant pair of molecular ions at m/z 1217/1239 is consistent with the presence of minor components carrying non-O-methylated Fuc. Significantly, the prominent A-type ion at m/z 239 confirmed the presence of a nonterminal component di-O-Me-HexA, whereas the pair at m/z 803/806 corresponded to (O-Me)HexA-2-(O-Me)Hex-O-Me-Hex⁻. A sodiated β-cleavage ion corresponding to (H₂O)(O-Me)Hex-(O-Me)Hex-dHex-(O-Me)Hex-titol was observed at m/z 799. To determine the identity of the (O-Me)HexA, the perdeuteromethylated sample was reduced with lithium aluminium hydride and subjected to GC-MS linkage analysis. In addition to the peaks already identified, a new peak was observed, eluting at a retention time similar to a 6-linked Glc standard. This was identified as a 6-linked 2,4-di-O-Me-Glc based on the ions observed (m/z 105, 118, 129, 165, 176, 189, and 236). The terminal residue of OSE-II was thus established as a 2,4-di-O-Me-GlcA involved in a predominantly 3-O-methylated Fuc. Taking all the data into account, it could be concluded that OSE-II shared the same core sequence with OSE-I but differed at the nonreducing terminus, where the Fuc in OSE-I was extended with a 2,4-di-O-Me-GlcA at position 4, and was largely O-methylated at position 3.

Linkage analysis on lithium aluminium hydride-reduced OSE-I revealed the same total in current profile as that from OSE-II. However, the electron impact-mass spectrum of the 6-linked Glc peak afforded ions at m/z 108, 121, 129, 168, 176, 189, and 236 consistent with a 1,5,6-O-Ac-4-O-Me-2,3-di-O-α-L-Rha-hexitol. Thus, OSE-III differed from OSE-II in having a 4-O-methylated glucuronic acid instead of a di-O-methylated residue. The 4-linked Fuc was detected in the FAB-MS analysis of the perdeuteracyclic and perdeuteromethylated derivatives confirmed this conclusion. The earlier yielded molecular ions 31 mass units higher than those of OSE-II (both before and after esterification) while the latter gave molecular ions 3 mass units higher, fully consistent with OSE-I having one O-Me group fewer than OSE-II. In addition, the esterified perdeuteracyclic OSE-I gave a prominent A-type ion at m/z 295, corresponding to a methyl esterified, mono-O-methylated HexA (cf. ion at m/z 264 for OSE-II). The difference in O-methylation between OSE-II and -III was therefore localized to the terminal HexA, corroborating the linkage analysis data.

Two-dimensional 1H-1H COSY NMR analysis of OSE-II revealed 4 distinct anomeric protons with corresponding cross-peaks for H-2 in the low field region (Fig. 2B, Table III) which were similar to those of OSE-I. An additional anomeric proton (somewhat obscure due to the residual water peak) with an intense cross-peak for H-2 (ω₂ = 4.82 ppm, ω₁ = 3.65 ppm, J₁₂ = 7.3 Hz) was detected in the spectrum, indicative of an additional β-hexosyl residue not present in the spectra of pGPL-I and OSE-I. The cross-peak connectivity of this residue lead to the assignment of H-3, H-4, H-5, but no other stereocentres could be determined. It was concluded that this additional glycosyl residue was not a 6-deoxy-Hex nor a Hex but was almost certainly attributable to di-O-Me-GlcA, occurring as a β-anomer.
but was only mono-O-methylated at position 4 of the GlcA, consistent with it being the most polar of all the GPLs characterized. After lithium aluminium hydride reduction followed by de-O-methylation, only D-Glc, L-Rha, and L-Fuc were detected in subsequent GC-MS analysis of the trimethylsilylated R-( )-2-butyl glycoside derivatives of OSE-II and -III, indicating a D-configuration for the additional terminal GlcA.

