The Carboxy Terminus of EmbC from Mycobacterium smegmatis Mediates Chain Length Extension of the Arabinan in Lipoarabinomannan

Libin Shi1,5, Stefan Berg1, Arwen Lee5, John S. Spencer1,1, Jian Zhang1, Varalakshmi Vissa1,1, Michael R. McNeil1, Kay-Hooi Khoo1,5, and Delphi Chatterjee1,3

From the 1Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, Colorado 80523, 2Department of Biochemistry and Molecular Biology, Dalian Medical University, Dalian, Liaoning, China, 116027, 3Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan , and 4Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan

D-Arabinofuranos, attached to either a galactofuran or a lipo- mannann, are the primary constituents of mycobacterial cell wall, forming the unique arabinogalactan (AG) and lipoarabinomannan (LAM), respectively. Emerging data indicate that the arabinans of AG and LAM are distinguished by virtue of the additional presence of linear termini in LAM, which entails some unknown feature of the EmbC protein for proper synthesis. In common with the two paralogous EmbA and EmbB proteins functionally implicated for the arabinosylation of AG, EmbC is predicted to carry 13 transmembrane spanning helices in an integral N-terminal domain followed by a hydrophilic extracytoplasmic C-terminal domain. To delineate the function of this C-terminal domain, the embC knock-out mutant of Mycobacterium smegmatis was complemented with plasmids expressing truncated embC genes. The expression level of serially truncated EmbC protein thus induced was examined by EmbC-specific peptide antibody, and their functional implications were inferred from ensuing detailed structural analysis of the truncated LAM variants synthesized. Apart from critically showing that the smaller arabinans are mostly devoid of the linear terminal motif, β-D-Ara(1→2)-α-D-Ara(1→5)-α-D-Ara(1→5)-α-D-Ara, our studies clearly implicate the C-terminal domain of EmbC in the chain extension of LAM. For the first time a full range of arabinan chains as large as 18–22 Ara residues and beyond could be released intact by the use of an endogenous endo-α-D-arabinanase from M. smegmatis, profiled, and sequenced directly by tandem mass spectrometry. In conjunction with NMR studies, our results unequivocally show that the LAM-specific linear termini are an extension on a well defined inner branched Ara-(18–22) core. This hitherto unrecognized feature not only allows a significant revision of the structural model of LAM-arabinan since its first description a decade ago but also furnishes a probable molecular basis of selectivity in biosynthesis, as conferred by the EmbC protein.

The major components of the mycobacterial cell wall are the mycolyl arabinogalactan-peptidoglycan complex and LAM,4 which have been the subject of intense research over the last 30 years (for review, see Refs. 1–4). Through this phenomenal volume of work, some unique features pertaining to these components have been clarified. The D-arabinan, a common constituent of both LAM and AG, is a unique homopolymer of about 40–70 d-arabinofuranose (Ara)f residues, assembled mostly into stretches of [-α-D-Ara(1→5)-α-D-Araf]n with critically spaced α3,5-Araf-branched sites. In AG, the nonreducing termini invariably form a well defined [β-D-Araf(1→2)-α-D-Araf]2β-5-Araf(1→5)-α-D-Araf motif (5), of which the terminal β-Araf and the penultimate α2-Araf serve as the anchoring points for the mycolic acids (6). Two such characteristic terminal Ara motifs with additional α5-Araf residues at the reducing end are assembled into a unique Ara22-mer, the largest structurally defined arabinan unit to date (7). Up to 3 such units are in turn attached in as yet unknown manner to the 5-position of the 6-β-D-galactofuranose residues of the D-galactofuran, located in close proximity to the linker region Rha(1→4)GlcNAc (8, 9).

In contrast to such an elaborate architecture of AG with highly structured arabinan, LAM on the other hand boasts a more heterogeneous structural entity based on several aspects such as charge distribution, size, and additional substituents within the molecule (10–12). A tripartite structure of LAM has been established and shown to consist of an arabinan chain on a mannan backbone with a phosphatidylinositol anchor that

---

8 The abbreviations used are: LAM, lipoarabinomannan; Ara, arabinofuranose; MS, mass spectrometry; EmbC<sub>ΔembC</sub>, recombinant strains with EmbC derivatives with C-terminal deletions; HSQC, heteronuclear single quantum correlation spectroscopy; HSQC-TOSY, proton-detected single quantum coherence-filtered total correlation spectroscopy; LM, lipomannan; MALDI-TOF, matrix-assisted laser desorption ionization-time of flight; Msm-arabinanase, endo-α-D-arabinanase from M. smegmatis; M. smegmatis<sub>ΔembC</sub>, M. smegmatis embC knock-out mutant; MOPS, 4-morpholinopropansulfonic acid; mAb, monoclonal antibody; PIM, phosphatidylinositol mannoside.
Arabinan Chain Length Extension in LAM

intercalates into the plasma membrane (Fig. 1). In Mycobacterium tuberculosis, the nonreducing ends of these arabinans are capped to varying degrees with mannose residues responsible for much of the LAM biological roles relating to host pathogen interaction. In M. smegmatis, a non-pathogen, this capping function is restricted to inositol phosphate on a minor fraction of the termini (13, 14). In addition to the branched nonreducing terminal Ara₆ motif as found in AG, a major difference that identifies the arabinan of LAM from that of AG is the distinctive presence of linear termini (15–17). A commonly held view is to attribute its formation to incomplete α-3-branching off the α-5-arabinan backbone before termination by β-2-arabinosylation, thus forming a linear Ara₄, β-D-Araf(1→2)-α-D-Araf(1→5)-α-D-Araf(1→5)-α-D-Araf, instead of the branched Ara₆ motif. The exact architectural setting and the underlying molecular basis governing their selective biosynthesis are largely unknown except that EmbA and EmbB are involved in this branching.

The genetics and biosynthesis of arabinan remain a gray area even with the availability of the M. tuberculosis genome sequence (18). The only biosynthetic information has come from the identification of a two-gene locus from Mycobacterium avium, the overexpression of which could render an otherwise susceptible M. smegmatis host resistant to ethambutol (19). Three paralogous genes in an operon encoding putative targets of ethambutol in M. smegmatis, M. tuberculosis, and Mycobacterium leprae were subsequently cloned and sequenced (20, 21). Two of these genes were similar to the embA and embB described in M. avium, as was the third paralog, embC, also similar to embA and embB. Although the Emb proteins have never been isolated nor their functions unequivocally established, critical biochemical analyses have demonstrated that (22) in individual knock-out mutants of embA and embB in M. smegmatis, AG defective at the nonreducing ends (because of the absence of the branched Ara₆ termini) was obtained; whereas the knock-out mutant of embC was unable to synthesize any LAM, synthesis of LM remained unaffected (23).

