Predominant Structural Features of the Cell Wall Arabinogalactan of *Mycobacterium tuberculosis* as Revealed through Characterization of Oligoglycosyl Alditol Fragments by Gas Chromatography/Mass Spectrometry and by $^1$H and $^{13}$C NMR Analyses*

(Received for publication, May 30, 1989)

Mamadou Daffe†, Patrick J. Brennan, and Michael McNeil§

From the Department of Microbiology, Colorado State University, Fort Collins, Colorado 80523

The peptidoglycan-bound arabinogalactan of a virulent strain of *Mycobacterium tuberculosis* was per-O-methylated, partially hydrolyzed with acid, and the resulting oligosaccharides reduced and O-pentadentioxyethylated. The per-O-alkylated oligoglycosyl alditol fragments were separated by high pressure liquid chromatography and the structures of 43 of these constituents determined by $^1$H NMR and gas chromatography/mass spectrometry. The arabinogalactan was shown to consist of a galactan containing alternating 5-linked \( \beta-\)D-galactofuranosyl (Gal\( f \)) and 6-linked \( \beta-\)D-Galf residues. The arabinan chains are attached to C-5 of some of the 6-linked Galf residues. The arabinan is comprised of at least three major structural domains. One is composed of linear 5-linked \( \alpha-\)D-arabinofuranosyl (Araf) residues; a second consists of branched 3,5-linked \( \alpha-\)D-Araf units substituted with 5-linked \( \alpha-\)D-Araf residues at both branched positions. The non-reducing terminal region of the arabinan was characterized by a 3,5-linked \( \alpha-\)D-Araf residue substituted at both branched positions with the disaccharide \( \beta-D-Araf-(1\rightarrow2)\)-\( \alpha-D-Araf \). $^{13}$C NMR of intact soluble arabinogalactan established the presence of both \( \alpha-\) and \( \beta-\)Araf residues in this domain. This non-reducing terminal motif apparently provides the structural basis of the arabinan in this domain. The arabinan chains are attached to the reducing terminal region of the galactan core and may link the arabinogalactan to the peptidoglycan. Evidence is presented for the presence of minor structural features involving terminal manno-pyranosyl units. Models for most of the heteropolysaccharide are proposed which should increase our understanding of a molecule responsible for much of the immunogenicity, pathogenicity, and peculiar physical properties of the mycobacterial cell.

* This work was supported by Grant AI-18357 and Contract NO1 AI-52582 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health. The NMR conducted at the Colorado State University Regional NMR Center was supported by National Science Foundation Grant CHE 86-16437. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Portions of this paper (including "Experimental Procedures," part of "Results," and Figs. 2, 3, 7, and 9) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

† Supported by funds from the Centre National de la Recherche Scientifique. Permanent address: Centre de Recherche de Biochimie et Génétique Cellulaires du C. N. R. S., 110 Route de Narbonne, 31062 Toulouse Cedex, France.

§ To whom correspondence should be addressed.

EXPERIMENTAL PROCEDURES

RESULTS

The strategy for the generation and analysis of per-O-alkylated oligoglycosyl alditols in order to establish the structures of complex polysaccharides has been described in detail (19, 20). Purified cell walls of *M. tuberculosis*, rather than arabinogalactan released from cell walls by degradative means (6, 11), were used in an attempt to obtain fragments contain-

The cell walls of members of the *Mycobacterium* genus and related genera contain a chemotype IV peptidoglycan (1, 2) to which an arabinogalactan is covalently attached. The arabinogalactan is further modified by esterification with mycolic acid residues (3, 4). Small amounts of tightly associated, highly immunogenic proteins are also present (5).

The arabinogalactan component of the cell wall of mycobacteria has been implicated in a range of biological responses associated with human and experimental mycobacterioses, such as the high titer IgG antibodies in tuberculous and leprosy sera (6, 7) and the state of T-cell-mediated immunological anergy evident in the multibacillary form of disease (8).

It has previously been demonstrated that the polymer is composed mostly of \( \alpha-\)arabinofuranosyl and \( \beta-\)galactosyl residues (9–12), that the arabinosyl residues were all furanoid (6, 13), and that the arabinosyl portions of the molecule were the dominant B-cell antigens (6). The linkages of many of the arabinosyl and galactosyl residues were determined (6, 14) and the disaccharide, \( \beta-Gal-(1\rightarrow6)-Gal \), isolated (15). More recently, additional linkages involving the galactosyl units have been identified (16, 17) and all of these galactosyl residues have been shown to be furanoid (17). However, the manner in which the variously linked glycosyl units are combined has not been established because of the complexity of the heteropolysaccharide which is not composed of a repeating unit. In this present, comprehensive study, 43 different oligosaccharide fragments of the polysaccharide were characterized which allowed the recognition of several structural motifs that were representative of the majority of the molecule. The proposed structure of the arabinogalactan should lead to a greater understanding of the role of the mycolylarabinogalactan-peptidoglycan complex in the immunogenicity, pathogenesis and physiology of mycobacteria (18).
TABLE I

Glycosyl-linkage analysis of arabinogalactan and percent of glycosyl linkages cleaved by partial acid hydrolysis

Conditions for methylation of arabinogalactan, complete acid hydrolysis, reduction with NaB\(^{14}H\), acetylation, and GC/MS analysis are described in the Miniprint. Partial acid hydrolysis was conducted in 2 M trifluoroacetic acid at 75 °C for 1 h.

| Glycosyl residue | Mol % | % cleaved by partial hydrolysis |
|-----------------|-------|---------------------------------|
| \(\alpha\)-Araf   | 11    | 70                              |
| 2-Araf           | 11    | 56                              |
| 5-Araf           | 37    | 38                              |
| 3,5-Araf         | 11    | 45                              |
| t-Gal\(^f\)      | 1     | ND                              |
| 5-Gal\(^f\)      | 17    | 26                              |
| 6-Gal\(^f\)      | 9     | 36                              |
| 5,6-Gal\(^f\)    | 2     | 50                              |
| 4-Rhap           | 1     | ND                              |
| t-Manp           | 0.8   | ND                              |
| 6-Manp           | 0.5   | ND                              |

* ND = not determined.

