The molecular biology of mycobacterial trehalose in the quest for advanced tuberculosis therapies

Ana Nobre,1† Susana Alarico,1† Ana Maranha,1,2 Vitor Mendes1,3 and Nuno Empadinhas1,4

1CNC–Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, Portugal
2Biosciences PhD Program, Department of Life Sciences, University of Coimbra, Coimbra, Portugal
3Department of Biochemistry, University of Cambridge, Cambridge, UK
4III/UC-Institute for Interdisciplinary Research, University of Coimbra, Coimbra, Portugal

Trehalose is a natural glucose disaccharide identified in the 19th century in fungi and insect cocoons, and later across the three domains of life. In members of the genus Mycobacterium, which includes the tuberculosis (TB) pathogen and over 160 species of nontuberculous mycobacteria (NTM), many of which are opportunistic pathogens, trehalose has been an important focus of research over the last 60 years. It is a crucial player in the assembly and architecture of the remarkable mycobacterial cell envelope as an element of unique highly antigenic glycolipids, namely trehalose dimycolate (‘cord factor’). Free trehalose has been detected in the mycobacterial cytoplasm and occasionally in oligosaccharides with unknown function. TB and NTM infection statistics and death toll, the decline in immune responses in the aging population, human immunodeficiency virus/AIDS or other debilitating conditions, and the proliferation of strains with different levels of resistance to the dated drugs in use, all merge into a serious public-health threat urging more effective vaccines, efficient diagnostic tools and new drugs. This review deals with the latest findings on mycobacterial trehalose biosynthesis, catabolism, processing and recycling, as well with the ongoing quest for novel trehalose-related mechanisms to be targeted by novel TB therapeutics. In this context, the drug-discovery pipeline has recently included new lead compounds directed toward trehalose-related targets highlighting the potential of these pathways to stem the tide of rising drug resistance.

The genus Mycobacterium

According to the LPSN (List of Prokaryotic Names with Standing in Nomenclature, http://www.bacterio.net/), the most comprehensive source for prokaryotic nomenclature, the genus Mycobacterium has been significantly enlarged in recent years and currently encompasses 169 valid species (as of May 2014). Mycobacterium tuberculosis is the most notorious among them as it causes life-threatening tuberculosis (TB), one of the most ancient infectious diseases known to man and still accountable for 1.3 million annual deaths worldwide (WHO – Tuberculosis 2013: http://www.who.int/mediacentre/factsheets/fs104/en). Some of these are caused by new multidrug-resistant (MDR), extensively drug-resistant (XDR) and totally drug-resistant (TDR) strains (Viveiros et al., 2012; Zumla et al., 2014).

Most mycobacteria are saprophic organisms that may opportunistically infect different animals including humans. They were designated nontuberculous mycobacteria (NTM) to be differentiated from those causing TB (Falkingham, 2009). About 50% of these organisms, including a high number of species causing pulmonary disease, were described in the last decade, allowing a closer look into mycobacterial biodiversity and the inherent infectious potential (Hoefsloot et al., 2013). Mycobacteria are rod-shaped aerobic bacteria that have traditionally been divided in two groups: the rapidly growing mycobacteria (RGM), which are able to form visible colonies on agar medium within 7 days, and the slowly growing mycobacteria (SGM), which require one to several weeks to grow colonies on plates (Falkingham, 2009; Legendre et al., 2012). With a few exceptions, phylogenetic studies generally support this division and have shown that while SGM are a strictly monophyletic group, RGM are a polyphyletic group and seem to evolutionarily precede slow growers (Devulder et al., 2005). An unusual resilience to low pH, high temperature, desiccation, heavy metals and disinfectants provide mycobacteria with a wide spectrum of adaptation to soils and aquatic environments, including man-made water-distribution systems (Billi & Potts, 2002; Falkingham, 2009; Wallace et al., 1998). All mycobacteria share a distinctive lipid-rich cell envelope (Fig. 1), which is a major factor...
Fig. 1. Schematic representation of the *M. tuberculosis* cell envelope and cytoplasmic compartment. The scheme includes the biosynthetic pathways and enzymes involved in the metabolism of trehalose (yellow) and trehalose-containing cell components, the standard drugs currently used for TB chemotherapies (shaded red) and recently described candidate drugs (boxed) which target enzymes of the mycobacterial trehalome (Swarts et al., 2012). Overall, only proteins with involvement in trehalose metabolism are represented. AG, arabinogalactan; DAT, diacyltrehalose; LAM, lipoarabinomannan; LM, lipomannan; MA, mycolic acid; OM, outer membrane or mycomembrane; PAT, pentacyltrehalose; PG, peptidoglycan; PIM, phosphoinositol mannoside; SL-1, sulfolipid-1; TDM, trehalose dimycolate; TMM, trehalose monomycolate. Ebselen, I3-AG85, C215, BM212, AU1235 and SQ109 are inhibitors of trehalose-related processes in *M. tuberculosis* (see main text and Table 4). Dashed arrows indicate putative function. Adapted from Jackson et al. (2013) and Zumla et al. (2013).
underlying their resilient character and success in adapting to harsh conditions in the environment and within the infected host (Falkinham, 2009; Yang et al., 2014). Some of the unique structures found in this protective shell are prime targets for some of the currently used anti-mycobacterial drugs (Favrot & Ronning, 2012; Zumla et al., 2014) (Fig. 1). Although our knowledge of the immunological and molecular mechanisms by which M. tuberculosis and NTM colonize their hosts has advanced considerably in recent decades, the emergence of intractable MDR strains, the aging population with the associated decline in immune responses (immunosenescence) and the increasing incidence of chronic diseases represent a significant challenge worldwide. This threat urges continued research into the molecular biology of these pathogens, from the identification of new gene functions and characterization of enzyme properties to the discovery of inhibitors that can be optimized into new lead compounds and used as efficient anti-mycobacterials (Cruz-Hervert et al., 2012; Hoefsloot et al., 2013; Viveiros et al., 2012).

The unique mycobacterial cell envelope

The mycobacterial cell envelope is a robust structure with three recognizable layers: an inner plasma membrane, a cell wall and an outer coat or capsule (Fig. 1) (Favrot & Ronning, 2012; Jackson et al., 2013). This cell envelope is rich in unique trehalose-containing glycolipids dominated by trehalose dimycolate (TDM; Fig. 2), which was associated with the rough morphology of mycobacterial colonies by Bloch and colleagues in 1953 and named ‘cord factor’ (Bloch et al., 1953). However, it was later demonstrated that ‘non-cording’ mycobacteria also produce TDM, confirming that this structure could not be solely responsible for cord formation (Glickman, 2008; Julián et al., 2010). Trehalose monomycolate (TMM; Fig. 2) is the TDM precursor and also a major component of the envelope (Etienne et al., 2009; Marrakchi et al., 2014). Other glycolipids found in the cell envelope but with species-specific distribution are the phenolglycolipids, glycopeptidolipids, phthiocerol dimycocerosates, sulfolipids and lipooligosaccharides (LOS) (Kaur et al., 2009; Minnikin et al., 2002).

**Fig. 2.** Trehalose-containing glycolipids from M. tuberculosis. LOS, lipooligosaccharides; PAT, pentacyltrehalose; SL-1, sulfolipid-1; TDM, trehalose dimycolate; TMM, trehalose monomycolate. Since LOS are absent from M. tuberculosis, the chemical structure reflects the mature form of LOS from M. marinum. Adapted from: Alibaud et al. (2014); Hatzios et al. (2009); Marrakchi et al. (2014); Rombouts et al. (2011); Seeliger et al. (2012).
The surface-exposed capsule (Fig. 1) is a loosely bound mesh dominated by the glycogen-like polysaccharide α-D-glucan and a few lipids, but also containing secreted proteins that include lipases, proteases, secretion systems and porins thought to be involved in mycobacterial–host interactions or in the maintenance of these organisms’ typical intracellular lifestyle (van der Wel et al., 2007). These components are among the important factors involved in the concerted strategy that allows *M. tuberculosis* to remain dormant within an individual for decades and reactivate into active TB later in life (Flynn & Chan, 2005). The bacilli evolved efficient strategies for the manipulation of host-macrophage biology in order to create a suitable niche where they can replicate. Nevertheless, highly virulent *M. tuberculosis* was reported to be able to translocate from phagolysosomes into the cytoplasm of nonapoptotic cells where it could replicate (Lee et al., 2011). More recently, these virulent strains were found to induce host-cell necrosis after phagolysosomal rupture, followed by evasion from macrophages and extracellular persistence within necrotic granulomas (Simeone et al., 2012). These observations allowed new perspectives upon *M. tuberculosis* infection strategies whereby bacilli evade innate immunity through recurring cycles of macrophage infections and killing. A recent theory sustains that *M. tuberculosis* selects permissive macrophages through surface-exposed phenolic glycolipids and evades microbicidal phagocytes using phthiocerol dimycocerosate lipids to mask the underneath structures with pathogen-associated molecular patterns (PAMPs) recognized by the innate immune system of the host (Cambier et al., 2014).

The mycobacterial cell wall is a complex sheath between the capsular coat and the inner plasma membrane (Fig. 1). It is made of a thick and densely packed hydrophobic layer (outer membrane, also known as mycomembrane) dominated by very long chain (C<sub>60</sub>–C<sub>90</sub>) branched fatty acids designated mycolic acids (MAs), which are covalently linked to the polysaccharide arabinogalactan (AG) in turn covalently attached to the inner grid of peptidoglycan (PG) that covers the phospholipidic plasma membrane (Fig. 1). Such heterogeneous macromolecular structure was designated mycolyl-arabinogalactan-peptidoglycan complex (MA-AG-PG, mAGP or MAPc) (Marrakchi et al., 2014; Verschoor et al., 2012). MAs intercalating with the hydrophobic moieties of solvent-extractable glycolipids largely dictate the fluidity of the envelope and uphold a permeability barrier that prevents the free access of polar compounds to the cytoplasm, including many commonly used antibiotics (Guenin-Maè et al., 2009; Korf et al., 2005). Although species-specific variation remains to be associated with defined events in host–pathogen crosstalk, for example how the distinct location of some virulence factors affects some biased immune responses to pathogens and opportunistic NTM, many of the structural and biosynthetic details of the remarkable cell wall have been comprehensively elucidated (Jackson et al., 2013; Kaur et al., 2009; Tahlan et al., 2012).

