Pathophysiological Implications of Cell Envelope Structure in *Mycobacterium tuberculosis* and Related Taxa


Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/59585

1. Introduction

Members of the *Mycobacterium tuberculosis* complex are successful pathogens due in a large part to the complex interactions of an array of very special lipid molecular classes and associated macromolecules that have been known for many years [1-6]. The cooperative assembly of these components results in the presence of very robust cell envelopes. In particular there is a specialised outer membrane; this can be termed a “mycomembrane”, but in this review the abbreviation, MOM, for “mycobacterial outer membrane” will be used. Subtle structural variations in mycobacterial lipid components point to important roles in the integrity and function of the mycobacterial cell envelope. Over three decades ago, an attempt was made to rationalise the role of the known constituent molecules with a proposal for a “chemical” model of the cell envelope [4]. This general model has been supported in subsequent studies but a wealth of new knowledge suggests that a significant upgrade is needed. Each class of cell envelope structural units will be introduced, reasons for “why” such structures are produced will be explored and an updated cell envelope model presented as a working hypothesis. Close comparisons of the cell envelope compositions of *M. tuberculosis* and related taxa will be interpreted in terms of evolution and pathogenicity. In various diseases, the term “pathophysiology” is usually used to describe physiological changes in affected hosts; here its meaning is turned around to explore the impact of aspects of mycobacterial cell physiology in pathogenic members of the *M. tuberculosis* complex.
Precise structures of the lipid moieties displayed in members of the *M. tuberculosis* complex have become established during the past five decades, so it is now possible to explore the significance of structure in the pathophysiological role of such lipids. Attention will focus on lipids, whose regular occurrence in significant proportions indicates that they are integral structural members of the mycobacterial cell envelope. Details of general biosynthetic pathways will not be considered, nor will particular biological effects of individual lipid classes. Recent cryo-electron microscopy studies have allowed the absolute dimensions of the mycobacterial cell envelope to be estimated [7-9] so all components must adopt conformations that allow their integration into an effective organelle. Tubercle bacilli also have a less clearly defined capsule external to the more coherent cell envelope [10,11]; a detailed discussion of this important region will not be attempted here. Similarly, a detailed exploration of protein content is not given, even for structurally vital proteins such as well-characterised porins [12]. The main chemical structures that must be accommodated in the cell envelope of *M. tuberculosis* are considered in the following sections, commencing with the long-chain mycolic acids that are mycobacterial signature components [4,13-18].

2. Mycolic acids

The 70 to 90 carbon mycolic acids (MAs) are very characteristic chemical components in the genus *Mycobacterium*. Members of the *M. tuberculosis* complex have three classes, the so-called α-mycolates, methoxymycolates and ketomycolates (Figure 1) [4,13-18]. The latter two varieties (Figure 1B) are comprised of subclasses having either *cis*-cyclopropane rings or *trans*-cyclopropane rings with an adjacent methyl branch. In the case of the α-mycolates (Figure 1A), representative C₈₀ mycolates from *Mycobacterium kansasii* and the *M. tuberculosis* complex are compared to highlight the importance of key structural differences, whose importance will be discussed later. All these MA classes occur naturally with at least five homologues and variations in the numbers of carbons between the various functional groups. The majority of MAs are covalently bound to arabinose termini of a mycoloylarabinogalactan-peptidoglycan (mAGP) macromolecule to form a lipid monolayer inner leaflet of the MOM [4,19], as will be described in detail later.

Recent physicochemical investigations have clearly demonstrated that mycolic acids characteristically adopt distinctly different folded conformations depending on structural niceties [20-24]. Ketomycolates in *M. tuberculosis* predominately fold to yield a compact “W” conformation, with four chains in parallel [20,21,23,24]. Such tight packing can provide the foundation for an effective hydrophobic permeability barrier in the inner leaflet of the MOM. In contrast, α- and methoxymycolates can form W-conformations but also more readily inhabit a range of more extended conformations [20-23], some of which can be visualised as “U” or “Z” shaped [24]. The α-mycolates tend towards an open fully extended U-conformation with the two distal chains extended and methoxymycolates show intermediate behaviour between the two other classes [22,24]. For the purpose of illustrating the importance of mycolate conformation in MOM inner leaflet function (see later), ketomycolates are restricted to W-
conformations, methoxymycolates are shown with both semi-folded sZ and fully folded W-shapes [24] and α-mycolates have extended eU-conformations [24] (Figure 2).

Figure 1. Representative structures of mycolic acids. A. α-Mycolates (C₈₀), comparing M. kansasii and M. tuberculosis complex. B, C. The main methoxy- and ketomycolates from M. tuberculosis.

