Carbohydrate Research, 120 (1983) 95–111
Elsevier Science Publishers B V., Amsterdam – Printed in The Netherlands

FAST-ATOM-BOMBARDMENT, NEGATIVE-ION MASS SPECTROMETRY OF THE MYCOBACTERIAL *O*-METHYL-D-GLUCOSE POLYSACCHARIDE AND LIPOPOLYSACCHARIDES*,†

ANNE DELL
Department of Biochemistry, Imperial College, London SW7 2AZ (Gt. Britain)

AND CLINTON E. BALLOU
Department of Biochemistry, University of California, Berkeley, California 94720 (U.S.A.)

(Received December 20th, 1982; accepted for publication, March 18th, 1983)

ABSTRACT

The mycobacterial *O*-methyl-D-glucose polysaccharide (MGP) and its acylated derivatives, the *O*-methyl-D-glucose lipopolysaccharides (MGLP), have been analyzed by fast-atom-bombardment mass spectrometry (F.a.b.m.s.). The molecular-ion peaks for MGP confirmed the degree of polymerization and, in addition, revealed a slight heterogeneity in the degree of methylation. Sequence ions were observed that are consistent with all known features of MGP structure. Spectra of the MGLP isomers confirmed the general distribution of the neutral and acidic acyl groups, and provided additional information regarding the specific location of individual acyl groups. The spectra of some MGLP preparations revealed that they were contaminated by lysophosphatidylinositol dimannoside. The study further documents the utility of F.a.b.m.s. for characterization of relatively large and complex carbohydrate derivatives.

INTRODUCTION

The ability to obtain precise molecular weights for large oligosaccharides or small polysaccharides is a great help in deciding between alternative plausible structures. If characteristic sequence data are also available, still more refined structural assignments can be made. Both of these criteria can now be satisfied by the technique of fast-atom-bombardment mass spectrometry[1] (F.a.b.m.s.), a method that has revolutionized the structural determination of relatively large and labile biomolecules[2,3].

*Dedicated to Professor Elvin A. Kabat.

†This work was supported by Grant AI-12522 from the National Institute of Allergy, Immunology, and Infectious Diseases, National Institutes of Health, and National Science Foundation Grant PCM 80-23388 (to C.E.B.), and by grants from the Medical Research Council and the Science and Engineering Research Council (to H.R Morris).

0008-6215/83/$ 03.00

1 R = $CH_3(CH_2)_6CO$, $R' = HO_2C(CH_2)_4CO$, $R'' = CH_3CO$ CH_3CH_2CO or Me_2CHCO

(MLP-I,-II -III, and -IV contain 0,1,2, or 3 residues R', respectively)

2 R = R' = H

3 R = acid

In this paper, we describe the application of F.a.b.m.s. to the mycobacterial *O*-methyl-D-glucose polysaccharide[3,4] (MGP, **2**) and its acylated derivatives, the *O*-methyl-D-glucose lipopolysaccharides[5] (MGLP, **1**). Although these molecules have complex structures and molecular weights that range from 3500 to 4000, the spectra are completely interpretable and are consistent with the previous structural asssignments. In addition, they reveal, in the structures of these macromolecules, new and subtle features that were not apparent from earlier studies[3–5].

EXPERIMENTAL

Materials. — The *O*-methyl-D-glucose polysaccharide and its acylated forms were prepared by Dr. L. S. Forsberg according to Gray and Ballou[6]. Glycerol was a standard analytical reagent from Fisons Scientific Apparatus (Leicestershire LE11 0RG, U.K.) and thioglycerol was from Sigma Chemical Co. (St. Louis, MO 63178, U.S.A.).

Methods. — F.a.b.m.s.[1,2] was performed with a VG Analytical ZAB 1F High-Field Magnet mass spectrometer. The compounds were dissolved in 5% acetic acid and the sample (1–10 μg in 1 μL) was loaded into a drop of glycerol (~2 μL) on the probe target. In some instances, thioglycerol (~1 μL) was also added. The atom gun was operated at 8 kV with Xenon as the bombarding gas. Negative-ion spectra were recorded at 6 or 8 kV accelerating voltage on an oscillograph, and masses were determined by counting the spectral lines. Recorded masses, therefore, are lower than calculated ones (based on C 12.00, H 1.01, and O 15.99) by ~1 mass unit at a M_r of 2500.

Gel filtration was carried out on a Bio-Gel P-4 column (0.8 × 100 cm) in 0.1M ammonium acetate to separate charged oligosaccharides, and on a Bio-Gel P-2 column (0.6 × 80 cm) in water to desalt the fractions. Oligosaccharides were hydrolyzed with 2M trifluoroacetic acid for 2 h at 120° in a sealed tube, and the monosaccharide products were chromatographed descending on Whatman No. 1 paper in 10:3:3 (v/v) 1-butanol–pyridine–water for 30 h. A silver nitrate–sodium hydroxide dip reagent was used to reveal the sugar-containing spots on the paper.

RESULTS AND DISCUSSION

The questions we have sought to answer by F.a.b.m.s. are: (a) What is the precise molecular weight of MGP (**2**) and is there evidence for heterogeneity in the number of hexose units or in the extent of methylation; (b) does the fragmentation pattern for MGP (**2**) eliminate alternative structures; (c) how are the five small acyl groups (three acetyl, one propionyl, and one isobutyryl) distributed at the non-reducing end of MGLP-I (**1**) and is there heterogeneity in the extent of such acylation; (d) does the fragmentation pattern confirm that the octanoate group in MGLP-I (**1**) esterifies the glyceric acid end-group of the chain; and (e) what structural information can be obtained concerning the more acidic forms of MGLP (**1**)