The mycobacterial cell envelope is also populated by mannose-rich lipoglycans such as lipomannan and lipoarabinomannan (LAM), and their likely precursors phosphatidylinositol mannosides, all bound to the plasma membrane and the mycomembrane through phosphatidylinositol (Fig. 1). In addition to their structural functions, these compounds have also been implicated in the modulation of host immunity (Dhiman et al., 2011; Jankute et al., 2012; Mishra et al., 2011). Overall, the intricacy of the cell-envelope layers and their elaborate macromolecular components impose a protective barrier to host defences and confer mycobacteria with an unusually high resistance to different biotic and abiotic stressors (Ehrt & Schnappinger, 2009; Favrot & Ronning, 2012).

**Trehalose: general information and biological roles**

In 1832 an unknown sugar was, for the first time, detected in the rye fungus *Claviceps purpurea* (Wiggers, 1832). Twenty-six years later the same sugar was isolated from ‘trehala manna’, the sweet-tasting cocoons of *Larinus* beetles, and named trehalose (Berthelot, 1858). A century after its original discovery, Mary Pangborn and Rudolph Anderson purified trehalose from a *Mycobacterium* strain for the first time (Pangborn & Anderson, 1933).

Trehalose is a non-reducing disaccharide of α(1→1)-linked glucose that is widely distributed in nature, including several eukaryotic organisms (fungi, nematodes, insects, crustaceans and plants) and members of the prokaryotic domains Bacteria and Archaea (Elbein et al., 2003). Free trehalose has multiple roles in unrelated organisms and may be associated with different roles in the same organism, from osmoadaptation to the regulation of energetic supplies or intracellular signalling (Iturriaga et al., 2009). This disaccharide is considered a universal stress-protector molecule, as it was shown to shield proteins, membranes, DNA and whole cells from a variety of hostile conditions, including heat stress, freezing, γ-radiation and desiccation (Iturriaga et al., 2009; Lee et al., 2013; Webb & DiRuggiero, 2012). Trehalose is also considered the molecule of dormancy as it accumulates to high levels in organisms enduring extreme desiccation or anhydrobiosis, a stable state of suspended animation characterized by a nearly complete absence of water (Crowe et al., 1992). In addition, trehalose is the primary sugar in insects’ haemolymph and the main source of energy required for flight (Liebl et al., 2010). This disaccharide was shown not only to stabilize *Saccharomyces cerevisiae* proteins against heat-induced denaturation and aggregation, but also to play an essential role in *Escherichia coli* resistance to cold (Kandror et al., 2002; Singer & Lindquist, 1998). Moreover, *Candida albicans* accumulated trehalose in response to oxidative stress and its absence was shown to accelerate *H<sub>2</sub>O<sub>2</sub>*-triggered apoptosis (Lu et al., 2011). In *Propionibacterium freudenreichii*, a bacterium with probiotic potential, the accumulation of trehalose was a common response to acid stress, osmotic pressure and...
oxidative insults (Cardoso et al., 2007). In the thermophilic bacterium Thermus thermophilus only trace levels of trehalose were detected during growth under optimal conditions, but this disaccharide became the dominant osmolyte during growth under severe salt stress (Alarico et al., 2007). Trehalose has also been implicated in the control of sugar metabolism in plants, in transcriptional regulation mechanisms and involved in plant–micro-organism interactions (Iturriaga et al., 2009). Notably, trehalose biosynthesis has recently been considered a potent virulence factor of the plant pathogen Pseudomonas aeruginosa PA14 by allowing its replication in the plant intercellular environment (Djonovíc et al., 2013). While in members of the genus Streptomyces trehalose is associated with spore germination and also in heat tolerance, other actinobacterial genera such as Mycobacterium, Corynebacterium and Nocardia use trehalose during mAGP assembly for cell wall edification and additionally as a ubiquitous structural element of their cell envelopes (Elbein et al., 2003). Overall, trehalose is an extremely versatile molecule, often an essential element in many organisms.

Functional and structural versatility of trehalose in mycobacteria

Trehalose is a constitutive carbohydrate in the mycobacterial cytoplasm and a precursor to various glycolipids found in their cell envelopes (Fig. 1) (Jackson et al., 2007; Woodruff et al., 2004). M. tuberculosis can remain viable for months in desiccated dust particles and after 2 years storage under vegetable oil. Likewise, some environmental NTM also exhibit a similar phenotype (Billi & Potts, 2002; Tsukamura et al., 1983). However, the contribution of intracellular trehalose to this extreme mycobacterial behaviour typical of dehydration-tolerant organisms is only hypothetical. The similarity between dehydration tolerance conveyed to model mycobacterial membranes by pure TDM, by TDM-rich membranes and by synthetic trehalose glycolipids suggests that mycobacterial resistance to desiccation may result, at least in part, from the proximity between TDM molecules and from the interactions of their trehalose groups with neighbouring molecules (Harland et al., 2008).

Early studies on the TDM chemical and structural nature exposed the tight association between trehalose and the mycobacterial cell envelope (Bloch et al., 1953). The TDM non-covalent intercalation with the AG-linked MAs in the outer membrane contributes to the envelope integrity and mediates many interactions with the host’s immune system (Etienne et al., 2009; Rousseau et al., 2003). Due to their surface-exposure, these highly antigenic glycolipids offered powerful tools for serodiagnosis of mycobacterial infections and some innovative immunotherapies (Brennan, 1989; Doherty, 2012). More recently, the dynamic properties of TDM as a function of temperature were investigated, indicating that the interaction between mycolates, their flexibility and stability can relate to the trehalose moiety itself. These properties also partially explain antibiotic impermeability and mycobacterial resistance to acid and thermal stresses (Migliardo et al., 2014). Additional links between TDM and stress resilience emerged from the recent identification of the function of TDM hydrolases with a role in M. tuberculosis intracellular growth (Yang et al., 2014). It was reported that the TDM-hydrolytic enzymes were induced by nutritional stress to favour nutrient influx across the remodelled cell envelope, with a concomitant decrease in the tolerance to multiple stresses in cultures and biofilms (Yang et al., 2014). Indeed, these enzymes were found to be crucial for balancing M. tuberculosis growth in the intracellular environment, TDM hydrolysis, nutrient acquisition and stress tolerance, both in immunocompetent and immunocompromised hosts (Yang et al., 2014).

In addition to TDM and its precursor TMM, M. tuberculosis possesses a panoply of other different surface-exposed acyltrehaloses, namely the sulfatides (SLs), diacyltrehaloses (DATs), triacyltrehaloses (TATs) and pentacyltrehaloses (PATs), all implicated in pathogen–host crosstalk (Fig. 2, Table 1) (Jackson et al., 2007; Rousseau et al., 2003). These glycolipids have been exclusively detected in members of the M. tuberculosis complex (MTBC), although they were absent from the avirulent M. tuberculosis H37Ra and from the Mycobacterium bovis BCG (Bacillus Calmette–Guérin) vaccine strain. Nevertheless, and regardless of their restricted distribution in virulent M. tuberculosis, the biological significance of their interactions with host phagocytes is still incompletely understood (Jackson et al., 2007; Rousseau et al., 2003).

A highly abundant trehalose-based glycolipid identified over a half-century ago in the M. tuberculosis cell envelope is the tetracylated trehalose-2-sulfate glycolipid named sulfolipid-1 (SL-1) (Fig. 2) (Middlebrook et al., 1959). Although this glycolipid’s levels have been proposed to directly correlate with strain virulence, its biological roles are still not completely defined since the phenotypes of mutants lacking the mature forms of SL-1 are mostly indistinguishable from wild-type strains in mice and guinea pig models (Gilmore et al., 2012; Rousseau et al., 2003). However, SL-1 has been proposed to be involved in the regulation of mycobacterial growth within phagocytes and to be dispensable for cell-envelope integrity (Gilmore et al., 2012). Since the chemical synthesis of SL-1 has been recently accomplished, studies with the pure glycolipid and its variants shall provide additional clues into the role of SL-1 and its components in the molecular interactions of M. tuberculosis with host immune cells (Geerdink & Minnaard, 2014).

LOS represent a different class of trehalose-containing glycolipids identified in some mycobacteria (Fig. 2, Table 1). These unusual glycolipids are composed of oligosaccharide chains linked to trehalose, in turn acylated with at least one polymethylated-branched fatty acid (Etienne et al., 2009; Hunter et al., 1983). Some LOS structures, differing from species to species, have been identified in the MTBC member ‘Mycobacterium canettii’ (but not in M. tuberculosis) and in some opportunistic NTM such as Mycobacterium
gordonae, M. kansasii, M. szulgai, M. gastri, M. mucogenicum and M. marinum (Alibaud et al., 2014). LOS comprise a series of glycolipids (LOS-I to LOS-IV represent sequential intermediates differing in the number of monosaccharide units) of which LOS-IV is the surface-exposed mature form with highly antigenic properties. Recently, a novel early precursor of the LOS series was identified in M. marinum and named LOS-0 (Alibaud et al., 2014). Unlike the surface-exposed LOS-IV identified in several Mycobacterium strains, the Mycobacterium smegmatis LOS were found to be located in a deeper layer of the cell envelope (Table 1) (Etienne et al., 2009). Although LOS have been implicated in virulence, colony morphology, sliding motility, biofilm formation and infection of murine macrophage-like cells, their biological functions remain only partially characterized (Alibaud et al., 2014; Rombouts et al., 2011). Although LOS have been implicated in virulence, colony morphology, sliding motility, biofilm formation and infection of murine macrophage-like cells, their biological functions remain only partially characterized (Alibaud et al., 2014; Rombouts et al., 2011). In the 1980s, two unusual trehalose-derived acylated oligosaccharides were isolated from M. smegmatis (Table 1) and considered putative precursors of LOS (Saadat & Ballou, 1983). Years later, Besra and colleagues characterized related glycosylated acyltrehaloses from M. smegmatis strains, which were also considered putative intermediates of the LOS series (Besra et al., 1994). Additional oligosaccharides elaborated from trehalose were further purified from this organism’s cytoplasm (Table 1), and due to their structural similarity to the sucrose derivatives raffinose and stachyose involved in maturation drying in plants they were proposed to be cellular stabilizers during mycobacterial dormancy (Ohta et al., 2002). While some of these oligosaccharides had previously been isolated from corynebacteria, yeasts and insects, and even from Mycobacterium fortuitum, two of the tetrasaccharides had never been detected in biological systems (Table 1) (Ohta et al., 2002). The identification of their metabolic and structural roles, as well as the key glycosyltransferases, will allow new insights into their roles in mycobacterial glycobiology.

