Truncated Structural Variants of Lipoarabinomannan in *Mycobacterium leprae* and an Ethambutol-Resistant Strain of *Mycobacterium tuberculosis**

Jordi B. Torrelles‡1, Kay-Hooi Khoo§2, Peter A. Sieling¶, Robert L. Modlin¶, Nannan Zhang‡, Angela M. Marques‡, Achim Treumann ||, Christopher D. Rithner**, Patrick J. Brennan‡ and Delphi Chatterjee‡§§

*This work was supported by Contract NO1 AI 55262 (to CSU; PJB, PI) from NIAID, NIH and Grants AI-47197 (PJB, PI), AI-22553 (RLM, PI) and AI-37139 (DC, PI) from the NIAID, NIH and in part by the Academia Sinica and National Science Council, Taipei, Taiwan. (NSC 89-2311-B-001-186) to KHK; AT was supported by the Wellcome Trust; JBT was affiliated with the Universitat Autonoma de Barcelona.*

1,2 These two authors have contributed equally

|| Present address: Royal College of Surgeons in Ireland, 123 St Stephen's Green, Dublin 2, Ireland

§§ To whom correspondence may be addressed. Tel: 970-491-7495; Fax: 970-491-1815; E-mail: delphi@lamar.colostate.edu

Abbreviations

Araf, arabinofuranose; 2 DE, 2 dimensional electrophoresis; DEAE, diethylaminoethyl; ESI-MS, electrospray ionization-mass spectrometry; ethambutol, Emb; GC, gas chromatography; HPAEC, high pH anion exchange chromatography; HSQC, heteronuclear single quantum correlation spectroscopy; IEF, isoelectric focussing; LepLAM, LAM from *M. leprae*; RvLAM, LAM from *M. tuberculosis*, H37Rv; CSU20 LAM, LAM from *M. tuberculosis* clinical isolate CSU20; LAM, lipoarabinomannan; LM, Lipomannan; Manp, mannopyranose; mAb, monoclonal antibody; MS, mass spectrometry; MTX, 5-deoxy-5-methyl-5-thio-α-xylofuranose; MSX, 5-deoxy-5-methyl-5-sulfoxy-α-xylofuranose; PAGE, polyacrylamide gel electrophoresis; PAS staining, periodic acid Schiff staining; PBS, phosphate buffered saline; PIMs, phosphatidylinositolmannosides.
SUMMARY

Current knowledge on the structure of lipoarabinomannan (LAM) results primarily from detailed studies on a few selected laboratory strains of *Mycobacterium tuberculosis*, *Mycobacterium bovis*, BCG and *Mycobacterium smegmatis*. Our previous work was the first to report on the salient structural features of *M. tuberculosis* clinical isolates and demonstrated significant structural variations. A prime effort is to correlate a particular structural characteristic with observed differences in eliciting immunobiological response, especially in the context of CD1-restricted presentation of LAM to T cells. T cell clones derived from the cutaneous lesion of leprosy patients have been shown to recognize specifically LAM from *Mycobacterium leprae* and not from *M. tuberculosis* Erdman or H37Rv. Herein we provide further fine structural data on LAM from *M. leprae* (LepLAM) and a tuberculosis clinical isolate, CSU20 (CSU20LAM), which was unexpectedly recognized by the supposedly LepLAM specific CD1-restricted T cell clones. In comparison to the *de facto* laboratory LAM standard from *M. tuberculosis* H37Rv (RvLAM), LepLAM derived from *in vivo* grown *M. leprae* is apparently simpler in its arabinan architecture with a high degree of exposed, non-mannose capped termini. CSU20, an ethambutol resistant clinical isolate, on the other hand makes a vastly heterogeneous population of LAM ranging from rather small and non-mannose capped, to full length and fully capped variants. LepLAM and CSU20LAM contain a higher level of succinylation than RvLAM which, in the context of truncated or less elaborated arabinan, may contribute to selective recognition by T cells. LAM from all species could be resolved into discrete forms by isoelectric focusing based apparently on their arabinan heterogeneity. In the light of our current and more recent findings, we reason that all immunobiological data should be cautiously interpreted and the actual LAM variants that may be present *in vivo* during infection and pathogenesis need to be taken into consideration.
INTRODUCTION

Substantial progress in leprosy control through multi-drug therapy (MDT) over the last two decades has led to a dramatic decline in prevalence from 10-12 million cases in the early 1980s to 534,000 today, as reported by 110 countries (1). However, the current prevalence of 3.4 cases per 10,000 populations in the top endemic countries at the beginning of 2003 still exceeds the targeted elimination (1 per 10,000) of leprosy as a public health problem by the year 2000, as declared by the World Health Organization in 1991. In reality, the significance of leprosy as a disease has been undermined by the resurgence of tuberculosis in the late 1990s which directed much of the focus of mycobacterial research away from \textit{M. leprae} to its close relative, the tubercle bacillus. \textit{M. leprae} is unique in that it is non-cultivable in the laboratory with an exceptionally slow growth rate, perhaps the longest doubling time of all bacteria. The bacillus can only be isolated from a host, the nine-banded armadillo or humans, and by cultivation in mouse foot-pad (2). Furthermore, leprosy is manifested as a spectrum of disease states depending on the responsiveness or unresponsiveness of various components in the host cellular immune system (3). It is therefore important to identify specific molecular determinants of \textit{M. leprae} that would modulate the host immune response by dictating the molecular nature of host-pathogen interactions, the spectrum of cytokines induced and importantly, the T cell response elicited.

Unlike the more restricted expression of species- and/or strain-specific surface glycolipids, the cell wall associated lipoarabinomannan (LAM) and its related lipomannan (LM) are potent immuno-modulators commonly found in all \textit{Mycobacterium} spp. investigated to date. One of the more significant recent immunobiological finding is the firm identification of LAM from \textit{M. leprae} (LepLAM), and not that from \textit{M. tuberculosis} H37Rv (RvLAM), as the specific non-peptide antigens recognized by CD1b and CD1c restricted \( \alpha \beta^+ CD4^- CD8^- \) double negative T cell lines derived from the skin lesion of a leprosy patient (4,5). The CD1 protein family has been proposed to comprise a distinct lineage of antigen presenting molecules involved specifically in the acquisition of cell mediated immunity to persistent intracellular pathogens (6). Notably, the binding site in CD1 is hydrophobic and flexibly accommodates a variety of anchoring hydrocarbon chains while exposing the diverse polar moieties of the presented lipid antigens in a region accessible for recognition by T cell receptors. The T cell specificity demonstrated against LepLAM over RvLAM therefore indicates a significant structural difference between the two LAM and confers on LepLAM an antigenic role in mediating leprosy specific immunopathological responses.

LepLAM could be found in copious amounts in \textit{M. leprae} (7) in comparison to \textit{M. tuberculosis}. Nevertheless, due to the problems in obtaining \textit{M. leprae} source material for experimental studies, very little is currently known of its fine structure in contrast to the large body of work investigating LAM from the slow growing laboratory strains of \textit{M. tuberculosis} and \textit{M. bovis} BCG, as well as the fast growing \textit{M. smegmatis}, \textit{M. chelonae} and \textit{M. kansasii} (reviewed in (8,9). Stemming from these studies, the salient and common structural features of the extremely heterogeneous LAM are now relatively well established (Fig. 1) although many of the fine details on the arabinan and mannan framework remain poorly understood. The genuine technical difficulty in defining the size distribution, branching pattern and to localize precisely
the many known or implicated charged substituents on the arabinomannan has inevitably led to a perceivable emphasis on the non-reducing terminal capping functions which has been accorded a myriad of receptor binding and immunomodulatory roles. In contrast, no particular function has as yet been attributed to the characteristic terminal arabinan motifs, in which the degree of branching off the linear \( \text{Ara}\beta 1\rightarrow 2\text{Ara}\alpha 1\rightarrow [5\text{Ara}\alpha 1\rightarrow 3]_n \) chains to produce the branched \( \text{Ara}_b \) termini, \( \text{Ara}\beta 1\rightarrow 2\text{Ara}\alpha 1\rightarrow 5(\text{Ara}\beta 1\rightarrow 2\text{Ara}\alpha 1\rightarrow 3)\text{Ara}\alpha 1\rightarrow 5\text{Ara}\alpha 1\rightarrow \), is probably as important in determining the spatial conformation and density of the cap and hence its biological attributes.