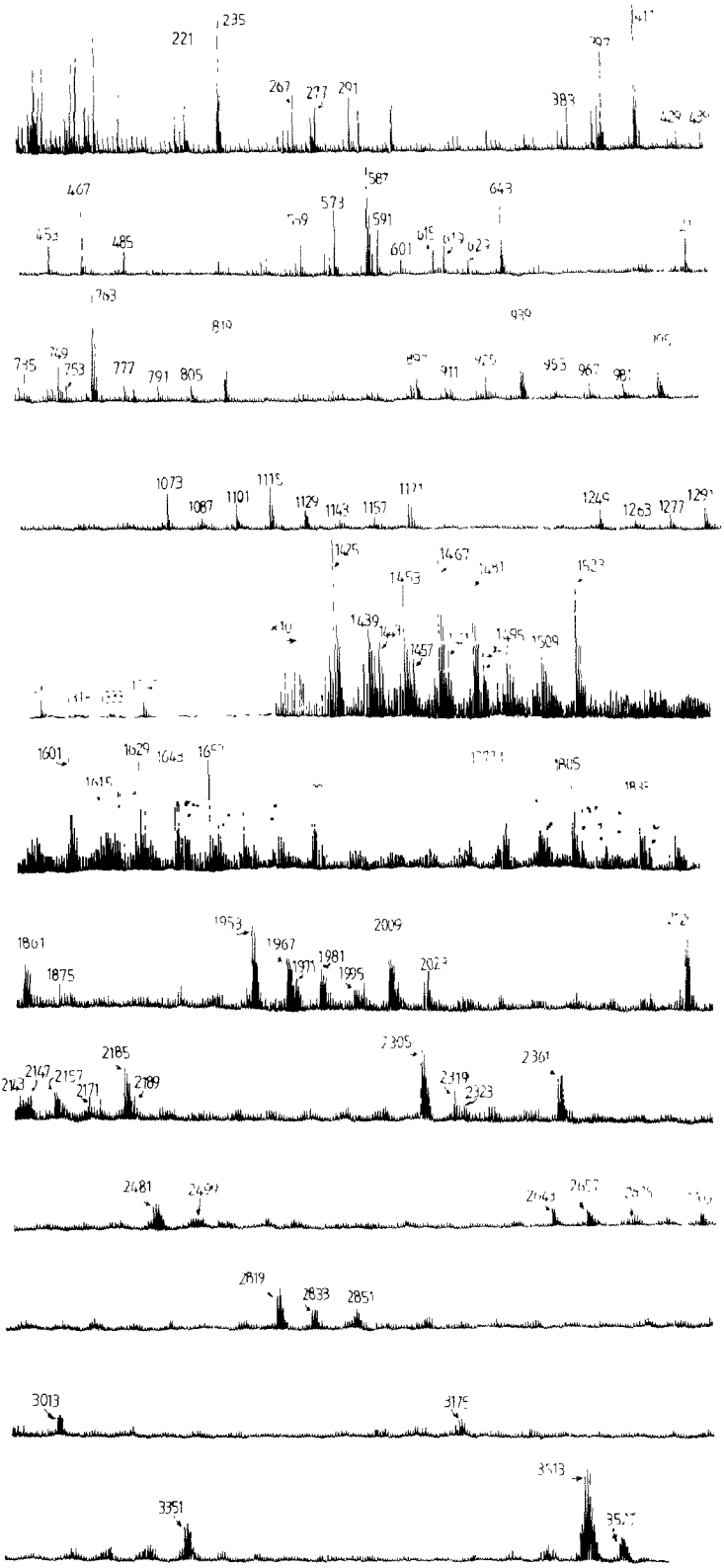


Fig. 1 Negative ion F.a b mass spectrum of MGP (**2**). The figure was prepared by tracing the oscillograph print-out, and masses were assigned by counting the lines in the spectrum

TABLE I

FRAGMENT IONS FROM MGP (2) CONTAINING GLYCERIC ACID

Observed ion	*Proposed composition*[a]
3513 (M − H)	$MG_{12}G_8Ga$
3351	$MG_{12}G_7Ga$
3337	$MG_{11}G_8Ga$
3175	$MG_{11}G_7Ga$
3013	$MG_{11}G_6Ga$
2851	$MG_{11}G_5Ga$
2675	$MG_{10}G_5Ga$
2499	MG_9G_5Ga
2323	MG_8G_5Ga
2189	$MG_{10}G_2Ga$
2175	MG_9G_3Ga
2147	MG_7G_5Ga
1971	MG_6G_5Ga
1837	MG_8G_2Ga
1823	MG_7G_3Ga
1809	MG_6G_4Ga
1795	MG_5G_5Ga
1661	MG_7G_2Ga
1647	MG_6G_3Ga
1633	MG_5G_4Ga
1619	MG_4G_5Ga
1485	MG_6G_2Ga
1471	MG_5G_3Ga
1457	MG_4G_4Ga
1443	MG_3G_5Ga
1295	MG_4G_3Ga
1267	MG_2G_5Ga
1119	MG_3G_3Ga
1091	MG_1G_5Ga
753	G_4Ga
591	G_3Ga
429	G_2Ga
267	G_1Ga

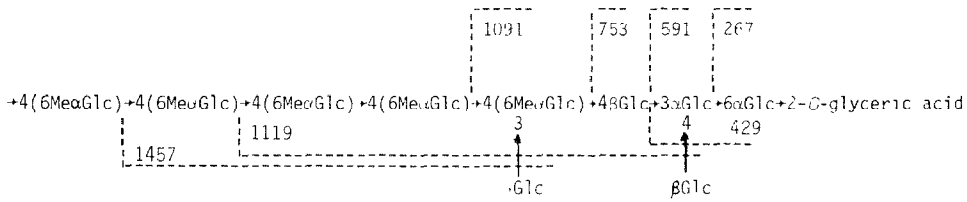
[a]Abbreviations: G, D-glucose; MG, *O*-methyl-D-glucose; and Ga, glyceric acid.

that contain one (MGLP-II), two (MGLP-III), or three (MGLP-IV) moles of esterified succinic acid? The following sections deal with each of these points.

The size of MGP. — The *calculated* molecular weight of MGP (**2**), based on previous studies[3], is 3515.2 (C 12.00). The *counted* negative-ion spectrum (Fig. 1) showed a major signal at *m/z* 3513, with a more intense line at *m/z* 3514 for the isomer containing a single ^{13}C atom, which correspond to M_r values of 3514 and 3515, respectively. A cluster of signals at 14 mass-units higher, with an intensity of ~0.25 relative to the cluster at *m/z* 3514, was also observed and suggests that there is heterogeneity in the degree of methylation of MGP (**2**), with ~20% of the molecules containing 13 methyl groups instead of 12. Thus, the observed M_r of MGP is consistent with a composition of 8 D-glucose, 12 *O*-methyl-D-glucose, and 1 glyceric acid units[3].

