Structural features of lipoarabinomannan from Mycobacterium bovis BCG. Determination of molecular mass by laser desorption mass spectrometry

Anne Venisse, Jean-Marc Berjeaud, Pierre Chaurand, Martine Gilleron, Germain Puzo

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1993 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 268, No. 17, Issue of June 15, pp. 12401-12411, 1993
Printed in U.S.A.

Structural Features of Lipoarabinomannan from *Mycobacterium bovis* BCG

DETERMINATION OF MOLECULAR MASS BY LASER DESORPTION MASS SPECTROMETRY*

(Received for publication, November 20, 1992, and in revised form, March 1, 1993)

Anne Venisse‡, Jean-Marc Berjeaud§, Pierre Chaurand¶, Martine Gilleron‡, and Germain Puzo‡‖

From the ‡Laboratoire de Pharmacologie et de Toxicologie Fondamentales du Centre National de la Recherche Scientifique, Département Glycoconjugués et Biomembranes, 118, route de Narbonne, 31062 Toulouse Cédex, France, the §Laboratoire de Pharmacologie et de Toxicologie Fondamentales du Centre National de la Recherche Scientifique, Département Biologie Structurale et Ingénierie des Protéines, 118, route de Narbonne, 31062 Toulouse Cédex, France, and the ¶Institut de Physique Nucléaire, 91406 Orsay Cédex, France

It was recently shown that mycobacterial lipoarabinomannan (LAM) can be classified into two types (Chatterjee, D., Lowell, K., Rivoire B., McNeil M. R., and Brennan, P. J. (1992) *J. Biol. Chem.* 267, 6234–6239) according to the presence or absence of mannosyl residues (Man*p*) located at the nonreducing end of the oligoarabinosyl side chains. These two types of LAM were found in a pathogenic *Mycobacterium tuberculosis* strain and in an avirulent *M. tuberculosis* strain, respectively, suggesting that LAM with Man*p* characterizes virulent and "disease-inducing strains."

We now report the structure of the LAM from *Mycobacterium bovis* Bacille Calmette-Guérin (BCG) strain Pasteur, largely used throughout the world as vaccine against tuberculosis. Using an up-to-date analytical approach, we found that the LAM of *M. bovis* BCG belongs to the class of LAMs capped with Man*p*. By means of two-dimensional homonuclear and heteronuclear scalar coupling NMR analysis and methylation data, the sugar spin system assignments were partially established, revealing that the LAM contained two types of terminal Man*p* and 2-*O*-linked Man*p*. From the following four-step process: (i) partial hydrolysis of deacylated LAM (dLAM), (ii) oligosaccharide derivatization with aminobenzoic ethyl ester, (iii) HPLC purification, (iv) FAB/MS-MS analysis; it was shown that the dimannosyl unit α-D-Man*p*-(1→2)-α-D-Man*p* is the major residue capping the termini of the arabinan of the LAM.

In this report, LAM molecular mass determination was established using matrix-assisted UV-laser desorption/ionization mass spectrometry which reveals that the LAM molecular mass is around 17.4 kDa. The similarity of the LAM structures between *M. bovis* BCG and *M. tuberculosis* H37Rv is discussed in regard to their function in the immunopathology of mycobacterial infection.

The so-called lipoarabinomannan (LAM),[1] ubiquituously present in the *Mycobacterium* genus, is a major component of mycobacterial cell walls (1). The structure and the antigenicity of LAM have been investigated extensively in the past by Azuma *et al.* (2) and Misaki *et al.* (3) whose works were conducted on arabinomannans (AMs) derived from mycobacterial cell walls by vigorous alkaline extraction. It was established that these AMs, whatever the investigated mycobacterial species (*Mycobacterium bovis* BCG, *Mycobacterium tuberculosis* H37Rv, etc), are composed of a backbone of (1→6)-linked D-Man*p* units to which are attached short side chains of (1→2)-linked D-Man*p* residues and of (1→5)-linked D-Ara*f* residues mostly present in the side chains. These arabinosyl side chains were shown to be the main antigenic determinants of the AMs.

Weber and Gray then demonstrated that "arabinomannan" from *Mycobacterium smegmatis* is phosphorylated and contains lactate and succinate residues (4). More recently, Hunter *et al.* (5) obtained native LAM from *M. tuberculosis*, also found to be phosphorylated and acylated by palmitoyl and tuberculostearoyl residues. A model was only recently proposed for this molecule (6) in which the internal mannan segment, basically similar to that proposed by Misaki *et al.* (3), formed a mannan "core" to which an arabinan chain is attached with oligoarabinosyl side chains. Moreover, it was also established that the reducing end is composed of a phosphatidylinositol unit esterified by the two fatty acids. It was assumed that this unit allowed the anchorage of LAM in the membrane (7). Thus, LAM from *M. tuberculosis* is a multiglycosylated form of the well-known mycobacterial mannosylphosphatidylinositols (8).

Native LAM from *M. tuberculosis* seems to be a key factor in host-pathogen relationships. It has been shown to be in-

* This work was supported by Grants 6718 from the Association pour la Recherche contre le Cancer and RECH/9000827 from the Region Midi-Pyrénées. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence and reprint requests should be addressed. Tel.: 61-33-59-12; Fax: 61-33-58-86.

[1] The abbreviations used are: LAM, lipoarabinomannan; dLAM, deacylated LAM; LM, lipomannan; dLM, deacylated LM; BCG, Bacille Calmette-Guérin; LPS, lipopolysaccharide; PAGE, polyacrylamide gel electrophoresis; UV-LDI, UV-laser desorption/ionization; TFA, trifluoroacetic acid; TOF, time-of-flight; GC/MS, gas chromatography-mass spectrometry; ^{1}H or ^{13}C NMR, proton or carbon nuclear magnetic resonance; DEPT, distortionless enhancement by polarization; COSY, ^{1}H-^{1}H correlated spectroscopy; HMBC, heteronuclear multiple bond correlation spectroscopy; HMQC, heteronuclear multiple quantum correlation spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; HPLC, high performance liquid chromatography; ABEE, aminobenzoic ethyl ester; FAB/MS, fast atom bombardment/mass spectrometry; MIKE, mass analysis ion kinetic energy; TNF, tumor necrosis factor; Man*p*, mannopyranose, Ara*f*, arabinofuranose, t, terminal; 2D, two-dimensional; 1D, one-dimensional.

volved in tuberculosis as a major B-cell immunogen (9), in suppression of antigen-dependent proliferation (10, 11), perhaps through inhibition of protein-antigen processing by antigen-presenting cells (12), in inhibition of γ-interferon activation of macrophages (13, 14), and in stimulation of tumor necrosis factor (TNF) release from macrophages (15, 16). According to Brennan *et al.* (17), by the evidence of interaction of LAM with *Mycobacterium*-infected host cells, it was suggested that LAM is a mycobacterial virulence factor.

More recently, Chatterjee *et al.* (18) reported a different terminal structural arrangement in LAM from the virulent (Erdman) strain of *M. tuberculosis* compared to the already described basic structure of LAM in an attenuated strain of *M. tuberculosis*, the difference being correlated with a functional difference (19). The majority of the terminal arabinan motifs in LAM from the Erdman strain are extensively substituted with either a single mannosyl residue, a dimannoside, or a trimannoside. Thus, Brennan and co-workers (18, 19) assumed that the Man-capped LAM is typical of disease-inducing mycobacterial strains especially, whereas the basic Ara-terminated kind is synonymous with "attenuated or rapidly growing strains."

In the present report, the structures of the LAMs from a virulent strain, *M. tuberculosis* H37Rv, which causes human tuberculosis, and from an attenuated strain of *M. bovis: M. bovis* BCG, which is used as a vaccine against tuberculous infection, were investigated. Structural differences between these two molecules may involve functional differences that could contribute to the knowledge of the molecular basis of the protective effect of *M. bovis* BCG or even the pathogenicity of *M. tuberculosis*.

Our study was first undertaken to determine the molecular mass of these LAMs by mass spectrometry thanks to a newly developed version of laser desorption: matrix-assisted UV-laser desorption/ionization (UV-LDI). Through the process of 2D homonuclear and heteronuclear scalar coupling NMR analysis of the whole molecule and the characterization of derivatized oligosaccharides by HPLC and FAB-MS after mild acid hydrolysis of LAM, a similar structure for the LAMs from the attenuated and virulent strains was reported in which arabinan side chains were substituted mostly with a dimannosyl residue. The implications of these structural observations will be discussed in the role of LAM as a virulence/pathogenicity factor.

MATERIALS AND METHODS

Preparation of LAM—Lipoarabinomannans (LAM) from two mycobacteria, *M. bovis* BCG strain Pasteur and *M. tuberculosis* H37Rv, were used. Briefly, the cells were delipidated using chloroform/methanol, 1:1, then refluxed in 50% (w/v) ethanol. The ethanol extract was concentrated, incubated in 2% Triton X-100, and dialyzed against 2% Triton X-100. This fraction was then extracted with chloroform/methanol, 2:1. The water-soluble fraction was freeze-dried, treated with trypsin, dialyzed against water, and freeze-dried again. LAM and LM were precipitated with 85% ethanol using NaCl as precipitant. These two acylated polysaccharides were tentatively separated onto a gel permeation Bio-Gel P-100 column (16 × 600 mm) eluted with 0.1 N CH_3COOH. All the fractions from the purification were analyzed by polyacrylamide gel electrophoresis. LAM was shown as a broad band located at an apparent molecular mass of 30–35 kDa and was still contaminated with LM at a molecular mass of 10–20 kDa, as they were described (12).

The sample containing LAM and LM was deacylated in 0.1 N NaOH, 2 h, 37 °C, neutralized with CH_3COOH, dialyzed, and loaded onto a permeation Bio-Gel P100 column (16 × 600 mm) eluted with 0.1 N CH_3COOH. Samples were assayed by gas-liquid chromatography. Analysis showed 100% carbohydrate content and a molar ratio of arabinose to mannose of 1.5:1 for dLAM from *M. bovis* BCG and from *M. tuberculosis* H37Rv. The fractions of dLAM and dLM were pooled and freeze-dried.

