The Physiology of *Mycobacterium leprae*

Patrick J Brennan
Department of Microbiology, Immunology and Pathology, Colorado State University

John S Spencer
Department of Microbiology, Immunology and Pathology, Colorado State University

Perspective

Physiology refers to a branch of biology that deals with the functions and activities of life. In this chapter, we address the physiology of *Mycobacterium leprae* and the physical and chemical phenomena involved. In terms of its structure, biogenesis, and underlying genetics, *M. leprae* is an obligate intracellular parasite incapable of growth in an axenic culture. Consequently, prior to the advent of the armadillo as a source of appreciable amounts, it was available only in minute quantities from human biopsies and mouse footpads. Therefore, an understanding of the physiology of *M. leprae* can be divided into a pre- and a post-armadillo-derived propagation of bacilli. Another major development arose from the pursuit of genomic sequences of members of the genus *Mycobacterium* including *M. leprae*, which provided a valuable data set for comparative studies of the physiology of mycobacteria. In particular, the data set is valuable for unraveling the nature and genetics of biosynthetic pathways, notably that of the unique cell wall of *M. leprae*. Accordingly, our comprehension of the physiology of *M. leprae*, due to the paucity of tangible quantities of the bacterium for classical chemical and enzymatic studies, is based on comparisons—chemical, enzymatic, genetic—with cultivatable mycobacteria, particularly prototypic *Mycobacterium tuberculosis*.

Previous reviews by us, our colleagues, and others cover aspects of the physiology of various mycobacteria (e.g., 1, 2, 3, 4, 5). Two reviews, in particular, are exceedingly helpful. The first is the signature publication by Cole et al. (6) on the definition of the *M. leprae* genome and its annotation. The second, by Vissa and Brennan (7), provides a comprehensive analysis of that

genome in the context of the composition, biosynthesis, and underlying genetics of the cell wall of *M. leprae*. The latter is extensively cited in this review.

Highlights

- It is difficult to study the composition, metabolism, and genetics of *Mycobacterium leprae* due to its inability to grow *in vitro*.
- Breakthroughs have only been accomplished with the availability of large quantities of bacilli grown *in vivo* from armadillos and the nude mouse footpad.
- A comprehensive understanding of the "minimal gene set" of *M. leprae* arose from the complete sequencing and annotation of the genome in 2001, allowing for comparative genomics with *M. tuberculosis* and other *Mycobacterium spp*.

Morphology and Cell Wall Ultrastructure

M. leprae is an acid-fast staining rod, 1–8 µm long and 0.3 µm in diameter, and thus does not differ remarkably from *M. tuberculosis* in that respect. Studies of the ultrastructure of *M. leprae* in sections and as a whole bacterium from man, mouse, and armadillo have been extensive (8), but have not shown any gross unique features compared to other mycobacteria. However, Draper (9) has described three ultrastructural features of the cell wall that may be characteristic of *M. leprae*: aberrant morphology, wall bands, and paracrystalline bodies. With respect to morphology, Draper observed a departure from the classical cylindrical shape of a bacillus in suspensions prepared from armadillos. Specifically, the cells had a tapered or double-tapered shape with hemispherical ends, which he attributed to a defect in the normal process of cell wall construction. Wall bands were first observed by Nishiura et al., who described them as "circumferential ridges on the outer surface of the cell, very numerous and positioned at random along the length of the cell" (10). These ridges may be scars left when the cell wall separated during the division process, and their randomness may reflect a defect in the cell wall construction. The paracrystalline, quasi-crystalline bodies seen in sectioned *M. leprae* probably correspond to the capsular matrices and foamy structures responsible for binding hundreds of bacilli into clumps or "globi" and into smaller clumps where the individual cells occur in parallel arrays, the noted "bundles of cigars" (11). The recent advent of cryoelectron microscopy (cryo-EM) has produced revolutionary images of the cell wall of several cultivable mycobacteria and confirmed that the 'lipid barrier', or the 'outer mycolate layer', is a *bona fide* mycobacterial outer membrane (MOM). Although this exciting development applies to the cell wall of *M. tuberculosis* in particular, we expect likewise for *M. leprae*. Clearly, a pressing research priority is the cryo-EM analysis of the highly purified cell wall/cell envelope of *M. leprae.*

Highlights

- Early ultrastructural studies of *M. leprae* grown *in vivo* showed morphological traits similar to those found in other mycobacteria grown *in vitro*.
- Researchers have described ultrastructural features of the cell wall that may be unique to *M. leprae*: aberrant morphology, wall bands, and paracrystalline bodies.

Cell Wall Composition, Structure, and Biogenesis

Comparison of the 3.27-megabase genome of *M. leprae* with that of *M. tuberculosis* (4.41-megabase) "reveals an extreme case of reductive evolution" by which "less than half of the genome contains functional genes but pseudogenes, with intact counterparts in *M. tuberculosis,* abound" (6). Yet Brennan and Vissa (11), in their analysis of the genome in concordance with the annotation provided by Cole (6), concluded that "cell wall biosynthesis capability remains relatively intact and comparisons with the genome of *M. tuberculosis* provide insights into the genetic basis of a minimal mycobacterial cell wall". Previously, we have also addressed the immunological significance of the cell wall of *M. leprae vis-a-vis* the disease of leprosy (12). Accordingly, comprehending the nature of the cell wall/cell envelope of *M. leprae* is paramount to understanding the physiology of *M. leprae* and its implications for the molecular basis of leprosy pathogenesis. Figure 1 provides a molecular depiction of the cell wall architecture of *M. leprae.*

STRUCTURE OF THE MAGP COMPLEX

Current knowledge on the biosynthesis and genetics of the major components of the mycobacterial cell wall (described below) has evolved primarily from studies on *M. smegmatis*, *M. tuberculosis*, *M. bovis BCG*, and *M. avium*. The studies have used a combination of approaches including structural analysis, metabolic labeling of cultured cells or cell free extracts, isolation and characterization of naturally occurring morphological variants or mutants, and, more recently, genetic manipulation such as mutagenesis of the genome and recombinant gene expression (see review in [2]). However, such approaches are limiting for *M. leprae* research as the organism cannot be cultivated *in vitro* and proven genetic tools are unavailable. Despite these limitations, sufficient information has been gathered on its chemical and structural composition using small amounts of cells obtained from animal or human sources. That information leads to the conclusion that the basic skeleton on which the cell wall of *M. leprae* is built is the covalently linked peptidoglycan-arabinogalactan-mycolic acid complex (MAGP), seen within all mycobacteria, and the related corynebacteria and nocardia, with some modifications (13, 14) (Figure 2).

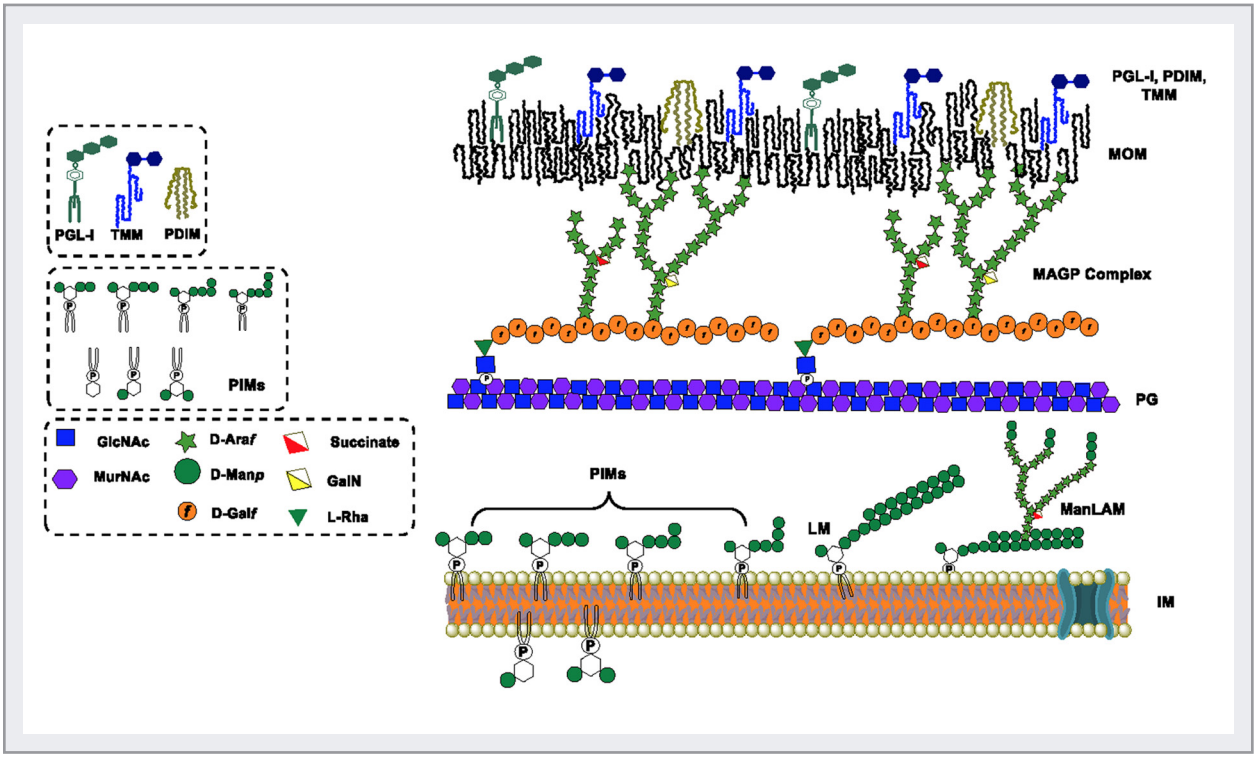

FIG 1 Model of the cell wall/envelope of *Mycobacterium leprae*.

Further elaboration is contained in the text. Abbreviations for molecular structures appearing in the cell wall: Phenolic glycolipid I (PGL-I); trehalose monomycolate (TMM); phthiocerol dimycocerosate (PDIM); phospho-inositol mannosides (PIMs); *N*-acetyl glucosamine (GlcNAc); D-arabinofuranose (D-Ara*f*); succinate; *N*-acetylmuramic acid (MurNAc); D-phospho-mannose (D-Man*p*); galactosamine (GalN); D-galactofuranose (D-Gal*f*); L-rhamnose (L-Rha).

The peptidoglycan (PG) of *M*. *leprae* is characteristic of the chemotype IV group, which includes mycobacteria, corynebacteria, and nocardia, because it contains *meso*-diaminopimelic acid (DAP) in the peptide side-chains (1) (Figure 2). However, distinctive features pertinent to mycobacteria include the presence of *N*-glycolylmuramic acid (MurNGlyc), as part of the repeating GlcNAc-MurNGlyc disaccharide backbone, and direct DAP-DAP cross-links and modifications at the free carboxylic functions of DAP and D-Glu. Apparently at least some of these features extend to the PG of *M*. *leprae.* In a combined liquid chromatography and mass spectrometry (LCMS) analysis of the PG of *M*. *leprae*, we established that, in contrast to earlier reports, the muramic acids of the sugar backbone are exclusively *N*-acetylated, i.e., MurNAc not MurNGlyc (15, 16, 17, 18). The un-cross-linked peptide side-chains of *M. leprae* PG consist of tetra- and tri-peptides containing either glycine or L-alanine. A feature unique to *M. leprae* is the substitution of L-Ala with glycine as the first amino acid of the peptide side-chain of peptidoglycan (19). The effect of this change on the physical properties of the peptidoglycan is not known.

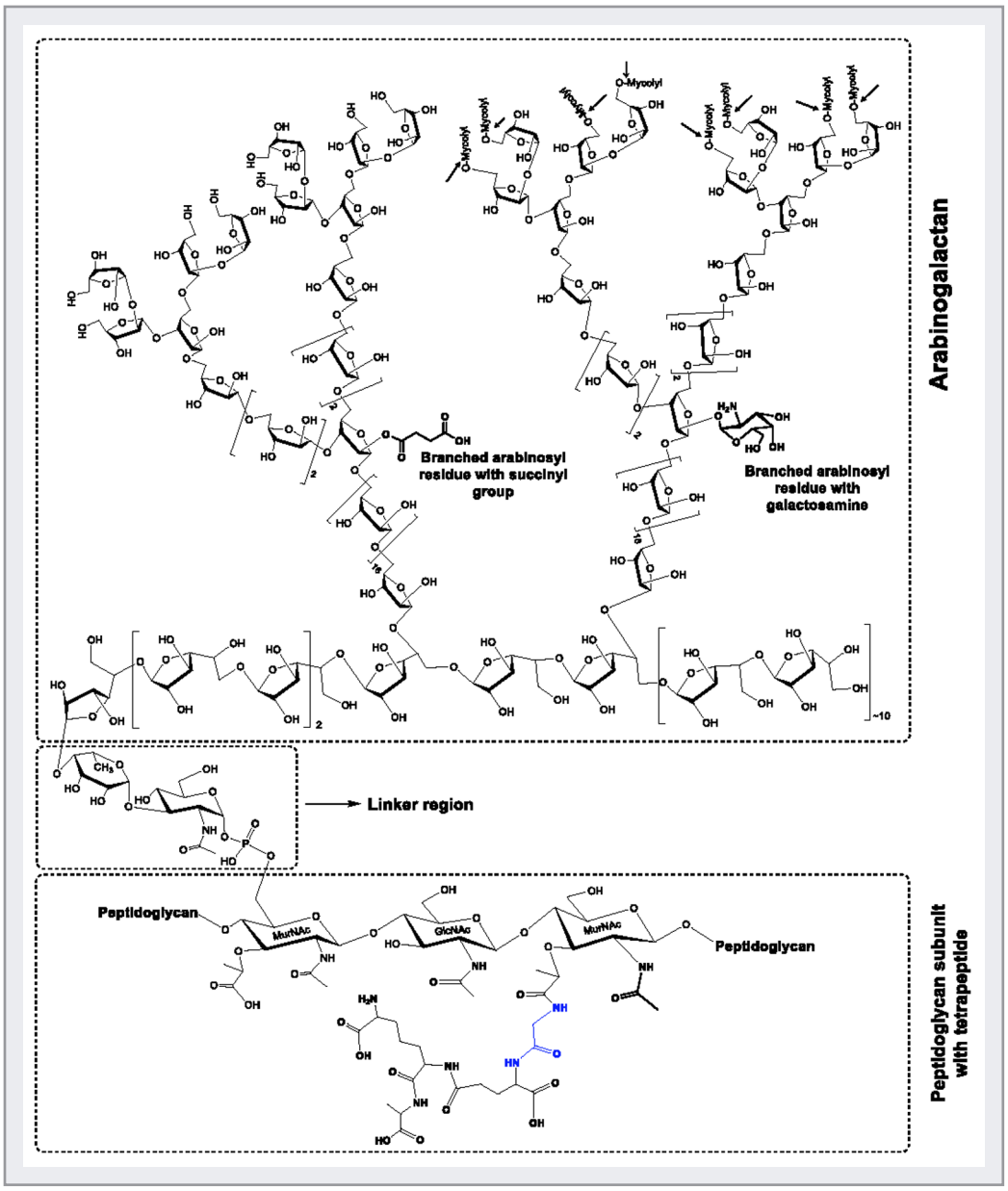

FIG 2 Structure of the peptidoglycan-linker unit-arabinogalactan portion of the mycobacterial cell wall.

This structure applies to *Mycobacterium tuberculosis*. Not all structural facets have been established for the cell wall of *M. leprae*, such as the succinyl and glucosaminyl residues attached to the arabinan. Divergences are discussed in the text.

The arabinogalactan (AG) of mycobacterial cell walls is a polymer of the furanose configurations of galactose (Gal*f*) and arabinose (Ara*f*), a configuration not found in human-derived carbohydrates (20). Typically, a homogalactan (~30 units in *M. tuberculosis*) composed of alternating 5 and 6 linked β-D-Gal*f* residues is linked to PG via a disaccharide-phosphoryl bridge (-L-Rha-D-GlcNAc-P-)

called the linker region/unit (LU) (21) (Figure 2). Two to three branches of 5-linked arabinan are attached near the reducing end of the galactan. The arabinan is composed of 5-linked Ara*f* with further internal branching (3- and 5-linked α-D-Ara*f*). The non-reducing ends are composed of the pentaarabinofuranosyl (Ara_5) motif [β-D-Ara*f*-(1,2)-α-D-Ara*f*]$_2$-3,5-α-D-Ara*f*-(1,5)-α-D-Ara*f*] (22). All of the major structural motifs established for the *M. tuberculosis* cell wall also pertain to *M. leprae*, with the exception that the AG of *M. leprae* is smaller and has 40–50% fewer Gal*f* residues (14, 23).

Two-thirds of the terminal arabinoses of the pentaarabinofuranosyl (Ara_5) motif within the arabinan chains are esterified with mycolic acids in *M. tuberculosis* (24) (Figure 2); however, the extent of such substitution in the case of the *M. leprae* arabinan has not been determined. Yet the nature of the mycolic acids of *M. leprae* cell walls has been extensively examined. In general, mycobacteria contain alpha mycolates, which lack oxygen functions, and the oxygenated keto- and methoxy-forms. The α-alkyl branch of the α-alkyl, β-hydroxy mycolic acids generally range in length from C_{14} to C_{26}; however, in *M. leprae* the chain length is about C_{20} (25). The main meromycolate chain is often modified with double bonds (cis and trans), cyclopropane, methyl, epoxy, keto, and methoxy groups that render flexibility (fluidity) to the cell wall. *M. leprae* does not have methoxymycolates (26, 27) due to the lack of a functional *mmaA3* gene (6), a feature shared with multi-passaged *M. bovis* BCG Pasteur (28). We have previously speculated on the implications of the absence of meromycolates (11).

