Carbohydrate Research, 251 (1994) 99–114
Elsevier Science B.V., Amsterdam

New pyruvylated, glycosylated acyltrehaloses from *Mycobacterium smegmatis* strains, and their implications for phage resistance in mycobacteria *

Gurdyal S. Besra [a], Kay-Hooi Khoo [b], John T. Belisle [a], Michael R. McNeil [a], Howard R. Morris [b], Anne Dell [b,†] and Patrick J. Brennan [a,†]

[a] *Department of Microbiology, Colorado State University, Fort Collins, Colorado 80523 (USA)*
[b] *Department of Biochemistry, Imperial College of Science, Technology, and Medicine, London (United Kingdom)*

(Received March 23rd, 1993; accepted July 16th, 1993)

ABSTRACT

Phage resistance and apparent lysogenization of *Mycobacterium smegmatis* due to infection with mycobacteriophage D29 results in the emergence of new variations of the pyruvylated, acylated trehaloses described by Saadat and Ballou, *J. Biol. Chem.* 258 (1983) 1813–1818. Thin-layer chromatography of the glycolipids from two strains of phage-resistant *M. smegmatis* ($mc^2$22 and $mc^2$11) and comparison with those from phage-sensitive strains revealed a new, more mobile glycolipid in each case. The structures of these acyltrehalose-containing lipooligosaccharides were elucidated by a combination of gas–liquid chromatography–mass spectrometry, methylation analysis, ^{1}H and ^{13}C NMR spectroscopy, and fast-atom bombardment mass spectrometry. The glycolipid from *M. smegmatis* $mc^2$22 is β-D-Glc*p*-(1 → 3)-4,6-*O*-(1-methoxycarbonylethylidene)-β-D-Glc*p*-(1 → 4)-β-D-Glc*p*-(1 → 6)-2-*O*-acyl-α-D-Glc*p*-(1 ↔ 1)-3,4-di-*O*-acyl-α-D-Glc*p* and that from *M. smegmatis* $mc^2$11 is 4,6-*O*-(1-methoxycarbonylethylidene)-3-*O*-Me-β-D-Glc*p*-(1 → 3)-4,6-*O*-(1-methoxycarbonylethylidene)-β-D-Glc*p*-(1 → 4)-β-D-Glc*p*-(1 → 6)-2-*O*-acyl-α-D-Glc*p*-(1 ↔ 1)-3,4-di-*O*-acyl-α-D-Glc*p*. These differ from the original pyruvylated glycolipids of Saadat and Ballou in the extent of their *O*-acylation and *O*-methylation. The findings are the first example of the definition of a chemical basis for phage resistance and presumed lysogeny in mycobacteria, and show parallels to related changes in gram-negative enteric bacteria.

INTRODUCTION

Members of the *Mycobacterium* genus have the ability to express their species and serotypic identity in variable glycosylated surface antigens, which in diversity equal the *O*-antigenic polysaccharides of the *Enterobacteriaceae*. Antigenic diversity occurs within the terminal nonreducing sugar arrangements on one of three

* Dedicated to Professor Clinton E. Ballou in honor of his outstanding contributions to our common discipline.

† Corresponding authors.

0008-6215/94/$07.00
SSDI 0008-6215(93)E0328-X

classes of glycolipids, namely the phenolic glycolipids, the glycopeptidolipids, and the trehalose-containing lipooligosaccharides[1,2]. Ballou and colleagues first described the presence of pyruvylated forms of glycosylated acyltrehaloses in mycobacteria[3,4], prior to full recognition of the extent of glycosylation, ubiquity, antigenicity, and species-specificity of this class of surface glycolipids, which we subsequently called lipooligosaccharides[2]. These glycolipids occur as surface appendages[5,6], and beneath them lies the cell wall skeleton, the mycoloyl–arabinogalactan–peptidoglycan complex, the compositional structure of which we[7] among others[8] have addressed. The research disciplines concerned with topics such as the full structural definition of the mycobacterial cell wall, its biogenesis, the identification of the site of action of antituberculosis drugs, such as isoniazid and ethambutol, and the role of cell-wall constituents in disease progression, would benefit from the availability of well defined mutants devoid of cell-wall components. As a first step in the acquisition of such mutants, we have analyzed the biochemical basis of phage resistance in strains of *Mycobacterium smegmatis*. In the process we have identified modified versions of the pyruvylated, glycosylated acyltrehaloses first described by Ballou and colleagues and have arrived at new insights into the molecular basis of phage resistance in mycobacteria that parallel those already described for the *Enterobacteriaceae*.

EXPERIMENTAL

Phage-resistant strains and growth conditions.—The origins of strains of *M. smegmatis* of the mc^2 designation have been described[9,10]. In the present instance, *M. smegmatis* 607 (ATCC, American Type Culture Collection) was streaked and picked for pure colonial forms as described[9]. Some of these (strains $mc^2$11 and $mc^2$22) were found to be spontaneously resistant to lysis by mycobacteriophage D29, whereas others were sensitive. The origins of phage D29 and procedures for phage-sensitivity testing have been described by Jones and White[11] and Jones and Beam[12]. Colony picking, phage-sensitivity testing, and the generation of strains were all the work of Jacobs and colleagues and will be described separately. Strains were grown[13] on 7H11 broth for 7–10 days at 37°C. Both cells and medium were evaporated to dryness at 50°C and the resulting solids extracted with 2:1 $CHCl_3$–MeOH. Dried extracts were dissolved in the biphasic mixture 4:2:1 $CHCl_3$–MeOH–water; the lower, organic phase was concentrated to dryness, and the contents were dissolved in 250 mL of hot acetone and allowed to precipitate at 4°C overnight. The precipitate was removed by centrifugation and the supernatant evaporated to yield the acetone-soluble lipids, the source of the glycolipids of interest.

Purification of glycolipids.—The acetone-soluble lipid preparations were applied to a column (2.5 × 30 cm) of Florisil (100–200 mesh, Fisher Scientific Co., Pittsburgh, PA) which was irrigated with 200 mL of $CHCl_3$ followed by 2, 5, 10, 20, and 50% MeOH in $CHCl_3$. Eluates (10 mL) were collected and subjected to TLC

in 30:8:1 $CHCl_3$–MeOH–water, and the plates were sprayed with 10% H_2SO_4 in EtOH then heated at 110°C for 5 min. The glycolipid-containing fractions (black response on spraying) were pooled and further purified by preparative TLC on silica gel (Merck 5735, silica gel $60F_{254}$, Darmstadt, Germany) with 30:8:1 $CHCl_3$–MeOH–water to afford the highly purified pyruvylated glycolipids PyrGL-I (R_f 0.78, 10 mg) from $mc^2$11 and PyrGL-I (R_f 0.35, 6 mg) from $mc^2$22 (Fig. 1B).

