Carbohydrate Research, 249 (1993) 383–398
Elsevier Science Publishers B.V., Amsterdam

Major structural features of the cell wall arabinogalactans of *Mycobacterium, Rhodococcus*, and *Nocardia* spp. †

Mamadou Daffe [1], Michael McNeil * and Patrick J. Brennan

Department of Microbiology, Colorado State University, Fort Collins, Colorado 80523 (USA)

(Received October 23rd, 1992; accepted in revised form February 19th, 1993)

ABSTRACT

The cell wall arabinogalactans of strains of *Mycobacterium, Rhodococcus*, and *Nocardia* were per-*O*-methylated, partially hydrolyzed with acid, and the resulting oligosaccharides were reduced and per-*O*-ethylated to yield per-*O*-alkylated oligoglycosyl alditol fragments. Analyses of these fragments by gas chromatography–mass spectrometry and of the intact solubilized polysaccharides by ^{1}H and ^{13}C NMR revealed the major structural features of the different arabinogalactans from representatives of the different genera. All of the mycobacterial products contained a homogalactan segment of alternating 5-linked α-galactofuranosyl (Gal*f*) and 6-linked β-Gal*f* residues. The arabinan segment consisted of three major domains, linear 5-linked α-arabinofuranosyl (Ara*f*) residues and branched (3 → 5)-linked Ara*f* units substituted with either 5-linked Ara*f* or the disaccharide β-Ara*f*-(1 → 2)-α-Ara*f* at both branched positions. The recognition of these features in in vivo grown *Mycobacterium leprae* is an important development. The arabinan from strains of *Nocardia* contains a nonreducing-end motif composed of the linear trisaccharide, β-Ara*f*-(1 → 2)-α-Ara*f*-(1 → 5)-Ara*f*, attached to linear 5-linked α-Ara*f* units. The galactan segment of the arabinogalactan of *Nocardia* sp. is composed of linear 5-linked β-Gal*f* units substituted in part at O-6 with terminal β-glucosyl units. The two representative strains of *Rhodococcus* also differed in the composition of the galactan moiety; in addition to the 5-linked Gal*f*, 2- and 3-linked β-Gal*f* units are present. The reducing end of the galactans, and therefore, apparently, of the entire arabinogalactans from all species from all genera, are apparently composed of the unit, rhamnosyl-(1 → 3)-*N*-acetyl-glucosamine, which, in turn, is apparently attached to peptidoglycan via phosphodiester linkage.

INTRODUCTION

The most distinguishing feature of actinomycetes, such as most members of the *Mycobacterium, Rhodococcus*, and *Nocardia* genera, is a chemotype IV cell wall marked by major quantities of arabinogalactan, apparently covalently attached at

* Corresponding author.

[1] Permanent address: Department des Glycoconjugues et Biomembranes, L.P.T.F. du C.N.R.S., 118, Route de Narbonne, Toulouse Cedex, France.

† Abbreviations used: BCG, Bacille Calmetter Guerin; ATCC, American type culture collection; TMC, Trudeau mycobacterial collection.

0008-6215/93/$06.00 © 1993 – Elsevier Science Publishers B.V. All rights reserved

one end to peptidoglycan and at the other to mycolic acids or mycolic-like acids. Arabinogalactan, which represents ~35% of the mass of cell walls from members of these genera, is thought to be their most dominant immunogen[1]. Unlike most such bacterial polysaccharides, arabinogalactan is devoid of a repeating unit. Although the polysaccharide has been the subject of intensive work, little was known about the arrangement of the variously linked arabinosyl and galactosyl units in their respective portions of the heteropolysaccharide. Recently[2], we addressed this fundamental question in a study of the arabinogalactan of a strain of *Mycobacterium tuberculosis* using methylation, partial acid hydrolysis, reduction, ethylation, and analysis of the resulting array of oligoglycosyl alditol fragments by GL–MS and HPLC. These analyses enabled the recognition of several novel structural motifs, which, in turn, allowed the exposition of revolutionary, although incomplete, structural models[2]. With the application of NMR analyses, it was possible to unambiguously assign the anomeric configuration to the predominant structural motifs and, accordingly, to the complex molecule[2]. Consequently, ^{13}C NMR appeared to be an excellent investigative tool, both as an aid in structural determination and in fingerprinting for taxonomic purposes. In the present study, we applied both NMR and GLC–MS analyses to the study of arabinogalactans from *Mycobacterium leprae* and *Mycobacterium avium*, the two other major mycobacterial pathogens, a strain of the attenuated bovine tubercle bacillus, bacillus of Calmette Guerin (BCG), a saprophytic *Mycobacterium* species (*Mycobacterium smegmatis*) and representative strains of related genera, *Rhodococcus* and *Nocardia*.

RESULTS AND DISCUSSION

Glycosyl linkage analysis.—Intact cell walls were per-*O*-methylated and small aliquots hydrolyzed, reduced, acetylated, and analyzed by GLC–MS to define glycosyl linkage compositions (Table I). The methylation data point to a close structural similarity in the arabinogalactans from the different mycobacterial species. A notable exception was *M. avium* serovar 4, the cell walls of which contained significant amounts (13%) of 4-linked mannopyranose (Man*p*). However, when the soluble cell wall arabinogalactan, released by alkaline treatment[2], and purified by DEAE-Sephadex column chromatography, was analyzed, it was shown to contain galactose and arabinose and no mannose, suggesting the presence, in the cell walls of some strains/serotypes of *M. avium*, of an unknown 4-linked, mannose-containing polysaccharide independent of the arabinogalactan. Incidentally, (1 → 4)-linked β-D-mannans have been described only in some red algae and higher plants[3]; thus, the possibility of bacterial contamination of the preparation is unlikely. In contrast, the cell walls of *M. avium* serovar 25 were devoid of such a polysaccharide (Table I).