We have also recently addressed the spatial organization of the known, chemically defined components of the cell wall core (the mycolyl-arabinogalactan-peptidoglycan [MAGP] complex) of some *Mycobacterium spp.* including *M. leprae* (23). We established that the cell wall of *M. leprae* contains significantly more mycolic acids attached to the AGP complex than does that of *M. tuberculosis* (i.e., mycolate to peptidoglycan [PG] ratios of 21:10 versus 16:10, respectively). In keeping with that characteristic, more of the arabinogalactan (AG), which lies between the mycolates and PG, exists in *M. leprae*. Also, the AG of *M. leprae* is smaller than that of *M. tuberculosis*, although those basic structural motifs that we established in 1990 do pertain (22).

Highlights

- The cell wall of *M. leprae* is built upon the covalently linked peptidoglycan-linker unit-arabinogalactan-mycolic acid complex (MAGP) evident in all mycobacteria, with some modifications.
- The peptidoglycan (PG) and arabinogalactan (AG) of the *M. leprae* cell wall show distinctive features compared to that of *M. tuberculosis*. The significance of these in the life/infectious cycle of *M. leprae* is a matter of conjecture.
- Methoxymycolates are absent in *M. leprae* due to the lack of a functional *mmaA3* gene. Again, the significance is not known.

Highlights (cont'd)

- Despite a severely compromised genome, of which less than half contains functional genes, cell wall biosynthesis capability remains relatively intact (see Chapter 8.2).

BIOSYNTHESIS AND GENETICS

Vissa and Brennan (7, 11) have written extensively on the biosynthesis and genetics of the cell wall of *M. leprae* based on their extensive fundamental research. The following discussion is extracted from their work. The biosynthetic pathway of the individual components and their assembly to form the MAGP complex, as deciphered from the approaches mentioned above, is depicted in Figure 3. We have previously identified most of the genes responsible for the enzymes catalyzing those anabolic reactions. Some were functionally characterized by the "cloning by homology approach" (7, 11), in which genes were identified by finding homologs for known enzymes in mycobacteria or other genomes. The homologs were then cloned, the proteins expressed, and their functions confirmed. Others are putative genes for reactions not yet biochemically characterized in mycobacteria.

Peptidoglycan

The classical peptidoglycan biosynthetic machinery was established for *E. coli* and appears to be remarkably well preserved in mycobacteria, including *M. leprae* (2, 3, 4, 29). The main events are the synthesis of the usual sugar-nucleotide, UDP-MurNAc, from UDP-GlcNAc (*murA*, *murB*); the sequential addition of amino acids to the MurNAc (*murC*, *murD*, *murE* and *murF*); the transfer of the MurNAc-pentapeptide to a polyprenyl-phosphate carrier (*murX*); and the attachment of GlcNAc from UDP-GlcNAc (*murG*). Glycan chains are formed by transglycosylation reactions of the disaccharide-pentapeptide chains (releasing the lipid carrier from the incoming unit), followed by the cross-linking of such chains via DAP-DAP and DAP-D-Ala bridges in the final stage of peptidoglycan synthesis. The enzymes for transglycosylation and transpeptidation are members of the penicillin-binding protein (PBP) family. Two *M. leprae* genes (*ponA* and *ponL*) have been cloned and expressed but their role in peptidoglycan synthesis has not been reported (30, 31). The *M. leprae* genome also contains several pseudogenes with homology to PBPs. In mycobacteria, amidation of the carboxyl side chains of glutamate and DAP occur at an undetermined stage in the pathway. The oxidation of the N-acetyl group of MurNAc to a glycolyl moiety in *M. tuberculosis* and other mycobacteria does not occur in *M. leprae* (rather, the N-acetyl function of MurNAc is retained). Finally, the *M. leprae* ortholog of *namH* (ML0085c) has been reported to be a possible pseudogene (7, 11).

As pointed out, previously many of the *M. leprae* peptidoglycan side chains have glycine rather than L-alanine, a substitution believed to have a genetic basis. However, Mahapatra et al. (17) analyzed the genome but could not find an appropriate ligase gene. In fact, they demonstrated by

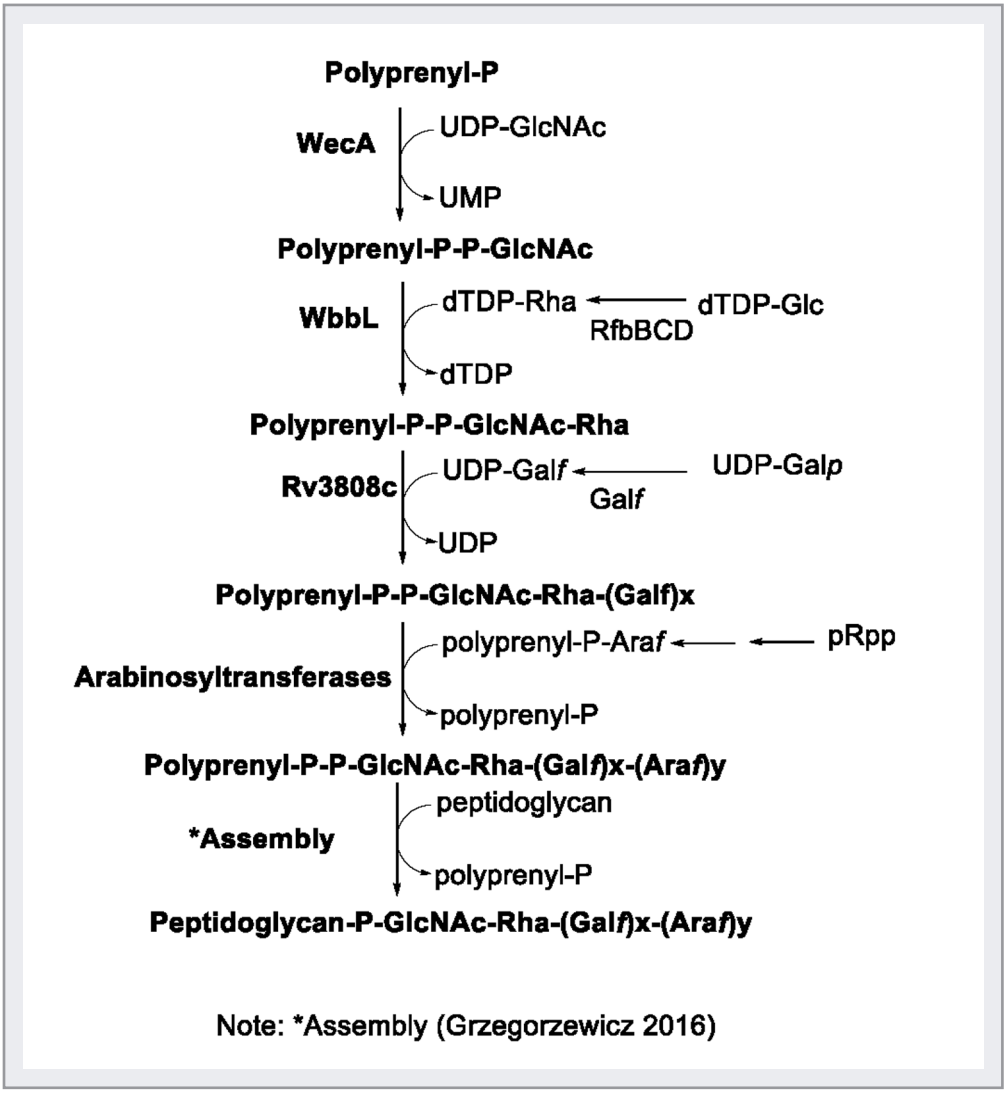

FIG 3 Mycobacterial arabinogalactan biosynthetic pathway.

The naming of the individual enzymes responsible for each step are in accord with the cognate gene identified in the genome of *M. tuberculosis*. The corresponding *M. leprae* genes/enzymes are identified in references (7, 11) and in the text. Final assembly of the two major heteropolysaccharides of the mycobacterial cell wall, arabinogalactan (AG) and peptidoglycan (PG), and their covalent attachment by specific ligases was recently elucidated (32).

in vitro studies with the recombinant MurC enzyme that it can use either the L-alanine or glycine substrate with comparable affinities. Therefore, we have speculated that the presence of glycine in *M. leprae* is perhaps an *in vivo* phenomenon driven by the ambient amino acid milieu.

Arabinogalactan

The synthesis of AG requires the sugar donors UDP-GlcNAc (33); TDP-rhamnose (*rmlA, rmlB*, *rmlC*, and *rmlD*) (33, 34); UDP-galactofuranose (UDP-Gal*f*) (*galE* and *galF*) (35); and decaprenyl-phospho-arabinose (DPA) (36) (Figure 3). The additions of GlcNAc (*rfe*), rhamnose (*wbbl*), Gal*f*, and Ara*f* donated by their respective sugar nucleotide, or DPA in the case of the Ara*f* units, are catalyzed by a series of sugar transferases. Regarding the galactosyltransferases, Mikušová et al. (37, 38) has shown that the gene Rv3808c (*glfT*) of *M*. *tuberculosis* is a galactosyltransferase in AG synthesis. There is an ortholog in *M. leprae*. Specifically, Kremer et al. (39) have suggested that this gene encodes a bi-functional transferase for the alternating 5 and 6 linked galactose residues of the galactan of the cell wall MAGP complex. Whether this enzyme (GlfT) can also catalyze the addition of the first galactose unit to the rhamnose sugar of the linker unit and the second galactose of the galactan has not been determined. GlfT contains the sugar nucleotide binding motif hhhhDxDxh, where 'h' represents an amino acid with a hydrophobic nature. With regard to the arabinosyltransferases responsible for the synthesis of the arabinan of MAGP, the work of Belanger et al. (40) suggests that the *embA* or *embB* genes of *M. avium* encode putative transferases for AG. A third gene, *embC,* also exists in all mycobacteria sequenced thus far, and Escuyer et al. (41) have shown that knocking out the *embA or embB* genes of *M. smegmatis* causes changes in the arabinan content and structure of AG. As pointed out, these Emb proteins are very homologous to each other; they are large and hydrophobic, and span membranes. The genes are well conserved among many mycobacteria and are intact in *M. leprae* in a gene cluster very similar to that seen in *M. tuberculosis.* However, the proteins they encode have no significant homology to any other proteins in the database and no known domains or motifs. It is not clear if the Emb proteins are involved in the actual glycosyltransferase catalysis or in the assembly of the arabinan (7). Such arabinoslytransferases have indeed been identified in work from our laboratory (42) and that of Besra (43). The stepwise synthesis of the 'linker unit' (LU) of -phospho-GlcNAc-Rha-(Gal*f*)$_n$-(Ara*f*)$_n$ occurs on the back of the decaprenyl-phosphate (DP) lipid carrier (37) before the entire -P-GlcNAc-Rha-(Gal*f*)$_n$-(Ara*f*)$_n$ unit is transferred (ligated) to approximately 1 in 10 of the MurNAc units of peptidoglycan. Evidence on the nature of this final transfer of the -P-GlcNAc-Rha-(Gal*f*)$_n$-(Ara*f*)$_n$ unit from its decaprenyl-P lipid carrier to PG has been published by our colleagues (32). Although the details of these intriguing and essential first steps in the synthesis of the core, the MAGP complex, of the mycobacterial cell wall have been illustrated with cultivable mycobacteria, mostly *M. tuberculosis,* the process must also apply to *M. leprae* since that essential core applies to all mycobacteria. A significant unknown at this point concerns the stage at which the mycolates are transferred to the -P-GlcNAc-Rha-(Gal*f*)$_n$-(Ara*f*)$_n$ unit. Specifically, the question is whether the mycolates are transferred before or after the ligation of the MAGP complex to PG.

Polyprenyl-phosphate lipid carriers including decaprenyl-phospo-arabinose (DPA)

The biosynthesis of many of the key cell wall polymers requires prenyl-phosphate as the means by which cell wall products originating in the cytoplasm are transported through the cytoplasmic/inner membrane to the cell wall. Thus, decaprenyl-P (DP) in the form of DP-mannose (DPM) is the

immediate donor of some of the mannoses for the synthesis of the mannan within LM and LAM. (All mannose residues, including those in DPM, originate in GDP-mannose, which can also be the direct donor of polymerized mannose residues.) DPA is the source of all arabinosyl residues in LAM and AG. DP is also the fulcrum on which the P-GlcNAc-Rha-(Gal*f*)$_n$-(Ara*f*)$_n$ unit is assembled (Figure 3) and the means by which it is transported through the membrane and introduced to PG for final cell wall core assembly (29, 32). Synthesis of DP has been described in detail (11) as follows:

> The synthesis begins with the formation of deoxyxylulose phosphate from pyruvate and glyceraldehyde-3P (*dxs-I*), which is converted to the 5-carbon isoprene compounds isopentenyl pyrophosphate (IPP) and its isomer dimethylallyl pyrophosphate (DMAPP) by the non-mevalonate pathway, also called the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (44). In *M. tuberculosis* there are two possible genes for this function (*dxs-I*, *dxs-II*). DXS-I has been shown to be functional by expression in *E. coli* (45). *dxs-II* may be redundant in *M. tuberculosis,* since the *M. leprae* genome has only one gene that is homologous to *dxs-I*. Homologous genes responsible for the complete essential synthetic pathway of IPP and DMAPP in *E. coli* (46, 47) have been found in *M. tuberculosis* and *M. leprae*. A non-essential IPP isomerase (*idi*) for the inter-conversion of IPP and DMAPP is present in *E. coli* (48) and *M. tuberculosis* but not in *M*. *leprae*. The gene responsible for condensation of IPP with DMAP to form geranyl diphosphate (GPP, C_{10}-PP) has not been identified. The addition of IPP to GPP results in farnesyl diphosphate (FPP, C_{15}-PP), which is subsequently elongated by seven cycles of polymerization to form decaprenyl diphosphate (DP) in a specific stereochemical configuration in *M*. *tuberculosis* (49) and also probably in *M. leprae* because homologous polymerase genes exist. In terms of synthesis of other isoprenoids, there are no homologs in *M. leprae* for four other prenyldiphosphate synthase genes found in the *M*. *tuberculosis* genome, except for the *grcC1* gene, which may be involved in the transfer of a prenyl moiety in the menaquinone synthetic pathway (11). There are also no homologs in *M*. *leprae* for the *M. tuberculosis* squalene synthase, monooxygenase, and cyclase genes probably involved in steroid synthesis.

The biosynthesis of DPA (Figure 4) itself has been defined in our laboratory and has proven to be a target for a new generation of putative anti-TB and anti-leprosy drugs. DPA originates in the pentose phosphate pathway as phospho-ribosyl pyrophosphate (PRPP) (50). Epimerization of the ribose to arabinose occurs after the transfer of the phosphor-ribose to a decaprenyl-phosphate carrier. The resulting epimerase (DprE1) is a flavoenzyme and a highly vulnerable drug target (51, 52). Many compounds with diverse scaffolds were found to act as covalent, such as the benzothiazinones and dinitrobenzamides (53), or non-covalent inhibitors in the replication of *M. tuberculosis* and are now being tested for anti-*M. leprae* activity.

