Chemistry & Biology, Vol. 9, 545–553, May, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S1074-5521(02)00142-4

The Methyl-Branched Fortifications of *Mycobacterium tuberculosis*

Review

David E. Minnikin,[1,3] Laurent Kremer,[2] Lynn G. Dover,[1] and Gurdyal S. Besra[1]
[1]School of Biosciences
The University of Birmingham
Edgbaston, Birmingham B15 2TT
United Kingdom
[2]INSERM U447
Institut Pasteur de Lille
1 rue du Pr. Calmette
BP245-59019 Lille Cedex
France

Mycobacterium tuberculosis **continues to be the predominant global infectious agent, annually killing over three million people. Recommended drug regimens have the potential to control tuberculosis, but lack of adherence to such regimens has resulted in the emergence of resistant strains.** ***Mycobacterium tuberculosis*** **has an unusual cell envelope, rich in unique long-chain lipids, that provides a very hydrophobic barrier to antibiotic access. Such lipids, however, can be drug targets, as exemplified by the action of the front-line drug isoniazid on mycolic acid biosynthesis. A number of these lipids are potential key virulence factors and their structures are based on very characteristic methyl-branched long-chain acids and alcohols. This review details the history, structure, and genetic aspects of the biosynthesis of these methyl-branched components, good examples of which are the phthiocerols and the mycocerosic and mycolipenic acids.**

Introduction

Tuberculosis (TB) remains a terrible scourge in many parts of the world and annually kills over three million people [1]. This fact clearly establishes *Mycobacterium tuberculosis* as the predominant global infectious agent. Two decades ago, it was generally considered that TB, and indeed the closely related mycobacterial disease of leprosy, were under control and that it was just a matter of time before they faded into history. Throughout the 1980s and 1990s, however, both TB and leprosy have been encouraged to reestablish themselves due to a number of factors. Regarding leprosy, it has been mainly a matter of problems in ensuring effective medical and nursing care. Compared with TB, it is much more difficult to catch leprosy, its main transmission route being through repeated bodily contact rather than an unselective aerosol mode. Therefore, expectations of bringing leprosy under control remain high, even though it is currently impossible to cultivate the causative agent, *Mycobacterium leprae*, and to directly study its biochemistry in detail. Alternatively, genomic comparison between the leprosy and tubercle bacilli offers an indirect approach for the development of new drugs and vaccines against leprosy [2, 3]. Thus, it is highly likely that any new successful drugs against common biochemical pathways in *M. tuberculosis* and *M. leprae* may also have an effect on *M. leprae*. However, it must be stated that different drug regimens have been required, so far, to combat these two pathogens.

The resurgence of TB springs mainly from inadequate drug compliance, the appearance of drug-resistant strains, and the facile infection of victims of the AIDS epidemic. The response by the World Health Organization (WHO) to the compliance problem has been the recommendation of "directly observed therapy" (DOT), an ambitious program in which patients are monitored individually. Drug-resistant strains appear to have arisen following inadequate adherence to recommended drug regimens; this is particularly serious in certain parts of the world as widespread as Eastern Europe, Central Africa, and pockets of USA. The only real solution to avoid AIDS patients succumbing to TB is to eradicate the viral condition, and this is bound to be a long haul, depending heavily on education and political directives.

It is currently postulated that the distinctive pathogenicity of *M. tuberculosis* can be correlated with the presence of an unusual cell envelope, rich in unusual polysaccharides and lipid components, as will be detailed below. To accelerate progress in the battle against TB, it is imperative that a full understanding of the structure, function, and biosynthesis of the mycobacterial cell envelope is attained without delay. The determination of the complete genome of *M. tuberculosis* [2] provides access to the full range of biosynthetic pathways leading to the assembly of the cell envelope. This review will highlight the structure and biosynthesis of multimethyl-branched fatty acids and methyl-branched long-chain diols in *M. tuberculosis*.

The Lipid Components of *M. tuberculosis*

The unusual lipophilic nature of the cell envelope of the tubercle bacillus was clearly indicated by pioneering studies in the 1930s coordinated by R.J. Anderson at Yale [4]. It was shown that high proportions of unique long-chain compounds were present, and the so-called mycolic acids were characterized for the first time. In the 1950s and 1960s, the essential structural details of these lipids were determined in the laboratories of E. Lederer (Paris), J. Asselineau (Paris and Toulouse), and N. Polgar (Oxford), as reviewed by Asselineau [5], Goren and Brennan [6], and Minnikin [7]. Research during the past thirty years has consolidated knowledge of the structure of the lipids of *M. tuberculosis*; details of the studies leading to these structures can be found in recent reviews [8, 9]. The principal structures of complex lipids from the *M. tuberculosis* complex are introduced in Figures 1 and 2, and a model for their possible interaction in the cell envelope is shown in Figure 3; further details will be discussed below.

The mycolic acids are the key lipid components in *M. tuberculosis*, and their essential structures are repre-