The linkage element between arabinogalactan and peptidoglycan.

Generation and Structures of Per-0-Alkylated Oligoglycosyl Alditols—The cell wall was per-0-methylated and an aliquot hydrolyzed, reduced, acetylated, and analyzed by GC/MS\(^2\) (Table I). The glycosyl linkage composition is almost identical to that previously reported (17), indicating full methylation and a stable, invariant structure.

Partially methylated oligosaccharide fragments were prepared from the per-0-methylated cell wall by partial acid hydrolysis which resulted in random cleavage of the various glycosidic linkages. The products were reduced with sodium borodeuteride, pentadecuterioethylated, and the resulting per-O-alkylated oligoglycosyl alditols isolated. The application of this series of degradation/modifications to a portion of the arabinan segment of the arabinogalactan is illustrated in Fig. 1, A-D.

To avoid ambiguities in the interpretation of mass spectral data, pentadecuterioethylated \(\text{C}_{15}\)\(^{2}H\)\(_{15}\) instead of pentahydroethylation \(\text{C}_{15}\)\(^{2}H\)\(_{15}\) was used; di-O-C\(_{5}\)\(^{2}H\)\(_{2}\)-arabinosyl and mono-O-C\(_{5}\)\(^{2}H\)\(_{2}\)-rhamnosyl units differ in mass and can be differentiated by MS, whereas the corresponding O-D\(_{5}\) anlogs have the same mass. The O-C\(_{5}\)\(^{2}H\)\(_{2}\) units on C-1 and C-4 of the arabinitol ("ald") end (Fig. 1D) arose because the arabinosyl residues in the arabinogalactan are furanoid and, when hydrolyzed and reduced, the hydroxyl groups at C-1 and C-4 are exposed. The presence of O-C\(_{5}\)\(^{2}H\)\(_{2}\) at positions other than at C-1 or C-4 of the pentitol indicates where other glycosyl residues were originally attached. Thus, the O-C\(_{5}\)\(^{2}H\)\(_{2}\) groups on C-1 and C-2 of the arabinitol ("ald") end (Fig. 1D) arose because the arabinosyl residues in the arabinogalactan are furanoid and, when hydrolyzed and reduced, the hydroxyl groups at C-1 and C-4 are exposed. The presence of O-C\(_{5}\)\(^{2}H\)\(_{2}\) at positions other than at C-1 or C-4 of the pentitol indicates where other glycosyl residues were originally attached. Thus, the O-C\(_{5}\)\(^{2}H\)\(_{2}\) group at C-5 of the arabinitol (Fig. 1D) established that this residue was substituted at C-5 in the arabinogalactan. The two arabinosyl residues ("a" and "b") in the sequence illustrated in Fig. 1D contain no O-C\(_{5}\)\(^{2}H\)\(_{2}\) groups. However, other oligoglycosyl alditols arising from the cleavage/modification of the methylated arabinogalactan did contain O-C\(_{5}\)\(^{2}H\)\(_{2}\) groups on residues "a" and "b" and provided considerable sequence information. Due to the complexity of the mixture of per-0-alkylated oligoglycosyl alditols, it was not possible to structurally characterize the components directly by GC/MS. Therefore, the per-0-alkylated oligoglycosyl alditols were separated by reverse-phase HPLC and the partially fractionated derivatives analyzed by GC/MS, \(^1\)H NMR, and, when appropriate, by glycosyl linkage composition analysis. Details of this methodology and the resulting data are presented in the Miniprint. The structures of 43 of the oligoglycosyl alditol fragments were established and are presented in Fig. 4.

Evidence for the Presence of Several Major Structural Motifs in Arabinogalactan—Examination of the structures 1 through 43 (Fig. 4) revealed that the majority of the oligoglycosyl alditols contain only Ara or Gal (rather than a mixture of both). The arabinogalactan appears to consist of arabinan and galactan regions, which is consistent with the conclusions arrived at from the analysis of enzymatic digests of the arabinogalactan (6). It was also possible to recognize oligosaccharide families and thereby to arrive at five major structural domains (Fig. 5).

\(^2\) The abbreviations used are: GC/MS, gas chromatography-mass spectrometry; EI/MS, electron intact-mass spectrometry; HPLC, high performance liquid chromatography; Ara, arabinosyl; Rha, rhamnosyl; d, \(^{1}\)H\(_{3}\), deuterium; Me, methyl; f, furanosyl; p, pyranosyl; SDS, sodium dodecyl sulfate; TMS, trimethylsilyl; BCG, Bacille Calmette-Guerin; RT, retention time; TMC, Trudeau mycobacterial collection; COSY, two-dimensional chemical shift correlated spectroscopy.
FIG. 4. Structures of the partially O-methylated, partially O-pentadeuterioethylated glycosyl alditols identified in this study. When not specified, the ring forms are furanoid. The nomenclature used is illustrated in the following example: 4,→5-Ara-(1→6)-Ara, is 2,3-di-O-methyl-5-pentadeuterioethyl arabinofuranosyl-(1→6)-2,3-di-O-methyl-1,4-di-O-pentadeuterioethyl arabinitol. A, structures of the nine partially O-methylated, partially O-pentadeuterioethylated monomannosyl alditols identified in this study. B, structures of the 11 partially O-methylated, partially O-pentadeuterioethylated diarabinosyl arabinosyl arabinotols arising from this study. C, structures of the four partially O-methylated, partially O-pentadeuterioethylated triarabinosyl arabinosyl arabinotols identified in this study. D, structures of the 12 partially O-methylated, partially O-pentadeuterioethylated oligogalactosyl galactitols identified in this work. The order of the 6-Gal and 5-Gal in 36 was not determined directly but was based on the structures of 27, 28, 31, and 32. The linkage of the internal residues in 33 and 34 was based on the structures of 29 and 30. E, structures of the seven partially O-methylated, partially O-pentadeuterioethylated oligoglycosyl alditols containing “mixed sugar” residues. The placement of the →5-Ara-(1→5)-Ara on C-5 and the →5-Gal on C-6 of the galactitol in 40 was not proven directly but rather was deduced from the structures of 29 and 39.
Individual per-O-alkyloligosaccharide alditols containing both arabinogalactan (see Fig. 9 and the text). All ring forms are furanoid characterized (Fig. 4E). These structures establish that the Ara and Galf residues, 39 and 40, were isolated and characterized. The linkage data in Table I indicate that 5-linked Ara residues are attached to C-5 of the 6-linked Gal/f residues and correspond to structural motif B (Fig. 5).