Electrophoresis—Gel electrophoresis of samples was performed on 15 × 9 cm gels according to the method of Laemmli (20). The separating gel was composed of 12.5% acrylamide cross-linked with 0.8% bisacrylamide. A 6% acrylamide stacking gel was layered above the separating gel. Individual lanes were loaded with 5 μg of sample and electrophoresed at a constant current of 20 mA.

Following electrophoresis, gels were silver-stained according to the method developed by Ansorge (21), with polysaccharide-specific periodic acid oxidation step after fixation, according to Tsai and Frasch (22).

Methylation—Methylation which used solid sodium hydroxide as the base was conducted according to a modified procedure from Ciucanu and Kerek described in Ref. 23. This procedure is experimentally easier and yields a cleaner product than the procedure from Hakomori. To the dried carbohydrate sample (1 mg) in a Teflon-lined screw-capped tube were added anhydrous dimethyl sulfoxide (0.3 ml) and finely powdered NaOH (two pellets). The mixture was sonicated for 15 min, and 0.1 ml of methyl iodide was added. After another sonication (15 min), the reaction mixture was placed at 25 °C for 1.5 h. The reaction was then quenched by the addition of about 1 ml of water, and sodium thiosulfate was necessary to neutralize I_2. Permethylated carbohydrates were extracted with chloroform. The chloroform layer was washed with water and dried under a stream of nitrogen. Two cycles of methylation are necessary to ensure a complete methylation of the polysaccharide.

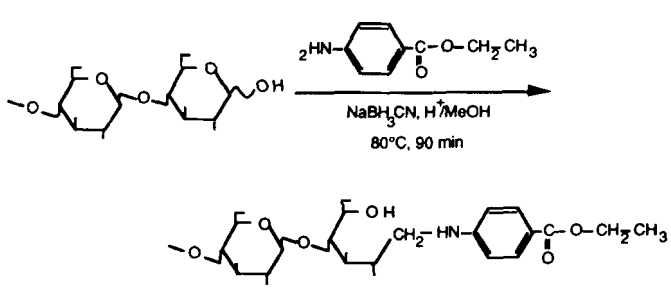
Glycosyl Linkage Composition—Fully methylated polysaccharides are often not soluble in hot aqueous solutions. For this reason, hydrolysis for 5 h with 2 N TFA at 100 °C was performed to obtain a complete hydrolysis. The methylated sugars were reduced with $NaBH_4$, 10 mg/ml, in NH_4OH, 1 M/C_2H_5OH, 1:1) freshly prepared. After 1 h at room temperature, the reaction was stopped by neutralization with CH_3COOH. The mixture was dried under N_2 twice with CH_3OH, 10% CH_3COOH and twice again with CH_3OH only. To peracetyl alditols, 50 μl of acetic anhydride were added and, after 1 h at 120 °C, 0.3 ml of H_2O and Na_2CO_3 were added to stop the reaction. The alditol acetates of methylated sugars were extracted with $CHCl_3$, dried, and solubilized in hexane before injection in GC and GC/MS. Quantitative determination of the glycosidic linkage was calculated according Sweet *et al.* (24).

Production of Oligosaccharides from LAM—dLAM (1 mg) was treated with 0.1 N HCl, 90 °C, 15 min and 0.075 N HCl, 75 °C, 15 min, and the hydrolysates were applied, respectively, to Bio-Gel P4 in 0.1 N acetic acid. The voided material contained, upon methanolysis and GC, Ara and Man in the ratio 1:20, and the included fractions were also shown to contain Ara and Man in different ratios. After lyophilization, these latter fractions containing oligosaccharides were derivatized.

Oligosaccharide Derivatization: Preparation of ABEE Derivatives—The ultraviolet-absorbing compounds were coupled to the oligosaccharides by reductive amination with cyanoborohydride. Shown below as an example is the coupling of polymer of maltose with *p*-aminobenzoic ethyl ester (ABEE) (Reaction 1).

Up to 1 mg of oligosaccharide was dissolved in 10 μl of H_2O and 40 μl of the reagent solution. Routinely, this reagent mixture was prepared by dissolving 1 mmol of derivatizing compound (165 mg of ABEE) and 35 mg of sodium borohydride in 41 μl of glacial acetic acid and 350 μl of methanol. The glass vials were capped, vortexed, and heated at 80 °C. After 90 min, the vial was cooled, equal volumes of chloroform and distilled water (1 ml) were added to the reaction mixture, and the contents were mixed. The derivatives were extracted into the aqueous phase which was lyophilized.

Separation of ABEE Derivatives—Derivatized oligosaccharides were both separated and effectively desalted by reversed-phase HPLC



REACTION 1

using a C18 column (ODS) and a water/acetonitrile solvent system with an elution gradient of 5–80% acetonitrile in 25 min with a flow rate of 2 ml/min. The oligosaccharide-ABEE derivatives were monitored by their absorbance at 304 nm, which is conferred by the *p*-aminobenzoate group. Eluted samples were lyophilized prior to mass spectrometric analysis.

HPLC was conducted on a Gilson (Gilson, France S. A., Villiers-le-bel, France) apparatus composed of two model 303 pumps, a 802 manometric module and a 811 dynamic mixer. The elution and the data acquisition were controlled with Gilson GME 714 Software and a Gilson model 621 data module interface, connected to an OPAT Normerel computer. Elution of the products was monitored with a Gilson model ultraviolet 116 variable wavelength detector.

Laser Desorption-Ionization Mass Spectrometry—The mass spectrometer consisted of a laser desorption ion source coupled to a time-of-flight (TOF) mass analyzer. A pulsed UV laser (N_2) of 337 nm wavelength and 3-ns pulse width type VSL337ND (Laser Science Inc., Cambridge, MA) equipped with a sensitive light absorber was used for ion desorption. The laser beam was focused on the target and moved to cover the entire surface. The spot size measured on the target was about 3.2 mm^2 with irradiance at the sample surface of 90×10^6 watts/cm^2. The operating voltage was 10 kV, and the accelerated ions were detected by microchannel plates.

The signals were recorded by a transient waveform recorder (LeCroy Corp., Chestnut Ridge, NY) coupled to a PC-AT microcomputer for the storage of the TOF mass spectra. All spectra shown here were accumulations of 10–60 single spectra and were obtained in a positive ion mode. Measurements typically took 10–15 min.

The samples were prepared using standard procedures by mixing appropriate volumes of matrix and analyte solution.

The matrix used was 2,5-dihydroxybenzoic acid dissolved to a concentration of 10 mg/ml in 10% (v/v) ethanol/water, 0.1% TFA solution.

3 μl of sample solution (10 mg/ml in doubly distilled water) were mixed with 30 μl of matrix solution to a final concentration of 1 mg/ml, and 5 μl of this mixture were applied on a stainless steel support and covered an area of about 5 mm in diameter. The droplet was allowed to dry at room temperature for a few minutes before introduction into the instrument.

FAB/MS and metastable ion spectra were recorded on a reverse geometry two-sector instrument ZAB-HS using an 8-kV xenon beam, in the 2000 to 100 mass/charge range. Samples of derivatized oligosaccharides were applied to a matrix of glycerol/thioglycerol, 1:1, on the probe tip.

GC and GC/MS of per-O-Alkylated Oligoglycosyl Alditols—Routine gas chromatography was performed on a Girdel series 30 equipped with an OV1 wall-coated open tubular, inner diameter 0.3-mm, 25-m capillary column using nitrogen gas at a flow rate of 2.5 ml/min with a flame ionization detector. A temperature program from 100 to 250 °C at a speed of 3 °C/min and an injector temperature of 260 °C were used for trimethylsilyl methyl glycoside analysis. The temperature of the detector was 310 °C.

GC/MS was performed on a Nermag model R 10/10 quadrupole mass spectrometer connected to a PDP-8M computer.

Nuclear Magnetic Resonance Spectroscopy—1D ^{13}C NMR spectra were recorded on a 250-MHz Brucker and 2D NMR spectra on a 500-MHz Brucker AMX-500 spectrometer equipped with an Aspect X32 computer. ^1H and ^{13}C chemical shifts are expressed in parts/million downfield from internal tetramethylsilane (0 ppm) and internal chloroform (77 ppm). Samples were lyophilized several times in D_2O to allow a better exchange with hydroxyls, dissolved in D_2O (Spin et Techniques, Paris, France, 99.96% purity) at a concentration of 20 mg/ml and analyzed in a 200 × 5 mm 535-PP NMR tube. All 2D NMR data sets were recorded at 30 °C without sample spinning for heteronuclear experiments. Data were acquired in the phase-sensitive mode by using the TPPI method (25) or the method of States *et al.* (26). COSY was recorded with standard pulse sequences (27). ^1H-^{13}C correlation spectra were recorded in the proton-detected mode with a Brucker 5-mm ^1H-broad band tunable probe with reversed geometry. The pulse sequence used for single-bond correlation spectra (HMQC) was that of Bax *et al.* (28). GARP sequence (29) at the carbon frequency was used as a composite pulse decoupling during the acquisition. Multiple-bond correlation spectra (HMBC) (30) were recorded and processed in the magnitude mode. The pulse sequence used for ^1H-detected heteronuclear relayed spectra (HMQC-HOHAHA) was that of Lerner and Bax (31).

The proton-decoupled ^{13}C DEPT spectrum was recorded with a recycle delay of 1.5 s, and 24078 scans were accumulated. The spectral width used was 4717 Hz and the data were digitized on 16K real points, giving a digital resolution of 0.28 Hz/point. The data are multiplied by a line-broadening function (LB = 1.5 Hz) prior to being Fourier-transformed.

2D $\{^1H,^1H\}$ COSY spectrum was recorded with presaturation during relaxation delay. The spectral width was 2128 Hz; the data matrix was 2K × 256 (TPPI) points, with 32 scans per t_1 value. Data were zero-filled in the F_1 dimension to obtain a final data matrix with 512 × 512 real points.

Phase-sensitive, ^{13}C-decoupled, ^1H-detected multiple quantum correlation spectrum (^1H $\{^{13}C\}$ HMQC) was recorded with presaturation during relaxation delay. The data matrix was 2K × 256 (TPPI) complex points with 32 scans per t_1 value. The spectral window was 15,092 Hz in the F_1 dimension (^{13}C) and 2,128 Hz in the F_2 dimension (^1H). Delay after BIRD pulse was 380 ms. For processing, a sine bell window shifted by p/2 was applied in both dimensions, followed by expansion of the data matrix to 2K × 512 real matrices. The resulting digital resolutions were 0.4 Hz/point in F_2 (^1H) and 7.4 Hz/point in F_1 (^{13}C).