[3]Correspondence: d.e.minnikin@bham.ac.uk

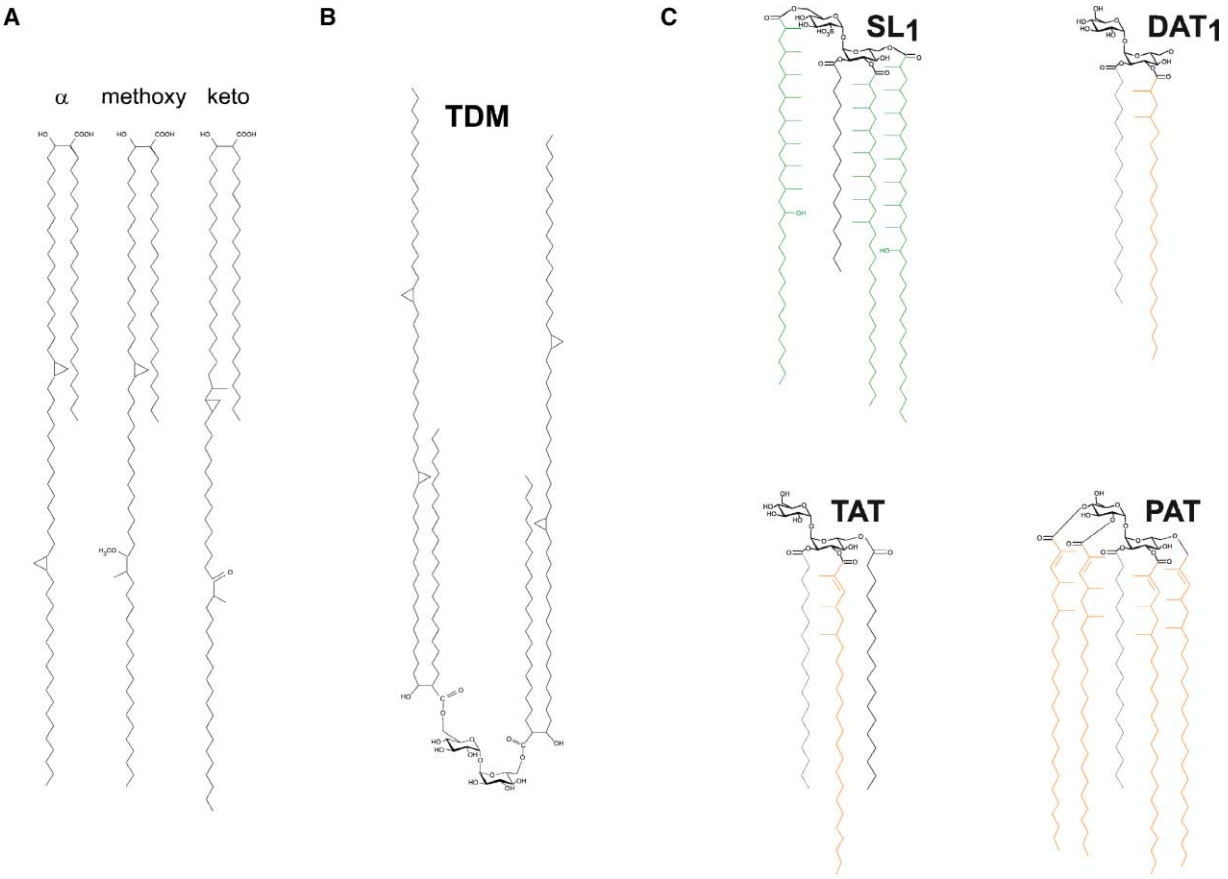

Figure 1. Structures of *M. tuberculosis* Mycolic Acids and Glycolipids, Based on Trehalose

(A) Schematic representation of the α, methoxy-, and ketomycolic acids from *M. tuberculosis*.

(B) Trehalose dimycolate (TDM) from *M. tuberculosis*.

(C) Complex glycolipids from *M. tuberculosis*. Sulfated tetra-acyl trehalose (SL_1) based on phthioceranic and hydroxyphthioceranic acids, diacyl trehaloses (DAT_1) based on mycosanoic acids, triacyl trehalose (TAT) and pentaacyl trehalose (PAT) based on mycolipenic acids.

sented in Figure 1A. Mycolic acids, with related structures, are also present in all mycobacteria and related actinomycetes [10]. Although the structures of these unique lipids have been well characterized, the enzymology and genetics leading to their biosynthesis have only recently begun to be elucidated [11]. As shown later, mycolic acids are predominately covalently bound to an arabinogalactan polysaccharide; some, however, are found as trehalose monomycolates (TMM) and dimycolates (TDM) (Figure 1B); TMM is involved in the transfer of mycolic acids into the cell wall [12].

M. tuberculosis is characterized by a range of complex glycolipids based on trehalose (Figure 1C). A family of sulfated acyl trehaloses (SL) were characterized in early studies by Goren and coworkers [6, 13], the principal glycolipid being termed SL_1 (Figure 1C). A pair of closely related diacyl trehaloses (DAT_1 and DAT_2) were isolated from *M. tuberculosis* [14], and DAT_1 was shown to have the structure in Figure 1C, based on a mycosanoic acid (Figure 4). The more polar DAT_2 glycolipid had a mycolipanolic acid (Figure 4) in place of the mycosanoate. These lipids cooccurred with more heavily acylated glycolipids [15], whose main component was shown [16] to be a pentaacyl trehalose (PAT) (Figure 1C). Some strains also had related triacyl trehaloses (TATs), with the possible structure shown in Figure 1C [17].

Other members of the *M. tuberculosis* complex can have alternative characteristic glycolipids antigens. *Mycobacterium bovis*, *Mycobacterium africanum*, *Mycobacterium microti*, and the attenuated *Mycobacterium bovis* BCG variant produce a phenolic glycolipid (Figure 2A), the so-called mycoside B [5–7, 18]. The unusual *M. tuberculosis* "Canetti" strains, with smooth colony morphology, are characterized by the presence of unique phenolic glycolipids [19] and lipo-oligosaccharides [20]. Two *M. tuberculosis* clinical isolates, labeled *O* and *K*, also produced phenolic glycolipids, although they had the usual rough morphology [21]. In this survey, attention will be restricted to lipids of the more usual strains of *M. tuberculosis* and *M. bovis*.

The dimycocerosates of the phthiocerol family (PDIMs) (Figure 2B) are major waxes, related to the phenolic glycolipids, which are found in most representatives of *M. tuberculosis* and some other related mycobacteria. These waxes (Figure 2B) and glycolipids based on trehalose (Figure 1C) are characterized by the presence of a

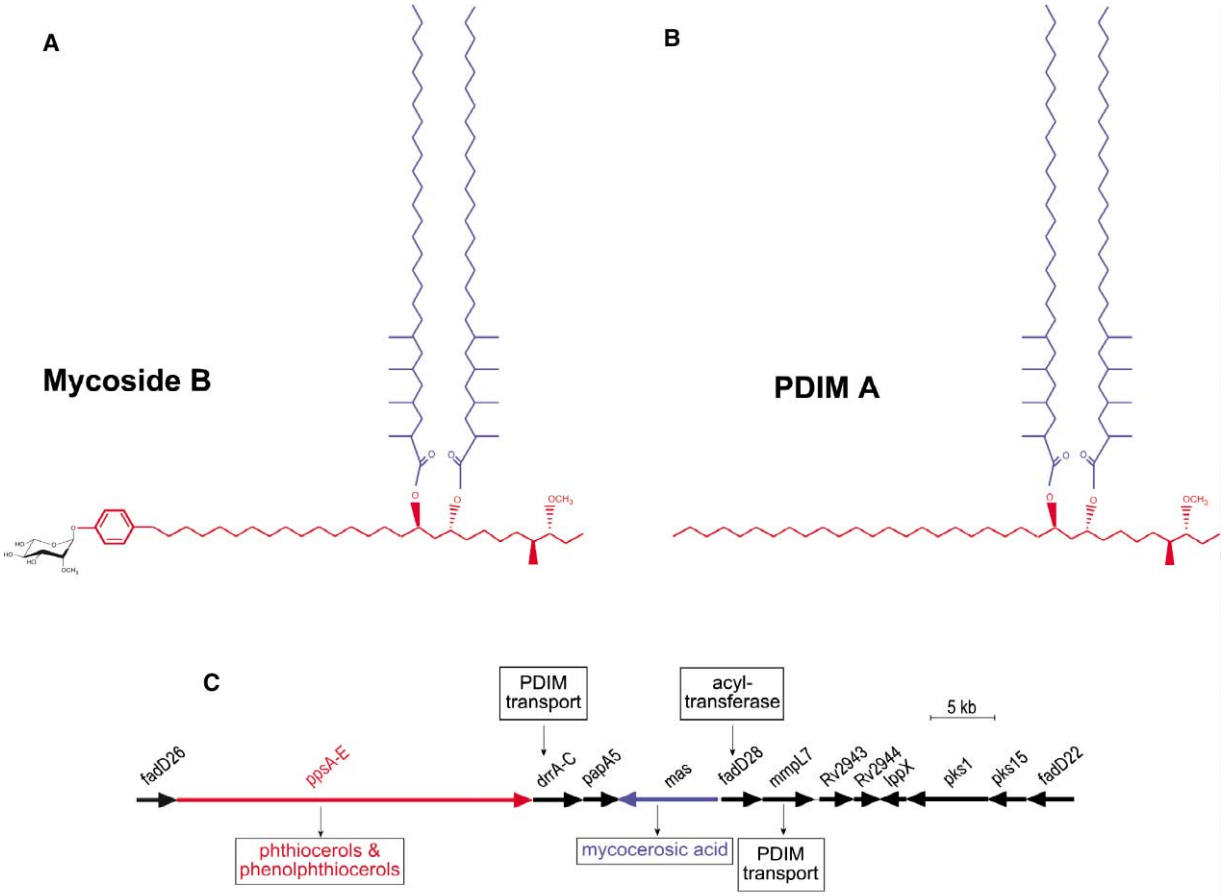

Figure 2. Structures of *M. tuberculosis* Complex Mycoside B, Phthiocerol Dimycocerosate, and Related Genomics

(A) Mycoside B (PGL) from *M. bovis* BCG.