The glycosyl linkage composition of the cell wall arabinogalactans, notably the galactan portion, of species of *Nocardia* and *Rhodococcus* were significantly

TABLE I

Glycosyl-linkage analysis of arabinogalactans from various species of *Mycobacterium*, *Nocardia*, and *Rhodococcus*

Glycosyl residue	*M. tuberculosis*	BCG	*M. avium* serovar 4	*M. avium* serovar 25	*M. smegmatis*	*M. leprae*	*N. asteroides*	*N. brasiliensis*	*R. equi*	*R. rhodochrous*
	(mol%)									
t-Ara*f*	11.0	13.2	7.5	11.0	12.0	14.6	8.5	9.3	14.2	7.8
2-Ara*f*	11.0	10.7	7.5	12.0	12.0	12.1	7.4	6.7	3.3	5.7
5-Ara*f*	37.0	38.3	29.0	36.0	38.0	39.7	44.0	38.2	37.2	36.7
3,5-Ara*f*	11.0	10.5	8.0	12.0	13.0	11.0	6.7	7.8	12.1	8.0
2,5-Ara*f*		0.4	tr [a]		tr		1.6	6.0	2.3	
5-Gal*f*	17.0	14.3	21.0	12.0	10.0	11.0	8.9	10.8	10.4	8.9
3-Gal*f*									6.5	
2-Gal*f*										11.0
6-Gal*f*	9.0	7.8	6.0	7.0	7.0	4.7	tr	tr	2.6	
5,6-Gal	2.0	2.5	3.0	4.0	2.5	2.3	12.0	7.0	1.0	1.3
t-Gal*f*	1.0	1.4	0.6	1.0	0.5	1.3	tr	tr	1.8	tr
4-Rha*p*	1.0	1.0	0.5	1.0	1.0	1.3	0.8	1.0	0.4	1.0
t-Glc*p*							8.0	10.0		
t-Man*p*	0.8	tr	2.5	3.0	2.0	0.4	4.0	3.0	3.3	7.8
4-Man*p*			13.0							
6-Man*p*	0.5		0.7	tr	tr	tr	tr	tr	2.6	3.4
2,6-Man*p*			tr	1.0	2.0	tr	tr	tr	0.9	3.2

[a] tr, Trace.

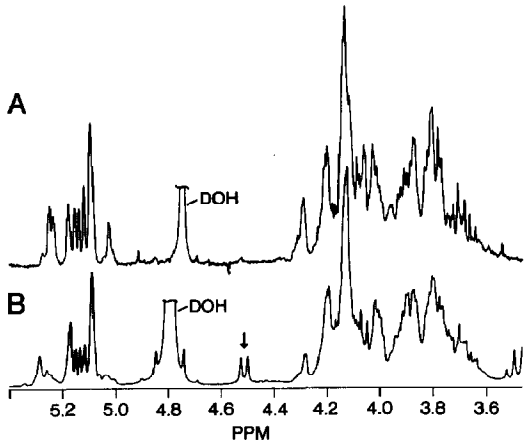

Fig. 1. Comparison of the ^1H NMR spectra of the base solubilized cell wall arabinogalactans of *M. leprae* (A) and *N. asteroides* (B). The arrow indicates the signal assigned to the H-1 of the β-glucosyl residue.

different from those of mycobacteria (Table I). The two *Nocardia* species contained additional terminal-glucopyranose (t-Glc*p*) (8–10%), and were devoid of 6-linked galactofuranose (Gal*f*), which was apparently replaced by (5 → 6)-linked Gal*f*. *Rhodococcus equi* contained 3-linked Gal*f* in addition to 5-linked Gal*f*, and 6-linked Gal*f*, whereas *Rhodococcus rhodochrous* contained 2-linked Gal*f* instead of 3-linked Gal*f*, and was devoid of 6-linked Gal*f*.

Gross comparison of the structure of the various arabinogalactans.—The ^1H NMR spectra of the soluble arabinogalactans from *M. leprae* (Fig. 1A), *M. smegmatis, M. bovis, M. avium*, and *R. equi* were similar to one another and to that of the product from *M. tuberculosis*. The exception was that from *N. asteroides* (Fig. 1B). This latter spectrum confirmed the presence of glucosyl residues, and their analyses allowed the attribution of the β configuration to them [δ 4.5, ($J_{1,2}$ 7–8 Hz); Fig. 1B]. The other anomeric protons of the various polysaccharides gave signals at δ 5.0–5.3, consistent with the presence of α- and/or β-furanosyl residues and/or α-pyranosyl residues.

The ^{13}C NMR spectra of the various mycobacterial arabinogalactans were similar to one another (Fig. 2). Based on the previous assignments of the signals of the arabinogalactan of *M. tuberculosis*[2], it was possible to identify resonances attributable to the terminal nonreducing end of the arabinan chains, i.e., terminal (t)-β-Ara*f* (C-1 at δ 101–102) linked to position 2 of an α-Ara*f* (C-1 at δ 106–108), this disaccharide being attached to both positions 3 and 5 of an α-Ara*f* residue[2]. The furanoid nature of the remaining Ara and Gal residues was also obvious from the chemical shift of the anomeric carbon resonances (δ 108–110).

Further assignments of the ^{13}C NMR resonances are presented in Table II. They are based on glycosyl linkage composition (Table I), ^{13}C–^1H correlation experiments, ^{13}C NMR spectroscopy in the DEPT mode, and on the literature.

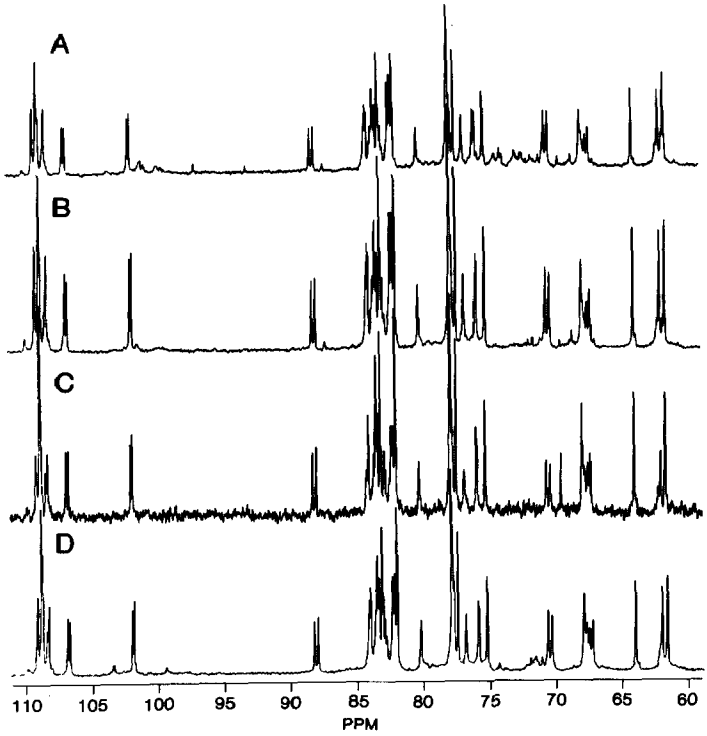

Fig. 2. Comparison of the ^{13}C NMR spectra of the base solubilized cell wall arabinogalactans of *M. tuberculosis* (A), *M. bovis* (B), *M. leprae* (C), and *M. smegmatis* (D). Assignments of the major signals are presented in Table II.