For the ^{13}C-decoupled, ^1H-detected multiple bond correlation spectrum (^1H $\{^{13}C\}$ HMBC), the data matrix was 2K × 256 complex points with 80 scans per t_1 value. The spectral window was 15092 Hz in the F_1 dimension (^{13}C) and 2128 Hz in the F_2 dimension (^1H). For processing, a sine bell window shifted by p/2 was applied in both dimensions, followed by expansion of the data matrix to 2K × 512 real matrices. The resulting digital resolutions were 0.4 Hz/point in F_2 (^1H) and 7.4 Hz/point in F_1 (^{13}C). The data are presented in the magnitude mode.

RESULTS

Lipoarabinomannan Purification and Molecular Mass Determination—Chromatography gel permeation was unsuccessfully applied using denaturing buffers in order to obtain a purified native LAM from the ethanol-insoluble fraction. The LAM remained contaminated by lipomannan (LM). This purification problem arises from the amphipathic nature of the LAM leading to the formation of micelles in aqueous solution with an approximate molecular mass of 10^6. It was overcome by deacylation of the fraction containing LAM under mild conditions. Fig. 1 shows the Bio-Gel P100 chromatography profile of the deacylated fraction using an aqueous solution containing 0.1 N acetic acid as eluate. This analysis revealed three broad peaks, namely A, B, and C, assigned by means of routine sugar analysis, to glucane, deacylated LAM (dLAM), and deacylated LM (dLM), respectively. Molecular masses of polysaccharides can be determined tentatively from gel permeation using eluates containing denaturing buffers like 6 M urea, guanidine HCl, or acid which disrupt molecular aggregates and prevent interaction with the matrix. By calibrating the gel by dextrans of different sizes, the apparent molecular mass of dLAM was found to be around 10 to 15 kDa.

Another approach used routinely for the characterization of proteins and lipopolysaccharides (LPS) is sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis (32, 33). The electrophoretic migration pattern of a LPS is characteristic of its structure and may reflect biochemical variations in its composition. However, using this method, the LPSs were classified only into high and low molecular weight species. In the literature, LAM from *M. tuberculosis* was described as a broad band at around 30–40 kDa compared to standard proteins. Fig. 2 shows the SDS-PAGE results of the *M. bovis* BCG LAM fraction (*B*) purified as discussed above, also containing LM. As expected, this gel is dominated by a typical broad band located at an apparent molecular mass of 30 kDa and a second band at 20 kDa assigned to LM. The rate of migration of the LAM from *M. tuberculosis* strain H37Rv (*A*) was close to that of *M. bovis* BCG, suggesting the similarity of their molecular masses. The dLAM (*C*) showed

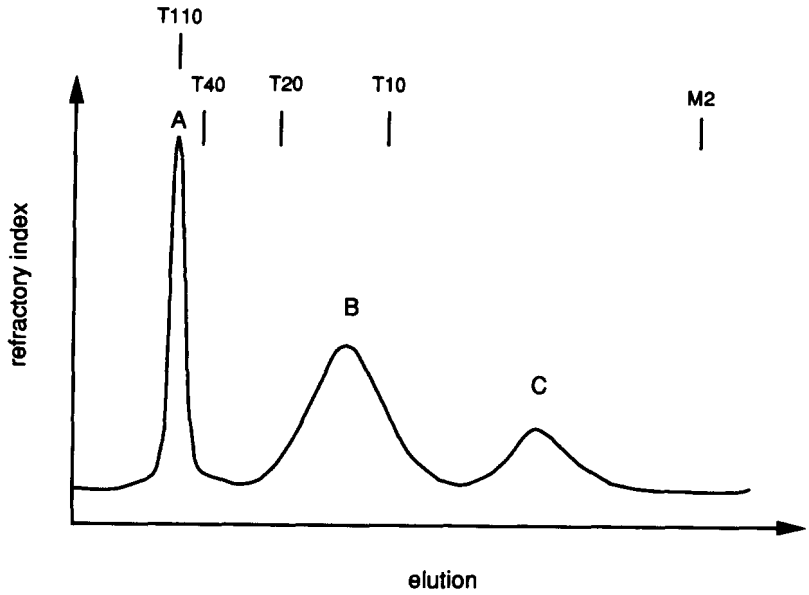


FIG. 1. **Purification of dLAM from *M. bovis* BCG on Bio-Gel P-100 column (16 × 600 mm) in 0.1 N CH_3COOH detected by refractory index.** Fractions from *peaks A, B,* and *C* were analyzed by gas-liquid chromatography. *Sample A* contains a high molecular weight contaminant composed of glucosyl residues, *sample B* contains only arabinosyl and mannosyl residues (Ara/Man = 1.5) and corresponds to dLAM, and *sample C* contains arabinosyl and mannosyl residues (Man/Ara = 20) and corresponds to dLM. The column was previously calibrated using dextrans (T40, T20, and T10) and maltose (M2) as size markers.

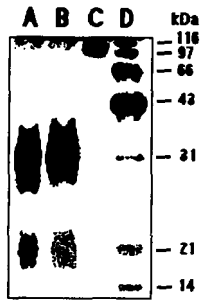


FIG. 2. **SDS-PAGE of enriched fractions of LAMs containing LM from *M. tuberculosis* H37Rv (*A*) and *M. bovis* BCG (*B*).** dLAM (*C*) did not migrate. The positions of proteins markers (*D*) are indicated. Sample amount was 5 μg. The gels (12.5%) (12 × 9 cm) were stained with a silver stain containing periodic acid.

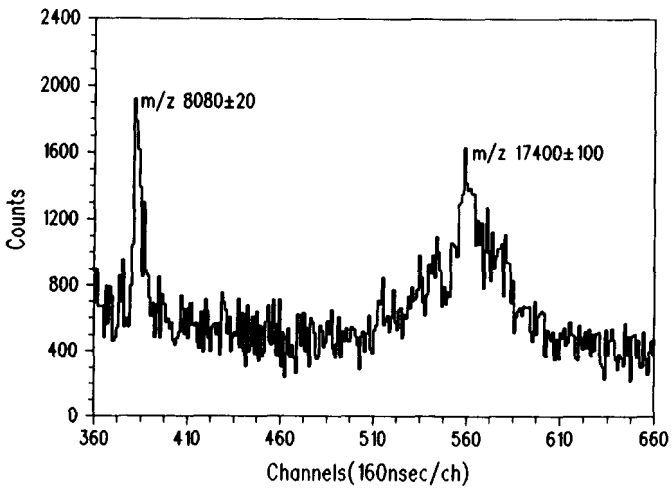


FIG. 3. **Matrix-assisted UV-laser desorption/ionization spectrum of a mixture of LAM and LM from *M. bovis* BCG in 2,5-dihydroxybenzoic acid as matrix.**

a band in the top of the gel with an apparent molecular mass of 100 kDa.

Fig. 3 shows the positive matrix-assisted UV-LDI mass spectrum of the LAM fraction contaminated by LM from the *M. bovis* BCG strain Pasteur. The matrix selected was 2,5-dihydroxybenzoic acid, since from literature data, this matrix is more convenient than sinnapinic acid for oligosaccharide (34, 35) and glycolipid (36) LDI analysis. The matrix was solubilized in a mixture of water/ethanol, 9:1, containing 0.1% of TFA to which 5 μg of sample in aqueous solution was added. The TFA was added to increase the protonation of the LAM and LM. The mass spectrum, recorded in the mass range 5 to 50 kDa, is dominated by two signals centered at m/z 8,080 ± 20 and m/z 17,400 ± 100. The peak at m/z 17,400 is extremely broad and appears as a hump, while the peak at m/z 8,080 is more narrow. The width of the high mass peak at half-height corresponds to about 4 kDa and could arise either from the heterogeneity of the analyzed sample or from glycosidic bond fragmentation occurring during the LDI process. However, LDI spectra of oligosaccharides (34, 35) are dominated by protonated or cationized molecular ions, and fragment ions are usually absent. Thus, the broad peak at m/z 17,400 revealed that the analyte sample consisted of a mixture of oligoglycans which could differ either by the number of mannose and arabinose units or by the presence of appendages already described like, for example, phosphate (5, 7), lactyl, or succinyl (4) residues. The low mass resolution of the direct TOF analyzer did not allow mass separation of the oligoglycans and led to the formation of a broad peak. LDI mass spectra of proteins of the same mass value showed narrow peaks.

As mentioned under "Materials and Methods," it is difficult to obtain LAM devoid of LM due to their micellar behavior; however, their separation becomes easier with the deacylated species. Fig. 4 shows the LDI mass spectrum of the deacylated LAM from *M. bovis* BCG. The mass spectrum is dominated by two broad peaks, an intense one centered at m/z 16,700 ± 200 and a weak one at m/z 33,350 ± 200. These peaks were assigned to monomeric and dimeric pseudomolecular ions, respectively. We have no evidence concerning their exact structure and, for instance, the formation of cluster ions containing the analyte alkaline cation and matrix adducts can occur. Nevertheless, the additional mass arising from the adduct molecules is weaker compared to the molecular mass of the analyte. Taken together, these experiments indicate that the molecular mass of the LAM from *M. bovis* BCG is 17.4 ± 0.1 kDa. Again, the width of the dLAM peak is similar to that of the LAM indicating that the LAM heterogeneity

was not due to the alkaline appendages. From these data, it was more difficult to conclude as to the molecular mass of the LM since no significant peaks were observed in the LDI spectrum of the dLM (not shown) and, consequently, we have no evidence for the assignment of the *m/z* 8,080 peak.

The LDI mass spectrum of a mixture of LAM and LM from *M. tuberculosis* strain H37Rv (not shown) showed a broad peak centered at *m/z* 17,250 Da with a width at half-height corresponding to about 5,000 mass units.

These experiments established that *M. tuberculosis*, the etiological agent of tuberculosis, and *M. bovis* BCG, the vaccine strains, are endowed with LAMs of similar heterogeneity and molecular mass at around 17.3 kDa. In order to characterize the structural features of both LAMs, a comparative structural analysis using conventional and recent analytical tools was undertaken.