(B) Phthiocerol dimycocerosate A (PDIM A) from *M. tuberculosis*.

(C) The genomic organization of surrounding open reading frames that produce and transport PGL/PDIM.

range of multimethyl-branched fatty acids, and a detailed consideration of these long-chain components is the main subject of this review. Lipids containing multimethyl-branched fatty acids were proposed [7] to interact with a covalently bound monolayer of mycolic acids to provide a characteristic mycobacterial outer-membrane organelle. Subsequent studies [22, 23] have supported this hypothesis, and direct evidence for two distinct envelope lipid domains has been given by preferential location of fluorescent probes of different lipophilicity [24]. The essence of the cell envelope organization is provided in Figure 3. Briefly, a relatively conventional plasma membrane and peptidoglycan is connected by a linker unit (LU) to an arabinogalactan, which bears the mycolic acids on some of its arabinose termini. Predominant components in the cell envelope of *M. tuberculosis* are the lipoarabinomannan (LAM) and related lipomannan (LM), which is not included in Figure 3. LAM and LM are lipopolysaccharides based on a phosphatidylinositol mannoside (PIM) anchors, which may locate into the plasma membrane (Figure 3).

The multimethyl-branched fatty acids from *M. tuberculosis* were characterized over an extended period following the pioneering studies of Anderson [4]. Prolonged hydrolysis of resistant waxes gave laevorotatory "mycocerosic" acids [25]; the trimethyl- and tetramethyl-branched (Figure 4) structures were elaborated by Marks and Polgar [26]. The principal components of Anderson's dextrorotatory "phthioic" acids [4] were shown by Chanley and Polgar [27] to be α,β-unsaturated "mycolipenic" acids (Figure 4), a finding echoed by Cason and Sumrell [28], who named them "phthienoic" acids. The same extracts yielded lesser proportions of "mycolipodienic" and "mycolipanolic" (Figure 4) acids [29, 30]. Related "mycosanoic" acids (Figure 4) were characterized by Cason et al. [31]. Acylated trehalose sulfates were found by Goren et al. [6, 13] to be based on dextrorotatory heavily branched "phthioceranic" (Figure 4) and "hydroxyphthioceranic" (Figure 4) acids.

The name "phthiocerol" was given to a long-chain methoxyglycol by Stodola and Anderson [32]. Key degradative studies by Hall and Polgar [33] lead to the essential structure of phthiocerol A (Figure 5) as 3-methoxy-4-methyl do- (and tetra)-triacontane-9,11-diol [34, 35]

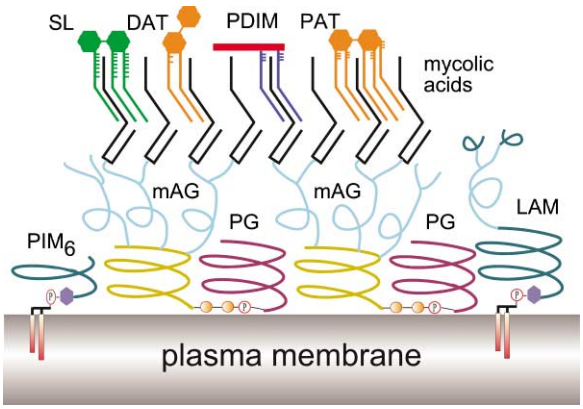

Figure 3. Arrangement of Structural Components in the Cell Envelope of *M. tuberculosis*, Showing Possible Interaction of Methyl-Branched Long-Chain Components with a Mycolic Acid Matrix

Mycolyl arabinogalactan (mAG) is connected by a phosphoryl linker unit to peptidoglycan (PG). Complex free lipids (PDIM, DAT, PAT, SL; see Figures 1 and 2) interact with mAG. Lipoarabinomannan (LAM) and phosphatidylinositol pentamannoside (PIM_5) are shown anchored in the plasma membrane. The mAG galactan is shown in yellow, and the LAM mannose components are in dark green; the arabinan of both these polysaccharides is represented in light blue. Modified from references [7] and [22]. The figure is principally designed to indicate the possible location and interactions of the methyl-branched components; the outermost carbohydrate/protein-rich layer [23] is omitted for clarity.

from the PDIM waxes (Figure 2B). Compounds related to phthiocerol A were thoroughly characterized as phthiocerol B (Figure 5) and phthiodiolone A (Figure 5) by Minnikin and Polgar [36, 37]. Smith and coworkers [38] detected a glycolipid from *M. bovis* that was shown [39] to be based on a phenolphthiocerol unit (Figure 5), 2-*O*-methylrhamnose, and mycocerosic acids (Figure 2A). As mentioned earlier, this glycolipid was termed mycoside B, the name mycoside being defined as a "type-specific glycolipid of mycobacterial origin" [5–7, 38]. This type of mycoside is more correctly termed a glycosylphenolphthiocerol dimycocerosate, and they are also known as "phenolic glycolipids" [8, 9].

Biosynthesis of *n*-Fatty Acids

M. tuberculosis has both a multifunctional fatty acid synthase (FAS) type I, which is found predominately in eukaryotes, and a dissociable multicomponent FAS type II, found exclusively in prokaryotes and plants [40]. The type I synthase catalyzes, in a bimodal distribution, the de novo synthesis of C_{16}-C_{18} and C_{24}-C_{26} fatty acids [41]. Mycobacterial *O*-methylated polysaccharides appear to stimulate the rate of fatty acid synthesis by FAS via formation of a ternary complex with FAS I. An acyl-CoA substrate noncovalently binds to the synthase, enhancing the rate-limiting diffusion of the product [42]. As a consequence, the release of C_{16}-C_{18} fatty acids caused by the *O*-methylated polysaccharides results not only in the increased rate of fatty acid biosynthesis, but also in a bimodal distribution in favor of C_{16}-C_{18} fatty acids. It has been suggested [40–42] that mycobacterial FAS I may function both as a de novo synthase and as an elongase; however, in vivo it may primarily elongate host fatty acids using malonyl-CoA. Mycobacterial FAS I is much larger than other multifunctional synthases, probably accounting for its unusual elongating activities. The domain organization of FAS I is like a head-to-tail fusion of two yeast FAS subunits: acyl transferase (AT), enoyl reductase (ER), dehydratase (DH), malonyl-palmitoyl transferase-acyl carrier protein (ACP) fused with β-ketoreductase (KR), and β-ketoacyl synthase (KS) [43]. No additional AT or KS domains, uniquely involved in chain elongation, were found. However, a putative active site sequence for a second nucleotide binding KR domain was apparent. The need for both NADH and NADPH for mycobacterial FAS I, in contrast to the single-reductant requirement of other FASs, presumably reflects this observation.