Glycosyl composition and linkage analysis.—Purified glycolipids and *O*-methylated/*O*-trideuteriomethylated glycolipids were hydrolyzed with 2 M CF_3CO_2H at 120°C for 1 h. The product sugars were reduced with $NaBD_4$ and the resultant alditols were *O*-acetylated and examined by GLC and GLC–MS[14]. The absolute configurations of the individual sugars were determined by GLC–MS analysis of the trimethylsilyl derivatives of their (*R*)-(−)- and (*S*)-(+)-*sec*-butyl glycosides, which were compared with authentic standards[15]. In order to establish the location of acyl functions on the oligosaccharide backbone, the native glycolipids were subjected to the neutral alkylating conditions of Prehm[16] followed by standard ethylation[17], as follows. The pure glycolipid was resuspended in 200 μL of trimethyl phosphate, under N_2, then 30 μL of 2,6-di-*tert*-butylpyridine and 20 μL of methyltrifluoromethanesulfonate were added. The mixture was stirred at room temperature for 5 h, following which 1 mL of water was added. The mixture was then applied to a C_{18} Sep-Pak cartridge (Waters, Milford, MA), which was eluted successively with water, MeCN, and EtOH. The EtOH eluate was dried to yield the naturally acylated *O*-methylated glycolipid, which was further *O*-ethylated using the Hakomori procedure[17], and purified on a C_{18} Sep-Pak cartridge as described above. The *O*-ethylated *O*-methylated glycolipid was recovered in the MeCN fraction.

Alternatively, NaOH-catalyzed per-*O*-methylation was carried out as follows. The lyophilized sample was dissolved in 1–2 drops of dry, redistilled dimethyl sulfoxide (Me_2SO) and ~1 mL of a Me_2SO–NaOH slurry (prepared by grinding 5 NaOH pellets in ~3 mL of Me_2SO in a dry mortar with pestle) was added, followed by ~0.5 mL of MeI (trideuteriomethyl iodide for *O*-deuteriomethylation). The mixture was placed on an automatic shaker for 10 min at room temperature after which the reaction was quenched by the careful and dropwise addition of ~1 mL of water. The *O*-methylated sample was then extracted into $CHCl_3$ (1 mL) and washed several times with water before drying down under a stream of N_2.

Per-*O*-deuterioacetylation was achieved by suspending lyophilized samples in 50 μL of pyridine, adding 50 μL of hexadeuterioacetic anhydride, and incubating at 80°C for 2 h. The reagents were then evaporated to dryness under a stream of N_2.

NMR analysis.—Routine ^1H, ^{13}C, and 2D COSY NMR spectra were recorded on a Bruker ACE-300 instrument at the Colorado State University Department of Chemistry Central Instrument Facility. Spectra were obtained for the purified glycolipids at concentrations of 2–10 mg/0.5 mL in 2:1 CD_3Cl–CD_3OD, before and after exchanging protons with the same solvent.

Gas–liquid chromatography–mass spectrometry.—GLC–MS of alditol acetates was performed on a Hewlett–Packard 5890 gas chromatograph connected to a Hewlett–Packard 5790 mass-selective detector as described previously[18]. GLC was routinely conducted on a fused silica capillary column of Durabond-1 (J&W Scientific, Rancho Cordova, CA) as described[18].

Fast-atom bombardment mass spectrometry.—FAB mass spectra were obtained using a VG Analytical ZAB-HF mass spectrometer fitted with an M-Scan FAB gun operated at 10 kV with Xe as the bombarding gas. Spectra were recorded on oscillographic chart paper and manually counted. FAB spectra were also obtained using a ZAB-2SE FPD mass spectrometer fitted with a cesium ion gun operated at

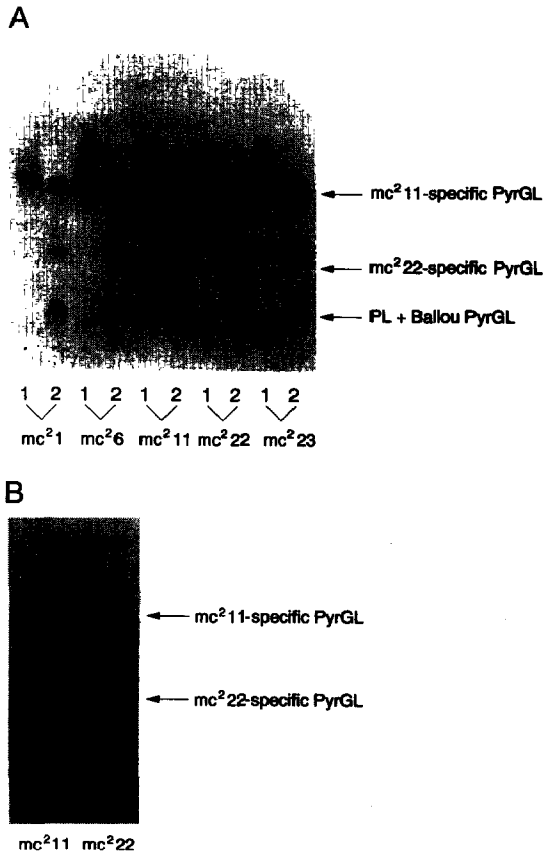

Fig. 1. A, TLC of the acetone-soluble lipids of phage-sensitive ($mc^2$1, $mc^2$6, and $mc^2$23) and phage-resistant ($mc^2$11 and $mc^2$22) strains of *M. smegmatis*. Lane 1 in each case corresponded to the 2% MeOH eluate and Lane 2 to the 5–10% MeOH eluate of a Florisil column (see text). B, TLC of the purified specific PyrGLs from $mc^2$11 and $mc^2$22. The TLC solvent in both cases was 30:8:1 $CHCl_3$–MeOH–water. Both plates were sprayed with 10% H_2SO_4 in EtOH and heated at 110°C for 5 min. In the case of the phage-resistant strain $mc^2$11 (Lane 1, plate B), the characteristic glycolipid, which gave a black color response, corresponded to the band (R_f 0.78) indicated, whereas other nonspecific lipids in this region gave a pink/reddish coloration. PL+Ballou PyrGL, a mixture of unresolved phospholipids and the pyruvylated glycolipids described by Ballou and colleagues[3,4].

25 kV. Data acquisition and processing were performed using the VG Analytical Opus software. Both underivatized and derivatized samples were dissolved in MeOH for aliquoting onto the matrix. All positive-ion spectra were acquired using *m*-nitrobenzyl alcohol as the matrix unless otherwise stated.

Fatty acid analysis.—Pure glycolipids were treated with 3 M HCl in MeOH at 80°C overnight. The hydrolyzates were dried under a stream of N_2, trimethylsilylated using TRI-SIL (Pierce Chemical Co., Rockford, IL) at 80°C for 20 min, again dried under a stream of N_2, redissolved in hexane, and analyzed by GLC–MS.

RESULTS

Isolation of new glycolipids from phage-resistant M. smegmatis.—A broad, TLC-based survey of extractable lipids from phage-sensitive ($mc^2$1, $mc^2$6, and $mc^2$23) and phage-resistant ($mc^2$11 and $mc^2$22) strains revealed the differences shown in Fig. 1A. The two indicated distinctive glycolipids, which were found only in $mc^2$11 and $mc^2$22, when charred with 10% sulfuric acid in ethanol produced the black color response indicative of lipids containing carbohydrate. Alkaline treatment of the total lipids of these organisms resulted in the loss of the acetone-soluble glycolipids, as also occured with the alkali-labile, pyruvylated glycolipids of *M. smegmatis*[3,4] and the lipooligosaccharides of *Mycobacterium kansasii, Mycobacterium szulgai, Mycobacterium malmoense, and Mycobacterium tuberculosis* Canetti[7]. The two apparent lipooligosaccharide-type glycolipids were chromatographically less polar than the rigorously characterized pyruvylated glycolipids described by Ballou and colleagues[3,4], which were also present in our strains of *M. smegmatis* and were the dominant acetone-soluble glycolipids of the phage-sensitive strains (Fig. 1A).