Thus, the methods used for the attribution of the C-1 resonances of t-β-Ara*f* and 2-linked α-Ara*f*, of C-2 of this latter arabinosyl residue, and of C-5 of the t-β-Ara*f*, have been described[2]. For purposes of assigning the remaining resonances, some principles were applied and a few conclusions reached as follows: (*i*) the DEPT experiments were used to differentiate nonglycosylated C-5 from glycosylated C-6 (of Gal*f* residues) which exhibited resonances in the same range of chemical shifts; (*ii*) the assignment of the resonances of the additional glucosyl residues of the arabinogalactan of *Nocardia* spp. was based on the literature[4]; (*iii*) the chemical shift of a resonance from a particular carbon atom of a sugar residue involved in glycosidic linkage was estimated to be shifted downfield by 4–10 ppm[5]; (*iv*) in the case of furanosyl residues, information from the literature on 2-, 3-, and 5-linked Ara*f* residues[6–8] and 5-linked Gal*f* residues[6] was used; (*v*) for 2- and 3-linked Gal*f* units, information on which resonances was not available in the literature, the upfield shift of the C-1 resonance of 2-linked Gal*f* was assigned by analogy to that of 2-linked Ara*f*[2,8]. The C-6 resonance of these sugars (δ 64.0) was similar to that of methyl-β-Gal*f* (δ 63.9)[4]. This particular signal was observed in the case of the spectrum of the product from the *Rhodococcus* sp. (Fig. 3B) but not

TABLE II

Assignments of the main ^{13}C NMR resonances of the arabinogalactans of the various actinomycetes [a]

δ (ppm)	Assignments
109.0	C-1 of 6-β-Gal*f*
108.7–08.5	C-1 of 5-α-Ara*f*, (3 → 5)-α-Ara*f*, 3-β-Gal*f*
108.3	C-1 of (5 → 6)-β-Gal
108.1	C-1 of 5-β-Gal*f*
106.8–106.6	C-1 of 2-α-Ara*f*, 2-β-Gal*f*
104.0	C-1 of β-Glc*p*
101.9–101.8	C-1 of β-Ara*f*
88.2–87.9	C-2 of 2-α-Ara*f*
87.5	C-2 of 2-β-Gal*f*
85.1	C-2 of 3-β-Gal*f*
	C-2 and C-3 of (3 → 5)-Ara*f*,
84.0–82.0	C-2 of β-Ara*f*, 5-α-Ara*f*, 5-β-Gal*f*, 6-β-Gal*f*, (5 → 6)-β-Gal*f*, 3-β-Gal*f*
	C-4 of β-Ara*f*, 2-α-Ara*f*, 5-α-Ara*f*, (3 → 5)-α-Ara*f*, 5-β-Gal*f*, 6-β-Gal*f*, 2-β-Gal*f*, 3-β-Gal*f*, (5 → 6)-β-Gal*f*
78.0–75.0	C-3 of β-Ara*f*, 2-α-Ara*f*, 5-α-Ara*f*, 5-β-Gal*f*, 6-β-Gal*f*, (5 → 6)-β-Gal*f*, 2-β-Gal*f*, β-Glc*p*
	C-5 of 5-β-Gal*f*, (5 → 6)-β-Gal*f*
74.2	C-2 of β-Glc*p*
70.8	C-5 of β-Glc*p*
70.7–69.7	C-5 of 6-β-Gal*f*, 2-β-Gal*f*, 3-β-Gal*f*
	C-6 of 6-β-Gal*f*, (5 → 6)-β-Gal*f*
68.0–67.0	C-5 of 5-α-Ara*f*, (3 → 5)-α-Ara*f*
64.1	C-5 of β-Ara*f*
64.0	C-6 of 2-β-Gal*f*, 3-β-Gal*f*
62.4–61.7	C-5 of 2-α-Ara*f*
	C-6 of 5-β-Gal*f*, β-Glc*p*

[a] The relevant ^{13}C NMR spectra are shown in Figs. 2 and 3.

in that of genera devoid of these glycosyl residues (Fig. 2 and 3; Table I). The resonance of C-6 of 5-linked Gal*f* was observed among those of the other nonglycosylated primary alcohol, at $\delta \sim 62$, in agreement with published data[6]; (*vi*) the majority of the C-1 resonances appeared between δ 108 and 109, corresponding to α-Ara*f*/β-Gal*f* units[4–8]; (*vii*) GLC evidence for the absence of some glycosyl residues (Table I) in some species correlated with the absence of corresponding signals (Figs. 2 and 3). For example, the signal at $\delta \sim 109$ was absent from the spectra of the arabinogalactan of *Nocardia* spp. Examination of Table I showed that the two *Nocardia* species were devoid of an internal 6-linked Gal*f* which was replaced by an internal (5 → 6)-linked Gal*f*. The absence of this signal at $\delta \sim 109$ in the spectrum of lipoarabinomannan[9,10] indicated that it corresponded to the C-1 of a galactosyl residue. Since the C-1 of 5-linked Gal*f* resonances at $\delta \sim 108$[6] and 5-linked Gal*f* was found in all the arabinogalactans (Table I), it follows that the resonance at $\delta \sim 109$ is assignable to C-1 of 6-linked Gal*f*, suggesting that this resonance, in contrast to that of 5-linked Gal*f* (δ 108.1), is not sensitive to glycosylation at position 6 (methyl galactofuranoside has its C-1 resonance at δ 109)[4].