Biosynthesis of Multimethyl-Branched Fatty Acids

The biosynthesis of multimethyl-branched fatty acids involves methylmalonyl-CoA as the chain-extending substrate, in contrast to malonyl-CoA for *n*-fatty acids. Vertebrate FAS I preparations are capable of synthesizing 2, 4, 6, 8-tetra methyl-branched fatty acids when methylmalonyl-CoA is provided as an elongation substrate from short (C_2 or C_3) primers [44]. The structures of methyl-branched acids in *M. tuberculosis* suggest that they are not formed via a multienzyme FAS I complex but by a unique family of *pks* enzymes. Rainwater and Kolattukudy developed a cell-free synthesis of mycocerosic acids from C_{20}-CoA and methylmalonyl-CoA and purified the enzyme to homogeneity from *M. bovis* BCG [45, 46]. The mycocerosic acid synthase (MAS) was found to be a multifunctional synthase, similar to FAS I, producing C_{27}-C_{32} mycocerosic acids from straight-chain precursors (Figure 4). Subsequent cloning and sequencing of *mas* revealed the domain structure KS-AT-DH-ER-KR-ACP [47]. Homology searching suggested that the AT and KS domains may be involved in methylmalonyl-CoA versus malonyl-CoA substrate selectivity. Both of the AT and KS domains of MAS showed selectivity for methylmalonyl-CoA over malonyl-CoA when overexpressed in *Escherichia coli* and *Mycobacterium smegmatis* (both hosts possess no MAS activity), producing multimethyl-branched acids from *n*-acyl primers [48].

Targeted disruption of *mas* from *M. bovis* BCG yielded a mutant lacking mycocerosic acids (C_{29} and C_{32}) and phenolic glycolipids, but still producing PDIMs with shorter mycocerosic acids (C_{23} and C_{26}) [49]. A recent study has implicated a second enzyme from *M. bovis* BCG, termed short-chain mycocerosic acids synthase (SMAS), catalyzing the synthesis of such acids [50]. Interestingly, no thioesterase domain was found, nor a mechanism, to account for chain termination in the synthesis of mycocerosic acids. Mycocerosic acids synthesized by MAS are not released, suggesting that a specific enzyme may be involved in transferring the mycocerosyl group from MAS to the hydroxyl groups of phthiocerol and phenolphthiocerol moieties (Figure 5). It had been postulated that an adjacent open reading frame (*orf3*) to *mas* would possibly catalyze this process. A recent study has established that *orf3* encodes a protein with in vitro acyl-CoA synthase (ACoAS) activity, thereby per-

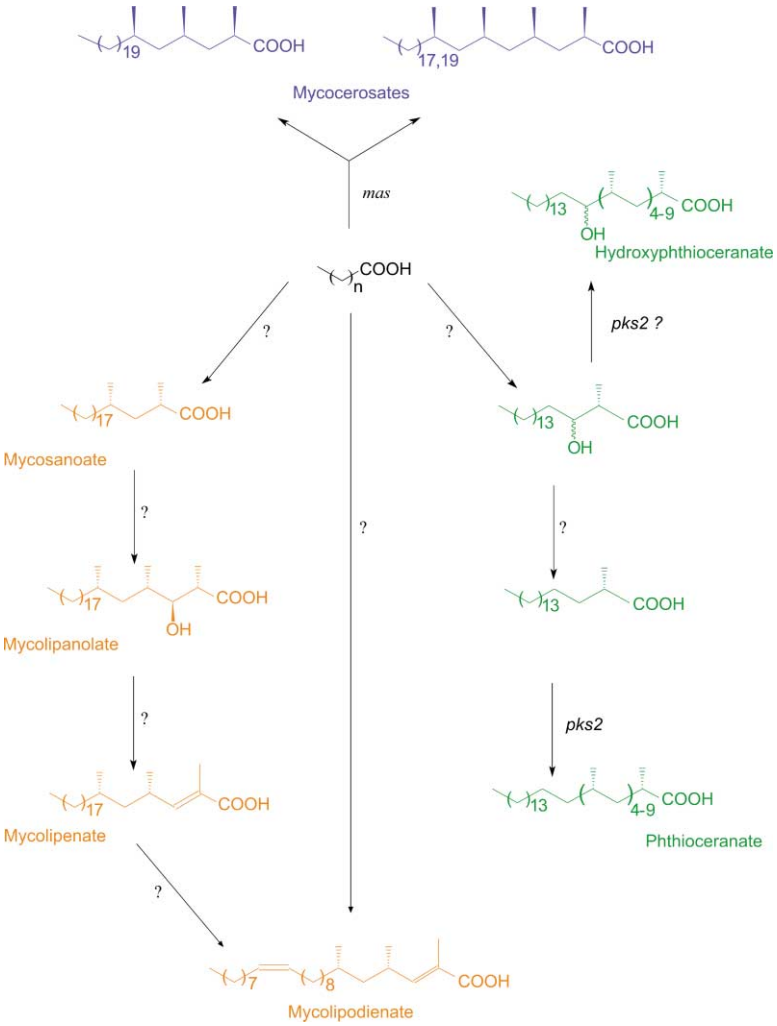

Figure 4. Structures of and Proposed Biosynthetic Pathways Leading to the Complex Multimethyl-Branched Fatty Acids from *M. tuberculosis*

See text for abbreviations.

forming CoA ester synthesis of *n*-C_{10}-C_{18} fatty acids but not mycocerosic acids [51]. In this study, the authors failed to observe both mycocerosic acid activation and mycocerosic acid release from MAS and its transfer to phthiocerols and phenolphthiocerols. More recent studies, where *acoas* was disrupted, produced a mutant that lacked phenolic glycolipids, even though it possessed normal levels of *mas* activity [52]. The authors suggested that this implicated ACoAS in the process of esterification of phenolphthiocerols (and phthiocerols) with mycocerosic acids. They argued against the possibility that disruption of *acoas* affected the supply of acyl-CoA primers to MAS. A number of highly similar *acoas* homologs are present near other polyketide synthase (PKS)-like genes in the *M. tuberculosis* genome and would compensate for the disruption of *acoas* unless a strict substrate-specificity existed. As discussed above, a second synthase, SMAS, exists in *M. bovis* BCG and releases its products directly as free fatty acids. Thus, esterification to acceptors can take place without the need for a separate ACoAS transferase activity. This would account for the continued synthesis of PDIMs in the case of both the *mas*- and *acoas*-disrupted mutants.