Methylation Analysis—In order to determine the glycosyl linkage composition of LAM from *M. bovis* BCG and from *M. tuberculosis* H37Rv, purified dLAMs (1 mg) were per-*O*-methylated as described by Ciucanu and Kerek, hydrolyzed, reduced, and per-*O*-acetylated, and the resulting alditol acetates were analyzed by GC and GC/MS. The data are summarized in Table I and indicate that mannosyl residues, on the one hand, and terminal and 2-*O*-linked arabinosyl residues, on the other hand, were in the pyranose and furanose forms, respectively. The furanose form of the 5-*O*-linked arabinosyl residues was established from the NMR study mentioned below.

The two LAMs studied showed similar ratios of partially methylated alditol acetates. The ratio of terminal Ara*f* (t-Ara*f*) (2%) to 2-*O*-linked Ara*f* (6 to 7%) indicates that around only about 30% of the 2-*O*-linked Ara*f* residues are substituted with t-Ara*f* in both LAMs since it has been reported for LAMs of *M. tuberculosis* (6) that 2-*O*-linked Ara*f* occurs in the arabinosyl side chains of the arabinan region. The number of t-Man*p* residues in the LAM from *M. bovis* BCG (24%) and from *M. tuberculosis* H37Rv (26%) exceeds the number of branched 2,6-linked Man*p* residues (8%). Similarly, it can be proposed that t-Man*p* residues were not only located on the mannan core but were also attached to the arabinan region. The higher abundance of 2-*O*-linked Man*p* in LAM from BCG and from H37Rv (10 and 9%, respectively) compared to that already described in LAM from H37Ra (6) (0.6%) suggested that these residues were also located on the arabinan core; however, the presence of additional 2-*O*-linked Man*p* in the mannan "core" cannot be excluded. From these data, at the nonreducing end of the branched arabinofuranosyl residues, we can propose not only the presence of t-Ara*f* but also that of t-Man*p* and maybe of 2-*O*-linked Man*p* residues.

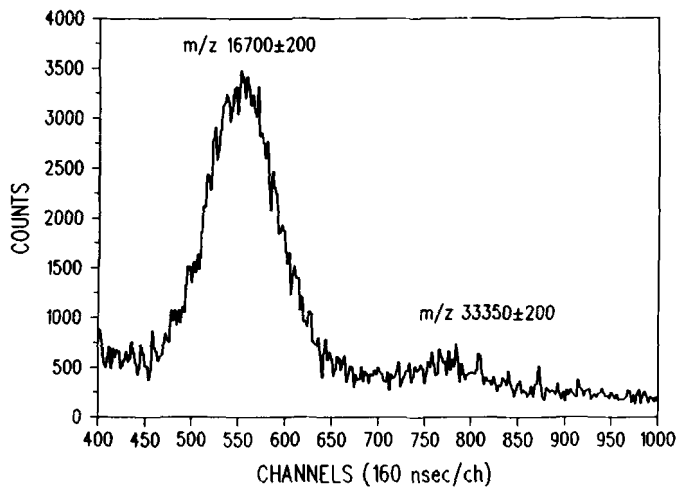


FIG. 4. **Matrix-assisted UV-laser desorption/ionization of dLAM from *M. bovis* BCG in 2,5-dihydroxybenzoic acid as matrix.**

NMR Analysis of LAMs from M. bovis BCG and from M. tuberculosis H37Rv—In order to determine the location of the Man*p* residues more precisely, the dLAMs of *M. bovis* BCG and *M. tuberculosis* H37Rv were analyzed by 1D and 2D homonuclear and heteronuclear scalar coupling NMR spectroscopy. Both dLAMs show similar NMR spectra, and only spectra of the dLAM from *M. bovis* BCG are shown and described. The ^1H NMR spectrum in D_2O at 500 MHz (not shown) indicated a complex anomeric proton region with multiple overlapping broad peaks whose direct assignment would be dubious.

The DEPT ^{13}C NMR spectrum of the dLAM is presented in Fig. 5. In the anomeric region, one can observe a broad and intense signal at 110 ppm (I_1) and poorly resolved signals at 109 (II_1) and 108.3–108 ppm (III_1), well defined signals at 104.8–104.7 (IV_1) and 100.8 (VII_1), and weak signals at 103.2 ppm (V_1) and 102 ppm (VI_1).

From literature data (37, 38), the anomeric carbon resonances between 108 and 110 ppm were assigned to α-Ara*f* residues. Thanks to data from arabinans (39), arabinogalactan (40), and LAM from H37Ra (6) and Erdman strains (18), and in agreement with the alditol acetate data, the downfield carbon resonances (I_1 and II_1) at 109–110 ppm were assigned to 5-*O*-linked α-Ara*f*. The signals at 108.3–108 ppm (III_1), with this upfield shift of 1 to 2 ppm, agreed with C-1 resonances of 2-*O*-linked α-Ara*f*. Indeed, an ether linkage at C-2 is known to cause a weaker upfield shift by 1–2 ppm of C-1 resonance (37). This assignment was also supported by 2D hetero- and homonuclear scalar couplings. The ^1H-detected heteronuclear multiple quantum correlation ($^{13}C/^1H$ HMQC) spectrum (Fig. 6) showed connectivities between the C-1

TABLE I

Methylation analysis of LAM from M. bovis BCG and M. tuberculosis H37Rv

Full name of the partially methylated alditol acetate	Abbreviated name of the glycosyl residue	LAM from *M. bovis* BCG	LAM from *M. tuberculosis* H37Rv
		mol %	
2,3,5-tri-*O*-Me-1,4-di-*O*-Ac-arabinitol	t-Ara*f*	1–2	2
3,5-di-*O*-Me-1,2,4-tri-*O*-Ac-arabinitol	2-Linked Ara*f*	7	6
2,3-di-*O*-Me-1,4,5-tri-*O*-Ac-arabinitol	5-Linked Ara*f*[a]	37	36
2-*O*-Me-1,3,4,5-tetra-*O*-Ac-arabinitol	3,5-Linked Ara*f*[a]	10	9
2,3,4,6-tetra-*O*-Me-1,5-di-*O*-Ac-mannitol	t-Man*p*	24	26
3,4,6-tri-*O*-Me-1,2,5-tri-*O*-Ac-mannitol	2-Linked Man*p*	10	9
2,3,4-tri-*O*-Me-1,5,6-tri-*O*-Ac-mannitol	6-Linked Man*p*	3	4
3,4-di-*O*-Me-1,2,5,6-tetra-*O*-Ac-mannitol	2,6-linked Man*p*	8	8

[a] Shown to be predominantly 5-*O*-linked Ara*f* rather than 4-*O*-linked Ara*p* (see text).

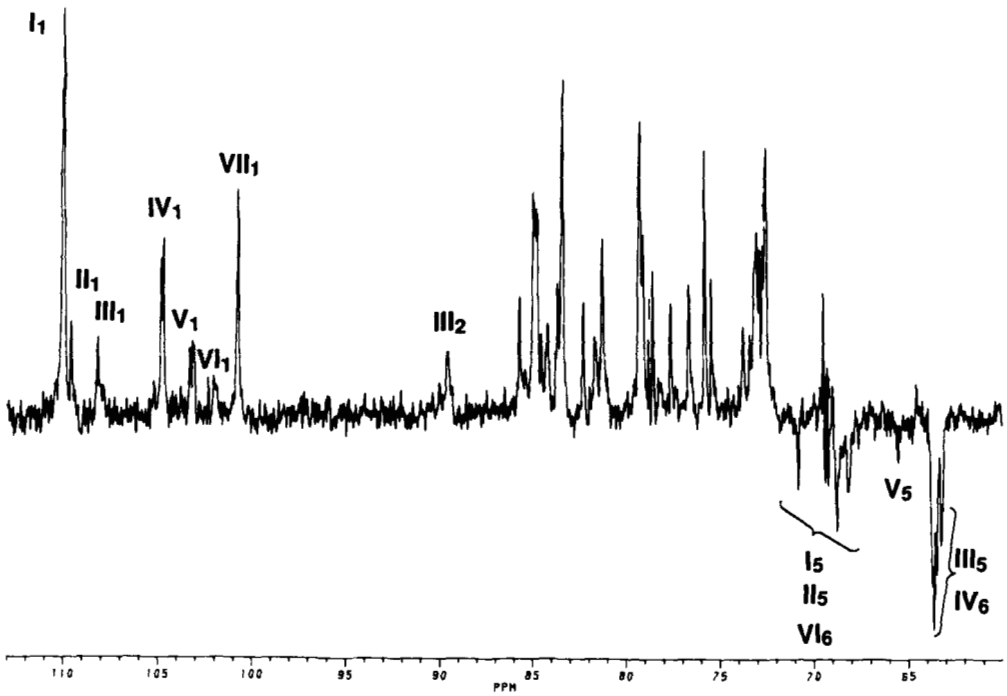


FIG. 5. **^{13}C NMR (DEPT) spectrum (135°) of dLAM from *M. bovis* BCG.** Glycosyl residues are labeled in *roman numerals* and their carbons in *arabic numerals*. *I-II*, 5-*O*-α-Ara*f*; *III*, 2-*O*-α-Ara*f*; *IV*, t-α-Man*p*; *V*, t-β-Ara*f*; *VI*, 6-*O*-α-Man*p*; *VII*, 2-*O*- and 2,6-*O*-α-Man*p*.