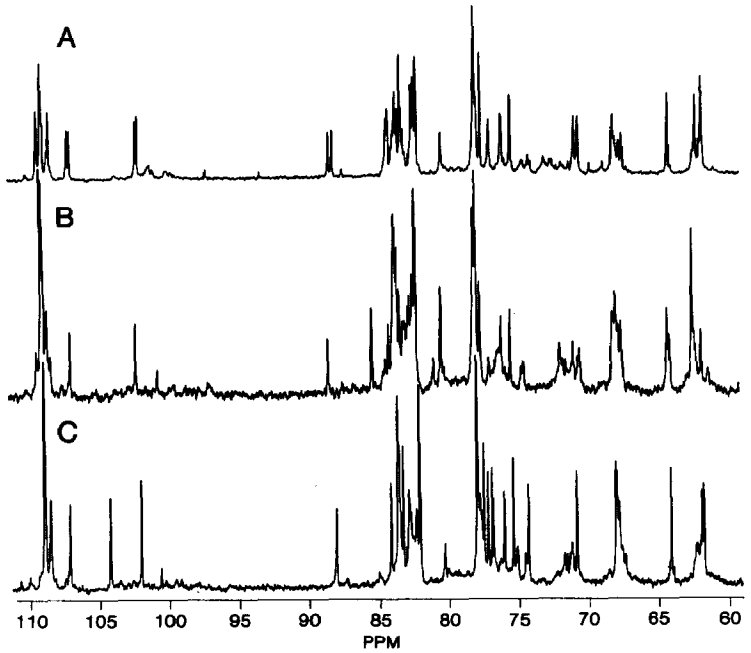

Fig. 3. Comparison of the ^{13}C NMR spectra of the base solubilized cell wall arabinogalactans of *M. tuberculosis* (A) (from Fig. 2A), *R. equi* (B), and *N. asteroides* (C). Assignments of the major signals are presented in Table II.

The similarity between the ^{13}C NMR fingerprints of the different mycobacterial arabinogalactans suggested that each contained the same major structural motifs already defined for the cell wall arabinogalactan of *M. tuberculosis*[2]. Distinctive differences between the spectra of the mycobacterial arabinogalactans and those of *Nocardia* and *Rhodococcus* (Fig. 3) lay in the number of resonances corresponding to t-β-Ara$f$ (δ 101–102) and 2-linked α-Ara$f$ (δ 106–107); in the latter two genera, only one signal was observed for each group of resonances, suggesting different arrangements in the termini of the arabinan chains.

Differences in the galactan segments of the heteropolysaccharides were also seen in the representatives of the different genera. The conclusions drawn from the analysis in Table I were supported by the absence of signals at $\delta \sim 109$ and ~ 70 in the spectra of the products from the *Nocardia* spp. corresponding respectively to C-1 and C-5 of 6-linked Gal$f$ (Fig. 3, Table II) and the presence, in the corresponding spectra, of resonances due to additional t-β-Glc$p$ units (Table II). The ^{13}C NMR spectra of the polysaccharide isolated from *Rhodococcus* were different from those of the *Myobacterium* and *Nocardia* spp. in the presence of an additional signal at $\delta \sim 64$ assignable to the C-6 of 2-linked Gal$f$ and 3-linked Gal$f$ (Table II).

Structures of comparative fragments from the arabinogalactans of various Mycobacterium spp.—Per-*O*-alkylated oligoglycosyl alditols were prepared from the

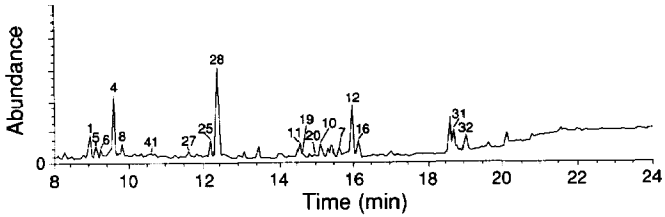

Fig. 4. The total ion chromatogram produced during GLC–MS analysis of the mixture of per-*O*-alkylated glycosyl alditols arising from the series of chemical reactions performed on the cell walls (see Experimental section) of *M. leprae*. The numbers given to some of the peaks correspond to the compound/fragment numbers in ref 2.

per-*O*-methylated cell walls of a few species of *Mycobacterium* not previously examined, by partial acid hydrolysis followed by $NaB[^2H]_4$ reduction, and ethylation of the resulting oligoglycosyl alditols[2]. The mixtures were analyzed by GLC–MS and the structures of the quantitatively major per-*O*-alkylated oligoglycosyl alditols were characterized in absolute terms, and by comparison with the well characterized products obtained previously from the cell wall of *M. tuberculosis*[2]. The total profile of the per-*O*-alkylated oligoglycosyl alditols obtained from the arabinogalactan of *M. leprae* is shown in Fig. 4. Details of the structures of each of these fragments is provided in ref 2.

Of considerable importance was the fact that the same structural arrangements were observed in *M. leprae*, demonstrating that they are not just a product of culture conditions but reflect the intracellular state of a true intracellular infectious agent. Indeed, the ^{13}C NMR and methylation analyses had already predicted the presence in all mycobacterial arabinogalactans of the same predominant structural motifs. Thus it appears that all of the mycobacterial arabinogalactans conform to the published pattern and consist of a homogalactan of alternating 5-linked β-Gal*f* and 6-linked β-Gal*f* units (motif D, Fig. 5). The arabinan component from all sources apparently consists of at least three major structural motifs: a linear 5-linked α-Ara*f* strand (motif C, Fig. 5), a branched (3 → 5)-linked α-Ara*f* unit substituted with 5-linked α-Ara*f* residues at both branched positions (motif B, Fig. 5), and a nonreducing terminal region characterized by (3 → 5)-linked α-Ara*f* residues substituted at both branched positions with the disaccharide β-Ara*f*-(1 → 2)-α-Ara*f* (motif A, Fig. 5).

Structures of arabinogalactan fragments from members of the Nocardia and Rhodococcus genera.—Likewise, the cell walls of *N. asteroides* and of *R. equi* were partially fragmented. The total ion profiles of the resulting per-*O*-alkylated oligoglycosyl alditols are shown in Figs. 6 and 7; the key fragments that were examined in detail are marked. The analytical data leading to a definition of these fragments are presented in the Experimental section, and the identity of the fragments is given in Table III.

From this information, we conclude that *Nocardia* and *Rhodococcus* differ from *Mycobacterium* in the nonreducing terminal region. The two *Nocardia* genera

t-Ara*f*-β-(1 → 2-Ara*f*-α-(1 → 3)
t-Ara*f*-β-(1 → 2-Ara*f*-α-(1 → 5) → Ara*f*-α-(1 →

AG motif A

Ara*f*-α-(1 → 3)
Ara*f*-α-(1 → 5) → Ara*f*-α-(1 →

AG motif A'

→ 5)-Ara*f*-α-(1 → 3)
→ 5)-Ara*f*-α-(1 → 5) → Ara*f*-α-(1 → 5)-Ara*f*-(1 →

AG motif B

$[\rightarrow 5)$-Ara*f*-α-(1 → $]_m$

AG motif C

$[\rightarrow 5)$-Gal*f*-β-(1 → 6)-Gal*f*-β-(1 → $]_n$

AG motif D

t-Ara*f*-β-(1 → 2)-Ara*f*-α-(1 → 5)-Ara*f*-α-(1 →

AG motif E

$[\rightarrow 5)$-Gal*f*-β-(1 → 5)-Gal*f*-(1 → $]_p$
6
↑
1
t-Glc*p*

AG motif F

$[\rightarrow 5)$-Gal*f*-β-(1 → 5)-Gal*f*-β-(1 → 3)-Gal*f*-(1 → $]_q$

AG motif G

→ 4)-Rha*p*-(1 → 3)-GlcNAc*p*-(1 →

AG motif H

Fig. 5. The major structural motifs present in the arabinogalactans of the various species from the actinomycete genera. The values for m, n, p, and q were not rigorously established.