Recent work has identified short amino acid sequences within the Emb proteins that conform to the proline-rich motif found in a family of bacterial membrane proteins known as polysaccharide chain length determinants (24, 25). In addition, Emb proteins have been determined to belong to the glycosyltransferase superfamily C (26). Site-directed mutagenesis of an aspartic acid in the glycosyltransferase superfamily C motif of EmbC led to ablation of LAM synthesis, suggesting an arabinosyltransferase activity (27). Interestingly, introduction of point mutations in the conserved proline motif of EmbC proximal to the C-terminal domain led to synthesis of smaller arabinans largely devoid of linear Ara₄, thus generating a structure similar to AG. This finding is reminiscent of an earlier work in which replacing the C-terminal domain of EmbC with the corresponding domain of EmbB likewise disfavored the synthesis of linear termini (23).

We hypothesize that the C-terminal domain of the Emb proteins, in contact and perhaps acting in concert with the proximal proline motif, are mainly responsible for selective AG versus LAM-like arabinan formation. Our present work aims to substantiate the contribution of this domain using a recombinant cloning approach to delete various numbers of amino acids from the C-terminal end of EmbC. We demonstrate that C-terminal truncation of EmbC halts LAM-specific linear extension from an AG-like inner arabinan core structure. Collectively, our data reveal a new model that implicates this inner branched Ara-(18–22)-mer core structure as a shared intermediate utilized by enzymes destined for either LAM or AG synthesis.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Culture Conditions—Escherichia coli strain DH5α (Invitrogen) was grown at 37 °C on Luria-Bertani medium with or without ampicillin (100 μg/ml), kanamycin (50 μg/ml), or hygromycin (50 μg/ml). M. smegmatis mc²155 (22) and recombinant strains were grown at 37 °C in either 7H9 (liquid) or 7H11 (solid, Difco) supplemented with oleic acid-albumin-dextrose-catalase enrichment and glycerol with or without kanamycin (50 μg/ml) and hygromycin (50 μg/ml). Transformation of E. coli was carried out according to standard procedure. Transformation of M. smegmatis mc²155 was performed by electroporation as described (28).

The status of LAM/LM/PIMs was examined in the total extracts from colonies grown and harvested over 4 days.
Arabinan Chain Length Extension in LAM

logarithmic phase to stationary phase for M. smegmatis). After 4 days of growth, the cellular content of LAM detected was minimal in all strains. After this observation, to delineate the LAM-arabinan in the mutants, cells were harvested only in log phase (2.5 days for cells cultured in liquid medium).

Construction of Recombinant Plasmids Encoding the Truncated EmbC Proteins—The 3225-bp open reading frame of embC from M. smegmatis was amplified from genomic DNA by PCR using primers (embC-f 5'-GTT AGC ATC TAG ACA TAT TAT GAC CGG TCC GCA TGC AG-3'; embC-r 5'-AAG CTT CAA CTC AGC CGG AGG GCC GCG GGG-3') including NdeI and HindIII restriction sites (underlined), respectively. The PCR product was cloned into pSTBlue-1 (Novagen) for sequencing confirmation and subsequently integrated into the following criteria. The peptide shares low or no sequence similarity to BLAST searches performed at The Institute for Genomic Research. Also, the peptide is part of the first predicted loop of EmbC, putatively expressed from the genome of M. smegmatis according to BLAST searches performed at The Institute for Genomic Research. Also, the peptide is part of the first predicted loop of EmbC. On one occasion, the above plasmids were also electroporated into M. smegmatis mc²155.

Polyclonal Antibody to M. smegmatis EmbC-specific Peptide—The sequence 162DDDDPGEAVRGERSGYDFR180 of M. smegmatis EmbC was selected for peptide synthesis based on the following criteria. The peptide shares low or no sequence identity with M. smegmatis EmbA, EmbB, and other antigens putatively expressed from the genome of M. smegmatis according to BLAST searches performed at The Institute for Genomic Research. Also, the peptide is part of the first predicted loop of M. smegmatis EmbC and contains a low number of hydrophobic residues. Peptide (5 mg) was synthesized and conjugated to carrier diphtheria toxoid via the linker maleimidocaproyl-N-hydroxysuccinimide (Mimotopes, Roseville, MN). Polyclonal antiserum was raised in BALB/c mice (Harlan, Indianapolis, IN) immunized with an emulsion containing 50 µg of the M. smegmatis EmbC peptide coupled to the carrier diphtheria toxoid in a 1:1 ratio of phosphate-buffered saline and incomplete Freund’s adjuvant injected at two abdominal sites subcutaneously. Booster injections of 50 µg of antigen in incomplete Freund’s adjuvant emulsion were repeated every 3 weeks, with test bleeds taken for enzyme-linked immunosorbent assay after the first and second booster. The titer of the antiserum after the third boost was >1:5000 as measured by Western blot against the native protein enriched in the membrane fraction of M. smegmatis or enzyme-linked immunosorbent assay (ELISA) assay by testing serial 2-fold dilutions of the serum against the free peptide (coated on ELISA plate wells at 1 µg/well).

Detection of EmbC in M. smegmatis Wild-type and Recombinant Strains by Western Blot—Cells were suspended in a buffer containing 50 mM MOPS, 10 mM MgCl₂, and 5 mM β-mercaptoethanol, pH 8.0, in the presence of protease inhibitor mixture (Sigma) and disrupted by intermittent probe sonication (Soni-prep 150, Sanyo Gallenkamp PLC) for 10 cycles (60-s bursts/90 s of cooling). Individual sonicated samples were digested with 10 mg/ml DNase and RNase for 1 h at 4 °C. Cell walls were obtained by centrifugation at 27,000 × g for 30 min, and the supernatants from this step were recentrifuged at 100,000 × g for 2 h to provide a sediment enriched in membrane. Protein concentration in each of the cell wall, membrane, and cytosolic fractions was determined by BCA assay before subjecting to 10% SDS-PAGE and transferred to nitrocellulose membranes as described (29). The membranes were blocked overnight with 2% bovine serum albumin in Tris-buffered saline incubated with polyclonal mouse serum (1:2000) for 4 h. After washing, the membranes were incubated with anti-mouse IgG-alkaline phosphatase-conjugated antibody (Sigma). The substrates nitro blue tetrazolium and 5-bromo-4-chloro-3-indolylphosphate (Sigma FAST BCIP) were used for color development.