Ions produced by fragmentation with loss of hexose from a nonreducing end-group of MGP (2). — Loss of hexose from a nonreducing end of MGP (**2**) by cleavage at glycosidic bonds yielded ions that carry the negative charge of the glyceric acid (Table I). (Although MGP is nonreducing, the glyceric acid end of the chain is the equivalent of the reducing end.) Because MGP (**2**) is branched and contains both terminal D-glucosyl and 3-*O*-methyl-D-glucosyl groups, ions at *m/z* 3351 (3513 − 162) and 3337 (3513 − 176) were expected and were observed (Fig. 1). An ion for the loss of two hexose units (one D-glucose and one *O*-methyl-D-glucose) was seen at *m/z* 3175. Finally, at *m/z* 3013 was observed an ion expected for the loss of two D-glucose and one *O*-methyl-D-glucose units. These results confirm the previous conclusion that MGP has a single terminal *O*-methyl-D-glucosyl group and at least one terminal D-glucosyl group[3]. An ion at *m/z* 2851 was also observed that corresponds to the loss of one *O*-methyl-D-glucose and three D-glucose units.

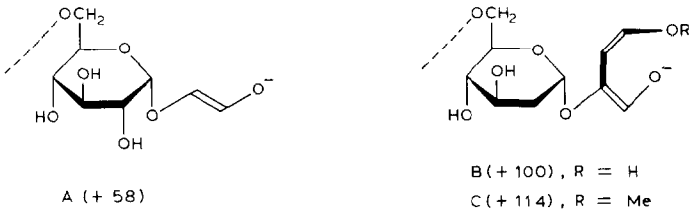
Lower-molecular-weight ions are also present in this series, an important point being that those with fewer than 11 *O*-methylhexose never have more than five D-glucose units (Table I). For some reason, all of the ions of this series were not observed, and the intensities of those seen varied considerably. The overall results, however, are consistent with the conclusions of earlier studies[3] that MGP (**2**) has two terminal D-glucosyl residues near the glyceric acid end-group of the chain and Scheme 1 rationalizes the fragmentation pattern leading to the smaller ions listed in Table I. Four of the ions in this series, at *m/z* 1485, 1661, 1837, and 2189, appear to contain only two D-glucose units, and they are assumed to arise from the minor MGP component that contains 13 *O*-methyl-D-glucose units. If this assignment is correct, it places the extra *O*-methyl-D-glucose unit near the glyceric acid end of the chain.



Scheme 1 Fragmentation pattern of the portion of the MGP (**2**) molecule near the glyceric acid terminal group

Characteristic, nonreducing-end fragment ions from MGP (2). — In the negative-ion spectrum, MGP (**2**) underwent fragmentation from the nonreducing end to produce ions with molecular weights that are consistent with the structures postulated in Scheme 2, in which the intact sugar ring represents any oligosaccharide sequence in the MGP molecule. The number under each structure indicates the weight increment to be added to each anhydrohexose unit for that fragment attached at C-1. These fragments are designated Types A, B, and C for the following discussion.

A structural feature of the nonreducing end of MGP (**2**), determined by classical methods[4], is the sequence 3-*O*-Me-α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-α-D-Glc*p*-(1→4)-6-*O*-Me-α-D-Glc*p*-(1→4)-6-*O*-Me-α-D-Glc*p*-(1→



Scheme 2. Postulated structures of the negative ions produced by fragmentation from the nonreducing end of MGP (**2**). The formation of such ions has been reported[7].

(see **2**), and a Type-A fragment with five intact hexose units arising from this end of the chain that contains more than three D-glucose units, *i.e.*, MeGlc→Glc→Glc→Glc→MeGlc→A (*m/z* 897) and the equivalent Glc→Glc→Glc→MeGlc→MeGlc→A should not be observed. In the region of this ion, however, one should also observe a series of pentasaccharide ions having an increasing *O*-methylhexose content, Glc→Glc→MeGlc→MeGlc→MeGlc→A (*m/z* 911), Glc→MeGlc→MeGlc→MeGlc→MeGlc→A (*m/z* 925), and MeGlc→MeGlc→MeGlc→MeGlc→MeGlc→A (*m/z* 939). Such a fragment series was observed (Fig. 1, Table II), and the absence of any ion at *m/z* 883 (897 − 14) was a strong indication that only three D-glucose units occur in the sequence flanked by *O*-methylhexose units. An analogous pattern was seen in the C-Type fragments (Table III), but the evidence is equivocal because the possible fragment Glc→Glc→Glc→Glc→MeGlc→C would have the mass 939 and is, therefore, indistinguishable from the A-Type fragment of the same mass. The C-Type fragments are important, however, for distinguishing whether units with the composition $(Glc)_3(MeGlc)_2$ come from the nonreducing end (MeGlc→Glc→Glc→Glc→MeGlc→ and Glc→Glc→Glc→MeGlc→MeGlc→) or from the reducing end [MeGlc→MeGlc→Glc→Glc→Glc→ and MeGlc→(Glc→)MeGlc→Glc→Glc→]. Only the nonreducing end of the polymer can yield the Type-C fragment of this nature, because this requires fragmentation of a *O*-methylhexose at the reducing end of the unit.

This pattern of four A-Type ions followed by four C-Type ions occurred repeatedly throughout the spectrum. At the low-mass end of the spectrum, ions were observed that are characteristic of fragments of Types A, B, and C with single hexose units: Glc→A (*m/z* 221), MeGlc→A (*m/z* 235), MeGlc→B = Glc→C (*m/z* 277), and MeGlc→C (*m/z* 291). All of these ions were seen at about equal intensities. At the pentasaccharide level, ions of Type A containing one to three D-glucose units have about half the intensity of the ion composed of five *O*-methylhexoses. As the sizes of the Type A fragment increase, the relative intensities of the four ions shifted so that, for the undecasaccharides, the ion with three D-glucose units had 2 to 3 times the intensity of the other three ions. For the dodecasaccharide, the ion for $(MeGlc)_{12}$→A was very weak and, in fact, this ion could represent the fragment $(Glc)_3(MeGlc)_9$→B (*m/z* 2171). At the tetradecasaccharide level, this Type A series showed a single ion for $(Glc)_3(MeGlc)_{11}$→A, which is consistent with other indications that the preponderant form of MGP contains eleven 6-*O*-methyl-D-glucose units.