Since the early detection of trehalose and characterization of the mycobacterial TDM, the portfolio of trehalose-based molecules in these organisms has expanded considerably although a significant fraction still remains to be associated with genuine functions or cellular structures. These include

<table>
<thead>
<tr>
<th>Glycolipid/oligosaccharide</th>
<th>Function</th>
<th>Organism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mycolates</td>
<td>Cell wall assembly/immune modulation/survival in macrophages</td>
<td>Mycobacteria</td>
<td>Ishikawa et al. (2009); Takayama et al. (2005); Welsh et al. (2013)</td>
</tr>
<tr>
<td>Trehalose monomycolate (TMM)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trehalose dimycolate (TDM)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polycyctrehaloses</td>
<td>Virulence modulation</td>
<td>M. tuberculosis</td>
<td>Hatzios et al. (2009); Jackson et al. (2007)</td>
</tr>
<tr>
<td>Diacyctrehalose (DAT)*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triacyctrehalose (TAT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pentaacyctrehalose (PAT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfolipid-1 (SL-1)</td>
<td>Virulence modulation</td>
<td>M. tuberculosis</td>
<td>Bhatt et al., (2007); Gilmore et al. (2012); Kumar et al. (2007)</td>
</tr>
<tr>
<td>Lipoooligosaccharides (LOS)</td>
<td>Modulation of phagocytosis/ virulence/colony morphology/ sliding motility/biofilm formation</td>
<td></td>
<td>Alibaud et al. (2014); Besra et al. (1993); van der Woude et al. (2012)</td>
</tr>
<tr>
<td>LOS-0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOS-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOS-II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOS-III</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOS-IV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trehalose-based oligosaccharides</td>
<td>Unknown</td>
<td>M. smegmatis</td>
<td>Besra et al. (1994); Narumi &amp; Tsunita (1967); Ohta et al. (2002); Saadat &amp; Ballou (1983)</td>
</tr>
<tr>
<td>Trehalose 6,6’-dimannosyl-phosphate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trehalose 6-mannosylphosphate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trehalose-containing pentasaccharide</td>
<td>Unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-Glucosyl-(1→4)-trehalose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-Glucosyl-(1→6)-trehalose*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-Glucosyl-(1→6)·β-glucose-(1→6)-trehalose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-Glucosyl-(1→4)·trehalose·(6→1)-galactose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-Galactosyl-(1→6)·galactosyl-(1→6)-trehalose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycosylated acyltrehaloses</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Also isolated from M. fortuitum.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>†Located in a deeper layer of M. smegmatis cell envelope, not surface-exposed.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
some unique trehalose-derived molecules identified in M. bovis BCG by Koyu Narumi and Toru Tsumita in the 1960s that have not been studied since. These rare molecules are apparently derived from trehalose 6-phosphate and include trehalose 6,6'-diphosphate, trehalose 6-mannosylphosphate and trehalose 6,6'-dimannosylphosphate (Table 1) (Narumi & Tsumita, 1964, 1965, 1967). Although these phosphorylated trehalose derivatives appear to be precursors of larger structures, additional research is required to reveal new clues on their biological roles.

### Trehalose biosynthesis in mycobacteria

Since trehalose is not synthesized in mammals (although trehalases for the specific hydrolysis of the trehalose present in food sources such as plants, mushrooms, crustaceans or insects are expressed in the intestinal epithelia), the mycobacterial resources for biosynthesis of this essential disaccharide and for the trehalose-containing glycolipids represent attractive targets for the development of new anti-mycobacterial agents. To date, five different pathways for trehalose biosynthesis have been identified in nature (Elbein et al., 2003; Nobre et al., 2008): the OtsA/OtsB, TreY/TreZ, TreS, TreP and TreT pathways. While the OtsA/OtsB pathway has the wider distribution in the three domains of life and is the only pathway available to eukaryotes, the extremely radiation-resistant bacterium *Rubrobacter xylanophilus* is the only known organism to possess archetypal genes of four of these pathways (Nobre et al., 2008).

### OtsA/OtsB

This widely distributed biosynthetic pathway operates through the enzyme OtsA (or Tps, trehalose-6-phosphate

---

**Table 2. Genes and enzymes involved in trehalose-related pathways in *M. tuberculosis* H37Rv**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv3490</td>
<td>OtsA/Tps, trehalose-6-phosphate synthase</td>
<td>Trehalose biosynthesis</td>
<td>Elbein et al. (2003)</td>
</tr>
<tr>
<td>Rv1563c</td>
<td>TreY, maltooligosyl-trehalose synthase</td>
<td>Trehalose biosynthesis</td>
<td>De Smet et al. (2000)</td>
</tr>
<tr>
<td>Rv1562c</td>
<td>TreZ, maltooligosyl-trehalose trehalohydrolase</td>
<td>Trehalose biosynthesis</td>
<td>De Smet et al. (2000)</td>
</tr>
<tr>
<td>Rv1564c</td>
<td>TreX (GlgX), glycoin-debranching enzyme</td>
<td>Glycogen debranching (putative)</td>
<td>Kalscheuer et al. (2010a); Seibold &amp; Eikmanns (2007)</td>
</tr>
<tr>
<td>Rv0126</td>
<td>TreS, trehalose synthase</td>
<td>Maltose biosynthesis</td>
<td>Miah et al. (2013)</td>
</tr>
<tr>
<td>Rv2402</td>
<td>Trehalase</td>
<td>Trehalose catabolism</td>
<td>Carroll et al. (2007)</td>
</tr>
<tr>
<td>Rv1235−Rv1238†</td>
<td>LpqY−SugA−SugB−SugC transporter system</td>
<td>Trehalose recycling</td>
<td>Kalscheuer et al. (2010b)</td>
</tr>
<tr>
<td>Rv3804c</td>
<td>FbpA, mycolyltransferase</td>
<td>TDM/mAGP assembly</td>
<td>Belisle et al. (1997)</td>
</tr>
<tr>
<td>Rv1886c</td>
<td>FbpB, mycolyltransferase</td>
<td>TDM/mAGP assembly</td>
<td>Armitige et al. (2000)</td>
</tr>
<tr>
<td>Rv0129c</td>
<td>FbpC, mycolyltransferase</td>
<td>TDM/mAGP assembly</td>
<td>Ibrahim et al. (2012)</td>
</tr>
<tr>
<td>Rv0206c</td>
<td>MmpL3, membrane-associated transporter</td>
<td>TMM translocation</td>
<td>Grzegorzekwicz et al. (2012)</td>
</tr>
<tr>
<td>Rv3451</td>
<td>Cut3, TDM hydrolase</td>
<td>TDM hydrolysis</td>
<td>Yang et al. (2014)</td>
</tr>
<tr>
<td>Rv3452</td>
<td>Cut4, TDM hydrolase</td>
<td>TDM hydrolysis</td>
<td>Yang et al. (2014)</td>
</tr>
<tr>
<td>Rv0295c</td>
<td>SfrO, sulfo transferase</td>
<td>SL-1 biosynthesis</td>
<td>Mougous et al. (2004)</td>
</tr>
<tr>
<td>Rv3822</td>
<td>Chp1, membrane-associated acyltransferase</td>
<td>SL-1 biosynthesis/export</td>
<td>Seeliger et al. (2012)</td>
</tr>
<tr>
<td>Rv3821</td>
<td>Sap, sulfolipid-1-addressing protein</td>
<td>SL-1 biosynthesis/export</td>
<td>Seeliger et al. (2012)</td>
</tr>
<tr>
<td>Rv3823c</td>
<td>MmpL8, membrane-associated transporter</td>
<td>SL-1 export</td>
<td>Converse et al. (2003); Domenech et al. (2005)</td>
</tr>
<tr>
<td>Rv3824c†</td>
<td>PapA1, acyltransferase</td>
<td>SL-1 biosynthesis</td>
<td>Bhatt et al. (2007);</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kumar et al. (2007)</td>
</tr>
<tr>
<td>Rv3820c‡</td>
<td>PapA2, acyltransferase</td>
<td>SL-1 biosynthesis</td>
<td>Bhatt et al. (2007);</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kumar et al. (2007)</td>
</tr>
<tr>
<td>Rv1182</td>
<td>PapA3, acyltransferase</td>
<td>PAT biosynthesis</td>
<td>Hatzios et al. (2009)</td>
</tr>
<tr>
<td>Rv1183‡</td>
<td>MmpL10, membrane-associated transporter</td>
<td>PAT export</td>
<td>Hatzios et al. (2009)</td>
</tr>
<tr>
<td>MMAR_2313</td>
<td>LosA, glycolyltransferase</td>
<td>LOS biosynthesis§</td>
<td>Burguère et al. (2005)</td>
</tr>
<tr>
<td>MMAR_2343</td>
<td>PapA4, polyketide-synthese-associated protein</td>
<td>LOS biosynthesis§</td>
<td>Rombouts et al. (2011)</td>
</tr>
<tr>
<td>MMAR_2309</td>
<td>UdglL, UDP-glucose dehydrogenase</td>
<td>LOS biosynthesis§</td>
<td>Ren et al. (2007)</td>
</tr>
<tr>
<td>MMAR_2333</td>
<td>WcaA, glycolyltransferase</td>
<td>LOS biosynthesis§</td>
<td>Sarkar et al. (2011)</td>
</tr>
<tr>
<td>MMAR_2342</td>
<td>MmpL12, putative membrane transporter</td>
<td>LOS export (putative)§</td>
<td>Rombouts et al. (2011)</td>
</tr>
</tbody>
</table>

*Essential for *M. tuberculosis* H37Rv *in vitro* growth (Griffin et al., 2011).
†Required for survival in murine macrophages or mouse model (Rengarajan et al., 2005).
‡Essential for *in vitro* growth on cholesterol (Griffin et al., 2011).
§Trehalose-containing LOS are absent from *M. tuberculosis*.
||The *M. tuberculosis* H37Rv orthologue (Rv1528c) encodes a truncated PapA4 protein.
synthase) that condenses NDP-glucose and glucose-6-phosphate into a phosphorylated compound, trehalose-6-phosphate (T6P). This intermediate is subsequently dephosphorylated into trehalose by a specific trehalose-6-phosphate phosphatase (OtsB or Tpp) (Fig. 1, Table 2). OtsA was initially identified in M. tuberculosis and its activity was found to be stimulated by some high molecular mass polyanions such as heparin (Elbein & Mitchell, 1974). In this organism, the second step in the pathway relies on OtsB2, a specific OtsB (Rv3372) that dephosphorylates T6P to free trehalose (Edavana et al., 2004). An OtsB parologue in this pathogen’s genome (Rv2006) was suggested to lack activity toward T6P and it was proposed instead to participate in the conversion of phosphorylated TMM (TMM-P) to TMM (Table 2) (Takayama et al., 2005). However, experimental evidence for this hypothesis is still missing (Tropis et al., 2005). Although trehalose can be synthesized via other pathways, genes otsA and otsB2 have been proposed to be essential for M. tuberculosis H37Rv growth by saturation transposon mutagenesis and required for survival in mice (Griffin et al., 2011; Murphy et al., 2005).

