MINI REVIEW

Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis

Dean C. Crick¹, Sebabrata Mahapatra, and Patrick J. Brennan

Department of Microbiology, Colorado State University, 200 W. Lake St., Fort Collins, CO 80523-1677, USA

Accepted on July 7, 2001

The compositional complexity of the mycobacterial cell envelope differentiates Mycobacterium species from most other prokaryotes. Historically, research in this area has focused on the elucidation of the structure of the mycobacterial cell envelope with the result that the structures of the mycolic acid-arabinogalactan-peptidoglycan complex from M. tuberculosis are fairly well understood. However, the current impetus for studying M. tuberculosis and other pathogenic mycobacteria is the need to identify targets for the development of new drugs. Therefore, emphasis has been shifting to the study of cell envelope biosynthesis and the identification of enzymes that are essential to the viability of M. tuberculosis. The publication of the complete M. tuberculosis genome in 1998 has greatly aided these studies. To date, thirteen enzymes involved in the synthesis of the arabinogalactan-peptidoglycan complex of M. tuberculosis have been identified and at least partially characterized. Eleven of these enzymes were reported subsequent to the publication of the M. tuberculosis genome, a clear indication of the rapid evolution of knowledge stimulated by the sequencing of the genome. In this article we review the current understanding of M. tuberculosis arabinogalactan-peptidoglycan structure and biosynthesis.

Key words: arabinan/isopentenyl diphosphate/Mycobacterium tuberculosis/peptidoglycan/polyprenyl phosphate

Mycobacterium tuberculosis cell envelope

Mycobacteria are classified as Gram-positive organisms, however, they have features of both Gram-positive and Gram-negative bacteria. The cell envelope of M. tuberculosis is made up of three major components: a plasma membrane, a covalently linked mycolic acid, arabinogalactan and peptidoglycan complex (MAPc), and a polysaccharide-rich capsule-like material. The unique nature of the MAPc led to the conclusion that the enzymes that synthesize this structure should yield a number of potentially unique drug targets (Crick and Brennan, 2000). Therefore, much research has focused on the elucidation of the structure of the MAPc, resulting in the publication of a multitude of recent review articles on the structure and function of the mycobacterial cell wall and capsule (Draper, 1982, Draper, 1998; Brennan and Draper, 1994; Brennan and Nikaido, 1995; Lee et al., 1996; McNeil et al., 1996; Besra and Brennan, 1997; Barry et al., 1998; Daffe and Draper, 1998; Baulard et al., 1999; McNeil, 1999).

The MAPc is made up of a cross-linked peptidoglycan, which is covalently linked to arabinogalactan (AG) chains via phosphoryl-N-acetylglucosaminosyl-rhamnosyl linkage units. The AG is in turn esterified to a variety of α-alkyl, β-hydroxy mycolic acids. The entire cell envelope is depicted schematically in Figure 1. The subsequent and more detailed figures show only the MAPc and do not include the noncovalently attached lipids, glycans, proteins and peptides. The model shown in Figure 1 is drawn with the peptidoglycan and galactan parallel to the plasma membrane. This orientation is consistent with traditional models of peptidoglycan (Ghuysen, 1968; McNeil, 1991), however, it should be noted that recent applications of molecular mechanics modeling indicate that it is possible that the peptidoglycan and AG strands are coiled and perpendicular to the plane of the plasma membrane (Dmitriev et al., 1999, 2000).

Known peptidoglycan classes have been identified by the type of peptide cross-linking that they demonstrate (Schleifer and Kandler, 1972), the peptidoglycan structure in M. tuberculosis is of the common A1γ type, although it does have some distinguishing features. The insoluble peptidoglycan consists of alternating units of N-acetylglucosamine (GlcNAc) and modified muramic acid residues. The typical N-acetyl functions on the muramic acid of bacterial peptidoglycan are further oxidized to N-glycolyl functions in M. tuberculosis. The muramic acid residues are also modified by tetrapeptide (L-alanyl-D-isoglutaminyl-meso-diaminopimelyl-D-alanine) side chains. Cross-linking can occur between two meso-diaminopimelic acid (DAP) residues as well as between DAP and D-alanine residues (Wietzerbin et al., 1974). Figure 2 is a representation of a small portion of M. tuberculosis peptidoglycan highlighting these features. The free carboxyl groups of the glutamate and DAP in the murein peptides from M. tuberculosis can be amidated in essentially any combination and a small percentage of the D-glutamate residues are substituted with glycine (Kotani et al., 1970).

