The Role of the embA and embB Gene Products in the Biosynthesis of the Terminal Hexaarabinofuranosyl Motif of Mycobacterium smegmatis Arabinogalactan*

Vincent E. Escuyer†, Marie-Annick Lety‡, Jordi B. Torrelles§, Kay-Hooi Khoo¶, Jyh-Bing Tang, Christopher D. Rither**, Claude Frehel†, Michael R. McNeil†, Patrick J. Brennan‡, and Delphi Chatterjee†‡‡

From the INSERM U411, 75723 Paris Cedex 15, France, the Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan, and the Departments of **Chemistry and ¶¶Microbiology, Colorado State University, Fort Collins, Colorado 80523

Received for publication, March 13, 2001, and in revised form, October 9, 2001

The emb genes are conserved among different mycobacteria. In Mycobacterium smegmatis and Mycobacterium tuberculosis, they belong to an operon comprising three genes, embC, embA, and embB. The EmbB protein has been proposed to be the target of ethambutol, a drug which is known to inhibit the synthesis of the arabinan portion of the mycobacterial cell wall arabinogalactan (AG). To further define the role of EmbB protein in arabinan biosynthesis, embA, -B, and -C genes were inactivated individually by homologous recombination in M. smegmatis. All three mutants were viable, and among the three, the slowest growing embB− mutant encountered profound morphological changes and exhibited a higher sensitivity to hydrophobic drugs and detergents, presumably due to an increase in cell wall permeability. Furthermore, chemical analyses showed that there was a diminution in the arabinose content of arabinogalactan from the embA− and embB− mutants. Specifically, in comparison with the wild-type strain, the crucial terminal hexaarabinofuranosyl motif, which is a template for mycolylation, was altered in both embA− and embB− mutants. Detailed nuclear magnetic resonance studies coupled with enzyme digestion, chromatography, and mass spectrometry analyses revealed that the disaccharide β-D-Araα-(1→2)-α-D-Araα extension from the 3-position of the 3,5-linked α-D-Araα residue is markedly diminished. As a consequence, a linear terminal β-D-Araα-(1→2)-α-D-Araα-(1→5)-α-D-Araα-(1→5)-α-D-Araα, formed, is a motif which is recognized, nonreducing terminal feature of lipoarabinomannan but not of normal AG. Upon complementation with the embB and embA wild-type genes, the phenotype of the mutants reverted to wild-type, in that normal AG was resynthesized. Our results clearly show that both EmbA and EmbB proteins are involved in the formation of the proper terminal hexaarabinofuranoside motif in AG, thus paving the way for future studies to identify the complete array of arabinosyltransferases involved in the synthesis of mycobacterial cell wall arabinan.

* This work was supported by financial support from INSERM, Paris, France (to V. E.), by NIAID, National Institutes of Health Grants AI-18357 (to P. J. B.), AI-37706 (to M. R. M.), AI-37139 (to D. C.), and Program Project NCDDG-OI PO1AI-46393; and in part by Grant NSC-89-2311-B-001–186 from the National Science Council, Taiwan (to K. H. K.). 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 may be addressed. E-mail: escuyer@insi.org.
‡ To whom correspondence may be addressed. Tel.: 970-491-7495; Fax: 970-491-1815; E-mail: delphi@lamar.colostate.edu.

The cell wall of mycobacteria consists of a highly impermeable layer of 70–90 unique carbon lipids, which are mycolic acids covalently linked to the polysaccharide arabinogalactan (AG). The central role of arabinan in the cell wall proper appears to be in maintaining its integrity by tethering the parallel packed outer mycolic acid lipid barrier perpendicular to the underlying peptidoglycan layer through the flexible glycosyl linkages of AG to form the mycolylarabinogalactan-peptidoglycan (mAGP) complex. The primary structure of the arabinan portion of AG has been elucidated (1–3). Unlike the comparable arabinan structure in lipoarabinomannan (LAM) (4), the arabinan of AG terminates only in a well defined hexaarabinofuranosyl (Ara6) structural motif, [(β-D-Araα-(1→2)-α-D-Araα-(1→5)]2-(3,5)α-D-Araα-(1→5)-α-D-Araα (1), where both the terminal β-D-Araα and the penultimate 2-α-D-Araα serve as the anchoring points for the mycic acids. In Mycobacterium tuberculosis, two such Ara6 units were found to be further assembled into a unique Ara6−mer, the largest structurally defined arabinan unit to date (5). Despite these advances, the precise structural details with respect to the number and further conjugation of the Ara6−mers to the (β-D-Galα-(1→5)-β-D-Galα-(1→6)-) scaffold remain unclear. Nor is it known whether the Ara6−mer architecture also applies to AG of all other Mycobacterium spp., although compositional analysis and NMR fingerprinting have implicated a similar terminal Ara6 motif as a well conserved feature among those examined to date (28), including the AG of Mycobacterium avium and Mycobacterium smegmatis.

One of the major antituberculous drugs, ethambutol (EMB), acts by inhibiting the biogenesis of the cell wall, apparently by direct action on arabinan biosynthesis (6–8). A two-gene locus (embAB) responsible for EMB resistance through target overexpression has been identified in M. avium, and an increase in minimal inhibitory concentration of EMB was demonstrated (9). Furthermore, these two genes were found to confer low levels of resistance to EMB when overexpressed in an otherwise susceptible M. smegmatis host (9). Three contiguous genes encoding putative target(s) for EMB in M. smegmatis and M. tuberculosis were subsequently cloned, sequenced, and characterized (10). Two of these genes were similar to the embA and embB genes described in M. avium, and the third one was termed embC. These genes are likely to be organized as an
operon in the order embC, embA, and embB. It is reasonable to expect that these genes are transcribed as a single polycistronic mRNA from a unique promoter, but this has yet to be shown. However, a previous study (11) strongly implies that at least the embB gene can be expressed from its own individual promoter, the exact location of which still remains to be characterized. The embCAB gene cluster was initially identified in an EMB-resistant strain of M. smegmatis (10, 11) and was subsequently characterized in both M. tuberculosis (10) and Mycobacterium leprae (12). The major virulence determinant of EMB is the EmbB protein, and thus, embB has been proposed as the main target of EMB. A large number of clinical isolates resistant to EMB have mutations in embB (10,13,14), and the EmbB protein has been proposed as the main target of EMB.