Extraction of LAM/LM/PIMs from M. smegmatis Wild-type and Recombinant Strains—Cells harvested from 4-liter lots of broth cultures in log phase were washed with phosphate-buffered saline and lyophilized. Dry cells were delipidated with organic solvents and disrupted mechanically with the Soniprep 150. The resulting homogeneous suspension was refluxed in 50% ethanol at 80 °C (3 × 2 h each). The combined supernatants were evaporated and digested with proteinase K (Invitrogen; 1 mg/ml). After dialysis, LAM/LM/PIMs fractions were dried (10, 30).

Size Fractionation, SDS-PAGE, and Immunoblotting—For purification of LAM/truncated LAM from all strains used in this study, size exclusion chromatography was performed on a Rainin SD 200 series liquid chromatography system fitted with a Sephacryl S-200 HiPrep 16/60 column in tandem with a Sephacryl S-100 column (Amersham Biosciences) at a flow rate of 1 ml/min (30). SDS-PAGE followed by periodic acid Schiff staining (31, 32) was used to monitor the elution profile of the fractions containing LAM and LM, which were then pooled and dialyzed. Fractions were re-analyzed by SDS-PAGE and immunoblotting with mAb CS-35 as described previously (31) to check for purity before detailed analysis.

Digestion with Endoarabinanase and Subsequent Analyses—Preparation of Cellulomonas gelida endoarabinanase has been described (5). To ensure complete digestion of LAMs, the reaction mixture was incubated overnight, and the digestion products were analyzed directly by Dionex analytical high pH anion exchange chromatography performed on a Dionex liquid chromatography system fitted with a Dionex Carbopac PA-1 column. The oligoarabinosides were detected with a pulse-amperometric detector (PAD-II) (Dionex, Sunnyvale, CA).

Monosaccharide Composition—Samples were hydrolyzed with 2 M trifluoroacetic acid, converted to alditol acetates, and analyzed by gas chromatography (33) on a Hewlett Packard gas chromatography model 5890 fitted with a SP 2380 column (30-m × 0.25-mm inner diameter) at an initial temperature of 50 °C for 1 min, increasing to 170 °C by 30°/min followed by 270 °C by 5 °C/min. Gas chromatography/mass spectrometry (MS) of the methylated alditol acetates were 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 before injection on a DB-5 column.
Digestion with Msm-arabinanase—Preparation of Msm-arabinanase has been described (34). More recently this enzyme has been sufficiently purified for use as a structural tool via ammonium sulfate precipitation and hydrophobic interaction chromatography (35). Purified LAM (2 mg) from each of the recombinant strains was digested with 250 μl of the Msm-arabinanase in 25 mM phosphate buffer at pH 7.0 and incubated for 4 h at 37 °C. The digestion products containing both the mannose core and the released oligoarabinosides were separated using Amicon Microcon YM-30 (Millipore) and centrifuging at 6,500 × g for 20 min. The flow-through was then directly permethylated (36) for MS analysis or after further purification by high performance liquid chromatography (Waters) on a Superdex™ Peptide 10/300 column (Amersham Biosciences) eluted with water at the rate of 0.5 ml/min. The elution profile was monitored using the Waters refractive index detector, model 2414. The fractions were pooled and analyzed.

MALDI-MS and MS/MS Analyses—The arabinan fragments released by Msm-arabinanase were either permethylated or prerred and permethylated using the NaOH/dimethyl sulfoxide slurry method as described (36). For MALDI-TOF MS glycan profiling, the permethyl derivatives in acetonitrile were mixed 1:1 with 2,5-dihydroxybenzoic acid matrix (10 mg/ml in water) for spotting onto the target plate. MS data acquisition in the reflectron mode was performed on a 4700 Proteomics Analyzer mass spectrometer (Applied Biosystems, Framingham, MA) equipped with an Nd:YAG laser (355-nm wavelength, <500-ps pulse, and 200-Hz repetition rate). MALDI-MS/MS was performed on a dedicated quadrupole TOF Ultima MALDI instrument (Micromass) in which case the permethylated samples in acetonitrile were mixed 1:1 with 2,5-dihydroxybenzoic acid matrix (10 mg/ml in water) for spotting onto the target plate. The nitrogen UV laser (337-nm wavelength) was operated at a repetition rate of 10 Hz under full power (300 μJ/pulse). MS survey data were manually acquired, and the decision to switch over to collision-induced dissociation MS/MS acquisition mode for a particular parent ion was made on-the-fly upon examination of the summed spectra. Argon was used as the collision gas with a collision energy manually adjusted (between 50–200 V) to achieve optimum degree of fragmentation for the parent ions under investigation.

NMR Spectroscopy of LAM Samples—Spectra were acquired after several lyophilizations in D₂O of 4–5 mg/0.6 ml in 100% D₂O. Two-dimensional 1H,13C(13C) heteronuclear single quantum correlation spectroscopy (HSQC) NMR spectra were acquired on a Varian Inova 500 MHz NMR spectrometer (NIH-SIG RR19811) using the supplied Varian pulse sequences. The HSQC data were acquired with a 7-kHz window for proton in F2 and a 15-kHz window for carbon in F1. The total recycle time was 1.65 s between transients. Composite pulse, GARP, 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 1000 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. Cos² weighting functions were matched to the time domain in both t1 and t2, and zero-filling was applied to both t1 and t2 before the Fourier transform. The final resolution was 3.5 Hz/point in F2 and 15 Hz/pt in F1. Two-dimensional 1H,13C(13C) proton-detected single quantum coherence-filtered total correlation spectroscopy (HSQC-TOCSY) NMR experiments were acquired on a Varian Inova-500 NMR spectrometer using the supplied Varian pulse sequence. The HSQC-TOCSY data were acquired with a 5.5-kHz window for proton in F2 and a 15-kHz window for carbon in F1. The TOCSY mixing time was 0.120 ms applied with a 7-kHz field. The total recycle time was 1.65 s between transients. GARP 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 1000 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. Cos² weighting functions were matched to the time domain in both t1 and t2, and zero-filling was applied to both t1 and t2 before the Fourier transform. The final resolution was 3 Hz/point in F2 and 15 Hz/point in F1.

