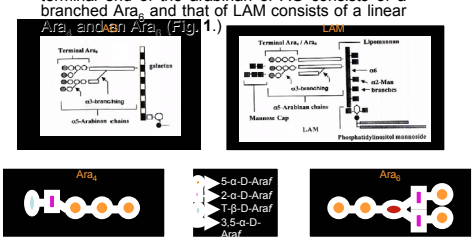
The Emb Proteins in *Mycobacterium smegmatis*

Toni Nagy, Varalakshmi Vissa, Delphi Chatterjee

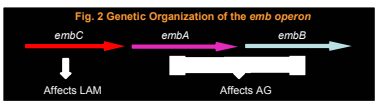
Department of Microbiology, Immunology, and Pathology, Colorado State University

Introduction

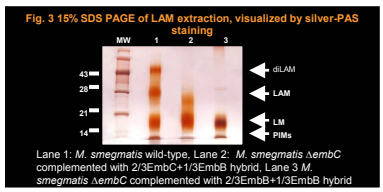
Numerous attempts have been made to develop improved strategies for the prevention and treatment of tuberculosis. The cell wall of *Mycobacterium tuberculosis*, the causative agent of this global disease has been the subject of numerous studies. Mycobacterial D-arabinans are complex molecules, predominantly found in the two major polysaccharides, arabinogalactan (AG) and lipoarabinomannan (LAM). Whereas LAM is a key molecule involved in immunopathogenesis, AG is attached to mycolic acids contributing to the integrity of the cell wall. Structural studies of LAM shows the arabinan attached to a mannan core which extends from a phosphatidylinositol mannoside anchor at the reducing end. The terminal end of the arabinan of AG consists of a branched Ara_6 and that of LAM consists of a linear Ara_4 and an Ara_2. (Fig. 1.)



Although the structure of LAM has been determined, the enhanced details of the assembly of arabinan molecules to form LAM is not known. Previous studies have led to the belief that lipomannan (LM) acts as the substrate in which arabinan is added to produce mature LAM. The Emb proteins encoded by the *embC*, *embA*, and *embB* genes in an operon, have been shown to be involved in the assembly of arabinan to form LAM and AG (Fig 2). Separate knockout mutants of *embA* and *embB* in *Mycobacterium smegmatis* reduced arabinosylation of AG, but had no effect on LAM. However, a knockout mutant of the *embC* gene resulted in cessation of LAM synthesis.



Further genetic and biochemical studies have been done to determine functional regions of the EmbC protein. Computer predictions suggest that EmbC is comprised of 11-13 transmembrane domains corresponding to the first 670 amino acids, and a soluble globular C-terminal domain corresponding to the remaining 430 amino acids. Specifically, gene fusions of *embC* and *embB* including varying numbers of transmembrane domains were generated to determine functional regions. The first hybrid gene involved a fusion of the first 703 amino acids of EmbB with the last 368 amino acids of EmbC. Upon transformation into the Δ*embC*, analysis showed that there was no complementation of the LAM defect. A second hybrid was formed by fusing the first 668 amino acids of EmbC with the last 407 amino acids of EmbB. After transformation into Δ*embC*, LAM was isolated and determined to have a truncated structure (Fig 3). Not only was the



Current work focuses on creating the ½:½ fusion protein which contains approximately 580 amino acids from both EmbC and EmbB. This fusion contains the first eight transmembrane domains of EmbC which will be integral in determining the contribution these domains make in biosynthesis of LAM. Focusing on the N-terminus of EmbC in the generation of hybrids will help establish the catalytic site of the *embC* gene that controls the arabinosylation of LM to give truncated/mature LAM.

Hypothesis

The N-terminal of EmbC recognizes the LAM precursor(LM) and the C-terminal of EmbC acts as the catalytic site for the synthesis of the complex arabinans of LAM.

Research

Gene Fusion

Using the full-length *embC* gene in pVV16 (vector) as the template, a 1.74Kb fragment of the N terminal half of *embC* was amplified by PCR

Fragment was purified on a 0.8% agarose gel and digested with *NdeI* and *BstXI*

Using full length *embB* gene in pBS KSII+ as the template, a 1.5Kb fragment of the C terminal half of *embB* was amplified by PCR

This fragment was purified on a 0.8% agarose gel and then digested with *HindIII* and *BstXI*

To prepare the pVV16 vector for ligation, digestion with *NdeI* and *HindIII* was necessary

A trimolecular ligation occurred with the digested pVV16 vector, the *NdeI-BstXI embC* and *BstXI-HindIII embB* fragments