TABLE II

TYPE-A-FRAGMENT IONS FROM MGP (**2**)

Observed ion	*Proposed structures*[a]
221	G-O-CH=CH-O$^-$
235	MG-O-
383	G→G-O-
397	MG→G-O-
411	MG→MG-O-
559	MG→G→G-O-
573	G→MG→MG-O-
587	MG→MG→MG-O-
721	MG→G→G→G-O-
735	G→G→MG→MG-O-
749	G→MG→MG→MG-O-
763	MG→MG→MG→MG-O-
897	MG→G→G→G→MG-O-
911	G→G→MG→MG→MG-O-
925	G→MG→MG→MG→MG-O-
939	MG→MG→MG→MG→MG-O-
1073	G_4MG_3-O-
1087	G_3MG_4-O-
1101	G_1MG_5-O-
1115	MG_6-O-
1249	G_4MG_4-O-
1263	G_3MG_5-O-
1277	G_1MG_6-O-
1291	MG_7-O-
1425	G_4MG_5-O-
1439	G_3MG_6-O-
1453	G_1MG_7-O-
1467	MG_8-O-
1601	G_4MG_6-O-
1615	G_3MG_7-O-
1629	G_1MG_8-O-
1643	MG_9-O-
1777	G_4MG_7-O-
1791	G_3MG_8-O-
1805	G_1MG_9-O-
1819	MG_{10}-O-
1953	G_3MG_8-O-
1967	G_3MG_9-O-
1981	G_1MG_{10}-O-
1995	MG_{11}-O-
2129	G_3MG_{10}-O-
2143	G_2MG_{10}-O-
2157	G_1MG_{11}-O-
2305	G_3MG_{10}-O-
2319	G_2MG_{11}-O-
2481	G_3MG_{11}-O-
2643	G_4MG_{11}-O-
2657	G_3MG_{12}-O-
2819	G_4MG_{12}-O-

[a]Abbreviations as in footnote to Table I.

TABLE III

TYPE-C-FRAGMENT IONS FROM MGP (2)

Observed ion	*Proposed structures*[a]
	CH=CH-OMe | G-O-C=CH-O$^-$
277	
291	MG-O-
439	G→G-O-
453	G→MG-O-
467	MG→MG-O-
601	G→G→G-O-
615	G→G→MG-O-
629	G→MG→MG-O-
643	MG→MG→MG-O-
777	MG→G→G→G-O-
	or G→G→G→MG-O-
791	G→G→MG→MG-O-
805	G→MG→MG→MG-O-
819	MG→MG→MG→MG-O-
953	G_3MG_2-O-
967	G_2MG_3-O-
981	G_1MG_4-O-
995	MG_5-O-
1129	G_3MG_3-O-
1143	G_2MG_4-O-
1157	G_1MG_5-O-
1171	MG_6-O-
1305	G_3MG_4-O-
1319	G_2MG_5-O-
1333	G_1MG_6-O-
1347	MG_7-O-
1481	G_3MG_5-O-
1495	G_2MG_6-O-
1509	G_1MG_7-O-
1523	MG_8-O-
1657	G_3MG_6-O-
1671	G_2MG_7-O-
1685	G_1MG_8-O-
1699	MG_9-O-
1833	G_3MG_7-O-
1847	G_2MG_8-O-
1861	G_1MG_9-O-
1875	MG_{10}-O-
2009	G_3MG_8-O-
2013	G_2MG_9-O-
2185	G_3MG_9-O-
2361	G_3MG_{10}-O-

[a]Abbreviations as in footnote to Table I.

The next larger ion of the Type-A series contained all of the *O*-methyl-D-glucose units in MGP plus three D-glucose units; it must represent the fragment 3MeGlc→Glc→Glc→Glc→$(6MeGlc)_{11}$→A. Finally, ions were observed that contain additional D-glucose units, and these must arise from fragmentation near the

glyceric acid end of the chain. The ion at *m/z* 2819 [$(Glc)_4(MeGlc)_{12}\rightarrow A$] must represent the sequence 3MeGlc→Glc→Glc→Glc→$(6MeGlc)_{11}$→Glc→A, and the ion at 2643 [$(Glc)_4(MeGlc)_{11}\rightarrow A$] could arise by loss of the 3-*O*-methyl-D-glucose unit. An ion at *m/z* 2833 would fit a fragment $(Glc)_3(MeGlc)_{13}\rightarrow A$, and it could arise from the minor MGP component having thirteen *O*-methyl-D-glucose units. The alternative explanation for this ion as coming from a Type-C fragmentation is not feasible because it would require $(Glc)_7(MeGlc)_9\rightarrow C$, and such an ion cannot be derived from MGP.

Type-C-sequence ions follow the same general pattern as those of Type A. As already mentioned, however, some ambiguity in the interpretation does arise because the less methylated Type-C ions overlap in molecular weight with the completely methylated Type-B ions. As expected from the structure of MGP, the largest of the Type-C ions contains fewer *O*-methyl-D-glucose units than does the Type-A ion, because one *O*-methyl-D-glucose unit is partially degraded during the fragmentation process that produces the first type. For some reason, the series ends at $(Glc)_3(MeGlc)_{10}\rightarrow C$, even though a $(Glc)_3(MeGlc)_{11}\rightarrow C$ ion is possible. This may be a consequence of the branching at the twelfth *O*-methyl-D-glucose unit.