It was known for many years that the major cell wall polysaccharide of mycobacteria is a branched-chain AG with the arabinose (Ara) residues forming the reducing termini. The AG was later shown to be attached to the muramic acid via a phosphodiester linkage (Lederer et al., 1975). More recently, it was confirmed that the polymer is unique in its elemental

¹To whom correspondence should be addressed
sugars (both the Ara and galactose [Gal] are in the furanose configuration; McNeill et al., 1987) and is comprised of a few distinct structural motifs (Figure 3). The galactan polymer is linked to the 6-position of some of the muramic acid residues of the peptidoglycan via an \( \alpha-L-Rha-(1\rightarrow3)-\alpha-D-GlcNAc-(1\rightarrowP) \) “linker region” (McNeil et al., 1990). The galactan itself is linear and consists of about 30 alternating 5- and 6-linked \( \beta-D-Galf \) residues (Daffe et al., 1990). The arabinan chains are attached to the 5 carbon of some of the 6-linked Galf residues (Daffe et al., 1990), apparently near the reducing end of the polymer (Besra et al., 1995), a conclusion based on the demonstration of the existence of large linear galactofurans. However, an alternate interpretation of the data could be that some of the galactan polymers are not arabinosylated, thus giving rise to large unbranched galactan chains. Most of the arabinan chains are composed of 5-linked \( \alpha-D-Araf \) with branching introduced by 3,5-\( \alpha-D-Araf \) residues. The nonreducing termini of the arabinan have a \( [\beta-D-Araf-(1\rightarrow2)-\alpha-D-Araf]-3,5-\alpha-D-Araf-(1\rightarrow5)-\alpha-D-Araf \) motif (Daffe et al., 1990; McNeil et al., 1994), and mycolic acids (long-chain \( \alpha \)-alkyl \( \beta \)-hydroxyl fatty acids) are located in clusters of four on two-thirds of the terminal arabinofuranosides (McNeil et al., 1990). Thus the primary structure of the arabinogalactan-peptidoglycan complex (APc) is quite well defined.

**Biosynthesis of the APc**

Research emphasis has recently been shifting from structural analysis to MAPc biosynthesis in mycobacteria. The complete sequencing of the *M. tuberculosis* genome (Cole et al., 1998) has provided a major impetus to the study of the enzymes of the organism. However, this resource has also resulted in over-confidence in our understanding of the biochemistry of *M. tuberculosis*. Many of the genes have an annotated function, but only a relatively small subset of these has had biochemical evidence of function demonstrated. Thus the annotations may provide the unwary with a false sense of functional insight. Therefore, in this review, only enzymes involved in APc synthesis for which there is evidence of function are listed (Table I). The enzymes and genes are referred to by their assigned names and by the unambiguous Rv number assigned by the genome project (Cole et al., 1998).

The following discussion of the biosynthesis of the APc is divided into three parts, involving the synthesis of polypropenyl phosphate (Pol-P), peptidoglycan, and AG. A discussion of mycolic acid biosynthesis is not included, as this topic is the subject of an excellent review by Barry et al. (1998).

**Pol-P synthesis**

It has long been known that Pol-P is involved in the biosynthesis of bacterial cell walls (Hemming, 1974), and Figure 4 makes it clear that the synthesis of many of the structures that are found in the *M. tuberculosis* MAPc are dependent on decaprenyl phosphate. The availability of Pol-P is rate limiting for cell wall synthesis in cell-free preparations from *Staphylococcus aureus* (Higashi et al., 1970) and *Bacillus* spp. (Anderson et al., 1972). In addition, it has been suggested that the rate of synthesis of lipid I (in peptidoglycan synthesis) of *Escherichia coli* may be dependent on the pool level of Pol-P (van Heijenoort, 1996). It has also been suggested that the rate of bacterial cell wall synthesis in vivo can be regulated by Pol-P levels (Baddiley, 1972), and we have recently shown that the slow-growing *M. tuberculosis* synthesizes Pol-P at much lower rates than the fast-growing *M. smegmatis* (Crick et al., 2000). Thus, the synthesis of the mycobacterial cell wall core can be thought to begin with the synthesis of decaprenyl phosphate.