TreY/TreZ

In this pathway, a terminal maltose unit in a glycogen chain is isomerized to trehalose by a maltooligosyltrehalose synthase (TreY) and subsequently released by a maltooligosyltrehalose trehalohydrolase (TreZ) (Fig. 1, Table 2). TreY catalyses the rearrangement of the χ(1→4)-linkage of the terminal disaccharide at a reducing end in glycogen into an χ(1→1)-linkage and TreZ acts on this terminal moiety to release free trehalose (Elbein et al., 2003). In most organisms with the TreY/TreZ pathway both genes are contiguous and in M. tuberculosis this operon-like cluster has an additional treX gene (Rv1564c) encoding a putative glycogen-debranching enzyme (De Smet et al., 2000). Although the involvement of TreX in trehalose and glycogen metabolism remains to be experimentally validated, the genomic association with the treY/treZ pair and the high sequence similarity with the TreX (GlgX) from Corynebacterium glutamicum encoding a debranching enzyme active towards glycogen, partially support the claim (Seibold & Eikmanns, 2007). Genetic inactivation of TreY/TreZ had no effect on M. tuberculosis growth, in vitro or in mice, indicating that trehalose supply for metabolic requirements was not compromised under the conditions tested (Murphy et al., 2005). Interestingly, TreX was considered essential for M. tuberculosis growth, suggesting that glycogen mobilization/recycling through production of linear χ(1→4)-glucose chains is indispensable for growth (Griffin et al., 2011; Seibold & Eikmanns, 2007). Most mycobacteria possess the TreY/TreZ (M. tuberculosis genes Rv1563c and Rv1562c), the OtsA/OtsB2 (Rv3490/Rv3372) systems and a TreS enzyme (Rv0126) initially proposed to catalyse trehalose synthesis (see below) (Fig. 1, Table 2) (De Smet et al., 2000). Evidence that these pathways were mutually redundant was obtained from the observation that mutants exhibited trehalose-independent growth rates comparable to wild-type strains (Woodruff et al., 2004). In contrast, triple mutants were unable to grow unless trehalose was exogenously supplied, which established the essential role for the disaccharide in mycobacteria (Murphy et al., 2005; Woodruff et al., 2004). Although the independent role for each trehalose pathway in mycobacterial physiology, from the assembly of crucial trehalose-based glycolipids like TDM to the turnover of glycogen for energy metabolism and the response to thermal or osmotic stress, is still a matter of debate, new important clues have recently emerged from in vivo genetic studies, X-ray crystallography and mechanistic details of TreS catalysis (Caner et al., 2013; Chandra et al., 2011; Elbein et al., 2010; Kalscheuer et al., 2010a; Miah et al., 2013; Roy et al., 2013; Zhang et al., 2011).

**TreS**

The enzyme trehalose synthase (TreS) was initially proposed to catalyse the isomerization of the χ(1→4)-linkage of maltose into an χ(1→1)-linkage in trehalose but, unlike TreY, using free maltose as a substrate (Elbein et al., 2003). TreS has also been shown to hydrolyse trehalose and maltose into glucose in vitro (Elbein et al., 2003; Miah et al., 2013; Pan et al., 2004). Moreover, because the incubation of TreS with radioactive glycogen in vitro released both radioactive maltose and trehalose, the enzyme was also proposed to catalyse the hydrolysis of glycogen with an α-amylase domain (Pan et al., 2008). However, this activity in mycobacteria has been recently disputed since the α-amylase activity was only observed in vitro and at a very low rate, indicative of a non-physiological role (Caner et al., 2013). Although it was considered one of the pathways for trehalose biosynthesis in mycobacteria, genetic studies complemented with 1H- and 19F-NMR analyses also refuted a role for TreS in trehalose synthesis from maltose in vivo and confirmed the reverse flow instead (De Smet et al., 2000; Miah et al., 2013). Maltose produced by TreS from free trehalose could then be processed by maltokinase (Mak or Pep2, Rv0127) into maltose-1-phosphate (M1P), which was found to be the substrate of a newly characterized maltosyltransferase (GlgE, Rv1327c) able to elongate glycogen chains and possibly other χ(1→4) glucose polymers such as the capsular x-glucan or the intracellular methyl-glucose lipopolysaccharides (Elbein et al., 2010; Kalscheuer et al., 2010a; Miah et al., 2013). Both Rv0127 (Mak) and Rv1327c (GlgE) genes were considered essential for M. tuberculosis growth (Griffin et al., 2011; Kalscheuer et al., 2010a). The enzyme Mak was biochemically characterized and the anticipated interdependence between trehalose and glycogen levels was further investigated (Chandra et al., 2011; Elbein et al., 2010; Mendes et al., 2010). Elbein and colleagues suggested that during heat stress excess trehalose would be used for glycogen production via TreS-Mak-GlgE pathway while under normal growth conditions glycogen would be synthesized by the GgA-GglB pathway (Elbein et al., 2010). These authors also proposed that a drop in trehalose levels...
below those required for metabolism would elicit glycogen hydrolysis via the ζ-amylase domain of TreS (Elbein et al., 2010; Pan et al., 2008). Whether this hypothetical and elegant trehalose-glycogen circuit involving TreS is regulated by stress, resembling the response of *C. glutamicum* to hyperosmotic challenges, during which the TreY/TreZ pathway acts as the main source of the osmolyte trehalose from stored glycogen, remains to be experimentally demonstrated (Seibold & Eikmanns, 2007). Since in *C. glutamicum* the constitutive recycling of glycogen depends on the activity of a debranching GlgX (TreX), the action of TreY/TreZ in *M. tuberculosis* is also likely to be coordinated with the TreX product encoded in the same operon, especially because the branching pattern of glycogen would block progression of isomerization and release of trehalose by TreY/TreZ beyond the ζ(1→6)-linked branches (Seibold et al., 2009). At this point, however, the data available cannot completely rule out the existence of a glycogen-degrading ζ-amylase activity of TreS under yet unidentified conditions in vivo (Elbein et al., 2010; Pan et al., 2008). New insights into the TreS reaction mechanism and complex formation recently emerged from the 3D structures of the *M. smegmatis* and *M. tuberculosis* TreS, the latter of which were found to form a hetero-octameric complex with Mak to enhance its activity (Caner et al., 2013; Roy et al., 2013). Moreover, the ζ-amylase activity measured by Pan and colleagues with purified TreS was refuted because the expected active site residues were absent from the acarbose binding site (Caner et al., 2013; Pan et al., 2008). It should be mentioned that in a previous study, TreS was proposed to have two different active sites that would justify the lack of ζ-amylase activity without affecting the observed trehalose-maltose interconverting activity (Pan et al., 2004). Interestingly, Caner and colleagues identified a C-terminal carbohydrate-binding module that could be relevant for glycogen recognition and assist polymerization or hydrolysis (Caner et al., 2013). Hypothetically, the mycobacterial TreS could undergo a stress-driven conformational change or a post-translational modification or even associate with an enzyme (or metabolite) that might activate its ζ-amylase activity, in a manner similar to the enhancement of Mak activity by association with TreS in vitro, which may be relevant in vivo (Roy et al., 2013).

In most mycobacterial species TreSs are encoded by genes (Rv0126 in *M. tuberculosis* H37Rv) adjacent to the mak gene (see above) and the TreSs from *M. smegmatis* and *M. tuberculosis* have been recently studied in detail (Pan et al., 2004, 2008; Zhang et al., 2011; Miah et al., 2013; Roy et al., 2013; Caner et al., 2013). Bifunctional TreS–Mak versions, with the TreS domain sharing 50–60 % amino acid identity to the mycobacterial TreSs, are common in many organisms including some *Pseudomonas* spp. where they were implicated in trehalose synthesis in vivo, although one of these bifunctional enzymes has been confirmed to synthesize maltose-1-phosphate from trehalose and ATP in vitro (our unpublished results) (Freeman et al., 2010). A few mycobacterial genomes, namely those from *Mycobacterium vanbaalenii*, *M. vaccae* or *M. gilvum*, lack the typical treS genes described above but each has a distantly located gene for a different TreS (Mvan_5178 in *M. vanbaalenii*). This putative TreS has very low amino acid identity (10–15%) with the ‘typical’ TreS but it is closely related (~50% identity) to TreS found in unrelated bacteria, including pseudomonads. The ‘atypical’ TreS from *Pseudomonas stutzeri* was characterized and also confirmed to catalyse the interconversion of trehalose and maltose (Lee et al., 2005). However, and despite the fact that both types of TreS co-exist in some of these bacteria with medical, agricultural and environmental significance, trehalose biosynthesis has not been investigated in detail (Freeman et al., 2010). Remarkably, sequences of both types of TreS co-exist in the hydrocarbon-degrading environmental *Mycobacterium* strains JLS, KLS and MCS. Although their catalytic and physiological functions remain to be examined, namely the involvement in glycogen hydrolysis with their ζ-amylase activity or synthesis via the TreS–Mak–GlgE route, these genomic rarities merit investigation (Miller et al., 2004). Not surprisingly, this distribution seems to suggest horizontal gene transfer between these genera, whose strains are increasingly detected in domestic water-distribution systems and in hospital settings, where MDR phenotypes increasingly disseminate (Fitzpatrick, 2009; Wang et al., 2012).