Size and composition of MGLP (1) isomers. — It was demonstrated previously[5] that the small neutral acyl groups were located at the 3-*O*-methyl-D-glucosyl end-group of the chain, that octanoic acid esterified the glyceric acid unit, and that the succinate groups were located at the glyceric acid end-group. MGLP-I lacks succinate residues whereas the more acidic forms, MGLP-II, -III, and -IV, contain one, two, and three succinate residues, respectively. Comparison of the spectrum obtained from MGLP-I (**1**) (Fig. 2) with that from MGP (**2**) (Fig. 1) showed several major ions in the region at *m/z* 3800 (3794, 3822, 3836, and 3850). In Table IV is

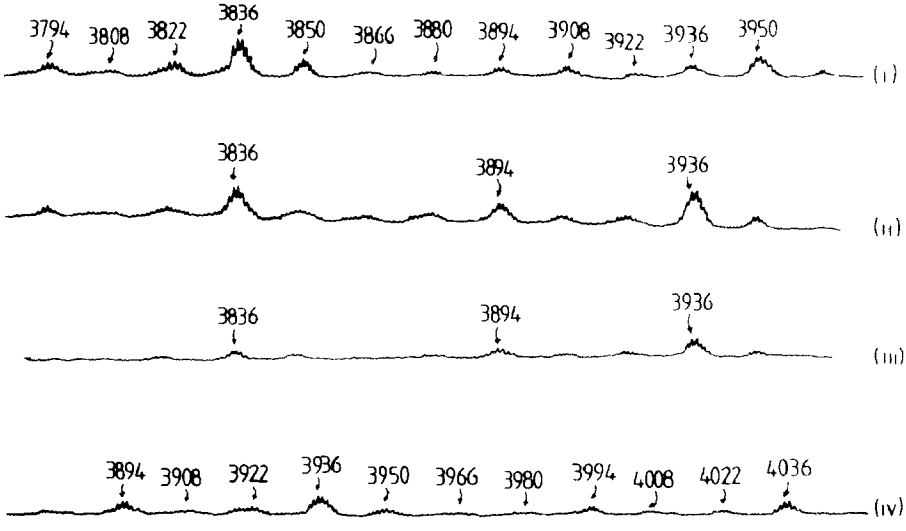


Fig. 2. Negative ion F.a b mass spectra of MGLP (**1**). Only the high mass ends of the spectra are reproduced which show the molecular ions for the isomers of MGLP-I (i), -II (ii), -III (iii) and -IV (iv)

TABLE IV

EXPECTED MOLECULAR IONS FOR MGLP (**1**)

Neutral acyl composition[a]					*Calc. MGLP* M_r *values*[b]			
Ac	*Pr*	*iBut*	*Oct*	*Sum*	*I*	*II*	*III*	*IV*
1	0	1	1	238	3752	3852	3952	4052
3	0	0	1	252	3766	3866	3966	4066
2	1	0	1	266	3780	3880	3980	4080
2	0	1	1	280	3794	3894	3994	4094
1	1	1	1	294	3808	3908	4008	4108
3	1	0	1	308	3822	3922	4022	4122
3	0	1	1	322	3836	3936	4036	4136
2	1	1	1	336	3850	3950	4050	4150
3	1	1	1	378	3892	3992	4092	4192

[a]Abbreviations: Ac, acetyl = 42; Pr, propionyl = 56; iBut, isobutyryl = 70; and Oct, octanoyl = 126.
[b]Succinyl = 100.

a listing of the calculated M_r values of MGLP-I isomers having different neutral acyl group compositions. There is correspondence between several of the observed ions (see Fig. 2) and the calculated formulas, the major ion at *m/z* 3836 agreeing with an MGLP-I containing three acetyl, one isobutyryl, and one octanoyl groups. Notably absent was an ion at *m/z* 3892 for the fully acylated form of MGLP-I, and this observation, plus the presence of other incompletely acylated forms, agrees with previous analyses revealing a variable and low content of propionyl and isobutyryl groups[8].

MGLP-II and -III gave very similar spectra in this high-mass region, but they differ from MGLP-I in giving several signals displaced to higher mass by 100 mass units (*m/z* 3894, 3936, and 3950). This is the expected consequence of the addition of a single succinyl group to MGLP-I isomers (Table IV). It was unexpected, however, that MGLP-II (one succinyl group) and MGLP-III (two succinyl groups) would give similar patterns, and the result here reported suggests that the molecules have undergone loss of succinyl group(s) during isolation or storage[8]. This would also account for the succinyl-free ion at *m/z* 3836 from both isomers. We have no explanation for the ion at *m/z* 3950 from MGLP-I*.

The spectrum of MGLP-IV showed ions for isomers with one (*m/z* 3894, 3936, and 3950) and two (*m/z* 3994 and 4036) succinyl groups. Again this result suggests that the sample must have lost succinyl group(s) since its isolation.

At least two important conclusions emerge from this analysis regarding MGLP structure (**1**). Firstly, several but not all possible isomers are found in the MGLP-I preparation, and one isomer preponderates (*m/z* 3836). Secondly, in the

*A plausible explanation offered by one of the referees is that some of the succinyl groups might exist as a methyl ester, thus reducing the charge on the MGLP molecule, and causing it to be eluted at a lower salt concentration and leading to a mixing of the MGLP isomers.