However, despite these observations relatively little is known about Pol-P biosynthesis in bacteria. With the recent realization that isoprenoid biosynthesis in some bacteria differs from that in animals, the synthesis of bacterial isoprenoids has...
begun to generate considerable interest. In animals, isoprenoids are synthesized in two stages (Rip et al., 1985). First, isopentenyl diphosphate (IPP) is synthesized by condensation of acetyl CoA to form β-hydroxy-β-methyl-glutaryl CoA, which is reduced to mevalonate and subsequently decarboxylated to form the five-carbon molecule IPP. IPP is then isomerized to the allylic diphosphate dimethyl allyl diphosphate (DMAPP, Figure 5 left panel). Second, molecules of IPP are sequentially added to DMAPP to form geranyl diphosphate (GPP) and farnesyl diphosphate (FPP). FPP can then be used as reaction primers for the synthesis of longer chain Pol-P via a 1→4 head to tail condensation (Rilling, 1985).

Relatively less is known about prokaryotic isoprenoid biosynthesis. Until recently it was assumed that the early stages of isoprenoid biosynthesis described for animals would also apply to prokaryotes. However, several studies (Zhou, 1991; Horbach et al., 1993; Duvold et al., 1997; Sprenger et al., 1997; Lois et al., 1998; Putra et al., 1998a, Putra et al., 1998b) have clearly shown that some bacteria use an entirely different pathway for IPP biosynthesis (Figure 5 right panel). Duvold et al. (1997) proposed a hypothetical biogenic scheme in which pyruvate is condensed with glyceraldehyde 3-phosphate (GAP) to form 1-deoxyxylulose 5-phosphate, which is then converted to 2-C-methyl-D-erythritol 4-phosphate and finally to IPP. Deuterium-labeled DL-methyl-erythritol was incorporated into the ubiquinone of E. coli (Duvold et al., 1997), and it has been shown that the synthesis of 2-C-methyl-D-erythritol-4-phosphate by the enzyme encoded...
by the dxr gene is the second step in the pathway (Kuzuyama et al., 1998; Takahashi et al., 1998). CDP-linked intermediates appear to be involved in IPP synthesis (Rohdich et al., 1999; Kuzuyama et al., 2000a; Luttgen et al., 2000). The remaining reactions in the synthesis of IPP are largely unknown, although 2-C-methyl-D-erythritol 2,4 cyclodiphosphate has been reported to be derived from the CDP-linked intermediates (Herz et al., 2000). Lange and Croteau (1999) proposed that phosphorylation of isopentenyl phosphate may be the last reaction in the synthesis of IPP although the function of the kinase reported has been disputed (Kuzuyama et al., 2000b).

The mevalonate-independent pathway for the biosynthesis of IPP has been demonstrated in a number of bacteria, including Zymomonas mobilis, Methylobacterium fujisawaense, E. coli, Alicyclobacillus acidoterrestris (Rohmer et al., 1993), Corynebacterium ammoniagenes (Duyvold et al., 1997), and Mycobacterium phlei (Putra et al., 1998a). Interestingly, not all bacteria use the mevalonate-independent pathway to synthesize IPP. Myxococcus fulvus, Staphylococcus carnosus, Lactobacillus planarum, and Halobacterium cutirubrum all appear to utilize the eukaryotic IPP biosynthetic scheme (Horbach et al., 1993).

Searches of the M. tuberculosis genome database revealed two open reading frames (Rv2682c and Rv3379c) in the genome encoding proteins with 38% identity and 33% identity (respectively) to the E. coli enzyme 1-deoxyxylulose-5-phosphate synthase (DXS). Rv2682c, which encodes an enzyme responsible for this activity, has been cloned (Bailey et al., 2000) and successfully expressed as an active recombinant protein in E. coli. The enzymatic activity was exploited to synthesize [14C]labeled deoxyxylulose and [14C]1-deoxyxylulose 5-phosphate to use as tools for metabolic labeling. Radiolabeled deoxyxylulose was utilized by M. smegmatis for the synthesis of a compound that had the chromatographic characteristics of menaquinone, confirming that mycobacteria utilize the nonmevalonate pathway for the synthesis of isoprenoids. However, the remaining genes involved in IPP synthesis in M. tuberculosis have still to be identified and the enzyme activities biochemically defined.
In eukaryotes once IPP has been synthesized some of it is isomerized to form DMAPP (Figure 5 left panel). In bacteria that utilize the mevalonate-independent pathway, it is still unclear how DMAPP is formed as the preponderance of evidence indicates that an isomerase may be involved but is not essential (Hahn et al., 1999; Charon et al., 2000; Rodriguez-Concepcion et al., 2000). Subsequently new molecules of IPP are sequentially added to the growing chain via a 1’-4 head to tail condensation (Rilling, 1985) until an FPP or a GGPP is the expected product (Figure 6). Once the appropriate chain length has been achieved, the prenyl phosphates undergo dephosphorylation to form Pol-P, although it is not clear whether this is a simple one-step dephosphorylation or whether both phosphates are removed and the resulting free alcohol is then rephosphorylated by a kinase.