FAB-MS Analysis of pGPLs—FAB-MS analysis of native pGPL-I afforded a cluster of \( [M + Na]^+ \) molecular ions which could be categorized into three major groups: (i) \( m/z \) 1821/1835; (ii) 1853/1867; and (iii) 1899/1913, reflecting the heterogeneity associated with the N-fatty acyl chains. As observed previously (Besra et al., 1993; Khoo et al., 1995), the pGPLs readily yielded protonated ions corresponding to the lipopeptide cores, especially when treated with mild methanolic-HCl. A comparison of the masses of these ions with the molecular ions for intact GPL enabled assignments of the compositions of the saccharide appendages. These compositions were consistent with those deduced from the data on OSE-I. More definitive sequencing of the intact pGPL-I was afforded by the key frag-

**Fig. 2.** A, \(^1\)H detected \([\text{H}, \text{C}]\) heteronuclear multiple-quantum correlation spectrum of pGPL-I; cross-peak labels correspond with those on the structure shown. The residues 3 and 4 are inter alia. B, two-dimensional, 500 MHz 1H-1H COSY spectrum of OSE-II.
m/z 1649/1652 and 1652/1655 in the FAB spectra of perdeuteromethylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the latter two terminally methylated pGPL-II and -III, respectively, consistent with the presence of an additional HexA residue in the later...
Thus, there was an indication that there were antigenic differences between the strains and that the whole cells of these two strains might readily be distinguished on the basis of their differential reactivities with the two antisera.

Whole ethanol/ether extracts containing GPLs were also well recognized by the antisera (results not shown), but the cross-reactivity was not fully reciprocal. It was observed that anti-M. simiae serum recognized M. habana glycolipid extracts quite well but the opposite reaction, i.e. anti-M. habana serum against M. simiae glycolipids, was weak and only significant at high serum concentrations. Using purified pGPLs, it could be further shown that, in addition to homologous recognition, anti-M. simiae serum recognized a mixture of pGPL-II and pGPL-III but not pGPL-I from M. habana, whereas anti-M. habana serum did not recognize pGPL-Sim from M. simiae (Fig. 5).

Absorption of antisera with purified GPLs before adding them to the ELISA plates, using a similar amount of GPLs for absorption as was coated on to the ELISA well, gave results as predicted from the above observations but with one notable anomaly (Fig. 6). Absorption of anti-M. simiae serum with pGPL-Sim reduced both homologous reactivity and reaction against pGPL-II+III of M. habana. However, absorption of this serum with pGPL-II+III reduced only the reaction against pGPL-II+III. These effects were statistically significant (p < 0.01 in a 2-way t test).

Taken together, the serological studies suggested that pGPLs were important antigenic determinants in these strains.
The cross-reactivity noted between anti-M. simiae serum and M. habana GPL extracts might partly be explained as recognition of the terminal O-methylated uronic acids on pGPL-II and -III. In contrast, anti-M. habana serum only cross-reacted weakly with M. simiae GPL extracts, consistent with it not recognizing pGPL-Sim. Furthermore, the anomaly noted with the cross-absorption experiments may be rationalized based on the subtle differences between the position of O-methylation on the terminal uronic acids. It may be hypothesized that anti-M. simiae serum contained antibodies which would recognize epitopes containing 3- and/or 4-O-Me-GlcA. Absorption with pGPL-II and -III would probably deplete antibodies against 4-O-Me-GlcA but leaving those which can still recognize the 3-O-Me-GlcA and other epitopes in M. simiae GPLs. In contrast, absorption with pGPL-Sim would deplete all antibodies against the 4-O-Me epitope which is the common determinant of pGPL-I and II.

Further evidence in support of this observation was obtained by extending the serological studies to polar GPL from M. avium serotype 19 (pGPL-I and -II) which shares a very similar structure with pGPL-II and -III, particularly in containing the terminal 3,4-di-O-Me-D-GlcA (Chatterjee et al., 1987). As expected, pGPL-I and II was recognized by both anti-M. habana and anti-M. simiae sera (results not shown). In the case of anti-M. simiae serum, the reaction was similar in magnitude to the reaction with the homologous pGPL-Sim and greatly reduced when the anti-sera was pre-absorbed with pGPL-I and -II. On the other hand, reactivity of pGPL-19 with anti-M. habana was readily detectable but only at about 25% of the magnitude of reaction against anti-M. simiae serum, and declined much more rapidly with increasing serum dilution than did the reaction with homologous pGPL-I and II.