**The Presence of Rhamnosse and Mannose in the Arabinogalactan and Evidence for Other Minor Structures**—There have been reports of the presence of rhamnose in the cell wall of mycobacteria (17, 21), however, its occurrence in mycobacterial arabinogalactan had not been demonstrated. The isolation and characterization of oligosaccharides 41–43 (Fig. 4E) established that some of the 5-linked galactosyl residues are linked to C-4 of a rhamnosyl residue. Thus, the reducing terminus of the arabinogalactan molecule is occupied by a rhamnose residue.

Two oligosaccharides, 37 and 38, (Fig. 4E) in which t-Manp or t-Manp-(1→6)-Manp are attached to 5-Araf were also characterized. These are minor components, as mannopyranosyl residues account for <2% of the arabinogalactan (Table I). It is unlikely that these mannopyranosyl residues arise from contaminating arabinomannan (22, 23), and they probably represent minor variations on the major structure. In addition, oligosaccharides 2 and 3 (Fig. 5A), which are present in small amounts, do not fit any of the dominant structural motifs and may represent incomplete versions of structural motif A.

**Confirmation of the Absolute Configuration of the Ara and Gal Residues**—GC/MS analysis of the (CH₃)₂Si derivatives of both R(-) and S(+)-2-butylyglycosides and appropriate standards (24) showed the Ara to be >98% D and the Gal to be >97% D and confirmed that both Ara and Gal in the arabinogalactan of *M. tuberculosis* are D (11). This also established that quantitatively minor amounts of L-sugars are not present and therefore do not account for immunological activity.

**Assignment of Acarabinogalactan by 'H NMR Analysis of the Per-0-Alkylated Glycosyl Alditol Fragments**—All of the major glycosyl linkages present in the arabinogalactan were represented in the per-0-alkyloxy glycosyl alditols analyzed by 'H NMR (Table II). The large coupling constant (J₁,₂) of 4.6 Hz for the t-Araf residue in 1 shows that the t-Araf units are β-linked. Analysis of 4 by 'H NMR shows that the 5-Araf residues are α-linked and that the 2-Araf residues (5 and 6) are α-linked, regardless of whether they are attached to C-3 or C-5 of the 5,6-branched Araf. The β configuration of the t-Araf and α configuration of the 2-Araf units were confirmed by 'H NMR analysis of 11. 'H NMR analyses of 9 showed that the 3,5-linked Araf residues are α-linked. 'H NMR analyses of 28, 31, and 36 demonstrated that all of the galactosyl residues are β-linked.

**Determination of the Anomeric Configurations of the Glycosyl Residues in the Arabinogalactan**—by ¹³C NMR—The ¹³C NMR spectrum of the entire arabinogalactan, solubilized by base treatment of whole mycobacterial cell walls, is shown in Fig. 6. The majority of the C-1 signals appear between δ108 and δ109, corresponding to α-Araf/β-Gal residues (25–28). The two signals at δ106.8 and δ106.6 were assigned to C-1 of α-linked 2-Araf residues. This assignment is possible because substitution at C-2 but not at C-3 or C-5 causes the chemical shift of C-1 to move upfield by 1–2 ppm (26, 26, 29). Since the only 2-linked glycosyl residues in the arabinogalactan are 2-Araf (Table I), the signals at δ106.8 and δ106.6 must arise from the C-1 of the 2-Araf residues. The assignment was

---

3 The non-ring oxygens on C-1 and C-2 of both α-Araf and β-Gal are trans, and thus α-Araf and β-Gal give similar chemical shift values for C-1 and H-1 and similar J₁,₂ coupling constant values. For α-Araf/β-Gal, (trans) C-1 has its resonance at about δ108 and J₁,₂ is approximately 3 Hz or less (25–29). For β-Araf/α-Gal (cis), the chemical shift of C-1 is δ102 and the J₁,₂ is 4 Hz or greater (25–29). The value of the chemical shift of II-1 cannot be used to distinguish α- or β-Araf or Gal.
TABLE II