RESULTS

Expression of C-terminal-truncated EmbC Proteins and Synthesis of Truncated LAM Variants—A prime focus of this work is to assign functions to the predicted C-terminal domain of the EmbC protein. To this end, primers were designed (supplemental Table I) to generate a series of EmbC proteins truncated from the C-terminal domain as depicted in Fig. 2. Recombinant plasmids were integrated into M. smegmatisΔembC (23), and the expression level of the respective differentially truncated EmbC protein was examined by Western blot using a polyclonal antibody generated against the peptide antigen from a part of the first predicted loop of M. smegmatis EmbC (residue 162–180, Fig. 2). LAM/LM synthesized were in turn extracted, analyzed by SDS-PAGE (Fig. 3A), and immunoblotted with mAb CS-35 (Fig. 3B), which recognizes the arabinan of LAM (37). As expected, the M. smegmatisΔembC (Fig. 3, A and B, lane 2) failed to synthesize LAM, evident by a lack of positive periodic acid Schiff staining and negative response with CS-35, which was further substantiated by neutral sugar composition (23), but when complemented with the wild-type embC gene (strain designated as EmbCWT from here onward), LAM of a larger size (lane 3) than that made by the wild-type strain (lane 1) was observed. Accordingly, a sharp protein band corresponding to 115.0 kDa (theoretical molecular mass of M. smegmatis EmbC) was evident by Western blot analysis of the membrane extracts from the wild-type M. smegmatis. This band was absent in M. smegmatisΔembC (Fig. 4, lane 2) or in the strain complemented with pVV16 (lane 3), whereas the EmbCWT strain (lane 4) showed an significantly increased level of protein expression (based on staining intensity of the protein band per the same amount of total proteins loaded on the SDS/PAGE) relative to that of the wild-type strain (lane 1).
An almost equivalent amount of truncated EmbC protein expression was observed for EmbC\textsubscript{Δ50c} and EmbC\textsubscript{Δ100c} (109.5 and 104.0 kDa) at the expected regions of the blot (Fig. 4, lanes 5 and 6). In contrast, protein bands of ~55.0 kDa (lanes 7–10) were obtained for EmbC\textsubscript{Δ200c}, EmbC\textsubscript{Δ300c}, EmbC\textsubscript{Δ368c}, and EmbC\textsubscript{Δ407c}. Thus, truncation beyond the first 200 amino acids from the C terminus apparently compromised its stability, leading to its proteolytic cleavage from the C-terminal end without significantly affecting its expression level. It is, however, most likely that the respective EmbC proteins were either degraded immediately post-translation or once extracted out from their natural environment after disruption, as all of the deletion derivatives led to some smaller proteolytic products. All the recombinant strains expressing truncated EmbC were found to synthesize truncated LAM (Fig. 3, lanes 4–9), which retained a strong positive response to mAb CS-35 (Fig. 3B). This antibody is known to react more avidly with the branched terminal Ara\textsubscript{f} moiety of mycobacterial arabinans present in both AG and LAM that encompasses the minimal recognition epitope, β-D-Araf(1→2)-α-D-Araf(1→5)-α-D-Araf(1→5)-α-D-Araf (37). Thus, C-terminal truncation of EmbC up to about 400 amino acids or so led to impaired synthesis of LAM but did not alter the underlying arabinosyl epitope as defined by mAb.

**FIGURE 2.** Schematic representation of the transmembrane topology of \textit{M. smegmatis} EmbC predicted by Phobius (50), predicted to be a 115.39-kDa protein (1074 amino acids) with 13 transmembrane helices (TMHMM 2.0, Phobius) and a 400-residue C terminus located outside of the membrane. The last four loops have an excess of positively charged residues and are predicted to be on the cytoplasmic side of the membrane (51). The amino acid deletion points from the C-terminal end in the six recombinant EmbC\textsubscript{\textsubscript{C} constructs are indicated with blue arrows. The Pro motif and the glycosyltransferase superfamily C motif are indicated with circles (27). The peptide sequence highlighted in pink is used to generate antibody and is in the first predicted extracytoplasmic loop. **FIGURE 3.** Analysis of LAM and LM of wild-type \textit{M. smegmatis} and the mutants by SDS-PAGE (A) and Western blot using mAb CS-35 (B). Molar mass markers are shown on the left. Lane 1, wild-type \textit{M. smegmatis}; lane 2, \textit{M. smegmatis}\textsubscript{ΔembC}; lane 3, EmbC\textsubscript{WT}; lane 4, EmbC\textsubscript{Δ50c}; lane 5, EmbC\textsubscript{Δ100c}; lane 6, EmbC\textsubscript{Δ200c}; lane 7, EmbC\textsubscript{Δ300c}; lane 8, EmbC\textsubscript{Δ368c}; lane 9, EmbC\textsubscript{Δ407c}. Two to three colonies were collected from a plate and extracted with 215 μl of solvent (1:1:0.1, chloroform/methanol/water) at 55 °C for 30 min. After centrifugation, the residual pellet was extracted with an equal volume of water (200 μl) and phosphate-buffered saline-saturated phenol (200 μl) at 80 °C for 2 h. The aqueous layer (containing LAM, LM, and PIMs) was dialyzed and analyzed by SDS-PAGE followed by periodic acid Schiff staining and Western blot with CS-35 antibody (23).

**FIGURE 4.** Western blot analyses with anti- \textit{M. smegmatis} EmbC peptide antibody. Membrane preparations from strains discussed in this study were used. SDS-PAGE was performed on 10% gel. Lanes 1–3, each with 20 μg of protein; lanes 4–10, each with 3 μg of protein. MW, molecular mass markers; lane 1, wild-type \textit{M. smegmatis}; lane 2, \textit{M. smegmatis}\textsubscript{ΔembC}; lane 3, \textit{M. smegmatis}\textsubscript{ΔembC} complemented with pVV16; lane 4, \textit{M. smegmatis}\textsubscript{ΔembC} complemented with the wild-type \textit{embC}; lane 5, EmbC\textsubscript{Δ50c}; lane 6, EmbC\textsubscript{Δ100c}; lane 7, EmbC\textsubscript{Δ200c}; lane 8, EmbC\textsubscript{Δ300c}; lane 9, EmbC\textsubscript{Δ368c}; lane 10, EmbC\textsubscript{Δ407c}.
To reflect the percentage change in Ara and Man content in LAM preparations, the Ara and Man in EmbCWT was considered as 100%.