The anti-tuberculosis drug, ethambutol (Emb) for example, could inhibit complete elaboration of the arabinan and consequently abolish its mannose capping functions in LAM from an Emb susceptible laboratory strain of \( M. \) tuberculosis at sub-minimal inhibitory concentration. Importantly, LAM from an Emb resistant \( M. \) tuberculosis clinical isolate (CSU19) was shown (10) to be more heterogeneous in the spread of size, ranging from full size to severely truncated, and have a significantly higher proportion of non-capped arabinan termini than RvLAM but retaining similar relative abundance of linear and branched arabinan termini. These and earlier studies demonstrated that the degree and chemical nature of capping functions on LAM is strain specific and dependent on the relative activities of the gene products involved in the biosynthesis of the arabinan which could be further affected by Emb. It also indicates that biological relevance of the capping function should be considered in the context of the underlying arabinan motifs.

Like LAM from \( M. \) tuberculosis, \( M. \) bovis BCG and \( M. \) avium, LepLAM is known to carry mannose caps but apparently at a very low level (11). As no detailed structural studies have ever been performed, it is unclear if its overall arabinan and mannan structural framework bear any resemblance to those of better characterized LAM such as RvLAM, especially in the light of the comparatively smaller genome of \( M. \) leprae due to massive gene deletion and decay (12,13). It is likely that one or more of the complement of glycosyltransferases in \( M. \) leprae and other gene products involved in the synthesis of LAM, may be absent or non-functional.

Based on revised sample preparation protocols for better conservation of sample materials, we were able to accumulate milligram quantities of pure LepLAM to enable a thorough structural study. We show in this work that LepLAM maintains most structural attributes of RvLAM but encapsulated within a simpler form. The low degree of mannose capping is paralleled by an overall smaller arabinan which echoes the concept of reductive evolution for an obligate intracellular mycobacterium. Similar to our previous findings, LAM from another Emb resistant \( M. \) tuberculosis clinical isolate, CSU20 (CSU20LAM), was likewise found here to be drastically altered both by size and by reduction in the degree of mannose capping in comparison to the RvLAM. Rather unexpectedly, CD1 restricted T cell clones responsive to LepLAM and not RvLAM was found to be equally responsive to CSU20LAM. This is the first evidence that differential cell mediated immunopathological response could be elicited by LAM differing not in the identity of the capping function but the overall arabinan framework. CSU20LAM from a clinical isolate is shown here to be more akin to LepLAM than RvLAM in both structural features and T cell immunogenicity.
EXPERIMENTAL PROCEDURES

Chemical reagents - All chemical reagents were of the highest grade from Sigma/Aldrich (Milwaukee, WI) unless otherwise specified. Milli-Q® water was used for all chemical reactions.

Growth conditions of M. tuberculosis H37Rv and CSU20 - M. tuberculosis CSU20 is a clinical strain isolated from a patient in Korea. The MIC (minimal inhibitory concentration) values of CSU20 for Emb, cycloserine and pyrazinamide were 17, 60 and 50 μg/ml respectively. Frozen stocks of the strains were plated on 7H11 agar plates (Difco) with oleic acid/albumin/dextrose/catalase enrichment supplement for 3 weeks. Early colonies were picked and transferred to 5 ml of glycerol-alanine salts (GAS) broth for two weeks at 37°C on a shaker. Large-scale cultures were initiated by using 1 ml of broth from the starter cultures to inoculate 50 ml of fresh media. After seven days at 37°C, 1 ml of the inoculated culture was transferred to 400 ml of broth at 37°C for two weeks, then following the same procedure the culture was upscaled to a litre and grown until late log phase and harvested. About 8-10 g of cells were obtained from 2 litres of culture. M. tuberculosis H37Rv was used as our reference standard.

Extraction of LAM/LM/PIMs from M. tuberculosis H37Rv and CSU20 - Wet cells (~ 10g) were delipidated with chloroform:methanol:water (10:10:3) for two hours (x2) at room temperature and dried. The dried biomass was suspended in breaking buffer containing protease inhibitor cocktail (pepstatin A, phenylmethylsulfonyl fluoride (PMSF), leupeptine), DNAse, and RNAse in PBS and disrupted mechanically using a French Press. Triton X-114 (Sigma/Aldrich) was added to the lysed cells to a final concentration of 8% (v/v) and, after cooling on ice, the solution was mixed at 4°C overnight. The cell wall was removed by centrifugation at 27,000 x g for 1 hour at 4 C, and the supernatant was incubated at 37 C to induce biphasic separation (14). The upper aqueous layer was mixed with the cellular debris and re-extracted as described above. The detergent layers were combined and the lipoglycans precipitated by the addition of 9 volumes of cold ethanol (95%, -20 °C). The precipitate was collected and treated with Proteinase K for 2 h at 60 C. The solution was dialyzed and lyophilized for further analysis.

Extraction of LAM/LM/PIMs from M. leprae - M. leprae was purified from nonirradiated armadillo livers by the Draper protocol (15). Pure wet cells (approximately 500 mg, obtained from 250 g of infected armadillo liver) were suspended in 2.5 ml PBS and disrupted mechanically using a probe sonicator (W-385 Heat Systems-Ultrasonic) to achieve over 90% breakage of cells. The suspension after cell breakage was centrifuged at 27,000 x g for 1 h at 4 C. The cell pellet was re-suspended in PBS and centrifuged as before. The combined supernatants from both centrifugations containing both cytosol and membrane were then centrifuged at 100,000 x g for 2 h. The resulting pellet, representing the membrane fraction of M. leprae, was re-suspended in PBS containing Triton X-114 (4% by volume), mixed well at 4 C, placed at 37 C to induce phase separation and treated with ethanol as described above. The ethanol precipitate was collected, dried, and partitioned between phenol and water (16). The aqueous layer from the phenol partition contained the majority of cellular LAM and LM. It was evaporated to dryness and was the basis for further purification steps.
Quantitation of LepLAM and RvLAM in whole bacteria - 30 mg (wet weight) of M. leprae and M. tuberculosis H37Rv, were washed twice with 10 mM NH₄HCO₃ and resuspended in 2 ml NH₄HCO₃ containing 1 mM PMSF. Cells were disrupted by intermittent probe sonication (Ultrasonic Homogenizer 4710; Cole and Palmer Instruments, Chicago, Ill.) for 25 and 35 min (60 secs bursts/60 secs of cooling). Approximately 95% of the cells were broken when the lysates were acid-fast stained and examined under the microscope. Protein concentrations of the whole cell lysates were measured by the bicinichoninic acid assay reagent (Pierce, Rockford, Ill). For quantitation of LAM/LM/PIMs, SDS-PAGE analysis was performed on the lysates and equal amounts of samples (based on proteins concentrations) were applied. Densitometry analysis of LAM and LM after silver-periodic acid Schiff (PAS) staining was performed using the public domain NIH Image Program (developed at the U.S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/).

Size fractionation, SDS-PAGE and immunoblotting - For LepLAM purification, a Sephacryl S-200 (Pharmacia) column (100 cm x 1.5 cm; Vt = 150 ml), was prepared by washing and suspending the gel in a buffer containing 0.2 M NaCl, 0.25% deoxycholate, 1 mM EDTA, 0.02% sodium azide, and 10 mM Tris, pH 8.0. For purification of LAM from the M. tuberculosis strains used in this study, HPLC was performed on a Rainin SD 200 series LC system fitted with a Sephacryl S-200 HiPrep 16/60 column in tandem with a HiPrep 16/60 Sephacryl S-100 column (Amersham Biosciences, Piscataway, NJ) equilibrated with the same eluent as for the open columns at a flow rate of 1 ml/min. SDS-PAGE and PAS staining (17) was used to monitor the elution profile of the fractions containing LAM and LM, which were then pooled and dialyzed at 37 °C without detergent followed by water for several days. LAM/LM fractions recovered were re-analyzed by SDS-PAGE to check for purity prior to detailed analysis. Sample concentrations were maintained at 2 μg in 10 μl of sample buffer.

Western blots were performed by transferring antigens (LAM/LM/PIMs, 5 μg each) from 15% SDS-PAGE to nitrocellulose membranes (Schleicher&Schuell, Keene, NH) and then probed with serum from patients with different clinical symptoms of leprosy and drug treatment status. Following incubation with the secondary anti-human IgG alkaline phosphatase, reactivity was detected with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium BCIP/NBT substrate (Sigma Fast Tablet substrate). Immunoblotting with mAb CS-35 and CS-40 were performed essentially as described (17). MAb CS-35 was generated against M. leprae and has been shown to cross react with LAM from M. tuberculosis as well as other mycobacteria. Recently, the primary epitope of recognition was defined as terminal branched Ara₆ motif present in both LAM and AG (18). MAb CS-40 on the other hand was raised against LAM from M. tuberculosis, Erdman (16) and reacts preferentially with the mannose capped LAMs, the precise epitope of recognition has not been established.