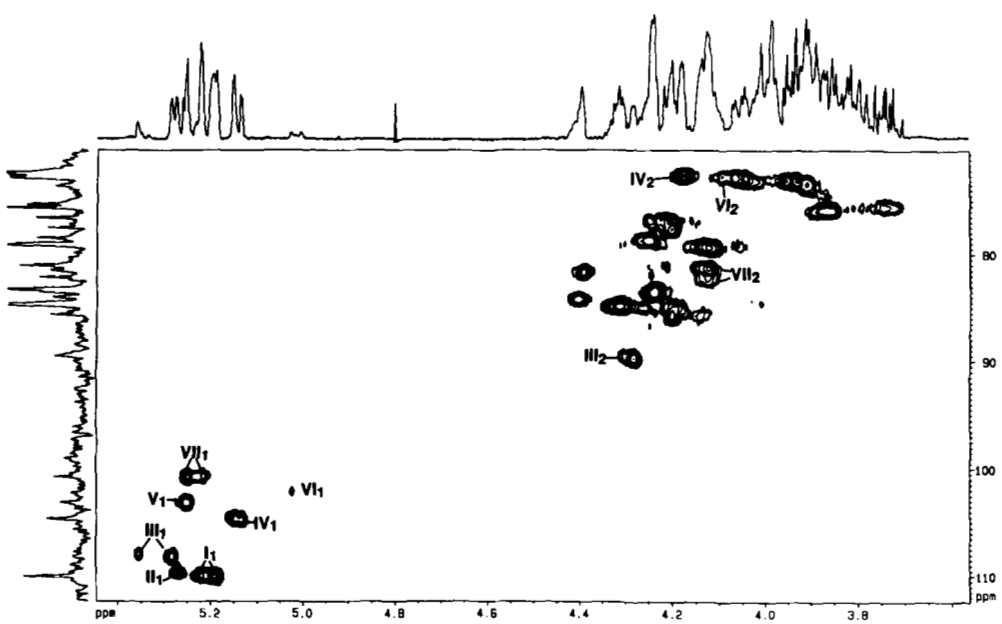


FIG. 6. **A region (δ ^{13}C: 70–110; δ ^{1}H: 5.0–5.4) of the phase-sensitive, ^{13}C-decoupled, ^{1}H-detected multiple quantum correlation spectrum (^{1}H {^{13}C} HMQC) of dLAM from *M. bovis* BCG at 500 MHz.** Glycosyl residues are labeled in *roman numerals* and their protons and carbons in *arabic numerals*.

resonances of (III), at around 108 ppm, and the two corresponding H-1s, at 5.36 and 5.28 ppm (III_1). The COSY spectrum (Fig. 7) reveals that both anomeric protons correlate with a unique resonance at δ4.28 assigned to the H-2s ($III_{1,2}$). Thus, from the HMQC spectrum, the C-2 resonances were localized at δ90 (III_2). This downfield chemical shift ($\Delta\delta = 5$ ppm) confirmed glycosyl substitution on C-2 of the arabinofuranosyl residue characterized by the C-1 resonance at δ108.3–108.

The resonance at 103.2 ppm (V) (Fig. 5) could be attributed to the C-1 of the t-β-Ara*f* unit, in agreement with literature data (39). Indeed, the 1D ^{13}C NMR spectrum of the LAM from an avirulent *M. tuberculosis* strain (H37Ra) (6), characterized by a large number of t-β-Ara*f*, exhibits an intense peak at 103 ppm. Moreover, the C-5 resonance at 67 ppm (V_5) in the 1D ^{13}C spectrum (Fig. 5) of dLAM revealed the β-anomeric configuration of the t-Ara*f* since the other primary alcohol carbons of 5-, 3,5-α-*O*-linked Ara*f* and 6- and 2,6-*O*-linked Man*p* resonate at 70 ppm and the primary alcohol carbon of 2-*O*-linked α-Ara*f* at around 65 ppm.

The remaining signals at δ100.8 (VII), δ104.8–104.7 (IV), and δ102 (VI) were assigned to the C-1s of the different types of *O*-linked Man*p*. The mannosyl residues of LAMs are known to be in the α configuration (4). The upfield resonance (VII)

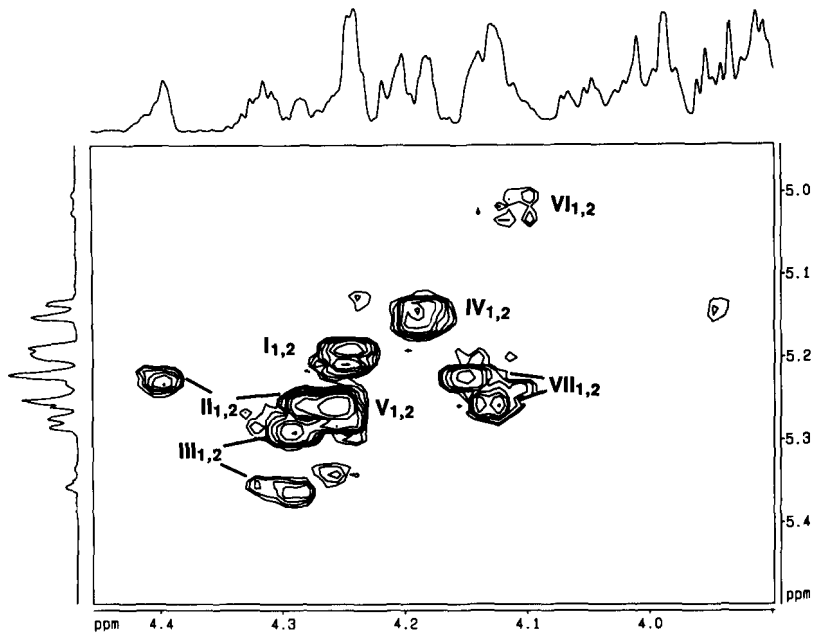


FIG. 7. **A region (F_1: δ5.0–5.4; F_2: δ4.0–4.4) of the 2D (^{1}H,^{1}H) COSY spectrum of dLAM from *M. bovis* BCG.** Glycosyl residues are labeled in *roman numerals* and their protons in *arabic numerals*.

at 100.8 ppm can be attributed to the anomeric resonance of either a 2-*O*- or 2,6-di-*O*-linked Man*p* in agreement with the shielding effect induced by glycosylation. The HMQC spectrum (Fig. 6) shows correlation of this upfield anomeric carbon resonance to two H-1 resonances at 5.25 and 5.22 ppm (VII_1). Thus, these protons characterized the two types of 2-*O*-linked-α-Man*p*. From the COSY and HMQC (Figs. 7 and 6), the H-2 resonances at δ4.13 and δ4.15 ($VII_{1,2}$), as well as the C-2 resonances between δ80.5 and δ81.25 (VII_2) were assigned. The downfield chemical shift values agreed with 2-*O*-linked α-Man*p* (41). In addition, HMQC-HOHAHA experiments (not shown) showed long range couplings from the H-1 resonances at 5.25 and 5.22 ppm (VII_1) to C-2 between 80.5 and 81.25 ppm, unambiguously supporting the spin system (VII) assignment of 2-*O*-linked Man*p* and 2,6-di-*O*-linked Man*p*.

The two other groups of signals at 104.8–104.7 ppm (IV) and 102 ppm (VI) can be assigned to t-α-Man*p* and 6-*O*-linked-α-Man*p*, respectively. By means of HMQC (Fig. 6), the connectivities of the two C-1s at 104.7 and 104.8 ppm (IV) with corresponding H-1 resonances at 5.13 and 5.145 ppm, respectively (IV_1), were established. From the COSY spectrum (Fig. 7), it was found that the H-2s resonated at 4.17 ($IV_{1,2}$) for both residues. The ^{13}C/^{1}H correlation showed connectivity of these H-2s with the C-2 at 72.65 ppm (IV_2). Likewise, for the last system (VI), the ^{13}C/^{1}H HMQC spectrum showed correlation between C-1 resonance at 102 ppm (VI) with the corresponding H-1 resonance at δ5.02 (VI_1). COSY and HMQC allowed the assignment of the H-2 resonance at δ4.1 ($VI_{1,2}$) and the C-2 resonance at δ72.6 (VI_2). These chemical shift values did not allow us to differentiate a t-Man*p* from a 6-*O*-linked Man*p*. Unfortunately, COSY allows attribution only up to H-2 since most of the remaining protons of the monosaccharides resonate in a very restricted zone, between 3 and 4 ppm, which is the cause of overcrowding in 2D COSY spectrum. In order to unambiguously assign these resonances, the mole percentage from the methylation analysis (3% for 6-linked Man*p* against 24% for t-Man*p*) and the intensity of the signals in the ^{1}H 1D spectrum (1 for the H-1 at δ5.02 (VI_1) and 7.5 for the H-1s at δ5.13 and δ5.15 (IV_1)) were correlated. Thus, the system (IV) with C-1s at 104.8 and 104.7 was assigned to two populations of t-Man*p* and the remaining system (VI) to the 6-*O*-linked α-Man*p*.

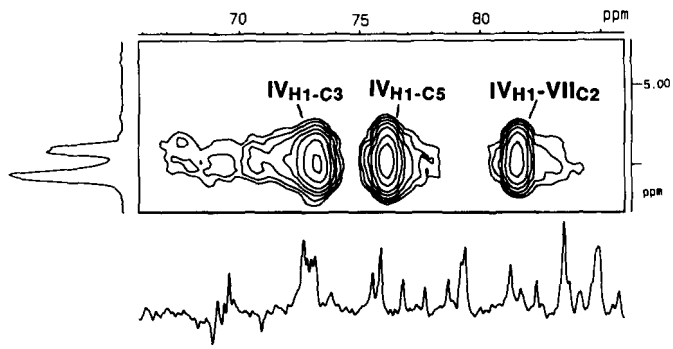


FIG. 8. **Expanded region (δ ^{13}C: 65–85; δ ^{1}H: 5.0–5.3) of the ^{13}C-decoupled, ^{1}H-detected multiple bond correlation spectrum (^{1}H {^{13}C} HMBC) of the dLAM from *M. bovis* BCG at 500 MHz.** Glycosyl residues are labeled in *roman numerals* and their protons in *arabic numerals*.

Both NMR spectroscopy and methylation data analysis demonstrated the occurrence of two types of t-Man*p* and 2-*O*-linked Man*p* in LAM from *M. bovis* BCG and *M. tuberculosis* H37Rv. These results allowed us to assume the presence of short side chains of (1→2)-linked D-Man*p* not only on the mannan, but also on the arabinan core.