Characterization of the sugar constituents of the PyrGLs.—The nature of the individual glycosyl units in the glycolipids was explored through hydrolysis, reduction with $NaBD_4$, *O*-acetylation, GLC, and GLC–MS of the alditol acetates as previously described[14,18]. GLC of the alditol acetates showed the presence of two sugars in the approximate ratio of 4:1 for PyrGL-I $mc^2$11, whereas PyrGL-I $mc^2$22 contained a single glycosyl entity that cochromatographed with the more abundant sugar component of PyrGL-I $mc^2$11. The alditol from the less abundant sugar component of PyrGL-I ($mc^2$11) was shown by GLC–MS (t_R 12.14 min; *m/z* 130, 190, 201, and 261) and cochromatography with an authentic standard to be 1,2,4,5,6-penta-*O*-acetyl-3-O-methylglucitol. The more abundant alditol was identified as 1,2,3,4,5,6-hexa-*O*-acetylglucitol by GLC–MS (t_R 12.51 min; *m/z* 103, 115, 128, 139, 145, 170, 187, 218, 260, 289, and 362) and cochromatography. Standard, trimethylsilylated 3-*O*-methyl-D-glucose and D-glucose (*S*)-(+)-*sec*-butylglycosides and (*R*)-(−)-*sec*-butylglycosides cochromatographed exclusively with the trimethylsilyl derivatives prepared from the glycolipids, indicating that all of the glycosyl residues were in the D absolute configuration. Thus, the glycolipid from $mc^2$11 contains 3-*O*-methyl-D-glucose and D-glucose, whereas $mc^2$22 contains only D-glucose, not unlike the sugar composition of the pyruvylated glycolipids of *M.*

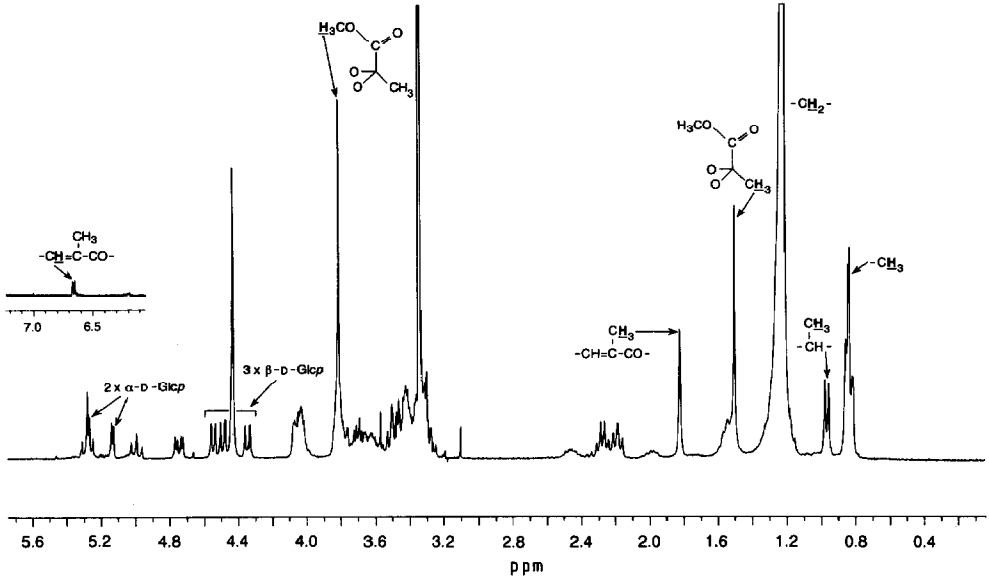

Fig. 2. ^{1}H NMR spectrum (300 MHz) of PyrGL-I of mc^{2}22 at a concentration of 6 mg in 0.5 mL of 2:1 $CDCl_3$–CD_3OD. The inset showing the olefinic region is reduced in scale.

smegmatis described by Ballou and colleagues[3,4]. The ^{1}H NMR spectra of the two glycolipids showed similar resonances and suggested that the glycolipids share the same basic oligosaccharide core but differ from each other and the known *M. smegmatis* lipooligosaccharides in the extent of either *O*-acylation or *O*-methylation.

Specifically, the ^{1}H NMR spectrum of PyrGL-1 mc^{2}22 (Fig. 2) revealed seven resonances, integral value of eight protons, in the anomeric region between δ 4.3 and 5.4. All resonances integrated for one proton except the complex signal centered at δ 5.28, which integrated for two protons. The three upfield resonances at δ 4.35 ($J_{1,2}$ 7.8 Hz), δ 4.48 ($J_{1,2}$ 7.8 Hz), and δ 4.53 ($J_{1,2}$ 7.5 Hz) were readily assigned to three β-D-Glc*p* residues. The remaining signals were assigned after 2D ^{1}H COSY NMR (results not shown). Thus, the resonance at δ 5.13 ($J_{1,2} < 3.8$ Hz) was assigned to H-1 of an α-D-Glc*p* residue. Further connectivities were found to H-2 at δ 3.65, H-3 at δ 4.99 ($J_{2,3} = J_{3,4} = 9.4$ Hz), and H-4 at δ 5.28, showing that this residue was acylated at positions O-3 and O-4. The 2D ^{1}H COSY NMR spectrum also allowed the assignment of a signal for H-1 of a second α-D-Glc*p* residue at δ 5.28 and H-2 at δ 4.75 ($J_{1,2} = 3.5$, $J_{2,3} = 9.4$ Hz), which clearly must be acylated at O-2. The location of these acylated glucosyl residues within PyrGL-I and the fatty acyl composition was determined by FABMS (see later).

The ^{1}H NMR spectra showed other significant signals: a single methine proton at δ 6.64 conjugated with an ester function and split by a neighboring proton; a single ester methyl group at δ 3.80 (~ 3 protons); a weakly split methyl group at δ 1.82 (~ 3 protons); the methyl group of a pyruvate ketal at δ 1.50 (~ 3 protons);

and a methyl strongly split by a single proton at δ 0.93 ($\sim$ 3 protons). The ^{1}H NMR spectrum of PyrGL-I mc^{2}11 yielded resonances very similar to those found for PyrGL-I mc^{2}22 (results not shown). However, a difference among the glycolipids was noted in that PyrGL-I mc^{2}11 yielded three additional signals; a second ester methyl (δ 3.76); an *O*-methyl group (δ 3.60); and a second methyl group of a pyruvate ketal at δ 1.48. It was also noted that the characteristic chemical shifts of the methyl protons of the 4,6-substituted pyruvate ketals (δ 1.50, mc^{2}22; δ 1.50 and δ 1.48, mc^{2}11) were indicative of an *S* configuration[19], where the methyl group is equatorial and the carboxyl acid is axial. ^{13}C NMR was informative in revealing the general nature of the glycolipids as being related to the nonreducing trehalose-containing glycolipids of *M. kansasii* and other species. For instance, the ^{13}C NMR spectrum of PyrGL-I mc^{2}22 showed characteristic resonances of δ 98 and 95 associated with the two C-1's of trehalose.