**TrebP**

In this reaction, an enzyme initially identified in fungi and named trehalose phosphorylase (TreP) can synthesize trehalose from glucose and glucose-1-phosphate (Eis et al., 2001). This unusual TreP activity has been observed in vitro, but its relevance in vivo is still unconfirmed. An archetypical TreP has not, thus far, been detected in mycobacteria, but according to the Cazy database (www. cazy.org) *M. tuberculosis* Rv3401 and orthologues are classified as glycoside hydrolases of family 65 (GH65), which groups different phosphorylases with specificities toward the related disaccharides maltose, kojibiose, nigerose and trehalose. Hence, characterization of Rv3401 orthologues is required to ascertain their substrate specificities and to try to confirm the possibility of putative TrePs in mycobacteria.

**TrebT**

Unlike OtsA, this rare trehalose glycosyltransferring synthase (TreT) characterized in hyperthermophilic archaea and in *R. xylanophilus* catalyses the direct synthesis of trehalose from NDP-glucose and glucose instead of glucose-6P (Nobre et al., 2008). To the best of our knowledge, *R. xylanophilus* is a rare example of an actinobacterium encoding a TreT, where it was shown to be active along with the OtsA/OtsB pathway. Whether the related mycobacteria possess a non-homologous TreT with similar substrate specificity remains unknown.

The expanding intricacy of trehalose biosynthesis in mycobacteria, the catalytic and structural details of the
interventional enzymes, as well as the regulatory phenomena involved in their expression, function and molecular interactions, are far from being resolved and warrant future research.

**Mycobacterial enzymes for trehalose catabolism**

*M. tuberculosis* and most saprophytic mycobacteria with sequenced genomes possess trehalases (Fig. 1, Table 2) that, in addition to TreSs with their proposed trehalose-hydrolysing and maltose-trehalose isomerization activities, may be important to balance the intracellular levels of trehalose during specific growth conditions (Elbein et al., 2010; Pan et al., 2004). As a natural consequence of *Mycobacterium leprae* genomic reductive evolution, trehase and TreS-encoding genes are absent from the genome, suggesting that trehalase could not be hydrolysed or, instead, that a glycoside hydrolase with trehalose-hydrolysing activity may split the disaccharide (Alarico et al., 2008; Singh & Cole, 2011).

Although the conversion of trehalose into glycogen through the TreS–Mak–GlgE anabolic pathway allows storage of energy and carbon, the activity of trehalase leads to the direct hydrolysis of trehalose, suggestive of energy and carbon mobilization mechanisms (Carroll et al., 2007; Elbein et al., 2010). However, the metabolic significance of trehalase function and a hypothetical coordinated regulation with the TreS–Mak–GlgE pathway in vivo are still elusive. Such hypothetical mechanisms could be important for rapid mobilization of carbon/energy reserves in a response to stress, or when the trehalose levels require tight regulation, for example during osmoadaptation (Elbein et al., 2010). Indeed, a trehalase-negative mutant *M. smegmatis* could not be recovered, suggesting that the trehalase-encoding gene is essential for growth (Carroll et al., 2007). Interestingly, the *M. tuberculosis* trehalase-encoding gene (*Rv2402*) was not considered essential by saturation transposon mutagenesis (Table 2), which could, hypothetically, point to an overlapping function with TreS (Griffin et al., 2011). Mycobacterial trehalases are atypical glycoside hydrolases whose activity is dependent on phosphate ions, although they did not exhibit phosphorylase activity in the conditions tested (Carroll et al., 2007). Although the role of phosphate in the activity was suspected to be related to conformational stability, the *M. smegmatis* trehalase was shown to be inhibited by the related polyanion polyphosphate. This observation and the reported stimulation of the TEP-synthesizing OtsA activity by polyanions could favour trehalose synthesis and accumulation under unknown conditions, provided the specific phosphatase OtsB could be functional under those conditions (Elbein & Mitchell, 1974). However, this remains a subject for future research.

**Trehalose recycling in mycobacteria**

Stress-resistant and drug-tolerant *M. smegmatis* biofilms formed during growth under hostile conditions are rich in free MAs, the hydrolytic products released by a specific serine esterase (MSMEG_1529) characterized recently as one of two *M. tuberculosis* TDM hydrolases (Rv3451 and Rv3452) (Ojha et al., 2010; Yang et al., 2014). Whether this esterase also elicits trehalose release from TMM remains to be demonstrated. Moreover, the transfer of mycolyl chains by Ag85 mycolyltransferases to AG during cell wall assembly, or from TMM to form TDM, also leads to the release of trehalose from the cells (Kalscheuer et al., 2010b; Takayama et al., 2005). Remarkably, this leakage is partially averted by a highly specific trehalose transporter system identified recently in *M. tuberculosis* where it was shown to recycle a fraction of the otherwise lost trehalose back into the cytoplasm (Kalscheuer et al., 2010b). Whether this mechanism recycles the trehalose released during TDM synthesis by the Ag85 complex or also captures the trehalose liberated from TDM hydrolysis by TDM hydrolases is not known (Yang et al., 2014). This ATP-binding cassette (ABC) transporter system, with partial homology to sugar transporters found in other bacteria, is composed of subunits LpqY–SugA–SugB–SugC (LpqY–SugABC) (*Rv1235–Rv1238*) (Fig. 1, Table 2) and spans the plasma membrane, the PG and AG layers (Fig. 1) (Kalscheuer et al., 2010b). Since trehalose is not synthesized in mammals, it is unlikely that this capture system evolved for its acquisition from the human host, although this crucial function might have prevailed as a selective advantage for ancient environmental RGM, presumed ancestors of *M. tuberculosis* (Devulder et al., 2005). In this evolutionary scenario, extant environmental mycobacteria may still rely on this system to uptake trehalose released by other organisms, for example during hypoosmotic shocks or due to cell death, and use it for glycolipids biosynthesis, as an energy and carbon supply or as a stress-counteracting solute (Alarico et al., 2013). However, NTM in general and *M. gilvum* and *M. vanbaalenii* in particular seem to lack one or several subunits of this system (Table 3), which suggests that trehalose recycling may be impaired in these organisms or that the substrate specificity of this putative transporter in these NTM species may be different. Anyhow, the high affinity of the LpqY–SugABC system for trehalose capture again highlights the extremely important role for this disaccharide in *M. tuberculosis* physiology within mammals in general and in the human host in particular (Kalscheuer et al., 2010b). Although the corresponding genes have not been considered essential for growth, their importance for bacterial survival in macrophages was confirmed. Indeed, a dramatic impairment of *M. tuberculosis* virulence arising from the interference with this trehalose recycling mechanism rendered the whole system an attractive target for drug discovery (Kalscheuer et al., 2010a; Renganarajan et al., 2005).

**Enzymes involved in the biosynthesis of trehalose-containing glycolipids**

A substantial amount of work has been carried out in the last decade to elucidate the details of how MAs and
### Table 3. Distribution of genes involved in trehalose-related pathways across selected species of the *M. tuberculosis* complex (MTBC), of NTM and in *M. leprae*

<table>
<thead>
<tr>
<th>H37Rv gene</th>
<th>Protein</th>
<th><em>M. tuberculosis</em></th>
<th><em>M. africanum</em></th>
<th><em>M. bovis</em></th>
<th><em>M. leprae</em></th>
<th><em>M. ulcerans</em></th>
<th><em>M. abscessus</em></th>
<th><em>M. avium</em></th>
<th><em>M. intracellulare</em></th>
<th><em>M. marinum</em></th>
<th><em>M. smegmatis</em></th>
<th><em>M. vanbaalenii</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Trehalose</td>
<td>Rv3490</td>
<td>OtsA</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3910</td>
<td>OtsB2</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Hydrolysis</td>
<td>Rv3372</td>
<td>OtsB2</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv1563c</td>
<td>TreY</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Recycling</td>
<td>Rv1562c</td>
<td>TreZ</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv1564c</td>
<td>TreX</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv0126</td>
<td>TreS</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv2402</td>
<td>Trehalase</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv1235-Rv1238</td>
<td>LpqY/SugA/B/C</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Mycolates</td>
<td>Rv3804c</td>
<td>FbpA</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv1886c</td>
<td>FbpB</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv0129c</td>
<td>FbpC</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv2060c</td>
<td>MmpL3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3451</td>
<td>TDM hydrolase</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3452</td>
<td>TDM hydrolase</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PAT / SL-1</td>
<td>Rv0295c</td>
<td>Sf0</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3822</td>
<td>Chp1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3821</td>
<td>Sap</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3823c</td>
<td>MmpL8</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3824c</td>
<td>PapA1</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv3820c</td>
<td>PapA2</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv1182</td>
<td>PapA3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>Rv1183</td>
<td>MmpL10</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>LOS</td>
<td>MMAR_2313</td>
<td>LosA</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>MMAR_2343</td>
<td>PapA4</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>MMAR_2309</td>
<td>UdgL</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>MMAR_2333</td>
<td>WcaA</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td></td>
<td>MMAR_2342</td>
<td>MmpL12</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

- ●, Highly conserved sequences (60–100% amino acid identity to the *M. tuberculosis* proteins). LOS biosynthetic proteins are compared to the *M. marinum* orthologues.
- ●, Sequences with moderate sequence identity (30–60%) (number of paralogues in parentheses); ○, Absence of conserved sequences (<30% identity).
- a, Strain AF2122/97 the N- and C-terminal domains of TreY are independent but the BCG strains encode full-length TreY orthologues.

Trehalose mycolates are synthesized (Marrakchi et al., 2014; Takayama et al., 2005). The model proposed by Takayama and colleagues argued that the mycolyl group of a newly synthesized MA could be transferred to T6P by unidentified mycolyltransferases through a mycolyl-mannosyl-phosphoheptaprenol intermediate to yield phosphorylated TMM (TMM-P). This intermediate would be subsequently dephosphorylated to TMM by an unknown phosphatase (hypothetically OtsB1) and shuttled across the plasma membrane by a transporter, which we now know to be assembled by MmpL3 encoded by Rv0206c in *M. tuberculosis* H37Rv (Table 2) (Grzegorzekicz et al., 2012; Takayama et al., 2005). However, the molecular events in this model based on bioinformatic analyses remain to be experimentally confirmed and it is possible that trehalose (and not T6P) may instead be the genuine substrate for TMM synthesis (Varela et al., 2012). Indeed, in vivo synthesis of mycolates in the closely related bacterium C.
Glutamicum does not seem to involve T6P (Tropis et al., 2005). It is also possible that the rare trehalose-6,6'-diphosphate, trehalose-6-mannosylphosphate and trehalose-6,6'-dimannosylphosphate isolated from M. bovis BCG may be putative intermediates in unknown steps of TDM biosynthesis (Narumi & Tsumita, 1965, 1967). However, this remains hypothetical and the chemically synthesized version of trehalose-6,6'-diphosphate could provide important clues to help solve this enigma (Ronnow et al., 1994).