The most common prenol (and therefore Pol-P) structures tend to be confined to four main groups as follows: (1) all-E-prenol, (2) di-E, poly-Z-prenol, (3) tri-E, poly-Z-prenol, and (4) all-Z-prenol (Anonymous, 1987). Bacteria typically contain a single predominant Pol-P composed of 11 isoprene units, with the di-E, poly-Z configuration as seen in undecaprenyl phosphate (bactoprenyl phosphate). *M. smegnatis* is known to contain two forms of Pol-P, a decaprenyl phosphate and a heptaprenyl phosphate (Takayama et al., 1973); these are structurally unusual in that the decaprenyl phosphate (Figure 6) contains one omega-, one E-, and eight Z-isoprene units (mono-E, poly-Z) (Wolucka et al., 1994) and the heptaprenyl phosphate consists of four saturated isoprene units on the omega end of the molecule and two E- and one Z-isoprene units (Besra et al., 1994) or four saturated and three Z-isoprene units (Wolucka and de Hoffman, 1998). It appears that *M. tuberculosis* may be more typical than *M. smegnatis* as a single predominant Pol-P (decaprenyl phosphate) was identified in this species, however, the stereochemistry of the individual isoprene units was not determined (Takayama and Goldman, 1970).

Recent experiments on the biosynthesis of Pol-P in mycobacteria show that *M. tuberculosis* synthesizes decaprenyl diphosphate but not heptaprenyl diphosphate (Crick et al., 2000), thus agreeing with the earlier structural data. The synthetic data suggest that geranyl diphosphate is the allylic substrate for two distinct prenyl diphosphate synthases, one located in the cell membrane that synthesizes α,E,Z-farnesyl diphosphate (E,Z-FPP) and the other in the cytosol that synthesizes α,E,E-farnesyl diphosphate (E,E-FPP) or α,E,E,E-geranylgeranyl diphosphate (E,E,E-GGPP). *M. tuberculosis* open reading frame Rv1086 encodes an α,E,Z-FPP synthase (Z-FPPS) (Schulbach et al., 2001). This enzyme is the first short chain Z-isoprenyl diphosphate synthase for which an amino acid sequence has been reported and adds a single isoprene unit to geranyl diphosphate in the Z-configuration. The cytosolic E,E-FPP synthase may be involved in mycobacterial sterol synthesis, which was recently reported (Lamb et al., 1998).

The *M. tuberculosis* open reading frame Rv2361c encodes a decaprenyl diphosphate synthase (Schulbach et al., 2000). The stereochemistry of each isoprene addition by this synthase has not been unequivocally established. However, based on amino acid sequence homology between Rv2361c and several known Z-isoprenyl diphosphate synthases (including Rv1086), it is safe to assume that each isoprene addition is in the Z-configuration. If the E,Z-FPP is the substrate for the enzyme encoded by Rv2361c a decaprenyl diphosphate having the mono-E, poly-Z configuration would be the expected product (Figure 6), a structure similar to that shown for the decaprenyl phosphate found in *M. smegnatis*. The subsequent step(s) necessary to form decaprenyl phosphate from decaprenyl diphosphate have, as yet, not been described for any *Mycobacterium* species.

### Peptidoglycan synthesis

Very little is known about the biosynthesis of the peptidoglycan of *M. tuberculosis* per se, however, it is generally assumed to be similar to that of *E. coli*. Peptidoglycan synthesis in *E. coli* has been reviewed by van Heijenoort (van Heijenoort, 1997).
D.C. Crick et al.