The Emb proteins are predicted to be integral membrane proteins with 11–13 transmembrane domains and a large carboxyl-terminal globular region of external location (10, 14). Furthermore, M. smegmatis strains overexpressing embB, embA, and embC had increased arabinosyltransferase activities implying that embA, -B, and -C could encode arabinosyltransferases (9). However, heterologous expression of functional Emb proteins for direct biochemical assays or mechanistic studies has not been successful to allow unequivocal attribution of enzymatic activities. As an alternative, we took the reverse genetic approach to inactivate embA, -B, and -C genes individually and report here the resulting phenotypes and structural alterations associated with the arabinan component of the cell wall AG. Interruption of embA and -B led to a marked decrease in the total arabinose content. In particular, the terminal disaccharide β-D-Ara(1→2)-α-D-Ara unit normally found on the 3-arm of the characteristic Ara₆ is largely missing in both mutants. This is the first report of AG-deficient mutants of mycobacteria with specific lesion in the terminal elaboration of arabinan that could be restored by complementing with the respective functional genes.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Culture Conditions—Escherichia coli strain DH5α (15) was grown at 37°C on Luria-Bertani medium supplemented either with ampicillin (100 μg/ml), kanamycin (50 μg/ml), or gentamicin (15 μg/ml). M. smegmatis mc²155 (16) was grown at 37°C in Middlebrook liquid, either 7H9 (liquid) or 7H10 (solid), supplemented with oleic acid-albumin-dextrose-catalase enrichment (Difco Laboratories), glycerol (0.2%), and Tween 80 (0.05%). Appropriate antibiotics were added when necessary at a final concentration of 20 μg/ml except for hygromycin (50 μg/ml). Transformation of E. coli was carried out according to standard procedure (17). Electromediation of M. smegmatis mc²155 was performed as described by Snapper et al. (16).

Plasmids and DNA Manipulations—pBluescript II KS+ was purchased from Stratagene. pOLYG and pPR23 were kindly supplied by Dr. O. Gaora and Dr. B. Gicquel, respectively (18, 19). All the restriction and nucleic acid-modifying enzymes were purchased from New England Biolabs. Taq DNA polymerase was purchased from Promega. Plasmid DNAs were purified on Qiagen columns according to the recommendations of the manufacturer (Qiagen, GmbH).

Insertional Mutagenesis—DNA fragments carrying internal sequences from embB and embC from M. smegmatis chromosome were amplified by PCR using the following primers: for embB, positions 1128–1257 and positions 2225–2357, according to the coding sequence and for embC, positions 44–73 and positions 2083–2112 (GenBank mascot accession number U68480). Amplified DNA fragments were cloned into pBluescript II and a cassette conferring resistance to kanamycin (aphA-3) (20) was inserted at the MscI site for embA and the NcoI site for embC. Resulting DNA fragments were finally subcloned into pPR23 (19), a suicide shuttle vector carrying the counterselection marker sacB to give, respectively, pMAL-A and pMAL-C. A 765-bp fragment was deleted and replaced with a promoterless cassette conferring resistance to kanamycin (aphA-3) ΔP. The resulting fragment was then subcloned into pPR23 to give pMAL-B (Fig. 1A). PMAL-A, -B, and -C were introduced into M. smegmatis by electroporation, and transfectants were selected at 30°C on Middlebrook 7H10 medium containing 5 μg/ml gentamycin. Transfectants were grown in 7H9 liquid medium without selection and then plated at 42°C on 7H10 supplemented with kanamycin (20 μg/ml) and sucrose (10%). Kanamycin-resistant and gentamicin-sensitive clones were analyzed by both PCR and Southern hybridization with appropriate probes. Southern blots confirm the double crossover events and replacement of wild-type sequences.

Electron Microscopy—For negative staining, samples of exponentially grown and stationary phase cells were fixed at room temperature for 10 min with formaldehyde to a final concentration of 1% and centrifuged, and the pellet was suspended in 0.5% phosphotungstic acid pH 7 (21). Formvar-carbon-coated grids were covered with one drop of bacterial suspension, air dried, and directly examined under the electron microscope (Jeol JEM-100CXII). For thin sectioning, samples were processed exactly as described previously (22). Intracellular polysaccharides were stained by the method of Thiery (23) as subsequently modified (22).

Drug Resistance Assays—Minimal inhibitory concentrations were determined by plating cells on solid Middlebrook 7H10 medium supplemented with various concentrations of the tested drug. The minimal inhibitory concentration was defined as the first concentration of drug that inhibited 100% of growth after 5 days of incubation (9).