In the subsequent steps of the currently accepted model, TDM is elaborated from TMM by mycolyltransferases of antigen 85 (Ag85) complex, a family of secreted and highly similar fibronectin-binding proteins (Fbps) encoded by genes fbpA (Rv3804c), fbpB (Rv1886c) and fbpC (Rv0129c) (Table 2). These enzymes catalyze the transesterification of the lipid moiety of TMM to another TMM molecule to yield TDM, or to AG to assemble the mAGP complex (Armitige et al., 2000; Belisle et al., 1997; Ibrahim et al., 2012; Kaur et al., 2009). In this process, a trehalose molecule is released and may be recycled back into the cytoplasm by the specific LpqY–SugABC transporter system identified recently (see above). A fourth fbpD (Rv3803c) parologue was found to be catalytically inactive and, in turn, it was suggested to play a role in host-tissue attachment (Wilson et al., 2004). The same way fbpA- and fbpB-negative mutants produce normally mycolylated cell walls, mycolate levels could also be independently restored by fbpA, fbpB or fbpC complementation, implying functional redundancy for the three fbp paralogues (Puech et al., 2002). Although the Ag85 proteins seem to be specific for TMM as mycolyl donor, a slight preference for different acceptors has been identified since Ag85C preferably transfers MAs to AG (see Table 2) (Mougous et al., 2009). In this process, a trehalose as consequence of inactivation of the corresponding biosynthesis pathways, mycolates can be synthesized from glucose, maltose or maltotriose (Tropis et al., 2005). Interestingly, and unlike mycobacteria, these mutants were viable, indicating that this disaccharide is not essential for their growth. In view of this sugar substrate relaxation and based on the clustering of the mycolyltransferase Ag85C-encoding gene (Rv0129c) with genes encoding TreS (Rv0126) and Mak (Rv0127) in M. tuberculosis, it is possible that maltose (or even maltose-1-phosphate) could be an intermediate in the synthesis of maltose monomycolate (MMM) or maltose dimycolate (MDM) under unknown conditions; however, this remains a hypothesis.

The biosynthesis of sulfolipid-1 (SL-1), a different trehalose-based glycolipid of M. tuberculosis, has been shown to be initiated by a sulfotransferase (Stf0) (Rv0295c) that specifically converts trehalose into trehalose-2-sulfate (Fig. 3, Table 2) (Mougous et al., 2004). The elucidation of the final steps of SL-1 assembly in M. tuberculosis revealed that biosynthesis was linked to transport across the plasma membrane (Seeliger et al., 2012), with the transmembrane protein MmpL8 (Rv3823c) playing a crucial role in the export of intermediate forms of SL-1 to the outer leaflet of the plasma membrane where maturation takes place (Domenech et al., 2005). Furthermore, the polyketide synthase Pks2 (Rv3825c), a membrane-associated acyltransferase (Chp1, Rv3822), and an SL-1-addressing integral membrane protein (Sap, Rv3821) in the same genetic cluster have also been implicated in SL-1 processing (Fig. 3, Table 2) (Seeliger et al., 2012). Genes papA1 (Rv3824c) and papA2 (Rv3820c) encode acyltransferases PapA1 and PapA2 involved in the acylation steps of SL-1 biosynthesis. Both genes are paralogues of papA3 (Rv1182), which encodes a similar acyltransferase (PapA3) that uses the products of polyketide synthase Pks3/4 to catalyse two esterifications of trehalose to DAT, the precursor of PAT (Fig. 3, Table 2) (Bhatt et al., 2007; Hatzios et al., 2009; Kumar et al., 2007). M. tuberculosis H37Rv strain does not synthesize PAT because Pks3/4 gene is interrupted by a stop codon,
N- and C-terminal domains are encoded by \textit{Rv1180} and \textit{Rv1181}. The disruption of Stf0 or PapA3 abolished SL-1 or PAT, respectively, from \textit{M. tuberculosis} lipid extracts (Hatzios et al., 2009; Mougous et al., 2004). Although their functions seem to overlap, the genetic machinery and biosynthetic systems for PAT and SL-1 assembly are functionally independent (Fig. 3) (Hatzios et al., 2009).

The pathway for LOS biosynthesis was only investigated later, possibly because these glycolipids were rarely detected among strains of the MTBC, with the exception of ‘\textit{M. canettii}’ (Daffe et al., 1991). Several glycosyltransferases involved in the glycosylation steps of LOS maturation (LOS-0 to LOS-IV) have been identified in a genetic cluster of \textit{M. marinum} (Fig. 4) (Alibaud et al., 2014; Burguière et al., 2005; Sarkar et al., 2011). The \textit{M. marinum} glycosyltransferase encoded by \textit{MMAR}_2313 (\textit{Rv1500} in \textit{M. tuberculosis} H37Rv) and designated LosA, transfers a dideoxygalactose derivative to form the mature oligosaccharide component of LOS-IV (Fig. 4) (Burguière et al., 2005). The acylation steps were assigned to the \textit{M. marinum} 478 amino acids polyketide synthase-associated protein PapA4 (\textit{MMAR}_2343), whose orthologue in \textit{M. tuberculosis} encodes a truncated enzyme with only 165 amino acids (Fig. 4, Table 2) (Rombouts et al., 2011). Also, the putative PapA3 gene (\textit{MMAR}_2355) located in the vicinity of PapA4 and sharing 50 % amino acid identity with the \textit{M. tuberculosis} \textit{Rv3824c} (PapA1), one of the SL-1 acyltransferases, was also considered to participate in LOS acylation (van der Woude et al., 2012). Additionally, and due to its clustering with papA4, the \textit{MmpL12} transporter encoded by \textit{MMAR}_2342 (\textit{Rv1522c} in \textit{M. tuberculosis} H37Rv) was proposed to catalyse translocation of LOS across the plasma membrane (Rombouts et al., 2011). Also in this cluster, the \textit{MMAR}_2352 product sharing 33 % amino acid identity with the \textit{M. tuberculosis} \textit{MmpL8}-associated protein \textit{Sap} (\textit{Rv3821}) involved in SL-1 biosynthesis could assist LOS translocation (Figs 3 and 4).

Interestingly, the presence of LOS in ‘\textit{M. canettii}’, the only member of the MTBC known to produce these glycolipids, could be partially explained by the presence of a full-length PapA4 in the genome, in addition to the truncated version found in \textit{M. tuberculosis} (Rombouts et al., 2011).

\textbf{Fig. 3.} Proposed pathways for the synthesis of PAT and SL-1 in \textit{M. tuberculosis}. Stf0, sulfotransferase; PapA1, PapA2, PapA3, acyltransferases; Chp1, membrane-associated acyltransferase; Sap, SL-1-addressing integral membrane protein; MmpL8 and MmpL10, transmembrane transporters. Dashed arrows indicate putative function. Adapted from Hatzios et al. (2009) and Seeliger et al. (2012).
Considering that *M. tuberculosis* H37Rv does not produce LOS, it is intriguing that gene *Rv1522c* (encoding MmpL12 transporter with 60% amino acid identity with MMAR_2342) was indicated as essential for growth (Griffin et al., 2011). It is possible that immature LOS are synthesized in this organism and partially acylated by one of the SL-1 acyltransferases PapA1 (*Rv3824c*) or PapA2 (*Rv3820c*) or even by the PapA3 of the PAT pathway (*Rv1182*) and translocated by MmpL12 (*Rv1522c*). Alternatively, MmpL12 may be involved in the translocation of a different glycolipid.

**Trehalose-containing glycolipids as immune modulators and vaccine adjuvants**

Trehalose-containing glycolipids found in the cell envelope seem to either protect mycobacteria from harsh environmental conditions or act as important virulence factors during pathogen–host interactions (Alibaud et al., 2014; Etienne et al., 2009; Harland et al., 2008; Hatzios et al., 2009; Stanley & Cox, 2013; Welsh et al., 2013). TDM is the most abundant glycolipid in the mycobacterial cell envelope and one of their major virulence effectors (Sakamoto et al., 2013; Welsh et al., 2013). Although all known mycobacteria (except *M. leprae*) synthesize TDM, there are many physico-chemical and antigenic structural differences between species, which affect the host immune responses and strongly correlate with strain virulence (Fujita et al., 2005; Glickman, 2008). In addition to this, TDM modulates macrophage gene expression and plays a role in preventing phagosome–lysosome fusion and inhibition of phagolysosome acidification that favours mycobacterial survival inside these macrophages’ compartments (Sakamoto et al., 2013; Welsh et al., 2013). TDM has also been shown to have potent immunostimulatory activity that imparts granuloma formation, adjuvant activity and tumour regression properties (Flynn & Chan, 2005; Lang, 2013; Welsh et al., 2013). Recently, two essential receptors responsible for the recognition of TDM by macrophages were identified. One of these, Mincle (macrophage inducible Ca²⁺-dependent C-type lectin), was demonstrated to be essential for activation of macrophages by TDM *in vitro* and for granuloma formation *in vivo* (Ishikawa et al., 2009). The macrophage C-type lectin (MCL), which seems to have emerged from Mincle by gene duplication (Lang, 2013; Miyake et al., 2013; Welsh et al., 2013). The 3D structures of these receptors were recently...
determined and the TDM-binding details deduced (Furukawa et al., 2013).

The TDM analogues glycerol monomycolate or trehalose-dicorynomycolate have been used as immunostimulants and vaccine adjuvants lacking the adverse effects of heat-killed mycobacteria (Andersen et al., 2009). Moreover, they can efficiently reproduce pathophysiological properties of M. tuberculosis and trigger an immunological memory comparable to live M. bovis BCG (Rosenkrands et al., 2005). The TDM synthetic analogue trehalose-dibehenate (TDB) is also a potent vaccine adjuvant in induction of T-cell immune responses through Mincle receptors (Schoenen et al., 2010). Clinical trials for TB subunit vaccination using TDB as adjuvant are in progress.