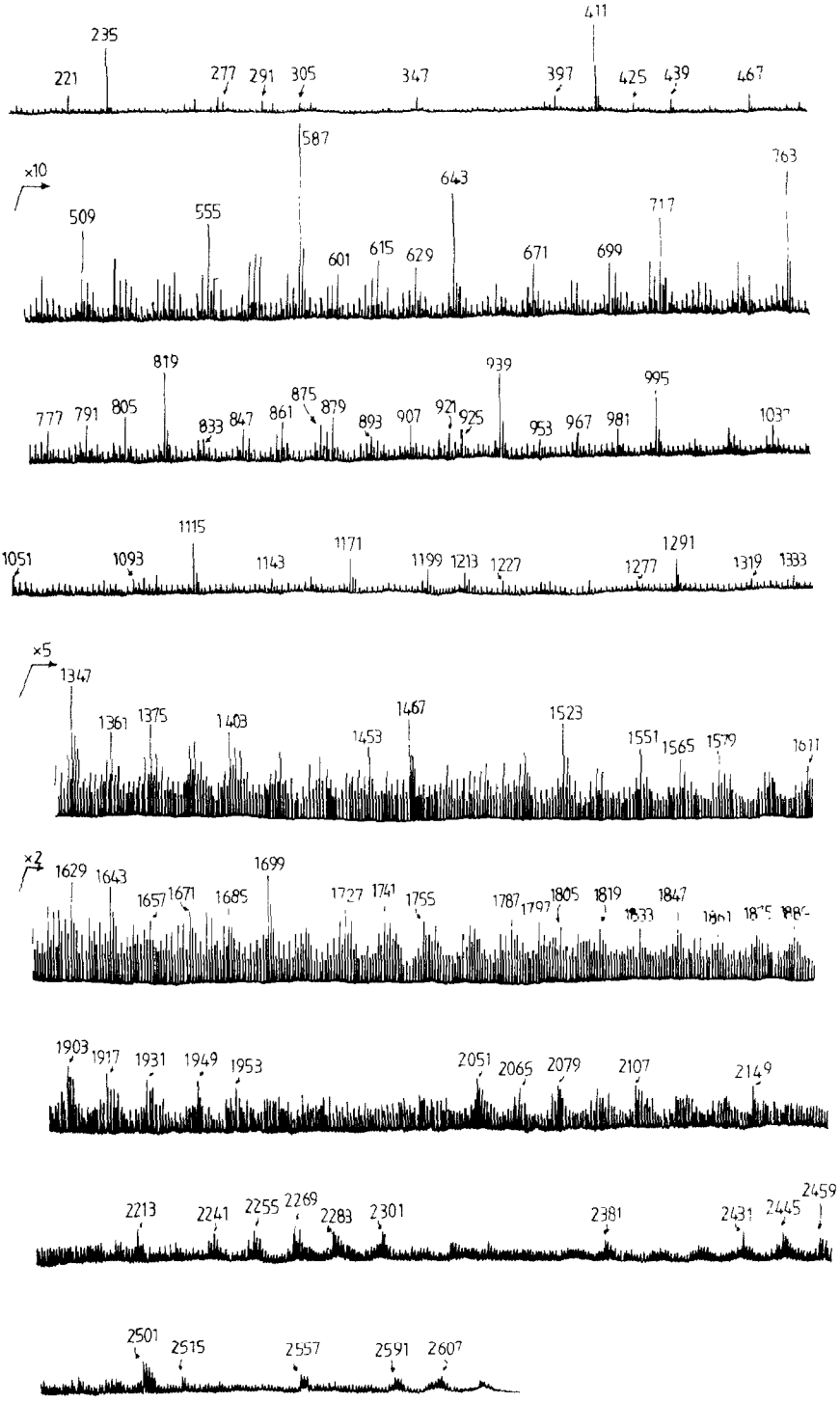


Fig. 3. Negative ion F.a.b. mass spectrum of MGLP-I. Only the spectrum below m/z 2620 is reproduced, although higher-mass ions were observed (see Fig. 2).

succinylated forms, the same isomer preponderates (*i.e.*, *m/z* 3836, 3936, and 4036). The second most abundant ion-species were seen at *m/z* 3794, 3894, and 3994, and these correspond to the form having two acetyl, one isobutyryl, and one octanoyl groups.

Location of neutral acyl groups on MGLP-I. — Type-A-fragment ions were observed in the MGLP spectra that support the previous conclusion[5] that the acetyl, propionyl, and isobutyryl groups are at the nonreducing end of the chain. Several ions in the spectrum of MGLP-I (Fig. 3) that were lacking in the spectrum of MGP were observed at *m/z* 305, 347, 425, 509, 671, 699, 833, 847, 861, 875, and 917, and in Table V are listed possible structures for these ions. In many instances, ambiguity exists because more than one structure can be drawn for a particular ion. There are reasons, however, that rule out all but one or two structures for each ion. Firstly, any ion having two propionyl or isobutyryl groups, or one of each, is un-

TABLE V

SMALL-TYPE-A FRAGMENT FROM MGLP-I

Observed ion	*Possible structures*[a]	
305	G-O-CH=CH-O$^-$	+ 2 C_2
	*MG-O-	+ C_4
347	G-O-	+ 3 C_2
	G-O-	+ C_3 + C_4
	*MG-O-	+ C_2 + C_4
	MG-O-	+ 2 C_3
425	*G→G-O-	+ C_2
509	G→G-O-	+ C_3 + C_4
	G→G-O-	+ 3 C_2
	*MG→G-O-	+ C_2 + C_4
	MG→G-O-	+ 2 C_3
671	G→G→G-O-	+ C_3 + C_4
	*G→G→G-O-	+ 3 C_2
	*MG→G→G-O-	+ C_2 + C_4
	MG→G→G-O-	+ 2 C_3
699	*G→G→G-O-	+ 2 C_2 + C_4
	G→G→G-O-	+ C_2 + 2 C_3
	*MG→G→G-O-	+ 2 C_2 + C_3
	MG→G→G-O-	+ 2 C_4
833[b]	*MG→G→G→G-O-	+ C_2 + C_4
847[c]	*MG→G→G→G-O-	+ 3 C_2
861[d]	*MG→G→G→G-O-	+ 2 C_2 + C_3
875[e]	*MG→G→G→G-O-	+2C_2 + C_4
917[f]	*MG→G→G→G-O-	+3C_2 + C_4

*Structures considered more probable. [a]Abbreviations as in footnote to Table I. [b]Other ions of this series are seen at *m/z* 1361 (+3 MG), 1889 (+6 MG), 2065 (+7 MG), and 2241 (+8 MG). [c]Other ions of this series are seen at *m/z* 1199 (+2 MG), 1375 (+3 MG), 1551 (+4 MG), 1727 (+5 MG), 1903 (+6 MG), 2079 (+7 MG), 2255 (+8 MG), 2431 (+9 MG), and 2607 (+10 MG). [d]Other ions of this series are seen at *m/z* 1037 (+1 MG), 1213 (+2 MG), 1565 (+4 MG), 1741 (+5 MG), 1917 (+6 MG), 2269 (+8 MG), and 2445 (+9 MG). [e]Other ions of this series are seen at *m/z* 1051 (+1 MG), 1227 (+2 MG), 1403 (+3 MG), 1579 (+4 MG), 1755 (+5 MG), 1931 (+6 MG), 2107 (+7 MG), 2283 (+8 MG), and 2459 (+9 MG). [f]Other ions of this series are seen at *m/z* 1093 (+1 MG), 1797 (+5 MG), 2149 (+7 MG), and 2501 (+9 MG).

likely because the MGLP molecules (**1**) usually contain less than one mol/mol of these two fatty acids. Secondly, an ion that requires more than one acyl group per D-glucose unit is ruled out by previous studies[5,9]. Applying these criteria, stars have been placed in front of the possible structures considered more probable for the observed ions. Several important conclusions can be drawn from these interpretations: (a) The isobutyryl group appears most often in association with *O*-methyl-D-glucose and may occur in MGLP (**1**) exclusively linked to the 3-*O*-methyl-D-glucose unit, as suggested by earlier studies[5]. (b) The other acyl group on this methylated sugar is usually an acetyl group.