2D-Gel Electrophoresis (2DE) of LepLAM/CSU20LAM/RvLAM - LAM (10 μg) was resuspended in 20 μl of IEF sample buffer (7M urea, 5% ampholyte 4.5-5.4 (Amersham) and ampholyte 3-10 (Amersham), 2% Nonidot P-40, 5% β-mercaptoethanol, and applied over the upper electrode buffer covering the top of a 6% polyacrylamide isoelectric focusing tube gel (2mm inside diameter x 70 mm) containing 1.6%
ampholytes. The tube gels were focused for 3 h at 1 kV with degassed 20 mM NaOH as the upper electrode buffer (catholyte) and 20 mM H3PO4 acid as the lower electrode buffer (anolyte), subsequently treated with transfer buffer (2.9% SDS, 71 mM Tris-HCl, 0.003% bromophenol blue) for 30 min, applied on a SDS-PAGE and stained with silver-PAS staining.

The pI value of each band was determined to be 4.7, 4.9, 5.1, 5.3, 5.5, 5.8 and 6.8, respectively, according to an IEF pH gradient profile as detected with a pH meter. All bands showed positive responses to the monoclonal anti-LAM antibody CS-35. LM and PIMs were also subjected to 2DE analysis and each gave only one band spanning pI range of 5.0-5.4.

**Digestion with endoarabinanase and subsequent analyses** – Preparation of endoarabinanase *Cellulomonas gelida* have been described (19). To ensure complete digestion of LAMs, the reaction mixture was incubated overnight and aliquots were withdrawn and analyzed by SDS-PAGE to ensure production of the mannan core as gauged by its electrophoretic mobility compared with LM. The digestion products that contained both mannan core and the released oligoarabinosides were analyzed directly by Dionex Analytical HPAEC performed on a Dionex LC System fitted with a Dionex Carbopac PA-1 column. The oligoarabinosides were detected with a pulse-amperometric detector (PAD-II) (Dionex, Sunnyvale, CA). For a preparative scale, CSU20LAM (a total of 4 mg) in 2x2 mg aliquots were taken in 100 µl of water and digested with 50 µl of enzyme for 4 h, after which fresh 30 µl of enzyme was added and digestion continued for another 4 h. The digestion mixture was applied to a Biogel-P4 column at 50 °C (1000 mm x 15 mm, flow rate 10 ml/h). Fractions were collected and aliquots were analyzed by HPAEC and subsequently by ESI-MS. For NMR analysis, 5 mg of CSU20LAM digest was applied to an octyl Sepharose (Amersham) column (200 x 20 mm) and eluted with 15, 45, and 60% propanol in 0.1 M ammonium acetate. Fractions (1 ml) were collected and monitored by carbohydrate detection on a TLC plate using α-naphthol spray. The endoarabinanase-resistant core obtained (45 and 60% propanol fractions) after four consecutive digestions was further desalted through a HiPrep 26/10 (Amersham) column prior to analyses.

**Digestion with α-mannosidase** - Digestion with α-1,2,3,6 mannosidase (Prozyme, San Leandro, CA), was carried out in sodium phosphate buffer (150 mM, pH 7.0) initially for 10 h at 37 °C using 1 µl of the enzyme per 50 µg of LAM. Aliquots were withdrawn and analyzed by Dionex HPAEC. To ensure complete digestion, a further 1 µl of the enzyme was added and the enzymatic digestion was continued for a further 16 h. The digestion was terminated by denaturing the enzyme.

**Monosaccharide, fatty acids, and succinate composition** - Samples were hydrolyzed with 2M TFA, converted to alditol acetates and analyzed by GC using scylo-inositol as internal standard (20). GC of the alditol acetates were performed on a Hewlett Packard Gas Chromatography Model 5890 fitted with a SP 2380 column (30 m x 0.25 mm id) at an initial temperature of 50 °C for 1 min, increasing to 170 °C by 30 °C/min followed by 270 °C by 5 °C/min.

Methanolation of LAM samples followed by trimethylsilylation afforded fatty acid methyl ester for quantitation by GC/MS (heptadecanoic acid, C17:0, was used as internal standard). GC/MS of the fatty
acid methyl esters as carried out on a ThermoQuest Trace Gas Chromatograph 2000 (ThermoQuest, Austin, TX) connected to a GCQ/Polaris MS mass detector (ThermoQuest). The derivatives were dissolved in hexanes prior to injection on a DB-5 column (10 m x 0.18 mm id, 0.18 micron film thickness, J&W Scientific, Folsom, CA) at an initial temperature of 60 °C for 1 min, increasing to 130 °C at 30 °C/min and finally to 280 °C at 5 °C/min.

Succinates were quantitated by GC/MS of the octyl-succinates using 50-100 μg of LAM after octanolysis with 3N 1-octanol (99%, 100 μl) at 120 °C for 30 min. The derivatives were dissolved in hexanes prior to injection on a DB-5 column at an initial temperature of 60°C held for 1 min. The temperature was increased to 330°C at the rate of 30°C/min.

NMR Spectroscopy of LAMs and LAM core – One dimensional 1H and 13C was performed on a Varian Inova AM-400 MHz and two-dimensional 1H-13C HSQC (heteronuclear single quantum correlation spectroscopy) NMR spectra were acquired on a Varian Inova 500 MHz NMR spectrometer using the supplied Varian pulse sequences. Spectra were acquired after several lyophilizations in D2O, of 6-7 mg/0.6 ml in 100% D2O. The HSQC data was acquired with a 7kHz window for proton in F2 and a 15kHz window for carbon in F1. The total recycle time was 1.65 seconds between transients. Adiabatic decoupling was applied to carbon during proton acquisition. Pulsed field gradients were used throughout for artifact suppression but were not used for coherence selection. The data set consisted of 1K complex points in t2 by 256 complex points in t1 using States-TPPI. Forward linear prediction was used for resolution enhancement to expand t1 to 512 complex points. A cosine-squared weighting function and zero-filling were applied to both t1 and t2 prior to the Fourier transform. The final resolution was 3.5 Hz/pt in F2 and 15 Hz/pt in F1.

Electrospray Ionization-Mass Spectrometry analyses - Samples were perdeuteroacetylated with pyridine:d6-acetic anhydride (1:1, v:v) at 80°C for 2 h after which the reagent was removed under a stream of nitrogen. Electrospray Ionization (ESI-) mass spectrometry (MS) analyses were performed on an Autospec orthogonal acceleration-time of flight (oa-TOF) mass spectrometer (Micromass, United Kingdom), fitted with an ESI source assembly and operated at 4 kV accelerating voltage. Samples were dissolved in methanol and 10 µl aliquots were injected through a Rheodyne loop into the mobile phase (methanol:water:acetic acid, 50:50:1, v:v:v), delivered at a flow rate of 5 µl/min into the ESI source by a syringe pump.

CD1-restricted T-cell assays - CD1b-restricted T-cell lines derived from a skin biopsy of a cutaneous lesion of a leprosy patient were co-cultured with antigen presenting cells pre-incubated with RvLAM, LepLAM and CSU20 LAM as described (4). T-cell proliferation was measured by incorporation of [3H]-thymidine in the medium.
RESULTS

**Extraction and fractionation of LAM** - When compared to the isolation of LAM from *M. tuberculosis* (16,21-23) several restrictions need to be considered prior to the derivation of LAM from *M. leprae*. Firstly, the amount of bacteria available is severely limited. Secondly, all components such as proteins, lipids, and carbohydrates, need to be conserved for further research needs. Hence, the protocol described for the isolation of LepLAM is a deviation from the procedures described in the literature. Protease inhibitors were excluded from the initial breaking buffer to allow the recovery of native proteins to be applied later for use in immunological studies. A sonicator was used for disruption of cells as opposed to a Bead Beater or French Press, since *M. leprae* is easy to disrupt in this way, and membrane fractions obtained after centrifugation could be used as the main source of LAM, LM and PIMs without compromising the cytosol and cell wall as a source of proteins. For a rough quantification, equal amounts of wet bacilli from *M. tuberculosis* and *M. leprae* were disrupted by intermittent probe sonication and when about 95% of cell breakage was obtained, protein concentrations of the cell lysates were measured. Lysates (based on protein estimation) were electrophoresed on an SDS-PAGE and densitometry analysis of LAM and LM after silver-PAS staining showed that the average of six different concentrations of LepLAM was 3-4 folds higher than LAM recovered from *M. tuberculosis* H37Rv (RvLAM) whereas the average of four different concentrations of LepLM was 5 folds more than RvLM (Fig. 2).