Sulfated trehalose esters or sulfatides and their primary representative SL-1 have been the focus of many studies indicating a role in M. tuberculosis virulence and in other members of the MTBC, to which they are restricted. Over the years, SL-1 has been proposed to affect phagosomes–lysosome fusion mechanisms in vitro, and also to modulate cytokines production and the oxidative response of human leukocytes, and even to suppress mitochondrial oxidative phosphorylation (Gilmore et al., 2012; Jackson et al., 2007). Surprisingly, in mice and guinea pig infection models, strains lacking mature SL-1 lacked obvious differences in virulence when compared to wild-type strains (Rousseau et al., 2003). In order to ascertain the genuine role of SLs in M. tuberculosis, an SL-1 deficient strain was constructed, in which the first committed step in the pathway was disrupted (Gilmore et al., 2012). The results confirmed that the SL-1 mutant had enhanced survival in human macrophages and increased resistance to the human antimicrobial peptide LL-37, but it also became evident that murine macrophages and murine models of TB lacked the human phenotype, suggestive of a host-specific contribution of SL-1 to pathogenesis.

SLs, as well as DAT and PAT chemical and structural features, play important roles in the surface properties of the mycobacterial cell envelope (Figs 1 and 2). For example, the antigenic properties of SLs are regulated by the size and location of the acyl chains (Guiard et al., 2009). Moreover, PAT and some forms of DAT have mycolipinic acids as the principal acyl substituents, which were found to be potent inhibitors of leukocyte migration in vitro and B-cell antigens (Jackson et al., 2007). Additionally, a mutant devoid of DAT and PAT was more prone to aggregation in liquid media and infection of host cells, indicating that the absence of these glycolipids altered cell-surface properties, possibly exposing other PAMP structures required for binding and microbial internalization by macrophages (Cambier et al., 2014; Dubey et al., 2002; Rousseau et al., 2003). Since these alterations did not affect the ability of the mutant to replicate and persist in mice, it is possible that other acyltrehaloses compensate for DAT and PAT functions or, like SL-1, they may have a host-specific phenotype not reproduced by the models used (Jackson et al., 2007).

Early work questioned the role of LOS in mycobacterial pathogenicity since an M. kansasii strain unable to synthesize LOS could cause a chronic systemic infection in mice, while strains containing LOS were rapidly eliminated (Belisle & Brennan, 1989). This observation led to the ‘masking hypothesis’ arguing that LOS may be a group of avirulence determinants that coat other glycolipidic components of the cell envelope, namely the antigenic virulence factors LAM and phenolic glycolipids, resembling the role of M. tuberculosis phthiocerol dimycoceroserate in the selection of non-microbicidal macrophages to avoid innate immune-response mechanisms (Belisle & Brennan, 1989; Cambier et al., 2014; Mishra et al., 2011). A similar role for M. smegmatis and M. abscessus glycopeptidolipids has been proposed, wherein the surface-exposed glycopeptidolipids cover the underneath phosphatidyl-mylo-inositol to prevent macrophage activation and facilitate colonization (Nessar et al., 2011; Villeneuve et al., 2003). The LOS masking hypothesis was recently supported by the work of Alibaud and colleagues, which observed an inverse correlation between M. marinum LOS production and phagocytosis by macrophages and amoebae (Alibaud et al., 2014).

Purified LOS were also proposed to inhibit TNF-α secretion by LPS-stimulated human macrophages, indicating a direct interference with the proinflammatory response (Rombouts et al., 2009). Moreover, in vitro studies confirmed that LOS-IV stimulates expression of macrophage cell surface proteins required for the formation of mature granulomas and restriction of bacterial growth (Rombouts et al., 2010). It was also later demonstrated that LOS-IV deficiency and LOS-III accumulation led to a hypervirulent strain, but it was unclear whether these effects were due to altered secretion of PE_PGRS protein, to the deficiency in LOS-IV or to the accumulation of LOS-III (van der Woude et al., 2012). The recently disclosed identity of novel genes of the LOS biosynthetic pathway in M. marinum shall provide additional important tools to examine the significance of these enigmatic glycolipids in mycobacterial pathogenesis (Alibaud et al., 2014).

**Enzymes of trehalose metabolism as targets for anti-mycobacterial therapies**

Mycobacteria are intrinsically resistant to many common antibiotics as they possess mechanisms to limit their access to the intracellular milieu and accumulation or interfere with the action of drugs that manage to enter cells (Viveiros et al., 2012). Examples of efficient mechanisms in this context are the efflux systems that promote extrusion of toxic metabolites including some antibiotics, the substrate promiscuity of acetyltransferases able to inactivate amino-glycosides that renders M. tuberculosis resistant to this class of drugs, and the endogenous protection of targets in a process involving molecular mimicry that, for example, elicits fluoroquinolone resistance (Chen et al., 2011; Hegde et al., 2005; Viveiros et al., 2012).
Since trehalose is an essential player in mycobacterial biology and is not synthesized in mammalian cells, the enzymes involved in the metabolism of trehalose-based elements in these organisms represent attractive targets for chemotherapeutic intervention provided their inactivation inhibits mycobacterial growth (Caner et al., 2013; Swarts et al., 2012). For optimal progression towards new efficient TB chemotherapies, the genetic, functional and structural characterization of validated targets will accelerate drug discovery and unquestionably shorten the path from the laboratory bench to the patient (Dias et al., 2011; La Rosa et al., 2012). Interestingly, some antibiotics effective against M. smegmatis, namely cathomycin, circulin, diumycin and moenomycin, inhibited OtsA activity in vitro (Pan & Elbein, 1996). Some of those antibiotics also inhibited the activity of OtsB, an immunodominant antigen that was considered a possible candidate for vaccine development (Zhang et al., 2007). In addition, the trehalose analogue trehalosamine (2-amino-2-deoxy-trehalose) was inhibitory of M. smegmatis growth as it probably disrupted normal trehalose metabolism (De Smet et al., 2000). However, the cytoplasmic location of these targets renders them less accessible to drugs and the efforts to optimize promising inhibitors into new lead compounds tend to target systems essential for mycobacteria that are preferably located in the outer compartments of cells. Nevertheless, the inner biosynthetic pathways for MAs, glycolipids and lipoglycans found in the mycobacterial cell envelope have long provided important targets for TB chemotherapies (Jackson et al., 2013). The synthetic drug isoniazid is one of the frontline anti-mycobacterials that inhibits the NADH-dependent enoyl reductase (InhA), an essential enzyme of the cytoplasmic FAS-II complex involved in MAs biosynthesis and one of the best validated targets for TB therapy (Morbidoni et al., 2006; Rozwarski et al., 1998). Remarkably, the natural antimicrobial compound pyridomycin produced by strains of Streptomyces and Dactylosporangium also selectively blocks InhA substrate-binding pockets and may be used as a strategy to overcome isoniazid resistance (Hartkoorn et al., 2014). Ethambutol is a bacteriostatic drug administered in combination with isoniazid (and with rifampicin and pyrazinamide) as the first-line therapeutic option for drug-susceptible TB and it disrupts the assembly of AG, a crucial structural network of the cell wall, by inhibiting a key arabinosyltransferase activity (EmbA and EmbB) (Stehr et al., 2014).

An ethambutol-related compound initially synthesized in 2005 and designated SQ109 (Table 4) was also found to impair mAGP assembly at a different target as it blocks the efficient transfer of MAs to the cell wall leading to TMM accumulation in the cytoplasm (Protopopova et al., 2005; Sacksteder et al., 2012). The target for SQ109 was identified as the plasma-membrane-embedded TMM transporter MmpL3 (Fig. 1) and neither mycolate synthesis nor mycolyltransferase activities were affected (Tahlan et al., 2012). Remarkably, SQ109 was also recently shown to inhibit respiration and ATP synthesis by targeting menaquinone biosynthesis enzymes and electron transport components, and also to disrupt the pH gradient and the membrane proton motive force (Li et al., 2014). As such, these results also seem to support the idea that SQ109 inhibition of TMM translocation results from the cumulative effect of direct and indirect inhibition of MmpL3 by disruption of respiration, membrane potential and pH gradient (Li et al., 2014). This highly efficient multitarget tropism of SQ109 leads to an extremely effective inhibition of M. tuberculosis growth. Since drug-combination regimens have long been the gold standard strategies for anti-TB therapies to restrain the emergence of spontaneous drug resistance, the ‘multi-targeting’ character of SQ109, currently under evaluation in phase II clinical trials, stands out as a huge advantage for advanced TB therapies (Sacksteder et al., 2012). The essentiality of mmpL3 in M. tuberculosis and the lethal phenotype arising from functional blocking were further confirmed with several newly identified mycobactericidal compounds, namely AU1235 (Table 4), with indoleamide compounds active against MDR M. tuberculosis in vitro, with indolcarboxamides, with the pyrrole derivative BM212 (Table 4), as well as some of its chemically modified versions, and with the benzimidazole C215 (Table 4) (Biava et al., 2007; Deidda et al., 1998; Grzegorzwicz et al., 2012; La Rosa et al., 2012; Lun et al., 2013; Poc et al., 2013; Rao et al., 2013; Remuñán et al., 2013; Stanley et al., 2012). All these chemicals target MmpL3 function and block translocation of TMM across the plasma membrane to the periplasmic compartment inhibiting the assembly of the cell wall mAGP complex or the mycomembrane TDM (Fig. 1). Again, a multitarget action for these compounds is anticipated. Anyhow, the possibility of a new generation of anti-TB drugs targeting the function of the TMM transporter MmpL3 has gained more than substantial support in the last few years (Rayasam, 2014; Sacksteder et al., 2012).

The crucial function of the Ag85 mycolyltransferases involved in TDM and mAGP synthesis and their secreted nature render this class of proteins also attractive drug targets. They are directly accessible to exogenous inhibitors and simultaneously avert intracellular inactivation mechanisms by enzymes and also efflux pumps (Chen et al., 2011; Favrot & Ronning, 2012; Viveiros et al., 2012). Their ‘targetable’ character is further reinforced by the fact that inhibition of a single M. tuberculosis Ag85 protein may readily have a direct consequence on cell wall biosynthesis, cell integrity and increased susceptibility to other antibiotics (Jackson et al., 1999; Nguyen et al., 2005). Moreover, since at least two Ag85 enzymes are required for mycobacterial viability, it is also highly unlikely that a ‘double’ drug resistance event arises (Favrot & Ronning, 2012). However, because Ag85 isoforms are structurally similar, an inhibitor of a single Ag85 protein may also synergistically affect the activity of the others with enhanced mycobactericidal outcome (Ronning et al., 2004). Interestingly, the enzymic mechanism of M. tuberculosis Ag85 enzymes probed with
6,6′-di-hexanoyltrehalose also seems to corroborate their different roles in vivo (Barry et al., 2011; Ronning et al., 2004).