**DISCUSSION**

The polar glycopeptidolipids of the M. avium complex were the first surface antigens used to identify mycobacterial strains (Schaefer, 1967; reviewed by Brennan, 1988). Species or strain-specific glycolipids structurally based on the same prototype mycoside C have since been identified in other mycobacterial species such as Mycobacterium peregrinum, Mycobacterium senegalense, Mycobacterium poenonicum, and M. cybotacticum (López-Marín et al., 1991, 1992, 1993; Besra et al., 1994; Kho et al., 1995) although in these species the GPLs are mainly of the apolar type without the haptenic oligosaccharide appendages which characterize the M. avium serotype.

The GPLs isolated from M. habana strain TMC 5135 and characterized in this study closely resemble those of the M. avium complex in, 1) the core structure of fatty acyl-D-Phe-(O-saccharide)-O-allyl-Or-Ala-α-alanyl-O-mono-saccharide; 2) the long fatty acyl chain length averaging at C-30; 3) the presence of apolar and polar subunits, both with a single 3-O-methyl or a 3,4-di-O-methyl-Rha residue at the alaninol end; and 4) complex oligosaccharide appendage attached to allyl-Thr via a 3-O-Me-6-deoxy-Tal in the case of the polar GPLs. The 4-O-Me- or 2,4-di-O-Me-GlcA at the nonreducing termini of the more polar pGPL-II and -III represent novel terminal epitopes not previously identified on mycobacterial surface glycolipids (Aspinall et al., 1995) although in these species the terminal 3,4-di-O-Me-GlcA on the GPL of M. avium serovar 19 (Chatterjee et al., 1988). Interestingly, the polar GPLs of M. simiae serotype I strain 26110 which is very similar to that of pGPL-II of M. habana was also found to terminate with a 3,4-di-O-Me-GlcA.

As is well known for M. avium (Brennan, 1988), the terminal moieties of the oligosaccharides of the polar GPLs are the dominant epitopes recognized by antisera against the whole bacterial cell. Our data show that the anti-M. habana serum includes distinct antibodies reacting specifically with pGPL-I and II and III and, among those recognizing the latter, some also cross-reacted with the polar GPLs from M. simiae serotype I strain 26110 and M. avium serovar 19. Thus the differentially O-methylated terminal GlcA residue, or the absence of it, in the pGPLs of M. habana may be sufficiently specific to confer the structural basis upon which it can be distinguished from the closely related M. simiae serotype I and M. avium complex through simple and sensitive serological and biochemical means of identification.

It is not known whether all the other strains identified as M. habana share its distinctive pattern of polar GPLs, but at least two others from Dr. J. Jenkins’ reference collection (referred to as habana-like M. simiae strain 26112 and 26113), were found by us to do so. Since these strains have a different history of culture from TMC 5135 over a period of at least 20 years, it may be concluded that the distinctive pattern is quite stable and that M. habana is distinct from M. simiae serotype I. It seems improbable that the GPLs characterized in the present work are in any way connected with the ability of the strains that produce them to protect against mycobacterial diseases. They are nonetheless important in this context because they apparently make possible a simple, rapid, and unequivocal identification of the proposed vaccine strain, essential for quality control in the eventual production and use of a vaccine.

**Acknowledgments**—We gratefully acknowledge P. F. Thurman for growing the mycobacteria; S. Lathwell for amino acid analysis; C. Rich for technical assistance in the two-dimensional NMR experiments; Dr. N. B. Singh and Dr. M. A. J. Jenkins for providing the mycobacterial strains.

**REFERENCES**

Anz, W., Mélissen, G., and Röder, U. (1969) Zbl. Bakt. I Orig. 221, 530–550.

Aspinall, G. O., Chatterjee, D., and Brennan, P. J. (1995) Adv. Carbohydr. Chem. Biochem. 51, 169–242.

Bax, A. and Summers, M. F. (1998) J. Am. Chem. Soc. 1018, 2093–2094.

Baes, I., and Magnusson, M. (1982) Acta Pathol. Microbiol. Scand. 90, 101–107.

Besra, G. S., McNeill, M. R., Rivero S., Kho, K.-H., Morris, H. R., Dell, A., and Brennan, P. J. (1993) Biochemistry 32, 347–355.