Preparations of AG—Cells (2 g of wet weight) were delipidated with hexane/ether (2:1) chloroform/methanol extraction as described previously (24). The dried residual biomass was washed with phosphate-buffered saline and disrupted mechanically with the Soniprep 150 (Sanyo Gallenkamp PLC). The pulse time was 60 s disrupting and 90 s relaxing repeated six times. The suspension after cell breakage was extracted three times with 50% ethanol, each time centrifuging at 27,000 × g. The supernatants were concentrated and partitioned between water and phenol to remove proteins (25). The aqueous layer containing the majority of cellular LAM, lipomannan, and phosphatidylinositol mannosides was freeze-dried, reconstituted in the same volume of water, and used for SDS-PAGE. The core wall mycolylarabinogalactan-peptidoglycan complex was obtained by extraction of the residual pellet with phosphate-buffered saline containing 2% SDS followed by extensive washing of the insoluble residue in the same buffer. The SDS was removed by sequential washing with phosphate-buffered saline, water, and finally acetone. The insoluble residue (mAGP complex) was then solubilized by alkaline treatment with 1 M sodium hydroxide (1) for 16 h at 80°C followed by neutralization and dialysis. The freeze-dried material was applied to a 26-μm × 1-m column of Sephacryl S-300 (Amersham Biociences, Inc.) with Milli-Q water as the elution buffer. Fractions were pooled on the basis of neutral sugar content and dried. Pools were analyzed by converting aliquots into alditol acetates after hydrolysis.

GC and GC/MS Analysis of Glycocalyx Composition—Alditol acetates and partially methylated alditol acetates for GC/MS analysis were prepared according to the procedure described previously (22). Briefly, AG samples or permethylated derivatives were hydrolyzed with 2 M HCl at 100°C, reduced with 10 mg/ml sodium borodeuteride (NaBD₄), 2 mM NH₄OH at room temperature for 2 h, and then acetylated with acetic anhydride at 100°C for 1 h. GC was carried out on an SP 2380 (30 μm × 0.25-mm internal diameter, Supelco) column at an initial temperature of 50°C held for 1 min. The temperature was then raised to 170°C at 30°C/min before increasing to 260°C at 5°C/min. GC/MS was carried out using a ThermoQuest Trace Gas Chromatograph 2000 (ThermoQuest, Austin, TX) connected to a GCQ/Polaris MS mass spectrometer (ThermoQuest). The partially methylated alditol acetates were dissolved in chloroform prior to injection on a DB-5 column (10 m × 0.18-mm internal diameter, J&W Scientific, Folsom, CA) at an initial temperature of 50°C held for 1 min. The temperature was raised to 150°C at 30°C/min before increasing to 260°C at 5°C/min.

Digestion with Endoarabinanase—Selective growth of a soil microorganism, Cellulomonas gelida, from which the endoarabinanase was purified, has been described previously (26). To ensure a consistent result, 10 μg of each AG sample was treated with 5 μl of the enzyme preparation for 8 h at 37°C. Aliquots from the digestion product mixtures that contained the released oligoarabinosides and oligogalactosides were analyzed by high performance liquid chromatography (HPLC). Analytical HPLC was performed on a Dionex LC system fitted with a Dionex Carboxpac PA-1 column, and the oligogalactosides were detected with a pulse-amperometric detector (PAD-II).

Electrospray Mass Spectrometry Analyses—Samples were permethylated using the NaOH/dimethyl sulfoxide slurry method or were perdeuterioacetylated with the pyridine-d₆-acetic anhydride method as described previously (22).
Alteration of Terminal Ara<sub>6</sub> in AG

scribed by Dell et al. (27). Electrospray MS analyses were performed on an Autospec orthogonal acceleration-time-of-flight mass spectrometer (Micromass, Manchester, United Kingdom) fitted with an electrospray 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 electrospray source by a syringe pump.

**RESULTS**

Gene Exchange Mutagenesis of the emb Genes of M. smegmatis—The three mutants emb<sup>A</sup>, emb<sup>B</sup>, and emb<sup>C</sup> were generated by homologous recombination using the thermosensitive vector pPR23 containing emb<sup>A</sup>, emb<sup>B</sup>, or emb<sup>C</sup> intragenic fragments disrupted by the aphA-3 cassette conferring resistance to kanamycin (20) (Fig. 1A). To confirm the allelic exchange, chromosomal DNA was prepared from the mutants, and PCR amplifications were performed using primers that yield amplification products of the expected size were obtained. Amplifications were performed using primers that yield amplification products only after the double recombination event occurred. Amplification products of the expected size were obtained with the three emb<sup>+</sup> mutants (Fig. 1B); no amplification was obtained with the wild-type (WT) strain. Southern blot analysis was also performed on the chromosomal DNA from the mutants, which confirmed the results obtained with PCR. Therefore, both PCR and Southern blot experiments demonstrated that the three emb<sup>−</sup> mutants were the result of allelic exchange involving replacement of the emb genes with the disrupted emb constructs and loss of the vector.

Inactivation of emb Genes Induces Morphological Alterations in M. smegmatis—WT, emb<sup>A</sup>, emb<sup>B</sup>, and emb<sup>C</sup> strains were examined during the exponential phase of growth for morphological changes. Thin sections of WT showed typical morphology with long rods, regular septation, and bacterial size of 2.5 μm (Fig. 2a). In contrast, the emb<sup>B</sup> mutant showed drastically altered morphology with size shortening (2.8 μm), swelling, and distortion (Fig. 2c). Accumulation of translucent cytoplasmic inclusions could also be observed (Fig. 2c). The emb<sup>B</sup> mutant was also altered in its morphology with size shortening (2.9 μm), slight distortion, and swelling but to a lesser extent than with the emb<sup>B</sup> mutant (Fig. 2b). The emb<sup>C</sup> mutant exhibited a shorter size (4.9 μm) but with a nonaltered shape (Fig. 2d). In this study, it was also found that interruption of the emb genes resulted in a severe alteration in the physiological behavior of the bacterium in several respects: for instance, complete loss of acid fastness for both emb<sup>B</sup> and emb<sup>A</sup> strains and only partial acid fastness in the case of the emb<sup>C</sup> mutant. In addition, susceptibility of these mutants to the hydrophobic antibiotics rifampicin and novobiocin was enhanced (minimal inhibitory concentrations of 125 and 60 μg/ml for the WT in comparison to minimal inhibitory concentrations of 25 and 4 μg/ml for the emb<sup>A</sup>, 1.5 and 0.25 μg/ml for the emb<sup>B</sup> and 40 and 25 μg/ml for the emb<sup>C</sup>), indicating altered permeability of the mutant strains. These results pointed to the probability that the emb<sup>−</sup> mutants had altered cell wall.