| Peralkylated oligoglycosyl alditol<sup>a</sup> | Chemical shift | J<sub>1,3</sub> | Assignment |
|----------------------------------------|---------------|------------|------------|
| Ara-(1→2)-Ara                           | 5.28          | 4.6        | t-β-Araf   |
| Ara-(1→3)-Ara                           | 5.08          | Br.s<sup>b</sup> | 2-α-Araf   |
| Ara-(1→5)-Ara                           | 4.98          | Br.s       | 2-α-Araf   |
| Ara-(1→5)-Ara                           | 5.00          | Br.s       | 5-α-Araf   |
| Ara-(1→2)-Ara-(1→3)-Ara                | 5.18          | 4.1        | t-β-Araf<sup>c</sup> |
| Ara-(1→2)-Ara-(1→3)-Ara                | 5.21          | Br.s       | 2-α-Araf<sup>c</sup> |
| Gal-(1→6)-Gal                           | 5.03          | 1.8        | 5-β-Gal<sup>d</sup> |
| Gal-(1→6)-Gal-(1→5)-Gal                | 4.9           | 1.8        | 6-β-Gal<sup>d</sup> |
| Gal-(1→6)-Gal-(1→5)-Gal-(1→0)-Gal      | 5.00          | Br.s       | 5-β-Gal<sup>d</sup> |
| Gal-(1→5)-Gal-(1→6)-Gal-(1→5)-Gal-(1→0)-Gal | 5.25  | Br.s       | 6-β-Gal<sup>d</sup> |

<sup>a</sup> See Fig. 4.
<sup>b</sup> Br.s = broad singlet.
<sup>c</sup> Assignment of β to the t-Araf and α to the 2-Araf was based on analysis of compounds 1, 5, and 6.
<sup>d</sup> These assignments may be reversed.

Fig. 6. The 13C NMR spectrum of the base-solubilized cell wall arabinogalactan of M. tuberculosis. The chemical shifts of resonances discussed in the text are labeled.

confirmed using 13C/<sup>1</sup>H correlation NMR and COSY NMR to trace the connectivity of the C-1s at δ106.8 and δ106.6 to the glycosyl substituted carbons at δ88.2 and δ87.9. The 13C/1H correlation spectrum (Fig. 7) showed connectivity between C-1 resonances at δ106.8 and δ106.6 and corresponding H-1 resonances at δ5.16 and δ5.25, respectively. The small coupling constant values (J<sub>1,2</sub> = 2.3 and 2.5 Hz, respectively) measured from the two-dimensional-COSY spectrum (Fig. 8) confirmed (27, 29) that the C-1 resonances at δ106.8 and δ106.6 were from α-Araf. Furthermore, the COSY 1H/1H correlation (Fig. 8) showed that the chemical shift of the H-2 was at δ4.19 for both residues. The 13C/1H correlation (data not presented) then allowed the resonances of the corresponding C-2s at δ88.2 and δ87.9 to be identified, and their chemical shift value confirmed glycosyl substitution at position 2 of the Ara (Table I and Ref. 29).

The signals at δ101.9 and δ101.8 in the 13C NMR spectrum of the arabinogalactan (Fig. 6) are consistent with the presence of β-Araf/α-Gal<sup>3</sup> residues (25, 26, 28). Pyranosyl residues would also give signals in this chemical shift region; however, the glycosyl-linkage analysis data (Table I), our previous data (17), and the structures of 1–36 and 39–43 show that all of the Ara and Gal residues are furanoid. The assignment of the signals at δ101.9 and δ101.8 to the C-1 of β-Araf/α-Gal is therefore unequivocal. The 1H NMR data (Table II) clearly established that the t-Araf residues are in
the β configuration, and, thus, the two C-1 signals at ~δ1012 must result from α-β-Araf. This assignment is supported by the signal at δ64.1 (Fig. 6) which corresponds to the C-5 (primary alcohol carbon) signal of α-β-Araf (δ64.2 for the methyl glycoside, Ref. 28) but not to that of α-α-Araf (δ62.4 for the methyl glycoside, Ref. 28). The signal at δ64.1 cannot result from any other primary alcohol carbon in the arabinogalactan; the primary alcohol carbons of 5- and 3,5-Araf and 6- and 5,6-Galf have chemical shifts at ~δ70 (25), and the primary alcohol carbons of 2-linked-α-Araf (29) and 5-linked-β-Galf (28) have their resonances at ~δ62.

13C NMR Confirms Structural Motif A—The different chemical shifts (δ106.6 and δ106.8) of the C-1 resonances of the two 2-linked-α-Araf residues (Fig. 6) of motif A are due to slight differences in their chemical environments. One 2-linked-α-Araf residue is attached to C-3 and the other to C-5 of the 3,5-linked-α-Araf unit. In addition, the C-2 resonances of the 2-linked-α-Araf are clearly observable and again are at two slightly different chemical shifts (δ88.2 and δ87.9), and, finally, the two C-1 resonances of the two t-β-Araf are also seen at δ101.9 and δ101.8. The intensities of the resonances at δ81.9, 88.2, 101.8, 101.9, 106.6, and 106.8 are approximately equal and are consistent with them being part of the same structure, structural motif A.

DISCUSSION

The arrangement of the arabinosyl and galactosyl residues within the arabinogalactan of the cell wall of mycobacteria has, until now, not been studied in detail. This arabinogalactan is unusual in that, unlike most bacterial polysaccharides (30, 31), it is not composed of an oligosaccharide repeating unit. The presence work confirms (6) the presence of distinctive galactan and arabian segments and established that the arabinogalactan is composed of a few distinct, defined structural motifs.

Structural motif A is the most unexpected and significant. It has been shown that arabinosyl residues are responsible for the antigenicity of arabinogalactan and that serological activity resides largely in a fraction containing 2-linked arabinosyl residues (6, 14). Thus, it is logical to speculate that part, or all, of structural motif A is the major humoral immunological epitope of arabinogalactan and, consequently, of whole mycobacteria (18). Monoclonal antibodies raised against lipopolyarabinomannan (22) also react with the purified cell walls, suggesting an arabinose-containing epitope common to lipopolyarabinomannan and arabinogalactan.