TABLE 1

Composition analysis of M. smegmatis wild-type LAM and truncated LAMs from the EmbCΔc constructs

| Strains       | Ara | Ara % change | Man | Man % change | Myo-Ins | Ara/Man ratio |
|---------------|-----|--------------|-----|--------------|---------|---------------|
| Wild-type LAM | 74  | NA           | 28  | NA           | 1       | 2.63          |
| EmbCWT        | 113 | 100          | 29  | 100          | 1       | 3.89          |
| EmbCΔ50c      | 40  | −64.6        | 25  | −10.7        | 1       | 1.60          |
| EmbCΔ100c     | 40  | −64.6        | 26  | −7.1         | 1       | 1.53          |
| EmbCΔ200c     | 31  | −72.4        | 22  | −21.4        | 1       | 1.41          |
| EmbCΔ300c     | 35  | −69.0        | 25  | −10.7        | 1       | 1.40          |
| EmbCΔ407c     | 75  | −33.6        | 35  | +25.0        | 1       | 2.14          |
| EmbCΔ407c     | 71  | −37.1        | 37  | +32.1        | 1       | 1.92          |

*To reflect the percentage change in Ara and Man content in LAM preparations, the Ara and Man in EmbCWT was considered as 100%.
Arabinan Chain Length Extension in LAM

sis revealed that the arabinan comprises t-Araf, 3,5-Araf, 5-Araf, and 2-Araf units (data not shown). By comparing the $^{13}$C,$^1$H heteronuclear NMR spectra from wild-type LAM (Fig. 7A) and EmbC$_{200c}$ LAM (Fig. 7B) and in agreement with previous studies that report two-dimensional NMR of wild-type LAM (38, 39), the overall arabinan domain seemed to hold in the mutant. The NMR parameters are consistent with the methylation analyses albeit with notable differences (Fig. 7B). There were altogether seven well differentiated and one overlapping spin system distributed in the anomic region ($\delta$1000–$\delta$109.3 ppm). Based on the literature (12, 38–41), the $^{13}$C resonance at $\delta$100.89 ppm that correlated to anomic protons at $\delta$5.13 ppm with $^1J_{H1,C1}$ coupling constants of $\sim$170 Hz was attributed to the 2,6-$\alpha$-mannopyranose. t-$\alpha$-mannopyranose resonated at $\delta$104.9 ppm and correlated to protons at $\delta$5.06 ppm. The cross-peak at $\delta$102.15 ppm correlating with $^1$H at $\delta$4.92 was assigned to 6-$\alpha$-mannopyranose and was weak due to saturation of the water peak around the same region. Two completely overlapping sets of cross-peaks centered at $\delta$103.35 with $^1$H at $\delta$5.16 corresponded to the t-$\beta$-Araf. Two well separated spin systems were identified for 2-$\alpha$-Araf (cross-peaks for overlapping C-2 were detected at 888.7), namely those attached to the 3-position (2-$\alpha$-Araf$\rightarrow$3, $\delta$108.19, $\delta$5.27 ppm) and 5-position (2-$\alpha$-Araf$\rightarrow$5, $\delta$108.44, $\delta$19.19 ppm) of the 3,5-$\alpha$-Araf.

As expected, multiple distinct spin systems were observed centered between $\delta$110.4 and $\delta$109.5 ppm, correlating with protons between $\delta$5.11 and $\delta$5.18 ppm. These were all attributed to the 5-$\alpha$-Araf in different chemical environments overlapping with one set of 3,5-$\alpha$-Araf ($\delta$5.18–5.11). Most of the resonances in the HSQC spectrum obtained from EmbC$_{200c}$ LAM (Fig. 7B) were in agreement with those assigned for wild-type LAM (Table 2) except for two additional cross-peaks centered at $\delta$110.7, $^1$H at $\delta$5.25 ppm and $^{13}$C at $\delta$110.63 with $^1$H at $\delta$5.04 ppm. The chemical shift ranging between $\delta$110 and $\delta$108 was supportive of the $\alpha$-anomeric configuration of the Araf residues. The cross-peaks at $\delta$5.25 and $\delta$5.04 ppm were new ($\delta$109.7 and 110.63), which could be attributed to terminal $\alpha$-Araf (42) arising from incomplete $\beta$-arabinosyl termination of a growing arabinan chain. Two distinct chemical shifts of t-$\alpha$-Araf could occur for the respective two arms of a branched

5 J. Zhang and D. Chatterjee, unpublished work.
reduced linear extension of an underlying branched arabinan framework, manifested as apparent lack or diminished expression of linear Ara\textsubscript{4} terminal motif relative to the branched Ara\textsubscript{6} motif. A comparative NMR analysis of LAM produced by a hybrid Emb protein carrying the two-thirds of the N terminus of EmbC and one-third of the C terminus of EmbB afforded a

FIGURE 7. Comparative partial two-dimensional NMR spectra of LAM purified from wild-type M. smegmatis (A) and EmbC\textsubscript{3200}. (B). \textsuperscript{1}H,\textsuperscript{13}C HSQC spectra were acquired in D\textsubscript{2}O. Only the expanded anomeric regions are shown (\textsuperscript{1}H, 6.46.60–65.05 ppm, and \textsuperscript{13}C, 98–110.5 ppm). t-\textalpha-Araf was assigned from chemical shift and a separate HSQC-TOCSY experiment (data not shown).
the mutants yielded mostly Ara17–20 as the major products, smaller arabinan size. Second and more important, instead of a Ara22 or Ara23, which would be consistent with an overall AG-like arabinan with the characteristic linear Ara4 motif largely missing (23).