contain structural motif E (Fig. 5) consisting of terminal β-Ara*f* linked to position 2 of an Ara residue, which, in turn, is linked to position 5 of an Ara unit within the arabinan chain. This conclusion was supported by the existence of fragments b [t-Ara*f*-(1 → 2)-Ara*f*-(1 → 5)-Ara*f*], and c [→ 2-Ara*f*-(1 → 5)-Ara*f*-(1 → 5)-Ara*f*] (Table III). Only traces of this motif are present in the mycobacterial arabinogalactans.

Interpretation of the data leading to elucidation of the structures of the nonreducing termini of the arabinan of *R. equi* is more complex. Significant

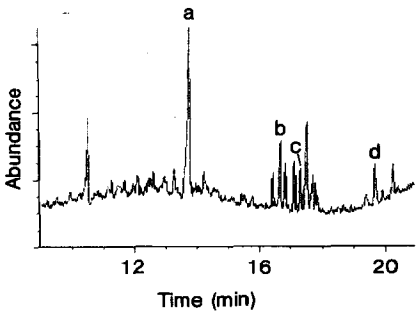

Fig. 6. The total ion chromatogram produced during GLC–MS analysis of the mixture of per-*O*-alkylated glycosyl alditols arising from the series of chemical reactions performed on the cell walls of *N. asteroides*. The letters given to some of the peaks correspond to the compounds/fragments whose structures are described in Table III.

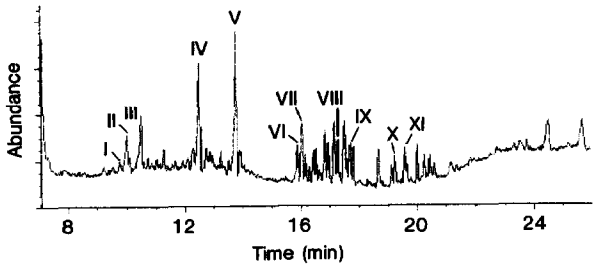

Fig. 7. The total ion chromatogram produced during GLC–MS analysis of the mixture of per-*O*-alkylated glycosyl alditols arising from the series of chemical reactions performed on the cell walls of *R. equi*. The numbers given to some of the peaks correspond to the compounds/fragments whose structures are described in Table III.

amounts of terminal Ara*f* were linked to both positions 3 and 5 of a (3 → 5)-linked Ara unit, as demonstrated by proof of the existence of fragments i, iii, vi, and viii (Table III). Presumably the nonreducing Ara*f* units in these fragments represent what in mycobacteria are 2-linked α-Ara*f* residues but in *R. equi* are t-α-Ara*f* residues lacking terminal β-Ara*f* units at position 2. The resultant motif A′ is shown in Fig. 5. However, the signals at δ 101.9 and δ 88.2 in the ^{13}C NMR spectrum of *R. equi* arabinogalactan (Fig. 3B) and the isolation of fragment vii (Table III) showed that, at times, motif A′ is modified by the addition of a single terminal β-Ara*f* unit. Whether this β-Ara*f* is always added to C-2 of the α-Ara*f* residue that is attached to C-3 of the branched α(3 → 5)-Ara*f* unit, as in fragment vii, is not known for sure, but the presence of single signals in the ^{13}C NMR spectrum in both the region of δ 101.9 and 88.2 suggest that this is the case. Also present are unbranched linear nonreducing termini, Ara-(1 → 5)-Ara, where, presumably, all of the terminal arabinosyl units are α, as shown by recognition of ii (Table III); note that this structure is different from that of motif E (Fig. 5) found in *Nocardia*. In addition to this large array of nonreducing-end arabinosyl structures in *R. equi*, there is evidence of a nonreducing-end t-Man*p* unit attached directly to C-5 of a branched (2 → 5)-Ara*f* unit as evidenced by the recognition of fragment ix (Table III).

In the homogalactan region, major differences were observed between the polysaccharides of mycobacteria, *Nocardia* and *Rhodococcus*. In addition to motif D (Fig. 5), established by the recognition of fragment v, *R. equi* contained substantial amounts of 5-linked Gal*f* linked to position 3 of a Gal*f* (motif G) as shown by recognition of fragments iv, x, and xi. *Nocardia* was found to contain 5-linked Gal*f*, (5 → 6)-linked Gal*f*, and t-Glc*p* (Table I), raising two possibilities concerning the structure of the galactan segment. Firstly, galactan could consist of a chain of alternating 5-linked Gal*f* and 6-linked Gal*f* units, as is the case in mycobacteria, with all of the 6-linked Gal*f* residues substituted at C-5 with a t-β-Glc*p*. Alternatively, the backbone may consist of linear 5-linked Gal*f* residues with about one-half of these substituted at C-6 with t-β-Glc*p*. These two possibili-

TABLE III

Per-*O*-alkylated oligoglycosyl alditols produced from the cell walls of *N. asteroides* and *R. equi* after a series of chemical derivatizations and fragmentations [a]

Species	Figure	Fragment	Structure of fragment
N. asteroides	Fig. 6	a	Glc*p*-(1 → 6)-Gal-ol [b] 5 ↑
N. asteroides	Fig. 6	b	Ara*f*-(1 → 2)-Ara*f*-(1 → 5)-Ara-ol
N. asteroides	Fig. 6	c	→ 2-Ara*f*-(1 → 5)-Ara*f*-(1 → 5)-Ara-ol
N. asteroides	Fig. 6	d	Glc*p*-(1 → 6) Gal-ol Gal*p*-(1 → 5)
R. equi	Fig. 7	i	Ara*f*-(1 → 3)-Ara-ol 5 ↑
R. equi	Fig. 7	ii	Ara*f*-(1 → 5)-Ara-ol
R. equi	Fig. 7	iii	Ara*f*-(1 → 5)-Ara-ol 3 ↑
R. equi	Fig. 7	iv	→ 5Gal*f*-(1 → 3)-Gal-ol
R. equi	Fig. 7	v	→ 5Gal*f*-(1 → 6)-Gal-ol
R. equi	Fig. 7	vi	Ara*f*-(1 → 3) Ara-ol Ara*f*-(1 → 5)
R. equi	Fig. 7	vii	Ara*f*-(1 → 2)-Ara*f*-(1 → 3)-Ara-ol 5 ↑
R. equi	Fig. 7	viii	Ara*f*-(1 → 5)-Ara*f*-(1 → 5)-Ara-ol 3 ↑
R. equi	Fig. 7	ix	Man*p*-(1 → 5)-Ara*f*-(1 → 5)-Ara-ol 2 ↑
R. equi	Fig. 7	x	→ 5-Gal*f*-(1 → 5)-Gal*f*-(1 → 3)-Gal-ol
R. equi	Fig. 7	xi	→ 5-Gal*f*-(1 → 3)-Gal*f*-(1 → 5)-Gal-ol

[a] Analytical data are presented in the Experimental section.