Selective Effects of emb Genes Inactivation on the Arabinin Synthesis of AG—The peculiar morphological variations observed among the emb<sup>−</sup> mutants and the evidence that the emb genes encode for key proteins in AG biosynthesis (9) led us to analyze the structure of the arabinin of AG from these mutants. Cell wall core (mAGP) was prepared from WT and the embA<sup>−</sup>, B<sup>−</sup>, and C<sup>−</sup> as described previously (24), and the ratio of Ara to Gal in mAGP was determined by gas chromatography of the alditol acetates prepared after hydrolysis (28). The glycosyl composition was calculated based on a single rhamnosyl (Rha) residue per AG chain (Table I) (2). The AG preparation from the WT mAGP was found to contain ~100 glycosyl residues with a relative molar ratio of Rha:Gal:Ara of 1.29:64 and an Ara:Gal ratio of 2.2:1 in accord with previous data (1). All three emb<sup>−</sup> mutants yielded AG with significant reduction in Ara content concomitant with an increase in the amount of Gal (Table I), although the effect was less dramatic with emb<sup>C</sup>. In accord with the extent of morphological alteration, the AG from the embB<sup>−</sup> mutant also gave the lowest Ara:Gal ratio (0.88:1) followed by embA<sup>−</sup> (1.01:1) and embC<sup>−</sup> (1.4:1). When comple-
from LAM also be synthesized. Further analysis of the LAM preparation polycistronic mRNA encoding the three Emb proteins could addition, the existence of an individual promoter for embA and treated as described under “Experimental Procedures.” The bar represents 0.5 μm.

**FIG. 2.** Morphological aspect of *M. smegmatis* WT and the emb mutants. Thin sections of *M. smegmatis* WT (a), embA− mutant (b), embB− mutant (c), and embC− mutant (d) were analyzed by electron microscopy. Cells were collected during the exponential phase of growth and treated as described under “Experimental Procedures.” The bar represents 0.5 μm.

| M. smegmatis mc2155 | Ara | Gal | Rha | Ara/Gal ratio |
|---------------------|-----|-----|-----|---------------|
| **WT**              | 64  | 29  | 1   | 2.21:1        |
| embA−               | 40  | 39  | 1   | 1.01:1        |
| embB−               | 40  | 45  | 1   | 0.88:1        |
| embC−               | 51  | 36  | 1   | 1.41:1        |

**TABLE I**
Neutral sugar composition of the AG component of insoluble mAGP from *M. smegmatis* WT and embA−, embB−, and embC−.

The neutral sugar composition (molar ratios) is an average from four different sets of experiments calculated based on a single rhamnosyl residue per AG chain. Varying amounts of Glc were also detected in the alditol acetates of all AG samples.

mented in trans with respective WT emb genes, the embA− and embB− mutants were able to synthesize AG with a glycosyl composition identical to the WT AG. This total synthesis of normal AG with a WT copy of the gene corroborates that the defect in AG observed in the embA− and embB− mutants was a direct consequence of the interruption of these genes. Synthesis of AG with embC complementation, however, was less complete, suggesting that the effect of embC inactivation on AG expression could be partly due to a polar effect on the expression of embA and/or embB located directly downstream. However, this remains to be directly demonstrated, and the exact cause of the AG defect in the embC− mutant is under further investigation. Based on the complementation experiments, the presence of the individual embB promoter is confirmed, and in addition, the existence of an individual promoter for embA is also strongly suggested. Thus, embA and embB genes can be expressed from their own individual promoters, but the genetic organization of the emb genes also indicates that, in addition, a polycistronic mRNA encoding the three Emb proteins could also be synthesized. Further analysis of the LAM preparation from embA− and embB− revealed no difference in composition from that of LAM from the WT, indicating that the embAB gene products are specifically involved in the biosynthesis of AG (data not shown). In contrast, full-size LAM could not be obtained from the embC− mutant, suggesting that the embC gene product is involved in the synthesis of LAM. Further detailed analysis of the latter aspects of embC gene inactivation will be reported elsewhere.2

Comparison of the Structures of AG from WT and embB− using NMR—To provide further insights into the chemical basis of alteration in the arabinan content, the base-solubilized AGs were subjected to NMR analyses. Since compositional analyses have indicated the presence of substantial amount of Glc in the samples, the crude preparation of AGs were applied to a Sephacryl S-300 size fractionation column to remove the majority of the contaminating glucan, which eluted at the void volume. A quantitative 1H NMR analysis of the AGs from WT and embB− clearly showed, as expected, that all the anomeric resonances corresponding to Araf residues were reduced in intensity with respect to the anomeric Galf resonances. This observation was corroborated by a 13C DEPT experiment (Fig. 3) for which the methods used to assign the C-1 resonances of terminal β-Araf, 2-linked α-Araf, and the C-2 of 2-linked α-Araf have been described previously (28). Further confirmation was sought using a two-dimensional 1H 13C HSQC experiment to obtain well resolved spectra (Fig. 4) from which resonances in highly overlapping anomeric regions of one-dimensional 1H NMR could be assigned. Thus, for the WT AG, C-1 resonance at δ 109 ppm correlated with anomeric protons at δ 5.07 ppm and was assigned to the 6-linked β-Galf. Two overlapping spin systems centered at δ 108.6, 108.7 ppm correlating with protons at ~δ 5.16 ppm were attributed to the 5-linked α-Araf and