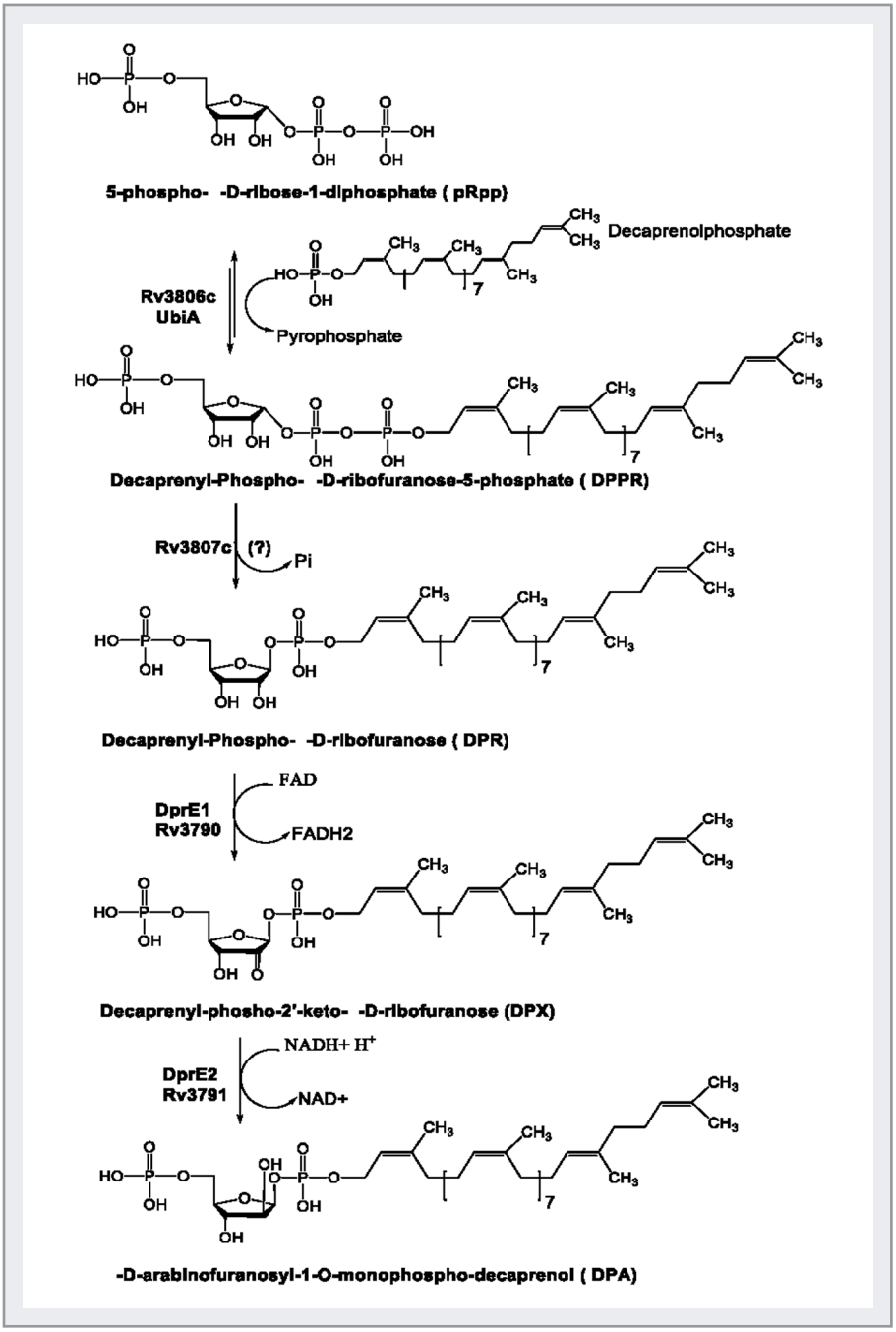

FIG 4 Mycobacterial biosynthetic pathway leading to the generation of decaprenylphosphoryl-d-arabinose, the universal source of cell wall arabinose units.

The naming of the individual enzymes responsible for each step are in accord with the cognate gene identified in the genome of *M. tuberculosis*. The corresponding *M. leprae* genes/enzymes are identified in references (7, 11) and in the text.

Biosynthesis of the mycolic acids and deposition on AG

We have previously reviewed the biosynthesis of the mycolic acids and deposition on AG (11) as follows:

> The synthesis of mycolates occurs in several stages: synthesis of the α-alkyl chain and the primer for the meromycolic acid by the multifunctional fatty acyl synthase FasI enzyme (*fas*) (54); linking by the β-keto acyl synthase III (*fabH*) (55) and elongation of the primer (25) by the disassociated fatty acyl synthase complex FasII (*accD6*, *fabD*, *kasA*, *kasB*, *acpM*, *mabA*, *inhA*) (56, 57); modification of the meromycolic acid (introduction of double bonds, cyclopropane rings, keto, methyl and methoxy groups), probably in parallel with elongation (58); and finally condensation of the α-alkyl chain and meromycolate. The condensation step and the carrier on which this reaction occurs are not known with certitude. The mycolates are then transferred to AG and trehalose to form MAGP and TMM/TDM respectively and may occur using the mycolyltransferases encoded by the members of the antigen 85 complex (*fbpA*, *fbpB* and *fbpC2*) (59, 60). However, Goren et al. (61) could find only TMM in *M. leprae*, indicating that TDM was absent. All of these steps have been characterized in *M. tuberculosis* and homologs for the genes are present in *M. leprae*, with the exception of the gene encoding the linking enzyme FabH; however, the mechanism of adding mycolates to AG in *M. leprae* is not clear.

Highlights

- The classical peptidoglycan biosynthetic machinery established for *E. coli* appears to be remarkably well preserved in mycobacteria, including *M. leprae*.
- Genome annotation and comparisons with similar anabolic pathways in *M. tuberculosis* indicates that construction of the linker unit (LU; P-Glc*N*Ac-Rha) and attached galactan occurs on the lipidic decaprenyl(C_{50})-phosphoryl unit and requires the appropriate sugar donors catalyzed by the corresponding sugar transferases: UDP-N-acetylglucosylaminyl (Glc*N*Ac)-, TDP-rhamnosyl(Rha)-, and UDP-galactosyl(Gal)-transferases. These events presumably take place in the interphase of the cytoplasm and cytolic membrane of *M. leprae*.
- The arabinofuranose (Ara*f*) units of the arabinan polymer arise in an off-shoot of the pentose phosphate cycle and are donated to the pre-formed C_{50}-P-LU-galactan *via* decaprenyl (C_{50})-P-arabinofuranose and a series of arabinosyl-transferases towards full arabinan formation. These events occur in the cytoplasmic membrane or periplasmic region of *M. leprae*; hence the need for lipid (C_{50})-P-based templates.
- Decaprenyl-P-Man (mannose) and GDP-Man are the donors of Man units required in conjunction with the corresponding mannosyl transferases for the synthesis of the mannan within lipomannan (LM) and lipoarabinomannan (LAM) of *M. leprae*.

Highlights (cont'd)

- *De novo* synthesis of the mycolic acids of mycobacteria is now well know. The mechanism for the addition of the mycolic acids of *M. leprae* to the terminal segments of the arabinan of the synthesized arabinan-galactan-LU complex is not known. We have previously reviewed the topic of mycolic acid biosynthesis as applied to *M. leprae* (11).

FREE, "EXTRACTABLE" LIPIDS AND LIPOGLYCANS

TMM, PI, PIMs, LM, and LAM

In addition to the mycolates esterified to the terminal arabinan of the core MAGP complex of the *M. leprae* cell wall, mycolic acids are also present in the extractable lipids as esters of trehalose: 6-mycolyl and 6,6-*O*-dimycolyl trehalose (TMM and TDM, respectively). Small amounts of TMM but not TDM have been identified in *M. leprae* (61). It appears that the other solvent extractable lipids characteristic of other species of the *Mycobacterium* genus are absent in *M. leprae*. Those other lipids include the glycopeptidolipids (GPLs) typical of the *M. avium* complex that define serotype/serovar specificity and colony morphology (62); trehalose based lipids such as acylated trehaloses (containing straight chain, mycerocerosic acids, mycolipanolic and mycolipenic fatty acids); and sulpholipids (trehalose 2 sulphate acylated with phthioceranic and hydroxyphthioceranic acids) present in strains of *M. tuberculosis*; and lipooligosaccharides (LOSs) isolated from several non-tuberculous mycobacteria (NTM) species.

The cell wall of *M. leprae* is also endowed with an unusually abundant proportion of the extractable intercalated lipoglycans, phosphatidylinositol mannosides (PIMs), lipomannans (LM), and lipoarabinomannans (LAM) (Figure 1) characterized in several mycobacteria. These products may be anchored in the plasma membrane and the outer mycolate membrane (OMM) via the phosphatidylinositol (PI) anchor, specifically, via their acyl chains, tuberculostearic and palmitic acid. This PI anchor is also found in the PIMs, LM, and LAM (63). LM and LAM are made up of linear α-1→6 linked mannan chains originating from PIM_2; most of those linear mannoses are further branched with α-1→2 D-mannoses (64) (Figure 1; Figure 5). LAM is a heterogeneous macromolecule arising from LM that contains arabinan branches similar in composition and structure to those of AG. In LAM, the branches may be terminated with "caps" of variable numbers of mannose or inositol or other residues (64, 65). The LAM of *M. leprae* has fewer of the terminal Ara_5 units and a lesser degree of mannose capping than that of *M. tuberculosis* (66).

Current interest in LAM stems largely from its role as a major molecular modulator of the immunopathogenesis of TB (67), yet our initial 'discovery' of LAM arose from an investigation of the major antigens of *M. leprae* (63). LAM had been known for years as a lipophilic soluble arabinose- and mannose-containing heteropolysaccharide (63); our significant contribution was

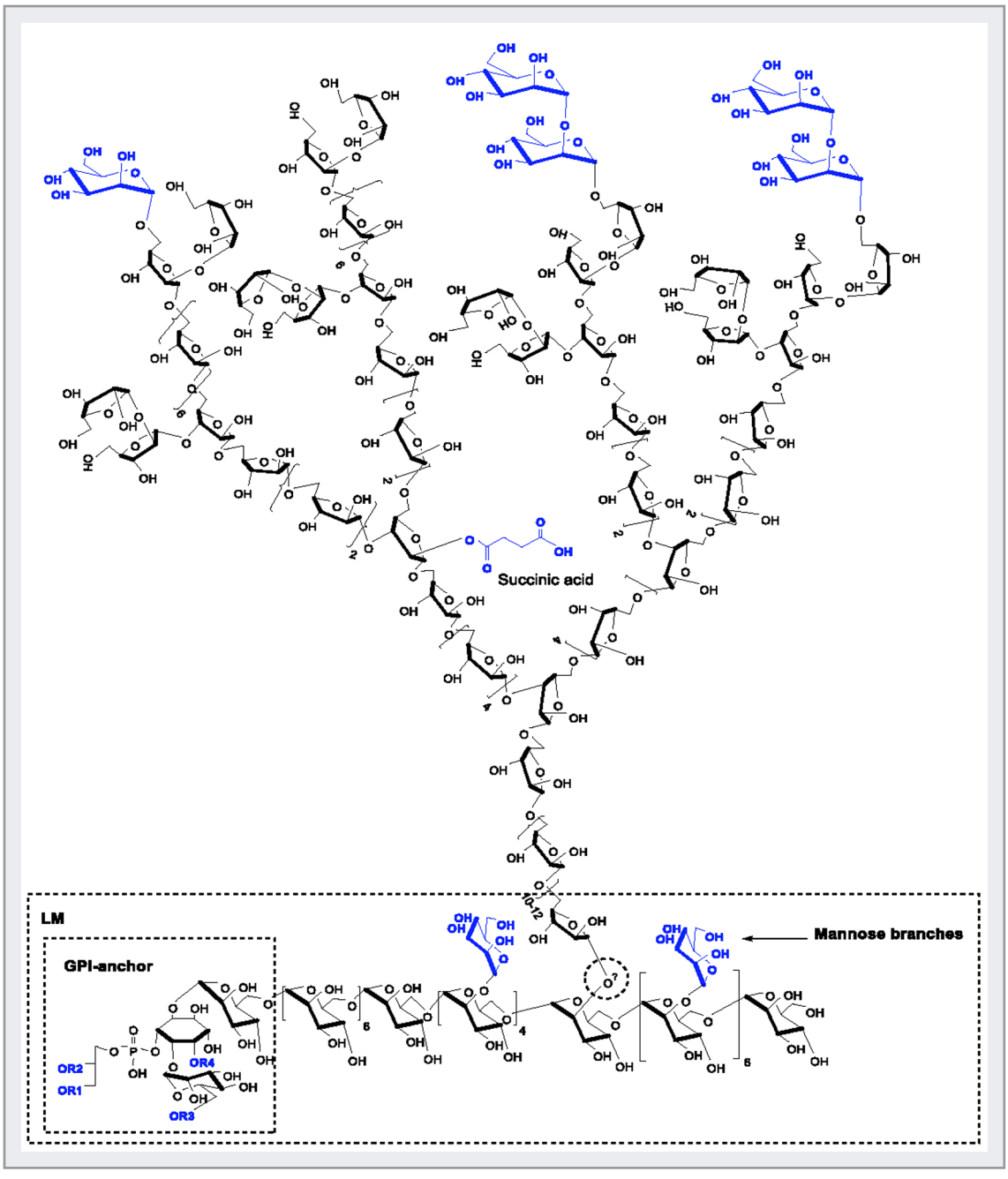

FIG 5 Generic structure of mycobacterial lipoarabinomannan.

Based on our research on the product from *M. leprae*, all of these structural features, thoroughly explored in *M. tuberculosis* and *Mycobacterium smegmatis,* do apply. It should be realized that LAM from all mycobacterial sources is a heterogeneous mixture; thus, a general version of the structure is presented here.

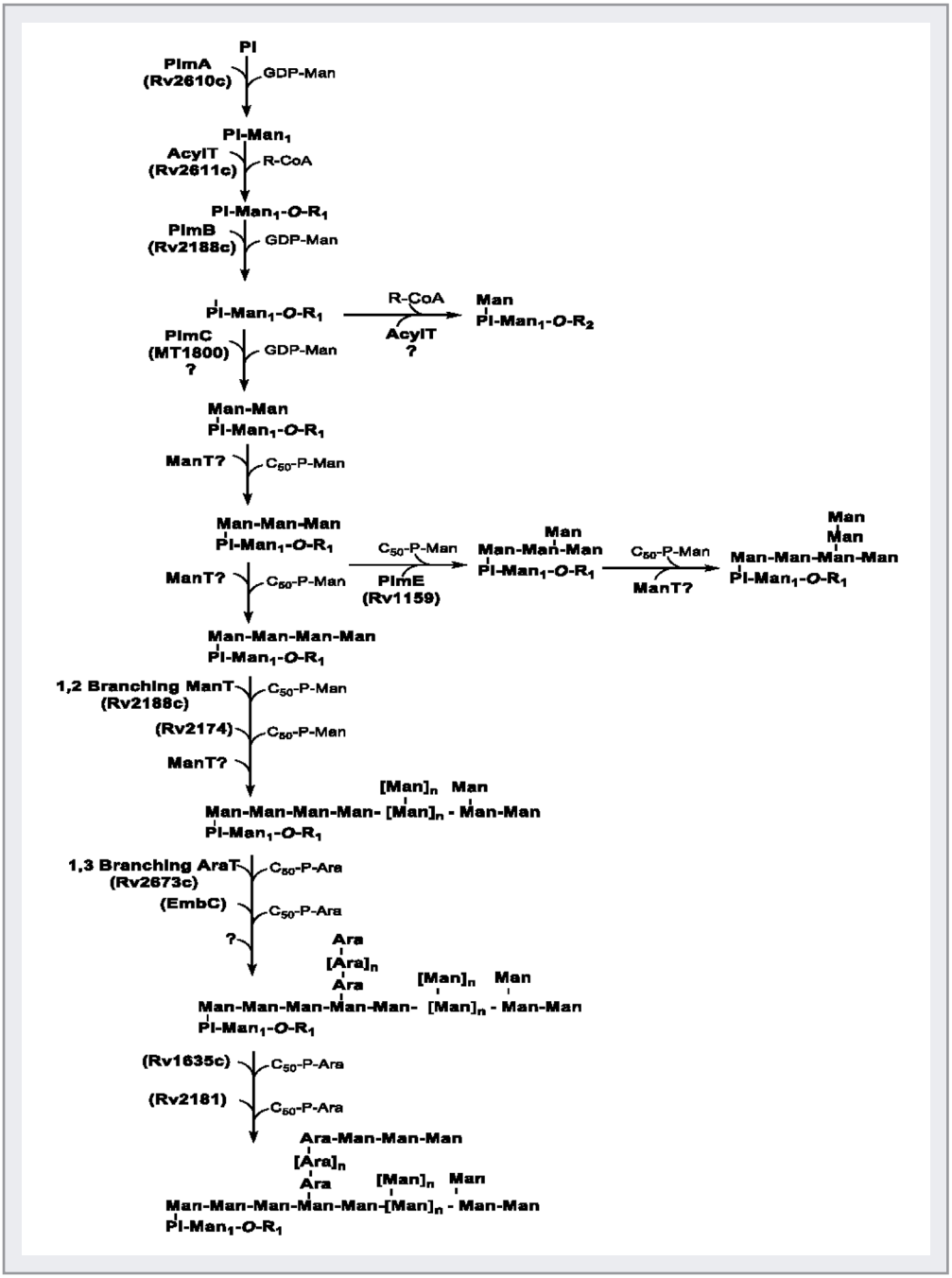

FIG 6 Mycobacterial lipoarabinomannan biosynthetic pathway.

The naming of the individual enzymes responsible for each step are in accord with the cognate gene identified in the genome of *M. tuberculosis*. The corresponding *M. leprae* genes/enzymes are partially identified in references (7, 11) and in the text.

discovering that it is structurally anchored on a phosphatidyl mannoside (PIM) core and thus is an extension of PIM_2 and LM (63). The biosynthesis of mannose-capped LAM (ManLAM) is initiated on the precursor PI using the gene product of *pgsA55,* followed by mannosylation using primarily GDP-mannose as the sugar donor but also DPM (68) (Figure 6).

Genes for a mannosyltransferase and an acyltransferase are linked to *pgsA* and may be required for the synthesis of PIM_1 and the acylation of mannose to form triacyl PIM_1, respectively. The gene product of Rv0557 (*pimB*) of *M. tuberculosis* has been identified as the second mannosyltransferase (69). Recombinant PimB has been shown to convert labeled tri-acylated PIM_1 to PIM_2 in the presence of GDP-mannose. PimB belongs to the family of glycosyltransferases that have a conserved C terminal motif EX7E (70). This gene is a pseudogene in *M. leprae* and we predict that one of the other genes sharing a homologous active site is used. Subsequent mannosylation with undefined enzymes generates LM. Of the many putative glycosyltransferases in the *M. tuberculosis* genome, several (Rv0539, Rv0696, Rv1781c, Rv1500, Rv1513, Rv1514c, Rv1516c, Rv1518, Rv1520, and Rv1525) do not have orthologs in *M. leprae*. We propose that these enzymes are not involved in the biosynthesis of mannan, arabinan, and galactan, since there are no significant structural differences between these molecules in *M. tuberculosis* and *M. leprae* (71, 72).