The positions of glycosyl linkages and fatty acyl substituents were inferred by analyzing the alditol patterns after various alkylation procedures. In the case of the PyrGL-I mc^{2}11, *O*-trideuteriomethylation was performed using the base catalyzed method[17], firstly to determine the linkage pattern and secondly the location of the 3-*O*-methyl-Glc*p* residue. The *O*-trideuteriomethylated oligosaccharides were successively hydrolyzed with 2 M CF_3CO_2H, reduced with $NaBD_4$, and *O*-acetylated. GLC–MS of the resulting alditol acetates revealed the presence of 1,5-di-*O*-Ac-2,3,4,6-tetra-*O*-(Me-d_3)-glucitol (t_R 9.59 min; *m/z* 108, 121, 132, 167, 168, and 214); 1,4,5-tri-*O*-Ac-2,3,6-tri-*O*-(Me-d_3)-glucitol (t_R 10.46 min; *m/z* 121, 168, and 239); 1,5,6-tri-*O*-Ac-2,3,4-tri-*O*-(Me-d_3)-glucitol (t_R 10.65 min; *m/z* 121, 168, 192, and 239); 1,4,5,6-tetra-*O*-Ac-2-*O*-(Me-d_3)-3-*O*-Me-glucitol (t_R 11.42 min; *m/z* 121, 165, and 261); and 1,3,4,5,6-penta-*O*-Ac-2-*O*-(Me-d_3)-glucitol (t_R 11.95 min; *m/z* 121 and 336). Thus, the linkage analysis suggested the presence of a terminal (*t*) Glc*p*, 4-substituted Glc*p*, 6-substituted Glc*p*, 4,6-disubstituted-3-*O*-methyl Glc*p*, and 3,4,6-trisubstituted Glc*p*. The complex substitution pattern associated with the latter two glucosyl residues indicated that these residues carry the two pyruvate ketal groups as first suggested by the ^{1}H NMR data. Therefore, the 4,6-disubstituted-3-*O*-methyl-Glc*p* is probably *t*-3-*O*-methyl-Glc*p* 4,6-pyruvate ketal, and the 3,4,6-trisubstituted Glc*p* is probably 3-linked with a 4,6-substituted pyruvate ketal substituent, as in the known *M. smegmatis* lipooligosaccharide[3,4].

The locations of the fatty acyl residues were determined by the mild methylation procedure of Prehm[16], followed by *O*-ethylation by the Hakomori method[17] in order to tag the positions of base-labile functions. Five new alditol acetates resulted from this treatment: 1,4,5,6-tetra-*O*-Ac-2,3-di-*O*-Me-glucitol; 1,5-di-*O*-Ac-3,4-di-*O*-Et-2-*O*-Me-glucitol; 1,4,5-tri-*O*-Ac-2,3,6-tri-*O*-Me-glucitol; 1,5,6-tri-*O*-Ac-2-*O*-Et-3,4-di-*O*-Me-glucitol; and 1,3,4,5,6-penta-*O*-Ac-2-*O*-Me-glucitol. Accordingly, the existence of a terminal 3-*O*-methyl-β-D-Glc*p* 4,6-pyruvate ketal; a terminal α-D-Glc*p* with acyl substituents at positions 3 and 4; a 4-linked β-D-Glc*p*; a 6-linked α-D-Glc*p* with an acyl substituent at position 2; and a 3-linked β-D-Glc*p* 4,6-pyruvate ketal was established.

A similar analytical strategy established that PyrGL-I mc^{2}22 has similar features except that the terminal 3-*O*-methyl-β-D-Glc*p* 4,6-pyruvate ketal was replaced by a *t*-β-D-Glc*p* residue, which was in accordance with the ^{1}H NMR evidence pointing to the presence of only one pyruvate ketal and the absence of an *O*-methylated sugar. It was suspected that the presence of the additional ester-methyl resonances in the ^{1}H NMR spectra of the glycolipids implied that the carboxyl groups of the 4,6-pyruvate ketal residues were in fact in methyl ester form. The fatty acyl groups of the glycolipids were liberated as methyl esters, and analysis by GLC–MS suggested the presence of C_{22}, C_{16}, and C_{14} residues. These were later shown by cochromatography and by their characteristic fragment ions in EIMS to be 2,4-dimethyl-2-eicosenoic (C_{22}), tetradecanoic (C_{14}), and hexadecanoic (C_{16}) acids, respectively. Additionally, octanoic acid (C_8), which due to its high volatility was not found by GLC–MS analysis, was unequivocally shown to be present by FABMS (see later).

The analysis to date suggests that PyrGL-I mc^{2}11 is, in fact, very similar in nature to the acidic oligosaccharides from *M. smegmatis* (ATCC 356). The distinguishing features are the methylation of the pyruvate ketal groups and the presence of a trehalose 3,4,2′-tri-*O*-acylated residue, whereas the pyruvylated glycolipids described by Ballou et al.[3,4] are characterized by nonmethylation of the 4,6-pyruvate ketals and a 6,4′-di-*O*-acylated trehalose moiety. This increase in

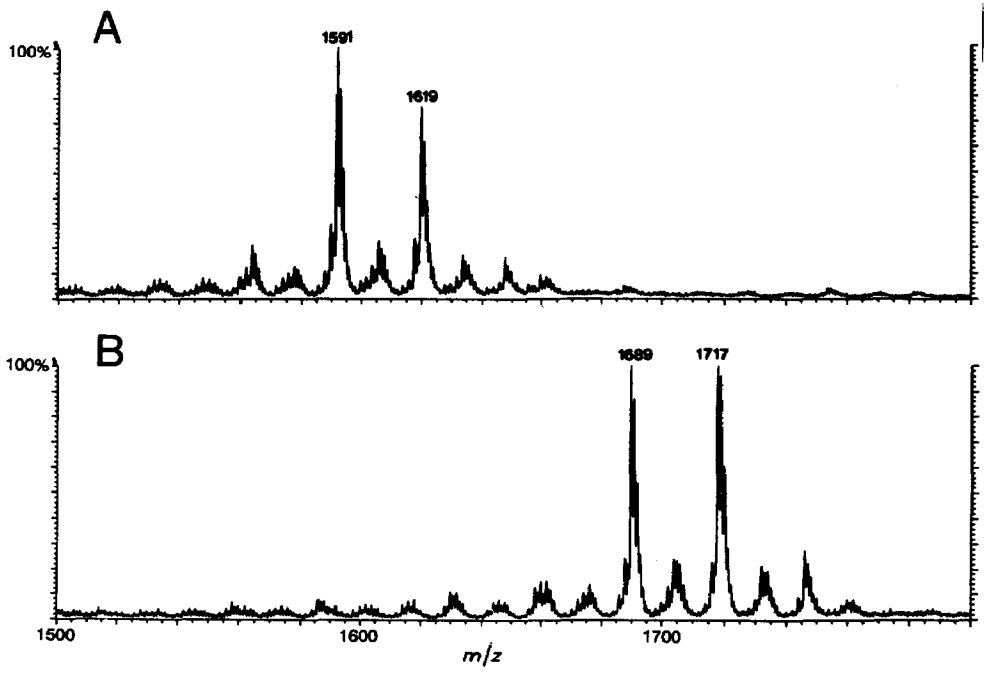

Fig. 3. The molecular ion regions of the positive FAB spectra of underivatized the PyrGLs from A, mutant mc^{2}22 and B, mc^{2}11.