For all other analysis, the LAM/LM/PIM enriched extracts were size-fractionated on a Sephacryl S-200 column and the eluates monitored by SDS-PAGE with silver-PAS staining. In comparison with the usual pattern afforded by fractionation of RvLAM/LM/PIMs, there was less overlap between LepLAM and LM in fractions such that LepLAM free of LM contamination could be readily isolated. This pattern of fractionation indicated that LepLAM had a narrower range of size distribution, the lower end of which did not overlap significantly with LM. In contrast, consecutive S-200 fractions of LAM isolated from the Emb-resistant clinical isolate CSU20 (CSU20LAM) gave increasingly “smaller” LAMs. This is reminiscent of the pattern reported for another Emb-resistant clinical strain, CSU19, which was shown to be distinct from that of RvLAM in that it was more heterogeneous in the spread of size, ranging from full size to severely truncated (10).

**Molecular composition and size** - Fractions containing pure LepLAM were pooled and were shown to react similarly to monoclonal antibodies CS-40 and CS-35 by Western blot analysis which indicated that LepLAM is structurally related to RvLAM and that both the terminal branched Ara₆ motif as defined by mAb CS35 (18) and the mannose-capped terminal arabinosyl epitopes as defined by CS-40 (16) are commonly present in both LAMs. However, LepLAM consistently exhibited a higher electrophoretic mobility indicative of a smaller size (Fig. 3). Monosaccharide composition analysis by GC showed an Ara to Man ratio of 1.2 to 1 which is comparable to that of the mannose-capped RvLAM. However, when the neutral sugar composition was quantified based on one inositol per mol of LAM, LepLAM yielded 44 Araf and 36 Manp residues in comparison to 64 Araf and 61 Manp residues in RvLAM (Table I). Thus, the
absolute neutral sugar composition is in support of an overall smaller LepLAM, consistent with its further migration on SDS-PAGE. The same analysis also led to an apparent conclusion that CSU20LAM was the smallest among the three LAM examined with only 28 Ara\(f\) and 24 Man\(p\) residues per inositol. However, this is an average figure from a full spectrum of CSU20LAMs with Ara:Man ratio ranging from approximately 1.5 to about 0.5.

The presence of additional charged substituents on LepLAM was first suggested by the higher salt concentration required (above 0.25M) to elute it off a DEAE Sephadex (Cl\(^-\) form) column; a 0-0.3M NaCl gradient in 0.2% Triton X 114 was required (data not shown). GC-MS analysis of the octyl ester derivatives of succinates obtained through octanolysis of intact LepLAM led to identification of a peak which co-eluted with authentic di-octyl succinate (24.6 min, [M+H]\(^+\) at \(m/z\) 343; fragment ions at \(m/z\) 213 and 157). Integration of this peak with respect to neutral sugar indicated that there could be in total an average of seven succinyl residues per mol of inositol in LepLAM and three to four in CSU20LAM. The fatty acids were analyzed by GC-MS after methanolation and trimethylsilylation. The ions at \(m/z\) 270 and 312 which signify hexadecanoic acids (C\(_{16}\)) and \(\Delta10\)-methyloctadecanoic acid (C\(_{19}\)) respectively were found in LepLAM and RvLAM, whereas CSU20LAM additionally contained an octadecanoic acid (C\(_{18}\), \(m/z\) 298) (Table I).

**Distinctive structural features of LepLAM by NMR analyses** - Purified LepLAM was analyzed using a combination of NMR experiments. One dimensional \(\[^{31}\text{P}\]\) NMR revealed one intense resonance at \(\delta0.06\) attributable to a phospho-diester belonging to the phosphatidylinositol anchor; there was no shift in the resonance of this peak when the pH was changed. The assignment of the resonances in a 1D \(\[^{1}\text{H}\]\) NMR is a considerable challenge as the protons are in a strongly overlapping region of the spectrum. However, two intense broad triplets of equal intensity centered at \(\delta2.45\) and \(\delta2.60\) ppm could be attributed to the methylene of the succinyl groups (22) in both LepLAM and CSU20LAM. Upon deacylation, the resonances for succinates and fatty acids (\(\delta1.4\) to 0.6 ppm) were removed, thus substantiating that these residues are all ester-linked. Integration of anomic protons (\(\delta4.8-5.4\) ppm) versus the succinate resonances indicated a total of 7 succinyl residues per molecule of LepLAM while the corresponding data indicated only 1 to 2 and 3 to 4 succinyl residues per molecule of RvLAM and CSU20LAM, respectively, based on their neutral sugar composition (125 residues for RvLAM, 80 for LepLAM, and 52 for CSU20LAM; Table I). Resonances of the overlapping anomic regions could be better assigned via the well-resolved 2D \(\[^{1}\text{H}\]^{13}\text{C}\) HSQC experiment and by referring to a body of published NMR data on the structure of LAM (22,24,25) and its endoarabinanase generated fragments (26). The glycosyl residue composition by NMR (Fig. 4) and the peak volumes are summarized in Table II. Spectral data on different LAMs were acquired on separate days and the chemical shift changed by 0.1-0.3 ppm. Chemical shifts for RvLAM were taken as a generic standard for simplification in presenting the data for LepLAM and CSU20LAM. Overall, as revealed in the figure, a reduction in volume intensity was apparent in all anomic carbons of LepLAM and a simplified anomic region in comparison to other ManLAMs typified.
the spectrum. In particular, there were noteworthy differences in the spectra of the LepLAM (Fig. 4A) and CSU20LAM (Fig. 4C) in comparison to that of RvLAM (Fig. 4B).

Based on the literature (22,26-28), evidence for the presence of mannose-capping on LepLAM was sought from the $^{13}$C resonance at $\delta 101.1$ ppm, correlating to anomeric protons at $\delta 5.17$ ppm of 2-$\alpha$-Manp from the mannose caps and anomeric protons at $\delta 5.13$ ppm for the 2,6-$\alpha$-Manp with $^{1}J_{H1,C1}$ coupling constants of $\approx 170$ Hz. The volume of the cross peaks for 2-$\alpha$-Manp ($^{13}$C $\delta 101.1$ ppm, $^{1}H$ $\delta 5.17$ ppm) was markedly reduced in LepLAM relative to RvLAM (Table II). Although the degree of mannose capping cannot be unambiguously extrapolated from NMR due to strong overlapping of 2-$\alpha$-Manp and 2,6-$\alpha$-Manp resonances, our previous studies by linkage analyses (11) agreed with the results obtained from the HSQC spectra. Overlapping anomeric carbons at $\delta 104.8$ correlated to protons at $\delta 5.08$ and was assigned to t-$\alpha$-Manp belonging to both the mannan core and the mannose caps. These cross peaks were less prominent in the LepLAM spectrum. Multiple distinct spin systems were observed centered between $\delta 110.4$-$\delta 109.5$ ppm correlating with protons between $\delta 5.20$-$\delta 5.10$. These were attributed to the 5-$\alpha$-Araf and 3,5-$\alpha$-Araf, which showed a marked complexity in RvLAM but yielded a much simpler spectra for LepLAM, indicating that the overall arabinan is simpler (Fig. 4 A, B). Two well-separated spin systems were identified for 2-$\alpha$-Araf (cross peaks for C-2 were detected at $\delta 88.7$ and $\delta 88.4$), namely those attached to the 3-position (2-$\alpha$-Araf→3, $\delta 108.3$, $\delta 5.26$) and 5-position (2-$\alpha$-Araf→5, $\delta 108.5$, $\delta 5.20$) of the 3,5-$\alpha$-Araf. As expected, the intensity of the latter peak, i.e. 2-$\alpha$-Araf→5, was stronger than that of 2-$\alpha$-Araf→3 due to incomplete branching off the 3-position which results in the terminal linear Ara$_4$ motif. LepLAM yielded only two sets of 5-$\alpha$-Araf along with the overlapping 3,5-$\alpha$-Araf, which is consistent with a simplification of the arabinan chain with lesser branching. The chemical shift ranging between $\delta 110$-$108$ was supportive of the $\alpha$-anomeric configuration of the Araf residues (24,29). The terminal $\beta$-Araf resonated at $\delta 103.4$, ($^{1}H$ at $\delta 5.16$ ppm) with overlapping 5-$\beta$-Araf (discussed below). The cross peaks corresponding to 5-deoxy-5-methyl-5-thio-$\alpha$-xylofuranose (5-MTX) (26,30) as has been described to be a common constituent of all $M$. tuberculosis LAMs, were clearly absent in the LepLAM spectrum. The HSQC spectrum also revealed two distinct sets of carbons at $\delta 33.5$ and $\delta 34.5$ with protons centered at $\delta 2.67$ and $\delta 2.5$ (data not shown). These were attributed to the two methylene groups of the succinyl derivatives according to Delmas et al., (22), although these were better resolved than reported.