Figure 2. Two-dimensional representative conformations of M. tuberculosis mycolic acids. A. Extended α-mycolate. B. Two alternatives for methoxymycolates. C. Fully folded ketomycolate.
3. Esters of mycolic acids

The so-called “cord factors” are the best known mycolic acid esters in mycobacteria; they are principally trehalose dimycolates (TDMs) with trehalose monomycolates (TMMs) also being encountered (Figure 3). The proportions of TDMs and TMMs vary widely in mycobacteria so an integral structural role is not indicated, their main importance lying in the key role as intermediates in the transfer of mycolic acids on to arabinosyl units in the cell envelope [13]. Glucose monomycolates (GMMs) are common in mycobacteria, but in highly variable proportions [25]. Consistent proportions, however, are recorded for monomycoloyl glycerols (MMGs) (Figure 3) in the Mycobacterium bovis members of the M. tuberculosis complex [26,27], thereby suggesting some cell envelope structural involvement. Mycobacteria also produce very complex mixtures of di- and triacylglycerols, some of which contain non-hydroxylated fatty acids that correspond to the meromycolate portion of mycolic acids [28,29], the so-called “mycobacteric” acids [2]. Triacylglycerols have a storage role in “lipid bodies” [30] but they have also been suggested as contributors to the MOM outer leaflet [31]; the complex mixtures of di- and triacylglycerols must be fully unravelled before precise roles can be properly defined.

A. Trehalose 6,6'-dimycolate (TDM)  B. Monomycoloyl glycerol (MMG)

Figure 3. Mycolic acid esters. A. Trehalose dimycolate (TDM). B. Monomycoloyl glycerol (MMG).

4. Phthiocerol and phenolphthiocerol dimycocerosate families

The phthiocerol and phenolphthiocerol long-chain diols are esterified by multimethyl-branched “mycocerosic” acids whose chiral centres have R absolute configuration in members of the M. tuberculosis complex and M. kansasii (Figure 4) [32-35]. Phthiocerol dimycocerosates (PDIMs) are mainly based on long-chain diols, the phthiocerol As and phthiodiolones (Figure 4). Related families of the phenolphthiocerol dimycocerosates have characteristic antigenic oligosaccharides linked to the phenolic residue and these are commonly known as “phenolic
glycolipids” (PGLs) (Figure 4) [33-36]. PDIMs are large (> 90 carbons) hydrophobic molecules that are considered to be “free lipid” constituents of the outer leaflet of the external MOM, interacting with the chains of the covalently bound MAs of the inner leaflet. The PDIMs from the *M. tuberculosis* complex are relatively large in comparison with those of other taxa. Other interesting cases are the so-called “Beijing” variants of *M. tuberculosis* [37] and *M. kansasii* [32, 33, 35] where only restricted selections of members of the phthiocerol family are encountered.

**Figure 4.** Representative structures of mycocerosate esters of phthiocerols and glycosyl phenophthiocerols. A. Phthiocerol family. B. Glycosyl phenophthiocerol (phenolic glycolipids). C. Simplified structures for illustrative use in Figures 10 and 11.
The oligosaccharide substituents of PGLs have demonstrable antigenicity [38] but the individual sugar units are relatively hydrophobic. Certain members of the *M. tuberculosis* complex, such as the rare smooth morphology tubercle bacilli termed “*Mycobacterium canettii*”, have two PGL types with different oligosaccharides but *M. bovis*, for example, has only a single type (Figure 4). The PGLs from *M. kansasii* have close structural similarities to those from “*M. canettii*” (Figure 4), but not to those from *Mycobacterium marinum* [33-36]. The chiral centres in the multimethyl-branched fatty acid substituents in both the PGLs and PDIMs from *M. marinum* and *Mycobacterium ulcerans* are of $S$ absolute configuration [33-36]. Significantly, the 1,3-diol units in the phthiocerol and phenolphthiocerol moieties (Figure 4) from *M. marinum* and *M. ulcerans* have erythro geometric configuration in contrast to the more common threo stereochemistry [33-36].

5. Acyl trehaloses

In addition to TDMs and TMMs, there are families of trehalose-based glycolipids acylated with multimethyl branched fatty acids with $S$ absolute configuration of their chiral centres (Figure 5) [34,39,40]. The main fatty acids encountered are $C_{24}$ mycosanoic, $C_{27}$ mycolipenic, $C_{27}$ mycolipanolic, $C_{37}$ phthioceranic and $C_{40}$ hydroxyphthioceranic acids (Figure 5). Diacyl trehaloses (DATs) are the simplest representatives, based on $C_{24}$ mycosanoic and $C_{27}$ mycolipanolic acids. The $C_{27}$ mycolipenates are the characteristic acyl components of pentaacyl trehaloses (PATs) (Figure 5). The exceptionally long phthioceranic and hydroxyphthioceranic acids are the fatty acids found in a family of sulfated trehalose glycolipids (SGLs) (Figure 5) [3,41].

6. Lipooligosaccharides

A highly polar series of lipids, which include trehalose in their saccharide core, are termed lipooligosaccharides (LOs) [42,43]. Such lipids are absent in many modern *M. tuberculosis* isolates but they are characteristic of “*M. canettii*” and *M. kansasii* (Figure 6). Lipooligosaccharides are associated with biofilms and motility [44]. Indeed, it has been shown that smooth variants of *M. kansasii*, containing LOs, are rapidly cleared from the organs of infected animals, but rough variants, lacking all LOs, produce chronic systemic infections [45].