2 D. Chatterjee and V. E. Escuyer, manuscript in preparation.
the 3,5-linked α-Ara₃. The C-1 resonance of 5,6-linked β-Gal₃ at δ 108.2 ppm (δH at δ 5.21 ppm) was resolved from that of 5-linked β-Gal₃ at δ 108.1 ppm, which correlated with protons at δ 5.28 ppm. The two sets of well separated spin systems were easily identified as the carbon-protons related to the C-1 resonances of 2-linked α-Ara₃, which otherwise overlapped in the one dimensional 1H NMR spectrum but separated into two singlets in the 13C DEPT experiment (Fig. 3A). The C-1 resonances of these carbons were at δ 106.8 (δH at δ 5.22 ppm) and δ 106.6 with δH at δ 5.29 ppm. The farthest upfield signals were the two sets of spin systems belonging to the terminal β-Ara₃ resonating at δ 101.8 ppm (δH at δ 5.19 ppm) and δ 100.8 ppm, respectively, correlating with protons centered at δ 5.44 ppm. The two clear singlets in the 13C DEPT spectrum, which correspond to C-2 of 2-linked α-Ara₃ (Fig. 3A), appeared to be completely overlapping in the HSQC spectrum. However, it could be resolved into two broad peaks in the slice spectrum of this region (see Fig. 4A, inset), indicative of a similar terminal arabinan structural motif in the WT AG as has been reported for the AG of M. tuberculosis (1).

There were noteworthy differences in the NMR spectra of the AG of embB gene (Fig. 3B) in comparison with those of WT. From the 13C DEPT spectra (Fig. 3), it is clear that one of the signals from each of the three resonance doublets attributed to C-1 of 2-linked α-Ara₃, C-2 of 2-linked α-Ara₃, and C-1 of terminal β-Ara₃ are missing in the spectra of the AG of embB mutant (Fig. 3B). Likewise, in the 1H 13C HSQC spectrum (Fig. 4B), the cross-peaks corresponding to terminal β-Ara₃ and 2-linked Ara₃, extending from the 3-arm of 3,5-linked Ara₃ in the WT are reduced to trace amounts. The slice spectrum in the HSQC (shown as an inset in Fig. 4B) further showed that the two broad singlets corresponding to C-2 of 2-linked Ara₃ have collapsed into one singlet. A reduction in the volume/Intensity of the resonance corresponding to 3,5-α-Ara₃ from spin systems at δ 108.7 ppm/δ 5.14 ppm was also evident. Together these data are indicative of the preferential albeit incomplete loss of a β-d-Ara₃-(1→2)-α-d-Ara₃ terminal disaccharide attached to the 3-OH of the otherwise branched 3,5-α-Ara₃ residue in the backbone of overall reduction in the arabinan content relative to the galactan.

Interestingly the relative peak volume of the anomic resonances as determined from the 1H 13C HSQC experiment (Table II) indicated that the amount of 5,6-linked Gal₃ relative to 5-linked or 6-linked Gal₃ was also reduced in the embB mutant AG. Since branching in the (5-β-Gal₃,1→6-β-Gal₃,5), galactan chain has been attributed to arabinan attachment, the lower degree of branching is consistent with reduction in the attached arabinan chains and/or extended chain length of galactan. In comparison, the AG from the embB mutant exhibited similar lesion in the arabinan with drastic reduction in the 2-linked α-Ara₃→3 but maintained the same degree of branching in the galactan as in WT AG. Thus, both embB and embB gene products may act together and specifically contribute to the proper synthesis of the terminal arabinan motif of AG. The embB gene may additionally be involved in the assembly of properly synthesized arabinan onto the galactan scaffold.

**HPAEC and MS Mapping of Endoarabinofuranase Digestion Products.—** It was reasoned that structural alteration within the terminal elaboration of the arabinan will primarily affect the terminal hexaarabinofuranosyl (Ara₉) structural motif, which would otherwise be released intact from AG by the crude endoarabinanases/galactanases preparation from C. gelida (26).

Specifically, if a terminal β-d-Ara₃-(1→2)-α-d-Ara₃ unit was indeed missing as suggested by NMR analyses, a terminal tetraarabinofuranoside (Ara₄) will be produced instead of Ara₆. The AG from the embB mutant AG could be resolved into two broad peaks in the slice spectrum of embB AG, the farthest upfield signals were δ 108.1 ppm, which correlated with protons centered at δ 5.44 ppm. The two sets of well separated spin systems were easily identified as the carbon-protons related to the C-1 resonances of 2-linked α-Ara₃, which otherwise overlapped in the one dimensional 1H NMR spectrum but separated into two singlets in the 13C DEPT experiment (Fig. 3A). The C-1 resonances of these carbons were at δ 106.8 (δH at δ 5.22 ppm) and δ 106.6 with δH at δ 5.29 ppm. The farthest upfield signals were the two sets of spin systems belonging to the terminal β-Ara₃ resonating at δ 101.8 ppm (δH at δ 5.19 ppm) and δ 100.8 ppm, respectively, correlating with protons centered at δ 5.44 ppm. The two clear singlets in the 13C DEPT spectrum, which correspond to C-2 of 2-linked α-Ara₃ (Fig. 3A), appeared to be completely overlapping in the HSQC spectrum. However, it could be resolved into two broad peaks in the slice spectrum of this region (see Fig. 4A, inset), indicative of a similar terminal arabinan structural motif in the WT AG as has been reported for the AG of M. tuberculosis (1).