Phenolic Glycolipid I-III; discovery, definition, roles, and biosynthesis of PGL-I

The dominant lipid is the unique and *M. leprae*-specific phenolic glycolipid PGL-I (73), a glycosylated derivative of the phenolphthiocerol dimycocerosate. We have previously reflected on the discovery of PGL-I, II, and III; their structural definitions; and their facility for the selective serodiagnosis of leprosy (74, 75, 76). The uniqueness and *M. leprae*-specificity of PGL-I is in its trisaccharide entity, 3,6-di-*O*-methyl-β-D-Glu-(1→4)-2,3-di-*O*-methyl-α-L-Rha-(1→2)-3-*O*-methyl-α-L-Rha (Figure 7), rendering PGL-I as highly and specifically antigenic for leprosy (77, 78). Synthetic glycoconjugates containing this trisaccharide or the terminal disaccharide are sensitive tools for the serodiagnosis of leprosy (79, 80, 81). More recently, PGL-I and its trisaccharide have been implicated in the induction of the peripheral neuropathy characteristic of leprosy (see Chapter 2.5; Chapter 9.2). Specifically, this debilitating morbidity has been attributed to axon demyelination due to direct interaction with the laminin of myelinating glia and subsequent infection (82, 83, 84, 85). There is now evidence that demyelination and axonal damage are not directly initiated by PGL-I/*M. leprae* but by "infected macrophages that patrol axons. PGL-I induces nitric oxide synthase in infected macrophages and the resulting increase in reactive oxygen species damages axons by injuring their mitochondria and inducing demyelination" (86).

As discussed by Brennan and Vissa (11),

> The phthiocerol moiety of the PGLs is synthesized using a set of multifunctional enzymes (*ppsA*, *ppsB*, *ppsC*, *ppsD*, and *ppsE*) that contain one or more of the acyltransferase, ketoacyl synthase, keto reductase, dehydratase, enoyl reductase, and acyl carrier modules for the polymerization of malonyl-CoA and methylmalonyl-CoA

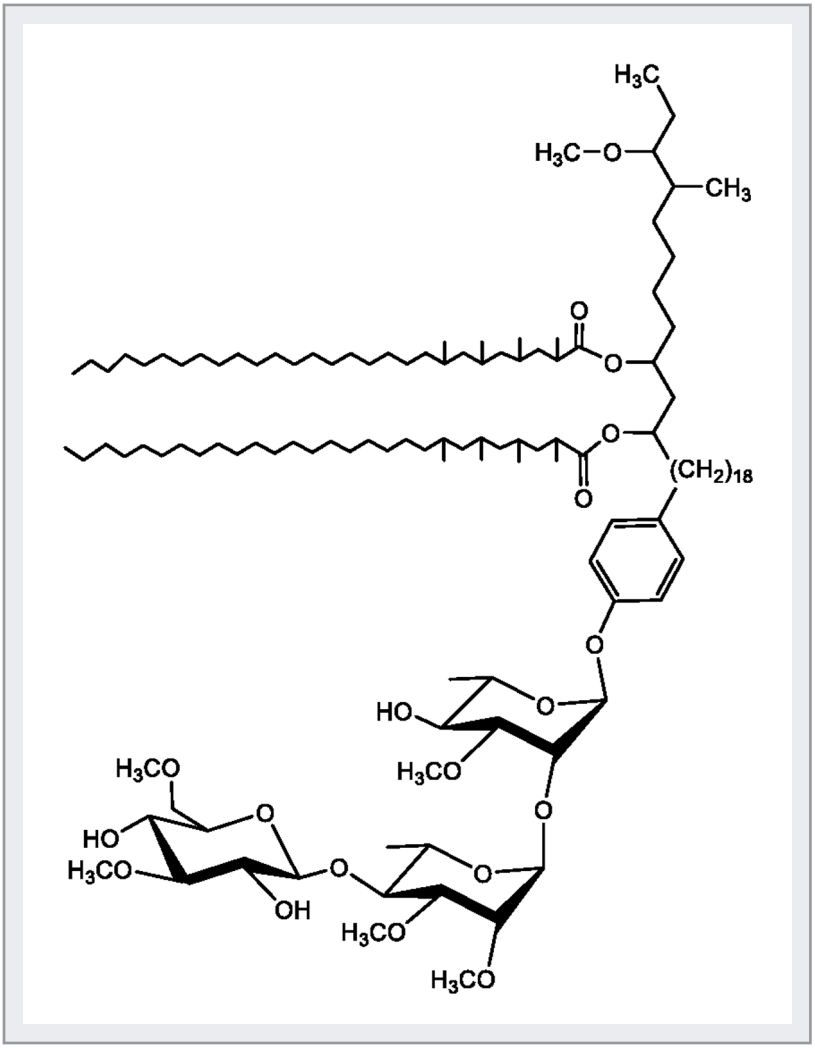

FIG 7 Structure of the species-specific phenolic glycolipid of *M. leprae*.

***M. leprae* specificity lies in the triglycosyl portion of the molecule, the 3,6-di-*O*-methyl-glucopyranosyl-(β1→4)-2,3-di-*O*-methyl-rhamnopyranosyl-(α1→2)-3-*O*-methyl-rhamnopyranosyl entity. The phenolic-phthiocerol-dimycocerosic lipid segment is a common feature of the cell walls of many mycobacteria.**

units on a C_{22}-CoA fatty acid precursor (87). Mycocerosic acid synthesis occurs by the elongation of fatty acyl-CoA primers with methylmalonyl-CoA (*mas*) (88), followed by the transfer to the phthiocerol using a specific acyl-CoA synthase (*fadD28*). Two membrane-associated proteins, MmpL7 and DrrC, have been shown to be responsible for the transport of the PDIM. In *M. tuberculosis,* genes for all of these functions are clustered on the genome, and mutations in these genes result in the disruption of PDIM synthesis and loss of virulence (89). In *M. leprae*, the *ppsA-E* genes are intact but have been separated from the *mas, fadD28*, and *mmpL7* genes (11). For the addition of the first rhamnosyl unit (attached to the

phenolic group) of the trisaccharide of PGL-I in *M. leprae*, we (11) have analyzed the genome for genes homologous to the rhamnosyltransferases. These genes include the *rtfA* of *M. avium*, which is responsible for the addition of the L-rhamnose to the 6-deoxy-talose in the case of its glycopeptidolipid (GPL) synthesis (90), and the *wbbL* gene of *M. tuberculosis*, which is involved in the synthesis of the 'linker unit' of the cell wall core (see above). Based on homology searches with *rtfA*, we have identified the gene *ml2348* as a candidate (11). Coincidentally, *ml2348* is located where the phthiocerol gene cluster of *pps/drr/papA5* has separated from the *mas/fadD28/mmpL7* of *M. leprae* and may indicate the gene cluster's role in PGL-I synthesis. *ml2348* is also homologous to genes responsible for the synthesis of glycosylated steroids in plants and antibiotics such as Balhiymicin and Tylosis (*tylN*, 6-deoxyallosyl transferase) in microbes (91). All of these homologs use sterol/phenol-like acceptors. As pointed out previously (7), *rtfA* and *tylN* encode glycosyltransferases responsible for 6-deoxyhexoses. Thus, we have proposed that ML2348 is likely the enzyme responsible for the addition of that first rhamnose unit to the phenol substituent of PGL-I, that ML0125 and ML0128 are the two other glycosyltransferases required for the synthesis of the complete trisaccharide appendage, and that ML0127 and ML0130 are the methylases responsible for the formation of the O-CH_3 appendages on the trisaccharide unit. These genes are located close to the mycocerosic acid gene cluster *mas/fadD28/mmpL7*. Genes that are highly homologous to *ml2348* are also present in *M. tuberculosis* (*rv1524* and *rv1526c*). These genes are clustered with another *pks* system, *pks5*. There are also homologs in *M. tuberculosis* for the candidate glycosyltransferases, Rv2958c and Rv2962c, and the methyltransferases, Rv2952 and Rv2959c, in a cluster similar to that in *M. leprae*. These homologs are likely responsible for the synthesis of the particular phenolic glycolipid found in select rare strains of *M. tuberculosis* such as *M. tuberculosis canetti* (11) but apparently occur in more common isolates.

Highlights

- The dominant lipid of *M. leprae* is the unique and species specific phenolic glycolipid I (PGL-I; some metabolic by-products PGL-II and -III also exist). PGL-I has proven to be a useful antigen for the specific serodiagnosis of leprosy. It has also been shown to indirectly damage the mitochondria of axons leading to demyelination of nerves, the principle cause of nerve damage in leprosy.
- Other lipids characteristic of NTMs (non-tuberculous mycobacteria) are absent from *M. leprae*, as they are from the *M. tuberculosis* complex. These lipids include the glycopeptidolipids (GPLs) that define serovar type and rough and smooth colony morphology in the *M. avium* complex, the trehalose-based LOS (lipooligosaccharides), and also, apparently, the sulpholipids and 'cord factor' (trehalose dimycolate).

Highlights (cont'd)

- *M. leprae* has an abundance of phosphatidylinositol mannosides (PIMs), lipomannan (LM), and lipoarabinomannan (LAM) as well as the conventional membrane-associated glycerophospholipids.

Identification of Antigens Using Monoclonal Antibodies

The key to understanding almost every aspect of the biology of *M. leprae* was the ability to grow the bacterium *in vivo* and to harvest enough bacilli to be able to perform experiments. The original source material had to come from a human patient, usually from a skin lesion biopsy. The host tissue had to be digested away, usually by enzymatic and chemical means, leaving behind enough viable bacilli to inoculate into the footpads of immunocompetent mice, although growth was limited. Once the strain was established in competent mice, it could be harvested and then expanded in footpads of athymic nude mice that lacked functional T-cell responses. Growth in the nude mouse footpad took six months to a year, yielding up to 10^{10} highly viable bacilli (92), roughly 5 mg of material, which was generally not enough to do substantial work. Established strains could then be harvested from nude mice to be inoculated intravenously into armadillos (93), where the bacilli grew to up to 10^{11} per gram of tissue, yielding between 100–1,000 mg from a single spleen. Techniques to liberate the bacilli using enzymatic and chemical digestion of host tissue followed by purification using polyethylene glycol/dextran sulfate density gradient were developed by Dr. Philip Draper (94). Although the viability of the bacilli from armadillo tissues was low, the yields of *M. leprae* whole cells free of host tissue allowed, for the first time, the ability to produce subcellular protein fractions and other native antigens as well as genomic DNA for study and formed the basis for defining protein antigens expressed *in vivo* by the bacillus using proteomic, immunologic, and genetic tools.

The WHO, through the Scientific Working Group on the Immunology of Leprosy (IMMLEP) and Immunology of Tuberculosis (IMMTUB), sponsored several workshops beginning in 1984. The purpose of the workshops was to begin to identify important protein, carbohydrate, and glycolipid constituents of mycobacteria that might be recognized by immune responses in animal models and humans. Monoclonal antibodies (mAb) generated against mycobacterial antigens in various laboratories around the world were submitted to the mAb bank at the Centers for Disease Control (CDC) in Atlanta, Georgia. The CDC sent the antibodies to laboratories for independent analysis, where a variety of methods were used to characterize them (95, 96, 97). Techniques such as polyacrylamide gel electrophoresis, Western blot, immunoprecipitation, and immunofluorescence of whole cells of various mycobacterial species were used to determine if the mAb were species-specific or cross-reactive. The antigens identified were characterized as proteins if they were

susceptible to proteases and, if resistant, as carbohydrates or lipids. Following the identification of the antigen targets, the mAb were used to screen DNA expression libraries, leading to the cloning of the genes that encoded these antigens and opening up a new era of molecular cloning towards generating specific mycobacterial antigens (98, 99).

Highlights

- The inability to grow *M. leprae in vitro* limited basic research into the biology of this mycobacterium.
- *In vivo* growth of *M. leprae* from patient biopsy material, first in mice and then in the armadillo, led to the establishment of various strains that, for the first time, enabled the growth of enough bacilli to study (see Chapter 10.2).
- The purification of larger quantities of *M. leprae* allowed the identification of important protein, carbohydrate, and glycolipid constituents as well as the basic structure of the cell wall.
- Antibody reagents developed in multiple laboratories were tested for specificity to mycobacterial species to determine which were *M. leprae* specific or cross-reactive.

Structural Biology and the Proteome

Aided by a relatively plentiful supply of infected armadillo tissues with a high bacillary load (supplied by the National Hansen's Disease Program), we developed techniques to define, for the first time, the dominant somatic proteins of the leprosy bacillus. Nearly 435 mg of whole cells were sonicated to >95% breakage, followed by several steps of differential centrifugation to prepare the three main subcellular fractions of *M. leprae*, namely, the soluble cytoplasmic proteins/antigens (MLSA), the cytoplasmic membrane proteins/antigens (MLMA), and the cell wall proteins/antigens (MLCwA) (100). It was established that the cell wall proteins could be selectively removed from the insoluble cell wall core (mycolyl-arabinogalactan-peptidoglycan complex [MAGP]) by extraction with 2% sodium dodecyl sulfate at 55°C. Triton X-114 bi-phasic separation was used to partition phospholipids and lipoglycans (such as LM and LAM) into the detergent phase, with the membrane proteins remaining in the aqueous phase. The identities of the major proteins in each subcellular fraction were derived by several methods, notably, separation by dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). SDS-PAGE allowed the determination of relative molecular weights, N-terminal sequence analysis, and reactivity to the substrate-defined mAbs generated in several WHO/IMMLEP workshops.

Three proteins were identified in the soluble MLSA fraction: MCP-I (the major cytoplasmic protein I, a 10 kDa chaperonin also known as GroES, or ML0380); MCP-II (the 23 kDa superoxide dismutase, ML0072); and the 28 kDa MCP-III, which was later characterized as a member of the mycolyltransferase family (Ag85 complex, ML2028). ML2028 is a key enzyme in mycolate transfer used to produce trehalose monomycolate (TMM) and trehalose dimycolate (TDM), components essential for the biogenesis and integrity of the cell wall (59). Two proteins dominated the membrane fraction: the 35 kDa MMP-I (major membrane protein I, ML0841); and the 18 kDa MMP-II (bacterioferritin, ML2038). For the SDS-soluble cell wall fraction, one major band that appeared as a doublet around 17 kDa was later identified as the 18 kDa heat shock protein (hsp18, ML1795). There was also evidence of the 65 kDa heat shock protein (ML0317) in this fraction, as had been previously described (101).

The development and use of two-dimensional (2-D) polyacrylamide gel electrophoresis to separate proteins by mass and charge in individual mycobacterial subcellular compartments led to new techniques of isolating discrete spots to perform either N-terminal amino acid sequencing or in-gel trypsin digestion for later mass spectroscopy analysis of peptide fragments (102). As more protein sequences were identified and placed into databases, maps of these protein spots by 2-D gel led to the identification of many more proteins than could be visualized by one-dimensional SDS-PAGE, as in the past (103). Once the genome of *M. leprae* was fully sequenced (6), the protein sequences of all of the 1,614 predicted coding genes allowed for the pre-determination of the theoretical molecular weight and isoelectric point (p*I*). That pre-determination, in turn, allowed for predictions about the location of unknown protein spots in certain regions of a 2-D gel based on molecular weight and charge (104). In addition, programs were developed that could identify all of the possible tryptic peptides and their theoretical fragment masses stored in an NCBI mycobacterial database. Experimentally generated MALDI-MS peptide mass fingerprint data were generated for each subcellular fraction, and proteins were positively identified with a required minimum of three peptides and 20% coverage for each. This high-throughput proteomics approach identified a total of 218 new proteins distributed by intracellular compartment as follows: 60 in the soluble cytosolic fraction (43 unique to this fraction), 98 in the membrane fraction (62 unique), and 104 in the cell wall fraction (73 unique) (105). These proteins could be assorted into specific functional categories, with the ability to compare each to the orthologs from *M. tuberculosis* to determine percent amino acid identity (Table A1 in Appendix). Proteins that overlapped in more than one compartment were in the minority, and only four proteins were common to all three subcellular fractions (ML1890, ML1891, ML2302, and ML2454c), indicating minimal cross-contamination of each of these compartments.

The identification of 13.5% of all of the proteins theoretically coded for by the *M. leprae* genome was a remarkable accomplishment, considering that only 200 mg of *in vivo* derived whole cells were used to prepare the subcellular fractions for the analysis. With the inclusion of these and all other proteins discovered in previous studies, a total of 284 *M. leprae* proteins (17.6%) were identified. Functional categorization of these proteins (lipid metabolism; intermediary metabolism and respiration; cell wall and cell processes; information pathways; virulence, detoxification, and adaptation; regulatory proteins; and conserved hypotheticals) led to a greater understanding of the metabolic and catabolic pathways that the intracellular pathogen required for survival and

multiplication within host cells. It also revealed antigens that were expressed at high enough levels that could be recognized by either cell-mediated or humoral immune responses (106, 107). The antigens have the potential to be developed into subunit vaccines or diagnostic tests, and important metabolic pathways that were identified could also be targeted with new drugs.