The ligation products were transformed into DH5α *E. coli* cells and transformants were selected on LB + Kanamycin (50μg/ml) + Hygromycin (100μg/ml) plates

The hybrid plasmid was isolated from several transformant colonies and subsequent DNA sequencing occurred to make certain there were no errors during PCR

Hybrid plasmids were transformed into *Mycobacterium smegmatis* Δ *embC* by electroporation and transformants were selected on 7H11+OADC+Kanamycin (50μg/ml)+Hygromycin (100μg/ml) plates

A colony was chosen and then grown in 7H9+ADC+Kanamycin (50μg/ml)+Hygromycin (100μg/ml) media

Performed LAM, LM, and PIM extraction and isolated LAM using the QuickLAM method

Biochemical analysis

QuickLAM Extraction

1-2 loopfuls of wet bacteria; spin down and discard supernatant

Add $CHCl_3$, MeOH, and H_2O to cell pellet; place in 55C H_2O bath for 30min

Spin down and discard supernatant; add H_2O and H_2O saturated phenol to pellet; vortex and incubate at 80C for 2 hours

Add CHCl3 and centrifuge; discard organic layer and transfer aqueous layer to new tube.

Dialyze aqueous layer against running DI H_2O for 24hrs

Results

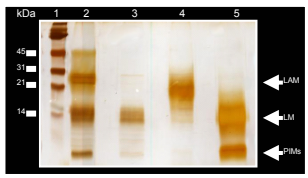


Fig. 5 15% SDS PAGE depicting LAM extraction, visualized using silver-PAS staining. Lane 1 shows a low range molecular weight marker (BIORAD). Lane 2 contains wild-type *Mycobacterium smegmatis*. Lane 3 contains the ½:½ hybrid in Δ*embC M. smegmatis*. Lane 4 contains purified hybrid LAM (2/3*embC* fused with 1/3*embB*). Lane 5 contains Δ*embC M. smegmatis*.

Results continued

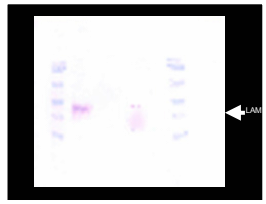
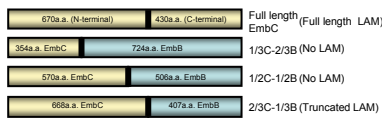


Fig 6. Western Blot using monoclonal antibody CS35 Lane 1: Molecular Weight; Lane 2: WT; Lane 3: ½:½ hybrid; Lane 4: 2/3C-1/3B hybrid; Lane 5: Δ*embC*; Lane 6: Molecular Weight

Summary and Conclusion

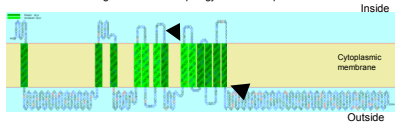
The results of SDS-PAGE and Western Blot show no detectable arabinosylation of LM to form LAM. It is observed that the first eight transmembrane domains cannot complement the LAM defect in Δ*embC*. The 2/3-1/3 hybrid contains 668 amino acids from EmbC, and the ½:½ hybrid contains 570 amino acids. This 98 amino acid difference may be where the mannan recognition site resides, allowing LAM biosynthesis to occur. Future efforts will be made to establish the site with in these 98 amino acids where catalytic activity begins. Also, work is in progress to create a new hybrid containing 7/9*embC*-2/9*embB* (815 amino acids from EmbC; 247 amino acids from EmbB). This construct contains all 13 transmembrane domains of the EmbC protein, so it is expected that full length LAM will be produced when complementing Δ*embC M. smegmatis*.

Fig 7 Summary of yeast and plasmid



Future Work

Fig. 8 Predicted Topology of the Emb proteins



Site of EmbC designated between arrows is the focus of future hybrids to determine the site of catalytic activity where LAM is produced.

- Purification of LM from the ½:½ hybrid using HPLC and comparative analysis to WT LM
- Creation of hybrid focusing on the N-terminal region between amino acid 570 and amino acid 668 to determine where recognition of LM to form LAM occurs
- Work to understand the mechanism of the C-terminal of EmbC, and how this enables the synthesis of the complex arabinans found in LAM.

Reference:

Zhang et al., Mol Microbiol, 2003

Acknowledgments:

- Supported by NIH-NIAID grant AI-37139
- Stefan Berg, Colorado State University
- Libin Shi, Colorado State University
- Vincent Escuyer
  Southern Research Institute, Birmingham, AL