7. Phosphatidylinositol mannosides and other polar lipids

The mycobacterial plasma membrane incorporates conventional polar lipids, such as phosphatidylethanolamine (PE), phosphatidylinositol (PI) and diphosphatidylglycerol (DPG) (Figure 7), which can interact together to form the basis of a typical membrane bilayer. However, most mycobacteria have a remarkably consistent family of four phosphatidylinosoli-
tol mannosides (PIMs); these comprise mono- (AcPIM<sub>2</sub>) and diacyl phosphatidylinositol dimannosides (Ac<sub>2</sub>PIM<sub>2</sub>) and mono- (AcPIM<sub>6</sub>) and diacyl phosphatidylinositol hexamannosides (Ac<sub>2</sub>PIM<sub>6</sub>) (Figure 7) [46-49]. Recent research has provided evidence that PIM<sub>2</sub> and PIM<sub>6</sub> classes may be unevenly distributed over the two leaflets of the mycobacterial plasma membrane [31]. These findings will be interpreted, later, as showing that PIMs may act to reinforce the plasma membrane, perhaps adding a further level of selective permeability to the mycobacterial cell envelope.

Figure 5. Acyl trehaloses. A. Diacyl trehalose (DAT). B. Pentaacyl trehalose (PAT). C. Sulfoglycolipid (SGL). D. Simplified structures for illustrative use in Figures 10 and 11.
8. Lipomannan and lipoarabinomannan

The basic structures of the PIMs polar lipid family (Figure 8) share the same manno-phosphatidylinositol (MPI) anchor with two classes of characteristic large lipoglycans, namely lipomannans (LMs) and lipoarabinomannans (LAMs) (Figure 8) [50-54].

9. Mycoloylarabinogalactan-peptidoglycan (mAGP)

The overall chemical structure of this complex macromolecule has been clarified during the past decade [55-58]. A specific linker unit covalently binds the proximal galactan portion of the arabinogalactan to peptidoglycan with the distal arabinose moieties providing anchorage for the 70 to 90 carbon long-chain mycolic acids (Figure 9) [19,55-58]. While chemical connec-
Figure 7. Phosphatidylinositol mannosides (PIMs). A. Phosphatidylinositol dimannosides. B. Phosphatidylinositol hexamannosides. C. Phosphatidylinositol. D. Phosphatidylethanolamine. E. Diphosphatidylglycerol. Simplified structures are included.
tivity is established, conformational preferences of the carbohydrate domains remain a matter of conjecture with diverse interpretations. It is becoming evident that versatile peptidoglycans can adopt different conformational arrangements, depending on the length of the polymeric disaccharide chains with helices being possible for shorter units [59]. It was shown that a synthetic peptidoglycan adopted a right-handed helical conformation [60]. A distinctive feature of mycobacterial peptidoglycan is the presence of a proportion of N-glycolyl muramic acid substituents, rather than the N-acetyl groups found in many other bacterial taxa [55-59].

The size and complexity of the mycoloylarabinogalactan-peptidoglycan, which is an extensive single macromolecule, provides a major challenge in perceiving how it can be coherently organised in three dimensions. This is not such a difficulty in many other bacterial taxa where no really major macromolecules are directly attached to peptidoglycan. It is now well established that the mycolic acids form a coherent inner leaflet of the MOM and this necessitates support from a well-integrated underlying platform.

The proposed mAGP arrangement (Figure 10) is based on a “scaffold” model [61,62], where peptide cross-linked helices are interspersed with helices of the galactan part of the arabinogalactan; the arabinan portion is then arranged to provide linkage points for mycolic acids. An

Figure 8. Essential structural topography of *M. tuberculosis* lipomannan and lipoarabinomannan. MPI is manno-phosphatidylinositol.
attractive arrangement can be envisaged with the galactan extending to a level similar to that of the peptidoglycan helix to produce an essentially level “mosaic platform” as a stable anchorage for mycolic acids. While the helical galactan can provide a relatively rigid base unit, the arabinan may be more flexible so that the bound mycolic acids can jostle for position and occupy optimal locations. Indeed, arabinan flexibility may be an important factor in allowing hydrophobic interactions to govern the relative location of mycolic acid chains and associated free lipids. Calculations [63] indicate that the arabinofuranose polymer is reluctant to adopt a rigid compact helical conformation, thereby allowing a degree of flexibility.

10. Cell envelope organisation

The original model [4], with an inner and outer membrane, was based mainly on chemical principles, supported by freeze-etching results [64] that showed two clear distinct parallel cleavage planes in the mycobacterial cell envelope. The dual membrane proposal was confirmed by a confocal microscopy study that showed differential location of two fluorescent dyes of different lipophilicity [65]. The outer membrane was visualised directly by cryo-electron microscopy and the essential dimensions of the mycobacterial cell envelope were revealed [7-9]. An updated model for the cell envelope organisation for tubercle bacilli is proposed in Figure 10. Justification for the details of the proposal will build outwards from the plasma membrane.