The sequence of these glycosyl residues was supported by ^{1}H-detected heteronuclear multiple bond correlation (^{13}C/^{1}H HMBC) experiments. This method allowed the determination of the carbons and protons involved in the interglycosidic linkage through three-bond correlation between the anomeric carbon or proton resonances and the carbon or proton resonances from the adjacent unit (42). Indeed, the partial HMBC spectrum on Fig. 8 showed connectivities between the two H-1s of t-Man*p* (IV_{H1}: δ5.13 and δ5.15) and the C-2s of 2-*O*-linked Man*p* (VII_{C2}: δ80.5–81.2). Besides this inter-residue correlation, this anomeric proton also exhibited long-range intra-residue correlation to C-3 (δC3 (IV): 73 ppm) and C-5 (δC5 (IV): 76 ppm) resonances ($^3J_{CH}$). This last NMR study established the interglycosidic linkage between the t-Man*p*s and the 2-*O*-Man*p*s.

*Characterization of Small Oligosaccharides Containing Man*p *from Hydrolyzed LAM from M. bovis BCG*—In order to confirm the presence of these mannosyl residues at the ends

of the arabinan side chains and to determine their size, a four-step analytical approach summarized below was successfully applied to the dLAM: (i) partial hydrolysis of LAM, (ii) derivatization of the reaction products with *p*-aminobenzoate ethyl ester (ABEE), (iii) purification by reversed-phase HPLC monitored by UV detection, and (iv) structural characterization by FAB/MS and FAB/MS-MS. LAM was partially hydrolyzed under two sets of conditions: 0.1 N HCl, 90 °C, 15 min (A) and 0.075 N HCl, 75 °C, 15 min (B). The subsets of oligosaccharides were separated from the nonhydrolyzed mannan by gel filtration chromatography as described under "Materials and Methods" and were derivatized with ABEE. The oligosaccharide-ABEE derivatives obtained after hydrolysis conditions (A) were analyzed by reversed-phase HPLC in a gradient of acetonitrile in water. The HPLC profile (Fig. 9*a*) revealed one major UV-absorbing *peak f* with a retention time of 20 min characteristic of a monosaccharide-ABEE. This assignment was confirmed by positive FAB/MS of *compound f* which exhibited both $(M + H)^+$ and $(M + Na)^+$ at *m/z* 300 and 322, respectively. Thus, from the molecular mass, it can be concluded that this major component, which represented 80% of the whole fraction, corresponds to Ara-ABEE. This method also allowed an estimation to be made of the relative amount of oligosaccharide by means of the peak integration, since only one chromophore is present per molecule of the derivatized oligosaccharide.

Besides *peak f*, less intense peaks were still well-resolved and one (*peak b*) predominates at 15.5 min and represents 12% of the whole fraction. The positive FAB/MS analysis of *fraction b* (Fig. 10*a*) showed two intense ions $(M + H)^+$ and $(M + Na)^+$ at *m/z* 624 and *m/z* 646, respectively, revealing a trisaccharide composed of 2 Man and 1 Ara unit. In order to sequence this trisaccharide, the unimolecular decompositions of the precursor ion *m/z* 624 were recorded (Fig. 10*b*). The MIKE spectrum exhibited two major ions at *m/z* 300 and *m/z* 462. The latter ion results from the loss of the *anhydro*-monosaccharide unit localized at the nonreducing end (Y_1 *m/z* 462). Thus, the mass difference unambiguously indicated that a Man*p* unit is localized at the nonreducing extremity of the trisaccharide. Likewise, fragment ion Y_2 at *m/z* 300 can be explained by the loss of an *anhydro*-disaccharide unit composed of 2 Man*p* residues. It allowed the localization of the arabinosyl residue at the reducing terminus of the oligo-

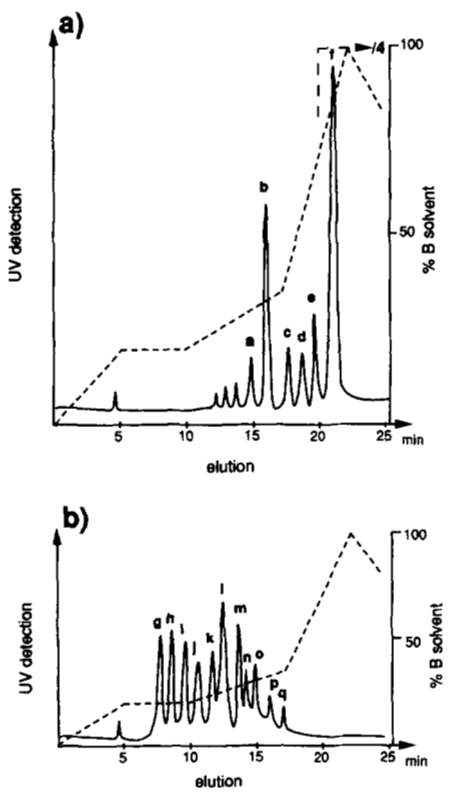


FIG. 9. **Reversed-phase HPLC profile (semipreparative column C_{18}, 9.5 × 250 mm, 5-μm Spherisorb ODS2) of oligosaccharide-ABEE derivatives obtained from LAM of *M. bovis* BCG after mild hydrolysis, gel filtration, and derivatization with aminobenzoic ethyl ester.** *a*, mild hydrolysis (0.1 N HCl, 15 min, 90 °C); *b*, mild hydrolysis (0.075 N HCl, 15 min, 75 °C). The column is eluted by a gradient of 5% acetonitrile in water (solvent A) to 80% acetonitrile in water (solvent B). *Fractions a-p* are characterized in the text.

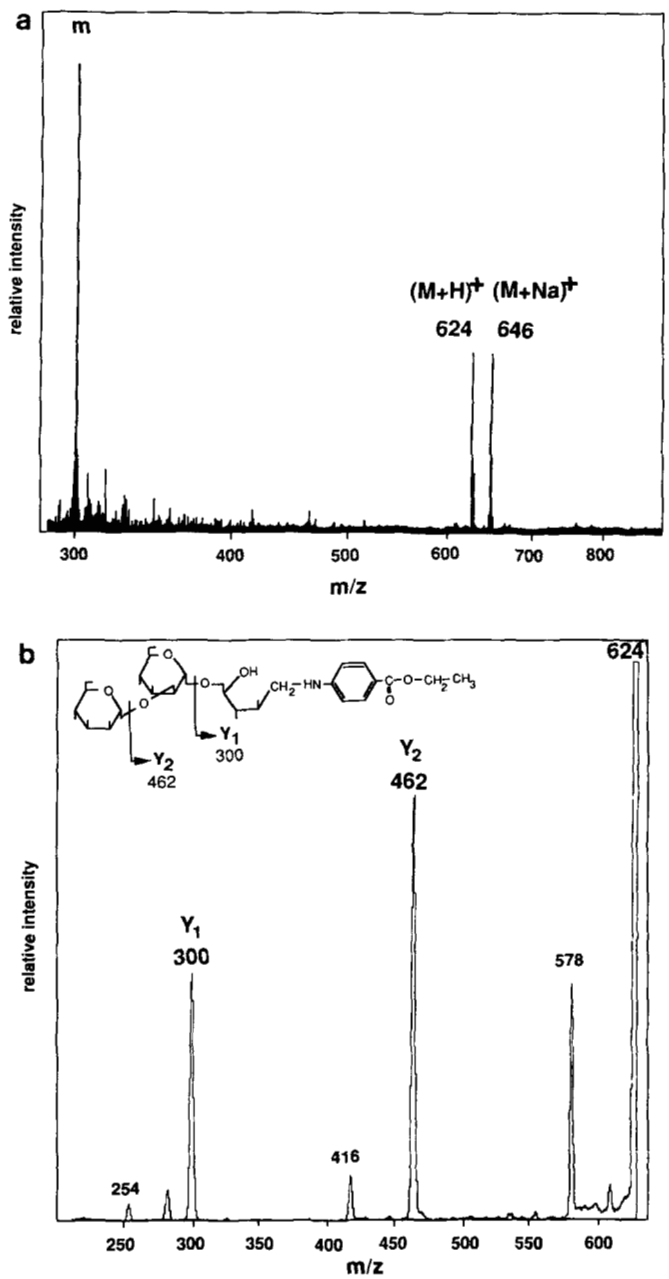


FIG. 10. **Characterization of *fraction b* from the HPLC profile (Fig. 9*a*) of oligosaccharide-ABEE obtained from LAM of *M. bovis* BCG.** *a*, positive ion FAB mass spectrum of *fraction b* in a matrix (*m*) of glycerol/thioglycerol. *b*, positive ion MIKE spectrum of the protonated molecular ion *m/z* 624 from *fraction b* in a matrix of glycerol/thioglycerol.

TABLE II

Cationized molecular ion mass values of fractions from the HPLC profile of Fig. 9b

Fractions	$(M + Na)^+$ *m/z*	Oligosaccharides
g	1306	6 Ara + 2 Man
h	1174	5 Ara + 2 Man
i	850	5 Ara
j	1042	4 Ara + 2 Man,
	808	1 Ara + 3 Man
k	910	3 Ara + 2 Man
l	778	2 Ara + 2 Man
m	880	4 Ara + 1 Man
n	646	1 Ara + 2 Man
o	616	2 Ara + 1 Man
p	586	3 Ara

saccharide. From these data, the trisaccharide sequence is as follows: Man*p*→Man*p*→Ara*f*. The relatively high abundance of the ion fragments suggests that the oligosaccharide was linear, in agreement with the methylation data (not shown). Associated with the loss of each hexose unit, additional ions appeared 46 mass units below each peak which may result from the loss of ethanol from the ABEE moiety.

Likewise, the purified HPLC oligosaccharide-ABEE *a* (Fig. 9*a*) was assigned to a tetrasaccharide composed of 3 Man and 1 Ara residues, and *fraction c* was attributed to a heterodisaccharide. In both cases, MIKE spectra revealed the location of an arabinofuranosyl unit at the reducing end.

Finally, *fraction d* corresponds to an arabinose homo-disaccharide and *sample e* to the monosaccharide Man. Hydrolysis procedure (A) therefore yields oligosaccharides with 1, 2, and 3 mannosyl residues and weak amounts of free mannose.

In the aim of detecting the presence of higher oligomannans at the end of the arabinan chains, milder hydrolysis according to protocol B (0.075 N HCl, 15 min, 75 °C) was applied to dLAM. Fig. 9*b* shows the partial reversed-phase HPLC profile revealing a dozen poorly resolved components, which were analyzed by positive mode FAB/MS. The spectra are dominated by cationized molecular ions whose mass values are summarized in Table II. They mostly correspond to Man-containing oligosaccharides. Seven oligosaccharides were composed of 2 mannosyl residues and 1 to 7 arabinosyl residues. Two others contained 1 mannosyl and another one 3 mannosyl units.