The MGLP spectra also support the conclusion that the A-Type fragments in the MGP spectrum derive from the nonreducing end of the chain. In the MGLP spectra, those Type-A fragments containing D-glucose are greatly reduced in intensity relative to the completely methylated fragments of each series (for example, *m/z* 383, 397, 559, 573, 721, 735, 749, 897, 911, 925, 1073, 1087, 1101, 1249, 1263, 1277, 1425, 1439, 1453, 1601, and 1615). This is expected if the D-glucose-containing fragments in MGLP are acylated and, thus, shifted in their molecular weights. Positive evidence for such a shift is seen in the new ions appearing at higher mass in the MGLP-I spectrum, listed in the footnotes to Table V. These ions can be accounted for as homologs of the lower-mass ions in Table V that are formed by addition of increasing numbers of *O*-methyl-D-glucose units.

Any group of small acyl substituents having a molecular-weight increment of 126 (three acetyl or one propionyl plus one isobutyryl groups) could also represent an octanoyl group. Previous studies, however, have provided strong evidence that this large acyl substituent is at O-3 of the glyceric acid unit[5]. Consistent with this, we observed a series of fragments (Table VI) that agree with the structure (D-glucose)$_x$→glyceric acid→octanoyl, where x = 1–4. This assignment is confirmed by the observation that the nonacylated glyceric acid-containing fragments in the MGP spectrum (at *m/z* 267, 429, 591, and 753) were not observed in the MGLP-I spectrum.

Other fragments from MGLP-III and -IV. — The spectra of these two preparations showed ions at *m/z* 895, 893, 733, 731, 571, 569, 409, 407, 255, 253, 171, and 153 that were not present in the other samples (Fig. 4). These are related in a way

TABLE VI

REDUCING-END FRAGMENTS FROM MGLP (**1**)

Observed ion	*Proposed structure*[a]
393	G→Ga-Oct
555	G→G→Ga-Oct
717	G→G→G→Ga-Oct
879	G→G→G→Ga-Oct ↑ G

[a]Abbreviations are as in footnotes to Tables I and IV.

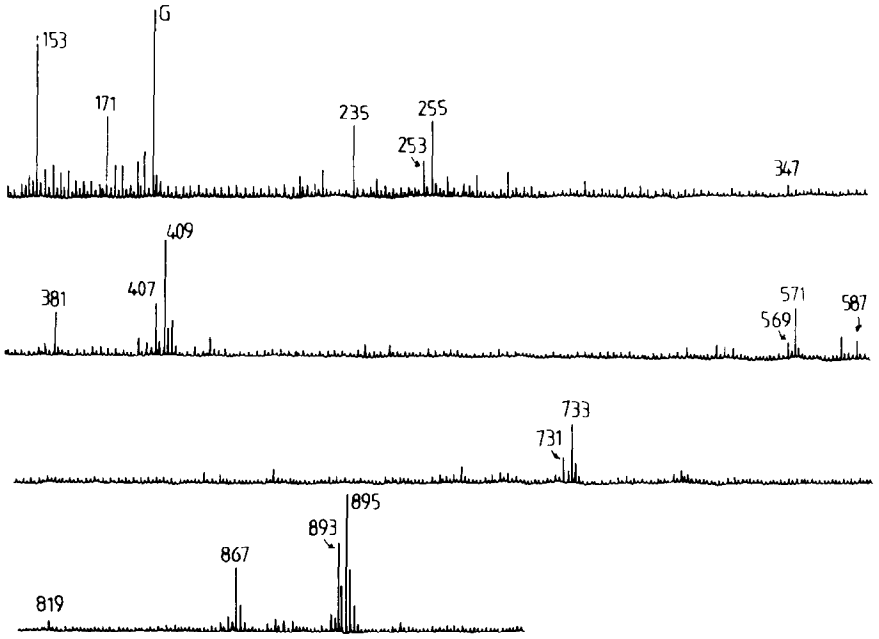


Fig. 4. Negative-ion F.a.b. mass spectrum of MGLP-III. Only the low-mass end of the spectrum is reproduced to show the ions for the glycophospholipid contaminant (**3**).

that suggests that they could represent sequential loss of three hexoses units from a molecule of M_r 895 (or 893) to leave a residue of M_r 409 (or 407). Because these signals showed a characteristic pattern of doublets separated by 2 mass units, we concluded that they might be isomers that differ in one double bond and, therefore, might contain a long-chain fatty acid. In fact, the ions at *m/z* 253 and 255 are correct for hexadecenoic and hexadecanoic acids, that at *m/z* 171 for glycerol phosphate, and that at *m/z* 153 for loss of water from 171. From a B/E linked scan[10], with the ion at *m/z* 895 selected as the parent ion, daughter ions were observed at *m/z* 733, 656, 570, 409, and 255 (data not shown), which confirms the common origin of these ions.

An attempt to fractionate the two components by binding the MGLP to a palmityl silicate affinity resin[11] according to Kari and Gray[12] failed. To facilitate separation, a sample of MGLP-III was deacylated with 0.1M sodium hydroxide and, after neutralization, the product was fractionated on Bio-Gel P-4 column (0.8 × 100 cm) by elution with 0.1M ammonium acetate. The MGP was excluded by the gel and appeared in the void volume of the column, whereas the deacylated glycophospholipid was included. The carbohydrate-containing material of the included component was collected and desalted by passage through a Bio-Gel P-2 column (0.6 × 80 cm) in water. Analysis of the resulting material by F.a.b.m.s. gave a major ion at *m/z* 657, which is consistent with removal of the fatty acid from the glycophospholipid. A second characteristic ion was seen at *m/z* 495 (657 −

162), apparently resulting from loss of a hexose unit. The spectrum was free of any ions related to MGP.