The inner plasma membrane in mycobacteria has been traditionally regarded as conventional, even though a significant role was lacking for the unusual phosphatidylinositol mannosides.
(PIMs) (Figure 7). A resolution of this conundrum has been indicated in a study [31], which showed strong evidence for locating Ac2PIM2 (Figure 7) as the sole polar component of the inner leaflet of the inner membrane. It was suggested that PIM6 would be present in the outer leaflet of the inner membrane, projecting into the periplasm. It is not clear why there are two versions of PIM2 and PIM6, with either three or four fatty acid chains (Figure 7), but as a working hypothesis both PIM2 lipids are placed in the inner leaflet and both PIM6 lipids in the outer leaflet of the plasma membrane (Figure 10). As demonstrated by two-dimensional thin-layer chromatography [26], the proportions of the principal four PIM types are remarkably consistent, as is the proportion of PI. It is possible that equal proportions of PIMs with three and four fatty acid constituents are optimal for close packing in membranes; detailed physical studies on these lipids would be instructive. The proportions of PIM2 exceed those of PIM6 so if PIM2 lipids are considered to predominate in the inner leaflet [31], then PI, PE and DPG (Figure 7) may complete the outer leaflet along with PIM6. There is a distinct possibility that mycobacterial inner plasma membranes, rich in PIMs with three and four fatty acid chain anchors, have special physical properties that enhance its stability and perhaps governs permeability. Indeed, it has been suggested that this inner membrane may be “a bilayer environment of unusually low fluidity” [31] contributing to drug resistance. It was also noted [31] that the behaviour of PIM2 liposomes had been found [66] to have behaviour suggestive of exceptional stability. It is now apparent that the inner mycobacterial plasma membrane is a highly specialised organelle, worthy of being distinguished with special nomenclature. Given the developing popularity of “MOM” for the mycobacterial outer membrane, a related simple suggestion might be “MIM” for the “mycobacterial inner membrane”. It was found that disruption of PIM2 production causes growth arrest [67,68] but the higher PIMs were dispensable [69], thereby indicating an important structural role for PIM2. It has also been indicated that the acylation state of PIMs is also significant [70].

The outer leaflet of the MIM inner plasma membrane is also a suggested location for the PIM-related LM and LAM (Figure 10), but unequivocal evidence is elusive with alternative MOM location being a possibility. In a well-balanced objective analysis [71], it was concluded that LAM had at least an initial anchorage in MIM. However, in some cases [72,73], the undoubted presence of LAM at the cell surface required invoking specific lipoglycan transport mechanisms that need to be fully defined. At least a transient MIM location for LAM is supported by the presence of related lipoglycans in other actinomycetes, which do not have mycolic acids and an outer membrane, as summarised recently [72]. The basic fact that the lipid anchors of LM and LAM are identical to those in PIMs (Figures 7,8) suggests very strongly that all these components have a common anchorage in MIM. This should not rule out possible interactions with the hydrophobic MOM surface, but such lipophilic binding is predictably less specific and it is very difficult to envisage LM and LAM as important integral components of the MOM outer leaflet. The PIM-based lipid anchors appear to be all very similar for related LM and LAM lipoglycans across the genus Mycobacterium and related mycolata; however, enormous variations in the surface of the MOM in such taxa would militate strongly against any specific incorporation of LM and LAM into the MOM outer leaflet. Immunogold atomic force microscopy failed to detect cell surface LAM in M. bovis BCG, but LAM was revealed after treatment with drugs that attack cell envelope targets [74]. For Corynebacterium glutamicum, it was found
Figure 10. Two-dimensional representation of the location and interaction of structural components in the cell envelope of *M. tuberculosis*. MOM and MIM are “mycobacterial outer membrane” and “mycobacterial inner membrane”, respectively. AGP is the arabinogalactan-peptidoglycan. Mycolic acid conformations (Figure 2) in the MOM are labelled “αₜ” for α-mycolate for fully extended “eU” shape, “Kₜ” and “Mₜ” for fully folded keto- and methoxymycolate “W” shape, respectively, and “Mₛ” for semi-folded “sZ” shape. Correlating dimensions of all components are drawn to fit within the spatial constraints imposed by cryo-electron microscopy [8,9]. Dimensions of peptidoglycan (PG) helices are derived from synthetic material [60]. Helices of the mannan sections of lipomannan (LM) and lipoarabinomannan (LAM) and both the arabinan and galactan components of arabinogalactan (AG) were modelled using GLYCAM. Details of simplified lipid structures are in Figures 4-6.
that LM and LAM are embedded in the plasma membrane, along with the PIMs [75]. Overall, the importance of lipophilic anchors in external LM and LAM is diminished by the observation that their lipid-free glycan counterparts, mannan and arabinomannan, may be one destination of these components [72].

The extensive supposed “periplasmic space” (Figure 10) will, in fact, be an area of intense activity as all cellular components of the MOM whose synthesis is initiated in the cytoplasm, and continued through MIM, may be assembled and organised within [51-54]. The fact that many of these cell envelope components are relatively large may explain why such a relatively extensive compartment is needed. Cryo-electron microscopy studies [7-9,76] gave indications of some rather indistinct structural elements within the periplasm, labelled as layers L1 and L2 [7,76]. The internal L1 layer is most probably associated with “granular” material of protein origin, with the outer L2 layer corresponding to some of the peptidoglycan-arabinogalactan matrix [76]. It has been proposed that the maintenance of the relatively low-density periplasmic space could be facilitated by the presence of large polymeric material [9]. The helical mannan polysaccharide moiety in LM (Figure 10) may have such a role, but it could also act as a scaffold or template to compartmentalise various biochemical activities. In M. smegmatis it was shown that LAM was dramatically reduced as the bacteria approached stationary phase, but LM, mycolic acids and arabinogalactan were unchanged; this indicates that all these latter three components have important structural roles [73].