Biosynthesis of Phthiocerols and Phenolphthiocerols

The biosynthesis of phthiocerol and phenolphthiocerol would involve elongation of *n*-C_{20} acyl chains or an acyl chain possessing a phenol residue at the ω end. Chain extension is presumed to occur initially using the appropriate precursor and malonyl-CoA, followed by two successive rounds using methylmalonyl-CoA, decarboxylation, and reduction. Overall, the biosynthesis probably involves multifunctional synthases similar to FAS and MAS, and analogous to those utilized for erythronolide biosynthesis [53]. Using an approach based on domain organization and Southern hybridization experiments, Azad et al. [54] identified a gene cluster, termed *pps1-5* (and *ppsA-E* in *M. tuberculosis* [2]), that is required for phthiocerol and phenolphthiocerol biosynthesis in *M. bovis* BCG (Figure 5, annotated in the context of *M. tuberculosis*). Further analyses established that the *pps* cluster was downstream of the *mas* locus and transcribed in the opposite direction (Figure 2C; see also Figure 6). Disruption of the *pps* cluster and subsequent biochemical analysis of the *pps* mutant established that the gene cluster was involved in both phthiocerol and phenolphthiocerol biosynthesis due to the absence of

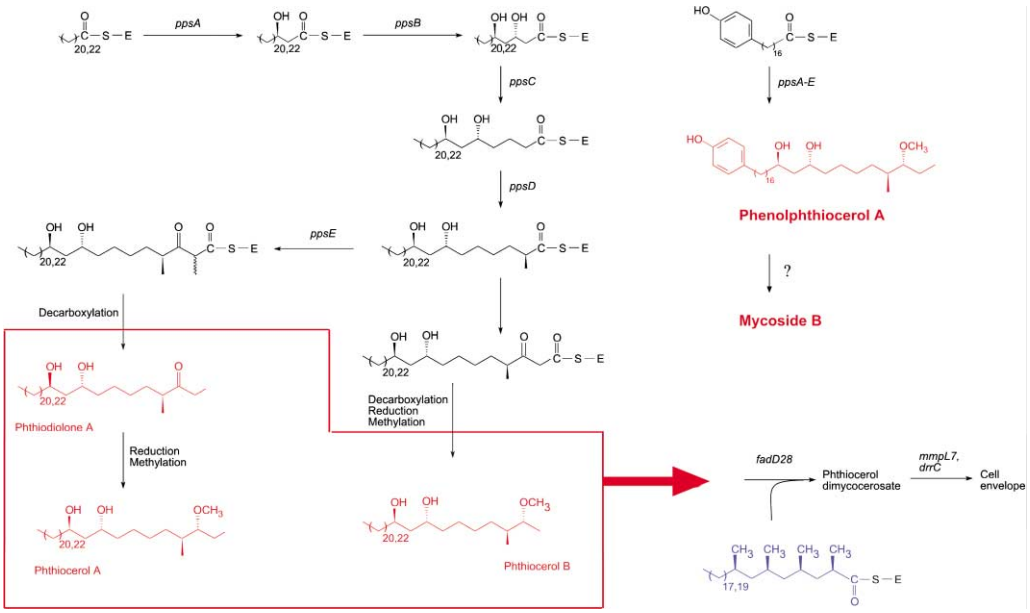

Figure 5. Structures of Phthiocerol and Phenolphthiocerol Families from the *M. tuberculosis* Complex and Proposed Pathways Leading to Their Biosynthesis

Adapted from Azad et al. [54]. See text for abbreviations.

PDIMs and PGL [54]. The mutant possessed normal MAS activity as compared to the wild-type, ruling out any indirect effect due to the disruption.

PDIMs and Virulent *M. tuberculosis*?

It has been recognized for some time that the distinctive nature of the cell wall of *M. tuberculosis* may be impor-

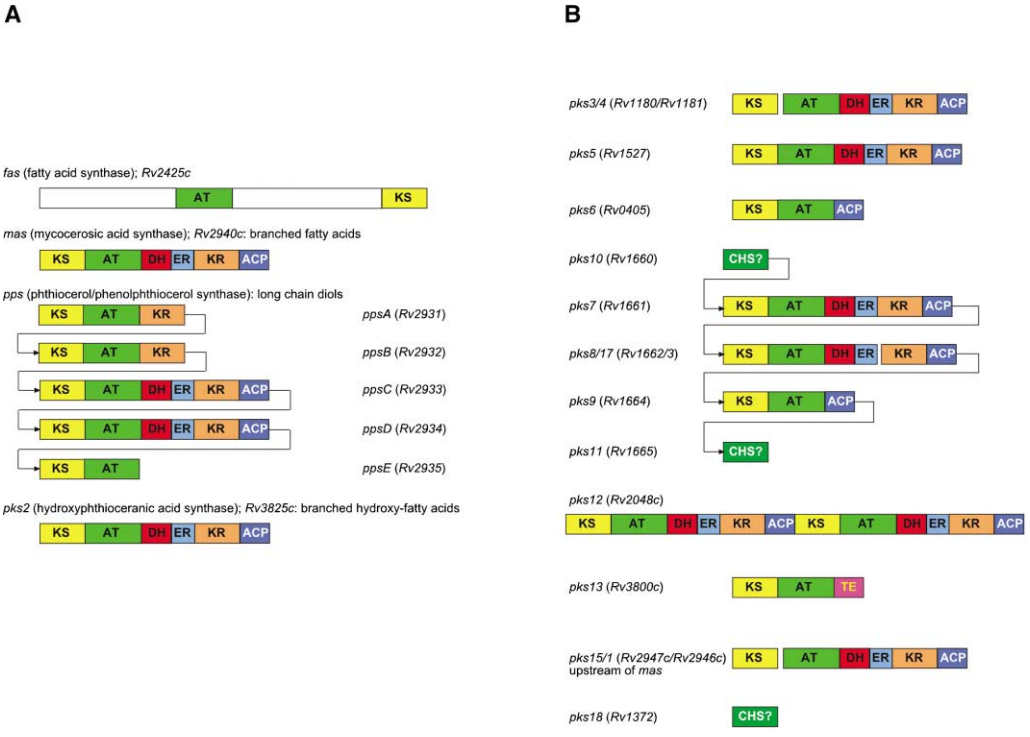

Figure 6. Domain Organization among *pks* Families from *M. tuberculosis*

(A) Known *pks* genes

(B) Unknown *pks* genes

The domain organization consists of the subunits β-ketoacyl synthase (KS), acyl transferase (AT), dehydratase (DH), enoyl reductase (ER), β-ketoreductase (KR), acyl carrier protein (ACP), thioesterase (TE), and chalcone synthase-like (CHS?).

tant for virulence. Cox et al. [55] and Camacho et al. [56] isolated a number of mutants based upon signature-tagged mutagenesis. The first mutant contained a transposon insertion in the promoter region of *ppsA-E*. The second contained an insertion in *fadD28*, presumably the ACoAS homolog described for *M. bovis* BCG involved in the transfer of mycocerosic acids to phthiocerols and phenolphthiocerols and adjacent to the *M. tuberculosis mas* (Figure 2C). Biochemical analysis of both mutants contained no detectable levels of PDIMs. The analyses presented for the *fadD28* mutant appear to contradict the earlier findings in *M. bovis* BCG, where disruption led only to the absence of PGLs and continued synthesis of PDIMs with shorter mycocerosic acids. This reaction is catalyzed by SMAS and an esterification process independent of ACoAS [51]. Recent studies by Camacho et al. [56] have shown that short-chain C_{26} mycocerosic acids are absent from *M. tuberculosis*, accounting for this discrepancy. A third mutant isolated by Cox and colleagues [55], with a transposon insertion in *mmpL7*, accumulated normal levels of PDIMs in cells. However, subsequent pulse-chase experiments established that, unlike wild-type *M. tuberculosis* that produces a considerable amount of PDIMs in the culture medium, the *mmpL7* mutant failed to liberate PDIM into the culture medium. Camacho et al. [56] have shown that culture fluid production of PDIM by the wild-type strain is due to the use of Tween in the 7H9 medium. This suggests that *mmpL7* is involved in PDIM transport. This is not surprising since *mmpL7* is similar to *actII-ORF3*, which is involved in polyketide export in *Streptomyces coelicolor* [57]. Interestingly, Camacho et al. [56] obtained a similar phenotype with a mutant in *drrC*, suggesting that for correct localization of PDIM, the MmpL7 and DrrABC transporters must interact for efficient translocation. In addition, the PDIM mutants exhibited a more permeable cell wall that was more sensitive to detergent, indicating that PDIM plays an important role in maintaining the integrity of the cell envelope.