There were noteworthy differences in the NMR spectra of the AG of embB strain in comparison with those of WT. From the 13C DEPT spectra (Fig. 3), it is clear that one of the signals from each of the three resonance doublets attributed to C-1 of 2-linked α-Ara₃, C-2 of 2-linked α-Ara₃, and C-1 of terminal β-Ara₃ are missing in the spectra of the AG of embB mutant (Fig. 3B). Likewise, in the 1H 13C HSQC spectrum (Fig. 4B), the cross-peaks corresponding to terminal β-Ara₃ and 2-linked Ara₃, extending from the 3-arm of 3,5-linked Ara₃ in the WT are reduced to trace amounts. The slice spectrum in the HSQC (shown as an inset in Fig. 4B) further showed that the two broad singlets corresponding to C-2 of 2-linked Ara₃ have collapsed into one singlet. A reduction in the volume/Intensity of the resonance corresponding to 3,5-α-Ara₃ from spin systems at δ 108.7 ppm/δ 5.14 ppm was also evident. Together these data are indicative of the preferential albeit incomplete loss of a β-d-Ara₃-(1→2)-α-d-Ara₃ terminal disaccharide attached to the
FIG. 4. Two-dimensional $^1$H $^{13}$C HSQC spectra of base-solubilized AGs of *M. smegmatis* WT (A) and *embB* mutant (B). Only the expanded anomeric regions are shown. The peaks of C-2 of 2-$\alpha$-Ara$_6$ are also included in the plots.
Alteration of Terminal Ara\(_6\) in AG

HSQC experiments on different AG samples were performed on separate days, and therefore, there were some variations (0.1–0.2 ppm) on the ppm values for \(^{13}\)C and \(^{1}H\). The chemical shifts reported for the three mutants have been corrected accordingly.

| Residues | \(^{13}\)C | \(^{1}H\) | Relative Peak Volume* |
|----------|-----------|-----------|---------------------|
| C-2 of 2-α-Ara\(_b\) | 88.0 | 4.25 | 1.00  |
| t-β-Ara\(_b\)[1] | 100.8 | 5.44 | 0.82  |
| t-β-Ara\(_b\)[2] | 101.8 | 5.19 | 1.00  |
| 2-α-Ara\(_b\) | 108.6 | 5.29 | 0.47 |
| 2-β-Ara\(_b\) | 108.6 | 5.22 | 0.53  |
| 3-α-Ara\(_b\) | 108.6 | 5.16 | 0.99  |
| 5-α-Ara\(_b\) | 108.7 | 5.14 | 2.62 |
| 5-β-Gal\(_b\) | 108.1 | 5.28 | 0.70 |
| 5-6-β-Gal\(_b\) | 108.2 | 5.21 | 0.45 |
| 6-β-Gal\(_b\) | 109.0 | 5.07 | 0.74 |

* The relative peak volumes are represented in two ways. Values in the first column for each of the AG samples were normalized to terminal β-Ara\(_b\) [2]. For the second column, the values were normalized to 6-β-Gal\(_b\).

Two sets of terminal and 2-linked α-Ara\(_b\) resonances were detected, denoted respectively as 1 and 2. The chemical shifts of the first set correspond to β-Ara\(_b\)-1-2-α-Ara\(_b\)-1-3 on the 3-arm, whereas the second were assigned to those on the 5-arm of the branched 3,5-α-Ara\(_b\) residue. Only about half of the terminal β-Ara\(_b\) detected was accounted for by direct conjugation to the 2-linked Ara\(_b\) in the WT AG based on this method of quantification.

### DISCUSSION

AG is the major constituent of the mycobacterial cell wall. Although the primary structure of this molecule has been established (30), very little is known about its biosynthetic pathway. One of the strategies to identify the genes involved in AG synthesis is based on the hypothesis that some of these genes will be similar to the genes involved in polysaccharide biosynthetic pathways in other bacteria. This approach proved to be highly successful for identifying genes responsible for the early steps of AG biosynthesis (31). However, due to the fact that β-Ara\(_b\) residues are rarely found in nature and the search for homologous enzymes is problematical, no arabinosyltransferases involved in the polymerization of mycobacterial AG arabinan have thus far been identified. An alternative approach to finding enzymes involved in AG biosynthesis is to use the antimycobacterial drug EMB, which inhibits the biosynthesis of the arabinan moiety of both AG and LAM (8, 24). It had been previously demonstrated that β-d-arabinofuranosyl-P-decaprenol is the precursor of the Ara\(_b\) residues of arabinan and accumulates when mycobacterial cells were treated with EMB (32). Further, the transfer of Ara\(_b\) from β-d-arabinofuranosyl-P-decaprenol to arabinan was inhibited by EMB, although only by about 50% in the context of a crude cell-free assay system (33). Based on this evidence, the hypothesis arose that the target(s) for EMB action were arabinosyltransferases. Previous studies had also implicated the Emb proteins, directly or indirectly, in arabinan biosynthesis (9, 10). It was demonstrated, by chemical analysis and radiolabeling experiments, that M. smegmatis cells transformed with emb\(_A\) on a multicopy plasmid and treated with EMB produce normal AG, whereas similar treated cells containing only the cloning vector showed a marked diminution in arabinan content of AG (9, 24). Taken together, these data suggested that Emb\(_A\) proteins either are themselves arabinosyltransferases or are implicated in arabinan synthesis, for example, in the export of the synthesized oligosaccharide precursors.