1996, van Heijenoort, 2001) and can be divided into four groups of sequential reactions. The initial reactions involve the synthesis of UDP-N-acetyl muramic acid (UDP-MurNAc) from UDP-GlcNAc by two reactions. First, enoylpyruvate (from phosphoenoylpyruvate) is added to the 3 position of the GlcNAc residue of the UDP-GlcNAc and subsequently the enoylpyruvate moiety is reduced to form UDP-MurNAc (Figure 7). MurA and MurB, respectively, catalyze these reactions. The murA gene of M. tuberculosis (Rv1315) has been overexpressed and partially characterized (De Smet et al., 1999). In this study, wild-type MurA in M. tuberculosis was found to have an aspartate at position 117 instead of cysteine as found in E. coli. This modification was demonstrated to confer resistance to the antibiotic fosfomycin in M. tuberculosis (De Smet et al., 1999).

UDP-MurNAc-pentapeptide is formed by the addition of L-alanine to the lactate moiety of the UDP-MurNAc, and the subsequent sequential addition of D-glutamic acid, DAP and a D-alanyl-D-alanine dipeptide to form UDP-MurNAc(penta-peptide). In E. coli MurC, MurD, MurE, and MurF catalyze these reactions (respectively). The MurC (L-alanine ligase, Rv2155c) of M. tuberculosis has also been overexpressed in active form (unpublished data) and is

Fig. 4. Biosynthetic pathway for mycolic acid-arabinogalactan-peptidoglycan complex (MAPc) of M. tuberculosis. This hypothetic pathway is based on the current understanding of the structure of the MAPc and the activities of known enzymes (indicated in italics with Rv numbers) expressed by M. tuberculosis. It should be noted that many of the enzymes involved in the synthesis of mycolic acids are also known but are not indicated here.
enzymatically similar to MurD from *E. coli*. All of the other enzymatic activities of this pathway await future investigation.

In *E. coli* Lipid I is formed by the activity of MurX (previously called MraY, Figure 7) which transfers phosphoryl-MurNAc(peptapeptide) to a molecule of undecaprenyl phosphate and a GlcNAc residue is subsequently added to form GlcNAc-MurNAc(pentapeptide)-diphosphoryl-undecaprenol (Lipid II) by MurG. The glycopeptide of Lipid II is thought to be translocated from the cytoplasmic face of the inner cell membrane to the periplasmic face, where it is used directly in the assembly of the peptidoglycan. Assembly is assumed to occur in the periplasm by a transglycosylation reaction that extends a glycan chain by the addition of the disaccharide pentapeptide with the release of undecaprenyl diphosphate.

The glycan chains are cross-linked by the formation of peptide cross bridges. The main cross-linkage in *E. coli* is between the penultimate D-alanine residue and the DAP residue of an adjacent peptide of a second glycan chain. The reaction is catalyzed by penicillin-binding proteins (PBPs) and involves the cleavage of the D-alanyl-D-alanine bond of the donor peptide, which provides the energy to drive the reaction (Ghuysen, 1991). The PBPs are classified primarily according to their molecular mass, and have penicilloyl serine transferase activity that catalyzes the cleavage of the cyclic amide bond of penicillin. The high molecular mass (HMM) PBPs are multi-modular enzymes with a N-terminal non–penicillin-binding module and a C-terminal penicillin-binding module. These PBPs are further subdivided into Class A and Class B based on their primary structures. The HMM Class A PBPs function as transglycosylases as well as transpeptidases. Chambers *et al.* (1995) detected four PBPs, including three of HMM in *M. tuberculosis* membranes, however, the transglycosylase or transpeptidase activity of these putative enzymes has not been demonstrated.

It has been reported (Wietzerbin *et al.*, 1974) that two-thirds of the peptide cross-links found in *M. tuberculosis* peptidoglycan are between the carboxyl group of a terminal D-alanine and the amino group at the D-center of DAP resulting a D,D bond (Figure 2). The remaining third of the linkages are between the carboxyl group at the L-center of a terminal DAP of one chain and the amino group at the D-center of the DAP of another peptide side chain, resulting in an L,D-cross-link. Similar L,D-peptide cross-links have been reported in stationary phase *E. coli* by Templin *et al.* (1999). These observations suggest the presence of both D,D and L,D-transpeptidase activities in *M. tuberculosis* and *E. coli*, however, no L,D-transpeptidases have been identified to date.