Besra, G. S., Gurcha, S. S., Kho, K.-H., Morris, H. R., Dell, A., Hamid, M. E., Mimmink, D. E., Goodfellow, M., and Brennan, P. J. (1994) Zbl. Mikrobiol. Hyg. 281, 413–418.

Bloom, B. R., and Murray, C. J. L. (1992) Science 257, 1055–1064.

Brennan, P. J. (1988) In Microbial Lipids (Rattlede, C., and Wilkinson, C., eds) Vol. 1, pp. 203–298, Academic Press, London.

Brennan, P. J., and Goren, M. B. (1979) J. Bacteriol. 135, 4205–4211.

Chatterjee, D., Aspinall, G. O., and Brennan, P. J. (1987) J. Biol. Chem. 262, 3528–3533.

Chatterjee, D., Bucz, C., Aspinall, G. O., and Brennan, P. J. (1988) J. Biol. Chem. 263, 4092–4097.

Dell, A., Reason, A. J., Kho, K.-H., Panico, M., McDowell, R. A., and Morris, H. R. (1994) Methods Enzymol. 230, 108–132.

Denner, J. C., Tsang, A. Y., Chatterjee, D., and Brennan, P. J. (1992) J. Clin.
References

Fine, P. E. M. (1995) Clin. Infect. Dis. 20, 11-14
Gerwig, G. J., Kamerling, J. P., and Vliegenthart, J. F. G. (1978) Carbohydr. Res. 62, 349-357
Gupta, H. P., Singh, N. B., Mathur, I. S., and Gupta, S. K. (1979) Indian J. Exp. Biol. 17, 1190-1193
Jolles, P., Bigler, F., Gendre, T., and Lederer, E. (1961) Bull. Soc. Chim. Biol. 43, 177-192
Kho, K.-H., Morris, H. R., McDowell, R. A., Dell, A., Maccarana, M., and Lindahl, U. (1993) Carbohydr. Res. 244, 205-223
Kho, K.-H., Suzuki, R., Morris, H. R., Dell, A., Brennan, P. J., and Besra, G. (1995) Carbohydr. Res. 276, 449-455
López-Marín, L. M., Laneéelle, M. A., Prome, D., Daffe, M., Laneéelle, G., and Prome, J. C. (1991) Biochemistry 30, 10536-10542
López-Marín, L. M., Lecture, M.-A., Prome, D., Laneéelle, G., and Daffe, M. (1992) Biochemistry 31, 11106-11111
López-Marín, L. M., Laneéelle, M.-A., Prome, D., and Daffe, M. (1993) Eur. J. Biochem. 215, 859-866

MacKenzie, S. L., and Tenaschuk, D. (1975) J. Chromatogr. 111, 413-415
Meissner, G., and Schröder, K.-H. (1975) Am. Rev. Resp. Dis. 111, 196-201
Noordeen, S. K. (1994) Int. J. Lepr. 62, 278-283
Ravigljone, M. C., Snider, D. E., and Kochi, A. (1995) J. Am. Med. Assoc. 273, 220-226
Rivière, M., Augé, S., Vercauteren, J., Wisingerova E., and Puza, G. (1993) Eur. J. Biochem. 214, 395-403
Schafer, W. B. (1967) Am. Rev. Resp. Dis. 96, 115-118
Singh, N. B., Srivastava, Gupta, H. P., Sreevatsa, and Desikan, K. V. (1985) Indian J. Lepr. 57, 278-281
Singh, N. B., Lowe, A. C. R. E., Rees, R. J. W., and Colston, M. J. (1989) Infect. Immn. 57, 653-655
Singh, N. B., Srivastava, A., Gupta, H. P., Kumar, A., and Srivastava, S. (1991) Int. J. Lepr. 59, 317-320
Valdivia-Alvarez, J., Suárez-Méndez, R., and Echemendía-Font, M. (1971) Bol. Hig. Epid. 9, 65-73
Yanaghara, D. L., Barr, V. L., Knisley, C. V., Tsang, A. Y., McClathy, J. K., and Brennan, P. J. (1985) J. Clin. Microbiol. 21, 569-574