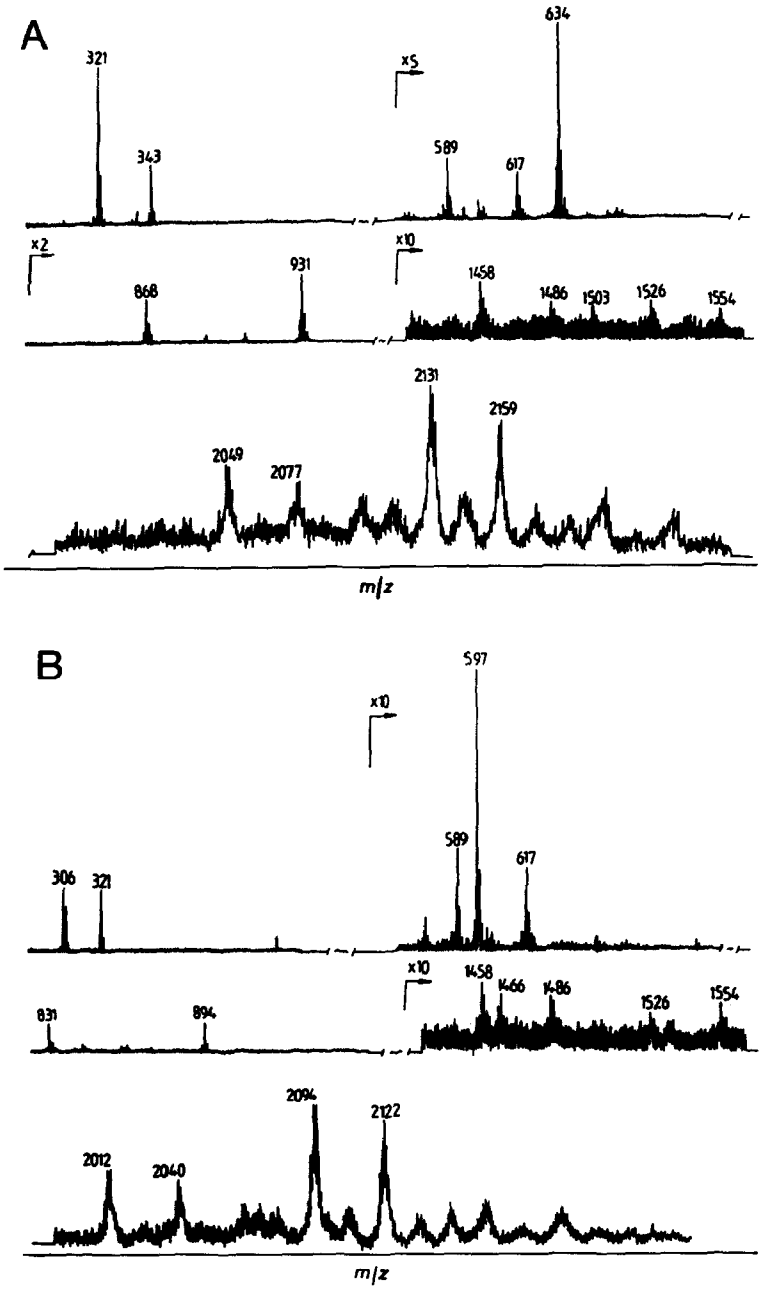

Fig. 4. The positive FAB spectra of *O*-deuterioacetylated PyrGLs from A, $mc^{2}22$ and B, $mc^{2}11$. The assignments of the major signals are summarized in Fig. 5. Signals at m/z 868 and 831 correspond to loss of a trideuterioacetic acid moiety from m/z 931 and 894, respectively. The minor molecular ion pairs at m/z 2049/2077 in A and m/z 2012/2040 in B probably correspond to a saturated C_{16} fatty acyl substituent being present in place of the unsaturated C_{22} fatty acid. These minor components were not observed in the native sample, a not uncommon phenomenon.

O-acylation and *O*-methylation explains the relatively nonpolar behavior, in migration patterns on TLC, of the glycolipids from mc^{2}11 and mc^{2}22 as compared with those studied by Ballou and colleagues[3,4].

The complete sequence of the glycosyl units in the oligosaccharide cores, the positions of the carboxyl-methylated pyruvated ketals, and the disposition of the fatty acyl residues were determined unequivocally by FABMS as discussed below.

FABMS analysis of the PyrGLs of M. smegmatis *mc^{2}11, mc^{2}22.*—The pyruvylated glycolipids isolated from *M. smegmatis* mutants mc^{2}22 and mc^{2}11 each afforded a major pair of molecular ions separated by 28 amu in the positive ion mode (Fig. 3). The $[M + Na]^+$ molecular ions afforded by mc^{2}11 at *m/z* 1689 and 1717 are 98 amu higher than those given by mc^{2}22 (*m/z* 1591 and 1619). In the negative ion mode, mc^{2}22 afforded $[M - H]^-$ molecular ion signals at *m/z* 1567 and 1595, while mc^{2}11 yielded signals at *m/z* 1665 and 1693 (data not shown), thus defining the molecular weights of the glycolipids from mc^{2}22 and mc^{2}11 as 1568/1596 and 1666/1694 Da, respectively. After *O*-deuterioacetylation with retention of the acyl substituents[18], mc^{2}22 yielded a similar pair of $[M + Na]^+$ molecular ions at *m/z* 2131 and 2159 (Fig. 4A), the mass shifts of which indicate the presence of 12 free OH groups as well as suggesting that the 28 amu heterogeneity was due to variation in the fatty acyl substituent(s). The *O*-deuterioacetylated mc^{2}11 (Fig. 4B) also retained the 28 amu heterogeneity, but the mass shifts of the molecular ion pairs correspond to the presence of only nine free