The accommodation and conformation of the mycoloyl arabinogalactan peptidoglycan macromolecular structure is a major challenge in the relatively limited space available (Figure 10). An informed choice has been made to use the “scaffold” approach of [61,62], with a helical peptidoglycan network interspersed with helical galactan units, as detailed in Figure 10. The relatively heavy peptidoglycan peptide cross-linking [52] may be a factor in favouring the scaffold arrangement in mycobacteria. This proposal echoes a previously advanced arrangement [52] that did not attempt to make precise spatial correlations and neglected to include the extensive periplasm. It would be interesting to explore the possibility that the N-glycolyl muramic acid substituents may have some influence on mycobacterial peptidoglycan conformation. The proposed coherent peptidoglycan – galactan “mosaic platform” layer (Figure 10) could provide a coherent anchorage for the arabinan moieties, some of which are esterified with mycolic acids. As noted previously, conformations of MAs vary with structural type and these are illustrated with W-conformations for keto-MAs, an extended U-conformation for α-MAs and equal numbers of W- and U-conformations for methoxy-MAs (Figure 2, 10). The MAs are comprised of approximately 50% α-MAs [17,18] as reflected in Figure 10. Such covalently bound MAs are instrumental in providing a stable MOM inner leaflet with potential for interaction with diverse free lipids to produce an effective outer membrane permeability barrier. Detailed spatial calculations attempted to simulate the accommodation of W-folded α-MAs into the MOM [77]; however, other MA conformations must be incorporated to provide a full picture. In a detailed quantitative study [31] of the cell envelope of M. smegmatis it was demonstrated that there were sufficient hydrocarbon chains in both inner and outer membranes to confirm the viability of the original dual membrane model [4].
A wide range of free lipid types are considered to form the outer leaflet of the outer mycomembrane of members of the *M. tuberculosis* complex. These range in polarity from the hydrophilic LOSs (Figure 6), through PGLs (Figure 4), DATs, SGLs and PATs (Figure 5), to the highly apolar hydrophobic PDIMs (Figure 4). All these unusual lipids have specialised fatty acid components, incorporating varying numbers of methyl branches located near to the carboxyl group (Figures 4-6). For the original chemical model of the mycobacterial cell envelope [4], it was conjectured that the presence of methyl branches might moderate the depth of insertion of the fatty acids into the MOM bilayer. This remains an attractive hypothesis, though detailed physical studies would be of value. The proximal multimethyl branches would favour a coiled conformation, leaving the distal straight-chain portion of such fatty acids to interact with mycolic acid chains. The straight-chain fatty acids in DATs, PATs (Figure 5) and LOSs (Figure 6) could possibly be layered parallel to the MOM surface providing elements of lateral stabilisation. In certain previous cell envelope models [9,78] there has been an unjustified tendency to include inverted TDMs as major components in the MOM outer leaflet. A proven role of TDMs is as an agent for the transfer of mycolic acids onto the arabinan matrix, presumably in the assembly of the inner leaflet of the MOM. It is more likely, therefore, that TDMs would align themselves with the trehalose unit adjacent to the arabinan, as previously indicated [4]. Similarly, monomycoloyl glycerols (MMGs) (Figure 3) in *M. bovis* would be more readily accommodated with an internal glycerol unit. Such a location might possibly act to modify MOM fluidity in *M. bovis* whose mycolic acid composition has high proportions of cis-cyclopropyl keto-MAs, in comparison with the predominant trans-cyclopropyl keto-MAs in *M. tuberculosis* [17,18]. However, in related taxa, such as *Corynebacterium glutamicum*, it is probable that the structurally much simpler mycolic acids, and perhaps their trehalose esters, are able to participate as members of the outer leaflet of the outer membrane [8,78].