MALDI-MS and MS/MS Mapping of Large Arabinan Fragments Released from LAM by Msm- arabinanase—Incomplete elaboration of the arabinan due to the deleted functional domain of EmbC would lead to overall truncation in size and, by implication, an altered branching pattern of the subterminal structural domains due to reduced linear extension. We next sought to demonstrate this effect by mapping the large arabinosyl oligomers that would be released from LAM by an endogenous Msm-arabinanase previously shown to release large oligoarabinosyl units from AG (34). The digestion products were filtered through Amicon Microcon YM-30 to remove any large goarabinosyl units from AG (34). The digestion products were permethylated for direct MALDI-MS profiling and MS/MS sequencing.

Although the wild-type LAM afforded a full range of arabinosyl oligomers from Ara9 to more than Ara30 (Fig. 8A), the profiles afforded by the EmbC200L and EmbC358C (Fig. 8, B and C) were distinctively different in several aspects. First, the released arabinosyl oligomers apparently extended only up to Ara22 or Ara23, which would be consistent with an overall smaller arabinan size. Second and more important, instead of a similar abundance for each of the oligomers as in the wild type, the mutants yielded mostly Ara17–20 as the major products, peaking at Ara18. The most pronounced effect was afforded by LAMs isolated from EmbC200L and EmbC358C, which correspond also to the most severely truncated LAMs as visualized by SDS-PAGE (Embc358c was not included in the SDS-PAGE because its electrophoretic mobility was similar to the EmbC300L). Interestingly, this drastic change in arabinosylation pattern was somewhat reversed with further truncation of the C-terminal domain. The corresponding MS profile for EmbC358c (Fig. 8D) exhibited a pattern more similar to that afforded by the wild-type LAM, albeit still with slightly more abundant Ara17–20. The data are, thus, consistent with the findings that its Ara content has increased back to the level approaching that of wild-type LAM (Table 1).

It is tempting to speculate that such an apparent change in the arabinosylation pattern was indeed due to impaired EmbC function in extending the arabinan in a linear manner. Thus, the Ara17–20 oligomers would correspond to a fundamental core structure commonly found in both AG and LAM, as defined by previous analysis to be an assembly of two uniquely branched Ara units (43). Supporting data were derived from MS/MS sequencing of each of the arabinosyl oligomers released from the wild-type LAM (Fig. 9) in comparison against a representative set from EmbC358C LAM (Fig. 10). To facilitate signal assignment, the samples were first reduced to convert the oligomers to oligoarabinosyl alditols and then permethylated. Each single cleavage induced by MS/MS would generate fragment ions with a free hydroxyl group. A fragment ion, which can only be derived by two cleavages, will be distinguished from the former by virtue of having an additional OH group.

In practice, a linear stretch of arabinosyl chain at the nonreducing end of an oligoarabinosyl alditol is identified by the ability to detect consecutive losses of Ara residues from nonreducing end with single cleavage events. This was indeed the case with each of the oligoarabinosyl alditols derived from wild-type LAM. For example, the loss of a fully methylated nonreducing terminal Ara residue from Ara9(m/z 1365) yielded the sodiated y ion at m/z 1191. Subsequent losses of Ara residues did not generate additional free OH groups. Thus, loss of 1–6 Ara residues could each occur by single cleavage to afford primary fragment ions at m/z 1191, 1031, 871, 711, 551, and 391, respectively. Nevertheless, loss of a third Ara residue (and onward) could also favorably be derived from double cleavages, giving rise to the fragment ion at m/z 857 or 14 units (a methyl group) lower than that produced by single cleavage. The data are, therefore, indicative of a significant degree of branching occurring at the third Ara residue from the nonreducing end, which exists along a linear structure (Fig. 9A). The same pattern was observed for all other oligoarabinosyl alditols (Fig. 9, B–D), where losses of as many Ara residues could all occur via single cleavage event, but the existence of branched points was also apparent.

In contrast, a distinctive fragmentation pattern could be observed for the corresponding oligoarabinosyl alditols derived from the LAM of EmbC358C (Fig. 10). For Ara14, a branched point apparently exists at the ninth Ara residue from the nonreducing end since losses of up to eight but not nine Ara residues from the nonreducing terminus could occur via single cleavage. Instead of prominent pairs of sodiated y ions at m/z 857/871 and 697/711, as afforded by the wild-type LAM (Fig. 9C), the truncated LAM gave mainly m/z 857 and 697 (Fig. 10A), indicating a stricter adherence to branching at the expense of linear extension. Alternatively, the branching pattern was made more apparent in the absence of linear extension. Even more strikingly for the Ara14, single-cleavage loss of three and four Ara residues from the nonreducing terminus was much more disfavored than that afforded by the corresponding Ara18 from wild-type LAM (Fig. 9D). A drop in intensity was associated with higher abundance of the double cleavage fragment ions at m/z 2459 and 2299 relative to those from single cleavage at m/z 2473 and 2313. As in Ara14, a branched point is again apparent at residue nine from the nonreducing terminus that led to detection of y ion at m/z 1497 instead of m/z 1511. The dominant expression of a character-

### Table 2: Arabinan Chain Length Extension in LAM

| Glycosyl residues | M. smegmatis LAM | EmbC358c LAM |
|-------------------|------------------|--------------|
| 13C | 1H | 13C | 1H |
| t-1-Araf | 103.35 5.16 | 103.44 5.16 | 103.35 5.16 | 103.44 5.16 |
| t-2-Araf | 108.19 5.27 | 108.26 5.27 | 108.19 5.27 | 108.26 5.27 |
| t-2-Araf | 108.44 5.19 | 108.46 5.20 | 108.44 5.19 | 108.46 5.20 |
| t-3-Araf | 110.22 5.11 | 110.28 5.11 | 110.22 5.11 | 110.28 5.11 |
| t-1-Araf* | ND | ND | 109.70 5.25 | ND |
| t-2-Araf* | ND | ND | 110.63 5.04 | 110.63 5.04 |
| t-3-Araf* | 110.15 5.14 | 110.2 5.14 | ND | ND |
| s-1-Araf | 109.88 5.18 | 109.9 5.19 | 109.88 5.18 | 109.9 5.19 |
| s-1-Manp | 104.9 5.06 | 104.95 5.07 | ND | ND |
| s-2-Araf | 102.15 4.92 | 102.19 4.92 | ND | ND |
| 2,6-s-Manp | 100.89 5.13 | 100.96 5.14 | 100.89 5.13 | 100.96 5.14 |
istic branched terminal Ara₈ structural motif was likewise observed for the Ara₁₇ and Ara₁₉ oligomers. In both cases (Fig. 10, B and D), loss of up to eight but not nine Ara residues from the nonreducing termini could occur via single cleavage. Thereafter, single cleavage loss of further Ara residues was mostly disallowed except for Ara₁₉ in which a loss of an Ara₁₇ moiety could again occur by single cleavage, giving rise to the ion at m/z 391. Thus, two Ara₈ moieties may be co-joined at another branched residue proximal to the reducing end to yield the implicated Ara₁₇ moiety, as schematically shown in Fig. 10.