Highlights

- Subcellular fractionation of *M. leprae* whole cells into soluble cytosol, membrane, and cell wall allowed for the identification of the proteins that make up each compartment.
- Two-dimensional protein electrophoresis of each subcellular fraction followed by trypsin digestion of spots revealed the identity of many dominant proteins in each fraction.
- Experimentally generated MALDI-MS peptide mass fingerprint data that could match sequences in the NCBI mycobacterial databases were used to identify 13.5% of all of the proteins theoretically coded for by the *M. leprae* genome.

Fundamental 'Metabolome' of *Mycobacterium leprae*

Tribute should be made to the now largely forgotten pioneering, pre-genomic definition of the central metabolic pathways of *M. leprae* by P.R. Wheeler in the early 1980s, who studied the metabolic capabilities of *M. leprae* grown in armadillo tissues ("live" in the sense of enzymatic activity). Wheeler showed that glucose was metabolized in *M. leprae* by glycolysis. Also, he showed that the hexose monophosphate shunt/pentose phosphate pathway was operative (108). A follow up of his work established that suspensions of *M. leprae* oxidized glycerol, 6-phosphogluconate, glucose, glucose 6-phosphate, and, at a low rate, gluconate to CO_2. The conversion rate of the reactions in these experiments was usually slow, 20 h, requiring 140 h to show measurable levels of activity. Studies with differentially radio-labeled glucose indicated that glycolysis and the hexose monophosphate pathway were used for glucose dissimilation. Key enzymes of glycolysis, the hexose monophosphate pathway, and glycerol catabolism were detected in cell-free extracts from purified *M. leprae*, but phosphoketolase, Entner-Doudoroff pathway activity, and gluconate kinase were apparently absent. While all of these enzymes were also present in host-tissue, convincing biochemical evidence was presented that indicated that all enzymes detected in extracts from *M. leprae* were authentic bacterial enzymes. Additionally, the enzymes could all be detected in extracts of *M. leprae* prepared by treatment with NaOH, in which host enzymes adsorbed to purified *M. leprae* were apparently inactivated (109).

Other enzymatic activities detected during this series of studies on cell-free extracts of *M. leprae* from armadillo liver included the following: superoxide dismutase and associated peroxidatic activity (110); N-acetyl-β-glucosaminidase, β-glucuronidase acid phosphatase, and glycosidase activities (111); and a NAD-dependent malate dehydrogenase and FAD-dependent malate-vitamin K reductase detected at about 1% the level of the NAD-dependent activity (112). Evidence of aspartate metabolism (113), *de novo* biosynthesis and scavenging of pyridines (114, 115), the conversion of lactate to pyruvate (116), and phospholipase activity (117) was also found. The significance of the detection of these isolated enzyme activities in *M. leprae* in the context of the holistic catabolism and anabolism of *M. leprae* was not clear other than demonstrating that carefully prepared cell-free extracts obtained from infected armadillo livers and spleens were metabolically active, contrary to the then prevailing belief that they were 'dead'. An important observation in the context of the pathogenicity of *M. leprae* (118) was that iron chelated to the exochelins from *Mycobacterium neoaurum* was taken up by a suspension of *M. leprae*. However, no uptake occurred when the iron was chelated with exochelins from *M. bovis* BCG or *M. smegmatis*, suggestive of a unique nature of *M. leprae* exochelins.

Other publications during the period by Wheeler et al. took a more global view of metabolism, such as the ability of *M. leprae* to complete the full synthesis of fatty acids (119, 120, 121). These investigators also generated thoughtful reviews on the capabilities and defects in the overall aspects of *M. leprae* metabolism (122, 123, 124, 125). Namely, the reviews noted the known obligate intracellular nature of *M. leprae* and its utter dependence on host metabolites and the consequent dominant question of the day: what were the precise details of *M. leprae* metabolic independence *versus* dependency?

Much more decisive answers on the details of host metabolic dependency of *M. leprae* were provided by Stewart Cole et al., the authors of the first complete definition of the genome of an isolate of *M. leprae* (6) (see Chapter 8.2). According to Cole et al., the genome of *M. leprae* provides a classic case of "extreme reductive evolution" compared to the genome of *M. tuberculosis*. Less than one-half of the genome (49.5%) contains functional genes, whereas "27% contains recognizable pseudogenes". Furthermore, "the remaining 23.5% of the genome does not appear to be coding, and may correspond to regulatory sequences or even gene remnants mutated beyond recognition" (6). Cole et al. (6) have analyzed the effects of this dramatic gene decay and genome downsizing on central and energy metabolism in particular. For instance, the crucial anaerobic respiratory chain is apparently severely compromised since "only the extreme 3' end of the NADH oxidase operon, *nuoA-N*, remains". According to the authors, "not only has the potential to produce ATP from the oxidation of NADH been lost, but also the regeneration of NAD^+ may be limited". They speculate on alternative routes for the oxidation of NADH, as well as emphasize evidence showing that *M. leprae* had apparently "lost microaerophilic electron transfer systems such as formate dehydrogenase, nitrate and fumarate reductase". However, it is not clear how important aerobic/microaerophilic electron transfer is to the *in vivo* life cycle of *M. leprae* and its pathogenesis.

In her separate analysis of the metabolic capability of *M. leprae* arising from definition of the genome, Vissa (7) concludes that

Mutations are found in genes involved in regulation (encoding repressors, activators, two-component systems, serine-threonine kinases and phosphatases), detoxification (genes encoding peroxidases), DNA repair (the *mutT*, *dnaQ*, *alkA*, *dinX*, and *dinP* genes) and transport or efflux of metabolites such as amino acids (arginine, ornithine, D-alanine, D-serine and glycine), peptides, cations (magnesium, nickel, mercury, ammonium, ferrous and ferric ions and potassium), and anions (arsenate, sulfate and phosphate).

She (7) also concludes, as had Cole et al. (6), that "In general, pseudogenes are found more frequently in degradative, rather than synthetic pathways". As examples, she reports that genes for the synthesis of the majority of such small metabolites as amino acids, purines, pyrimidines, and fatty acids was generally intact, as were the genes necessary for the synthesis of major macromolecules such as ribosomes, aminoacyl RNAs, RNA, and proteins.

Addressing gene reduction as distinct from mutations, both Cole et al. (6) and Vissa (7) stress genes/coding sequences that were particularly affected, such as insertion sequences and the acidic and glycine-rich families of proteins containing proline-glutamic acid (PE) or proline-proline-glutamic acid (PPE) motifs, which are important in antigenic variation and substrates for the Type VII mycobacterial secretion system (126, 127, 128). In addition, Vissa mentions that expected *M. leprae* repressors, activators, oxidoreductases, and oxygenases were also affected by gene reduction. Her conclusion from this analysis is as follows: "Thus, while preserving genes required for its transmission, establishment and survival in the host, *M. leprae* has discarded genes that can be compensated for by a host-dependent parasitic lifestyle".

Highlights

- Early studies by P.R. Wheeler defined the central metabolic capabilities of *M. leprae* using armadillo-derived bacilli.
- Key enzymes involved in glycolysis, the hexose monophosphate pathway, and glycerol catabolism were present, while those of the Entner-Doudoroff pathway and other enzymes were apparently absent.
- The complete sequencing of the *M. leprae* genome in 2001 revealed extreme reductive evolution, with only 49.5% of the genome retaining functional coding capacity, leading to a full understanding of the dependency of *M. leprae* on host metabolites.
- Dramatic gene decay and genome downsizing has an impact on central and energy metabolism, in particular. The crucial anaerobic respiratory chain is apparently severely compromised, and the ability to produce ATP from the oxidation of NADH is lost.
- Genes for the synthesis of the majority of small metabolites (amino acids, purines, pyrimidines, and fatty acids) were generally intact, as was the synthesis of major macromolecules such as ribosomes, aminoacyl RNAs, RNA, and proteins.

Conclusions

In this comprehensive review of the physiology of *M. leprae,* we try to describe and analyze all that is known of its structure, metabolism, and underlying genetics. Other than the early studies on the morphology of *M. leprae*, our knowledge of its physiology dates only from the 1970s with the advent of infected armadillo tissues as a source of appreciable quantities of the bacterium itself and its genome. Access to milligram quantities of highly purified *M. leprae* allowed actual chemistry-based definition of the cell wall of *M. leprae* and its proteome, the major topics of this review. That definition was aided considerably by established definitions of the cell walls and proteomes of cultivable mycobacteria derived from pioneering research in these and other laboratories over the past 35 years.

In order to present images of the complex metabolic pathways of *M. leprae*—whether the central catabolic and energy-generating routes leading to fundamental 'building blocks' or the specialized anabolic pathways resulting in the formation of the *M. leprae* cell envelope—we rely on a number of sources. In particular, we rely on the pioneering efforts of P.R. Wheeler to define central metabolism in the classical way, using cell-free extracts of armadillo-derived *M. leprae*. In addition, S.T. Cole's pioneering annotation of the genome of *M. leprae*, which he had just defined, has proven invaluable. Finally, the excellent independent annotation of the genome by V. Vissa comprises a body of work underlying much of this present review. The annotation of the *M. leprae* genome by S.T. Cole is particularly helpful in defining the peculiarities of electron transport and ATP generation by *M. leprae*. Perhaps deficiencies in crucial electron flow through the respiratory chain of *M. leprae* may explain the age-old dilemma: the obligate intracellular nature of the leprosy bacillus.

The genomes of *M. leprae*, *M. lepromatosis*, and multiple other mycobacteria have been defined and settled. Annotations have been accomplished but will be an ongoing process. The compositional definition of the cell wall and its biogenesis and genetics, described here in detail, is largely accomplished, but definitions of further minutiae will continue. We realize that much of what we describe is not directly relevant to those in the front line of leprosy treatment and ultimate eradication; however, the topic is an essential part of tomes such as this that strive for a comprehensive understanding of leprosy and its etiological agent.

Acknowledgements

The authors are indebted to the over 30 years of support for the former NIH/NIAID Leprosy Contract to Colorado State University that began in 1979 and ended in 2012. This support largely supplied the raw materials derived from *M. leprae*-infected armadillo tissues that led to the production and free distribution of *M. leprae* whole cells, native subcellular cytosolic, membrane and cell wall protein fractions, whole cell lipids, lipoarabinomannan (LAM), lipoglycans (including

PGL-I), cell wall MAGP, DNA, and derivatives (monoclonal and polyclonal antibody reagents, recombinant proteins, synthetic PGL-I in the form of ND-O-BSA). That distribution allowed leprosy researchers all over the world to gain incredible insight into the host pathogen interaction and to begin to develop new drugs, vaccines, and diagnostic tests that may eventually eliminate this disease. We are grateful to Dr. Shiva Angala for the creation of all of the original figures depicting cell wall structures and metabolic pathways. Our thanks also to all of the patients all over the world who have participated in research studies to advance our knowledge.