Structural motifs B and C account for the bulk of the internal portions of the arabian segments of the arabinogalactan and are consistent with previous studies (6) that had established the presence of regions containing 5-linked α-arabinofuranosyl residues. Structural motif D, composed of alternating 6-linked and 5-linked galactofuranosyl residues, is supported by the existence of the diaccharide, 6 O β D galactofuranosyl-D-galactose, among the products of the degradation of arabinogalactan (15). The occurrence of structural motif D was unexpected since the molar ratio of 5-Galf to 6-Galf, including the 5,6-branched Galf, was about 1.5. No detectable amounts of 5-Galf/1→6-Galf or 6-Galf/1→5-Galf-(1→6)-Galf were revealed by close examination of the mass spectra of the individual per-O-alkylated oligogalactoeyl alditols. The possibility that the 2,3,5-tri-O-methyl galactose (6-Galf) released by acid hydrolysis was preferentially degraded was examined by subjecting the permethylated arabinogalactan to various hydrolytic conditions, but the ratio of 5-Galf to 6-Galf remained constant. In addition, no new scissorable sugars were revealed by mild acid hydrolysis. The evidence, from oligogalactosyl alditol analysis, for structural motif D is strong, but the existence of an undetected arrangement involving 5-Galf is still possible.

FIG. 9. An illustration of some of the ways in which the major structural motifs of the cell wall arabinogalactan may be assembled. The purpose of this illustration is to stimulate ideas on how the structural units may be joined together. This is not intended to be a final representation of the structure of arabinogalactan.

1 Incidentally, the presence of the two environmentally different 3,5-α-Araf residues in motifs A and B (Fig. 5 and 9) is evident in the 13C NMR spectrum of periodate oxidized and NaBH4 reduced arabinogalactan. The complex group of signals at δ108.9–110.3 is replaced by two signals at δ109.1 and δ108.7, corresponding to the periodate-stable 3,5-α-Araf. The signals for the C-1s of the oxidized/reduced residues present in acyclic acetals are shifted to ~δ104.

2 M. Daffe, P. J. Brennan, and M. McNeil, unpublished results.
a framework for the design of future experiments. The complete elucidation of the structure of the arabinogalactan will require the isolation of oligoglycosyl fragments encompassing branched and linear regions of both the arabinan and galactan. In addition, the complete structural characterization of the arabinogalactan requires that oligoglycosyl units (37 and 38 and 1 and 2, Fig. 4) that do not conform to the model proposed in Fig. 9 are accounted for. However, 37 and 38 are minor products and probably represent minor populations of arabinogalactan in which mannosyl or mannobiosyl rather than 2-linked arabinofuranosyl residues occupy non-reducing termini; indeed, mannose is a component of the arabinogalactan of some corynebacteria (16).

The proposed structures (Fig. 9) are also of value in determining where the mycolyl residues, that are part of the native mycolylarabinogalactan, are attached. It has been reported that the mycolic acids are attached to O-5 of arabinosyl residues (3, 4, 33-35), and it can be seen from Fig. 9 that position 5 of both the t-Araf and 2-Araf, both present in structural motif A, are available for such attachment. The diarabinosidyl mycolate (mycolyl-5-Araf-(1→3)-Araf), isolated by Markovits and Vilkas (33), is in accord with structural motif A only if the mycolic acid is attached to C-5 of the 2-Araf.

Considerable progress has recently been made in defining the individual, extractable entities of mycobacterial cell walls and relating these to aspects of the pathogenesis and immunogenesis of disease (18). However, the massive insoluble mycolylarabinogalactan peptide-glycan framework had defied even primary structural definition. Clearly, the structure of these molecules must be understood before their role in the refractoriness of mycobacteria to chemotherapy, the peculiar persistence of mycobacteria within the macrophage, and the propensity for disease recrudescence can be addressed.

Acknowledgments—We thank Marilyn Hein for preparing the manuscript, Carol Marander for the graphics, and Cynthia Bosic and Dr. Delphi Chatterjee for critical technical help and valuable scientific discussions.