Taken together, both MS profile and MS/MS data are consistent with a structural model for the arabinan of LAM in which additional linear α-5-Ara extensions occur on a common Ara₂₂ oligomeric base, a function largely mediated by the last 350 amino acids of the EmbC. In wild-type LAM, such a structural motif would also be present, but the fragmentation pattern of the corresponding Ara₁₈ was not as distinct since other isomeric structures with linear stretches of oligoarabinosyl chains coexist. A reduced linear extension due to truncation in the C-terminal functional domain of EmbC also reduced the

![Arabinan Chain Length Extension in LAM](image)

FIGURE 8. MALDI-MS profiling of arabinosyl oligomers released by Msm-arabinanase. From wild-type LAM (A) or truncated LAMs from mutants EmbC₁₂₀c (B), EmbC₁₃₅c (C), and EmbC₁₅₀c (D). The corresponding MS profiles of EmbC₁₂₀c and EmbC₁₅₀c are very similar to that of EmbC₁₃₅c which is shown here as a representative spectrum. For all mutants, Ara₁₇–₁₉ became the most prominent Msm-arabinanase digestion products, and of those investigated, the most pronounced change was afforded by EmbC₁₃₅c (C). As shown in B and C, severe truncation led to detection of another cluster of signals at around m/z 4000 corresponding to Ara- and Hex-containing components as could be inferred from their mass intervals. These are likely to be derived from mannan core with remaining Ara attached, but their exact identity was not further investigated in this study. With EmbC₁₃₅c (D), the profile apparently reverted back to resemble more that afforded by wild-type LAM (A), although Ara₁₇–₁₉ still registered as the major products. All major arabinosyl oligomers from the digests of wild-type LAM and LAM from EmbC₁₃₅c were selected for MALDI-MS/MS analyses, but only the MS/MS spectra for those peaks labeled as Araₙ in bold in A and C are shown in Fig. 7 and 8, respectively. Manp, mannopyranose.
overall heterogeneity, leading essentially to a more AG-like arabinan.

**DISCUSSION**

The arabinan in mycobacteria has important structural and pathogenic implications. The covalent modification of the arabinan termini of AG with mycolic acids results in an effective lipid barrier, whereas mannose capping of *M. tuberculosis* LAM generates crucial immunomodulatory properties in the host. It is, therefore, a valid therapeutic target because blocking its biosynthesis would disrupt both the cell wall *per se* and its associated biologically active LAM (44–46). However, many of the structural features of arabinan remain incompletely understood, largely due to unavailability of defined mutants and appropriate analytical tools to address its structural complexities.

Until very recently little was known about the genes involved in the arabinan synthesis. Through genetic and biochemical studies, the EmbC protein specifically has now emerged as the candidate LAM-arabinan transferase (23, 27), which in contrast to the related EmbA and EmbB proteins (destined for AG-arabinan biosynthesis) (21, 22), is reckoned to carry an additional recognition motif that would accommodate the extra degree of flexibility necessary for the synthesis of the extended LAM-type arabinan. To explain the findings that interruption of the *embC* gene leads to a “complete cessation” in the synthesis of the arabinan in LAM but not AG, whereas LM synthesis remains unaffected, we propose that in the LAM arabinan biosynthetic pathway segments of the EmbC are involved in the transfer of crucial preformed polyprenyl-linked arabinan chains. Further delineation of the structure-function aspects of the Emb proteins has, thus, become vital in understanding the enzymology of biosynthesis of LAM and AG.

In our present study we have exploited the “LAMless” phenotype of the knock-out strain *M. smegmatis*Δ*embC* to artificially generate LAM variants by a recombinant cloning approach. Thus, when the LAMless phenotype of the knock-out mutant was rescued with plasmids expressing EmbC derivatives with C-terminal deletions, truncated and structurally altered LAM yet recognized...
by LAM-specific antibody CS-35 were detected. The significant decrease in the abundance of linear Ara4 in LAM as obtained by the Cellulomonas endoarabinanase digestion, however, suggested a dramatic alteration in the terminal end. More importantly, the use of the endogenous Msm-arabinanase provided us with an advantage to examine for the first time the arabinan structures beyond just the non-reducing terminal motifs.

Larger oligoarabinosyl fragments could be obtained and readily defined by MALDI-MS and MS/MS analyses, which enabled us to show that the branched terminal Ara6 motif is carried on an Ara18 that descends from the Ara22-mer described many years ago in a seminal study on AG (7). For years it was believed that this Ara18 (Fig. 11) was only present in AG, and three such motifs completed the AG-arabinan, whereas the arabinan of LAM was less structured. On the other hand, the terminal Ara4 motif arises from linear stretches of arabinan, which is the hallmark of LAM arabinan only. All the structural evidence now points to the arabinan of LAM as a composite of both linear terminal stretches and an AG-like branched Ara18 core unit. However, how these are distributed...
on the mannan backbone is unclear. Our present data establish for the first time that the nonreducing terminal linear extension from any branched point (α3,5-Araf) is up to eight or nine Araf residues (as shown in Fig. 11). In wild-type LAM, a vast range of heterogeneity in size is apparent. The detection of up to Ara30 among the Msm-arabinanase digest indicates that a portion of the precursor 1 (which is also drawn as 18 Araf residues but could be of a different size) is further extended and would be

FIGURE 11. Model for assembly of arabinan of LAM. The model is based on the information obtained from various structural annotations and functional consequence of genetically manipulated EmbC protein and, therefore, is a "proof of concept" at this stage of our work. Biochemically this has not been established yet. Ara18 and Ara30 can be released intact by digestion with Msm-arabinanase. The specificity of this enzyme with respect to possible cleavages at other sites is currently not well established. Ara4 and Ara6 can be liberated with the Cellulomonas endoarabinanase. There is no structural evidence for Precursor 1, although it is drawn showing 18 Araf residues. This must not be confused with Ara30 (referred to the structure liberated by Msm-arabinanase). Precursor 1 has no β-Araf capping, a crucial event of chain termination. DPA, decaprenylphosphorylarabinose.
consistent with a structural model depicted in Fig. 11. Because various degrees of linear extension coexist with the branched Ara\textsubscript{a} terminal motif carried on the quintessential Ara-18 unit, a full range of digestion products would, thus, be predictably derived as demonstrated in this study. Each Ara\textsubscript{a} detected is furthermore a composite of various isomers and, therefore, could not give a distinctive MS/MS fragmentation pattern.