In this study, we found that inactivation of the emb\(_A\) and emb\(_B\) genes resulted in a severe alteration in the physiological behavior of the bacterium. Further, we showed that inactivation of these genes resulted in a marked reduction of the total arabinose content of AG compensated by a partial increase in galactose content. More specifically, upon detailed structural analyses by HPAEC, MS, and NMR, the amount of the nonreducing terminal disaccharide β-d-Ara\(_b\)-1-(1-2)-α-d-Ara\(_b\) that could be derived from AG was shown to be dramatically diminished in both emb\(_A\)- and emb\(_B\)-mutants. Since a substantial amount of the otherwise predominant Ara\(_b\) branched structural unit was found to be converted to linear Ara\(_b\), it can be further concluded that a primary lesion in the emb\(_A\)- and emb\(_B\)- mutants resides in the 3-arm branching off the α-5-arabinan chain proximal to the nonreducing end and the attachment site for the mycolates. It has been established that only two-thirds of the mycolylation sites of AG are substituted with mycolates either at the terminal β-d-Ara\(_b\) or the penultimate 2-α-d-Ara\(_b\) (3). The β-d-Ara\(_b\)-1-(1-2)-α-d-Ara\(_b\)-deficient polymer formed as a consequence of emb\(_A\) or emb\(_B\) gene inactivation is expected to contain fewer mycolates resulting in an increase of the cell wall permeability. The amounts of mycolic acids were indeed reduced in both emb\(_A\)- and emb\(_B\)-mutants (data not shown).
and there was an increased permeability to hydrophobic antibiotics. However, the ability of these mutants to grow, albeit more slowly than the WT, demonstrates that the emb genes are nonessential. As for the increase in galactose content in AG from the emb/H11002 mutants, we hypothesize that it is due to an abnormal synthesis of the galactan when restraints imposed by arabinose addition are lessened.

The total number of arabinosyltransferases that are required for arabinan biosynthesis is a matter of speculation, depending on how the arabinan chains are assembled. Based on the structural information available, one can predict that as many as five to six enzymes are needed for the formation of the different linkages, and these could very well vary for LAM and AG. The presence of the common terminal hexaarabinofuranosyl motif in AG and LAM suggests that similar sets of arabinosyltransferases are involved in their assembly, but they must be sufficiently distinct to allow divergent processing. For instance, the embA− and embB− mutants have no discernible effect on arabinosylation of LAM even though they both dramatically alter the arabinosylation of AG. In contrast, preliminary data indicate that in the embC− mutant, the arabinosylation of LAM is abolished (data not shown). Therefore, although very similar in sequence, the Emb proteins are able to discriminate between the arabinan synthesis of AG and LAM. However, the structural basis for this specificity remains to be elucidated.

Based on our results, it is now clear that the Emb proteins are involved in the proper branching of the terminal Ara6 motif. The fact that the inactivation of embA or embB resulted in a very similar defect strongly suggests that EmbA and EmbB proteins share a very similar if not identical function. We propose that these proteins could both act as α-(1→3) arabinosyltransferases. However, several questions remain unanswered. For example, β-D-arabinofuranosyl-P-decaprenol is the only precursor/donor identified in arabinogalactan/arabino-
mannan synthesis (32), and because lipid-linked sugar donors are conventionally transported across the membranes (34), one can postulate that some intermediates need to be translocated from the inner leaf to the outer side of the plasma membrane in order for complex arabinosylation to occur. The 11–13 transmembrane domains present in the three Emb proteins (10, 14) suggest that they may also operate, at least in part, as proteins involved in transporting LAM and AG arabino precursors across the plasma membrane. In this case, the Emb proteins would work in close relation with the true arabinosyltransferases so far unidentified or might be bifunctional proteins with both transferase and transport activities.

Until recently the Emb proteins were known exclusively in mycobacteria. However, progress in the sequencing of the genome of Corynebacterium diphtheriae revealed the presence of a gene having 40% identity with the mycobacterial emb genes (Sanger Center, Cambridge, UK). C. diphtheriae cell wall structure presents many similarities with that of mycobacteria, including the presence of AG (35, 36). However, the glycosyl linkage composition of AG from C. diphtheriae differs from published mycobacterial AG composition by the absence of 3,5-linked Ara (37). This observation strongly suggests that the Emb proteins are members of a family of proteins involved in the formation of various linkages in the arabinan molecule. The fact that they are not confined to mycobacteria offer the possibility of comparative studies among the bacteria sharing similar cell wall structure (38, 39). The fine specificities and hence functional roles of each of the homologous Emb proteins are likely to be encoded by differences in a few key amino acid residues within the catalytic or active sites. This and additional emb gene sequence information derived from various EM-resistant mutants will aid in the final delineation of the precise role of the Emb proteins in the arabinan biosynthesis as well as the mechanism of EM drug action and resistance.

Acknowledgments—We acknowledge Drs. Julia Inamine and Dean Crick for many helpful discussions and Dr. Yi Xin for providing the endoarabinanase arabinan degrading enzymes.