REFERENCES

1. Draper, P. (1982) in The Biology of the Mycobacteria (Hattledge, C., and Stanford, J., eds) Vol. 1. pp. 9-52, Academic Press, New York
2. Petit, J. F., and Lederer, E. (1984) in The Mycobacteria: A Sourcebook. Part A (Kubica, G. P., and Wayne, L. H., eds) pp. 301-313, Marcel Dekker, New York
3. Azuma, I., Yamamura, Y., and Fukushima, K. (1968) J. Bacteriol. 96, 1885-1887
4. Amar-Nacach, C., and Vilkas, E. (1970) Bull Soc Chim Biol 52, 145-151
5. Melancon-Kaplan, J., Hunter, S. W., McNeil, M., Stewart, C., Modlin, R. L., Rea, T. H., Convit, J., Salgama, P., Mahra, V., Bloom, B. R., and Brennan, P. J. (1988) Proc. Natl. Acad. Sci. U. S. A 85, 1917-1921
6. Misaki, A., Seto, N., and Azuma, I. (1974) J. Biochem. (Tokyo) 76, 15-27
7. Miller, R. A., Harnisch, J. P., and Buchanan, T. M. (1984) Int. J. Legum. 52, 139-139
8. Kleinhenz, M. E., Ellner, J. J., Spagnuolo, P. J., and Daniel, T. M. (1981) J. Clin. Invest. 68, 153-162
9. Stacey, M., and Kent, P. W. (1948) Adv. Carbohydr. Chem. 3, 311-336
10. Misaki, A., and Yukawa, S. (1966) J. Biochem. (Tokyo) 59, 511-520
11. Azuma, I., Kimura, H., Niinaka, T., Aoki, T., and Yamamura, Y. (1968) J. Bacteriol. 95, 263-271
12. Kanetsuna, F. (1968) Biochim. Biophys. Acta 158, 130-149
13. Misaki, A., Ikawa, N., Kato, T., and Kotani, S. (1970) Biochim. Biophys. Acta 215, 405-408
14. Kotani, S., Kato, T., Matsuda, K., Kato, T., and Misaki, A. (1971) Biken J. 14, 379-387
15. Vilkas, E., Amar, C., Markovits, J., Viegenthart, J. F. G., and Kamerling, J. P. (1973) Biochim. Biophys. Acta 297, 425-435
16. Abou-Zeid, C., Volband, A., Michel, G., and Cocito, C. (1982) Eur. J. Biochem. 128, 363-370
17. McNeil, M., Walliner, S. J., Hunter, S. W., and Brennan, P. J. (1987) Carbohydr. Res. 166, 299-308
18. Brennan, P. J. (1989) Rev. Infect. Dis. 11 (Suppl. 2), S420-S430
19. McNeil, M., Davill, A., Aman, P., Franzen, L., and Albersheim, P. (1982) Methods Enzymol. 83, 3-45
20. Valet, B. S., Davill, A. G., McNeil, M., Robertson, B. K., and Albersheim, P. (1980) Carbohydr. Res. 79, 165-192
21. Haworth, N., Kent, P. W., and Stacey, M. (1948) J. Chem. Soc. 1111-1120
22. Hunter, S. W., Gaylord, H., and Brennan, P. J. (1986) J. Biol. Chem. 261, 12345-12351
23. Aspinall, G. O. (1985) in The Polysaccharides, Vol. 2, p. 345, Academic Press, New York
24. Gerwig, G. J., Kamerling, J. P., and Viegenthart, J. F. G. (1978) Carbohydr. Res. 62, 349-357
25. Bradbury, J. H., and Jenkins, G. A. (1984) Carbohydr. Res. 126, 125-156
26. Gorin, P. A. J., and Mazurek, M. (1976) Carbohydr. Res. 48, 171-186
27. Joseleau, J. P., Chambat, G., Vignon, M., and Barmoud, F. (1977) Carbohydr. Res. 59, 160-175
28. Gorin, P. A. J., and Mazurek, M. (1975) Can. J. Chem. 53, 1212-1222
29. Mizutani, K., Kasai, R., Nakamura, M., Tanaka, O., and Matsuura, H. (1989) Carbohydr. Res. 185, 27-58
30. Sutherland, I. W. (1982) Adv. Microbial Physiol. 23, 79-149
31. Kenna, L., and Lindberg, B. (1983) in The Polysaccharides (Aspinall, G. O., ed) Vol. 2, pp. 287-363, Academic Press, New York
32. Fujioka, M., Koda, S., and Morimoto, Y. (1985) J. Gen. Microbiol. 131, 1323-1329
33. Markovits, J., and Vilkas, E. (1969) Biochim. Biophys. Acta 192, 49-54
34. Azuma, I., and Yamamura, Y. (1962) J. Biochem. 50, 200-206
35. Azuma, I., and Yamamura, Y. (1963) J. Biochem. 53, 275-281
36. McNeil, M., Tsang, A. Y., and Brennan, P. J. (1987) J. Biol. Chem. 262, 2630-2635
37. Neelkj, K., Saito, H., and Hakomori, S. (1973) Arch. Biochem. Biophys. 155, 464-472
38. York, W. S., Davill, A. G., McNeil, M., Stevenson, J. T., and Albersheim, P. (1986) Methods Enzymol 118, 3-40
39. Fournet, B., Dhalluin, J.-M., Streek, G., Montreuil, J., Bosso, C., and Defaye, O. (1980) Anal. Biochem. 108, 35-96
40. Lennartss, G., Lundblad, A., Sjohlad, S., Svensson, S., and Ockerman, P. A. (1976) Biomed. Mass Spectrom. 3, 51-54
41. Hallgren, P., and Lundblad, A. (1977) J. Biol. Chem. 252, 1023-1033
42. Kochertov, N. K., and Chizhov, O. S. (1966) Adv. Carbohydr. Chem. 21, 39-92
43. Sharp, J. K., and Albersheim, P. (1964) Carbohydr. Res. 128, 193-202
Arabinogalactan of M. tuberculosis

EXPERIMENTAL PROCEDURES

Growth of M. tuberculosis and Production of cell walls.

M. Tuberculosis TMC 9741 (®) Erdman) was grown for two months on Sauton's medium in 2.4-liter Fernbach flasks (22). The bacteria were killed by autoclaving at 121°C for 1.5 hr, centrifuging, and freeze-drying for 10 days at -30°C. The bacterial harvests in 100 ml of 50 mM Tris (pH 7.2) containing 0.1% Tween 20 were centrifuged at 10,000 x g for 15 min. The resulting cell suspension was sonicated using a Heat Systems-Ultrasonics, Inc., W-180W sonicator (Farmingley, N.Y.). The sonicator had a 0.7-mm inch probe and was used at full power with 150-watt output power for 15 min. Microscopic examination revealed that the vast majority of bacteria were broken and no mycobacteria could be seen. Cell walls were recovered by centrifugation at 20,000 x g for 30 min. Centrifugation was repeated until the supernatant was clear. The pellet was washed with chilled 0.85% saline solution and finally lyophilized. These cell walls, weighing 0.8 g, were rehydrated with 70% ethanol in water at 56°C. The suspension, which contained the cell wall arabinogalactan, was precipitated with acetonecold, centrifuged, and dissolved in saline to reconstitute samples. The precipitate formed was removed by centrifugation. The supernatant was made to 80% with ethanol and kept overnight at 4°C to precipitate the polysaccharide. The pellet was recovered by centrifugation and reconstituted with approximately 99% of arabinogalactan which is approximately 20% of arabinogalactan in the original cell wall.