Appendix

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
Class 0 – Virulence, Detoxification, Adaptation						
ML0862	*ephD*	putative oxidoreductase	M	Rv2214c	P[d]	80.4%
ML1623[c]	*htpG*	heat shock protein Hsp90 family	CW	Rv2299c	P	80.8%
ML2269	*bpoC*	possible peroxidase BPOC (non-heme-peroxidase)	C/CW	Rv0554	–	83.5%
ML2490[c]	*clpB*	heat shock protein	CW	Rv0384c	P[d]	89.0%
ML2588	*yrbE1B*	conserved membrane protein	CW	Rv0168	T	85.1%
Class I – Lipid Metabolism						
ML0134	*fadD22*	putative acyl-CoA synthetase	M	Rv2948c	T	75.9%
ML0135	*pks1*	putative polyketide synthase	CW	Rv2946c Rv2947c	P[d] P/T[d]	90.5% 82.5%
ML0138c	*fadD28*	acyl-CoA synthetase	M/CW	Rv2941	T	81.9%
ML0139	*mas*	putative mycocerosic synthase	M/CW	Rv2940c	P/T[d]	88.0%
ML0323	*cysS*	putative cysteinyl-tRNA synthase	M/CW	Rv3580c	–	86.5%
ML0389	*choD*	putative cholesterol oxidase	CW	Rv3409c	T	88.6%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML0452c		probable alpha-mannosyltransferase PimA	M	Rv2610c	–	82.3%
ML0726c	*bccA*	acetyl/propionyl CoA carboxylase [alpha] subunit	M	Rv3285	P[d]	88.7%
ML0731c	*accD5*	acetyl/propionyl CoA carboxylase [beta] subunit	M	Rv3280	P[d]	88.7%
ML1158	*atpA4*	probable acetyl-CoA acetyltransferase	C	Rv1323	P/T	86.9%
ML1191[c]	*fas*	fatty acid synthase	M/CW	Rv2524c	P/T[d]	85.8%
ML1229	*pks4*	type-I polyketide synthase	CW	Rv1180 Rv1181	– P/T[d]	81.3% 75.3%
ML1234c	*fadD21*	possible acyl-CoA synthase	CW	Rv1185	T	72.4%
ML1657	*accD6*	acetyl/propionyl CoA carboxylase [beta] subunit	CW	Rv2247	T	92.8%
ML1806c[c]	*inhA*	NADH-dependent enoyl-[acyl-carrier-protein] reductase	C/CW	Rv1484	–	90.3%
ML1996	*nrp*	probable peptide synthase	CW	Rv0101	P[d]	65.7%
ML2349	*papA5*	possible conserved polyketide synthase associated protein	CW	Rv2939	T	83.8%
ML2353c	*ppsE*	polyketide synthase	M/CW	Rv2935	P/T[d]	82.0%
ML2355c	*ppsC*	polyketide synthase	M/CW	Rv2933	P/T[d]	82.3%
ML2358c	*fadD26*	probable acyl-CoA synthase	CW	Rv2930	P/T[d]	76.8%
ML2563[c]	*fadE5*	acyl-CoA dehydrogenase	C/M	Rv0244c	P[d]	89.4%
ML2565	*fabG4*	possible oxidoreductase	M/CW	Rv0242c	P/T[d]	84.8%
ML2655[c]	*fbpC2*	secreted antigen 85C	C	Rv0129c	P	81.3%
Class II – Information Pathways						
ML0492c	*ppiB*	putative peptidyl-prolyl cis-trans isomerase B	M	Rv2582	–	72.5%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML0548	*priA*	putative primosomal protein N' (Replication factor Y)	CW	Rv1402	P[d]	80.7%
ML0826	*glyS*	putative glycyl-tRNA synthase	M	Rv2357c	P[d]	90.2%
ML0854	*gpsI*	putative polyribonucleotide phosphorylase / guanosine pentaphosphate synthetase	M/CW	Rv2783c	P[d]	90.0%
ML0987	*recA*	recA protein	C	Rv2737c	P/T[d]	46.7%
ML1174	*rphA*	ribonuclease RPHA tRNA (tRNA nucleotidyl transferase)	C	Rv1340	–	88.8%
ML1381c	*polA*	DNA polymerase I	M	Rv1629	P[d]	85.2%
ML1382	*rpsA*	30S ribosomal protein S1	C/M	Rv1630	P/T[d]	93.8%
ML1498c	*typA*	possible GTP-binding, protein elongation factor	CW	Rv1165	P[d]	85.2%
ML1553	*proS*	prolyl tRNA synthetase	M/CW	–	–	–
ML1590c	*frr*	ribosome recycling factor	C	Rv2882c	P/T	90.8%
ML1598c	*rpsB*	30S ribosomal protein S2	M	Rv2890c	P/T[d]	91.5%
ML1659c	*rnc*	RNAse III	CW	Rv2925c	–	86.7%
ML1702c	*gatA*	probable glutamyl-tRNA amidotransferase	C	Rv3011c	–	88.5%
ML1735c	*nrdI*	NrdI-family protein	C	Rv3052c	–	79.7%
ML1844c	*rplF*	50S ribosomal protein L6	M	Rv0719	P[d]	90.5%
ML1847c	*rplE*	50S ribosomal protein L5	M/CW	Rv0716	P/T[d]	91.4%
ML1848c	*rplX*	50S ribosomal protein L24	CW	Rv0715	P/T[d]	93.3%
ML1857c	*rpsC*	30S ribosomal protein S3	M	Rv0707	P/T[d]	95.0%
ML1860c	*rplB*	50S ribosomal protein L2	M	Rv0704	P/T[d]	92.1%
ML1861c	*rplW*	50S ribosomal protein L23	M	Rv0703	P/T[d]	85.0%
ML1890c[c]	*rpoC*	[beta]' subunit of RNA polymerase	C/M/CW	Rv0668	P/T[d]	95.6%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML1891c[c]	*rpoB*	[beta] subunit of RNA polymerase	C/M/CW	Rv0667	T	94.9%
ML1906c[c]	*nusG*	transcription antitermination protein nusG	C	Rv0639	P	93.4%
ML1956c	*rplQ*	50S ribosomal protein L17	M	Rv3456c	P[d]	89.9%
ML1959c	*rpsK*	30S ribosomal protein S11	CW	Rv3459c	T	81.9%
ML2393	*greA*	probable transcriptional elongation factor	C	Rv1080c	P[d]	94.5%
ML2684c	*ssb*	single strand binding protein	C	Rv0054	P	89.3%
Class III – Cell Wall and Cell Processes						
ML0042		putative conserved membrane protein	CW	Rv3882c	P[d]	72.1%
ML0047c		putative conserved transmembrane protein	CW	Rv3877	–	74.0%
ML0092	*glf*	putative UDP-galactopyranose mutase	CW	Rv3809c	P[d]	86.6%
ML0093	*glfT*	bifunctional UDP-galactofuranosyl transferase	M	Rv3808c	P/T[d]	85.4%
ML0096		putative conserved transmembrane protein	M/CW	Rv3805	–	80.9%
ML0105c	*embA*	putative arabinosyl transferase	CW	Rv3794	P[d]	64.2%
ML0136	*llpX*	putative conserved lipoprotein	M/CW	Rv2945c	P/T	76.4%
ML0203		putative conserved transmembrane protein	CW	Rv3635	–	77.2%
ML0227		hypothetical protein	C	Rv3605c	–	85.4%
ML0319c	*lpqE*	putative lipoprotein	CW	Rv3584	P	63.4%
ML0488[c]	*secF*	putative protein-export membrane protein	M	Rv2586c	–	72.1%
ML0557c	*lprG*	putative lipoprotein	M/CW	Rv1411c	P/T[d]	68.1%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML0569c		possible conserved exported protein	M	Rv1433	–	68.3%
ML0595		probable conserved ATP-binding protein ABC transporter	C	Rv1463	P	85.8%
ML0644[c]		putative integral membrane protein	M/CW	Rv3193c	P	86.1%
ML0691c	*dacB1*	putative D-alanyl-D-alanine carboxypeptidase	M	Rv3330	–	77.2%
ML0779[c]	*secA*	putative preprotein translocase subunit	CW	Rv3240c	P[d]	87.5%
ML0810c		possible secreted protein	CW	Rv3212	P/T[d]	79.9%
ML0857c		putative conserved transmembrane protein	CW	Rv2219	–	80.8%
ML0910	*murF*	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase (D-alanine:D-alanine-adding enzyme)	CW	Rv2175c	–	75.5%
ML0977[c]	*ftsK*	cell division transmembrane protein	CW	Rv2748c	P[d]	78.1%
ML1248		probable macrolide ABC-transporter protein, ATP binding component	CW	Rv2477c	P[d]	92.3%
ML1334c		possible conserved membrane protein	CW	Rv2091c	T	62.9%
ML1425c		probable sugar ABC-transport integral membrane protein	M	Rv2039c	–	79.2%
ML1562c	*efpA*	putative transmembrane efflux protein	CW	Rv2846c	T	86.5%
ML1622	*ffh*	signal recognition particle protein	M	Rv2916c	P[d]	87.8%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML1633c[c]		probable exported protease	M	Rv2224c	P/T[d]	84.8%
ML1667c		possible conserved membrane or secreted protein	M	Rv2969c	P	65.5%
ML1769	*uspB*	probable sugar ABC transporter integral membrane protein	C	Rv2317	–	85.0%
ML1808c		possible conserved membrane protein	M	Rv1481	P	91.0%
ML1839c	*sppA*	protease IV, signal peptide peptidase	M	Rv0724	P	80.2%
ML2095c	*psts3*	periplasmic phosphate-binding lipoprotein	M	Rv0928	P	77.3%
ML2296		putative conserved transmembrane protein	M	Rv3669	–	77.9%
ML2305		probable anion transporter ATPase	M	Rv3679	T	84.2%
ML2378	*mmpL4*	conserved large transmembrane protein	CW	Rv0450c	T	78.4%
ML2410c		conserved transmembrane protein.	CW	Rv0528	–	80.9%
ML2433		conserved membrane protein	M	Rv0497	–	54.2%
ML2454c	*hbhA*	possible hemagglutinin	C/M/CW	Rv0475	P/T[d]	81.4%
ML2687c		conserved transmembrane protein	CW	Rv0051	–	78.7%
Classe VII – Intermediary Metabolism and Respiration						
ML0065	*ethA*	putative monooxygenase	CW	Rv3854	P/T[d]	80.1%
ML0074	*glpQ1*	putative glycerophosphoryl diester phosphodiesterase	CW	Rv3842c	T	88.8%
ML0108c		putative oxidoreductase	M	Rv3791	–	89.0%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML0109c		putative FAD-linked oxidoreductase	CW	Rv3790	P[d]	86.6%
ML0127[c]		possible methyltransferase	M/CW	Rv2959c	–	76.1%
ML0150c	*pgi*	glucose-6-phosphate isomerase	CW	Rv0946c	–	85.5%
ML0155	*sucC*	succinyl-CoA synthase [beta]	CW	Rv0951	P[d]	86.7%
ML0161	*purH*	putative phosphoribosyl-amino-imidazolecarboxamide formyl-transferase/ IMP cyclohydrolase	M	Rv0957	–	88.1%
ML0183	*moeA1*	putative molybdopterin biosynthesis protein	CW	Rv0994	P[d]	88.3%
ML0210c[c]	*ppa*	putative inorganic pyrophosphatase	M	Rv3628	P[d]	89.5%
ML0235	*clpC*	putative ATP-dependent Clp protease ATP-binding subunit ClpC	M/CW	Rv3596c	P[d]	97.2%
ML0255	*eno*	putative enolase	M	Rv1023	P	88.8%
ML0269c	*fgd1*	putative F-420-dependent glucose-6-phosphate dehydrogenase	C	Rv0407	P	89.3%
ML0297c[c]	*thiG*	putative thiamine biosynthesis protein	CW	Rv0417	–	86.8%
ML0348		possible coenzyme F420-dependent oxidoreductase	C/M	Rv3520c	–	86.8%
ML0365	*rpsI*	30S ribosomal protein S9	M	Rv3442c	–	83.2%
ML0379	*gcp*	probable O-sialoglycoprotein endopeptidase	C	Rv3419c	–	86.1%
ML0388	*guaB3*	putative inosine-5'-monophosphate dehydrogenase	C/M	Rv3410c	–	90.5%
ML0393c		probable hydrolase	C	Rv3400	P	74.4%
ML0418		putative oxidoreductase	CW	Rv3368c	–	81.9%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML0429c		probable short-chain type dehydrogenase/reductase	C	Rv2509	P[d]	88.8%
ML0485c	*gabT*	putative 4-aminobutyrate aminotransferase	CW	Rv2589	–	83.7%
ML0521	*pepQ*	putative cytoplasmic peptidase PepQ	C	Rv2535c	–	83.0%
ML0537	*pyrF*	orotidine 5'-phosphate decarboxylase	CW	Rv1385	–	77.3%
ML0544	*metK*	putative S-adenosylmethionine synthase	C/M	Rv1392	P/T[d]	91.3%
ML0551		putative methyltransferase	C	Rv1405c	–	77.7%
ML0559	*ribA*	putative GTP cyclohydrolase II/3,4-dihydroxy-2-butanone-4-phosphate synthase	M	Rv1415	P	93.4%
ML0570	*gap*	glyceraldehyde 3-phosphate dehydrogenase	M	Rv1436	P[d]	89.1%
ML0572	*tpi*	triosephosphate isomerase	C	Rv1438	P[d]	83.9%
ML0674[c]	*folD*	putative methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	C	Rv3356c	P	86.5%
ML0685c		putative hydrolase	M	Rv3337 Rv3338	– –	74.6% 65.9%
ML0763	*pmmA*	putative phosphomannomutase	CW	Rv3257c	–	83.1%
ML0839c	*cysK*	putative cysteine synthase	C/M	Rv2334	P	85.8%
ML0861c	*sucB*	putative dihydrolipoamide acyltransferase	C	Rv2215	P[d]	81.8%
ML0873	*adk*	putative adenosine kinase	C	Rv2202c	–	83.3%
ML0925	*glnA*	glutamine synthase class I	M	Rv2220	P[d]	91.8%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML0961c	*metE*	5-methyltetrahydropteroyltriglutamate-homocystein methyltransferase	C/CW	Rv1133c	P[d]	85.0%
ML0985		possible glycosyltransferase	CW	Rv2739c	–	80.9%
ML0995	*miaA*	tRNA [delta](2)-isopentenyl-pyrophosphate transferase	CW	Rv2727c	–	82.2%
ML1059	*dapE*	succinyl-diaminopimelate desuccinylase	M	Rv1202	–	86.7%
ML1091	*mdh*	probable malate dehydrogenase	C	Rv1240	P[d]	89.1%
ML1095c	*odhA/sucA*	2-oxoglutarate dehydrogenase, E1 components	M	Rv1248c	P[d]	89.0%
ML1129	*hom*	homoserine dehydrogenase	M	Rv1294	–	99.8%
ML1142	*atpH*	bi-domained ATP synthase B, [delta] chain protein	M	Rv1307	P/T[d]	77.8%
ML1143	*atpA*	ATP synthase [alpha] chain	C	Rv1308	P/T[d]	69.0%
ML1144	*atpG*	ATP synthase [gamma] chain	M/CW	Rv1309	P/T[d]	83.9%
ML1145	*atpD*	ATP synthase [beta] chain	C/M	Rv1310	P/T[d]	92.6%
ML1167		possible hydrolase	M	Rv1333	–	76.7%
ML1257	*hisD*	histidinol dehydrogenase	CW	Rv1599	–	84.8%
ML1307[c]	*metH*	probable 5-methyltetrahydrofolate-homocystein methyltransferase metH (methionine synthase, vitamin-B12 dependent isozyme)	C	Rv2124c	P[d]	88.3%
ML1340c	*ribD*	possible bifunctional enzyme riboflavin biosynthesis protein RibD + 5-amino-6-uracil reductase	C	Rv2671	P[d]	72.9%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML1469c	*ndk*	nucleoside diphosphate kinase NDK	C	Rv2445c	P	87.4%
ML1480	*clpP*	ATP-dependent Clp protease proteolytic subunit	C	Rv2461c	–	95.0%
ML1513c	*dapA*	probable dihydrodipicolinate synthase DAPA (DHDPS)	C	Rv2753c	–	86.7%
ML1669		probable oxidoreductase	C	Rv2971	P	79.8%
ML1685c	*leuC*	3-isopropylmalate dehydratase large subunit	CW	Rv2988c	P/T[d]	88.7%
ML1691c	*leuB*	3-isopropylmalate dehydrogenase	M	Rv2995c	T	83.9%
ML1694c	*ilvC*	ketol-acid reductoisomerase IlvC	C	Rv3001c	P/T[d]	86.2%
ML1696c	*ilvB*	probable acetolactate synthase (large subunit) IlvB	C	Rv3003c	P	87.1%
ML1712c	*fixA*	probable electron transfer flavoprotein (beta subunit) FixA	C	Rv3029c	P[d]	95.1%
ML1730	*adhA*	alcohol dehydrogenase	CW	Rv3045	P[d]	85.8%
ML1832c	*kad*	probable adenylate kinase Kad	C	Rv0733	P	83.4%
ML1942		probable cholesterol dehydrogenase	CW	Rv1106c	P	75.5%
ML1950c[c]		carbohydrate degrading enzyme	M	Rv1096	–	81.3%
ML1985c[c]	*aceA*	isocitrate lyase	M/CW	Rv1916	T	87.9%
ML2011	*aao*	probable D-amino acido oxidase AAO	C	Rv1905c	–	71.4%
ML2046	*lldD2*	L-lactate dehydrogenase	M	Rv1872c	P/T[d]	85.0%
ML2088c		putative cytochrome P450	C/M	Rv1880c	–	38.0%
ML2130c	*gltA2*	citrate synthase GLTA2	C/M	Rv0896	P	91.0%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML2226c	*ptrB*	protease II	CW	Rv0781 Rv0782	– –	69.9% 77.8%
ML2227c	*purC*	phosphoribosylaminoimidazole-succinocarboxamide synthase	M	Rv0780	–	81.5%
ML2322c	*asd*	aspartate semialdehyde dehydrogenase	M	Rv3708c	–	84.4%
ML2324	*leuA*	probable 4-dihydroxy-2-naphthoate octaprenyltransferase 2-isopropylmalate synthase	CW	Rv3710	P	80.6%
ML2387[c]	*lpd*	dihydrolipoamide dehydrogenase Lpd	M	Rv0462	P	83.1%
ML2394c	*metB*	cystathionine [gamma]-synthase	M	Rv1079	P	87.3%
ML2396c	*cysM2*	probable cystathionine [beta]-synthase CBS	C	Rv1077	P	90.1%
ML2406[c]	*menA*	UDP-galactose 4-epimerase	CW	Rv0534c	–	80.6%
ML2419c	*hemB*	hemB, [delta]-aminolevulinic acid dehydratase	CW	Rv0512	–	87.5%
ML2428c	*galE2*	probable UDP-galactose 4-epimerase GALE2	C	Rv0501	P	84.0%
ML2430c	*proC*	pyrroline-5-carboxylate reductase	M/CW	Rv0500	–	82.4%
ML2441c[c]	*gpm*	phosphoglycerate mutase	C	Rv0489	–	89.0%
ML2449c		putative amidohydrolase	CW	Rv0480c	P	81.9%
ML2474		possible cytidine/deoxycytidylate deaminase	CW	Rv3752c	–	87.4%
ML2501[c]		probable iron-sulphur-binding reductase	CW	Rv0338c	P	78.1%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML2573	*gabD*	probable succinate-semi-aldehyde dehydrogenase [NADP+] dependent	C	Rv0234c	P	82.9%
ML2659c[c]	*pepA*	probable secreted serine protease	M	Rv0125	T	70.1%
ML2672c	*icd2*	probable isocitrate dehydrogenase [NADP+]	M	Rv0066c	P	85.4%
Class VIII – Hypothetical Unknown						
ML0959		hypothetical protein	CW	–	–	–
Class IX – Regulatory Proteins						
ML0639[c]	*whiB7*	putative transcriptional regulator	CW	Rv3197A	P	69.3%
ML1286		possible two-component response regulatory protein	CW	Rv1626	P	90.7%
ML1511	*npdA*	probable NAD-dependent deacetylase NpdA	M	Rv1151c	–	85.2%
ML2302		putative Crp/Fnr-family transcriptional regulator	C/M/CW	Rv3676	P	96.0%
ML2440c	*senX3*	probable two-component system sensor histidine kinase	M	Rv0490	P	84.0%
Class X – Conserved Hypotheticals						
ML0129c		conserved hypothetical protein	M	Rv2953	P/T	86.6%
ML0333c		conserved hypothetical protein	C	–	–	–
ML0369c		conserved hypothetical protein	CW	–	–	–
ML0405		conserved hypothetical protein	M	Rv3616c	T	62.7%
ML0447		conserved hypothetical protein	M/CW	–	–	–

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML0593		conserved hypothetical protein	M	Rv1461	P	54.3%
ML0703		conserved hypothetical protein	M	Rv3311	–	77.5%
ML0781		conserved hypothetical protein	C	Rv3237c	T	80.6%
ML0901		conserved hypothetical protein	CW	Rv2172c	T	80.9%
ML1040		conserved hypothetical protein	M	Rv2681	P	77.4%
ML1330		conserved hypothetical protein	CW	Rv2095c	–	74.2%
ML1368		conserved hypothetical protein	CW	Rv1709	–	78.9%
ML1369		conserved hypothetical protein	M/CW	Rv1710	–	77.9%
ML1380		conserved hypothetical protein	M/CW	Rv1780	P	82.4%
ML1418c		conserved hypothetical protein	C/CW	Rv1732c	P	86.3%
ML1423c		conserved hypothetical protein	M	Rv2037	P/T	80.5%
ML1444		conserved hypothetical protein	M	Rv2054	P	76.1%
ML1637c		conserved hypothetical protein	CW	Rv2228c	–	76.8%
ML1661c		conserved hypothetical protein	CW	Rv2927c	–	90.2%
ML1704		conserved hypothetical protein	CW	Rv3013	–	84.3%
ML1706c		conserved hypothetical protein	C	Rv3015c	–	78.3%

TABLE A1 Functional distribution of *M. leprae* proteins identified by proteomic analysis and expression of *M. tuberculosis* orthologs (cont'd)

Sanger ID[a]	Gene	Protein function	Fraction	*M. tuberculosis* Sanger ID[b]	Expression by *M. tuberculosis*	Identity
ML1782c		conserved hypothetical protein	CW	Rv2257c	–	76.4%
ML2031		conserved hypothetical protein	CW	Rv1883c	–	76.2%
ML2070		conserved hypothetical protein	M	Rv1836c	P	66.0%
ML2113c		conserved hypothetical protein	CW	Rv0910c	–	83.9%
ML2336		conserved hypothetical protein	CW	Rv3722c	T	87.6%
ML2465		conserved hypothetical protein	M	Rv0464c	P/T	67.4%
ML2491		conserved hypothetical protein	M/CW	Rv1754c	P	61.5%
ML2627		conserved hypothetical protein	CW	Rv0216	P	35.8%
ML2640c		conserved hypothetical protein	CW	Rv0146	P	78.1%

a Functional classification and ID refer to the original annotated *M. leprae* gene number assigned by Sanger ID from Cole ST et al. 2001. Nature 409:1007–1011.

b *M. tuberculosis* gene number refers to H37Rv Sanger ID from Cole ST et al. 1998. Nature 393:537–544.

c Proteins also shown to be expressed by transcriptional analysis.

d Proteins identified by subcellular protein profiling from Mawuenyega KG et al. 2005. Mol Biol Cell 16:396–404.