[b] Nomenclature: e.g., Glc*p*-(1 → 6)-Galol, 2,3,4,6-tetra-*O*-methyl-glucopyranosyl-(1 → 6)-2,3-di-*O*-methyl-1,4,5-tri-*O*-pentadeuterioethylgalactitol. The presence of a $C_2^2H_5$ unit on C-5 of the 6-linked Gal-ol unit indicates that it is the point of further glycosylation, and is so demonstrated by the arrow.

ties were readily resolved by recognition of fragment a, t-Glc-(1 → 6)-Gal[5 →]-ol, which requires a homo-(1 → 5)-linked β-Gal*f* backbone.

Fragment d (Table III) supports this arrangement, although its structure was based, in part, upon knowledge of the structure of fragment a. Thus, the existence of motif F (Fig. 5) in *Nocardia* was tentatively established.

Nature of the linkage between the arabinogalactans and the peptidoglycan.—The arabinogalactans of mycobacteria are attached to peptidoglycan through a phosphodiester link and involve a special linker disaccharide, → 4-Rha*p*-(1 → 3)-GlcNAc[14], where the galactan is attached at the 4 position of the Rha and the GlcNAc is attached at C-1 to a phosphate present on C-6 of muramic acid. To determine whether this principle applied to the related genera, the cell walls were

hydrolyzed with a mild acid[14] to selectively cleave glycosyl phoshates, reduced, fully hydrolyzed, and the trimethylsilyl derivatives were analyzed by GLC–MS Glucosaminitol, arising from GlcNAc-1 phosphate residues, was detected for all strains. In addition, a 4-linked Rha*p* unit was observed in all of the cell walls (Table I), further supporting the evidence[14,21] that this particular linkage is common to all actinomycetes. The actual intact disaccharide, Rha-(1 → 3)-GlcNAc, and galactosyl-containing extensions were previously recognized in *M. leprae, M. tuberculosis, M. bovis*[14], and *Nocardia*. Incidentally, the galactosamine observed in some mycobacterial cell walls[14,15] was not rendered reducible by the mild acid and, indeed, was entirely absent from the cell walls of *M. smegmatis* and both *Nocardia* species, indicating that this amino sugar does not participate in the linkage between arabinogalactan and peptidoglycan.

Chemotaxonomy supported by morphology is the most definitive means for defining actinomycete genera[16]. Aerobic actinomycetes of the nocardioform morphological group have in common a chemotype IV cell wall, the major distinguishing characteristics if which are *meso*-diaminopimelic acid, arabinose, galactose, and high molecular weight, 2-hydroxy, 3-branched fatty acids, of the mycolic acid type[17,18]. Further differentiation can be provided by the length of the mycolic acids[1,19]; mycobacteria are endowed with mycolic acids of 60–90 carbons; nocardia are typified by 46–60 carbon mycolates; and rhodococcus by 34–66 carbon mycolates. However, there was little information on the composition of cell wall arabinogalactan other than that they may contain additional sugars[20].

Until recently, little substantial information was available on the arrangement of the arabinosyl and the galactosyl residues within the heteropolysaccharide. Application of acid catalyzed partial depolymerization to the per-*O*-alkylated cell wall polysaccharide of *M. tuberculosis* had allowed the formulation of novel structural features[2]. Independent proof of the postulated structural features of arabinogalactan has since been obtained through enzymatic cleavage (unpublished results by the authors). In the present work, application of some of these technique to other mycobacterial species clearly established close structural similarity between the arabinogalactans isolated from virulent, attenuated, and saphrophytic strains of *Mycobacterium*. However, the most weighty finding was the realization that the heteropolysaccharides isolated from related genera have some shared and unique structural features in both the arabinan and galactan segments. The arabinan structural motifs B and C were found in all genera and must provide the bulk of the internal portions of the arabinans. However, the nonreducing termini (shown in motifs A, A′, and E) differ. Mycobacteria appear distinctive in the elaboration of some termini in the form of a pentasaccharide, whereas *Rhodococcus* and *Nocardia* genera present their termini primarily in the form of trisaccharides (motifs A′ and E, respectively), although other nonreducing termini are present in *R. equi*.

The structures of the galactan component differ according to the genus and within the genus *Rhodococcus*. The alternating 6-linked and 5-linked galactofuranosyl residues that typify the *Mycobacterium* genus is only partially valid for *R.*

equi and not at all for *Nocardia*. Thus *R. equi* produces motif G and the *Nocardia* strains elaborate a homogalactan consisting of 5-linked galactofuranosyl residues heavily substituted by terminal β-glucosyl residues. Finally, as predicted from the work of Fujioka et al.[21] and ourselves[14], mycobacteria and related genera apparently share the same reducing terminal motif involved in the linkage of the heteropolymer to peptidoglycan.

It should be noted that the motifs detected in *Nocardia* (motifs B, C, E, F, and H) (Fig. 5) and in *R. equi* (motifs A', B, C, D, G, and H) (Fig. 5), although correct in themselves, do not necessarily represent all of those present in these cell wall arabinogalactans. Elucidation of the arrangements of the motifs, i.e., how they are linked together, remains a challenging problem and one currently being addressed for the mycobacterial arabinogalactan by structurally characterizing larger molecular weight fragments (unpublished results of the authors).