These partial hydrolysis experiments supported a mannosylated extension of the arabinan of *M. bovis* BCG and established that the major structural motif located at the nonreducing end of the arabinan of this LAM was a dimannosyl residue. Besides this disaccharide, a few mannose and mannotriose residues also capped the arabinan side chains.

DISCUSSION

Knowing the molecular mass of biopolymers constitutes key information for their structural characterization but also for a better understanding of their functions. In the case of proteins, genetic approaches and electrophoretic analysis permit the determination of their molecular weight. However, for glycoproteins, and *a fortiori* for lipopolysaccharides and for polysaccharides, the only precise approach remains mass spectrometry. New ionization methods such as electrospray (43) and matrix-assisted laser desorption (44, 45) appear to hold the greatest promise for the mass spectrometric analysis of large biopolymers. Today, using these ionization modes, routine mass spectrometric analysis of proteins in the mass range up to 200 kDa (45) with only picomole amounts of sample (46) has become possible. However, glycoproteins, glycolipids, and carbohydrates still remain quite refractory to mass spectrometric analysis.

In the present paper, we report that matrix-assisted LDI-MS is a powerful tool for determining the molecular mass of lipoarabinomannans (LAMs). Matrix-assisted LDI spectra of LAMs and LMs from *M. tuberculosis* and from *M. bovis* BCG are dominated by two peaks at around *m/z* 17,300 and 8,000. By means of SDS-PAGE, all the investigated native LAMs showed similar electrophoretic behavior characterized by broad bands between 30 and 35 kDa (Fig. 2). The molecular mass assignment of the native LAM from *M. bovis* BCG was supported by the LDI mass spectrum of the deacylated LAM showing a broad peak centered at 16.7 kDa. These data unambiguously revealed that the molecular masses of the LAMs from *M. bovis* BCG and from *M. tuberculosis* are quite similar at around 17.3 kDa. Thus, knowing the LAM molecular mass, established by mass spectrometry, provides key structural information for determining the molecular architecture of LAM. From the molecular mass and methylation data, it can be proposed that the major molecular species of both LAMs are composed of approximately 110 glycosyl residues: 70 assigned to arabinose and the remaining ones to mannose.

The peaks assigned to the pseudomolecular ions are extremely broad, and their width at half-maximum corresponds to 4 to 5 kDa, while the LDI spectra of the proteins having a similar molecular mass, used to calibrate the time-of-flight analyzer, exhibit a narrow peak with a width at half-maximum of 10 mass units. This phenomenon cannot be reasonably explained by an interglycosidic fragmentation process since oligosaccharide LDI spectra are dominated by cationized molecular ions (45). Thus, in agreement with the broad band observed with SDS-PAGE and the large peak on gel permeation, it was proposed that the width of the pseudomolecular signal mainly arises from the molecular heterogeneity of the LAM preparation. Moreover, since the LDI mass spectrum of the deacylated LAM does not show a larger peak width than that of the native LAM, it can be proposed that the molecular mass heterogeneity of LAM arises from molecular species containing a number of glycosyl units estimated to vary by around 40 suggesting that LAM heterogeneity arises from molecules having between 90 and 130 glycosyl units. However, the low mass resolution of the TOF analyzers does not allow their separation even using a TOF analyzer equipped with an ion reflector. Similar data have been reported in the literature when glycoproteins were analyzed by matrix-assisted LDI (44, 47). Nevertheless, it is the first time that the molecular mass of large bacterial lipopolysaccharides can be precisely measured by mass spectrometry giving a mass value which drastically differs from the one usually reported. The second relevant structural feature revealed by this analysis concerns the determination of the mass difference around 800 ± 100 mass units between the molecular mass of the LAM from *M. bovis* BCG and its deacylated form, respectively. This mass difference is almost in agreement with the structure proposed for the LAM of *M. tuberculosis* by Brennan and co-workers (7) in which the alkali groups are restricted to two fatty acids (tuberculostearate and palmitate) esterifying the phosphoglyceryl unit, but does not preclude the presence, in small amounts, of other alkali groups. The molecular mass and the molecular heterogeneity of LAMs from the protective strain, *M. bovis* BCG, and from the pathogenic strain, *M. tuberculosis*, are quite similar.

In order to determine the structure of LAM from *M. bovis* BCG, besides the routine methylation analysis, a novel analytical approach was successfully applied based on 2D NMR

spectroscopy and FAB/MS. 2D homo- and heteronuclear scalar coupling analysis of the *M. bovis* BCG LAM in combination with methylation data revealed, besides the 2,6-di-*O*-linked Man*p* present in the mannan domain, the presence of another sort of 2-*O*-linked Man*p* and also of two kinds of terminal Man*p*. One type occurred in the mannan domain (Fig. 11), while the second type was localized at the nonreducing end of the arabinan side chain by the following analytical approach. After mild hydrolysis, the oligosaccharides were derivatized with aminobenzoic ethyl ester (ABEE). The derivatized oligosaccharides were purified by reversed-phase HPLC under UV control and further sequenced by FAB/MS and FAB/MS-MS analysis. From the pseudomolecular ion and the sequencing fragment ions, it was clearly established that the major oligosaccharide is the trisaccharide: Man*p*(1→2)Man*p*(1→)Ara*f*. Besides this trisaccharide, a tetrasaccharide with the following structure Man*p*(1→2)Man*p*(1→2)Man*p*(1→)Ara*f* and the disaccharide Man*p*(1→)Ara*f* were found in small amounts. Thus, these data unambiguously demonstrate that the arabinan side chains are encapped by mannosyl residues allowing the proposition of a model for the molecular architecture of the *M. bovis* BCG LAM (Fig. 11). This structure is in contradiction with the results, based on methylation analysis, obtained by Misaki *et al.* (3) in 1977 who concluded that all of the 2-*O*-linked Man residues were part of side chains attached to the 6-*O*-linked mannan backbone.

Our analytical approach was also applied to the LAM of *M. tuberculosis* strain H37Rv leading to similar structural features for both LAMs. Further experiments were undertaken to purify large oligoglycosyl fragments with branched and linear regions of the arabinan and mannan segments after mild hydrolysis of LAM in order to elucidate its structure completely. Compared to methods with numerous chemical steps, the ease of preparation and separation of derivatized oligosaccharides and the resolution and sensitivity of this technique greatly influenced the choice of our strategy. Derivatization of the reducing termini with suitable chromophores such as *p*-aminobenzoate permits UV detection for purification by reversed-phase HPLC with a solvent gradient. The minimal increase in molecular mass of such derivatives and the conferred hydrophobic character enhance the FAB mass spectrometry sensitivity and improve assignment of the sequencing fragment ions, compared to underivatized carbohydrates. Webb *et al.* (48) reported also predictable differences in fragment ion abundance between linear and branched oligomers, suggesting that in addition to sequencing, branching information may also be determined by this method.

This study establishes that LAMs from an attenuated strain, *M. bovis* BCG, and from a virulent strain, *M. tuberculosis* H37Rv, share the same structure presenting mannosyl residues at the end of arabinan side chains. From the literature data, it was found that LAMs from other strains of *M. tuberculosis*, strain Erdman (18), and Ayoama B (49) have their arabinan side chains encapped with Man*p* residues. It can also be concluded that, when compared to our own results, the 1D ^{13}C NMR spectrum of LAM from *M. paratuberculosis* (50) indicates that this LAM was endowed with the Man-capped arabinan. However, it was established by Brennan and co-workers (6) that the avirulent strain, *M. tuberculosis* H37Ra, is devoid of mannosyl residues at the nonreducing end of the arabinan domain.

LAM from an avirulent strain of *M. tuberculosis* H37Ra induces, at low concentrations, TNF secretion (16), whereas LAM from the virulent Erdman strain, which possesses Man at the termini of the arabinosyl side chains, has 100-fold less activity (19). TNF is thought to play a central role in the pathogenesis of mycobacterial diseases through antimycobacterial activity as well as immunopathologic effects such as granuloma and tissue necrosis (51–55). A deficit of this cytokine would no longer hinder the proliferation of the virulent mycobacteria within host macrophages. So, for these authors, Man-capped LAM, which did not elicit TNF secretion, was synonymous of virulent and "disease-inducing strains."

In this study, the lack of structural difference between the two Man-capped LAMs from *M. bovis* BCG and *M. tubercu-*

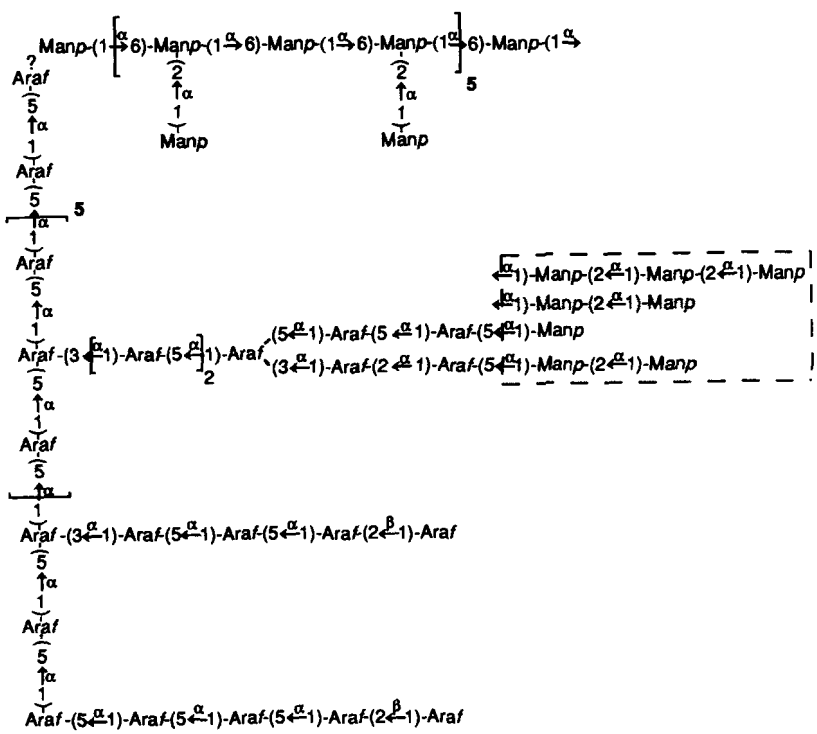


FIG. 11. **A model of the arabinomannan component of LAM from *M. bovis* BCG.** This structural hypothesis takes into account the molar ratios of the linked glycosyl residues and the molecular mass and shows the major structural motifs that occupy the nonreducing termini of the arabinofuranosyl side chains. The other structural features shown here were established previously (6, 18). Details of the attachment of the arabinan segments to the mannan core are unknown.

losis H37Rv apparently does not support the hypothesis that LAM is a key factor in the immunopathology of tuberculosis. This structure of Man-capped LAM from a given mycobacterial strain may influence the ability of the organism to survive but not to give rise to active infection. The Bacille Calmette-Guérin has been employed as a tuberculosis vaccine for 7 decades, and, after the infection with *M. bovis* BCG, the host may acquire protective immunity against subsequent tuberculous infection. The bacillus has to multiply or to survive in the macrophage without causing any pathological damage. Nevertheless, mechanisms involved in the acquired immune response to *M. bovis* BCG remain unclear. Man-capped LAM can hence be considered as a factor of intramacrophagic mycobacterial survival.