REFERENCES

1. Daffe, M., Brennann, P. J., and McNeill, M. (1990) J. Biol. Chem. 265, 6726–6734

2. McNeill, M., Daffe, M., and Brennann, P. J. (1990) J. Biol. Chem. 265, 16200–16206

3. McNeill, M., Daffe, M., and Brennann P. J. (1991) J. Biol. Chem. 266, 1217–1223

4. Chatterjee, D., and Khoo, K.-H. (1998) Glycobiology 8, 113–120

5. Besra, G. S., Khoo, K. H., McNeil, M. R., Dell, A., Morris, H. R., and Brennann, P. J. (1995) Biochemistry 34, 4257–4266

6. Takayama, K., and Kilburn, J. O. (1989) Antimicrob. Agents Chemother. 33, 493–499

7. Deng, L., Mikusová, K., Robuck, K. G., Schermer, M., Brennann, P. J., and McNeill, M. R. (1995) Antimicrob. Agents Chemother. 39, 694–701

8. Mikusová, K., Slayden, R. A., Besra, G. S., and Brennann, P. J. (1995) Antimicrob. Agents Chemother. 39, 2484–2489

9. Besra, G. S., Brennann, P. J., and Inamine, J. M. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 1191–1194

10. Telenti, A., Philipp, W. J., Sreevatans, S., Bernasoni, C., Stockhauers, K. E., Brennann, P. J., and Jacobs, W. R. Jr. (1997) J. Biol. Chem. 272, 269–275

11. Lety, M. A., Nair, S., Berche, P., and Escuyer, V. (1997) Antimicrob. Agents Chemother. 41, 2629–2633

12. Cole, S. T., Gogarten, K., et al. (2001) Nature 409, 1007–1011

13. Sreevatans, S., Stockhauers, K. E., Pan, X., Kreiswirth, B. N., Moghazeh, S. L., Nair, S., Frehel, C., and Leduc, M. (1987) J. Biol. Chem. 262, 291–297

14. Ramaswamy, S. V., Amin, A. G., Goksel, S., Stager, C. E., Dou, S. J., El Sahly, H., Moghazeh, S. L., Nair, S., Frehel, C., and Leduc, M. (1987) J. Biol. Chem. 262, 363–368

15. Hanahan, D. (1983) J. Mol. Biol. 166, 557–580

16. Snapper, S. B., Melton, R. E., Mustafa, S., Kiesler, T., and Jacobs, W. R., Jr. (1990) Mol. Microbiol. 4, 1911–1919

17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., pp. 1.82–1.84, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

18. Garbe, T. R., Barathi, J., Barnini, S., Zhang, Y., Ahou-Zeid, C., Tang, D., Mukherjee, H., and Young, D. B. (1994) Microbiology 140, 133–138

19. Pelicic, V., Jackson, M., Meyrat, J. M., Jacobs, W. R., Jr., Gicquel, B., and Guilhot, C. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 10955–10960

20. Tran-Cot, P., Goubau, G., Lambert, T., and Courvalin, P. (1985) EMBO J. 4, 3583–3587

21. Nair, S., Frehel, C., Nguyen, L., Escuyer, V., and Berche, P. (1999) Mol. Microbiol. 31, 1186–1196

22. Frehel, C. and Leduc, M. (1987) J. Bacteriol. 169, 210–217

23. Thierry, J. P. (1976) J. Microsc. 6, 987–1018

24. Khoo, K.-H., Douglas, E., Azadi, P., Inamine, J. M., Besra, G. S., Mikusová, K., Brennann, P. J., and Chatterjee, D. (1996) J. Biol. Chem. 271, 28682–28689

25. Chatterjee, D., Roberts, A. D., Lowell, K., Brennann, P. J., and Orme, I. M. (1992) Infect. Immun. 60, 1249–1253

26. McNeill, M., Robuck, K. G., Harker, M., and Brennann, P. J. (1994) Glycobiology 4, 165–173

27. Dell, A., Reason, A. J., Khoo, K.-H., Panico, M., McDowell, R. A., and Morris, H. R. (1994) Methods Enzymol. 230, 108–132

28. Daffe, M., McNeill, M., and Brennann, P. J. (1995) Carbohydr. Res. 249, 383–398

29. Khoo, K.-H., Tang, J. B., and Chatterjee, D. (2001) J. Biol. Chem. 276, 3863–3871

30. Daffe, M. and Draper, P. (1998) Adv. Microb. Physiol. 39, 131–203

31. McNeill, M. R. (1999) in Genetics of Bacterial Polysaccharides (Goldberg, J. B., ed) pp. 207–223, CRC Press, Boca Raton, FL

32. Wolucka, B. A., McNeill, M. R., de Hoffmann, E., Chojnacki, T., and Brennann, P. J. (1994) J. Biol. Chem. 269, 23328–23335

33. Lee, R. E., Brennann, P. J., and Besra, B. S. (1997) Glycobiology 7, 1121–1128

34. Rocchetta, H. L., Burrows, L. L., and Lam, J. S. (1999) Microbiol. Mol. Biol. Rev. 63, 521–553

35. Marienfeld, S., Uhlmann, E. M., Schmid, R., Kramer, R., and Burkovski, A. (1997) Antonie Leeuwenhoek 72, 291–297

36. Sutcliffe, I. C. (1997) Vet. Microbiol. 56, 287–289

37. Puech, V., Chami, M., Lemassu, A., Laneelle, M. A., Schiffer, B., Gounon, P., Bayan, N., Benz, R., and Daffe, M. (2001) Microbiology 147, 1365–1382
The Role of the *embA* and *embB* Gene Products in the Biosynthesis of the Terminal Hexaarabinofuranosyl Motif of *Mycobacterium smegmatis* Arabinogalactan

Vincent E. Escuyer, Marie-Annick Lety, Jordi B. Torrelles, Kay-Hooi Khoo, Jyh-Bing Tang, Christopher D. Rithner, Claude Frehel, Michael R. McNeil, Patrick J. Brennan and Delphi Chatterjee

*J. Biol. Chem.* 2001, 276:48854-48862.  
doi: 10.1074/jbc.M102272200 originally published online October 24, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M102272200

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

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

This article cites 35 references, 14 of which can be accessed free at http://www.jbc.org/content/276/52/48854.full.html#ref-list-1