EXPERIMENTAL

Strains and growth conditions.—Mycobacterium avium serovar 4 (TMC 1463), *M. avium* serovar 25 (strain 72-888), *M. bovis* BCG (ATCC 35733), *M. tuberculosis* Erdman (TMC 107), and *M. smegmatis* (ATCC 14468) were grown on 7H11 broth for 2–6 weeks at 37°C. *Nocardia brasiliensis* (ATCC 19296), *N. asteroides* (a clinical isolate), *Rhodococcus rhodochrous* (ATCC 13808), and *Rhodococcus equi* (a clinical isolate) were grown on tryptic soy broth for 2 weeks at 37°C. The suspensions were autoclaved and the cells were recovered by centrifugation, washed, and frozen at −20°C. *M. leprae* was isolated from experimentally infected armadillo tissues[22].

Production of cell walls.—The bacterial cells were suspended in phosphate buffer (pH 7.2) and sonicated[2]. Cell walls were recovered by centrifugation at 27000*g*, extracted with 2% SDS (w/v) at 95°C for 1 h to remove soluble proteins and lipoarabinomannan[2], and finally pelleted at 27000*g*. Cell walls so prepared were washed several times with water and 80% (v/v) acetone to remove SDS and lyophilized. Cell walls prepared in this way were the object of most of the analyses. However, to release arabinogalactans from peptidoglycans for NMR analyses, cell walls were treated with 2 M NaOH for 16 h at 80°C and the soluble arabinogalactans purified as described[2].

Methylation of the cell wall arabinogalactans and production and analysis of per-O-alkylated oligoglycosyl alditols.—The intact cell walls were methylated by a slight modification of the Hakomori procedure using CH_3I as described[2]. In order to obtain partially methylated oligoglycosyl alditol fragments, the per-*O*-methylated cell walls were subjected to partial acid hydrolysis with 2 M CF_3CO_2H for 1 h at 75°C, followed by reduction with $NaB[^2H]_4$, and pentadeuterioethylation[2]. GLC–MS of the per-*O*-alkylated oligoglycosyl alditol fragments were performed on a Hewlett–Packard 5890 gas chromatograph connected to a Hewlett–Packard 5970 mass selector detector as described[2].

Glycosyl linkage composition.—For the purposes of establishing glycosyl linkage composition, aliquots of the per-*O*-alkylated arabinogalactans were hydrolyzed with 2 M CF_3CO_2H at 120°C for 1 h, reduced, acetylated, and analyzed by GLC–MS.

For analysis of mild acid labile (i.e., phosphate-linked) GlcNAc, cell walls were hydrolyzed with mild acid, reduced, *N*-acetylated, trimethylsilylated, and analyzed by GLC–MS as described[14].

NMR analysis.—The alkali-released solubilized arabinogalactans were suspended in 1 mL of $[^2H]_2O$. Acetone (1 μL) was added as an internal standard, and NMR was performed on a Bruker ACE-300 at the Department of Chemistry Central Facility or the Regional NMR Center at Colorado State University. The acetone signals were assigned at δ 31.4 and 2.04 for ^{13}C and ^1H NMR, respectively, corresponding to δ 0 for Me_4Si. Proton-decoupled ^{13}C spectra were obtained with composite pulse proton decoupling (Bruker supplied program).

Analytical data on the per-O-alkylated oligoglycosyl alditols.—The pertinent per-*O*-alkylated oligoglycosyl alditols used in establishing structural motifs shared by other genera (motifs B and C; Fig. 5) have been characterized as previously described in detail[2]. The analytical data listed below concerns only key fragments leading to recognition of specific motifs found in *Rhodococcus* (motifs A′ and G) and *Nocardia* (motifs E and F). The mass spectral nomenclature used is that of Kochetkov and Chizhov[11].

Compound **a**: GLC, t_R: 13.7 min. EIMS ions at *m/z* 88, 187 (aA_2), 219 (aA_1), 293 ($aldJ_2$), and alditol cleavage ions 216, 184 (216–CH_3OH), 375, 463. Proposed structure t-Glc*p*-(1 → 6)-Gal-ol-5 ← .

Compound **b**: GLC t_R: 16.7 min. EIMS ions at *m/z* 101, 143 (aA_2), 175 (aA_1), 230 ($aldJ_2$), 276 [$aldJ_0$ (refs. 2 and 12)], 290 ($aldJ_1$), 303 (baA_2), 335 (baA_1), 428, 472 (alditol cleavage). Proposed structure t-Ara-(1 → 2)-Ara-(1 → 5)-Ara-ol. It was determined in our earlier studies[2] that 2-linked Ara*f* residues result in the presence of both $aldJ_0$ and $aldJ_1$ ions.

Compound **c**: GLC t_R: 17.3 min. EIMS ions at *m/z* 120, 162 (aA_2), 194 (aA_1), 230 ($aldJ_2$), 290 ($aldJ_1$), 354 (baA_1), 491 (alditol cleavage). Proposed structure → 2)-Ara-(1 → 5)-Ara-(1 → 5)-Ara-ol.

Compound **d**: GLC t_R: 19.7 min. EIMS ions at *m/z* 88, 101, 187 ($a'A_2$), 219 ($a'A_1$), 238 (aA_1), 288 [alditol rearrangement[13]], 478 ($a'aldJ_2$), 497 ($a'aldJ_2$), 538 ($a'aldJ_1$). Proposed structure

t-Glc-(1 → 6)\
→ 5-Gal-(1 → 5)/ Galol

The location of the t-Glc and 5-linked Gal is based upon knowledge of the structure of compound **a**.

Compound **i**: GLC t_R: 9.6 min. EIMS ions at *m/z* 101, 143 (aA_2), 175 (aA_1), 249 ($aldJ_2$), 309 ($aldJ_1$), 376 (alditol cleavage). Proposed structure t-Ara-(1 → 3)-Ara-ol ← .

Compound **ii**: GLC t_R: 9.9 min. EIMS ions at m/z 101, 143 (aA_2), 175 (aA_1), 230 ($aldJ_2$), 290 ($aldJ_1$), 268 (alditol cleavage), 312 (alditol cleavage). Proposed structure t-Ara-(1 → 5)-Ara-ol.

Compound **iii**: GLC t_R: 10.1 min. EIMS ions at m/z 101, 143 (aA_2), 175 (aA_1), 249 ($aldJ_2$), 309 ($aldJ_1$), 268 (alditol cleavage), 331 (alditol cleavage). Proposed structure t-Ara-(1 → 5)-Ara-ol-3 ← .

Compound **iv**: GLC t_R: 12.5 min. EIMS ions at m/z 101, 155 (aA_3), 206 (aA_2), 238 (aA_1), 274 ($aldJ_2$), 334 ($aldJ_1$), 439 (alditol cleavage), 407 (439-methanol) (no alditol cleavage ions). Proposed structure → 5)-Gal-(1 → 3)-Gal-ol. Based also on the methylation analysis (Table I).