The noticeable reduction of a significant portion of the linear arabinan stretches could be inferred as the primary lesion leading to truncation in size of LAM made by the recombinant strains containing EmbC derivatives with C-terminal deletions. As a consequence the overt phenotype is manifested as AG-like since the Ara\textsubscript{a} unit became the primary default product in the absence of further extension. It should, however, be noted that unlike the arabinan of AG, a significant portion of the truncated arabinans thus formed was not terminated with β-Araf. Our NMR data clearly showed the appearance of terminal α-Araf otherwise not found on wild-type LAM or AG. Thus, irrespective of the contribution of in vivo arabinanase activity, the wild-type LAM digest product presented is quite clear from Fig. 8 where up to Ara\textsubscript{30} oligomers were obtained. The fact remains that, with EmbC derivatives, we see a distinct profile, which is a derived as demonstrated in this study. Each Ara\textsuperscript{n}NaC\textsubscript{50}-P-Ara\textsubscript{f} on the arabinan assembly involving a long chain of 5-linked Ara\textsubscript{(1-8)} and its linear extension, the existence of a coordination and number of arabinosyltransferases required in forming the arabinan in LAM, perhaps in conjunction with other Emb proteins.

Given the complexity of the structure of the arabinan of LAM and number of arabinosyltransferases required in forming the Ara-(18–22) and its linear extension, the existence of a coordinated multienzyme complex in growing mycobacterial culture seems to be logical. Decaprenylphosphorylarabinose is the only known Ara\textsubscript{f} donor (9, 47) described. If we draw an analogy of the LAM synthesis with LPS synthesis, the Emb proteins could act in arabinoxylan assembly involving a long chain of 5-linked Ara\textsubscript{f} on a C\textsubscript{50}-P-Araf, which has been already translocated across the plasma membrane (48, 49).

In our current working biosynthesis model (Fig. 11), we postulate that an intermediate comprising of α-linked Ara\textsubscript{f} (Precursor 1 in Fig. 11) units is pre-synthesized from several decaprenylphosphorylarabinose units and ligated to either LM- or lipid-linked galactan. The ligation is dependent on the N-terminal domain of the Emb proteins. AG versus LAM biosynthetic pathways may then segregate, as it would require distinct enzymatic attributes to linear extend versus β capping of the α-linked Ara\textsubscript{f} precursor. We have shown that for LAM, this linear extension is dependent on the C-terminal domain of EmbC. In the wild-type strain this would give rise to as large as an Ara\textsubscript{30} oligomer upon Msm-arabinanase digestion corresponding to the upper end of size distribution of the digestion products detected by MALDI-MS. Other LAM-type arabinan variants may have less extensive linear extension or with none or only one of the possible four termini of Precursor 1 extended.

Based on this working model, we further speculate that the Emb proteins may function as dimers (or higher multimers). The preferred dimers in the wild-type strain are likely to be EmbC/EmbC and EmbA/EmbB for the synthesis of LAM and AG, respectively. For the latter, we have been able to demonstrate that EmbA or EmbB alone are unable to transfer the disaccharide β-(1→2)-α-D-Araf onto a synthetic linear acceptor, but this transferase activity is restored when the enzyme source was combined from \emph{M. smegmatis}\textsubscript{embB} and \emph{M. smegmatis}\textsubscript{embA} mutants.\textsuperscript{6}

However, other dimers may exist under circumstances such as the lack of the ideal partner or changes in the C-terminal domain. Conceptually, then, for the EmbC C-terminal deletion mutants reported herein, the active enzymes could conceivably include the N terminus of EmbC that would ligate the arabinan to LAM, but the C terminus of EmbA or EmbB failed to commit the further elongation steps. It is surprising that in \emph{M. smegmatis}\textsubscript{embC}, which is unable to synthesize LAM, we have never been able to find any arabinan-containing precursors such as the Precursor 1 in Fig. 11. This can be rationalized as (a) there could be a rapid turnover of these metabolites, (b) degradation of the intermediate arabinan occurs with the endogenous Msm-arabinanase, and most likely, (c) EmbC works in concert with other proteins and utilizes the intermediates. The third possibility is currently being addressed.

All glycosyltransferases with dependence for a polyprenol-linked sugar substitute (proteins of glycosyltransferase superfamily C) have so far been predicted to contain multimembrane-spanning domains. A stage is now set to track down these elusive arabinan intermediates and arabinosyltransferases. We expect most or all of the proteins involved in the arabinan synthesis to be integral membrane proteins with membrane-spanning domains.

Acceptors with structures resembling to Precursor 1 are now being chemically synthesized, so that assays can be developed in the future for chain length elongation or capping to substantiate the hypothetical model presented in this work.

Acknowledgments—We gratefully acknowledge Dr. Vincent Escuyer for providing the \emph{M. smegmatis}\textsubscript{embC}, Dr. Dean Crick for valuable scientific discussions, Dr. Yufang Ma for initial help with genetics, and Christopher Rithner for performing the NMR experiments. Mass spectrometry analyses were performed at the Core Facilities for Proteomics Research located at the Institute of Biological Chemistry, Academia Sinica, supported by a Taiwan National Science Council Grants 93-3112-B-001-010-Y and 94-3112-B-001-009-Y and the Academia Sinica.

REFERENCES
1. Crick, D. C., Mahapatra, S., and Brennan, P. J. (2001) \emph{Glycobiology} \textbf{11}, 107–118

\textsuperscript{6} S. Khasnobis, J. Zhang, S. Angala, A. Amin, M. McNeil, D. Crick, and D. Chattorjee, unpublished result.
Arabinan Chain Length Extension in LAM