**Distinctive structural features of CSU20LAM by NMR analysis** - NMR analyses of a heterogeneous population of CSU20LAM (Fig. 4C) showed no marked discrepancy with the spectrum of the RvLAM and other published spectra from $M$. bovis, BCG (25), $M$. chelonae, $M$. kansasii (27,28) and thus indicate that the overall branching pattern and mannose capping were fairly similar throughout (Fig. 4B). The presence of 5-MTX in CSU20LAM has already been shown in our earlier work (26). However, unlike RvLAM, during the course of the purification by size exclusion chromatography, a subpopulation of LAM shorter in size (by SDS-PAGE) was also isolated from this strain. In order to examine if the diminution in length was due to shorter arabinan (i.e. decrease in 5-$\alpha$-Araf) or lack of mannose capping (i.e. increase in t-Araf) as seen
in a previous clinical isolate (10), the 'truncated' or 'short' CSU20LAM was also analyzed by NMR for comparison (results not shown). It was reasoned that if most of the non-reducing termini in CSU20LAM were substituted with Manp, the vast majority of t-β-Araf will be substituted and thus appear as 5-β-Araf. Quantitation of 5-Araf by methylation analysis could be misleading due to the presence of 5-α-Araf. In the NMR spectra, there is a perfect overlap in the anomeric signals of t-β-Araf and 5-β-Araf, and hence they cannot be distinguished in this region. However, due to the substitution of the mannose caps on the 5-OH, there is a significant shift for the C-5 of the β-Araf. Comparing the HSQC spectrum of 'normal' CSU20LAM against that defined as 'short' structures showed that almost all the anomeric resonances were comparable (data not included). The only noticeable differences were in the intensities of the peaks belonging to Manp residues. In 'normal' CSU20LAM, there were more 2-α-Manp and t-α-Manp from the caps than from the core. In contrast, in the anomic region of 'short' LAM, 2-α-Manp and 2,6-α-Manp were found to afford approximately the same peak intensity as the t-α-Manp from core and from caps. From the same HSQC spectra, detailed analyses of the hydroxymethylene region showed that both LAMs had 5β-Araf (cross peaks at δ71 ppm) but the intensity ratio of C5 of 5-β-Araf/t-β-Araf (two sets of cross peaks centered at δ71 and δ66 ppm) was higher for 'normal' CSU20LAM than for 'short' LAM, indicating that 'short' LAM had more (two fold) free, non-capped, non-reducing terminal β-Araf. In summary, the data from 1H-13C-HSQC spectra provided evidence that 'short' LAM contained less α-Manp cap units, but the underlying arabinan structure was similar to that of 'normal' LAM, i.e. with a similar degree of branching.

**Endoarabinanase resistant core of CSU20LAM** - In order to examine the internal regions of the arabinan built on the mannan core, CSU20LAM was subjected to exhaustive digestion with endogenous arabinanase. After four consecutive digestions, a core was obtained that gave a diffused band on SDS-PAGE and was slightly larger than LM (results not shown). 2D-NMR analysis of the core (Fig. 5) revealed very distinct spin systems different from CSU20LAM. The core presented only two sets of cross peaks centered between carbons at δ110-δ110.2 ppm correlating with protons between δ5.25-δ5.18 ppm. The cross peaks corresponding to β-Araf, 2-α-Araf, 3,5-α-Araf and 2-α-Manp that were present in CSU20LAM prior to digestion were clearly absent (Table II), indicating that the non-reducing termini were removed. The cross peak at δ5.25 ppm was new, and was attributed to terminal α-Araf, which would arise from termination of the internal 5-α-Araf (δ5.18 ppm) chain not due to β-Araf capping but trimming by the endoarabinanase. The cross peak of carbon at δ101 ppm with proton at δ5.2 ppm corresponded to 2,6-α-Manp, whereas t-α-Manp centered at δ104.8 ppm and δ5.17 ppm. Due to the saturation of the water peak, which overlapped with 6-α-Manp, cross peak for this residue was not visible in this experiment. The NMR results were corroborated by HPAEC analysis, as no Ara2, Ara4 and Ara6 were produced when the core was repeatedly digested with the enzyme.

Interestingly, the 1H, 13C-HSQC NMR spectrum also revealed two distinct sets of carbons, with four cross peaks at δ32.5, δ34.0, δ36.1 and δ37.7 correlating with protons at δ2.72 and δ2.65, δ2.9 and δ2.8. These were attributed to the nonequivalent methylene groups of the succinyl derivatives. Because these
cross peaks were well separated, we assume that one is on the arabinan, as on previously reported 5-α-Araf (28) and the other one on the mannan backbone, as LM seems to also contain one succinate (data not shown). Thus, NMR analysis of the endoarabinanase resistant mannan core revealed the presence of succinyl residues in the internal structural domains of CSU20LAM.

**Mannose cap and the nonreducing terminal arabinan motifs** - The compositional analysis, reactivity against mAb CSU35 and 40, and the assignment of various NMR resonance signals provided a picture of the arabinan motifs of LepLAM that is consistent with previously reported linkage analysis data (11) in which the approximate 1:1 ratio of terminal Ara:f:2-linked Ara first indicated a very low level of capping. The presence of 2-linked Manp, together with detection by mass spectrometry (MS) analysis of ions corresponding to Man2Ara2 and Man3Ara2 among the acetolysates of LepLAM (11) nonetheless supported the presence of di- and tri-mannose caps, albeit at low abundance. As for CSU20LAM, the data obtained was indicative of its close resemblance to another Emb-resistant clinical isolates, CSU19, which we previously characterized (10) by mapping its nonreducing terminal arabinan motifs. To determine if the underlying α5-arabinan chains also terminate in either one of the two well defined motifs, namely a branched Ara5 or a linear Ara4, like all other LAMs examined to date (Fig. 1), intact LepLAM, RvLAM and CSU20LAM were subjected to endoarabinanase digestion and the resulting products were mapped directly by HPAEC and, after Bio-Gel P6 clean up, followed by MS analysis.

As documented previously (10), mannose-capped RvLAM yielded predominantly Ara2, Man2Ara4 and Man4Ara6 as the major products, since a dimannosyl cap is most commonly found on each β-Ara→2-α-Ara→ terminus. Using RvLAM as a standard reference, the HPAEC profile of the endoarabinanase digestion products gave the aforementioned three major peaks which were subsequently converted to Ara2, Ara4 and Ara6 upon further α-mannosidase digestion to remove the caps (Fig. 6A). ESI-MS analysis of the perdeuteracetol derivatives of the endoarabinanase digestion products afforded major [M+Na]+ molecular ions for the Man2Ara4 and Man3Ara6 at m/z 1613 and 2652, respectively (Fig. 7A). Non-capped Ara4, Ara5, and Ara6 were also present but at much lower abundance (m/z 1019, 1241, 1463) along with Man1Ara4, Man2Ara5, Man2Ara6, Man2Ara7, Man3Ara4, Man3Ara5, Man3Ara6, Man4Ara7, and Man5Ara6 at m/z 1316, 1835, 2057, 2280, 1910, 2133, 2355, 2874 and 2949, respectively. As characterized previously, this profile is consistent with the presence of smaller amounts of Man1 and Man3 cap, in combination with additional minor Ara5 and Ara7 digestion products, to give the overall heterogeneity pattern observed.

A drastically different profile was afforded by LepLAM in which the spectrum (Fig. 7B) was dominated by Ara4 (m/z 1019) and Ara6 (m/z 1463) with smaller amount of Ara5 (m/z 1241), Man2Ara4 (m/z 1613) and Ara7 (m/z 1685). This is in full agreement with the conclusion drawn from HPAEC analysis in which mainly the non-capped Ara2, Ara4 and Ara6 were detected along with a small amount of Man2Ara4 which, as expected, disappeared after α-mannosidase digestion (Fig. 6B). It was also evident that Ara2 was less abundant in LepLAM in comparison to the digestion products obtained from RvLAM. This could imply that
LepLAM has a shorter $\alpha$-5-arabinan chains or that its arabinan is less susceptible to further trimming into $\alpha$-Ara$\rightarrow$5-$\alpha$-Ara units. Interestingly, the severely limited mannose capping led to MS detection of the otherwise seldom observed ion series of Ara$_8$ ($m/z$ 1907), Ara$_9$ ($m/z$ 2130), Ara$_{10}$ ($m/z$ 2352), Ara$_{11}$ ($m/z$ 2574), and Ara$_{12}$ ($m/z$ 2796) which extended as high up as Ara$_{20}$ (data not shown), accompanied mainly by those with a Man$_2$ cap. Most of the species afforded by RvLAM could be found in the digest of LepLAM but at different relative abundance.