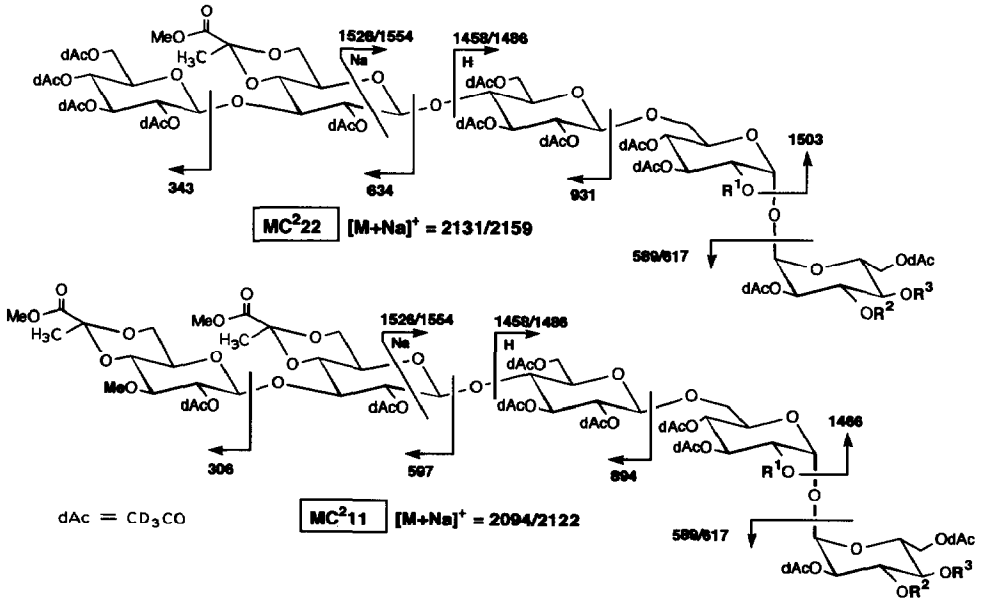

Fig. 5. Assignments of the fragment-ion signals in the positive FAB spectra of *O*-deuterioacetylated PyrGLs from mc^{2}22 and mc^{2}11 (Fig. 4). dAc represents the trideuterioacetyl group; R^1–R^3 represent fatty acyl substituents.

OH groups. The 98-amu difference between underivatized glycolipids from $mc^2 22$ and $mc^2 11$ which translated into a minus 37-amu difference after *O*-deuterioacetylation was therefore consistent with $mc^2 11$ having an additional methyl-esterified pyruvate (Me-Pyr, 84 amu) and an *O*-methyl (14 amu) substituent, as indicated by the glycosyl composition and NMR analyses. The locations of these additional substituents were unambiguously defined by the characteristic A-type ions afforded by the positive FAB spectra of the *O*-deuterioacetylated derivatives (Fig. 4).

As schematically summarized in Fig. 5, the ions at *m/z* 343 and 306, afforded by $mc^2 22$ and $mc^2 11$, respectively, unambiguously localized both the additional methyl-esterified pyruvate and the *O*-methyl substituents on the terminal hexose of $mc^2 11$, while ions at *m/z* 634 (for $mc^2 22$) and 597 (for $mc^2 11$) localized the common pyruvate group onto the penultimate residue. Further A-type ions at *m/z* 931 (for $mc^2 22$) and 894 (for $mc^2 11$) were consistent with a third unsubstituted hexose residue, while a fourth pair of ions at *m/z* 1503 ($mc^2 22$) and 1466 ($mc^2 11$), which also differed by 37 amu, corresponded to a hexose acylated by 2,4-dimethyl-2-eicosenoate (C_{22}). This fatty acyl group yielded an acylium ion at *m/z* 321 in the FAB spectrum (Fig. 4), with no apparent 28-amu heterogeneity. Conversely, all three pairs of fragment ions inferred to contain the terminal Glc*p* of the acyltrehalose unit also consistently exhibited the heterogeneity. Thus, signals at *m/z* 589 and 617 may be assigned as the A-type ions of the common acylated terminal hexose, while the pairs of ions at *m/z* 1458/1486 and 1526/1554 correspond to protonated glycosidic cleavage ions and sodiated ring cleavage ions, respectively, as schematically illustrated in Fig. 5.

The delineation of the fatty acyl substituents on this terminal Glc was facilitated by FABMS analysis of the *O*-methylated derivatives. Under the NaOH per-*O*-methylation[20] conditions[21], the alkali-labile fatty acyl substituents are partially or completely retained while an *O*-acetyl substituent is completely replaced by a methyl group[22]. The molecular ions afforded by each of the partially de-*O*-acylated products thus provide an informative "map" of the fatty acyl substituents, especially when compared with analogous data obtained after Prehm methylation[16] in which the acyl and acetyl groups are fully retained. Thus, the Prehm-methylated $mc^2 22$ sample yielded a single major pair of $[M + Na]^+$ molecular ion signals at *m/z* 1759 and 1787, corresponding to fully methylated $mc^2 22$ PyrGL retaining all of the fatty acyl functions (results not shown). In contrast, after NaOH per-*O*-methylation, the signals at *m/z* 1759/1787 were very weak. A minor pair of signals was observed at *m/z* 1647/1675, while a dominant molecular ion signal was seen at *m/z* 1451. The former pair of signals corresponds to the replacement of a C_8 fatty acyl substituent with an *O*-methyl group (an increment of 112 amu), while the latter major signal corresponds to the additional replacement of a C_{14}/C_{16} moiety with an *O*-methyl group (an increment of 196/224 amu). The information obtained confirms that the heterogeneity is associated with the C_{14}/C_{16} fatty acyl group and not, for example, C_8/C_{10}. Also, the retention of the C_{22} substituent (signal at *m/z* 1451) is consistent with it being attached to the less