Earlier studies demonstrated that the trehalose analogue 6-azido-6-deoxy-trehalose (6-TreAz) inhibited the activity of all three members of the Ag85 complex in vitro and mycobacterial growth (Belisle et al., 1997; De Smet et al., 2000). Due to the Ag85 proteins’ crucial roles, many subsequent studies focused on the identification or design of different Ag85 inhibitors and have disclosed promising lead compounds. Several synthetic trehalose analogues with long hydrocarbon chains were shown to inhibit Ag85 proteins and the growth of M. smegmatis (Belisle et al., 2012). In contrast, trehalose analogues with sulfonamide -alkylamine groups impaired the activity to isoniazid was also inversely proportional to the growth (Sanki et al., 2014). An unexpected development in 6-TreAz on Ag85 proteins was demonstrated in vitro, but most TreAz analogues internalized by the LpqY–SugABC recycling system (Fig. 1) were incorporated into glycolipids via de novo synthesis and not by the transfer activities of mycolyltransferases (Belisle et al., 1997; Swarts et al., 2012). A newly identified Ag85C inhibitor designated I3-AG85 (Table 4) was recently found to inhibit the growth of sensitive and MDR-TB through the inhibition of TDM biosynthesis and cell envelope integrity without affecting the abundance of MAs in the mAGP complex (Warrier et al., 2012). Although this seems to challenge the association of Ag85C with the preferential transfer of MAs to AG, TDM synthesis by the remaining Ag85 proteins was also affected by I3-AG85, corroborating the inhibitory effect of 6-TreAz on Ag85 proteins.

### Table 4. New TB drug candidates: inhibitors of M. tuberculosis trehalose-related pathways and growth

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>Chemical structure</th>
<th>Pathway</th>
<th>Target</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM212</td>
<td><img src="image" alt="BM212 Chemical Structure" /></td>
<td>TMM translocation</td>
<td>MmpL3</td>
<td>(La Rosa et al., 2012)</td>
</tr>
<tr>
<td>SQ109</td>
<td><img src="image" alt="SQ109 Chemical Structure" /></td>
<td>TMM translocation</td>
<td>MmpL3</td>
<td>(Sacksteder et al., 2012)</td>
</tr>
<tr>
<td>AU1235</td>
<td><img src="image" alt="AU1235 Chemical Structure" /></td>
<td>TMM translocation</td>
<td>MmpL3</td>
<td>(Grzegorzewicz et al., 2012)</td>
</tr>
<tr>
<td>C215</td>
<td><img src="image" alt="C215 Chemical Structure" /></td>
<td>TMM translocation</td>
<td>MmpL3</td>
<td>(Stanley et al., 2012)</td>
</tr>
<tr>
<td>I3-AG85</td>
<td><img src="image" alt="I3-AG85 Chemical Structure" /></td>
<td>TDM synthesis</td>
<td>Ag85C</td>
<td>(Warrier et al., 2012)</td>
</tr>
<tr>
<td>Ebselen</td>
<td><img src="image" alt="Ebselen Chemical Structure" /></td>
<td>TDM synthesis mAGP assembly?</td>
<td>Ag85C</td>
<td>(Favrot et al., 2013)</td>
</tr>
</tbody>
</table>
their isofunctional role in vitro, despite the increasing evidence of a different function in vivo (see above).

Recently, a marked plasticity of Ag85 proteins for a number of fluorescein-labelled trehalose analogues was experimentally confirmed by their ability to link these unnatural disaccharides to TMM and to TDM, leading to the development of imaging probes to selectively visualize live M. tuberculosis (Backus et al., 2011; Koester et al., 2011). The azide-modified trehalose analogues (2-, 3-, 4- and 6-TreAz) designed to incorporate trehalose-containing glycolipids also allowed labelling of the cell envelope with fluorescent probes and visualization of surface glycolipids (Swarts et al., 2012). Further progress on the development of trehalose-based imaging tools extends to the synthesis of the first TDM probe for affinity-based proteome profiling, which was designed to reveal the molecular interactions of this glycolipid with its binding proteins, for example the components of macrophage receptors Mincle and MCL (Furukawa et al., 2013; Khan et al., 2013). Probing the molecular mechanisms underlying the interactions between trehalose-related enzyme targets and their ligands will add new insights into the mycobacterium-host paradigm and further our knowledge on the mycobacterial ‘trehalome’, an enlightening term coined in the Bertozzi lab (Swarts et al., 2012).

**Missing links in M. abscessus trehalose metabolism**

An emerging cause of pulmonary disease increasingly reported in the recent literature, not only in immunocompromised patients but also in immunocompetent individuals worldwide, is M. abscessus, one of the most virulent RGM known (Hoefsloot et al., 2013; Varghese et al., 2013). This health threat is aggravated by the organism’s natural and acquired resistance to multiple antibiotics and disinfectants, as well as by its remarkable ability to survive in stable biofilms in artificial environments, namely domestic and hospital water-distribution systems (Mullis & Falkinham, 2013; Nessar et al., 2012; Wallace et al., 1998). Intensification of research on the genetics, metabolism and lifestyle of this pathogen in natural and man-made environments, as well as during host infection, is therefore mandatory.

An essential role for trehalose in the biology of M. abscessus is also anticipated, but the metabolism of this disaccharide in this pathogen has been poorly investigated. The identity of the genes coordinating trehalose pathways can be inferred from the fully sequenced genomes available (http://patricbrc.org/portal/portal/patric/Home). Moreover, the development of specific genetic tools will promote the assessment of unique phenotypic traits of M. abscessus and their roles in pathogen–host interactions (Klotzsche et al., 2009; Nessar et al., 2011). However, the nature of the putative trehalose-based glycolipids in this organism, as well as the genetic/enzymatic biosynthetic resources available, remain largely unknown. A preliminary survey within the M. abscessus genome revealed a rather exotic genetic map for trehalose biosynthesis as this pathogen does not seem to rely on the archetypal three-pathway scheme. Typical TreY/TreZ and TreS were not detected (Table 3), rendering the OtsA/OtsB system as the putative exclusive source of endogenous trehalose. How the glycogen-to-trehalose cycle is operating in this organism remains a subject for future research. Hypothetically, the available TreX, whose catalytic function remains to be experimentally established, may play an unprecedented role in trehalose synthesis from glycogen. Intriguingly, we could also not detect typical otsB genes, suggesting that the conversion of T6P into trehalose may rely on a different type of phosphatase (Table 3). The enzymes for biosynthesis of SL-1 and PAT as well as those for LOS and the putative systems for their translocation to the outer layers of the cell envelope are also not conserved between M. tuberculosis and M. abscessus (Table 3). Investigation of their genuine functions in this organism in particular, and in NTM in general, will certainly advance our understanding about their metabolic abilities, their infection strategies and eventually the identification of novel drug targets, upon which we will need to intervene to fight these pathogens (Dias et al., 2011; Nessar et al., 2011). Application of the elegant trehalose-based imaging tools described above will be of paramount importance to assess the contribution of these apparent genomic ‘incidents’ to the M. abscessus successful lifestyle and pathogenesis.

**Concluding remarks**

Regardless of the research efforts of decades, the quest for more efficient and safer strategies to fight M. tuberculosis and NTM continues. The high global numbers of new infections and the inexorable dissemination of MDR strains call for more efficient vaccines and new anti-mycobacterial therapies. While the mycobacterial cell envelope remains a major target for antibiotics, the deeper we understand the biosynthesis and turnover of its unique components, the closer we will be to find new therapeutic solutions. While, for example, the drugs isoniazid or ethambutol targeting essential enzymes of MAs biosynthesis and structural assembly of the cell, respectively, have been validated by their therapeutic efficacy, new efficient drugs able to target additional functions in the mycobacterial envelope are urgently needed. These encompass the biogenesis of the inner PG and the MAs-rich mycomembrane, and the biosynthesis and export of structurally and immune relevant glycoconjugates. The essential transporter MmpL3 for the trehalose-containing TMM is a new valid target for TB chemotherapy with a panoply of promising leads already in preclinical and clinical pipelines. While some overlap between TB and NTM trehalose-related systems can be anticipated, NTM research will naturally reveal novel trehalose-based molecules and pathways, as well as evolutionary creative survival strategies, that are worth exploring from fundamental and pharmacological perspectives.

While among the enzymes of trehalose metabolism in M. tuberculosis some still remain to be associated with genuine
functions, and only the Ag85 proteins and the TreS have 3D structures determined, the disclosure of novel scaffolds for inhibitors’ screening and rational drug design are also mandatory. Considering the rapidly growing numbers of Mycobacterium species described and their impending clinical impact in modern societies, the dissection of novel enzyme functions, structures and catalytic mechanisms will unquestionably afford a deeper understanding of these organisms’ molecular biology and metabolism, and contribute to the quest for efficient strategies to fight the resilient and infectious mycobacteria.

Acknowledgements
We are grateful to Milton S. da Costa for enlightening discussions on microbial diversity and trehalose biology, and also to the reviewers for sharing their invaluable ideas and suggestions to improve the substance of this review. We acknowledge the Mizutani Foundation for Glycoscience, Japan, and also Fundação para a Ciência e a Tecnologia (FCT), Portugal, Programa Operacional Factores de Competitividade (POFC) and FEDER/COMPETE for financial support through projects PTDC/BIA-BCM/112459/2009, PTDC/BIA-MIC/2779/2012 and PEst-C/SAU/LA0001/2013-2014. S.A., A.M. and V.M. acknowledge the FCT grants SFRH/BPD/43321/2012 and PEst-C/SAU/LA0001/2013-2014. S. A., Mandal, P., Singh, P., Golla, U., Baranwal, S., Chauhan, S. & Tomar, R. S. (2014). Ebselen induces reactive oxygen species (ROS)-mediated cytotoxicity in Saccharomyces cerevisiae with inhibition of glutamate dehydrogenase being a target. FEBS Open Bio 4, 77–89.


References


Miyake, Y., Toyonaga, K., Mori, D., Kakuta, S., Hoshino, Y., Oyamada, A., Yamada, H., Ono, K., Suyama, M. & other authors (2013). C-type lectin MCL is an FcRγ-coupled receptor that mediates the adjuvanticity of mycobacterial cord factor. *Immunity* 38, 1050–1062.


Edited by: S. Spiro