Compound **v**: GLC t_R: 13.8 min. EIMS ions at m/z 101, 155 (aA_3), 206 (aA_2), 238 (aA_1), 274 ($aldJ_2$), 334 ($aldJ_1$), and alditol cleavage ions of 184 (216 − 32), 216, 375, 482. Proposed structure → 5-Gal-(1 → 6)-Gal-ol.

Compound **vi**: GLC t_R: 15.9 min. EIMS ions at m/z 101, 143 (aA_2, $a'A_2$), 175 (aA_1, $a'A_1$), 244 [alditol rearrangement[13]], 268 (alditol cleavage), 390 ($aaldJ_2$, $a'aldJ_2$), 450 ($aaldJ_1$, $a'aldJ_1$). Proposed structure

t-Ara-(1 → 3)\
Ara-ol\
t-Ara-(1 → 5)/

Compound **vii**: GLC t_R: 16.0 min. EIMS ions at m/z 101, 143 (aA_2), 175 (aA_1), 295 ($aldJ_0$), 303 (baA_2), 309 ($aldJ_1$), 469 ($baldJ_2$). No alditol cleavages. Proposed structure t-Ara-(1 → 2)-Ara-(1 → 3)-Ara-ol-5 ← .

Compound **viii**: GLC t_R: 17.3 min. EIMS ions at m/z 101, 143 (aA_2), 175 (aA_1), 230 ($aldJ_2$), 309 ($aldJ_1$), 354 (baA_1), 447 (alditol cleavage). Proposed structure t-Ara-(1 → 5)-Ara-[3 ←]-(1 → 5)-Ara-ol.

Compound **ix**: GLC t_R: 17.7 min. EIMS ions at m/z 88, 101, 187 (aA_2), 219 (aA_1), 230 ($aldJ_2$), 290 ($aldJ_1$), 398 (baA_1), 469 ($baldJ_1$), 535 (alditol cleavage). Proposed structure t-Man-(1 → 5)-Ara-[2 ←](1 → 5)-Ara-ol.

Compound **x**: GLC t_R: 19.3 min. EIMS ions at m/z 101, 120, 155 (aA_3), 206 (aA_2), 238 (aA_1), 274 ($aldJ_2$), 334 ($aldJ_1$), 442 baA_1), 643 (alditol cleavage). Proposed structure → 5)-Gal-(1 → 5)-Gal-(1 → 3)-Gal-ol.

Compound **xi**: GLC t_R: 19.6 min. EIMS ions at m/z 101, 155 (aA_3), 206 (aA_2), 238 (aA_1), 274 ($aldJ_2$), 320 [$aldJ_0$ (ref 12)], 410 (baA_2), 442 (baA_1), 478 ($baldJ_2$). Proposed structure → 5)-Gal-(1 → 3)-Gal-(1 → 5)-Gal-ol.

ACKNOWLEDGMENTS

The authors are grateful to Marilyn Hein for preparing the manuscript, Carol Marander for the graphics, and Cynthia Bozic and Kimberly Robuck for technical help. This work was supported by Grant AI-18357 and Contract NO1 AI-05074 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health. M.D. was also supported by funds from C.N.R.S., France.

REFERENCES

1 P.J. Brennan, in C. Ratledge and S.G. Wilkinson (Eds.), *Microbial lipids*, Academic Press, London, 1988, pp 203–298.
2 M. Daffe, P.J. Brennan, and M. McNeil, *J. Biol. Chem.*, 265 (1990) 6734–6743.
3 G.O. Aspinall, in G.O. Aspinall (Ed.), *The Polysaccharides*, Vol. 2, Academic Press, New York, 1983, pp 1–9.
4 P.A.J. Gorin and M. Mazurek, *Can. J. Chem.*, 53 (1975) 1212–1222.
5 J.H. Bradbury and G.A. Jenkins, *Carbohydr. Res.*, 126 (1984) 125–156.
6 P.A.J. Gorin and M. Mazurek, *Carbohydr. Res.*, 48 (1976) 171–186.
7 J.-P. Joseleau, G. Chambat, M. Vignon, and F. Barnoud, *Carbohydr. Res.*, 58 (1977) 165–175.
8 K. Mizutani, R. Kasai, M. Nakamura, and O. Tanaka, *Carbohydr. Res.*, 185 (1989) 27–38.
9 P.L. Weber and G.R. Gray, *Carbohydr. Res.*, 74 (1979) 259–278.
10 E.A. Sudgen, B.S. Samagh, D.R. Bundle, and J.R. Duncan, *Infect. Immun.*, 55 (1987) 762–770.
11 N.K. Kochetkov and O.S. Chizhov, *Adv. Carbohydr. Chem.*, 21 (1966) 39–92.
12 J.K. Sharp and P. Albersheim, *Carbohydr. Res.*, 128 (1984) 193–202.
13 B. Fournet, J.-M. Dhalluin, G. Strecker, J. Montreuil, G. Bosso, and J. Defaye, *Anal. Biochem.*, 108 (1980) 35–56.
14 M. McNeil, M. Daffe, and P.J. Brennan, *J. Biol. Chem.*, 265 (1990) 18200–18206.
15 P. Draper, *J. Gen. Microbiol.*, 65 (1991) 313–324.
16 S.T. Williams and E.M.H. Wellington, in M. Goodfellow and R.G. Board (Eds.), *Microbiological Classification and Identification*, Academic Press, London, 1980, pp 139–165.
17 M.P. Lechevalier and H. Lechevalier, *Int. J. Syst. Bacteriol.*, 20 (1970) 435–444.
18 M. Goodfellow and D.E. Minnikin, *Annu. Rev. Microbiol.*, 31 (1977) 153–180.
19 M. Goodfellow and D.E. Minnikin, in M. Goodfellow and R.G. Board (Eds.), *Microbiological Classification and Identification*, Academic Press, London, 1980, pp 189–256.
20 C. Abou-Zeid, A. Voiland, G. Michel, and C. Cocito, *Eur. J. Biochem.*, 128 (1982) 363–370.
21 M. Fujioka, S. Koda, and Y. Morimoto, *J. Gen. Microbiol.*, 131 (1985) 1323–1329.
22 S.W. Hunter, T. Fujiwara, and P.J. Brennan, *J. Biol. Chem.*, 257 (1982) 15072–15078.
23 W.S. York, A.G. Darvill, M. McNeil, J.T. Stevenson, and P. Albersheim, *Methods Enzymol.*, 118 (1986) 3–40.