11. Evolutionary and pathogenicity aspects of cell envelope composition

It is of particular interest to attempt to obtain an understanding of the influence and importance of cell envelope composition in mycobacterial pathogenicity and evolution. A consensus is developing that an attractive evolutionary pathway can be envisaged from environmental *Mycobacterium kansasii*, through “*M. canettii*” to all the modern biotypes of the *M. tuberculosis* complex [79-83]. *M. kansasii* is the environmental organism that phenotypically resembles *M. tuberculosis* most closely and this relationship has been supported by genomic comparisons [79-81]. Cogent arguments have been advanced to associate the evolution of ancient tubercle bacilli, such as “*M. canettii*”, with bacteria similar to *M. kansasii*, including indications of horizontal gene transfer between these taxa [80,81]. Key genes acquired by horizontal gene transfer include those coding for mycobacterial lipids, transferases and proteins related to adaptation to anaerobic conditions [80,81]. *M. kansasii* continues to cause pulmonary disease in Silesian and South African miners, the bacterium being contracted from water in showers [81]. A detailed study has shown that the unusual smooth morphology “*M. canettii*” strains appear to form a pivotal role in the evolution of tuberculosis [84]. Although extant strains of
“M. canettii” still cause human tuberculosis, they differ significantly in infectivity and appear to be relatively ancestral [85,86]. Genomic studies indicate that the very diverse “M. canettii” isolates appear to coalesce into a form of bottleneck after which all the modern human and animal biotypes evolved in a relatively linear manner [83,84,87,88]. A plausible working hypothesis for the evolution of *M. tuberculosis sensu stricto* is outlined in Figure 11, highlighting the possible contribution of cell envelope lipid composition.

The most fundamental underlying difference between members of the *M. tuberculosis* complex, broadly including “M. canettii”, and *M. kansasii* and related environmental taxa is seen in the mycolic acids. The α-mycolic acids from *M. kansasii* have the regular spacing of the proximal and distal cis-cyclopropyl groups (Figure 1), common to a wide range of mycobacteria [17,18]. In clear contrast, α-mycolates from “M. canettii” and members of the *M. tuberculosis* complex have the chain between the hydroxyl group and the proximal cyclopropyl group shortened most significantly from 17 methylene groups in, for example, *M. kansasii* to 11 and 13 methylenes in members of the *M. tuberculosis* complex [4,17,18,79]. Additionally, the chains in 2-position and the terminal meromycolate chain are both relatively extended by two carbons. As noted above [20,22,24,79], *M. tuberculosis* α-mycolic acids extend more readily in molecular dynamics simulations with apparent interaction of the chain in 2-position with the chain between the two cyclopropyl groups, in a “fully extended shape” (Figure 2) [22] or “eU” conformation [24].

The methoxymycolates and ketomycolates of “M. canettii” and *M. tuberculosis* (Figure 1) conform to the general pattern of these components in related mycobacteria, such as *M. kansasii*, but, significantly, these oxygenated mycolates are slightly larger than any others [17,18]. The particular ability of α-mycolates to adopt extended flexible U-conformations is probably significant for interactions with free lipid components of the MOM outer leaflet. Indeed it is possible that the exceptionally long terminal chain in *M. tuberculosis* α-mycolic acids may penetrate right to the MOM outer leaflet to contribute to the hydrophobicity of the cell envelope surface.

In this context, the balance of the three main types of mycolates is probably significant; the ratios of the α-, methoxy- and ketomycolates are, respectively, ~10:5:8 for *M. kansasii*, ~10:6:8 for “M. canettii” and ~10:5:5 for *M. tuberculosis* [17,18,79]. It is conceivable that having 50% α-mycolates may optimise hydrophobic interactions with the particular range of free lipids in the MOM outer leaflet in *M. tuberculosis*. In a detailed consideration [79], it was found possible to discern quite a range of distinct differences in the overall mycolic acid composition of these three taxa; the overall conclusion was an apparent simplification and tightening up of mycolate composition in modern *M. tuberculosis*.

There are also very significant changes in cell envelope MOM free lipid composition between all the taxa, shown in Figure 11. *M. kansasii* has characteristic polar LOSs (Figure 6), relatively polar PGLs (Figure 4) and PDIMs based only on phthiodiolones (Figure 4). These three lipid families appear to be the principal components of the outer leaflet of the MOM of *M. kansasii*, possibly contributing to a relatively polar cell surface compatible with a native hydrophilic environment. The PGLs of *M. kansasii* and “M. canettii” provide an appealing phenotypic link, with a loss of a single sugar to give a triglycosyl PGL and three sugar units to give a monoglycosyl PGL (Figures 4 and 11), as also highlighted previously [81]. However, some most significant new lipid structural principles are introduced in “M. canettii” (Figure 11). Two
major classes of acyl trehalose glycolipids, the DATs and PATs (Figure 5), are encountered [86]. The relatively polar DATs are good antigens, which probably contribute to the cell surface properties of tubercle bacilli [38]. In contrast, the barely antigenic PATs are very non-polar and are likely to increase cell surface hydrophobicity quite significantly. The reduced 9-sugar oligosaccharide in the LOSs of “M. canettii”, compared with the 13-sugar oligosaccharide in M. kansasii might also result in reduced hydrophilicity. It is not known if all the diverse isolates of “M. canettii” have similar lipid profiles, but it certainly appears that the MOM free lipid composition of this taxon has some redundancy with perhaps an over-generous provision of lipid types (Figure 11). This redundancy might correlate with a high propensity for horizontal gene transfer, positively accumulating new mycobacterial lipid principles at the expense of a rather overblown excess of lipid biosynthesis. This might support a concept that “M. canettii” is a low pathogenicity intermediate taxon in evolutionary transit, rather than an efficient pathogen with a particular niche. In contrast, as will be detailed below, the modern M. tuberculosis complex has emerged out the ancestral melange as a group of efficient pathogens with a range of specialised hosts.