For the extremely heterogeneous CSU20LAM analyzed as a whole, it was clear that the proportion of non-capped Ara$_4$ and Ara$_6$, relative to those mannose-capped are higher than those in RvLAM. The MS spectrum produced was extremely complicated (data not shown) and could be considered as a superimposition of that from LepLAM on RvLAM. Three sets of mannose-capped peaks were prominent, namely i) Man$_2$Ara$_4$ and Man$_3$Ara$_4$, ii) Man$_2$Ara$_6$ and Man$_3$Ara$_6$, and iii) Man$_4$Ara$_6$, Man$_5$Ara$_6$, and Man$_6$Ara$_6$ which together, can be rationalized as corresponding to the linear Ara$_4$ termini with either a di- or trimannosyl cap, and branched Ara$_6$ termini with di- or trimannosyl caps on only one or both termini. To better resolve the MS analysis, the endoarabinanase digestion products from CSU20LAM were fractionated on a BioGel P4 column and collected individual fractions were screened through by ESI-MS after perdeuteroacetylation. Mass spectra from representative fractions are shown in Fig. 8. Analyzing individual fractions afforded detection of higher components from early eluting fractions otherwise obscured by the more abundant peaks when analyzing the total digest. Even so, at high mass end, the three mass units difference between four Ara residues and three Man residues require relatively accurate mass measurement to ascertain the molecular composition which is only possible for more abundant signals. In comparison with similarly fractionated RvLAM digestion products pooled for the larger components, it is obvious that most species afforded by RvLAM could be identified in CSU20LAM which additionally contained components such as Man$_2$Ara$_{12}$, Man$_3$Ara$_{11}$, Man$_3$Ara$_{12}$, Man$_3$Ara$_{13}$, Man$_3$Ara$_{14}$, and Man$_3$Ara$_{15}$. These are more similar to those of LepLAM which could be rationalized as deriving from either linear arabinan chains or branched chain capped only at one non-reducing terminus. In contrast, most of the larger arabinan components afforded by RvLAM contained Man$_4$, Man$_5$ or Man$_6$, consistent with the presence of Man$_2$ and/or Man$_3$ caps on each terminus of a branched arabinan motifs. Thus the overall nonreducing termini presented by CSU20LAM are very similar to CSU19LAM reported previously. The reproducible pattern given by both Emb-resistant clinical isolates indicate that this phenomenon of synthesizing a wide range of LAM, from those similar to mostly mannose-capped RvLAM to those simpler, under-capped LepLAM may be a common feature applicable to many M. tuberculosis clinical isolates. This heterogeneity in LAM is however in contrast with the observed structural features of a M. kansasii clinical isolate (28).

**Humoral and cell mediated immune response to LepLAM and CSU20LAM** - Despite significant downsizing and severe lack of mannose capping among its arabinan termini, LepLAM was still recognized by both CS-35 and CS-40 monoclonal antibodies. As expected, CSU20LAM was also recognized by these antibodies. Importantly, LepLAM was also recognized by individual serum samples
from leprosy patients. Patients across the spectrum of leprosy were selected at random and all sera showed positive reactions to LAM albeit differentially (Fig. 9). In contrast, LM was completely negative and interestingly, some sera reacted with the isolated PIMs. That the leprosy patients have high titre antibodies to LepLAM could be due to the fact that LepLAM is secreted or shed from the bacilli during the course of infection and is thereby readily recognized by the host. In fact, the presence of significant amounts of LepLAM in liver homogenate (supernatant obtained after the bacilli have been sedimented during the processing of *M. leprae*) has been demonstrated (results not shown).

Reactivity against sera indicated that LepLAM presents B-cell epitopes probably through the exposed arabinan. It is interesting to note that CD1b-restricted T cell lines derived from a skin biopsy of a cutaneous lesion of a leprosy patient has been shown to be reactive against the presented LepLAM. In the light of current findings in which CSU20LAM was demonstrated to be extremely heterogeneous and to some extent resembles LepLAM in containing populations of short and/or under-capped LAM, its reactivity against the leprosy derived CD1-restricted T cell clones were tested. The results showed that T cell clones LDN4 and LCD4, when treated with different concentrations of RvLAM, LepLAM and CSU20LAM, proliferated preferentially in response to LepLAM and CSU20LAM (Fig 10). Proliferation against LepLAM is of no surprise and simply acted as positive control. Failure to recognize RvLAM argues that certain specificity was maintained against the presented non-lipid epitope. From the same rationale, recognition of CSU20LAM indicated that it is more comparable to LepLAM than RvLAM. That two and not just one independent CD1-restricted T cell clones gave similar discriminating specificity suggested that the shared characteristic is immunodominant which could not be ascribed simply to presence or absence of mannose cap. Thus, for the first time, an immunogenic trait of LAM has been attributed to not the capping or fatty acyl functions but possibly, the underlying arabinan constructs.

**Additional isoforms in LepLAM and CSU20LAM** - 1D SDS-PAGE coupled with silver-PAS staining or Western blot with anti-LAM monoclonal antibodies have been useful in assessing the heterogeneity of LAM in size which typically gives broad diffuse band. Interestingly, if SDS-PAGE analysis was preceded by a first dimensional IEF separation, both LepLAM and CSU20LAM, as well as RvLAM, could be similarly resolved into several discrete isoforms equilibrating at different pH instead of a mere horizontal spread (Fig. 11A). Pre-fractionation into populations differing in the number of fatty acyl functions by hydrophobic interaction column chromatography did not alter the IEF blueprint as each of the pooled fractions gave the same number of isoforms across the 2DE. The data thus indicated that each LAM isoforms on the 2DE gel not only carried the same spread in size but also similar fatty acid composition and all contained the hexaarabinofuranoside Ara₆ motifs (18), demonstrated as positive reactivity against the monoclonal antibody CS-35 (data not shown). Focusing on CSU20LAM, treatment with α-mannosidase or removal of the caps with α-(1→2)-mannosidase did not collapse the isoforms except some size shifting (Fig 11B, top panel). However, as the arabinan was trimmed away by endoarabinanase digestion, the resulting mannan core could still be resolved into isoforms with more trailing of the bands and leaning towards lower pH (Fig. 11B, second panel). LM, on the other hand (Fig.
11 B, third panel), yielded one isoform.
DISCUSSION

LAM has been implicated in a diverse range of immunological attributes (31). The precise structural requirements for the recognition of LAM remain speculative and has been a basis of fundamental research in numerous laboratories. One consistency seemed to be the ability of LAM to induce TNF-α and IL-1β, its dependency on CD-14, and its synergization with IFN-γ in the induction of NO (31,32) in its native acylated form. Even recently it has been demonstrated that LAM served as a better ligand in its multiacylated forms in binding to human surfactant protein hSP-A (33). It has also been established that all immunological attributes of LAM are abolished upon treatment with mild alkali, which would essentially remove all ester linkages (i.e. fatty acyl and succinyl residues). There is also evidence suggesting that the mannose-capped RvLAM is generally less effective in stimulating macrophages than the non-capped LAM which may owe its potency to unusual inositol phosphate capping or even the exposed arabinan ((27) as reviewed in (9,31)). Schlesinger et al. noted differences in the level of uptake between ManLAMs isolated from virulent or attenuated strains of M. tuberculosis (34,35). Moreover, it has been shown that certain CD1b and CD1c-restricted T-cells recognize LAM and PIMs specifically, and one cell line BDN2e was able to distinguish between RvLAM and LepLAM, suggesting subtle structural differences (4,5).

Overall then, the body of evidence implicating LAM in modulating immune response such as that manifested by macrophage cytokine induction is compelling. However, despite many of the structural studies emerging from this laboratory and many others (reviewed in (8)), there has been a general deficit in two arenas, namely the structure to function relationship of LAM and the identification of the biosynthetic steps in its formation. The raison d’être being the inherent extreme heterogeneity of LAM not constructed out of regular repeating units like most polysaccharides, and a general lack of functional mutants. With respect to the mutants of LAM, a recent major achievement has been the identification of the Emb proteins (36,37) and the observation that with a dysfunctional EmbC, (a protein identified as a transmembrane arabinosyltransferase that seemed to add the Ara residues to the pre-existing 2 to 3 Ara residues in LM), nonpathogenic M. smegmatis is unable to synthesize LAM (38). In M. tuberculosis, the EmbC has proven to be essential (Parish and Chatterjee, unpublished work). These studies could not however be extended to address the biological roles of the respective structural motifs of LAM, in particular the arabinan.