The hydrophobic Mmp proteins appear to be restricted to mycobacteria, hence the designation "mycobacterial membrane protein." MmpL7 and other members of the MmpL family are predicted to contain 12 membrane-spanning α helices organized in three clusters [58]. This arrangement is reminiscent of that seen in the proton-dependent efflux proteins belonging to the RND (resistance, nodulation, division) superfamily, sharing similar size and topology [59]. The RND superfamily includes ActII-ORF3, involved in *Streptomyces coelicolor* polyketide export [57], and the SecDF type II secretion protein of *Bacillus subtilis* [60].

M. tuberculosis replicates extensively in the lungs during the first three weeks of infection. However, the *M. tuberculosis* strains mutated in the *pps*-promoter, *fadD28* or *mmpL7*, showed severely restricted replication in the lungs of infected mice [55, 56]. Although replication is greatest in the lungs, infection by intravenous injection also leads to seeding of *M. tuberculosis* into other organs, such as the spleen and liver. All three mutants had growth rates comparable with the wild-type in both of these organs. Thus, PDIM was suggested as being a virulence factor specifically required for growth of *M. tuberculosis* in the lungs of infected mice.

Table 1. A Survey of Genomic Organization Surrounding *pks* Genes in *M. tuberculosis*

Gene	Designation	Gene Length
ppsA-E	*Rv2931-Rv2935*	267,555
mas	*Rv2940c*	6,333
mmpL7	*Rv2942*	2,760
pks1	*Rv2946c*	4,848
pks15	*Rv2947c*	1,488
mmpL8	*Rv3823c*	3,267
pks2	*Rv3825c*	6,378
fadD23	*Rv3826*	1,752
pks3	*Rv1180*	1,464
pks4	*Rv1181*	4,746
mmpL10	*Rv1183*	3,006
fadD21	*Rv1185c*	1,734
mmpL12	*Rv1522c*	3,438
pks5	*Rv1527c*	6,324
fadD24	*Rv1529*	1,752
fadD25	*Rv1521*	1,749
mmpL1	*Rv0402c*	2,874
mmpS1	*Rv0403c*	426
pks6	*Rv0405*	4,206
fadD30	*Rv0404*	1,755

The related phenolic glycolipid-1 (PGL-1) from *M. leprae* has recently been shown to be a virulence factor through invasion of Schwann cells and determining the bacterial predilection to the peripheral nerve [61].

Biosynthesis of Phthioceranates, Mycosanoates, and Mycolipenates

An ORF, which possessed the catalytic domains found in MAS, was predicted to be involved in methyl-branched fatty acid biosynthesis by Sirakova et al. [62]. The *mas*-like gene *pks2* annotated in the *M. tuberculosis* database [2] was subsequently disrupted and the mutant analyzed for altered incorporation of [1-^{14}C]propionate into lipids. Two-dimensional thin-layer chromatography, followed by autoradiography, established that a lipid coinciding with sulfolipids was missing in the *pks2* mutant. Gas chromatography mass spectrometry (GC/MS) and radio-GC established that the hydroxyphthioceranic acids (heptamethyl C_{30}, heptamethyl C_{32}, octamethyl C_{32}, and octamethyl C_{34}) and phthioceranic acids found in sulfolipids (Figure 5) were the principal multimethyl-branched fatty acids absent in the *pks2* mutant. It is possible that the initial transformation of a straight-chain acid to the postulated α-methyl-β-hydroxy acid intermediate (Figure 4) is catalyzed by *pks2*. Disruption of *pks2* activity would then stop synthesis of both phthioceranic and hydroxyphthioceranic acids. Alternatively, the same *pks2* may be involved in the final elongation steps of both phthioceranic and hydroxyphthioceranic acids (Figure 4). The precise site of action of *pks2* remains to be clarified. Another intriguing aspect concerns how the balance between the synthesis of phthioceranic and hydroxyphthioceranic acids is controlled. An attractive common pathway is shown in Figure 4, but separate routes could operate. In such a common pathway, the regulation of the dehydration and reduction of the α-methyl-β-hydroxy acid intermediate (Figure 4) would be an important key process. The isolation of a sulfolipid mutant should pave the way forward to examining the

role played in pathogenesis of this lipid, which is only found in virulent *M. tuberculosis* strains [6, 13].

Little is know about the details of the biosynthesis of the mycosanoate/mycolipenate pathway (Figure 4). These fatty acids have the same chirality as the phthioceranates, so their pathways may have related genes. It is possible that the mycolipodienates could derive from mycolipenates by desaturation, but it is more likely that an unsaturated fatty acid, such as oleate, is incorporated at an earlier stage (Figure 4).

M. tuberculosis Genome and Future Prospects

To date only three *pks* families have been assigned a known function in *M. tuberculosis*, notably the *pps* cluster (phthiocerol and phenolphthiocerols), *mas* (mycocerosic acids), and *pks2* (hydroxyphthioceranic and phthioceranic acids?). It is clear from the genome that many more *mas*-like genes exist than multimethyl-branched fatty acids, assuming one *pks* for each class, according to Figure 4. This raises the possibility that key cell wall lipids or polyketide metabolites exist, possibly only expressed under certain culture conditions or during infection. These polyketides, including chalcone and stilbene synthases that are phylogenetically divergent from other polyketides and fatty acid synthases, could generate unreduced polyketides that are typically associated with pigments and flavonoids. For instance, the gene cluster *pks10*, *pks7*, *pks8*, *pks17*, *pks9*, and *pks11* includes two of the chalcone-synthase-like enzymes and two modules of an apparent type I system (see Figure 6). The biological significance of these metabolites is unclear, but it is anticipated that some of these will be key for intracellular growth. Interestingly, of the 18 *pks* identified in *M. tuberculosis*, only six were found in *M. leprae*, and the number of *mmpL* was restricted to only five instead of 16 in *M. tuberculosis* [2, 3]. The marked absence of so many *pks* clusters may be due to the lack of dextrorotatory multimethyl-branched acids, such as the mycolipenic, mycolipodienic, mycolipanolic, mycosanoic, phthioceranic, and hydroxyphthioceranic acids (Figure 4) in *M. leprae*. Laevorotatory mycocerosic acids are produced by *M. leprae* for incorporation into its PDIMs and phenolic glycolipids [8, 9]. A key feature that has arisen from the earlier studies of *pps* and *mas* has been the observation of the adjacent ORFs that play a key role in the biosynthesis and proper localization of PDIMs, such as *fadD28* and *mmpL7*. It could be argued from just a brief survey (Table 1) of the genome that some *pks* genes are located around similar ORFs, suggesting functional activities associated with acyltransferases (*fabD*-like) and transporters (*mmpL*-like). Thus, further studies are warranted to establish this connectivity as well as the function of the remaining *pks* genes in the *M. tuberculosis* genome.