Aspects of the cell envelope lipid composition of “M. canettii” are essentially the same as for M. tuberculosis sensu stricto, including mycolic acid composition, but there are a number of key differences. Significantly, the LOSs and PGLs present in “M. canettii” are absent in M. tuberculosis, but the latter is uniquely characterised by the presence of SGLs [89]. As noted above, the highly polar LOSs are associated with hydrophilic interactions, such as motility and biofilm formation [44]. The PGLs are also relatively polar, but in addition to having a structural role in the outer leaflet of the MOM, particularly specific functions are elusive. SGLs behave on chromatography as quite polar lipids, presumably due to the presence of the sulfated trehalose unit. However, SGLs have very large nonpolar phthioceranic and hydroxyphthioceranic fatty acid acyl chains (Figure 5), which are likely to provide enhanced hydrophobicity. In two informative studies, it was shown that deficiencies in SGLs [90] and DATs and PATs [91] did not have decisive effects on replication and persistence. In bacterial disease, clear distinctions must be made between transmissibility and the pathogenic process once the infection has been established. It would appear that the very hydrophobic SGLs and PATs (Figure 5) may have primary value in the transmission process, perhaps accompanied by some secondary roles in pathogenicity.

Previous studies have shown an important link between hydrophobicity and aerosol performance in Mycobacterium avium [92]. Preliminary studies are under way to compare the hydrophobicity of “M. canettii” and M. tuberculosis. For cultures grown on solid media in the presence of Congo Red [93], the smooth colonies of M. kansasii and “M. canettii” resisted staining by the dye, but two different rough strains of M. tuberculosis absorbed so much dye that they almost merged into the background (Figure 12). In preliminary standard partitioning experiments between hexadecane and water [94], it was found that M. tuberculosis was approximately one fifth more hydrophobic than “M. canettii” (details not shown). These simple experiments indicate a distinct difference in hydrophobicity between one member of the smooth supposed ancestral “Cannetti” taxon and rough modern human tubercle bacilli. It follows, therefore, that M. tuberculosis sensu stricto may be specifically adapted for aerosol
transmission, whereas “M. canettii” is less favoured for this mode of transmission [85,86]. A link between cell envelope hydrophobicity and aerosol transmission is also probably signifi-

Figure 11. A generalised scenario for the significance of lipid composition in the evolution of *M. tuberculosis*. Details of the simplified lipid structures are in Figures 4-6. The numbers of sugars in particular LOSs and PGLs are shown in brackets with an asterisk, e.g. LOS (13)*.
cant in *Mycobacterium abscessus* infections [95,96]. Airway infections are often associated with rough variants that lack polar glycopeptidolipids and are likely to be more hydrophobic and aerosol transmissible than the usual smooth strains [95,96]. It was also shown that rough *Gordona* species were much more hydrophobic than smooth morphology variants of the same strain [97]. By use of chemical force microscopy the cell envelope surface of *M. bovis* BCG was found to be uniformly hydrophobic, but drugs disrupting mycolic acid synthesis destroyed this feature [74].

A jump from environmental *M. kansasii* to the obligate *M. tuberculosis* human pathogen requires the cooperation of suitable animal hosts, in which evolutionary changes can take place. The total dearth of any evidence for tuberculosis in *Homo sapiens*, prior to the Holocene, suggests zoonotic evolutionary hosts [79,98,99]. Characteristic skeletal lesions in a variety of Pleistocene megafauna are suggesting the presence of tuberculosis over a wide time period, going back to at least 120 ka BP [100,101]. In one well-documented example of a 17ka extinct bison, diagnosis by lesion was conclusively confirmed by amplification of *M. tuberculosis* complex DNA and recovery of pristine mycolipenate (Figure 5) and mycocerosate (Figure 4) tuberculosis lipid biomarkers [99,102,103]. The complex digestive organs in megafauna, such as bison and mastodons are potentially ideal vessels to facilitate evolutionary development, as the rich population of other microbial species present would be well situated to participate in horizontal gene transfer. There is substantial evidence for Pleistocene horizontal gene transfer in ancestral strains of the tubercle bacillus before linear evolution in the Holocene [80,81,104].

It is conceivable that, during the Pleistocene, thinly spread members of *Homo sapiens*, and also *Homo neanderthalensis*, may have contracted evolving ancestral tuberculosis from infected megafauna. However, there is no evidence to suggest that modern tuberculosis developed from ancestral strains solely in the human vehicles, mentioned above. Indeed, recent studies indicate that Neanderthals all became extinct before about 40ka BP [105], which precedes the perceived ~30 to 12ka BP evolutionary bottleneck in tubercle bacilli [84,87,106]. A feasible scenario, therefore, for the emergence of all the modern biotypes of the *M. tuberculosis* complex through the bottleneck would involve a complex web of interactions between *H. sapiens* and Pleistocene animal reservoirs until the dramatically ameliorated climatic conditions at the
beginning of the Holocene allowed humans to form settlements and promulgate communal tuberculosis. The apparent ready spread of human *M. tuberculosis* in early Holocene settlements was probably a result of the ability of the tubercle bacillus to fly in aerosols. This could be a consequence of the dramatically enhanced hydrophobicity of *M. tuberculosis sensu stricto* as compared with the most likely ancestral strains, labelled “*M. canettii*”. It will not be easy to pinpoint a precise event when modern *M. tuberculosis* made its decisive debut on the human stage. As noted above, the digestive systems of Pleistocene megafauna were probably very efficient vessels for the evolution of tubercle bacilli and the change from the relatively hydrophilic “*M. canettii*”, in the digestive tract, to the hydrophobic *M. tuberculosis* that could be expelled from the lung. The enormous diversity of the extant surviving “*M. canettii*” taxon [84] probably points to an even greater range of smooth ancestral strains being passaged, recycled and adapted through animals and the environment. Indeed it probably reasonable to suggest that the great genetic diversity of smooth strains of the tubercle bacilli may be a result of their emergence over a wide time scale and geographical area.