The specific recognition of LepLAM by the sera of leprosy patients and in a CD1-restricted manner by T cells indicated that some as yet undefined structural features of LepLAM contributed to immunogenic discrimination against RvLAM but not CSU20LAM from M. tuberculosis. In essence, we are thus looking for some common structural signatures on the two LAM (LepLAM and CSU20LAM) that may be positively recognized, and that are apparently absent or masked in the third LAM (RvLAM). Founded on this basis, we demonstrated, in this study, that the arabinan in LepLAM is shorter and simpler, with minimal amount of mannose-capping. While fully elaborated and capped arabinan similar to that of RvLAM is also present in CSU20LAM, the latter is distinguished by the extra heterogeneity which, at the other end of the
structural variants of arabinans in LAM

spectrum, comprises a 'short' or 'truncated' population very similar to that of LepLAM. The exact location and degree of succinylation on the arabinan of LAM has yet to be resolved (22, 28) but our current findings indicated that it is not dependent on full arabinan chain elongation or elaboration. On the contrary, the simpler arabinan of LepLAM actually carries more succinates per molecule on average and likewise, more succinylation could be detected on CSU20LAM in comparison to RvLAM. It is likely that the succinates are preferably carried on the internal $\alpha$-5-Ara chain and/or become more exposed with less elaborated arabinan. Our data did indeed show that the CSU20LAM core obtained after exhaustive endoarabinanse digestion still retain the succinates despite losing the characteristic nonreducing terminal Ara motifs. Arguably, a more exposed succinylation proximal to the mannan core may constitute part of the conformational epitope recognized by T cell receptor in the context of CD1 presentation.

Other structural indices such as the common presence of mannose capping on all LAM, albeit at varying degrees, or the $\alpha$-methylthioxylofuranose in LAM from both M. tuberculosis strains but not from M. leprae, are less convincing candidates for common recognition of LepLAM and CSU20LAM by T cells. It should be further reminded that the identity of constituent Araf residues in the setting of $\alpha$5-elongation, $\alpha$3-branching off a 3,5-$\alpha$-Araf residue, and $\beta$2-chain termination, are essentially conserved between M. tuberculosis and M. leprae; only the relative amount of each is altered. Conventional analytical methods, as applied here and many other similar structural studies on total LAM population, often could not discriminate against such subtleties while attempts to resolve each LAM into distinct populations with differing biological activities based solely on differences in arabinan have consistently failed in the past. Our 2DE experiments present an attractive alternative route and indicate that IEF resolution of LAM into isoforms is perhaps based on the arabinan architecture rather than capping or fatty acyl functions. It suggests that the arabinan exerts a primary and significant effect on the effective pi of a heterogeneous molecule like LAM. Despite apparent pi similarity, all of the discrete isoforms of LAM from the three sources may not be the same and further studies are under way to delineate their structure-function correlation, in conjunction with new analytical and genetic tools.

In comparison with that of M. tuberculosis, the genome of M. leprae is unique in containing a large number of pseudogenes (nearly 40% of the genome), and a minimal set of functional genes. The significance of such a genome structure has been investigated in the context of understanding the physiology of the organism and pathogenesis (12,39-41). Despite a drastic gene reduction and decay, genes for synthesis of cell wall structures are reasonably intact (40), including the embC gene implicated in the synthesis of the arabinan of LAM (38). Comparative functional genomics studies should facilitate the quest in identifying all of the arabinosyltransferases involved, as well as how they may be regulated. Such mechanism may further shed light on how and why a large sub-population of CSU20LAM carries severely truncated or incompletely elaborated arabinan. Immunopathogenesis studies of tuberculosis and leprosy should also take into consideration that previous immunobiological findings based on RvLAM alone may not be a true reflection of pathophysiological conditions and that the actual in vivo activity is critically dependent on the size and structural variations of the arabinan elaborated on LAM.
REFERENCES

1. WHO. (2004), WHO
Structural variants of arabinans in LAM


Table I. Compositional analysis of LepLAM and CSU20LAM in comparison to RvLAM.
The sugars, succinates and fatty acyl composition were determined by separate GC-MS analysis while the presence of MTX (methylthioxylofuranose) was identified by NMR analysis.

<table>
<thead>
<tr>
<th></th>
<th>Ara</th>
<th>Man</th>
<th>myo-Ins</th>
<th>succiny</th>
<th>MTX</th>
<th>fatty acyl</th>
</tr>
</thead>
<tbody>
<tr>
<td>RvLAM</td>
<td>64</td>
<td>61</td>
<td>1</td>
<td>2</td>
<td>+</td>
<td>C\textsubscript{16,0}, C\textsubscript{19}</td>
</tr>
<tr>
<td>LepLAM</td>
<td>44</td>
<td>36</td>
<td>1</td>
<td>7</td>
<td>-</td>
<td>C\textsubscript{16,0}, C\textsubscript{19}</td>
</tr>
<tr>
<td>CSU20LAM</td>
<td>28</td>
<td>24</td>
<td>1</td>
<td>4</td>
<td>+</td>
<td>C\textsubscript{16,0}, C\textsubscript{18, C19}</td>
</tr>
</tbody>
</table>

Table II. Anomeric resonances of LepLAM, RvLAM and CSU20LAM.
Signal volumes were determined from \textsuperscript{1}H\textsuperscript{13}C HSQC NMR experiments performed on separate days using the same 500MHz NMR spectrometer (see Experimental section). Variations (0.01-0.4 ppm) on the ppm values for the reported chemical shifts of \textsuperscript{13}C and \textsuperscript{1}H have been corrected accordingly.

<table>
<thead>
<tr>
<th>Residues</th>
<th>LepLAM</th>
<th>RvLAM</th>
<th>CSU20LAM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>\textsuperscript{12}C</td>
<td>\textsuperscript{1}H</td>
<td>Vol\textsuperscript{a}</td>
</tr>
<tr>
<td>2,6-\alpha-Man\textsubscript{\text{p}} (core)</td>
<td>101.1</td>
<td>5.13</td>
<td>3.79</td>
</tr>
<tr>
<td>2-\alpha-Man\textsubscript{\text{p}} (cap)</td>
<td>101.1</td>
<td>5.17</td>
<td>2.43</td>
</tr>
<tr>
<td>?</td>
<td>102.6</td>
<td>5.39</td>
<td>0.38</td>
</tr>
<tr>
<td>6-\alpha-Man\textsubscript{\text{p}}</td>
<td>102.2</td>
<td>4.92</td>
<td>1.00</td>
</tr>
<tr>
<td>?</td>
<td>102.6</td>
<td>4.92</td>
<td>0.30</td>
</tr>
<tr>
<td>t-\beta-Ara\textsubscript{\text{f}} + 5-\beta-Ara\textsubscript{\text{f}}</td>
<td>103.4</td>
<td>5.16</td>
<td>11.06</td>
</tr>
<tr>
<td>?</td>
<td>103.4</td>
<td>5.30</td>
<td>0.43</td>
</tr>
<tr>
<td>t-\alpha-Man\textsubscript{\text{p}} (core+caps)</td>
<td>104.8</td>
<td>5.08</td>
<td>6.57</td>
</tr>
<tr>
<td>?</td>
<td>104.9</td>
<td>5.14</td>
<td>0.06</td>
</tr>
<tr>
<td>t-\alpha-Man\textsubscript{\text{p}} (core+caps)</td>
<td>105.0</td>
<td>5.03</td>
<td>3.25</td>
</tr>
<tr>
<td>?</td>
<td>105.0</td>
<td>5.16</td>
<td>0.05</td>
</tr>
<tr>
<td>MTX</td>
<td>105.2</td>
<td>5.46</td>
<td>0.41</td>
</tr>
<tr>
<td>?</td>
<td>106.3</td>
<td>5.38</td>
<td>0.07</td>
</tr>
<tr>
<td>2-Ara\textsubscript{\text{f}}\rightarrow 3</td>
<td>108.3</td>
<td>5.26</td>
<td>2.47</td>
</tr>
<tr>
<td>2-Ara\textsubscript{\text{f}}\rightarrow 5</td>
<td>108.5</td>
<td>5.20</td>
<td>8.62</td>
</tr>
<tr>
<td>5-\alpha-Ara\textsubscript{\text{f}}</td>
<td>109.5</td>
<td>5.26</td>
<td>0.35</td>
</tr>
<tr>
<td>5-\alpha-Ara\textsubscript{\text{f}}</td>
<td>110.0</td>
<td>5.19</td>
<td>5.97</td>
</tr>
<tr>
<td>5-\alpha-Ara\textsubscript{\text{f}}</td>
<td>110.2</td>
<td>5.14</td>
<td>9.73</td>
</tr>
<tr>
<td>3, 5-\alpha-Ara\textsubscript{\text{f}}\textsuperscript{b}</td>
<td>110.3</td>
<td>5.12</td>
<td>1.22</td>
</tr>
<tr>
<td>5-\alpha-Ara\textsubscript{\text{f}}</td>
<td>110.1</td>
<td>5.12</td>
<td>8.21</td>
</tr>
<tr>
<td>5-\alpha-Ara\textsubscript{\text{f}}</td>
<td>110.4</td>
<td>5.11</td>
<td>23.71</td>
</tr>
<tr>
<td>5-\alpha-Ara\textsubscript{\text{f}}</td>
<td>110.5</td>
<td>5.03</td>
<td>0.26</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Signal volumes were integrated and normalized to 6-\alpha-Man, taken as the reference signal. Signals not yet assigned are denoted by "?", most of which could only be detected in the RvLAM spectrum but not afforded by the simpler or truncated LepLAM and CSU20LAM.