Acknowledgments

This review is dedicated to the memory of Nicholas Polgar (1904–1996), the principal scientist in the determination of the essential structures of the phthiocerol family and dextro- and laevorotatory fatty acids from the tubercle bacillus; these lipids are now the focus of key studies related to the pathogenesis and virulence of *Mycobacterium tuberculosis*.

G.S.B., a Lister Institute-Jenner Research Fellow, and D.E.M. acknowledge support from the Medical Research Council, The Wellcome Trust, and the National Institutes of Health. The authors thank Dr. Steve Martin (GlaxoSmithKline) for the annotation of domain organizations represented in Figure 6.

References

1. W.H.O. (1996). TB: Groups at Risk, WHO Report on the Tuberculosis Epidemic. (Geneva: World Health Organization).
2. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., et al. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature *393*, 537–544.
3. Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Garnier, T., Churcher, C., Harris, D., et al. (2001). Massive gene decay in the leprosy bacillus. Nature *409*, 1007–1011.
4. Anderson, J. (1941). Structural peculiarities of acid-fast bacterial lipids. Chem. Rev. *29*, 225–243.
5. Asselineau, J. (1966). The Bacterial Lipids (Paris: Hermann).
6. Goren, M.B., and Brennan, P.J. (1979). Mycobacterial lipids: chemistry and biologic activities. In Tuberculosis, G.P. Youmans, ed. (Philadelphia: W.B. Saunders Co), pp. 69–193.
7. Minnikin, D.E. (1982). Lipids: complex lipids, their chemistry, biosynthesis and roles. In The Biology of the Mycobacteria, C. Ratledge, and J. Stanford, eds. (London: Academic Press), pp. 95–184.
8. Brennan, P.J. (1988). *Mycobacterium* and other actinomycetes. In Microbial Lipids, Volume I, C. Ratledge and S.G. Wilkinson, eds. (London: Academic Press), pp. 203–298.
9. Besra, G.S., and Chatterjee, D. (1994). Lipids and carbohydrates of *Mycobacterium tuberculosis*. In Tuberculosis: Pathogenesis, Protection and Control, B.R. Bloom, ed. (Washington, D.C.: American Society for Microbiology), pp. 285–306.
10. Minnikin, D.E., and Goodfellow, M. (1980). Lipid composition in the classification and identification of acid-fast bacteria. In Microbiological Classification and Identification, R.G. Board and M. Goodfellow, eds. (London: Academic Press), pp. 189–256.
11. Kremer, L., Baulard, A.R., and Besra, G.S. (2000). Genetics of mycolic acid biosynthesis. In Molecular Genetics of Mycobacteria, G.F. Hatfull and W.R. Jacobs, eds. (Washington, D.C.: American Society for Microbiology), pp. 173–189.
12. Belisle, J.T., Vissa, V.D., Sievert, T., Takayama, K., Brennan, P.J., and Besra, G.S. (1997). Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. Science *276*, 1420–1422.
13. Goren, M.B. (1970). Sulfolipid I of *Mycobacterium tuberculosis*, strain H37Rv. Purification and properties. Biochim. Biophys. Acta *210*, 116–126.
14. Besra, G.S., Bolton, R.C., McNeil, M.R., Ridell, M., Simpson, K.E., Gluska, J., van Halbeek, H., Brennan, P.J., and Minnikin, D.E. (1992). Structural elucidation of a novel family of acyltrehaloses from *Mycobacterium tuberculosis*. Biochemistry *31*, 9832–9837.
15. Minnikin, D.E., Dobson, G., Sesardic, D., and Ridell, M. (1985). Mycolipenates and mycolipanolates of trehalose from *Mycobacterium tuberculosis*. J. Gen. Microbiol. *131*, 1369–1374.
16. Daffé, M., Lacave, C., Lanéelle, M.-A., Gillois, M., and Lanéelle, G. (1988). Polythienoyl trehalose, glycolipids specific for virulent strains of the tubercle bacillus. Eur. J. Biochem. *172*, 579–584.
17. Muñoz, M., Lanéelle, M.A., Luquin, M., Torrelles, J., Julian, E., Ausina, V., and Daffé, M. (1997). Occurrence of an antigenic triacyl trehalose in clinical isolates and reference strains of *Mycobacterium tuberculosis*. FEMS Microbiol. Lett. *157*, 251–259.
18. Minnikin, D.E., Parlett, J.H., Dobson, G., Goodfellow, M., Magnusson, M., and Ridell, M. (1986). Lipid profiles of members of the *Mycobacterium tuberculosis* complex. In *Mycobacteria* of Clinical Interest, M. Casal, ed., (Amsterdam: Elsevier), pp. 75–78.
19. Daffé, M., Lacave, C., Lanéelle, M.-A., and Lanéelle, G. (1987). Structure of the major triglycosyl phenol-phthiocerol of *Mycobacterium tuberculosis* (strain Canetti). Eur. J. Biochem. *167*, 155–160.