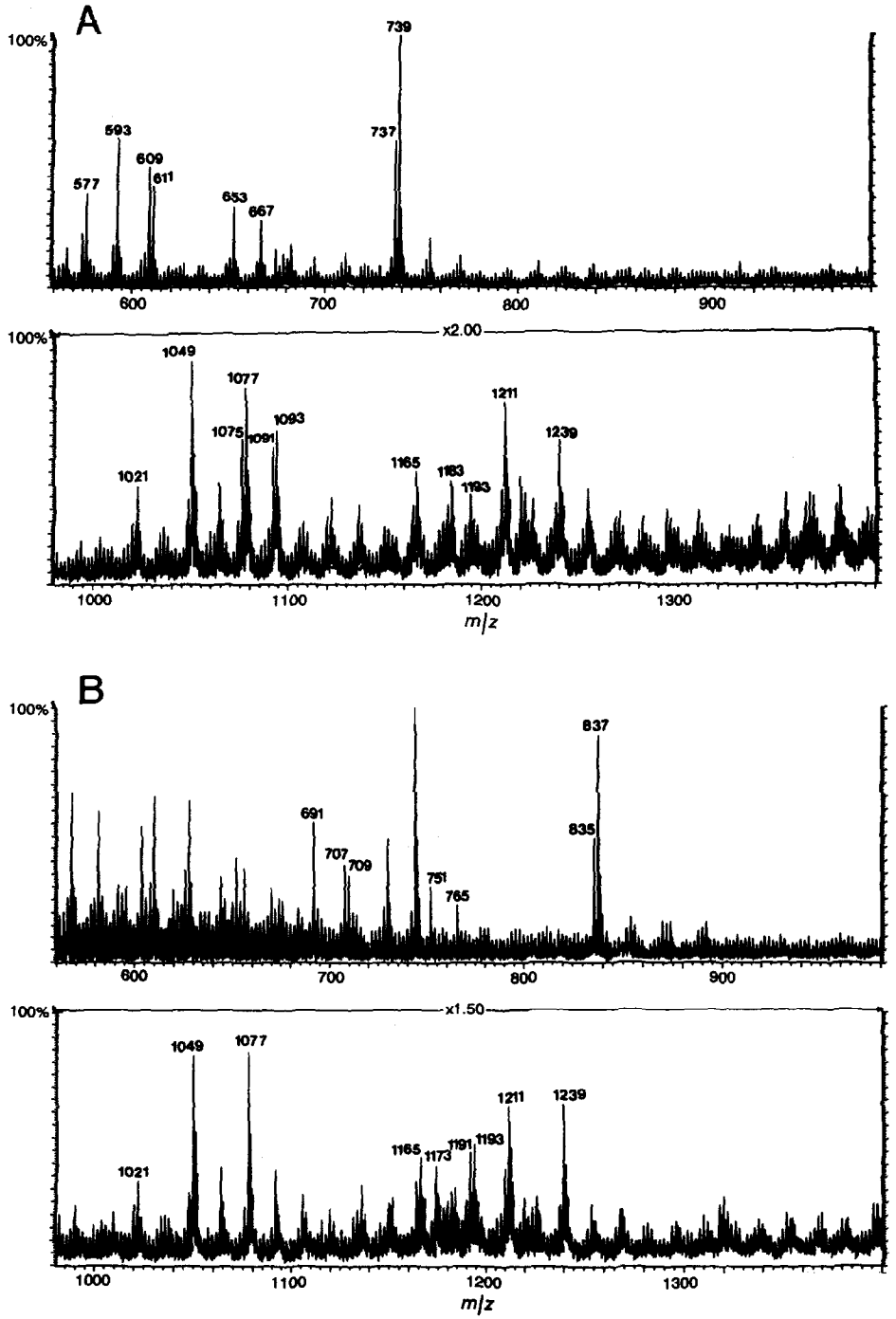

Fig. 6. Positive FAB spectra of underivatized PyrGLs from A, $mc^2 22$ and B, $mc^2 11$ showing the major fragment ions observed in the region m/z 560–1400. The majority of signals are assigned in the text. The peaks below m/z 750 in B that are not labeled are attributed to impurities. Signals at 2 amu below assigned signals (e.g., 737 and 835) correspond to structures having an additional double bond.

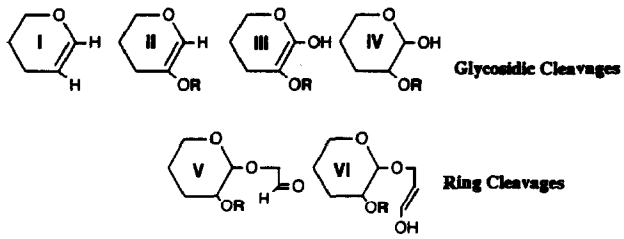

Fig. 7. A schematic illustration of the various modes of glycosidic (I–IV) and ring cleavage (V and VI) associated with cationized molecular ions.

accessible inner residue. Similar data were obtained from the NaOH-per-*O*-methylated mc²11 PyrGL. Only very weak $[M + Na]^+$ molecular ions were obtained after *O*-methylation which corresponded to the retention of all the fatty acyl groups (*m/z* 1815/1843); however, $[M + Na]^+$ molecular ions at *m/z* 1507 and 1703/1731 were also clearly present. The latter pair of ions corresponds to the replacement of a C_8 fatty acyl with an *O*-methyl group, while the ion at *m/z* 1507 corresponds to the additional loss of the C_{14}/C_{16} heterogeneity. In support of the terminal Glc*p* being di-*O*-acylated with a C_8 and a C_{14}/C_{16} fatty acid are the A-type ions at 589 and 617 afforded by the *O*-deuterioacetylated samples (Fig. 4) mentioned earlier.

Further corroborative data on the sequence was provided by detailed analysis of the fragment-ion signals observed in the positive FAB spectra of underivatized PyrGL samples (Fig. 6). The cluster of ions afforded by mc²22 in the region of *m/z* 560–680 (Fig. 6A) may be rationalized as resulting from the various modes of glycosidic and ring cleavages often associated with cationized molecular ions[22,23]. Thus, signals at *m/z* 577, 593, 609, and 611 correspond to sodiated fragment ions [Hex-(MePyr)Hex-Hex] of the structures **I**, **II**, **III** and **IV**, respectively (Fig. 7), resulting from cleavages at either side of the glycosidic oxygen of the third hexose residue, whereas signals at *m/z* 653 and 667 correspond to structures **V** and **VI** (Fig. 7), resulting from ring cleavages at the fourth hexose residue. Significantly, the mass difference between structures **I** and **II** is dependent on the substituent at position 2, indicated as -OR in Fig. 7.

Thus, the analogous set of cleavages at the next hexose residue, which is acylated at position 2, yielded ions at *m/z* 739, 1075, 1091, and 1093. Notably, the mass difference between 739 (structure **I**) and 1075 (structure **II**), viz 336 amu, is consistent with R being the C_{22} substituent (320 amu increment). These two sets of ions confirmed the partial sequence Hex-(MePyr)Hex-Hex-(2-*O*-C_{22} acyl)Hex for mc²22 as shown in Fig. 5. For mc²11 (Fig. 6B), both sets of ions, i.e., *m/z* 593, 609, 611, 653, 667 and *m/z* 739, 1075, 1091, 1093 were shifted to 98 amu higher at *m/z* 691, 707, 709, 751, 765 and *m/z* 837, 1173, 1189, 1191, respectively, consistent with the presence of additional *O*-methyl and MePyr groups. Finally, common fragment ions relating to the acyltrehalose terminal and carrying the 28

amu heterogeneity were observed for the PyrGLs of both mutants (Figs. 6A and 6B). Sodiated β-cleavage * hexose-triacyltrehalose ions yielded signals at m/z 1183/1211, while the related ring-cleavage † ions yielded signals 28 amu higher at m/z 1211/1239. Loss of a water molecule from the former pair gave rise to signals at m/z 1165/1193. Similar sodiated β-cleavage and ring-cleavage ions of the triacyltrehalose core afforded signals at m/z 1021/1049 and 1049, 1077, respectively.

In summary the FABMS data unambiguously defined the sequence of PyrGL-I $mc^2 22$ as Hex-(MePyr)Hex-Hex-(2-*O*-C_{22} acyl)-Hex-(di-*O*-C_{10}, $C_{14,16}$ acyl)Hex and PyrGL-I $mc^2 11$ as having additional methyl-esterified pyruvate (MePyr) and *O*-methyl substituents on the nontrehalose terminal hexose. The GLC–MS analysis presented earlier demonstrated that the fatty acyl residues on the terminal α-D-Glc*p* residue reside at O-3 and O-4; however, the specific placement of particular acyl residues (C_8 or C_{14}/C_{16}) at these positions was not deduced in this present study.

DISCUSSION

The discovery of pyruvylated acylated trehaloses by Ballou and colleagues arose through a typical astute observation during the course of an entirely different pursuit, namely a search for biosynthetic precursors of the 6-*O*-methylglucose-containing lipopolysaccharides. In the present instance, we were seeking chemical manifestations of phage resistance with predetermined notions and observed only by accident alterations in the basic structures described by Saadat and Ballou[3,4]. The apolar *C*-mycoside glycopeptidolipids, prominent in *M. avium* and a few other mycobacteria[1], including *M. smegmatis*[24], and, indeed, in the present strains of *M. smegmatis* (results not shown), demonstrate decided activity for adsorption and inactivation of mycobacteriophage D4 and were long thought to serve as the primary receptors for this lytic phage[25,26]. Indeed, Goren and colleagues[27] concluded that the terminal methylated rhamnose unit of the core structure of the glycopeptidolipids is the primary D4 phage receptor site. Accordingly, we approached the question of the molecular basis of phage resistance in the belief that structural alterations of the glycopeptidolipid molecule would be involved, and with little thought of the "newer" pyruvylated glycolipids.