C – cytosolic/soluble; M – membrane; CW – cell wall; T – gene expression demonstrated by transcriptional analysis; P – gene expression demonstrated by proteomic analysis.

References

1. **Brennan PJ, Nikaido H.** 1995. The envelope of mycobacteria. Annu Rev Biochem **64:**29–63.

2. **Angala SK, Belardinelli JM, Huc-Claustre E, Wheat WH, Jackson M.** 2014. The cell envelope glycoconjugates of *Mycobacterium tuberculosis*. Crit Rev Biochem Mol Biol **49**(5)**:**361–399.

3. **Jackson M, McNeil MR, Brennan PJ.** 2013. Progress in targeting cell envelope biogenesis in *Mycobacterium tuberculosis.* Future Microbiol **8:**10.2217/fmb.13.52. doi: 10.2217/fmb.13.52

4. **Kaur D, Guerin ME, Škovierová H, Brennan PJ, Jackson M.** 2009. Chapter 2: Biogenesis of the cell wall and other glycoconjugates of *Mycobacterium tuberculosis*. Adv Appl Microbiol **69:**23–78. doi: 10.1016/S0065-2164(09)69002-X

5. **Angala SK, Palčeková Z, Belardinelli JM, Jackson M.** 2018. Covalent modifications of polysaccharides in mycobacteria. Nat Chem Biol **14:**193–198. doi: 10.1038/nchembio.2571

6. **Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007–1011.

7. **Vissa VD, Brennan PJ.** 2001. The genome of *Mycobacterium leprae*: a minimal mycobacterial gene set. Genome Biol **2:**1023–1030.

8. **Hirata T.** 1985. Electron microscopic observations of cell wall and cytoplasmic membrane in murine human leprosy bacilli. Int J Lepr Other Mycobact Dis **53:**433–440.

9. **Draper P.** 1983. The bacteriology of *Mycobacterium leprae*. Tubercle **64:**43–56.

10. **Nishiura M, Okada S, Izumi S, Takizawa H.** 1969. An electron microscope study of the band structure of the leprosy bacillus and other mycobacteria. Int J Lepr Other Mycobact Dis **37:**225–238.

11. **Brennan PJ, Vissa VD.** 2001. Genomic evidence for the retention of the essential mycobacterial cell wall in the otherwise defective *Mycobacterium leprae*. Lepr Rev **72:**415–428.

12. **Melancon-Kaplan J, Hunter SW, McNeil M, Stewart C, Modlin RL, Rea TH, Convit J, Salgame P, Mehra V, Bloom BR, Brennan PJ.** 1988. Immunological significance of *Mycobacterium leprae* cell walls. Proc Natl Acad Sci USA **85:**1917–1921.

13. **Draper P, Kandler O, Darbre A.** 1987. Peptidoglycan and arabinogalactan of *Mycobacterium leprae*. J Gen Microbiol **133:**1187–1194.

14. **Daffé M, McNeil M, Brennan PJ.** 1993. Major structural features of the cell wall arabinogalactans of Mycobacterium, Rhodococcus, and Nocardia spp. Carbohydr Res **249:**383–398.

15. **Mahapatra S, Crick DC, Brennan PJ.** 2000. Comparison of the UDP-N-acetylmuramate:L-alanine ligase enzymes from *Mycobacterium tuberculosis* and *Mycobacterium leprae*. J Bacteriol **182:**6827–6830.

16. **Mahapatra S, Crick DC, McNeil MR, Brennan PJ.** 2008. Unique structural features of the peptidoglycan of *Mycobacterium leprae*. J Bacteriol **190:**655–661. doi: 10.1128/JB.00982-07

17. **Mahapatra S, Scherman H, Brennan PJ, Crick DC.** 2005a. N glycolylation of the nucleotide precursors of peptidoglycan biosynthesis of *Mycobacterium* spp. is altered by drug treatment. J Bacteriol **187:**2341–2347. doi: 10.1128/JB.187.7.2341-2347.2005

18. **Mahapatra S, Yagi T, Belisle JT, Espinosa BJ, Hill PJ, McNeil MR, Brennan PJ, Crick DC.** 2005b. Mycobacterial lipid II is composed of a complex mixture of modified muramyl and peptide moieties linked to decaprenyl phosphate. J Bacteriol **187**:2747–2757. doi: 10.1128/JB.187.8.2747-2757.2005

19. **Draper P.** 1976. Cell walls of *Mycobacterium leprae*. Int J Lepr Other Mycobact Dis **44:**95–98.

20. **McNeil M, Wallner SJ, Hunter SW, Brennan PJ.** 1987. Demonstration that the galactosyl and arabinosyl residues in the cell-wall arabinogalactan of *Mycobacterium leprae* and *Mycobacterium tuberculosis* are furanoid. Carbohydr Res **166:**299–308.

21. **McNeil M, Daffe M, Brennan PJ.** 1990. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. J Biol Chem **265:**18200–18206.

22. **Daffé M, Brennan PJ, McNeil M.** 1990. Predominant structural features of the cell wall arabinogalactan of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by ^{1}H and ^{13}C NMR analyses. J Biol Chem **265:**6734–6743.

23. **Bhamidi S, Scherman MS, Jones V, Crick DC, Belisle JT, Brennan PJ, McNeil MR.** 2011. Detailed structural and quantitative analysis reveals the spatial organization of the cell walls of *in vivo* grown *Mycobacterium leprae* and *in vitro* grown *Mycobacterium tuberculosis*. J Biol Chem **286:**23168–23177. doi: 10.1074/jbc.M110.210534

24. **McNeil M, Daffe M, Brennan PJ.** 1991. Location of the mycolyl ester substituents in the cell walls of mycobacteria. J Biol Chem **266:**13217–13223.

25. **Barry CE 3rd, Lee RE, Mdluli K, Sampson AE, Schroeder BG, Slayden RA, Yuan Y.** 1998. Mycolic acids: structure, biosynthesis and physiological functions. Prog Lipid Res **37:**143–179.

26. **Minnikin DE, Dobson G, Goodfellow M, Draper P, Magnusson M.** 1985. Quantitative comparison of the mycolic and fatty acid compositions of *Mycobacterium leprae* and *Mycobacterium gordonae*. J Gen Microbiol **131:**2013–2021.

27. **Daffé M, Draper P.** 1998. The envelope layers of mycobacteria with reference to their pathogenicity. Adv Microb Physiol **39:**131–203.

28. **Behr MA, Schroeder BG, Brinkman JN, Slayden RA, Barry CE 3rd.** 2000. A point mutation in the mma3 gene is responsible for impaired methoxymycolic acid production in *Mycobacterium bovis* BCG strains obtained after 1927. J Bacteriol **182:**3394–3399.

29. **Crick DC, Mahapatra S, Brennan PJ.** 2001. Biosynthesis of the arabinogalactan-peptidoglycan complex of *Mycobacterium tuberculosis*. Glycobiology **11:**107R–118R.

30. **Lepage S, Dubois P, Ghosh TK, Joris B, Mahapatra S, Kundu M, Basu J, Chakrabarti P, Cole ST, Nguyen-Distèche M, Ghuysen JM.** 1997. Dual multimodular class A penicillin-binding proteins in *Mycobacterium leprae*. J Bacteriol **179:**4627–4630.

31. **Basu J, Mahapatra S, Kundu M, Mukhopadhyay S, Nguyen-Distèche M, Dubois P, Joris B, Van Beeumen J, Cole ST, Chakrabarti P, Ghuysen JM.** 1996. Identification and overexpression in *Escherichia coli* of a *Mycobacterium leprae* gene, ponl, encoding a high-molecular-mass class A penicillin-binding protein, PBP1. J Bacteriol **178:**1707–1711.

32. **Grzegorzewicz AE, de Sousa-d'Auria C, McNeil MR, Huc-Claustre E, Jones V, Petit C, Angala SK, Zemanová J, Wang Q, Belardinelli JM, Gao Q, Ishizaki Y, Mikušová K, Brennan PJ, Ronning DR, Chami M, Houssin C, Jackson M.** 2016. Assembling of the *Mycobacterium tuberculosis* cell wall core. J Biol Chem **291:**18867–18879. doi: 10.1074/jbc.M116.739227

33. **Mikusová K, Mikus M, Besra GS, Hancock I, Brennan PJ.** 1996. Biosynthesis of the linkage region of the mycobacterial cell wall. J Biol Chem **271:**7820–7828.

34. **Ma Y, Stern RJ, Scherman MS, Vissa VD, Yan W, Jones VC, Zhang F, Franzblau SG, Lewis WH, McNeil MR.** 2001. Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. Antimicrob Agents Chemother **45:**1407–1416.

35. **Weston A, Stern RJ, Lee RE, Nassau PM, Monsey D, Martin SL, Scherman MS, Besra GS, Duncan K, McNeil MR.** 1997. Biosynthetic origin of mycobacterial cell wall galactofuranosyl residues. Tuber Lung Dis **78:**123–131.

36. **Wolucka BA, McNeil MR, de Hoffmann E, Chojnacki T, Brennan PJ.** 1994. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. J Biol Chem **269:**23328–23335.

37. **Mikušová K, Yagi T, Stern R, McNeil MR, Besra GS, Crick DC, Brennan PJ.** 2000. Biosynthesis of the galactan component of the mycobacterial cell wall. J Biol Chem **275:**33890–33897.

38. **Mikušová K, Beláňová M, Korduláková J, Honda K, McNeil MR, Mahapatra S, Crick DC, Brennan PJ.** 2006. Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. J Bacteriol **188:**6592–6598. doi: 10.1128/JB.00489-06

39. **Kremer L, Dover LG, Morehouse C, Hitchin P, Everett M, Morris HR, Dell A, Brennan PJ, McNeil MR, Flaherty C, Duncan K, Besra GS.** 2001. Galactan biosynthesis in *Mycobacterium tuberculosis*. Identification of a bifunctional UDP-galactofuranosyltransferase. J Biol Chem **276:**26430–26440.

40. **Belanger AE, Besra GS, Ford ME, Mikusova K, Belisle JT, Brennan PJ, Inamine JM.** 1996. The embAB genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. Proc Natl Acad Sci USA **93:**11919–11924.

41. **Escuyer VE, Lety MA, Torrelles JB, Khoo KH, Tang JB, Rithner CD, Frehel C, McNeil MR, Brennan PJ, Chatterjee D.** 2001. The role of the embA and embB gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of *Mycobacterium smegmatis* arabinogalactan. J Biol Chem **276:**48854–48862.

42. **Zhang J, Angala SK, Pramanik PK, Li K, Crick DC, Liav A, Jozwiak A, Swiezewska E, Jackson M, Chatterjee D.** 2011. Reconstitution of functional mycobacterial arabinosyltransferase AftC proteoliposome and assessment of decaprenylphosphorylarabinose analogues as arabinofuranosyl donors. ACS Chem Biol **6:**819–828. doi: 10.1021/cb200091m

43. **Birch HL, Alderwick LJ, Bhatt A, Rittmann D, Krumbach K, Singh A, Bai Y, Lowary TL, Eggeling L, Besra GS.** 2008. Biosynthesis of mycobacterial arabinogalactan: identification of a novel α(1→3) arabinofuranosyltransferase. Mol Microbiol **69:**1191–1206. doi: 10.1111/j.1365-2958.2008.06354.x

44. **Rohmer M, Knani M, Simonin P, Sutter B, Sahm H.** 1993. Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. Biochem J **295:**517–524.

45. **Bailey AM, Mahapatra S, Brennan PJ, Crick DC.** 2002. Identification, cloning, purification, and enzymatic characterization of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose 5-phosphate synthase. Glycobiology **12:**813–820.

46. **Takahashi S, Kuzuyama T, Watanabe H, Seto H.** 1998. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. Proc Natl Acad Sci USA **95:**9879–9884.

47. **Campos N, Rodriguez-Concepción M, Sauret-Güeto S, Gallego F, Lois LM, Boronat A.** 2001. *Escherichia coli* engineered to synthesize isopentenyl diphosphate and dimethylallyl diphosphate from mevalonate: a novel system for the genetic analysis of the 2-C-methyl-D-erythritol 4-phosphate pathway for isoprenoid biosynthesis. Biochem J **353:**59–67.

48. **Rodríguez-Concepción M, Campos N, Lois LM, Maldonado C, Hoeffler JF, Grosdemange-Billiard C, Rohmer M, Boronat A.** 2000. Genetic evidence of branching in the isoprenoid pathway for the production of isopentenyl diphosphate and dimethylallyl diphosphate in *Escherichia coli*. FEBS Lett **473:**328–332.

49. **Schulbach MC, Brennan PJ, Crick DC.** 2000. Identification of a short (C15) chain Z-isoprenyl diphosphate synthase and a homologous long (C50) chain isoprenyl diphosphate synthase in *Mycobacterium tuberculosis*. J Biol Chem **275:**22876–22881.

50. **Scherman MS, Kalbe-Bournonville L, Bush D, Xin Y, Deng L, McNeil M.** 1996. Polyprenylphosphate-pentoses in mycobacteria are synthesized from 5-phosphoribose pyrophosphate. J Biol Chem **271:**29652–29658.

51. **Piton J, Foo CS, Cole ST.** 2017. Structural studies of *Mycobacterium tuberculosis* DprE1 interacting with its inhibitors. Drug Discov Today **22:**526–533. doi: 10.1016/j.drudis.2016.09.014

52. **Landge S, Mullick AB, Nagalapur K, Neres J, Subbulakshmi V, Murugan K, Ghosh A, Sadler C, Fellows MD, Humnabadkar V, Mahadevaswamy J, Vachaspati P, Sharma S, Kaur P, Mallya M, Rudrapatna S, Awasthy D, Sambandamurthy VK, Pojer F, Cole ST, Balganesh TS, Ugarkar BG, Balasubramanian V, Bandodkar BS, Panda M, Ramach-**

andran V. 2015. Discovery of benzothiazoles as antimycobacterial agents: synthesis, structure-activity relationships and binding studies with *Mycobacterium tuberculosis* decaprenylphophoryl-β-D-ribose 2'-oxidase. Bioorg Med Chem **23:**7694–7710 doi: 10.1016/j.bmc.2015.11.017

53. **Bailey MA, Na H, Duthie MS, Gillis TP, Lahiri R, Parish T.** 2017. Nitazoxanide is active against *Mycobacterium leprae*. PLoS One **12**(8)**:**e0184107. doi: 10.1371/journal.pone.0184107

54. **Fernandes ND, Kolattukudy PE.** 1996. Cloning, sequencing and characterization of a fatty acid synthase-encoding gene from *Mycobacterium tuberculosis* var. bovis BCG. Gene **170:**95–99.

55. **Choi KH, Kremer L, Besra GS, Rock CO.** 2000. Identification and substrate specificity of beta-ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. J Biol Chem **275:**28201–28207.

56. **Mdluli K, Slayden RA, Zhu Y, Ramaswamy S, Pan X, Mead D, Crane DD, Musser JM, Barry CE 3rd.** 1998. Inhibition of a *Mycobacterium tuberculosis* beta-ketoacyl ACP synthase by isoniazid. Science **280:**1607–1610.

57. **Quémard A, Sacchettini JC, Dessen A, Vilcheze C, Bittman R, Jacobs WR Jr, Blanchard JS.** 1995. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. Biochemistry **34:**8235–8241.

58. **Yuan Y, Mead D, Schroeder BG, Zhu Y, Barry CE 3rd.** 1998. The biosynthesis of mycolic acids in *Mycobacterium tuberculosis*. Enzymatic methyl(ene) transfer to acyl carrier protein bound meromycolic acid *in vitro*. J Biol Chem **273:**21282–21290.

59. **Belisle JT, Vissa VD, Sievert T, Takayama K, Brennan PJ, Besra GS.** 1997. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. Science **276:**1420–1422.

60. **Puech V, Bayan N, Salim K, Leblon G, Daffé M.** 2000. Characterization of the *in vivo* acceptors of the mycoloyl residues transferred by the corynebacterial PS1 and the related mycobacterial antigens 85. Mol Microbiol **35:**1026–1041.

61. **Dhariwal KR, Yang YM, Fales HM, Goren MB.** 1987. Detection of trehalose monomycolate in *Mycobacterium leprae* grown in armadillo tissues. J Gen Microbiol **133:**201–209.

62. **Jackson M.** 2014. The mycobacterial cell envelope—lipids. Cold Spring Harb Perspect Med **4**(10)**:**a021105. doi: 10.1101/cshperspect.a021105

63. **Hunter SW, Brennan PJ.** 1990. Evidence for the presence of a phosphatidylinositol anchor on the lipoarabinomannan and lipomannan of *Mycobacterium tuberculosis*. J Biol Chem **265:**9272–9279.

64. **Khoo KH, Dell A, Morris HR, Brennan PJ, Chatterjee D.** 1995. Structural definition of acylated phosphatidylinositol mannosides from *Mycobacterium tuberculosis*: definition of a common anchor for lipomannan and lipoarabinomannan. Glycobiology **5:**117–127.