12. Conclusion

The cell envelope of *M. tuberculosis* and related taxa is an assemblage of efficiently coordinated and closely interacting macromolecules and exquisitely designed lipid moieties. One of the most characteristic organelles is the “mycobacterial outer membrane” or “MOM”, but recent studies have shown that the “mycobacterial inner membrane”, or “MIM”, may be comparable in importance (Figure 10). Indeed the MIM, if it has a very characteristic inner leaflet composed predominately of PIM$_2$ phospholipids, may be a formidable barrier protecting the cytoplasm (Figure 10). The MOM is composed of an outer leaflet of “free lipids”, which interacts with a monolayer of mycolic acids covalently bound to the relatively flexible arabinan portion of an arabinogalactan that, in turn, is linked to peptidoglycan. Utilising the hypothesis of the “scaffold” model, the peptidoglycan and the galactan section of the arabinogalactan form helices perpendicular to the cell surface resulting in a stable type of “mosaic platform” from which the MOM can be linked via an arabinan cushion (Figure 10). The space between the MIM and peptidoglycan, usually termed the “periplasm”, is envisaged as being the initial place for the linkage of the characteristic lipoglycans, namely “lipomannan” (LM) and “lipearabinosomannan” (LAM) (Figure 10). However, LAM certainly is transported to the external capsular regions, but LM may linger in the periplasm affording a degree of organisation and stability. Overall, it is important to note that the mycobacterial cell envelope is a dynamic three-dimensional organelle; the hypothetical arrangement pictured in Figure 10 can only be a two-dimensional snapshot illustrating likely locations and interactions.

Given the hypothesis that tubercle bacilli evolved from an environmental organism, such as *M. kansasii*, via a mammalian adapted semi-environmental ancestral taxon, such as “*M. canettii*”, parallel linked changes in the lipid components of the MOM can be discerned (Figure 11). A clear phenotypic link between *M. kansasii* and “*M. canettii*” is provided by clearly comparable “phenolic glycolipids” (PGLs) (Figure 4). The introduction of distinct mycolic acid
(Figure 1) and acyl trehalose (DATs, PATs, Figure 5) structures shows that “M. canettii” has an indisputable link with modern members of the M. tuberculosis complex (Figure 11). Refinement of the MOM lipid composition of “M. canettii”, leads to the simplified MOM free lipids of M. tuberculosis, comprising PDIMs, DATs, PATs (Figures 4,5) and the very characteristic sulfoglycolipids (SGLs, Figure 5). The apparent enhanced hydrophobicity of the cell envelope of M. tuberculosis sensu stricto correlates well with the ability to be transmitted in aerosols. In broad evolutionary terms, a plausible scenario involves various Pleistocene megafauna passaging environmental mycobacteria, such as M. kansasii, until horizontal gene transfer resulted in the diverse family of smooth morphology strains labelled “M. canettii”. Further evolutionary development, with possible human involvement for the first time, produced modern M. tuberculosis whose novel ability to fly in aerosols probably coincided with congregating settled humans multiplying together along with tubercle bacilli as new devastating companions.

Acknowledgements

The Leverhulme Trust Project Grant F/00 094/BL (OY-CL, DEM, GSB). GSB was supported by a Personal Research Chair from Mr. James Bardrick and the UK Medical Research Council. The UK Medical Research Council and the Medical Research Foundation provided support to AB.

Author details

David E. Minnikin1, Oona Y-C Lee1, Houdini H.T. Wu1, Vijayashankar Nataraj1, Helen D. Donoghue2, Malin Ridell3, Motoko Watanabe4, Luke Alderwick1, Apoorva Bhatt1 and Gurdyal S. Besra1

*Address all correspondence to: d.e.minnikin@bham.ac.uk

1 Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Birmingham, UK

2 Centre for Clinical Microbiology, Division of Infection and Immunity, University College London, London, UK

3 Department of Microbiology and Immunology, Institute of Biomedicine, University of Gothenburg, Gothenburg, Sweden

4 School of Pharmacy, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan
References


[31] Bansal-Mutalik R, Nikaido H. Mycobacterial outer membrane is a lipid bilayer and the inner membrane is unusually rich in diacyl phosphatidylinositol dimannosides. PNAS 2014;111: 4958-4963.