MS Analysis

The sonicated arabinogalactan solution was suspended in 1 ml of [NH₄]OH and the nonsedimented material removed. Acetonitrile (2%) was added as an internal mass standard. The per-2,3,4-tri-O-methylarabinosyl-α-L-arabinofuranosyl (2) and 3,4,6-tri-O-methylarabinofuranosyl (3) arabinogalactans were analyzed by reverse-phase liquid chromatography (HPLC) and mass spectrometry as described (18). The results were expressed as absorbance (A) at 510 nm. The mass spectrometry was used to confirm the structure described for the per-2,3,4-tri-O-methylarabinosyl-α-L-arabinofuranosyl (2) and 3,4,6-tri-O-methylarabinofuranosyl (3) arabinogalactans.

N-TOOC/MS Analysis of the per-2,3,4-tri-O-methylarabinosyl-α-L-arabinofuranosyl (2) arabinogalactan.

An aliquot of per-2,3,4-tri-O-methylarabinosyl-α-L-arabinofuranosyl (2) arabinogalactan was partially hydrolized in 1 M TFA for 1 hr at 70°C, reduced with NaBH₄, and analyzed by HPLC. The TFA was added to the sample at 1:1 ratio and the mixture was then incubated at 70°C for 1 hr. The reaction mixture was then slowly neutralized with an equal amount of water and the mixture dialyzed against water. The dialyzed mixture was lyophilized and the per-2,3,4-tri-O-methylarabinosyl-α-L-arabinofuranosyl (2) arabinogalactan was reconstituted with an acetone insoluble pellet which was described (18).

Preparation and Analysis of Partially 3,4,6-Tri-O-Methylated, Partially O-Linked, Partially N-Linked arabinogalactan.

arabinogalactan fractions were hydrolyzed, reduced and acetylated as described (18). All drying steps were conducted at room temperature to avoid evaporation of the volatile, highly acetylated alditols. The samples were analyzed by N-TOOC/MS as described above. The temperature program was 50°C for 5 min, followed by a 20°C/min rise to 280°C, where it was held for 1 min. The ion source was maintained at 70 eV. The samples were injected as trimethylsilyl (TMS) ethers. The mass spectrometer was set to scan from 50 to 800 amu with 0.11 scans per sec. Mass spectra were acquired in positive ion mode, with the instrument interfaced to a Hewlett-Packard 5995A mass spectrometer.

The ion source was maintained at 70 eV. The samples were injected as trimethylsilyl (TMS) ethers. The mass spectrometer was set to scan from 50 to 800 amu with 0.11 scans per sec. Mass spectra were acquired in positive ion mode, with the instrument interfaced to a Hewlett-Packard 5995A mass spectrometer.

The interpretation of the mass spectra of the per-2,3,4-tri-O-methylarabinosyl-α-L-arabinofuranosyl (2) arabinogalactan is shown in the figure. For the TMS derivative of 2, the TMS (m/z 153) (254) ion is reduced with 40% methanol (M) to form an internal residue, which is then further reduced with (M) to give a terminal residue of (m/z 127) (28). The terminal residue is further reduced with 25% acetic acid to form a terminal residue of (m/z 109) (29). The subsequent step is formic acid (m/z 93) (30) and 2-hydroxypropionic acid (m/z 97) (30). The subsequent step is formic acid (m/z 93) (30) and 2-hydroxypropionic acid (m/z 97) (30). The subsequent step is formic acid (m/z 93) (30) and 2-hydroxypropionic acid (m/z 97) (30). The subsequent step is formic acid (m/z 93) (30) and 2-hydroxypropionic acid (m/z 97) (30). The subsequent step is formic acid (m/z 93) (30) and 2-hydroxypropionic acid (m/z 97) (30). The subsequent step is formic acid (m/z 93) (30) and 2-hydroxypropionic acid (m/z 97) (30). The subsequent step is formic acid (m/z 93) (30) and 2-hydroxypropionic acid (m/z 97) (30).
Arabinogalactan of M. tuberculosis

The presence of G[4]A[5]G groups on C[6] or C[7] on nonreducing terminal galactosyl residues were also readily distinguished by mass spectrometry. The G[4]A[5]G group at C[6] but not at C[7] is readily lost from the ion (A) arising from the nonreducing galactosyl residue. As ions, regardless of the G[4]A[5]G; group is at C[6] or C[7], have a m/z value of 320, and both ions remain to produce an A ion with an m/z value of 262. The A ion with an m/z value of 262, in G[4]A[5]G at C[6] eliminates pentadentateesterol on to produce an A ion at m/z 155, while the A ion containing the G[4]A[5]G group at C[6] eliminates pentadentateesterol and produces an A ion at m/z 184. Composition 28, 31 and 36 (Fig. 4) all gave an ion at m/z 155 while compound 27 and 34 (Fig. 4) gave an ion at m/z 184.