Thus, with the exceptions of the amidation of free carboxyl groups and glycolylation of the muramic acid residues, it is reasonable to assume that murein synthesis in *M. tuberculosis* is similar to that seen in *E. coli*. It is unknown at what stage these modifications occur in the biosynthetic scheme of mycobacterial peptidoglycan and their physiological significance is

---

**Fig. 5.** A comparison of the mevalonate-dependent and -independent isopentenyl diphosphate synthetic pathways.
unclear. It is known, however, that peptidoglycan from *M. tuberculosis* is more resistant to degradation by enzymes such as lysozyme than peptidoglycan from *E. coli*.

### AG synthesis

The synthesis of the AG proper in *M. tuberculosis* begins with the transfer of a GlcNAc-1-phosphate from UDP-GlcNAc to prenyl phosphate followed by an addition of rhamnose (Rha) from dTDP-Rha, thus forming the “linker region” of the AG (Figure 3) as shown in *M. smegmatis* (Mikusova et al., 1996). The enzyme that catalyzes the transfer of the GlcNAc 1-phosphate has yet to be biochemically identified in mycobacteria, although there is at least one gene that has significant homology to known GlcNAc-1-phosphotransferases (Dal Nagare et al., 1998).

**RmlA** (Rv0334), **RmlB** (Rv3464), **RmlC** (Rv3465), and **RmlD** (Rv3266c) are responsible for the synthesis of dTDP-Rha from glucose-1-phosphate (Ma et al., 2001; McNeil, 1999) encoding an α-D-glucose-1-phosphate thymidyl transferase (Ma et al., 1997), dTDP-D-glucose 4,6-dehydratase (Ma et al., 2001), dTDP-4-keto-6-deoxy-D-glucose 3,5 epimerase (Stern et al., 1999) and dTDP-rhamnose synthetase (Hoang et al., 1999) respectively. The *M. tuberculosis wbbL* gene (Rv3265c) was shown to complement an *E. coli* mutant deficient in Wbbl that was consequently unable to synthesize O-antigen (McNeil, 1999). Therefore it is clear that Rv3265c encodes the rhamnosyl transferase that adds Rha to the prenyl diphosphoryl GlcNAc.

UDP-galactofuranose (UDP-Galf) is the donor of the gal/ residues of the galactan (Weston et al., 1997). The first reaction in the synthesis of UDP-Galp is the conversion of UDP-glucose to UDP-galactopyranose (UDP-Galp) catalyzed by UDP-galactopyranose epimerase. The gene that encodes this enzyme in *M. tuberculosis* has been tentatively identified as Rv3634 (Weston et al., 1997). The enzyme encoded by glf then converts UDP-Galp to UDP-Galf, which is the donor of the gal residues of the galactan in *M. tuberculosis* (Weston et al., 1997). The *M. tuberculosis* glf is very similar to the glf enzymes from *E. coli* (Nassau et al., 1996) and Klebsiella pneumoniae (Koplin et al., 1997).

The synthesis of a prenyl diphosphate linked oligosaccharide containing GlcNAc, Rha, Gal, and Ara in cell-free preparations from *M. smegmatis* has been demonstrated (Mikusova et al., 2000). Therefore it is believed that subsequent to the formation of the linker region, the Galp and Ara residues are added to form the mature lipid-linked AG, which is then transferred to peptidoglycan en bloc by a ligase, which releases decaprenyl phosphate.

Based on the structure of the galactan portion of the AG one would predict that there are two to four Galp transferases. One that adds the initial Galp residue to the linker region, possibly a specific enzyme to add the second Galp residue and either one enzyme for each of the 1→5 and 1→6 linkages or a dual function enzyme that catalyzes the formation of both types of linkage. Mikusova et al. (2000) first cloned and overexpressed a Galp transferase gene (Rv3808c) from *M. tuberculosis*, the data presented indicated that the enzyme catalyzed the formation of both 1→5 and 1→6 linkages. This observation has since been substantiated and extended using assays based on chemically synthesized neoglycolipid acceptors (Kremer et al., 2001).

The Ara residues of the arabian originate from the pentose phosphate pathway/hexose monophosphate shunt (Scherman et al., 1995, Scherman, 1996) and the immediate precursor of the polymerized Ara is decaprenylphosphoryl-Araf (Wolucka et al., 1994). Given the complexity of the arabinofuran found in the AG a number of arabinosyl transferases are expected to be found. As yet, no arabinosyl transferases have been unequivocally identified in *M. tuberculosis*. The proteins encoded by the embA and embB genes have been reported to function as an Araf transferase in *Mycobacterium avium* (Belanger et al., 1996). However, this activity remains to be demonstrated for the enzymes encoded by the homologous genes from *M. tuberculosis*.