65. **Chatterjee D, Khoo KH.** 1998. Mycobacterial lipoarabinomannan: an extraordinary lipoheteroglycan with profound physiological effects. Glycobiology **8:**113–120.

66. **Torrelles JB, Khoo KH, Sieling PA, Modlin RL, Zhang N, Marques AM, Treumann A, Rithner CD, Brennan PJ, Chatterjee D.** 2004. Truncated structural variants of lipoarabinomannan in *Mycobacterium leprae* and an ethambutol-resistant strain of *Mycobacterium tuberculosis*. J Biol Chem **279:**41227–41239.

67. **Torrelles JB, Knaup R, Kolareth A, Slepushkina T, Kaufman TM, Kang P, Hill PJ, Brennan PJ, Chatterjee D, Belisle JT, Musser JM, Schlesinger LS.** 2008. Identification of *Mycobacterium tuberculosis* clinical isolates with altered phagocytosis by human macrophages due to truncated lipoarabinomannan. J Biol Chem **283:**31417–31428. doi: 10.1074/jbc.M806350200

68. **Besra GS, Morehouse CB, Rittner CM, Waechter CJ, Brennan PJ.** 1997. Biosynthesis of mycobacterial lipoarabinomannan. J Biol Chem **272:**18460–18466.

69. **Schaeffer ML, Khoo KH, Besra GS, Chatterjee D, Brennan PJ, Belisle JT, Inamine JM.** 1999. The pimB gene of *Mycobacterium tuberculosis* encodes a mannosyltransferase involved in lipoarabinomannan biosynthesis. J Biol Chem **274:**31625–31631.

70. **Geremia RA, Petroni EA, Ielpi L, Henrissat B.** 1996. Towards a classification of glycosyltransferases based on amino acid sequence similarities: prokaryotic α-mannosyltransferases. Biochem J **318**(Pt 1):133–138.

71. **Guerin ME, Korduláková J, Alzari PM, Brennan PJ, Jackson M.** 2010. Molecular basis of phosphatidyl-myo-inositol mannoside biosynthesis and regulation in mycobacteria. J Biol Chem **285:**33577–33583. doi: 10.1074/jbc.R110.168328

72. **Dianišková P, Korduláková J, Skovierová H, Kaur D, Jackson M, Brennan PJ, Mikušová K.** 2011. Investigation of ABC transporter from mycobacterial arabinogalactan biosynthetic cluster. Gen Physiol Biophys **30:**239–250. doi: 10.4149/gpb_2011_03_239

73. **Hunter SW, Brennan PJ.** 1981. A novel phenolic glycolipid from *Mycobacterium leprae* possibly involved in immunogenicity and pathogenicity. J Bacteriol **147:**728–735.

74. **Spencer JS, Brennan PJ.** 2011a. The role of *Mycobacterium leprae* phenolic glycolipid I (PGL-I) in serodiagnosis and in the pathogenesis of leprosy. Lepr Rev **82:**344–357.

75. **Spencer JS, Kim HJ, Wheat WH, Chatterjee D, Balagon MV, Cellona RV, Tan EV, Gelber R, Saunderson P, Duthie MS, Reece ST, Burman W, Belknap R, Mac Kenzie WR, Geluk A, Oskam L, Dockrell HM, Brennan PJ.** 2011b. Analysis of antibody responses to *Mycobacterium leprae* phenolic glycolipid I, lipoarabinomannan, and recombinant proteins to define disease subtype-specific antigenic profiles in leprosy. Clin Vaccine Immunol **18:**260–267.

76. **Spencer JS, Duthie MS, Geluk A, Balagon MF, Kim HJ, Wheat WH, Chatterjee D, Jackson M, Li W, Kurihara JN, Maghanoy A, Mallari I, Saunderson P, Brennan PJ, Dockrell HM.** 2012. Identification of serological biomarkers of infection, disease progression and treatment efficacy for leprosy. Mem Inst Oswaldo Cruz **107**(Suppl. I)**:**79–89.

77. **Hunter SW, Fujiwara T, Brennan PJ.** 1982. Structure and antigenicity of the major specific glycolipid antigen of *Mycobacterium leprae*. J Biol Chem **257:**15072–15078.

78. **Cho SN, Yanagihara DL, Hunter SW, Gelber RH, Brennan PJ.** 1983. Serological specificity of phenolic glycolipid I from *Mycobacterium leprae* and use in serodiagnosis of leprosy. Infect Immun **41:**1077–1083.

79. **Fujiwara T, Hunter SW, Cho SN, Aspinall GO, Brennan PJ.** 1984. Chemical synthesis and serology of disaccharides and trisaccharides of phenolic glycolipid antigens from the leprosy bacillus and preparation of a disaccharide protein conjugate for serodiagnosis of leprosy. Infect Immun **43:**245–252.

80. **Chatterjee D, Cho SN, Brennan PJ, Aspinall GO.** 1986. Chemical synthesis and sero-reactivity of *O*-(3,6-di-*O*-methyl-β-D-glucopyranosyl)-1→4)-*O*-(2,3-di-*O*-methyl-α-L-rhamnopyranosyl)-(1→9)-oxynonanoyl-bovine serum albumin—the leprosy-specific, natural disaccharide-octyl-neoglycoprotein. Carbohydr Res **156:**39–56.

81. **Brennan PJ, Chatterjee D, Fujiwara T, Cho SN.** 1994. Leprosy specific neoglycoconjugates: synthesis and application to serodiagnosis of leprosy. Methods Enzymol **242:**27–37.

82. **Ng V, Zanazzi G, Timpl R, Talts JF, Salzer JL, Brennan PJ, Rambukkana A.** 2000. Role of the cell wall phenolic glycolipid-1 in the peripheral nerve predilection of *Mycobacterium leprae*. Cell **103:**511–524.

83. **Rambukkana A.** 2001. Molecular basis for the peripheral nerve predilection of *Mycobacterium leprae*. Curr Opin Microbiol **4:**21–27.

84. **Rambukkana A, Zanazzi G, Tapinos N, Salzer JL.** 2002. Contact-dependent demyelination by *Mycobacterium leprae* in the absence of immune cells. Science **296:**927–931.

85. **Rambukkana A.** 2010. Usage of signaling in neurodegeneration and regeneration of peripheral nerves by leprosy bacteria. Prog Neurobiol **91:**102–107. doi: 10.1016/j.pneurobio.2009.12.002.

86. **Madigan CA, Cambier CJ, Kelly-Scumpia KM, Scumpia PO, Cheng TY, Zailaa J, Bloom BR, Moody DB, Smale ST, Sagasti A, Modlin RL, Ramakrishnan L.** 2017. A macrophage response to *Mycobacterium leprae* phenolic glycolipid initiates nerve damage in leprosy. Cell **170:**973–985.e10. doi: 10.1016/j.cell.2017.07.030.

87. **Azad AK, Sirakova TD, Fernandes ND, Kolattukudy PE.** 1997. Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. J Biol Chem **272:**16741–16745.

88. **Mathur M, Kolattukudy PE.** 1992. Molecular cloning and sequencing of the gene for mycocerosic acid synthase, a novel fatty acid elongating multifunctional enzyme, from *Mycobacterium tuberculosis* var. bovis Bacillus Calmette-Guerin. J Biol Chem **267:**19388–19395.

89. **Camacho LR, Constant P, Raynaud C, Laneelle MA, Triccas JA, Gicquel B, Daffe M, Guilhot C.** 2001. Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. J Biol Chem **276:**19845–19854.

90. **Eckstein TM, Silbaq FS, Chatterjee D, Kelly NJ, Brennan PJ, Belisle JT.** 1998. Identification and recombinant expression of a *Mycobacterium avium* rhamnosyltransferase gene (rtfA) involved in glycopeptidolipid biosynthesis. J Bacteriol **180:**5567–5573.

91. **Wilson VT, Cundliffe E.** 1998. Characterization and targeted disruption of a glycosyltransferase gene in the tylosin producer, *Streptomyces fradiae*. Gene **214:**95–100.

92. **Truman RW, Krahenbuhl JL.** 2001. Viable *M. leprae* as a research reagent. Int J Lepr Other Mycobact Dis **69:**1–12.

93. **Kirchheimer WF, Storrs EE.** 1971. Attempts to establish the armadillo (*Dasypus novemcinctus* Linn.) as a model for the study of leprosy. I. Report of lepromatoid leprosy in an experimentally infected armadillo. Int J Lepr Other Mycobact Dis **39:**693–702.

94. **Draper P.** 1980. Purification of *M. leprae*. Annex 4 of Report of the Fifth Meeting of the Scientific Working Group on the Immunology of Leprosy (IMMLEP), Geneva, 24 to 26 June 1980. World Health Organization document TDR/IMMLEP-SWG (5)/80.3. World Health Organization, Geneva, Switzerland.

95. **Engers HD, Abe M, Bloom BR, Mehra V, Britton W, Buchanan TM, Khanolkar SK, Young DB, Closs O, Gillis T, Harboe M, Ivanyi J, Kolk AHJ, Shepard CC.** 1985. Results

of a World Health Organization-sponsored workshop on monoclonal antibodies to *Mycobacterium leprae*. Infect Immun **48:**603–605. (Letter to the editor.)

96. **Engers HD, Houba V, Bennedsen J, Buchanan TM, Chaparas SD, Kadival G, Closs O, David JR, van Embden JDA, Godal T, Mustafa SA, Ivanyi J, Young DB, Kaufmann SHE, Khomenko AG, Kolk AHJ, Kubin M, Louis JA, Minden P, Shinnick TM, Trnka L, Young RA.** 1986. Results of a World Health Organization-sponsored workshop to characterize antigens recognized by mycobacterium-specific monoclonal antibodies. Infect Immun **51:**718–720. (Letter to the editor.)

97. **Khanolkar-Young S, Kolk AHJ, Andersen AB, Bennedsen J, Brennan PJ, Rivoire B, Kuijper S, McAdam KPWJ, Abe C, Batra HV, Chaparas SD, Damiani G, Singh M, Engers HD.** 1992. Results of the third immunology of leprosy/immunology of tuberculosis antimycobacterial monoclonal antibody workshop. Infect Immun **60:**3925–3927.

98. **Young RA, Bloom BR, Grosskinsky CM, Ivanyi J, Thomas D, Davis RW.** 1985. Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. Proc Natl Acad Sci USA **82:**2583–2587.

99. **Young RA, Mehra V, Sweetser D, Buchanan T, Clark-Curtiss J, Davis RW, Bloom BR.** 1985. Genes for the major protein antigens of the leprosy parasite *Mycobacterium leprae*. Nature **316:**450–452.

100. **Hunter SW, Rivoire B, Mehra V, Bloom BR, Brennan PJ.** 1990. The major native proteins of the leprosy bacillus. J Biol Chem **265:**14065–14068.

101. **Gillis TP, Miller RA, Young DB, Khanolkar SR, Buchanan TM.** 1985. Immunochemical characterization of protein associated with *Mycobacterium leprae* cell wall. Infect Immun **49:**371–377.

102. **Sonnenberg MG, Belisle JT.** 1997. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. Infect Immun **65:**4515–4524.

103. **Marques MAM, Chitale S, Brennan PJ, Pessolani MCV.** 1998. Mapping and identification of the major cell wall-associated components of *Mycobacterium leprae*. Infect Immun **66:**2625–2631.

104. **Marques MAM, Espinosa BJ, da Silveira EKX, Pessolani MCV, Chapeaurouge A, Perales J, Dobos KM, Belisle JT, Spencer JS, Brennan PJ.** 2004. Continued proteomic analysis of *Mycobacterium leprae* subcellular fractions. Proteomics **4:**2942–2953.

105. **Marques MAM, Neves-Ferreira AGC, da Silveira EKX, Valente RH, Chapeaurouge A, Perales J, Bernardes R da S, Dobos KM, Spencer JS, Brennan PJ, Pessolani MCV.**

2008. Deciphering the proteomic profile of *Mycobacterium leprae* cell envelope. Proteomics **8:**2477–2491.

106. **Geluk A, Duthie MS, Spencer JS.** 2011. Postgenomic *Mycobacterium leprae* antigens for cellular and serological diagnosis of *M. leprae* exposure, infection and leprosy disease. Lepr Rev **82:**402–421.

107. **Kim HJ, Prithviraj K, Groathouse N, Brennan PJ, Spencer JS.** 2013. Gene expression profile and immunological evaluation of unique hypothetical unknown proteins of *Mycobacterium leprae* by using quantitative real-time PCR. Clin Vacc Immunol **20:**181–190. doi:10.1128/cvi.00419-12.

108. **Wheeler PR.** 1982. Metabolism of carbon sources by *Mycobacterium leprae*: a preliminary report. Ann Microbiol (Paris) **133:**141–146.

109. **Wheeler PR.** 1983. Catabolic pathways for glucose, glycerol and 6-phosphogluconate in *Mycobacterium leprae* grown in armadillo tissues. J Gen Microbiol **129:**1481–1495.

110. **Wheeler PR, Gregory D.** 1980. Superoxide dismutase, peroxidatic activity and catalase in *Mycobacterium leprae* purified from armadillo liver. J Gen Microbiol **121:**457–464.

111. **Wheeler PR, Bharadwai VP, Gregory D.** 1982. N-acetyl-β-glucosaminidase, β-glucuronidase and acid phosphatase in *Mycobacterium leprae*. J Gen Microbiol **128:**1063–1071.

112. **Wheeler PR, Bharadwai VP.** 1983. Enzymes of malate oxidation in *Mycobacterium leprae* grown in armadillo livers. J Gen Microbiol **129:**2321–2325.

113. **Sritharan V, Wheeler PR, Ratledge C.** 1990. Aspartate metabolism in *Mycobacterium avium* grown in host tissue and axenically and in *Mycobacterium leprae*. J Gen Microbiol **136:**203–209.

114. **Wheeler PR.** 1990. Biosynthesis and scavenging of pyrimidines by pathogenic mycobacteria. J Gen Microbiol **136:**189–201.

115. **Wheeler PR.** 1989. Pyrimidine biosynthesis de novo in *M. leprae*. FEMS Microbiol Lett **48:**185–189.

116. **Wheeler PR, Bharadwaj VP, Katoch VM, Kannan KB.** 1984. Lactate dehydrogenase in *Mycobacterium leprae* grown in armadillo liver. Int J Lepr Other Mycobact Dis **52:**371–376.

117. **Wheeler PR, Ratledge C.** 1991. Phospholipase activity of *Mycobacterium leprae* harvested from experimentally infected armadillo tissue. Infect Immun **59:**2781–2789.

118. **Hall RM, Wheeler PR, Ratledge C.** 1983. Exochelin-mediated iron uptake into *Mycobacterium leprae*. Int J Lepr Other Mycobact Dis **51:**490–494.

119. **Wheeler PR, Bulmer K, Ratledge C.** 1990. Enzymes for biosynthesis de novo and elongation of fatty acids in mycobacteria grown in host cells: is *Mycobacterium leprae* competent in fatty acid biosynthesis? J Gen Microbiol **136:**211–217.

120. **Wheeler PR, Bulmer K, Ratledge C.** 1991. Fatty acid oxidation and the beta-oxidation complex in *Mycobacterium leprae* and two axenically cultivable mycobacteria that are pathogens. J Gen Microbiol **137:**885–893.

121. **Wheeler PR, Ratledge C.** 1988. Use of carbon sources for lipid biosynthesis in *Mycobacterium leprae*: a comparison with other pathogenic mycobacteria. J Gen Microbiol **134:**2111–2121.

122. **Wheeler PR.** 1989. Biosynthetic pathways in *Mycobacterium leprae*. Acta Leprol **7**(Suppl 1)**:**21–24. Review.

123. **Wheeler PR, Ratledge C.** 1988. Metabolism in *Mycobacterium leprae*, *M. tuberculosis* and other pathogenic mycobacteria. Br Med Bull **44:**547–561. (Review.)

124. **Wheeler PR.** 1986. Metabolism in *Mycobacterium leprae*: possible targets for drug action. Lepr Rev **57**(Suppl 3)**:**171–181.

125. **Wheeler PR.** 1984. Metabolism in *Mycobacterium leprae*: its relation to other research on *M. leprae* and to aspects of metabolism in other mycobacteria and intracellular parasites. Int J Lepr Other Mycobact Dis **52:**208–230.

126. **Ates LS, Brosch R.** 2017. Discovery of the type VII ESX-1 secretion needle? Mol Microbiol **103:**7–12. doi: 10.1111/mmi.13579.

127. **Gröschel MI, Sayes F, Simeone R, Majlessi L, Brosch R.** 2016. ESX secretion systems: mycobacterial evolution to counter host immunity. Nat Rev Microbiol **14:**677–691. doi: 10.1038/nrmicro.2016.131.

128. **Ates LS, van der Woude AD, Bestebroer J, van Stempvoort G, Musters RJ, Garcia-Vallejo JJ, Picavet DI, Weerd Rv, Maletta M, Kuijl CP, van der Wel NN, Bitter W.** 2016. The ESX-5 system of pathogenic mycobacteria is involved in capsule integrity and virulence through its substrate PPE10. PLoS Pathog **12:**e1005696. doi: 10.1371/journal.ppat.1005696.