The position of substitution on the alditol could be determined for some of the mono- and digalactosyl alditols by ions produced by cleavage between carbons of the alditol. Compound 6 (Fig. 4) shows an ion at m/z 178 (cleavage between C[4] and C[5] of the alditol with the charge on C[3]), while compound 5 (Fig. 4) produced no ion. To identify the pyranose residues present in the pyranose alditol compounds, 27 and 24 (Fig. 4), the presence of an intense ion at m/z 88, a radical cation containing C[4] and C[5] and their attached G[4]A[5]G groups, was sought. The intensity of the ion at m/z 88 is weak when only furanose alditol residues are present. Many of the compounds isolated by HPLC were also characterized by GC/MS of the partially acetylated, partially methylated, partially pentadentateesterolyzed (alditol) produced by hydrolysis, reduction and acetylation, as illustrated in Fig. 5 for compound 11. These analyses confirmed structural features deduced by direct GC/MS analysis and also allowed elucidation of features not readily obtainable by direct GC/MS analysis. The most important information revealed by this analysis was the nature of the linkage of the internal galactosyl residues in di- and tri-galactosyl alditols, as illustrated in Fig. 5, compound "a," at 3.91 min, 1,4-di-O-acetyl-2,5,5-tri-O-methyl arabinosyl, and compound "b," at 4.47 min, 1,2,4-tri-O-acetyl-5,5-di-O-methyl arabinosyl, proving that the internal residue in 2-derivative, not 5-derivative. Compound "c," at 5.04 min, 1,5-b-D-arabinofuranosyl-2,5,5-tri-O-methyl arabinosyl, thus proving that the linkage is 1-2 instead of 1-5, 3-1 linked to the arabinosyl residue. Therefore, the structural features of compound 11 established by direct GC/MS on the intact compound were confirmed, and, most importantly, the internal was shown to be 1-2-deriv. Likewise, the 6-linked and 5-linked internal galactosyl residues in compounds 12 and 13 (Fig. 4) were distinguished in this fashion as were components of other compounds where the need arose.

Analytical Data on the Par-Acetylated Digalactosyl Aalditol.

In the analytical data listed below, the GC retention time quoted are those obtained with GC program 1: 1% on glycosyl linkage composition data are given, it is because the compounds in question co-eluted from the HPLC with other compounds of such structures as to render glycosyl linkage analysis assignments ambiguous.

**Compound 1:**
- GC R: 8.6 min, HPLC R: 17.3 min, Yields ions at m/z 101, 143 (A), 155 (A), 175 (A), 230 (A), 290 (A), Glycosyl linkage composition: 1,4-di-O-acetyl-1,2,5-tri-O-methyl arabinosyl, and 1,4-di-O-acetyl-2,3,5-tri-O-methyl, 1,4-di-O-acetyl-2,3,5-tri-O-methyl arabinosyl in the ratio of 11:1.

**Compound 2:**
- GC R: 8.9 min, HPLC R: 18.9 min, Yields ions at m/z 101, 143 (A), 175 (A), 230 (A), 290 (A), 351 (Alditol cleavage).

**Compound 3:**
- GC R: 9.1 min, HPLC R: 20.7 min, Yields ions at m/z 122, 143 (A), 180 (A), 194 (A), 230 (A), 290 (A), 351 (Alditol cleavage) and 351 (Alditol cleavage). Glycosyl linkage composition: 1,4-di-O-acetyl-2,3,5-tri-O- methyl-2-pentadentateesterol arabinosyl and 2,3,5-di-O-acetyl-2,3,5-tri-O-methyl-2-pentadentateesterol arabinosyl in the ratio of 11:1.

**Compound 4:**
- GC R: 9.3 min, HPLC R: 21.4 min, Yields ions at m/z 122, 143 (A), 194 (A), 230 (A), 290 (A), 351 (Alditol cleavage) and 351 (Alditol cleavage) Glycosyl linkage composition: 1,4-di-O-acetyl-1,2,5-tri-O- methyl-2-pentadentateesterol arabinosyl and 2,3,5-di-O-acetyl-2,3,5-tri-O-methyl-2-pentadentateesterol arabinosyl in the ratio of 11:1.

**Compound 5:**
- GC R: 9.7 min, HPLC R: 22.2 min, Yields ions at m/z 120, 143 (A), 194 (A), 230 (A), 290 (A), 351 (Alditol cleavage) and 351 (Alditol cleavage). Glycosyl linkage composition: 1,4-di-O-acetyl-1,2,5-tri-O-methyl-2-pentadentateesterol arabinosyl and 2,3,5-di-O-acetyl-2,3,5-tri-O-methyl-2-pentadentateesterol arabinosyl in the ratio of 11:1.

**Compound 6:**
- GC R: 10.1 min, HPLC R: 26.5 min, Yields ions at m/z 120, 143 (A), 194 (A), 230 (A), 290 (A), 351 (Alditol cleavage) and 351 (Alditol cleavage).

**Compound 7:**
- GC R: 15.0 min, HPLC R: 34.5 min, Yields ions at m/z 101, 143 (A), 175 (A), 230 (A), 290 (A), 351 (Alditol cleavage). Glycosyl linkage composition: 1,4-di-O-acetyl-2,3,5-tri-O-methyl arabinosyl, 1,4-di-O-acetyl-2,3,5-tri-O-methyl arabinosyl, 1,4-di-O-acetyl-2,3,5-tri-O-methyl arabinosyl and 1,4-di-O-acetyl-2,3,5-tri-O-methyl arabinosyl in the ratio of 11:1.

**Compound 8:**
- GC R: 14.5 min, HPLC R: 27.5 min, Yields ions at m/z 101, 143 (A), 175 (A), 230 (A), 290 (A), 351 (Alditol cleavage). Glycosyl linkage composition: 1,4-di-O-acetyl-1,2,5-tri-O-methyl arabinosyl, 1,3,4,5,6-pentadentateesterol arabinosyl in the ratio of 11:1.
Arabinogalactan of M. tuberculosis

Fig. 7. A portion of the carbon hydrogen two-dimensional correlation spectrum of arabinogalactan. The correlation of the carbon at δ 100.8 and δ 104.6 to hydrogen at δ 3.16 and δ 3.19 are marked A and B, respectively.
Predominant structural features of the cell wall arabinogalactan of Mycobacterium tuberculosis as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by 1H and 13C NMR analyses.

M Daffe, P J Brennan and M McNeil

*J. Biol. Chem.* 1990, 265:6734-6743.

Access the most updated version of this article at [http://www.jbc.org/content/265/12/6734](http://www.jbc.org/content/265/12/6734)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/265/12/6734.full.html#ref-list-1](http://www.jbc.org/content/265/12/6734.full.html#ref-list-1)