20. Daffé, M., McNeil, M., and Brennan, P.J. (1991). Novel type-specific lipooligosaccharides from *Mycobacterium tuberculosis*. Biochemistry *30*, 378–388.
21. Watanabe, M., Yamada, Y., Iguchi, K., and Minnikin, D.E. (1994). Structural elucidation of new phenolic glycolipids from *Mycobacterium tuberculosis*. Biochim. Biophys. Acta *1210*, 174–180.
22. Dmitriev, B.A., Ehlers, S., Rietschel, E.T., and Brennan, P.J. (2000). Molecular mechanics of the mycobacterial cell wall: from horizontal layers to vertical scaffolds. Int. J. Med. Microbiol. *290*, 251–258.
23. Daffé, M., and Draper, P. (1998). The envelope layers of mycobacteria with reference to their pathogenicity. Adv. Microb. Physiol. *39*, 131–203.
24. Christensen, H., Garton, N.J., Horobin, R.W., Minnikin, D.E., and Barer, M.R. (1999). Lipid domains of mycobacteria studied with fluorescent molecular probes. Mol. Microbiol. *31*, 1561–1572.
25. Ginger, L.G., and Anderson, R.J. (1945). The chemistry of the lipids of tubercle bacilli. LXXII. Fatty acids occurring in the wax prepared from tuberculin residues. Concerning mycocerosic acid. J. Biol. Chem. *157*, 203–211.
26. Marks, G.S., and Polgar, N. (1955). Mycoceranic acid. Part II. J. Chem. Soc., 3851–3857.
27. Chanley, J.D., and Polgar, N. (1950). Dextrorotatory acids of tubercle bacilli lipids. Nature *166*, 693–695.
28. Cason, J., and Sumrell, G. (1950). Structural features of an acid of the phthioic type. J. Am. Chem. Soc. *72*, 4837–4838.
29. Coles, L., and Polgar, N. (1968). The mycolipaniolic acids. J. Chem. Soc., 1541–1544.
30. Coles, L., and Polgar, N. (1969). The mycolipodienic acids. J. Chem. Soc., 23–27.
31. Cason, J., Lange, G.L., and Urscheler, H.R. (1964). Isolation of 2,4-dimethyldocosanoic acid from the lipids of tubercle bacilli. Tetrahedron *20*, 1955–1961.
32. Stodola, F.H., and Anderson, R.J. (1936). The chemistry of lipids of tubercle bacilli. XLVI. Phthiocerol, a new alcohol from the wax of the human tubercle bacillus. J. Biol. Chem. *114*, 467–472.
33. Hall, J.A., and Polgar, N. (1954). The constitution of phthiocerol. Chem. and Ind., 1293.
34. Drayson, P.K., Lewis, J.W., and Polgar, N. (1958). Experiments relating to phthiocerol. Part III. Degradative studies of a C_{11} oxidation product of phthiocerol. J. Chem. Soc., 430–433.
35. Demarteau-Ginsburg, H., Lederer, E., Ryhage, R., Ställberg-Stenhagen, S., and Stenhagen, E. (1959). Structure of phthiocerol. Nature *183*, 1117–1119.
36. Minnikin, D.E., and Polgar, N. (1966). Studies relating to phthiocerol. Part V. Phthiocerol A and B. J. Chem. Soc., 2107–2112.
37. Minnikin, D.E., and Polgar, N. (1967). Studies relating to phthiocerol. Part VII. Phthiodiolone A. J. Chem. Soc., 803–807.
38. Smith, D.W., Randall, H.M., Maclennan, A.P., Putney, R.K., and Rao, S.V. (1960). Detection of specific lipids in mycobacteria by infrared spectroscopy. J. Bacteriol. *79*, 217–229.
39. Demarteau-Ginsburg, H., and Lederer, E. (1963). Sur la structure chimique du mycoside B. Biochim. Biophys. Acta *70*, 442–451.
40. Bloch, K. (1975). Fatty acid synthases from *Mycobacterium phlei*. Methods Enzymol. *35*, 84–90.
41. Petersen, D.O., and Bloch, K. (1977). *Mycobacterium smegmatis* fatty acid synthetase, long chain transacylase chain length specificity. J. Biol. Chem. *252*, 5735–5739.
42. Wood, W.I., Petersen, D.O., and Bloch, K. (1977). *Mycobacterium smegmatis* fatty acid synthetase, a mechanism based on steady state rates and product distributions. J. Biol. Chem. *252*, 5745–5749.
43. Fernandes, N.D., and Kolattukudy, P.E. (1996). Cloning, sequencing and characterization of a fatty acid synthase-encoding gene from *Mycobacterium tuberculosis* var. *bovis* BCG. Gene *170*, 95–99.
44. Kolattukudy, P.E., Fernandes, N.D., Azad, A.K., Fitzmaurice, A.M., and Sirakova, D. (1997). Biochemistry and molecular genetics of cell wall lipid biosynthesis in mycobacteria. Mol. Microbiol. *24*, 263–270.
45. Rainwater, D.L., and Kolattukudy, P.E. (1983). Synthesis of mycocerosic acids from methylmalonyl coenzyme A by cell-free extracts of *Mycobacterium tuberculosis* var. *bovis* BCG. J. Biol. Chem. *258*, 2979–2985.
46. Rainwater, D.L., and Kolattukudy, P.E. (1985). Fatty acid biosynthesis in *Mycobacterium tuberculosis* var. *bovis* Bacillus Calmette-Guerin. Purification and characterization of a novel fatty acid synthase, mycocerosic acid synthase, which elongates *n*-fatty acyl-CoA with methylmalonyl-CoA. J. Biol. Chem. *260*, 616–623.
47. Mathur, M., and Kolattukudy, P.E. (1992). Molecular cloning and sequencing of the gene for mycocerosic acid synthase, a novel fatty acid elongating multi-functional enzyme, from *Mycobacterium tuberculosis* var. *bovis* Bacillus Calmette-Guerin. J. Biol. Chem. *267*, 19388–19395.
48. Fernandes, N.D., and Kolatukudy, P.E. (1997). Methylmalonyl coenzyme A selectivity of cloned and expressed acyltransferase and β-ketoacyl synthase domains of mycocerosic acid synthase from *Mycobacterium bovis* BCG. J. Bacteriol. *179*, 7538–7543.
49. Azad, A.K., Sirakova, T.D., Rogers, L.M., and Kolattukudy, P.E. (1996). Targeted replacement of the mycocerosic acid synthase gene in *Mycobacterium bovis* BCG produces a mutant that lacks mycosides. Proc. Natl. Acad. Sci. USA *93*, 4787–4792.
50. Fernandes, N.D., and Kolattukudy, P.E. (1998). A newly identified methyl-branched chain fatty acid synthesizing enzyme from *Mycobacterium tuberculosis* var. *bovis* BCG. J. Biol. Chem. *273*, 2823–2828.
51. Fitzmaurice, A.M., and Kolattukudy, P.E. (1997). Open reading frame *3*, which is adjacent to the mycocerosic acid synthase gene, is expressed as an acyl coenzyme A synthase in *Mycobacterium bovis* BCG. J. Bacteriol. *179*, 2608–2615.
52. Fitzmaurice, A.M., and Kolattukudy, P.E. (1998). An acyl-CoA synthase (*acoas*) gene adjacent to the mycocerosic acid synthase (*mas*) locus is necessary for mycocerosyl lipid synthesis in *Mycobacterium tuberculosis* var. *bovis* BCG. J. Biol. Chem. *273*, 8033–8039.
53. Katz, L., and Donadio, S. (1993). Polyketide synthesis, prospects for hybrid antibiotics. Annu. Rev. Microbiol. *47*, 875–912.
54. Azad, A.K., Sirakova, T.D., Fernandes, N.D., and Kolattukudy, P.E. (1997). Gene knock-out 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.
55. Cox, J.S., Chen, B., McNeil, M., and Jacobs, W.R. (1999). Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. Nature *402*, 79–83.
56. Camacho, L.R., Constant, P., Raynaud, C., Lanéelle, M.A., Triccas, J.A., Gicquel, B., Daffé, M., and 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.
57. Fernandez-Moreno, M.A., Caballero, J.L., Hopwood, D.A., and Malpartida, F. (1991). The *act* cluster contains regulatory and antibiotic export genes, direct targets of translational control by the bldA tRNA gene of *Streptomyces*. Cell *66*, 769–780.
58. Tekaia, F., Gordon, S.V., Garnier, T., Brosch, R., Barrell, B.G., and Cole, S.T. (1999). Analysis of the proteome of *Mycobacterium tuberculosis* in silico. Tuber. Lung Dis. *79*, 329–342.
59. Saier, M.H., Tam, R., Reizer, A., and Reizer, J. (1994). Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport. Mol. Microbiol. *11*, 841–847.
60. Bolhuis, A., Broekhuisen, C.P., Sorokin, A., van Roosmalen, M.L., Venema, G., Bron, S., Quax, W.J., and van Dijl, J.M. (1998). SecDF of *Bacillus subtilis*, a molecular Siamese twin required for the efficient secretion of proteins. J. Biol. Chem. *273*, 21217–21224.
61. Ng, V., Zanazzi, G., Timpl, R., Talts, J.F., Salzer, J.L., Brennan, P.J., and Rambukkana, A. (2000). Role of the cell wall phenolic glycolipid-1 in the peripheral nerve predilection of *Mycobacterium leprae*. Cell *103*, 511–524.
62. Sirakova, T.D., Thirumala, A.K., Dubey, V.S., Sprecher, H., and Kolattukudy, P.E. (2001). The *Mycobacterium tuberculosis pks2* gene encodes the synthase for the hepta- and octamethyl-branched fatty acids required for sulfolipid synthesis. J. Biol. Chem. *276*, 16833–16839.