### Summary

The structures of the APc from *M. tuberculosis* have largely been worked out, and each demonstrates a number of unusual characteristics. To date, 13 enzymes involved in the synthesis of the APc of *M. tuberculosis* have been identified and at least partially characterized (Table I). These enzymes are also indicated in italics on the biosynthetic pathway leading to the formation of the APc shown in Figure 4. Overall, this represents a significant portion, perhaps as much as a third, of the total enzymes involved.
The biosynthesis of peptidoglycan is quite well understood in *E. coli*, and there is every reason to believe that peptidoglycan synthesis in *M. tuberculosis* is similar although the enzymology and essentiality of the unique features of *M. tuberculosis* murein remain unexplored. The donors of the sugars involved in the AG synthesis have been identified. All of the genes involved in the formation of TDP-Rha and UDP-Gal are known. The enzyme that uses TDP-Rha to synthesize the linker region is known and a Galf transferase has been identified.

The major impetus for studying the cell walls of *M. tuberculosis* and other pathogenic mycobacteria is the need to identify targets for the development of new drugs. Therefore, identification of which enzymes are essential to the viability of *M. tuberculosis* is of paramount importance. However, many other significant questions also remain unanswered. It is not yet clear what the complete biosynthetic pathway for IPP in mycobacteria is, although we now understand the early steps. The nature of the enzymes that synthesize arabinosyl phosphoryl decaprenol remain unknown as does the nature and identity of the genes encoding the Araf transferases and additional Galf transferases.

Nothing is known about the enzyme(s) that ligate the AG to the peptidoglycan. Another issue that remains to be examined is the physical organization of the biosynthetic enzymes. What is their topology? Are there multienzyme complexes? In addition, it is unclear if or how these synthetic enzymes and/or complexes are regulated. The nature of the early precursors (nucleotide-sugars, amino acids, malonyl CoA, etc.) of the peptidoglycan, AG, and mycolic acids suggest that they originate in the cytoplasm, therefore, components of the mature structures must be transported across the cell membrane. How is this accomplished, and what is the nature of the components that do move across the membrane? In a related question, does the mycolic acid layer enclose a functional periplasm in which cell wall core assembly can take place? What is the nature of the ultrastructure of the cell wall, do the strands of peptidoglycan lie parallel to the cell wall as traditionally thought, or do they form perpendicular coils as suggested by molecular modeling? A picture of *M. tuberculosis* APc synthesis is beginning to take shape; however, many more pieces must be added before the puzzle is complete.
Acknowledgments

We thank Dr. Michael McNeil for stimulating discussions and careful scrutiny of Figures 2 and 3. Work from the authors' laboratories was funded by grants AI-18357 and AI-49151 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, and by grant AI-46393 under the National Cooperative Drug Discovery Groups–Opportunistic Infections program from the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Abbreviations

AG, arabinogalactan; APC, arabinogalactan-peptidoglycan complex; Ara, arabinose; Araf, arabinofuranose; DAP, d-meso-diaminopimelic acid; DMAPP, dimethylallyl diphosphate; DXS, 1-deoxyxylulose 5-phosphate synthase; E,E-FPP, α,α,E,E-farnesyl diphosphate; E,Z-FPP, α,α,E,Z-farnesyl diphosphate; E,E,E-GGPP, α,α,E,E-geranylgeranyl diphosphate; Gal, galactofuranose; GalA, galactopyranose; GAP, D-glyceraldehyde 3-phosphate; GlcNAc, N-acetylgalactosamine; GPP, geranyl diphosphate; HMM, high molecular mass; IPP, isopentenyl diphosphate; MAFPc, mycolyl-arabinogalactan-peptidoglycan complex; MurNAc, N-acetylmuramic acid; PBP, penicillin-binding protein; Pol-P, polyprenyl phosphate; Rha, rhamnose; Rhap, rhamnopyranose; TDP, thymidine diphosphate; UDP, uridine diphosphate; Z-FPPS, α,α,E,Z-farnesyl diphosphate synthase.

References


Higashi, Y., Siewert, G., and Strominger, J.L. (1970) Biosynthesis of the peptidoglycan of bacterial cell walls. XIX. Isoprenoid alcohol phospho-

D.C. Crick et al.

116R