The nature of the hexose component was determined by paper chromatography of an acid hydrolyzate of the deacylated glycophospholipid. Treatment of the chromatogram with a silver nitrate–sodium hydroxide reagent revealed two components, one of which migrated with D-mannose (R_{Glc} 1.22), and one with *myo*-inositol (R_{Glc} 0.31). These are expected for phosphatidylinositol di-D-mannoside, which is known to occur in *Mycobacterium smegmatis*[13]. Thus, a probable structure for the contaminant of these MGLP samples is the lysoderivative of this phospholipid (**3**). The occurrence of this compound only with the more acidic MGLP isomers is consistent with its strongly acidic phosphate group.

Conclusion. — Although polysaccharides may have uniform structures, they usually vary in chain length or in numbers of repeating units. The *O*-methyl-D-glucose polysaccharide from *M. smegmatis* is unusual in that it has a precisely defined structure that is characteristic of molecules made by some template mechanism. For MGP, however, the structural definition probably results from the specificity of biosynthetic enzymes and the conformational properties of MGP, as modulated by its novel ability to interact strongly with long-chain fatty acids[14,15]. Even for MGP, however, F.a.b.m.s. reveals a slight heterogeneity in the degree of methylation, and perhaps in the degree of polymerization, although the latter point is equivocal. In other respects, the spectra fully document the structure of MGP as previously reported[3].

With regard to the acylated forms, the mass spectra revealed extensive heterogeneity in short-chain acyl group substitution, the preponderant form of MGLP-I containing three acetyl and one isobutyryl groups, and a slightly less abundant form containing three acetyl and one propionyl groups. Whether this heterogeneity is natural or an artifact of the manipulations required for isolation is unclear. No evidence was obtained for heterogeneity in acylation of the glyceric acid by octanoic acid.

Analyses of the succinylated forms of MGLP were less satisfactory, and the results indicate that extensive loss of succinyl groups had occurred since the molecules were isolated. This may have happened during storage of the samples or during lyophilization of the fractions to remove the ammonium hydrogen carbonate used in elution from the DEAE–Sephadex column[6,8]. Regardless, the spectra did show that the major form of MGLP-I is also observed in its succinylated states in MGLP-II and MGLP-III. This is consistent with a biosynthetic pathway MGLP-I → MGLP-II → MGLP-III. The actual acyl composition of MGLP (**1**) probably reflects the availability of acylcoenzyme A donors in the cell that, in turn, may reflect the nutritional state of the culture[16].

An unexpected bonus of this study was the detection of lyso(monoacyl)phosphatidylinositol di-D-mannoside[13] as a contaminant of the more acidic MGLP preparations. The presence of this derivative may reflect the ability of the polysaccharides to bind long-chain fatty acids[14,15].

Although we show, here, that F.a.b. mass-spectral data of the underivatized polysaccharide can provide documentation for all sequence features of MGP structure, it does not follow necessarily that one can derive the structure from the spectra. This is because many fragments result from two-bond cleavages and, consequently, it is difficult to distinguish terminal- from internal-sequence ions. In the case of MGP, the presence of glyceric acid identifies a reducing-end fragment and, as a general procedure, conversion of reducing polysaccharides to the (pentafluoro)benzyloxime derivative will facilitate recognition of ions that arise from the reducing end of the moecule[17]. Permethylation of the branched, high-mannose oligosaccharides from glycoproteins has also been used to facilitate the identification of sequence ions[18].

ACKNOWLEDGMENT

This study was carried out in the laboratory of Professor H. R. Morris, Imperial College.

REFERENCES

1 M. BARBER, R. S. BORDOLI, R. D. SEDGWICK, AND A. N. TYLER, *Chem. Commun.*, (1981) 325–327.
2 H. R. MORRIS, A. DELL, A. T. ETIENNE, M. JUDKINS, R. A. MCDOWELL, M. PANICO, AND G. W. TAYLOR, *Pure Appl. Chem.*, 54 (1982) 267–279.
3 L. S. FORSBERG, A. DELL, D. J. WALTON, AND C. E. BALLOU, *J. Biol. Chem.*, 257 (1982) 3555–3563.
4 C. E. BALLOU, *Acc. Chem. Res.*, 1 (1968) 366–373.
5 W. L. SMITH AND C. E. BALLOU, *J. Biol. Chem.*, 248 (1973) 7118–7125.
6 G. R. GRAY AND C. E. BALLOU, *Methods Enzymol.*, 35 (1975) 90–95.
7 A. DELL AND C. E. BALLOU, *Biomed. Mass Spectrom.*, 10 (1983) 50–56.
8 J. M. KELLER AND C. E. BALLOU, *J. Biol. Chem.*, 243 (1968) 2905–2910.
9 G. R. GRAY AND C. E. BALLOU, *J. Biol. Chem.*, 247 (1972) 8129–8135.
10 I. HOWE, D. H. WILLIAMS, AND R. D. BOWEN, *Mass Spectrometry*, 2nd edn. McGraw–Hill, New York, 1981, p. 13.
11 O. HINDSGAUL AND C. E. BALLOU, unpublished results.
12 B. E. KARI AND G. R. GRAY, *J. Biol. Chem.*, 254 (1979) 3354–3357.
13 Y. C. LEE AND C. E. BALLOU, *Biochemistry*, 4 (1965) 1395–1404.
14 K. BLOCH, *Adv. Enzymol.*, 45 (1977) 1–84.
15 C. E. BALLOU, *Pure Appl. Chem.*, 53 (1981) 107–112.
16 K. NARUMI, J. M. KELLER, AND C. E. BALLOU, *Biochem. J.*, 132 (1973) 329–340.
17 A. DELL, J. E. OATES, H. R. MORRIS, AND H. EGGE, *Int. J. Mass Spectrom. Ion Phys.*, 46 (1983) 415.
18 H. EGGE, J. C. MICHALSKI, AND G. STRECKER, *Arch. Biochem. Biophys.*, 213 (1982) 318–326.