\textsuperscript{b} There is a significant overlap of 3, 5-\alpha-Ara\textsubscript{\text{f}} with 5-\alpha-Ara\textsubscript{\text{f}} and thus the integration of signal volumes of 3, 5-\alpha-Ara\textsubscript{\text{f}} is impossible to assign accurately. This could be the reason for a lower value of this peak in this Table.
FIGURE LEGENDS

Fig. 1. Schematic structure of ManLAM. Representation of the tripartite structure of ManLAM showing only the mannose-capped non-reducing termini, the mannan core, and the PI anchor. This is a generalized figure and does not depict modifications like MTX, succinates or inositol phosphate capping which are known to be present in LAM from various biological sources.

Fig. 2. SDS-PAGE analysis of the M. leprae and M. tuberculosis H37Rv sonicates. 30 mg of cells (wet weight) were re-suspended in 2 ml 10 mM NH4HCO3 and disrupted by sonication. (A) Electrophoresis was performed on 6% stacking gel, 15% running gel with 10 µl of the M. leprae and M. tuberculosis total sonicates (lane 1 and 2) and the proteins stained with silver. (B) 20, 10, 5, 2.5, 1.25 and 0.65 µl of M. leprae (lanes, 1, 2, 3, 4, 5 and 6) and M. tuberculosis (lanes 7, 8, 9, 10, 11 and 12) sonicates were electrophoresed and PAS staining was applied immediately following fixation to enhance the appearance of LAM, LM and PIMs.

Fig. 3. SDS-PAGE analyses of LAMs. The size difference amongst the different LAMs can be approximately gauged by aligning the respective front ends of the afforded broadly diffused electrophoretic bands. It is evident from the SDS-PAGE profile that RvLAM exhibited a lower electrophoretic mobility than the other two LAMs, consistent with it being about 20 glycosyl residues larger as estimated by neutral sugar composition analysis.

Fig. 4. Comparative partial 2D NMR spectra of LepLAM, RvLAM and CSU20LAM. The 2D NMR 1H-13C HSQC spectra of LepLAM (A), RvLAM (B) and CSU20LAM (C) were acquired in D2O. Only the expanded anomic regions are shown. The intensity of peak volumes were measured and the data presented in Table II.

Fig. 5. 2D NMR spectrum of the mannan core of CSU20LAM. The HSQC spectrum of the core was acquired in D2O. Only the low field region is shown. The anomic region is simpler than the CSU20LAM (Fig.4C) because of the loss of resonances for 3,5-α-Araf, multiple 5-α-Araf, 2-α-Araf, t-β-Araf and 2-α-Manp post-endoarabinanase digestion. The two sets of paired crosspeaks in the inset between δ32-δ38 ppm are the succinates. In the parent LAM, there is only one set. We reason that trimming of the arabinan causes environmental changes, thus exposing succinates from the inner region (residual arabinan and mannan) of LAM. The basis of the difference in intensity of the two sets of peaks in not yet known. RvLAM core gave similar pattern, except the peaks were of equal intensity and LepLAM was not put through the digestion.

Fig. 6. Dionex HPAEC profile of endoarabinanase digested RvLAM and LepLAM. RvLAM (A) and LepLAM (B) before (top panel) and after (lower panel) α-mannosidase digestion. The lack of characteristic mannose-capped structures is evident in the non α-mannosidase treated LepLAM. For
Structural variants of arabinans in LAM

direct comparison, the digestion products are dried and injected directly without further purification.

**Fig. 7. ESI-MS analysis of RvLAM and LepLAM.** Analysis of the perdeuteroacetyl derivatives of the endoarabinanase digestion products of RvLAM (A) and LepLAM (B). Assignment of the major peaks detected were as annotated (M, Man; A, Ara). Mass spectrum for LepLAM was magnified from the mass region above \( m/z \) 1700 to show the minor larger sized components. Both spectra also contained components at above \( m/z \) 3000 and a prominent Ara\(_2\) peak at the lower end of the spectra.

**Fig. 8. ESI-MS analysis of CSU20LAM.** Analysis of the perdeuteroacetyl derivatives of the endoarabinanase digestion products of CSU20LAM after fractionation on BioGel-P4 column. The mannan core with a few Ara\(_f\) attached eluted in the void pool of the P4 column. Each of the collected fractions were individually screened through and only representative spectra from #40, 42, 45, 51 and 55 are shown. There is significant overlap of components collected within each of the consecutive fractions. As a consequence of fractionation into several successive fractions, the relative abundance of each component could not be estimated from any single spectra. Major signals were assigned as annotated (M, Man; A, Ara). For comparison, the spectrum for a pooled fraction containing the larger components from RvLAM was included. The smaller and more abundant components from RvLAM were as shown in Fig. 7A.

**Fig. 9. Humoral immune response of leprosy patients to LepLAM.** Western blots were developed using individual patients sera (1:5,000 dilution). 5 \( \mu \)g of LepLAM was used as the antigen (only lane 1 in blots ‘b’ through ‘g’ is shown). LM and PIMs (5 \( \mu \)g each) were also used in this experiment (lanes 2 and 3 in serum ‘a’). LM was consistently negative (‘b-g’), and PIMs were faintly positive for only sera ‘a’ and ‘h’. Sera ‘a,b,d,g’ and ‘h’ were from lepromatous (LL) forms of leprosy and sera c,e, and f were from borderline lepromatous (BL) patients.

**Fig. 10. Proliferative responses of T cell lines to LAM.** CD1-b restricted T cell lines, LCD4 and LDN4, were cultured with antigen presenting cells pre-incubated with different concentrations of RvLAM, LepLAM and CSU20LAM. T cell proliferation was measured by incorporation of \(^{[3}\text{H}]\)-thymidine incorporation on day 3 from the medium.

**Fig. 11. 2DE of LAM.** (A) Comparative 2DE profiles of LepLAM (upper panel), RvLAM (middle panel) and CSU20LAM (lower panel). 12 \( \mu \)g of each LAM sample (minimum amount required to obtain separation) were applied to the IEF gel. The pl value of each band is determined by IEF pH gradient profile (pl 4.7-5.8). (B) Contribution of the various structural motifs to resolution on 2DE was investigated by treating CSU20LAM with \( \alpha \)-mannosidases (upper panel) and endoarabinanase (middle panel). The relative running position of LM is shown in the lower panel for comparison.
Fig. 1

Terminal $\text{Ara}_4 / \text{Ara}_6$

$\alpha$-3-branching

$\alpha$-5-Arabinan chains

Mannose Cap

Lipomannan

$\alpha$-6 mannan

$\alpha$-2-Man branches

Phosphatidylinositol mannoside

LAM
Fig. 4A

- 2,6-α-Manp (Core)
- 2-α-Manp (Caps)
- t-β-Araf + 5-β-Araf
- t-α-Manp (core + caps)
- 2-α-Araf → 5
- 5-α-Araf
- 3,5-α-Araf
Fig. 10
Truncated structural variants of lipoarabinomannan in mycobacterium leprae and an ethambutol-resistant strain of mycobacterium tuberculosis
Jordi B. Torrelles, Kay-Hooi Khoo, Peter A. Sieling, Robert L. Modlin, Nannan Zhang, Angela M. Marques, Achim Treumann, Christopher D. Rithner, Patrick J. Brennan and Delphi Chatterjee

J. Biol. Chem. published online July 19, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M405180200

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