To confirm this immunoregulatory role of LAM, further experiments are now being undertaken to determine the dose effect of LAMs from *M. tuberculosis* H37Rv and *M. bovis* BCG in their different biological functions, such as immunosuppression of T cell proliferation and immunomodulation of the release of cytokines. For example, the role of LAM from *M. bovis* BCG in the secretion of TNF remains to be defined and quantitatively compared with the effect of LAM from a virulent strain.

LAM exhibits a wide spectrum of immunomodulatory effects. Identification of the structure, in correlation with the function, of such immunomodulators from different pathogenic or nonpathogenic mycobacteria will enhance our understanding of the pathogenesis of tuberculosis and of the protective effect of *M. bovis* BCG and is essential for the development or improvement of a vaccine against the disease.

Acknowledgments—We thank Dr. M.-A. Lanéelle for the kind gift of the *M. tuberculosis* H37Rv cultures, Dr. M. Gheorghiu for the *M. bovis* BCG cultures, Prof. Y. LeBeyec for the use of the TOF mass spectrometer with pulsed UV laser and data system, and Prof. J. Vercauteren for access to the 500-MHz NMR spectrometer and data system.

REFERENCES

1. Daniel, T. M. (1984) in *The Mycobacteria. A Sourcebook* (Kubica, G. P., and Wayne, L. G., eds) pp. 417–465, Marcel Dekker Inc., New York
2. Azuma, I., Ajisaka, M., and Yamamura, Y. (1970) *Infect. Immun.* **2,** 347–349
3. Misaki, A., Azuma, I., and Yamamura, Y. (1977) *J. Biochem. (Tokyo)* **82,** 1759–1770
4. Weber, P. L., and Gray, G. R. (1979) *Carbohydr. Res.* **74,** 259–278
5. Hunter, S. W., Gaylord, H., and Brennan, P. J. (1986) *J. Biol. Chem.* **261,** 12345–12351
6. Chatterjee, D., Bozic, C. M., McNeil, M., and Brennan, P. J. (1991) *J. Biol. Chem.* **266,** 9652–9660
7. Hunter, S. W., and Brennan, P. J. (1990) *J. Biol. Chem.* **265,** 9272–9279
8. Chatterjee, D., Hunter, S. W., McNeil, M., and Brennan, P. J. (1992) *J. Biol. Chem.* **267,** 6228–6233
9. Gaylord, H., Brennan, P. J., Young, D. B., and Buchanan, T. M. (1987) *Infect. Immun.* **55,** 2860–2863
10. Ellner, J. J., and Daniel, T. M. (1979) *Clin. Exp. Immunol.* **35,** 250–257
11. Kaplan, G., Gandhi, R. R., Weinstein, D. E., Lewis, W. R., Patarrayo, M. E., Brennan, P. J., and Cohn, Z. A. (1987) *J. Immunol.* **138,** 3028–3034
12. Moreno, C., Mehlert, A., and Lamb, J. (1988) *Clin. Exp. Immunol.* **74,** 206–210
13. Barnes, P. F., Fong, S.-J., Brennan, P. J., Twomey, P. E., Mazumder, A., and Modlin, R. L. (1990) *J. Immunol.* **145,** 149–154
14. Sibley, L. D., Hunter, S. W., Brennan, P. J., and Krahenbuhl, J. L. (1988) *Infect. Immun.* **26,** 1232–1236
15. Sibley, L. D., Adams, C. B., and Krahenbuhl, J. L. (1990) *Clin. Exp. Immunol.* **80,** 141–148
16. Moreno, C., Taverne, J., Mehlert, A., Bate, C. A. W., Brealey, R. J., Meager, A., Rook, G. A. W., and Playfair, J. H. L. (1989) *Clin. Exp. Immunol.* **76,** 240–245
17. Brennan, P. J., Hunter, S. W., McNeil, M., Chatterjee, D., and Daffé, M. (1990) in *Microbial Determinants of Virulence and Host Response* (Ayoub, E. M., Cassel, G. H., Branche, W. C., and Henry, T. J., eds) pp. 55–75, American Society for Microbiology, Washington, D.C.
18. Chatterjee, D., Lowell, K., Rivoire, B., McNeil, M. R., and Brennan, P. J. (1992) *J. Biol. Chem.* **267,** 6234–6239
19. Chatterjee, D., Roberts, A. D., Lowell, K., Brennan, P. J., and Orme, I. M. (1992) *Infect. Immun.* **60,** 1249–1253
20. Laemmli, U. K. (1970) *Nature* **227,** 680–685
21. Ansorge, W. (1983) in *Electrophoresis 82* (Stathakos, D., ed) 4th Ed, pp. 235–242, Walter de Gruyter, Berlin
22. Tsai, C. M., and Frasch, C. E. (1982) *Anal. Biochem.* **119,** 115–119
23. Thomas-Oates, J. E., and Dell, A. (1989) *Biochem. Soc. Trans.* **17,** 243–245
24. Sweet, D. P., Shapiro, R. H., and Albersheim, P. (1975) *Carbohydr. Res.* **40,** 217–225
25. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113,** 967–974
26. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magnetic Res.* **48,** 286–292
27. Aue, W. P., Bartholdi, E., and Ernst, R. R. (1976) *J. Chem. Phys.* **64,** 2229–2246
28. Bax, A., Griffey, R. H., and Howkins, B. L. (1983) *J. Magnetic Res.* **55,** 301–315
29. Shaka, A. J., Barker, P. B., and Freeman, R. (1985) *J. Magnetic Res.* **64,** 547–552
30. Bax, A., and Summers, M. F. (1986) *J. Am. Chem. Soc.* **108,** 2093–2094
31. Lerner, L., and Bax, A. (1986) *J. Magnetic Res.* **69,** 375–380
32. Hitchock, P. J., and Brown, T. M. (1983) *J. Bacteriol.* **154,** 269–277
33. Caroff, M., Chaby, R., Karilian, D., Perry, J., Deprun, C., and Szabo, C. (1990) *J. Bacteriol.* **172,** 1121–1128
34. Stahl, B., Steup, M., Karas, M., and Hillenkamp, F. (1991) *Anal. Chem.* **63,** 1463–1466
35. Mock, K. K., Davey, M., and Cottrell, J. S. (1991) *Biochem. Biophys. Res. Commun.* **177,** 644–651
36. Egge, H., Peter-Katalini, J., Karas, M., and Stahl, B. (1991) *Pure Appl. Chem.* **63,** 491–498
37. Gorin, P. A. J., and Mazurek, M. (1976) *Carbohydr. Res.* **48,** 171–186
38. Mizutani, K., Kasai, R., Nakamura, M., Tanaka, O., and Matsuura, H. (1989) *Carbohydr. Res.* **185,** 27–38
39. Hervé du Penhoat, C., Michon, V., and Goldberg, R. (1987) *Carbohydr. Res.* **165,** 31–42
40. Daffé, M., Brennan, P. J., and McNeil, M. (1990) *J. Biol. Chem.* **265,** 6734–6743
41. Gorin, P. A. J. (1981) *Adv. Carbohydr. Chem. Biochem.* **38,** 13–104
42. Abeygunawardana, C., and Bush, C. A. (1990) *Biochemistry* **30,** 378–388
43. Whitehouse, C. M., Dreyer, R. N., Yamashita, M., and Fenn, J. B. (1985) *Anal. Chem.* **57,** 675–679
44. Karas, M., Bahr, U., and Giessmann, U. (1991) *Mass Spectrom. Rev.* **10,** 335–357
45. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, T. (1991) *Anal. Chem.* **63,** 1193–1203
46. Karas, M., Ingendoh, A., Bahr, U., and Hillenkamp, F. (1989) *Biomed. Environ. Mass Spectrom.* **18,** 841–843
47. Chait, B. T., and Kent, S. B. H. (1992) *Science* **257,** 1885–1894
48. Webb, J. W., Jiang, K., Gillece-Castro, B. L., Tarentino, A. L., Plummer, T. H., Byrd, J. C., Fisher, S. J., and Burlingame, A. L. (1988) *Anal. Biochem.* **169,** 337–349
49. Ohashi, M. (1970) *Jpn. J. Exp. Med.* **40,** 1–14
50. Sugden, E. A., Samagh, S. S., Bundle, D. R., and Duncan, J. R. (1987) *Infect. Immun.* **55,** 762–770
51. Rook, G. A. W., and Attiyah, R. (1991) *Tubercle* **72,** 13–20
52. Bermudez, L. E., and Young, L. S. (1988) *J. Immunol.* **145,** 3006–3013
53. Denis, M., and Gregg, E. O. (1990) *Immunology* **71,** 139–141
54. Schnittman, S., Lane, H. C., Witebsky, F. G., Gosey, L. L., Hoggan, M. D., and Fauci, A. S. (1988) *J. Clin. Immunol.* **8,** 234–243
55. Kindler, V., Sappino, A. P., Grau, G. E., Piguet, P. F., and Vassalli, P. (1989) *Cell* **56,** 994–998