There are precedents for the present study in the extensive work conducted on the chemical basis of lysogenization in strains of *Salmonella typhimurium*. Apparently all of the A phages adsorb to the *O*-polysaccharide chains of the lipopolysaccharides of susceptible strains. In particular, the changes caused by a lysogenic infection with phages of groups A1 and A2 resulted in substitution by an α-linked glucose at C-6 of the galactosyl residue of the *O*-antigen repeat unit[28]. This

* Cleavage on the 'C-1 side' of the glycosidic oxygen, which becomes part of the charged fragment.
† Cleavage of the type giving m/z 1526/1554 from the trideuterioacetylated lipids (Fig. 5).

substitution, which amounts to the expression of the *O*-1 character, interfered with adsorption of phage P22. From this and other work[29], it can be concluded that lysogeny for several of the A phages causes changes in the composition of the *O*-polysaccharide chain which in turn reduce or abolish the adsorption of other phages. Of greater relevance to the present study is the work of Wollin et al.[30], who demonstrated that lysogenization of *S. typhimurium* with either of the bacteriophages A3 or A4 resulted in acetylation of the L-rhamnosyl residues of the O-antigen. This lysogenic conversion then prevented the absorption of the A3 and A4 phages themselves. In the present context, the lessons learned from extensive study of the chemical basis of lysogenization in *S. typhimurium* suggest that the additional O-acylation and/or O-methylation of the pyruvylated glycolipids observed here is the basis of D29 phage resistance in *M. smegmatis*. It is therefore a gratifying thought, on the occasion of this tribute to C.E. Ballou, that he may have opened the door to a molecular understanding of lysogenization in mycobacteria.

ACKNOWLEDGMENTS

We thank Dr. William Jacobs for the gift of mc^2 strains of *M. smegmatis*, and Russell Suzuki and Steve Rivoire for valuable technical assistance. This work was supported by Grant AI-18357 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (to PJB); a Medical Research Council Programme Grant; Wellcome Trust Grants Number 030826 (to HRM and AD) and 036485 (Prize Fellowship to KHK). The NMR conducted at the Colorado State University Regional NMR Center was supported by National Science Foundation Grant CHE 86-16437.

REFERENCES

1 P.J. Brennan, in C. Ratledge and S.G. Wilkinson (Eds.), *Microbial Lipids*, Vol. I, Academic Press, London, 1988, pp 203–298.
2 M. McNeil, D. Chatterjee, S.W. Hunter, and P.J. Brennan, *Methods Enzymol.*, 179 (1989) 215–242.
3 S. Saadat and C.E. Ballou, *J. Biol. Chem.*, 258 (1983) 1813–1818.
4 K.-I. Kamisango, S. Saadat, A. Dell, and C.E. Ballou, *J. Biol. Chem.*, 260 (1985) 4117–4121.
5 K.-S. Kim, M.R.J. Salton, and L. Barksdale, *J. Bacteriol.*, 125 (1976) 739–743.
6 J.T. Belisle and P.J. Brennan, *J. Bacteriol.*, 171 (1989) 3465–3470.
7 M.R. McNeil and P.J. Brennan, *Res. Microbiol.*, 142 (1991) 451–463.
8 P. Draper, in C. Ratledge and J. Stanford (Eds.), *The Biology of the Mycobacteria*, Vol. 1, Academic Press, New York, 1982, pp 9–52.
9 S.B. Snapper, R.E. Melton, T. Kieser, S. Mistafa, and W.R. Jacobs, Jr., *Mol. Microbiol.*, 4 (1990) 1911–1919.
10 J.T. Belisle, L. Pascopella, J.M. Inamine, P.J. Brennan, and W.R. Jacobs, Jr., *J. Bacteriol.*, 173 (1991) 6991–6997.
11 W. Jones and A. White, *Can. J. Microbiol.*, 14 (1968) 551–555.
12 W.D. Jones, Jr. and R.E. Beam, *Can. J. Microbiol.*, 15 (1969) 1112–1114.
13 P.J. Brennan and M.B. Goren, *J. Biol. Chem.*, 254 (1979) 4205–4211.
14 M. McNeil, A.Y. Tsang, and P.J. Brennan, *J. Biol. Chem.*, 262 (1987) 2630–2635.
15 G.J. Gerwig, J.P. Kamerling, and J.F.G. Vliegenthart, *Carbohydr. Res.*, 62 (1978) 349–357.

16 P. Prehm, *Carbohydr. Res.*, 78 (1980) 372–374.
17 K. Stellner, H. Saito, and S. Hakomori, *Arch. Biochem. Biophys.*, 262 (1973) 2630–2635.
18 G.S. Besra, M.R. McNeil, B. Rivoire, K.-H. Khoo, H.R. Morris, A. Dell, and P.J. Brennan, *Biochemistry*, 32 (1993) 347–355.
19 P.J. Garegg, P.-E. Jansson, B. Lindberg, F. Lindh, J. Lönngren, I. Kvarnstrom, and W. Nimmich, *Carbohydr. Res.*, 78 (1980) 127–132.
20 I. Ciucano and F. Kerek, *Carbohydr. Res.*, 131 (1984) 209–217.
21 A. Dell, *Methods Enzymol.*, 193 (1990) 647–660.
22 A. Dell, K.-H. Khoo, M. Panico, R.A. McDowell, A.T. Etienne, A.J. Reason, and H.R. Morris, in M. Fukuda (Ed.), *Glycobiology: A Practical Approach*, IRL Press, Oxford, in press.
23 A. Dell, P. Azadi, P. Tiller, J. Thomas-Oates, H.J. Jennings, M. Beurret, and F. Michon, *Carbohydr. Res.*, 200 (1990) 59–76.
24 M. Daffe, M.A. Laneelle, and G. Puzo, *Biochim. Biophys. Acta*, 751 (1983) 439–443.
25 A. Furuchi and T. Tokunaga, *J. Bacteriol.*, 111 (1972) 404–411.
26 M.B. Goren, J.K. McClatchy, B. Martens, and O. Bronl, *J. Virol.*, 9 (1972) 999–1003.
27 K.R. Dhariwal, A. Liav, A.E. Vatter, G. Dhariwal, and M.B. Goren, *J. Bacteriol.*, 168 (1986) 283–293.
28 O. Lüderitz, A.M. Staub, and O. Westphal, *Bacteriol. Rev.*, (1966) 192–255.
29 R. Wollin, U. Eriksson, and A.A. Lindberg, *J. Virol.*, 38 (1981) 1025–1033.
30 R. Wollin, B.A.D. Stocker, and A.A. Lindberg, *J. Bacteriol.*, 169 (1987) 1003–1009.